COLLEGE DOCTORAL PAYS DE LA LOIRE
ED MATIERE, MOLECULES, MATERIAUX & GEOSCIENCES

Université Gustave Eiffel

THESE DE DOCTORAT

UNIVERSITE GUSTAVE EIFFEL

ECOLE DOCTORALE N° 596
Matière, Molécules, Matériaux et Géosciences
Spécialité : *Sciences de la Terre et de l'Environnement*

Par
Juliette GODET

Evaluation d'une chaîne de prévision immédiate des impacts des crues soudaines à l'échelle de l'arc méditerranéen français

Thèse présentée et soutenue à Bouguenais, le 5 décembre 2025
Unité de recherche : Laboratoire Eau Environnement

Rapportrices avant soutenance :

Isabelle RUIN Directrice de recherche CNRS, Institut des Géosciences de l'Environnement
Hélène ROUX Professeure, Institut de Mécanique des Fluides de Toulouse

Composition du Jury :

Examinateurs :	Pierre BRIGODE	Professeur Junior (CPJ), ENS Rennes
	Maria del Carmen LLASAT-BOTIJA	Professeure, Université de Barcelone
	Maria-Helena RAMOS	Directrice de recherche, INRAE-HYCAR
Dir. de thèse :	Eric GAUME	IGPEF, Université Gustave Eiffel
Co-dir. de thèse :	Olivier PAYRASTRE	IDTPE, Université Gustave Eiffel
Co-dir. de thèse :	Pierre JAVELLE	IDAE, INRAE-RECOVER
Membre invité :	Bruno JANET	IGPEF, Service Central Vigicrues

Behind every warning is a family preparing to evacuate, a farmer determining about crops, principals deciding whether to close their schools. Real people. Real lives. Real smiles.

Prof Celeste Saulo, SG WMO

Préface

Pendant plus de trois ans, j'ai dédié l'ensemble de mon temps de travail à ce sujet de thèse, et ce manuscrit concrétise mes efforts. Il peut paraître impersonnel ; le langage scientifique l'est souvent, et l'intelligence artificielle a parfois été utilisée pour clarifier le discours. Pourtant, je tenais à écrire ici que ce travail de thèse est une partie de mon histoire personnelle, et que la conviction d'apporter quelque chose d'utile à la société m'a toujours portée - même si ce n'est qu'une petite brique parmi tant d'autres dans la muraille symbolisant la lutte contre les "catastrophes naturelles".

À l'aube d'un nouveau chapitre de vie, c'est avec une immense affection que je repense à mes années d'école, de collège, de lycée, de classe préparatoire, d'école d'ingénieurs et de doctorat, et je souhaite dédier ce manuscrit à l'ensemble des enseignant.es auprès desquel.les j'ai appris durant ces vingt-cinq premières années de ma vie. J'ai tant aimé apprendre, questionner, partager et participer dans tous ces lieux d'enseignement. Ce manuscrit constitue l'aboutissement de mon parcours éducatif, et j'espère qu'il participe à rendre ne serait-ce qu'une infime partie des connaissances que j'ai reçues.

Enfin, les deux enseignants auxquels je pense le plus en écrivant ces lignes sont mes parents. Chrystel Godet-Lainé, la plus merveilleuse des mamans du monde, et ma toute première institutrice à l'école. Laurent Godet, la personne la plus savante et cultivée de mon univers, qui j'espère aurait été fier de sa fille, sur le point de devenir docteure.

Remerciements

J'adresse mes remerciements les plus profonds à mes encadrants de thèse ; mon directeur Eric Gaume, pour sa grande disponibilité malgré ses responsabilités, et pour ses suggestions toujours très justes et pertinentes ; mon encadrant Olivier Payrastre, pour son accompagnement sans faille, sa pédagogie et sa rigueur scientifique qui m'inspireront toujours ; mon encadrant aixois Pierre Javelle pour son écoute attentive, sa bienveillance, et ses conseils judicieux et rassurants. On ne choisit pas vraiment son équipe encadrante quand on démarre une thèse, alors j'ai la sensation d'avoir été très chanceuse de recevoir un trio d'encadrants très complémentaire et me correspondant parfaitement.

Je remercie Pierre Nicolle, qui a très étroitement suivi mon avancement, et qui m'a toujours donné un coup de main dès que j'en avais besoin.

Je remercie Céline Cattoën-Gilbert, pour son formidable accueil au NIWA (Earth Sciences NZ), sa disponibilité et sa grande gentillesse.

Je remercie les membres de mon comité de suivi de thèse, et les membres du jury, pour leur temps et l'intérêt qu'ils ont porté à mes travaux.

Je remercie l'ensemble des acteurs impliqués dans la prévention des inondations qui ont contribué à mon travail par des discussions, des conseils, du partage de codes ou de données. Ils sont si nombreux que je ne pourrai pas tous les citer : la Caisse Centrale de Réassurance, les Services de Prévision des Crues et les SDIS de l'arc méditerranéen, le Service Central Vigicrues, Météo-France, l'INRAE, le CEREMA, Enedis, Orange...

Je remercie tous les collègues et ami.es que j'ai rencontré.es durant ce parcours, entre le laboratoire Eau Environnement de l'Univ. Eiffel à Nantes, l'équipe RHAX de l'INRAE à Aix-en-Provence, le NIWA à Christchurch en Nouvelle-Zélande, et à toutes les conférences et écoles d'été auxquelles j'ai participé, en France et à l'étranger. Je n'oublierai jamais les moments chaleureux passés à vos côtés tout au long de ces formidables années.

Enfin, alors que je m'apprête à soumettre ce manuscrit, j'adresse une tendre pensée à tous mes proches. À ceux qui sont arrivés dans ma vie au cours de cette période. À ceux qui étaient déjà là et le sont encore. À ceux qui sont partis pour toujours. Merci pour tout.

Résumé

Les crues soudaines, qui touchent les petits cours d'eau, représentent une part importante des dommages liés aux inondations en France. Mieux les anticiper est essentiel pour en limiter les conséquences. Cette thèse construit et évalue une chaîne innovante de prévision immédiate des crues soudaines, intégrant une représentation des impacts attendus. La chaîne repose sur des outils existants de modélisation hydrologique et hydraulique, enrichis par des réflexions sur l'articulation des blocs de simulation. Elle a été appliquée sur environ 20 000 km de cours d'eau (bassins versants > 5 km^2) des départements côtiers de l'arc méditerranéen français, sur une période continue de plus de dix ans (2010–2022).

Dans un premier temps, la chaîne de simulation a été évaluée bloc par bloc, sur sa capacité à prévoir l'intensité des crues et à modéliser les zones inondées. Par la suite, la chaîne complète a été évaluée à l'aide de plusieurs jeux de données d'impacts : arrêtés CatNat, données assurantielles et interventions des services de secours. Enfin, la pertinence des avertissements basés sur l'intensité des crues, dans la logique du système opérationnel Vigicrues Flash actuel, a été comparée à celle d'avertissements basés sur les impacts. Les résultats ont montré que la prévision des impacts permet de réduire sigificativement le nombre de fausses alarmes et d'améliorer la hiérarchisation des zones impactées.

Table des matières

I Introduction générale **13**

1 Positionnement de la thèse dans la prévention des risques naturels 14

1.1 Risques naturels et prévention : définitions, chiffres et réflexions 14

1.2 Une approche focalisée sur un aléa unique : les crues soudaines 19

1.3 Anticiper les crues soudaines : la dernière ligne de défense contre le risque . . . 26

2 Prévision immédiate des impacts des crues soudaines 32

2.1 Structuration par blocs des chaînes de prévision des crues 32

2.2 Opportunités et verrous posés par les avertissements basés sur les impacts 39

3 Objectifs de la thèse et organisation du manuscrit 44

3.1 Objectifs de la thèse 44

3.2 Organisation du manuscrit 45

II Évaluation d'un système d'avertissement aux crues soudaines basé sur l'intensité de l'aléa **47**

0 Synthèse en français 48

1 Introduction 50

1.1 Flash floods and operational early warning systems worldwide 50

1.2 The flash flood warning system in France 50

1.3 Assessing the performance of flood warnings 52

2 Validation data 53

2.1 Continuous measurements from hydrometric stations 53

2.2 Peak discharge estimates from post-event surveys 53

2.3 CatNat decrees 54

3 Methods 55

3.1 Rainfall-runoff model 55

3.2 Contingency-based validation 57

3.3 Flood intensity-based contingencies 57

3.4 Impact-based contingencies 59

3.5 Determination of rare discharge quantiles based on the SHYREG method 59

4 Results 61

4.1 Flood intensity-based validation 61

4.2 Impact-based validation 62

5 Discussion and conclusions 65

5.1 Study limitations 65

5.2 Key results and interpretations 66

5.3 Perspectives 67

III Évaluation d'une modélisation hydraulique couvrant les petits cours d'eau à une échelle régionale **69**
0 Synthèse en français . . . 70
1 Introduction : context and objectives . . . 72
2 Data description . . . 73
2.1 Study areas . . . 74
2.2 Computation and validation of flood hazard maps . . . 75
2.3 Computation and validation of actual extreme flood events simulations . . . 78
3 Description of data prepocessing methods . . . 79
3.1 DTM preprocessing . . . 79
3.2 Hydrograph calculations . . . 80
3.3 Floodos computations . . . 81
4 Illustration of the validation methods and of the results obtained with the Floodos model 82
4.1 Validation of flood hazard maps using hydrometric data (rating curves) . . . 82
4.2 Validation of flood event inundation maps using observed flood extents and high water marks . . . 85
5 Discussions . . . 87
6 Conclusions . . . 89

IV Implémentation de la chaîne de prévision immédiate basée sur les impacts **91**
1 Connecter l'hydrologie et l'hydraulique : associer des points-rivière aux mailles d'une grille de modélisation hydrologique . . . 92
1.0 Synthèse en français . . . 92
1.1 Introduction . . . 94
1.2 Allocation methods and validation metrics . . . 95
1.3 Case study and data . . . 98
1.4 Results . . . 100
1.5 Discussion and conclusions . . . 104
2 Mise en œuvre d'une chaîne de prévision basée sur les impacts . . . 106
2.1 Reconstitution de l'emprise inondée . . . 106
2.2 Calculs des impacts . . . 109
2.3 Application de la chaîne complète pour le besoin la thèse et transférabilité en contexte opérationnel . . . 111
3 Discussions et conclusions . . . 113

V Évaluation de la plus-value d'une chaîne de prévision immédiate basée sur les impacts **115**
0 Synthèse en français . . . 116
1 Introduction . . . 119
1.1 Hazard-based and impact-based warnings . . . 119
1.2 Impact-based warnings in the context of flash floods . . . 120
1.3 Verification of impact-based warnings . . . 121
2 Flood Impact data . . . 122
2.1 Collecting flash flood impact data in South-Eastern France . . . 122
2.2 CatNat decrees . . . 125
2.3 Insurance data . . . 126
2.4 Fire and rescue service operations records . . . 127

3 Nowcasting chains and validation methods . . . 129
3.1 Hazard-based nowcasting chain . . . 129
3.2 Impact-based nowcasting chain . . . 130
3.3 Validation approach . . . 132
4 Results and discussions . . . 134
4.1 Comparison between hazard-based and impact-based models : significant reduction of false alarms . . . 135
4.2 Discrepancy between CatNat data and the IB model . . . 136
4.3 Low hit rates for the fire and rescue services operations . . . 137
4.4 Flood-based and hybrid models . . . 138
4.5 Model uncertainties and data limitations . . . 139
4.6 Validation results at the event scale . . . 140
5 Conclusions . . . 141

VI Incertitudes : prévisions d'ensemble de pluie et propagation des erreurs 145
1 Plus-value de prévisions d'ensemble de pluies dans une chaîne de type HBW . . . 146
1.0 Synthèse en français . . . 146
1.1 Introduction . . . 148
1.2 The short-range hydrometeorological forecasting chains . . . 149
1.3 Case studies and verification method . . . 153
1.4 Results and discussion . . . 157
1.5 Conclusions . . . 164
2 Mobilisation de prévisions d'ensemble de pluies en entrée d'une chaîne de type IBW . . . 165
2.1 Mise en œuvre de la chaîne IBW à l'échelle de l'événement du 1er décembre 2019 165
2.2 Analyse de la propagation des incertitudes de prévision de pluie au sein de la chaîne IBW . . . 167
3 Conclusions et perspectives . . . 173

VIIConclusions et perspectives 175
1 Enseignements tirés en lien avec les objectifs de la thèse . . . 176
1.1 Mise en place d'une chaîne de prévision immédiate des impacts des crues soudaines pour l'Arc Méditerranéen . . . 176
1.2 Évaluation détaillée sur une large fenêtre spatio-temporelle . . . 178
2 Perspectives . . . 181
2.1 Prévision des pluies : allonger les délais d'anticipation . . . 181
2.2 Prévision des débits : améliorer les modèles hydrologiques . . . 181
2.3 Prévision des zones inondées : complexifier et interpoler . . . 182
2.4 Prévision des impacts : mieux répondre aux besoins des utilisateurs . . . 182
2.5 Chaîne de prévision complète : mieux comprendre, mieux évaluer, mieux communiquer . . . 182

Bibliographie 185

Annexes 211

Chapitre I

Introduction générale

Objectifs et organisation du chapitre :

Ce chapitre introductif vise à poser les fondations du travail de recherche présenté dans cette thèse. Il s'articule en deux grandes parties. La première inscrit la thématique dans le champ de la prévention des risques naturels, en commençant par une présentation générale des risques naturels en France et dans le monde, de leurs impacts et des dispositifs mis en place pour leur prévention, avec un recentrage progressif sur le risque inondation. Elle décrit ensuite le phénomène des crues soudaines, souvent brutales et dévastatrices lorsqu'elles affectent les bassins versants réactifs de l'arc méditerranéen français. Enfin, cette première partie détaille le fonctionnement des systèmes de prévision des crues et des crues soudaines actuellement en vigueur en France.

La seconde partie s'intéresse plus concrètement aux outils et à la structuration des chaînes de prévision des crues, depuis le forçage météorologique jusqu'à l'estimation des impacts, en passant par la modélisation hydrologique et hydraulique. Elle met en lumière l'intérêt croissant, dans la littérature scientifique comme dans les applications opérationnelles, pour des systèmes capables de produire des avertissements intégrant explicitement les impacts potentiels. Les bénéfices attendus et les défis associés à cette évolution sont discutés. Le chapitre se conclut par la présentation des objectifs poursuivis dans cette thèse, ainsi que de son organisation générale.

1 Positionnement de la thèse dans la prévention des risques naturels

1.1 Risques naturels et prévention : définitions, chiffres et réflexions

1.1.1 Le cadre sémantique des risques "naturels"

Évoquer le concept de "risques naturels" nécessite en premier lieu de définir la notion de "risque". Les racines étymologiques de ce mot sont assez incertaines, mais semblent renvoyer assez fréquemment aux notions de "danger", "hasard" et "incertitude" [Saint-Martin, 2018]. Ici, la définition retenue est celle du Bureau des Nations Unies pour la prévention des catastrophes, traduite de l'anglais : *"le risque est la probabilité qu'un événement ait des conséquences négatives sur les personnes, les systèmes ou les biens"* [UNDRR, 2019]. Cette définition correspond donc à l'équation généralement associée à la notion de risque :

$$Risque = Aléa \times Exposition \times Vulnérabilité \tag{I.1}$$

Les risques naturels peuvent se définir comme les risques associés aux aléas d'origine naturelle. Mais cette définition a des limites, puisque l'on sait que le changement climatique lié aux activités humaines a un impact sur les événements pluvieux, que les excavations minières peuvent déclencher des séismes, et que les incendies de forêt sont rarement d'origine naturelle. UNDRR [2020] a proposé une classification des aléas en plusieurs types, qui n'utilise pas le terme "naturel" et évite donc ces ambiguïtés :

- Météorologiques et hydrologiques : des exemples spécifiques incluent les cyclones, tsunamis, crue soudaine, sécheresse.
- Géophysiques : séismes, glissements de terrain, éruptions volcaniques, etc.
- Environnementaux : feux de forêt, contamination des sols par les pesticides, pollution plastique, etc.
- Extraterrestres : impacts de météorites, tempêtes solaires, etc.
- Chimiques : intoxication au monoxyde de carbone, contamination des aliments, etc.
- Biologiques : épidémies, espèces invasives, etc.
- Technologiques : accidents nucléaires, ruptures de digue, cyberattaques, etc.
- Sociétaux : guerres, chocs financiers, etc.

Par ailleurs, l'expression de "risque naturel" est de plus en plus critiquée par les géographes, avec notamment Leone et al. [2021] qui la qualifient comme *"un raccourci plus commode que juste"*, car la société contribue très largement à l'élaboration de sa propre vulnérabilité [1]. Ce mouvement est traduit à l'international par la campagne "#NoNaturalDisaster".

Restent dans l'équation I.1 les termes "Exposition" et "Vulnérabilité". Si la notion d'"Exposition" se rapporte assez clairement à tout enjeu localisé dans une zone géographique exposée à l'aléa (e.g. les personnes, les biens, les activités, le patrimoine), la notion de "Vulnérabilité" est complexe. La définition de UNDRR [2023] sera une nouvelle fois conservée pour cette thèse : *"Les conditions déterminées par des facteurs ou processus physiques, sociaux, économiques et environnementaux qui augmentent la susceptibilité d'un individu, d'une communauté, de biens ou de systèmes à être impactés par un événement"*. Cette notion est un paramètre essentiel de l'équation I.1, sans lequel le risque serait considéré comme égal si un événement identique en termes d'intensité et d'enjeux affectés touchait deux zones géographiques avec des caractéristiques différentes [2].

1. Par exemple, en construisant en zone inondable, ou en ne respectant pas les normes parasisimiques.
2. Par exemple, un cyclone d'une catégorie fixée aurait un impact plus dévastateur en Haïti qu'en Floride, car ces deux pays disposent de moyens de préparation, de secours et de ressources bien différents.

1.1.2 Les impacts des aléas naturels en France et dans le monde

La notion d'impact est également une notion clé qui sera omniprésente dans cette thèse. Elle sera utilisée de manière indifférenciée avec le terme "dommage", qui désigne les aspects négatifs des impacts. Les dommages directs (directement induits par l'aléa, notamment par destruction physique) et les dommages tangibles (qui peuvent être traduits en valeur monétaire) sont généralement utilisés pour classifier les différents types de dommages [Merz et al., 2020]. Par exemple :

- Les impacts directs tangibles peuvent désigner la destruction de routes, d'immeubles, le déplacement d'équipes de secours, l'interruption de certains services et activités au sein de la zone affectée par l'aléa.
- Les impacts directs intangibles peuvent correspondre aux blessures, décès, traumatismes, ou pertes de patrimoine environnemental, culturel ou historique.
- Les impacts indirects tangibles peuvent qualifier les problèmes liés à des coupures d'électricité, de trafic ou à des pertes d'emploi qui s'étendent au delà de la zone affectée.
- Les impacts indirects intangibles peuvent inclure les pertes de confiance envers les autorités lorsqu'il est observé, au delà de la zone affectée, que la crise a été mal gérée [3].

Les chiffres les plus fréquemment utilisés pour caractériser des impacts directs tangibles sont les coûts des dommages matériels, qui peuvent être estimés via les données assurantielles. L'un des principaux réassureurs mondiaux, Munich Re, a ainsi estimé que le coût moyen annuel des catastrophes naturelles entre 1993 et 2022 est d'environ 180 Md US$, et de 250 Md US$ (soit un montant proche du PIB de la Grèce) sur la période 2018-2022 [MunichRe, 2024], ce qui démontre une hausse significative sur la période récente.

En termes de dommages directs intangibles, MunichRe [2024] a également calculé un nombre moyen annuel de 11,400 décès liés aux catastrophes naturelles sur la période 2018-2022 [4], et UNICEF [2023] a fait savoir qu'entre 2016 et 2021 et dans 44 pays, 43.1 millions d'enfants ont été forcés à fuir leur domicile à cause de catastrophes naturelles. 95% de ces déplacements étaient causés par des inondations et/ou tempêtes, 3% par des sécheresses, un peu moins de 2% par des incendies, et minoritairement par d'autres aléas tels que les séismes.

À l'échelle mondiale, les inondations sont classées en tête selon plusieurs indicateurs d'impact, comme la fréquence ou le nombre de personnes affectées, d'après un rapport du CRED [2024] établi à partir des données EM-DAT [5], et synthétisé dans le tableau I.1. Ce même rapport souligne que si les pertes économiques sont plus élevées dans les pays développés, ce sont les pays les moins avancés qui paient le plus lourd tribut en termes de victimes et de populations affectées. C'est ce que montre la figure I.1 : à l'exception des États-Unis et du Japon, les catastrophes naturelles provoquent beaucoup plus de déplacements dans les pays du Sud, notamment en Asie : plus de 70 millions en dix ans en Chine, et plus de 30 millions sur la même période en Inde et aux Philippines.

En France, la Caisse Centrale de Réassurance (CCR) dresse chaque année un bilan du régime d'indemnisation des risques de catastrophes naturelles (dit régime Cat Nat). Ce régime, reposant sur le principe de "solidarité nationale", permet aux particuliers et aux entreprises d'être indemnisés lorsque leur commune a été déclarée en état de "catastrophe naturelle" par un arrêté interministériel publié au

3. Cela fait notamment penser aux inondations catastrophiques en Espagne fin octobre 2024, comme le montre un article du journal Le Monde [Morel, 2024].

4. Bien que le recensement des décès liés aux aléas naturels pose toujours question ; les chiffres d'EM-DAT dans le tableau I.1 ne sont pas cohérents avec cette estimation.

5. Un événement est comptabilisé dans la base de données EM-DAT s'il respecte les critère suivant : 10 victimes ou plus, 100 personnes affectées ou plus, déclaration de l'état d'urgence, demande de l'aide internationale. Ainsi, beaucoup d'événements sont omis dans le calcul.

	Nb. de catastrophes	Nb. de décès	Nb. de personnes affectées (Md)	Dommages économiques (Md US$)
Inondations	170	5518	74.6	41.1
Sécheresse	16	1157	57.1	9.6
Tempêtes	104	10 017	32.2	95.6
Séismes	27	35 124	5.6	39.9
Températures extrêmes	21	11 470	4.8	3.4
Mouvements de terrain	19	838	0.2	0.3
Eruptions volcaniques	5	80	0.3	0.2
Feux de forêt	11	86	0.7	6.3

TABLEAU I.1 – Chiffres moyens annuels sur les impacts des catastrophes naturelles dans le monde, calculés sur la période 2003-2022 par CRED [2024] à partir des données EM-DAT.

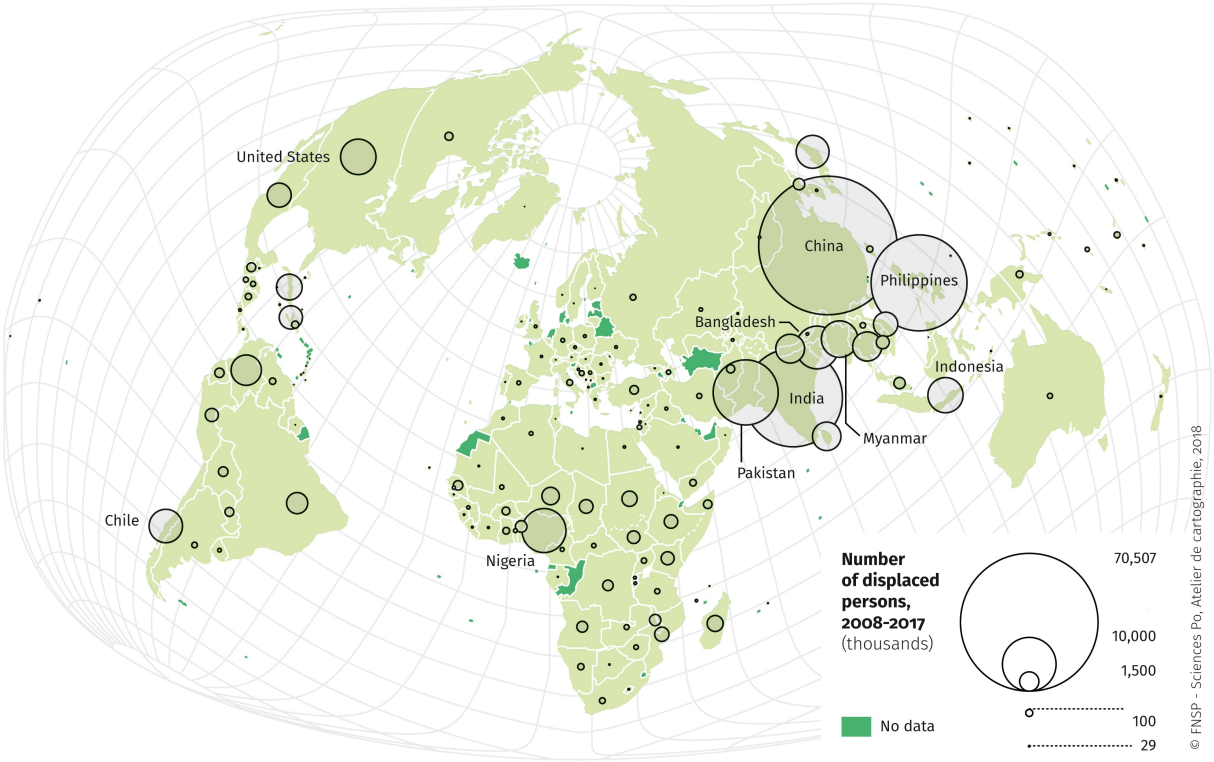


FIGURE I.1 – Personnes déplacées suite à des catastrophes naturelles entre 2008 et 2017, selon les données du Internal Displacement Monitoring Centre.

Journal officiel de la République française (JO). Le régime Cat Nat n'a pas réellement d'équivalent dans le monde, et est généralement regardé avec envie depuis l'étranger [MRN, 2012].

Dans son dernier bilan, CCR [2024] a calculé que les inondations représentaient 49% - soit près de 28 Md€ - de la sinistralité Non-Auto cumulée sur la période 1982-2023, faisant des inondations le risque majeur en France sur cette période, devant la sécheresse (42%) et plus particulièrement les dommages liés au retrait-gonflement des argiles, les vents cycloniques (5%) et les séismes (2%). [6] Les figures I.2a et b montrent d'une part que la quasi-totalité des communes françaises ont été reconnues au moins une fois en état de catastrophe naturelle "Inondation", et que l'arc méditerranéen français se distingue nettement par sa singularité, en termes de fréquence mais également en termes de coût des sinistres. La figure I.3 montre l'évolution du coût moyen d'une reconnaissance Cat Nat inondation, avec une nette tendance à la hausse (coût moyen multiplié par deux entre la période 1982-2023 et la période 2013-2023).

Dans un recueil de chiffres clés sur les risques naturels [SDES, 2024], le service des données et

6. Cependant, sur les 10 dernières années, la sécheresse (54%) devance les inondations (31%).

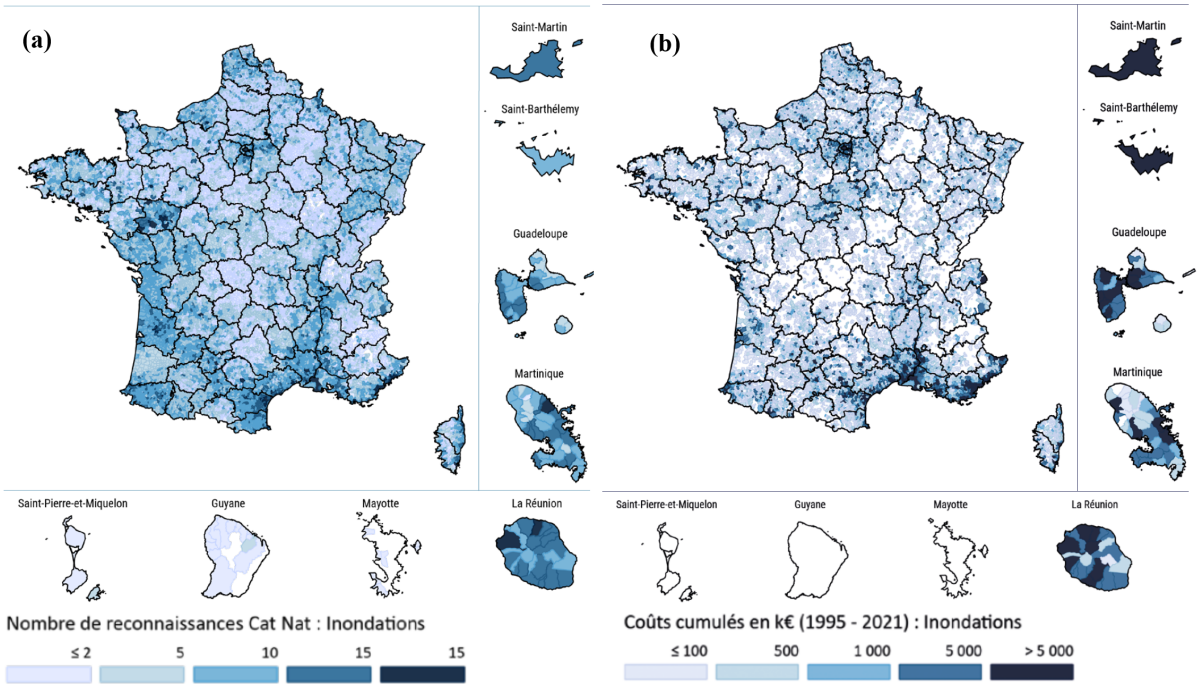


FIGURE I.2 – (a) Nombre d'arrêtés CatNat "Inondation" sur la période 1982 - 2023 et (b) coûts associés cumulés sur la période 1995 - 2021 [CCR, 2024]

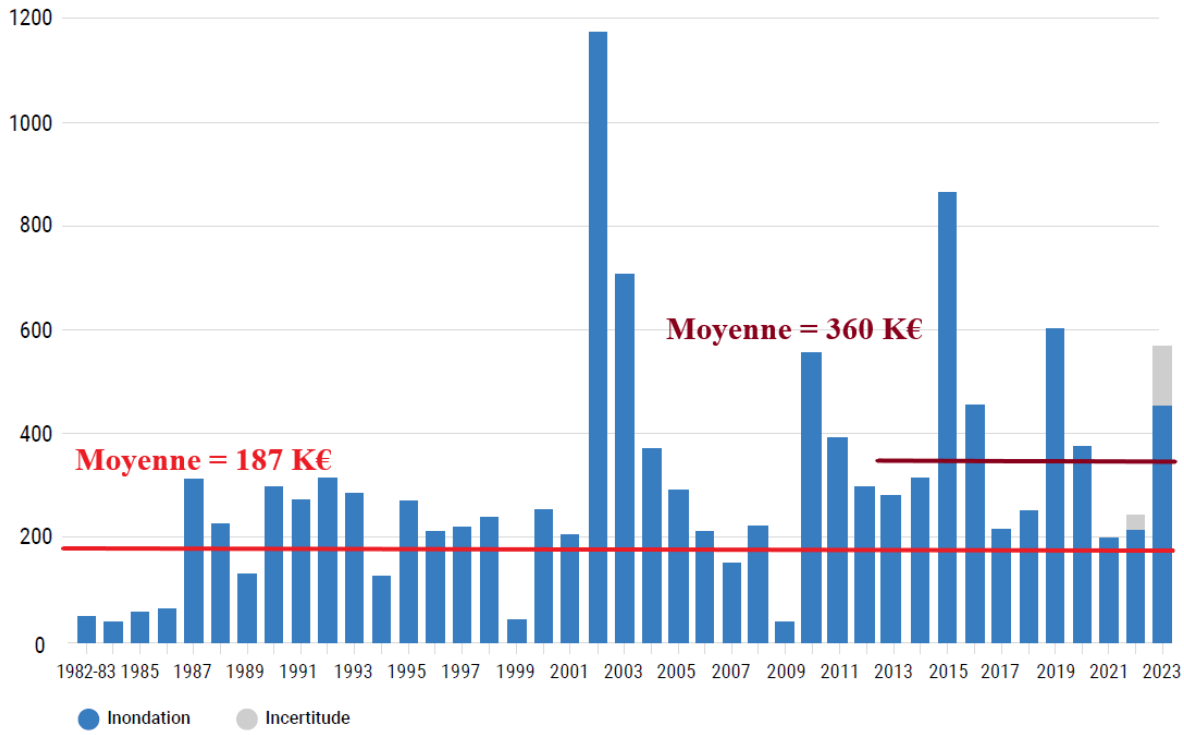


FIGURE I.3 – Evolution du coût moyen d'une reconnaissance Cat Nat inondations (en milliers d'euros d'indemnisation 2023) [CCR, 2024].

études statistiques du Ministère de l'Environnement a estimé que 32 418 décès entre 1900 et 2021 en France ont été causés par des catastrophes naturelles, dont près de 90% par l'éruption de la montagne Pelée en Martinique en 1902, la plus meurtrière du 20è siècle à l'échelle mondiale. En dehors de cet événement cataclysmique, ce sont les phénomènes atmosphériques tels que les tempêtes et les ouragans qui comptabilisent le plus de décès (1 983 décès, soit 58% des décès non liés à l'éruption), puis les inondations (25%, soit 865 décès au total - i.e. 7 décès par an en moyenne).

1.1.3 Les politiques de prévention des risques naturels aux échelles globale et nationale

Les risques naturels ont des répercussions profondes sur les sociétés humaines, incitant à la mise en place de politiques de prévention pour en réduire les impacts. À l'échelle mondiale, le Cadre de Sendai pour la réduction des risques de catastrophes (2015-2030) constitue la référence centrale. Adopté lors de la troisième Conférence mondiale des Nations Unies à Sendai, au Japon, il repose sur quatre priorités fondamentales : comprendre les risques, renforcer la gouvernance, investir dans la résilience et améliorer l'anticipation, la réponse et la reconstruction. L'UNDRR (United Nations Office for Disaster Risk Reduction) est l'organisme chargé de coordonner la mise en œuvre du Cadre de Sendai. Il fournit des orientations stratégiques et techniques, facilite la coopération internationale et promeut l'intégration des risques dans les politiques de développement durable. En mobilisant à la fois les gouvernements, les acteurs privés, les ONG et les communautés locales, l'UNDRR vise à renforcer la résilience des sociétés face aux catastrophes naturelles.

En France, la gestion des risques naturels repose sur une législation solide et articulée autour de sept piliers fondamentaux :

- La connaissance des risques : Inventaires, cartographies et études techniques.
- La surveillance, la prévision et l'alerte : Météo-France et le Service Central Vigicrues (SCV, ex SCHAPI) pour la prévision des crues, et les maires et les préfectures pour les alertes.
- La prise en compte des risques dans l'aménagement : Plans de prévention des risques naturels (PPRn).
- La réduction de la vulnérabilité : Programmes de sensibilisation et normes de construction.
- L'information préventive : Outils de communication pour le grand public.
- La préparation à la gestion de crise : Plans ORSEC (Organisation des Secours), et PCS (Plans communaux de sauvegarde).
- Le retour d'expérience : Évaluation systématique après chaque événement majeur.

Cependant, malgré ces avancées vers une gestion cohérente et efficace des risques, des conséquences imprévues et non souhaitées peuvent survenir. Par exemple, des mesures de prévention visant à réduire la vulnérabilité des bâtiments face à un type d'aléa spécifique (par exemple, des constructions sur pilotis pour réduire le risque d'inondation) peuvent avoir des effets non désirés sur d'autres types d'aléas, comme l'augmentation de la vulnérabilité aux séismes. Ce phénomène, qualifié d'"asynergies" dans les mesures de prévention des risques par de Ruiter et al. [2021], met en lumière les dilemmes que peuvent rencontrer les gestionnaires de risques. D'autre part, la surprotection de certaines zones peut attirer davantage de populations et d'infrastructures, augmentant ainsi la densité et la vulnérabilité, créant un effet pervers à long terme. Par exemple, la construction de digues le long des cours d'eau peut donner un faux sentiment de sécurité, encourageant l'urbanisation en zone inondable. Si ces digues venaient à céder lors d'un événement extrême, les conséquences pourraient être bien plus catastrophiques que si les populations avaient été tenues à distance.

La figure I.4 illustre ce système complexe, où les interactions entre hydrologie, société et politiques de prévention sont influencées par des facteurs externes tels que les changements climatiques et les tendances socio-économiques globales. L'interdépendance entre ces éléments montre que toute décision prise dans un domaine peut avoir des répercussions inattendues sur un autre. Face à l'imprévisibilité des phénomènes naturels, la gestion des risques nécessite une humilité collective. Il est essentiel de combiner expertise scientifique, implication collective et flexibilité politique pour adapter continuellement les stratégies aux évolutions environnementales et sociétales.

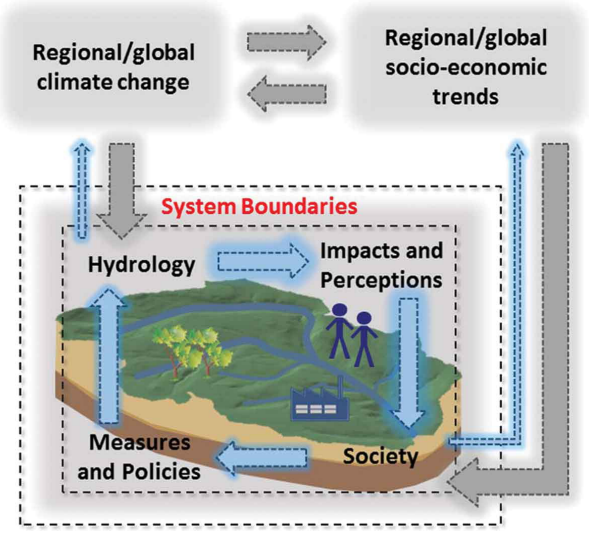


FIGURE I.4 – Illustration de la boucle de rétroaction interne entre les processus hydrologiques et sociaux, qui forme un challenge dans la prévention des inondations [Westerberg et al., 2017].

1.2 Une approche focalisée sur un aléa unique : les crues soudaines

1.2.1 Typologie des inondations et définition des crues soudaines

Dans les sections précédentes, le phénomène d'inondation s'est distingué des autres aléas naturels, apparaissant en tête selon plusieurs indicateurs d'impact, tant à l'échelle nationale que mondiale. Toutefois, le terme inondation reste générique et recouvre une grande diversité de phénomènes différents, dont une typologie est proposée en tableau I.2 [7].

Ce travail de thèse s'est concentré sur les crues soudaines affectant les bassins versants réactifs de l'arc méditerranéen français. Les crues soudaines, ou crues éclair, se caractérisent par une montée très rapide des eaux, généralement en quelques dizaines de minutes à quelques heures après un épisode pluvieux intense [Georgakakos, 1986]. Elles affectent principalement des bassins versants de petite à moyenne superficie, le plus souvent inférieure à 2000 km^2. Les débits spécifiques peuvent y être exceptionnellement élevés, atteignant jusqu'à 15 $m^3/s/km^2$ sur des bassins d'environ 100 km^2, voire 30 $m^3/s/km^2$ sur des bassins d'environ 10 km^2, en région méditerranéenne européenne, ce qui classe ces crues parmi les plus extrêmes au monde pour des bassins de moins de 500 km^2 [Gaume and Borga, 2013]. En raison de leur survenue brutale, de leur échelle spatiale réduite et de leur durée très courte, elles échappent fréquemment aux réseaux de mesure pluviométrique et hydrométrique classiques, ce qui les rend particulièrement difficiles à anticiper [Gaume et al., 2009].

Il peut toutefois s'avérer difficile de définir les frontières entre les crues soudaines et certains autres phénomènes. Par exemple, bien que le terme ruissellement désigne généralement tout écoulement de surface en dehors du réseau hydrographique pérenne, les crues soudaines peuvent également concerner de très petits ruisseaux, habituellement à sec. De la même manière, la distinction entre crues torrentielles et crues soudaines dépendra du seuil au delà duquel on considère que le phénomène de transport solide prédomine. Les frontières entre ruissellements, crues torrentielles et crues soudaines sont donc parfois floues, et il n'est pas toujours pertinent de chercher à différencier ces phénomènes. Dans le cadre de ce travail, un réseau hydrographique est considéré comme défini dès lors que la superficie du bassin versant dépasse 5km^2. Tout écoulement survenant en dehors de ce réseau est ainsi assimilé à du ruissellement,

7. Il est important de souligner que les crues et inondations majeures résultent très souvent de la combinaison de plusieurs phénomènes. Selon l'échelle d'analyse, un même événement peut relever de plusieurs catégories présentées dans le tableau. Les exemples cités sont donc à considérer de manière indicative, car ils pourraient aussi être classés dans d'autres types d'inondation.

Catégorie principale	Phénomène associé	Mécanisme / contexte	Exemple en France
Inondations par débordement de cours d'eau	Crues fluviales	Augmentation lente du débit d'une rivière causant des débordements, après des cumuls/ intensité de pluie importants	Oct.–fév. 2023–2024 : crues historiques dans le Pas-de-Calais (> 300 communes touchées)
	Crues soudaines	Crues rapides formées sur de petits bassins en réaction à des pluies intenses localisées	20 mai 2025 : crue-éclair au Lavandou, montée de la Môle de 0,26 m à 5.75 m entre 08h00 et 09h50
	Crues torrentielles	Crues formées sur des bassins à forte pente, causant débordements et transport solide important	2 oct. 2020 : tempête Alex, vallées de la Vésubie, Roya et Tinée (Alpes-Maritimes)
	Crues nivales	Fonte rapide de la neige combinée à des pluies	Juin 2024 : crue torrentielle à La Bérarde (massif de l'Oisans, Isère)
	Crues d'embâcles / débâcles	Obstruction puis libération brutale du lit du cours d'eau (glace, débris, ponts, etc.)	Juin 1940 : villages des Pyrénées-Orientales dévastés par la rupture d'un barrage naturel dû à un glissement de terrain après des fortes pluies
Inondations hors débordement de cours d'eau	Remontée de nappe	Nappe phréatique affleurante après saturation durable des sols	Janv. 2018 : remontées de nappe en région parisienne lors des crues Seine / Marne
	Ruissellement pluvial urbain	Précipitations dépassant la capacité d'infiltration / drainage en zone imperméabilisée	23 sept. 2024 : épisode orageux à Cannes (Alpes-Maritimes)
	Submersion marine	Surcote temporaire (vent, houle, marée) envahissant le littoral	28 févr. 2010 : tempête Xynthia, Vendée et Charente-Maritime (47 morts, > 1 milliard € de dégâts val. 2022)

TABLEAU I.2 – Typologie des principaux types d'inondation et exemples en France

et n'est pas pris en compte par le modèle. A l'inverse, toute inondation constatée au delà de cette surface drainée sera assimilée à un débordement de petit cours d'eau, et donc à une crue soudaine. Ceci constitue une limitation importante, qui sera largement discutée par la suite.

1.2.2 Origine et géographie des crues soudaines

En Europe et en France, les crues soudaines sont particulièrement fréquentes et violentes sur le pourtour Méditerranéen [Gaume et al., 2009, Llasat et al., 2010]. Cette récurrence s'explique par une combinaison de facteurs climatiques, topographiques, hydrogéologiques et anthropiques.

Facteurs climatiques Le climat méditerranéen, en particulier à l'automne, est propice à des épisodes de précipitations intenses et localisées [Tramblay and Somot, 2018, Mastrantonas et al., 2021]. Cette saison est marquée par un contraste thermique important entre une mer encore chaude, source d'humidité, et une atmosphère qui commence à se refroidir, favorisant l'instabilité. Les remontées d'air chaud, humide et instable en provenance de la Méditerranée rencontrent alors des conditions propices à la convection atmosphérique. Ces masses d'air peuvent générer des systèmes très actifs, parfois stationnaires, qui déversent des pluies diluviennes sur des secteurs restreints. Deux mécanismes principaux sont à l'origine de ces épisodes :

- Le forçage orographique, où l'air humide s'élève au contact d'un relief (Cévennes, Alpes, Pyrénées. . .), se refroidit et condense, entraînant des précipitations intenses.
- Le forçage convectif, où des cellules orageuses se forment du fait de l'instabilité atmosphérique et peuvent s'auto-entretenir sur plusieurs heures, produisant localement des intensités de pluie extrêmes (> 100 mm/h).

Ces systèmes peuvent être très localisés, générant des cumuls exceptionnels en un point, tandis que quelques kilomètres plus loin, les précipitations restent modérées. Cela complique considérablement la prévision de ces événements.

Facteurs topographiques Les crues soudaines touchent majoritairement des cours d'eau de faible ou moyenne taille, au lit étroit et au dénivelé amont assez important. Le relief peut jouer un rôle d'amplificateur : il accentue le soulèvement orographique, et favorise les écoulements rapides vers l'exutoire. Les zones les plus exposées se situent donc à proximité immédiate des reliefs : Cévennes, Alpes-Maritimes, Pyrénées, etc.

Facteurs hydrogéologiques La nature des sols, la couverture végétale, l'humidité initiale et la structure géologique du bassin versant influencent fortement la réponse hydrologique à un événement pluvieux. Les sols peu profonds, ou déjà saturés en eau, favorisent un écoulement rapide et direct, limitant le stockage d'eau dans les sols. Dans certains cas extrêmes, les rendements d'écoulement peuvent atteindre 100% de la pluie tombée [Marchi et al., 2010], tandis qu'en conditions plus classiques, ces taux se situent généralement entre 10 et 30% [Gaume et al., 2016]. Les figures I.5 et I.6 illustrent le rôle de ces facteurs à travers deux cas d'étude [Gaume and Borga, 2013, Payrastre et al., 2019]. D'une part, la cartographie des cumuls de précipitations et des débits de pointe observés lors de la crue d'octobre 2008 dans les bassins versants des Gardons met en évidence une réponse très hétérogène des bassins aux pluies intenses. Ainsi, des débits spécifiques de pointe compris entre 15 et 35 $m^3/s/km^2$ ont été localement enregistrés dans le secteur de l'Amous, un bassin versant dont la réactivité semble supérieure à celle d'autres zones plus touchées par les fortes pluies. Cette différence pourrait s'expliquer par un substratum géologique moins fracturé et altéré, limitant l'infiltration profonde. D'autre part, les hydrogrammes modélisés pour la Nartuby lors des crues de l'Argens en juin 2010 montrent que les débits maximums estimés (traits bleus) correspondent à des rendements d'écoulement estimés à l'aide de la méthode du Curve Number plutôt faibles, traduisant un stockage d'eau important dans le bassin amont, vraisemblablement liée à la présence dominante de formations calcaires très perméables.

Facteurs anthropiques L'imperméabilisation des surfaces (bâtis, voiries) aggrave généralement les écoulements localisés. Par ailleurs, les activités agricoles peuvent altérer significativement la réponse des sols aux pluies : le phénomène de battance, qui correspond à la formation d'une croûte superficielle sous l'action des gouttes de pluie, entraîne une baisse de l'infiltration de l'eau dans le sol et ainsi une augmentation des phénomènes d'écoulements rapides.

1.2.3 Illustrations des dégâts causés par des crues-éclair sur l'arc Méditerranéen français

Les crues soudaines provoquent régulièrement des pertes humaines et des dégâts matériels majeurs. Leur survenue brutale, combinée à des dynamiques hydrologiques rapides et localisées, limite fortement les capacités de réaction des populations et des services de secours. Au-delà de leur intensité, ces événements se distinguent aussi par la diversité des territoires touchés : milieux urbains denses, vallées isolées de montagne, territoires périurbains ou ruraux. Trois événements emblématiques survenus dans

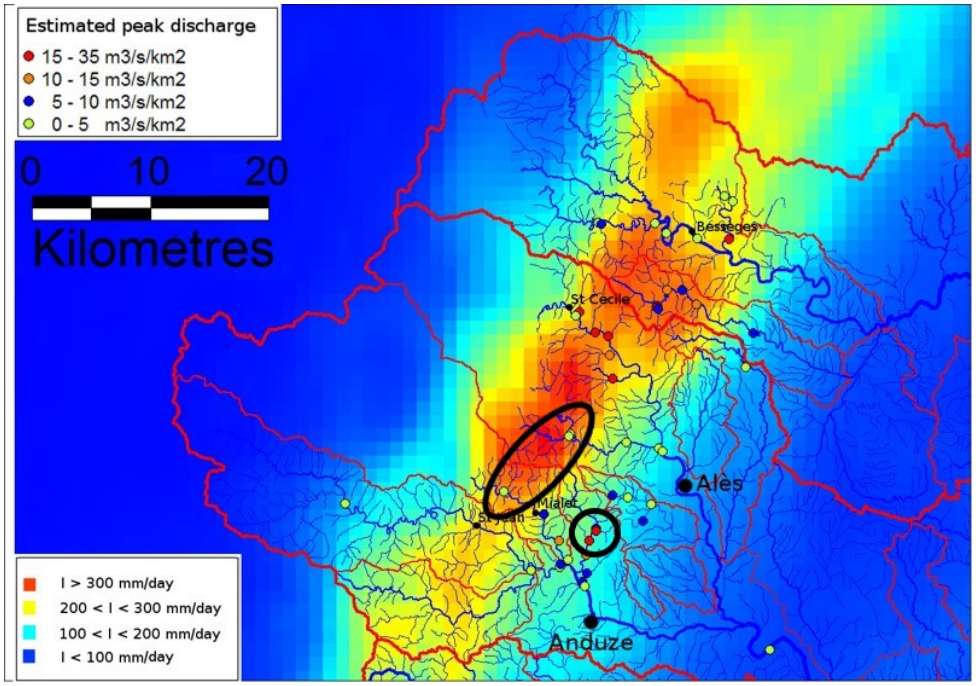


FIGURE I.5 – Répartitions spatiales des cumuls pluviométriques et des débits sur les bassins versants des Gardons lors de la crue d'octobre 2008 [Gaume and Borga, 2013]

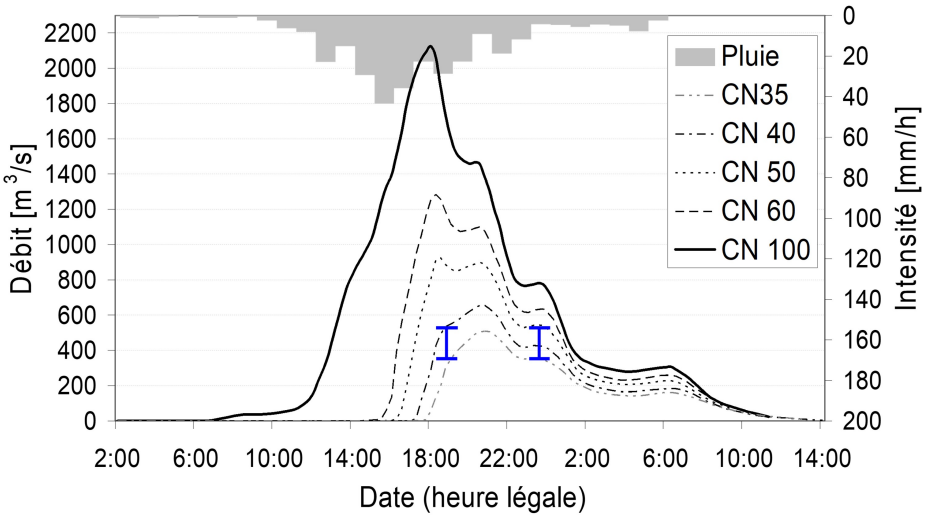


FIGURE I.6 – Crue de la Nartuby à Trans-en-Provence, Le 15 juin 2010 [Payrastre et al., 2019]
L'hydrogramme en trait gras représente la réponse théorique avec un ruissellement de 100%. L'écart avec les débits maximaux estimés (fourchettes bleues) montre que les coefficients de ruissellement sont restés modestes durant cette crue remarquable.

le sud de la France vont être présentés ci-après : ils illustrent cette variété de dommages causés aux infrastructures, aux réseaux, aux activités économiques et au vivant [8].

8. Il est important de noter que les impacts des crues soudaines ne dépendent pas uniquement des facteurs climatiques, topographiques, hydrogéologiques et anthropiques cités précédemment. Leur gravité résulte aussi de conditions sociales et de vulnérabilités spécifiques. Plusieurs travaux, notamment ceux de Ruin et al. [2008], Terti et al. [2015], ont par exemple montré le rôle déterminant de certains facteurs de vulnérabilité dans la survenue d'accidents mortels liés à ces événements.

Nîmes – Octobre 1988 : une ville submergée par ses cadereaux. Le 3 octobre 1988, Nîmes connaît une crue-éclair d'une intensité exceptionnelle. En quelques heures, il tombe plus de 400 mm de pluie sur la ville et son territoire amont - soit l'équivalent de plus de six mois de précipitations habituelles [Fabre, 1990]. Les cadereaux, cours d'eau intermittents drainant le plateau des garrigues au nord de Nîmes et traversant le centre-ville, ont débordé, atteignant des débits spécifiques de l'ordre de $25 m^3/s/km^2$ et transformant les rues en torrents.

Les dégâts sont colossaux : des milliers de logements sont inondés, des dizaines de milliers d'habitants privés d'électricité et de téléphone, des centaines de commerces et d'entreprises sinistrés, des milliers de véhicules endommagés ou emportés. Neuf victimes sont à déplorer. Les infrastructures urbaines sont profondément touchées : voiries éventrées, réseaux d'eau détruits, écoles inutilisables, et transports à l'arrêt [Fabre, 1990].

Malgré le choc, la mobilisation des secours est immédiate. Le plan ORSEC est déclenché dès le matin, coordonné depuis la mairie. Militaires, pompiers, agents communaux, bénévoles, entreprises : plus de 10 000 personnes participent à l'effort de remise en état. Sous l'impulsion du maire de l'époque, la ville se fixe un objectif symbolique : rouvrir Nîmes en huit jours. La solidarité locale, relayée par la création du fonds « Nîmes Solidarité », jouera un rôle clé dans cette reconstruction.



FIGURE I.7 – Illustration des dégâts causés par l'inondation du 3 octobre 1988 de Nîmes © France 3

Aude – Octobre 2018 : une crue éclair aux conséquences lourdes. Le 15 octobre 2018, un épisode pluvieux intense affecte la partie intermédiaire du bassin versant de l'Aude. Plus de 300 mm de pluie sont enregistrés localement en moins de 24 heures et plusieurs affluents réagissent violemment : le Lauquet (196 km²) atteint un débit de crue d'environ 880 m³/s, le Trapel (55 km²) dépasse les 300 m³/s, et l'Orbiel (253 km²) atteint 490 m³/s [Hocini et al., 2021]. Les réactions de ces affluents ont causé un fort débordement de l'Aude juste à l'aval de Carcassonne.

L'événement cause d'importants dégâts humains et matériels [CCR, 2019]. Quinze personnes perdent la vie et 99 sont blessées. Le nombre de sinistres déclarés dépasse les 29 000, dont 27 000 dans le département de l'Aude. Près de 19 000 habitations sont touchées, ainsi que 5 700 véhicules, pour un coût estimé à plus de 250 millions d'euros.

Les réseaux sont gravement affectés : plus de 10 000 personnes se retrouvent sans eau potable, 6 000 foyers privés d'électricité. Les infrastructures de transport subissent également des dommages considérables : plus de 1 000 km de routes départementales sont coupées ou détruites, 50 routes sont totalement impraticables le 15 octobre, et plusieurs restent fermées plus d'une semaine. Quatre ponts sont détruits (notamment à Villegailhenc et Trèbes), cinq autres endommagés, et plusieurs lignes ferroviaires sont interrompues.

Plusieurs bâtiments publics sont touchés : 17 collèges endommagés, 800 élèves affectés, et l'hôpital

de Carcassonne, pourtant récent, fortement impacté. Le barrage du lac de la Cavayère présente également des dégradations structurelles. Sur le plan agricole, des vergers et des parcelles de vignoble sont détruits dans plusieurs secteurs.



FIGURE I.8 – Illustration des dégâts causés par les inondations dans l'Aude en octobre 2018, à Villegailhenc et Trèbes © Sébastien Gominet

Alpes-Maritimes – Octobre 2020 : crues torrentielles et isolement massif liés à la tempête Alex

Le 2 octobre 2020, la tempête Alex génère des précipitations exceptionnelles sur les Alpes-Maritimes, avec plus de 500 mm en 24 heures à Saint-Martin-Vésubie, et jusqu'à 663 mm enregistrés au pluviomètre EDF des Mesches. Cet épisode entraîne la mort de 10 personnes et la disparition de 8 autres dans le département. Les vallées de la Roya, de la Vésubie, et dans une moindre mesure celle de la Tinée, subissent des crues torrentielles particulièrement destructrices, provoquant l'effondrement de nombreux ouvrages et l'isolement de plusieurs communes. D'après Payrastre et al. [2022], le caractère exceptionnel de cet événement tient davantage aux cumuls de pluie sur 24 heures et aux volumes de matériaux mobilisés qu'aux valeurs de débit de pointe observées, qui restent comparables à celles de crues antérieures dans le même secteur.

Les dommages sont importants et multiples [CCR, 2021]. Dans 70 communes reconnues en état de catastrophe naturelle, 470 bâtiments sont gravement endommagés ou détruits, principalement à Saint-Martin-Vésubie, Tende et Roquebillière. Une soixantaine de kilomètres de routes disparaissent partiellement ou totalement, et plus de 60 ouvrages d'art sont touchés, dont 20 ponts détruits. Les réseaux sont aussi sévèrement impactés : 200 km de canalisations d'eau et d'assainissement endommagés ou arrachés, 17 stations d'épuration et plus de 10 installations de traitement d'eau potable hors service. Côté énergie, 31 km de lignes haute tension, 20 transformateurs et 22 km de réseau basse tension sont détruits, affectant jusqu'à 12 500 clients.

Les infrastructures critiques sont également touchées : plusieurs hôpitaux, bâtiments publics, stations d'épuration, gares et lignes ferroviaires sont temporairement inaccessibles ou endommagés. Le musée d'histoire de Saint-Martin-Vésubie est détruit à 98 %. Deux cimetières, à Tende et Saint-Martin-Vésubie, sont ravagés par les crues, emportant plusieurs dizaines de sépultures, dont certaines seront retrouvées sur les côtes italiennes.

Les impacts concernent aussi le milieu naturel et animalier. Le parc Alpha, dédié à la faune sauvage, est en grande partie détruit. Plusieurs loups sont portés disparus, certains retrouvés vivants après plusieurs semaines. Le cheptel d'élevage local est également durement touché, avec la perte de nombreux animaux dans les vallées.

Face à la désorganisation des réseaux terrestres, un pont aérien par hélicoptère est mis en place entre Nice et les communes isolées. Jusqu'à 15 appareils sont mobilisés simultanément pour l'évacuation des habitants, l'acheminement de vivres, et le transport de matériel. Plus de 2000 bâtiments font l'objet d'un diagnostic structurel d'urgence dans les jours qui suivent.



FIGURE I.9 – Illustration des dégâts causés par la tempête Alex en octobre 2020 à Saint-Martin-Vésubie © France Bleu

1.2.4 Les crues soudaines dans un contexte de changement climatique

À l'échelle globale, de nombreux travaux mettent en évidence une intensification des précipitations extrêmes dans le contexte du changement climatique. Depuis les années 1950, une augmentation de la fréquence et de l'intensité des pluies intenses est observée dans la plupart des régions du monde, et cette tendance devrait se poursuivre dans les décennies à venir [Masson-Delmotte et al., 2021]. Cette évolution s'explique en grande partie par l'influence des activités humaines sur le climat. Il a également été montré que les événements de pluie les plus rares devraient s'intensifier plus fortement que les événements courants : pour les pluies centennales, les extrêmes journaliers pourraient augmenter de 13 à 38% d'ici la fin du siècle, contre 10.5% and 28.2% pour les pluies annuelles, d'après une étude effectuée à l'échelle mondiale [Gründemann et al., 2022].

Pour autant, la relation entre intensification des précipitations et augmentation des crues n'est pas systématique. Plusieurs travaux soulignent que les débits de crue n'augmentent pas partout, voire diminuent dans certaines régions, malgré l'augmentation des précipitations extrêmes [Sharma et al., 2018]. Parmi les facteurs pouvant expliquer ce découplage, on retrouve la baisse de l'humidité initiale des sols, la réduction de l'étendue spatiale des pluies, ou encore une moindre contribution de la neige à l'hydrologie de certains bassins [Tramblay et al., 2025].

En parallèle, l'évolution de l'occupation du sol modifie les réponses hydrologiques. Une étude globale menée sur plus de 100 grands bassins fluviaux [Mazzoleni et al., 2022] montre que l'urbanisation contribue aux variations annuelles des surfaces inondées, en complément de la variabilité hydroclimatique.

Selon Dottori et al. [2023], les pertes économiques annuelles liées aux inondations en Europe pourraient atteindre 44 milliards d'euros d'ici 2100 dans un scénario à +3 °C, contre 8 milliards si des stratégies d'adaptation efficaces sont mises en place. Ces travaux montrent que les mesures fondées sur la réduction des pics de crue (zones de rétention) sont parmi les plus efficaces, tant en termes de réduction de l'exposition que de rapport coût-bénéfice.

Sur le pourtour méditerranéen, l'intensification des précipitations extrêmes est également bien documentée. En France, l'analyse de longues séries de mesures montre une augmentation significative des cumuls journaliers les plus élevés : les pluies dépassant 200 mm en une journée sont aujourd'hui deux fois plus fréquentes qu'il y a 50 ans, et l'intensité moyenne des extrêmes a augmenté de +22 % sur la période 1961–2015 [Ribes et al., 2019]. Toutefois, les tendances de précipitations sur le long terme restent débattues. Certaines études font état d'une baisse des cumuls annuels sur les dernières décennies, souvent attribuée à des influences anthropiques, tandis que d'autres soulignent une forte variabilité interannuelle et l'absence de tendance significative à grande échelle [Vicente-Serrano et al., 2025].

Pour ce qui est des crues, les évolutions sont contrastées. Dans de nombreux bassins du sud de la

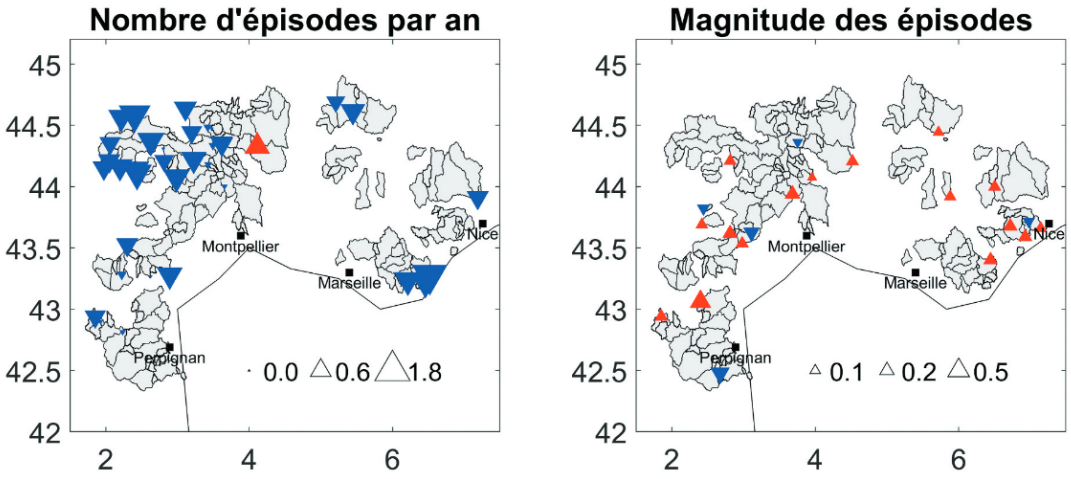


FIGURE I.10 – Tendances observées sur l'occurrence annuelle des débits de crue (gauche) et leur magnitude (droite), en prenant pour seuil le 99ème centile des débits journaliers. Les triangles bleus indiquent une tendance significative à la baisse, les triangles rouges une tendance à la hausse, suivant le test de Mann-Kendall au seuil de significativité de 10%. D'après Tramblay et al. [2019]

France, la fréquence des crues a diminué ces dernières décennies, comme le montre la figure I.10. Cette observation, issue de l'étude de Tramblay et al. [2021], suggère que la baisse de l'humidité des sols pourrait atténuer les effets de l'intensification des pluies extrêmes, entraînant ainsi une diminution de la fréquence des crues. Toutefois, d'autres travaux mettent en évidence une tendance à l'augmentation de la magnitude des crues les plus extrêmes mais une diminution de la magnitude des crues plus courantes [Poncet et al., 2024]. Par ailleurs, une étude menée à l'échelle de la France sur la période 1975-2100 indique que la proportion d'inondations générées par saturation des sols tendrait à diminuer alors que la proportion d'inondations générées par excès d'infiltration lors de précipitations intenses tendrait à augmenter, notamment en zone méditerranéenne [Tramblay et al., 2025].

Les impacts des crues-éclair sont fortement conditionnées par l'évolution de l'occupation du sol. Sur l'arc méditerranéen, la pression foncière et démographique a favorisé l'urbanisation de zones vulnérables [Erdlenbruch and Garcia, 2022]. Pour anticiper l'évolution du risque, il est essentiel de croiser l'analyse des tendances de l'aléa avec celle de l'exposition et de la vulnérabilité. Selon les projections de la CCR, le montant annuel moyen des dommages liés aux crues rapides pourrait augmenter de 6% à 19% à l'horizon 2050 selon les scénarios climatiques du GIEC, en considérant uniquement l'évolution de l'aléa. De plus, la croissance des biens assurés (augmentations attendues du nombre de bien assurés et des valeurs assurées) pourrait entraîner une hausse supplémentaire d'environ 20% [CCR, 2023b].

1.3 Anticiper les crues soudaines : la dernière ligne de défense contre le risque

L'anticipation des phénomènes constitue l'un des piliers essentiels de la prévention des inondations, aussi bien à l'échelle nationale que globale. Une information anticipée sur l'occurrence imminente d'une inondation permet d'actionner les derniers leviers d'action pour limiter les effets de l'aléa sur les territoires. Cependant, le champ lexical lié à l'anticipation (*prévision*, *vigilance*, *avertissement*, *alerte*) peut porter à confusion, ainsi, une première section vise à clarifier ces notions, à la fois en anglais et en français, étant donné que ce manuscrit mobilise les deux langues.

1.3.1 Clarification sémantique : de la prévision à l'alerte

Systèmes d'avertissement – *Warning systems*. Dans ce manuscript, nous nous baserons sur la définition de Kuller et al. [2021], selon laquelle un système d'avertissement repose sur trois composantes clés : le contenu du message, son format, et le canal de communication. Le *contenu* correspond à l'information transmise, incluant la nature du phénomène attendu, son emplacement, son intensité, sa durée, les conséquences anticipées, les incertitudes associées et les actions recommandées. Le *format* peut être textuel, graphique ou cartographique ; le *canal* peut mobiliser des supports variés comme la radio, la télévision, les réseaux sociaux ou les SMS.

Prévision, prévision immédiate – *Forecast*, *nowcast*. Dans ce cadre, la *prévision* (en anglais : *forecast*) désigne l'estimation des conditions futures. En matière d'inondation, il peut s'agir, par exemple, d'estimations de hauteur ou de débit à une échéance donnée au niveau de stations hydrométriques, ou encore de dépassements de seuils de débit. En progressant dans la chaîne de modélisation, les prévisions peuvent également concerner des estimations du nombre de bâtiments impactés ou de personnes affectées par l'inondation. La différence entre les termes *prévision* et *prévision immédiate* sera plus spécifiquement abordée par suite et notamment dans la deuxième partie de l'introduction. On peut néanmoins noter ici que la prévision immédiate se rapporte généralement à des échéances très courtes (de l'ordre de quelques heures) et à des phénomènes à cinétique rapide.

La prévision constitue donc une composante essentielle du contenu d'un message d'avertissement, et elle constitue le cœur de cette thèse. Le terme *avertissement* (ou *warning* en anglais) englobe la prévision, et le raccourci sera parfois fait de parler d'*avertissement* alors qu'il est en réalité question de la prévision sous-jacente.

Enfin, un *hindcast* est un rejeu du modèle permettant de simuler des événements déjà passés. Utilisés principalement comme outils d'évaluation, les *hindcasts* permettent de comparer les résultats de simulation à des données historiques pour évaluer la précision des modèles. Le terme sera rarement utilisé dans ce manuscrit, car il est moins courant dans la littérature, mais il correspond essentiellement à ce qui est proposé dans ce travail.

La Vigilance : une notion française. En France, la *vigilance* désigne un dispositif institutionnel d'information du public, visant à anticiper un risque naturel susceptible de menacer la sécurité des personnes. Deux cartes en constituent les principaux outils : la carte de vigilance météorologique (élaborée par Météo-France) et la carte de vigilance crues (produite par le Service central Vigicrues (ex SCHAPI) et les Services de Prévision des Crues). Ces cartes classent le danger selon une échelle de quatre couleurs (vert, jaune, orange, rouge), et sont mises à jour régulièrement. Elles ne constituent pas une alerte en soi, mais représentent une première étape de sensibilisation.

Le terme *vigilance*, propre au contexte français, n'a pas d'équivalent direct en anglais. Toutefois, les *Flood Guidance Statements* utilisés en Angleterre et au Pays de Galles constituent un produit s'en rapprochant [Cole et al., 2016].

L'alerte : une responsabilité de l'autorité de police. Dans la littérature internationale, le terme *alert* est parfois utilisé de manière interchangeable avec *warning*, bien qu'il soit moins fréquent. En France, en revanche, le terme *alerte* revêt une signification spécifique : il renvoie à une décision relevant de l'autorité de police. Elle peut être déclenchée par un maire ou un préfet, sur la base des informations techniques reçues, et selon le contexte local. L'alerte vise à enclencher une réponse opérationnelle (activation du

Plan Communal de Sauvegarde, envoi de messages via FR-Alert, mise à l'abri, etc.) afin de protéger concrètement la population.

La figure I.11 synthétise les liens entre ces différentes composantes dans le contexte français. Elle illustre la circulation de l'information, depuis les producteurs de prévisions jusqu'aux citoyens, en passant par les niveaux préfectoral et communal. Ce schéma met en évidence le découplage entre l'information scientifique (*prévision*, *vigilance*) et la décision politique (*alerte*), ainsi que les responsabilités spécifiques associées à chaque acteur.

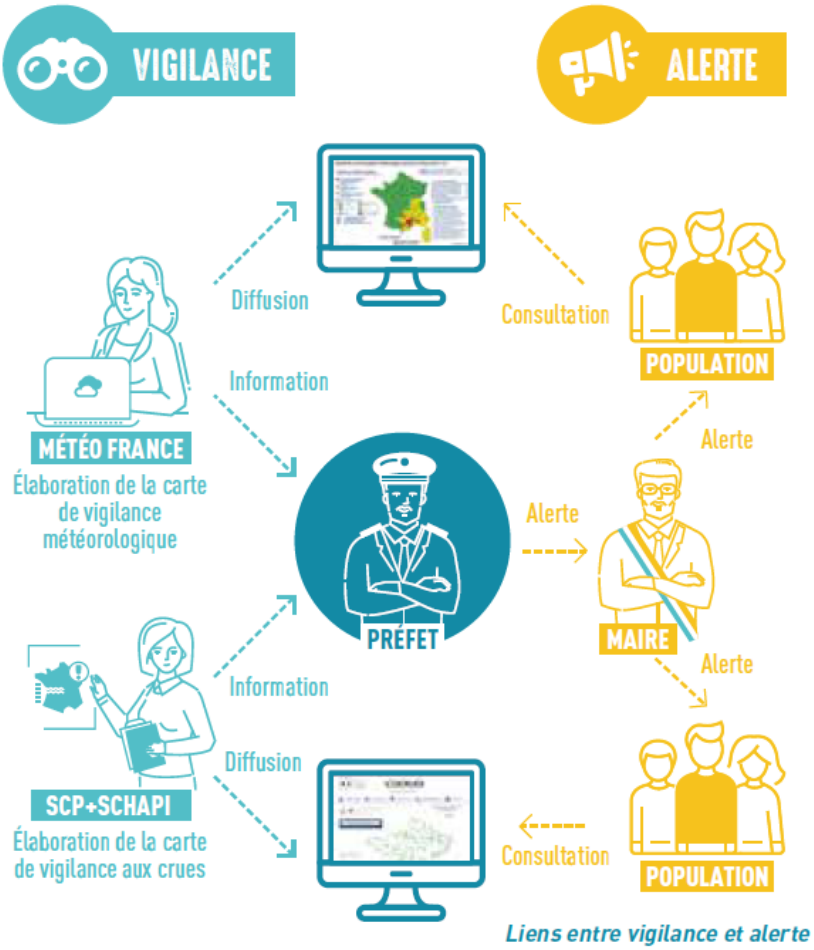


FIGURE I.11 – Liens entre vigilance et alerte. Source : Centre Européen de Prévention du Risque d'Inondation

1.3.2 Mise en place de systèmes de prévisions des crues en France

Dans de nombreux pays, les systèmes et services dédiés à la prévision des inondations se sont développés de manière réactive, à la suite de catastrophes ayant révélé les failles des systèmes existants. Ces événements ont souvent servi de déclencheurs politiques, techniques et institutionnels. Au Royaume-Uni, les inondations généralisées de l'été 2007 ont été à l'origine de la mise en place du Flood Forecasting Centre (Angleterre et Pays de Galles) et du Scottish Flood Forecasting Service en Écosse [Cole et al., 2016]. Au Brésil, les inondations et coulées de boue meurtrières survenues en 2011 dans la région de Rio de Janeiro ont motivé la création du Centro Nacional de Monitoramento e Alertas de Desastres Naturais (CEMADEN) [de Moraes, 2023].

En France également, l'organisation de la surveillance des cours d'eau et la prévision des crues est le fruit d'une construction progressive, largement marquée par des événements traumatiques ayant

mis en évidence les lacunes des dispositifs existants. Si les premières réflexions remontent au XIXe siècle, à la suite des crues majeures de 1840, 1846 et surtout 1856, ayant notamment occasionné de graves inondations dans le bassin de la Loire, il faut attendre les années 1980 pour voir émerger une véritable structuration de la prévision hydrologique à l'échelle nationale. Cette décennie est marquée par des avancées scientifiques en hydrologie statistique, en modélisation, et par l'intégration croissante des radars météorologiques, mais aussi par une prise de conscience politique réactivée à chaque événement d'ampleur.

Le cadre institutionnel se renforce progressivement, avec la réorganisation des Services d'annonce de crues (SAC) en 1984. Cependant, c'est à la suite des crues meurtrières de l'Ouvèze (1992), de l'Aude (1999) et du Gard (2002) qu'un tournant s'opère. Dès 2003, l'État fonde le SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations, aujourd'hui appelé SCV pour Service Central Vigicrues), bras technique et scientifique du ministère de l'Écologie. Le réseau national de prévision des crues se consolide autour de trois piliers : les Services de Prévision des Crues (SPC), les unités d'hydrométrie et les cellules de veille hydrologique. En parallèle, la base de données hydrométriques nationale (Hydroportail, https://hydro.eaufrance.fr/) évolue, alimentée par près de 3000 stations de mesure actives.

En 2006, le lancement de la carte de vigilance crues, accessible via le site www.vigicrues.gouv.fr, marque un jalon important dans la communication de la prévision des crues auprès du grand public. Elle marque aussi le passage officiel de l'annonce à la prévision des crues, avec un objectif d'anticipation fixé à 24h sur les cours d'eau principaux. Inspirée du système de vigilance météorologique de Météo-France, cette carte vise à informer quotidiennement les autorités, les médias et le grand public sur le risque de crue à une échéance de 24 heures, à l'aide d'une échelle de couleur allant du vert (pas de vigilance particulière) au rouge (risque de crue majeure), comme illustré en figure I.12.

Les niveaux de vigilance sont définis, dans un premier temps, à partir de l'intensité des phénomènes physiques attendus sur les tronçons surveillés ; ils sont ensuite modulés - à la hausse ou à la baisse - en fonction des enjeux identifiés sur les territoires concernés (population, biens, infrastructures, activités économiques). Ce dispositif ne couvre toutefois que les tronçons instrumentés des principaux cours d'eau, soit environ 23 300 km de linéaire en 2025, ce qui représente une couverture d'environ 75% de la population vivant en zone inondable. En revanche, de nombreux petits bassins versants, notamment en amont des cours d'eau surveillés, demeurent en dehors du périmètre de surveillance.

1.3.3 Mise en place de systèmes dédiés aux zones non couvertes par la Vigilance Crues

Surveiller l'ensemble du territoire face au risque de crues soudaines constitue un défi majeur, notamment en raison du grand nombre de cours d'eau non jaugés, c'est-à-dire dépourvus de stations de mesure. Si des systèmes de prévision calibrés localement, s'appuyant sur des données fines et une connaissance approfondie du territoire, peuvent être mis en place à petite échelle (comme c'est le cas à Nîmes [Pla et al., 2019]) leur déploiement généralisé serait extrêmement coûteux et difficile à maintenir. En pratique, ces dispositifs locaux restent donc exceptionnels.

Ce constat s'est imposé avec force en 2010, à la suite de deux événements dramatiques survenus à quelques mois d'intervalle : la tempête Xynthia, en février, qui a provoqué une submersion marine meurtrière sur la côte atlantique, notamment à La Faute-sur-Mer (Vendée), et les inondations de juin dans le Var, ayant touché gravement la ville de Draguignan et ses alentours. En réaction, les pouvoirs publics ont lancé en 2011 le Plan Submersions Rapides, fixant de nouvelles priorités centrées sur les phénomènes à cinétique rapide.

Deux services complémentaires ont alors été développés à l'échelle nationale : APIC (Avertissement



FIGURE I.12 – Carte de vigilance crue du 1er avril 2024, à 10h. Source : https://www.vigicrues.gouv.fr/

Pluie Intense à l'échelle des Communes), mis en place par Météo-France en 2011, et Vigicrues Flash, piloté par le SCV, et opérationnel depuis 2017 [Piotte et al., 2020]. Ce dernier couvre les petits bassins versants à temps de réponse court (généralement entre 3h et 6h), grâce à une modélisation hydrologique automatique alimentée par les lames d'eau radar en temps réel. Plus de 30 000 km de linéaires sont ainsi surveillés, avec des avertissements envoyés toutes les 15 minutes aux collectivités abonnées.

Vigicrues Flash constitue une avancée notable dans la couverture des crues soudaines. Ce dispositif repose sur une modélisation hydrologique qui estime les débits à partir des pluies observées, puis les compare à des seuils de référence pour attribuer un niveau de vigilance : « crue forte » ou « crue très forte » [Javelle et al., 2016, Piotte et al., 2020] (voir figure I.13). Le fonctionnement détaillé de ce système sera présenté au chapitre II. Une caractéristique essentielle de Vigicrues Flash mérite toutefois d'être

soulignée : contrairement à la Vigilance crues déployée sur les grands cours d'eau, ce système entièrement automatisé ne prend actuellement en compte, à aucun moment du calcul, les enjeux susceptibles d'être affectés sur le territoire. Cette limite - qui a motivé l'orientation du sujet de thèse - sera largement discutée dans la seconde partie de l'introduction.

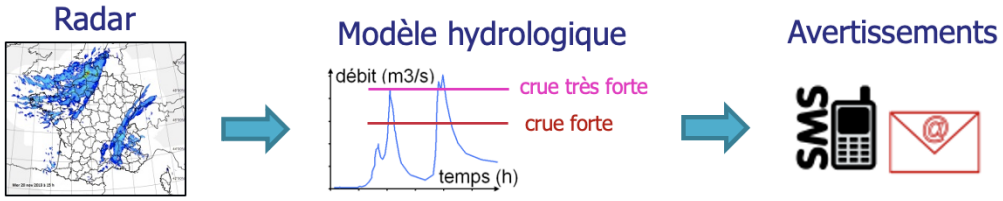


FIGURE I.13 – Fonctionnement simplifié de Vigicrues Flash. Source : Service central Vigicrues

Une autre limite importante de Vigicrues Flash réside dans sa faible capacité d'anticipation : le délai d'avertissement est inférieur à 3 heures dans environ 65% des cas, selon une première estimation (voir chapitre II). Cela s'explique principalement par le fait que le système repose sur des données d'observation de pluie — les lames d'eau radar — en entrée du modèle hydrologique, sans recourir à des prévisions de pluie qui permettraient d'allonger les horizons de prévision. Toutefois, l'intégration de prévisions soulève plusieurs défis, tant sur le plan technique que méthodologique. Il convient notamment de choisir des produits adaptés (modèles numériques de prévision du temps, extrapolations radar, ou approches hybrides [Poletti et al., 2019, Lovat et al., 2022]), et de prendre en compte l'incertitude intrinsèque à ces produits. Cette incertitude peut être mieux représentée grâce aux prévisions d'ensemble [Descamps et al., 2015, Charpentier-Noyer et al., 2023], mais leur interprétation et leur intégration opérationnelle posent encore des difficultés.

2 Prévision immédiate des impacts des crues soudaines

2.1 Structuration par blocs des chaînes de prévision des crues

Pour limiter les pertes humaines et matérielles liées aux crues, la mise en œuvre de systèmes de prévision et d'avertissement s'est fortement développée au cours des dernières décennies. Cette dynamique s'observe aussi bien dans les publications scientifiques que dans les initiatives opérationnelles, qu'elles soient locales (à l'échelle d'un bassin versant), nationales (services de prévision hydrométéorologique), ou supranationales (systèmes continentaux voire globaux).

Les dispositifs d'avertissement reposent généralement sur des chaînes de prévision combinant plusieurs blocs de modélisation : prévision des pluies, modélisation hydrologique, simulation hydraulique, et anticipation des impacts (cf. figure I.14).

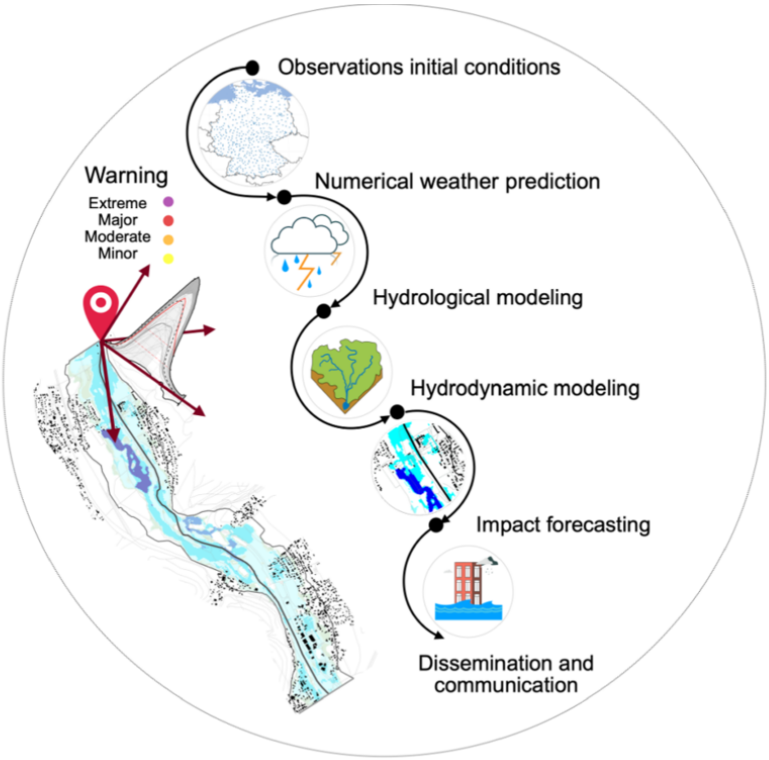


FIGURE I.14 – Structure d'une chaîne de prévision des impacts des inondations [Najafi et al., 2024].

Néanmoins, tous ces blocs ne sont pas nécessairement présents dans les dispositifs opérationnels. Certains systèmes émettent ainsi des avertissements aux crues soudaines directement à partir des prévisions de pluie [Corral et al., 2008, 2019, Shamir et al., 2013, Clark et al., 2014]. À l'inverse, Vigicrues Flash en France intègre un modélisation hydrologique pluie-débit, mais repose uniquement sur des observations de pluie. Dans les sous-sections suivantes, nous présentons successivement chacun des blocs de modélisation qui peuvent être intégrés à une chaîne de prévision des crues, en mettant en évidence leurs principes, leurs apports et leurs limites dans le cadre spécifique de la prévision des crues soudaines.

2.1.1 La prévision immédiate des pluies

La prévision immédiate des pluies, ou *nowcasting*, constitue généralement le premier bloc des chaînes de prévision des crues soudaines. Les extrapolations radar, basées sur des approches dites lagrangiennes, consistent à identifier les structures pluvieuses à partir des observations radar, puis à prolonger leur trajectoire par simple advection en utilisant les vitesses et directions de déplacement estimées sur le passé

récent [Berenguer et al., 2011, Silvestro and Rebora, 2012]. Ces méthodes offrent de bonnes performances jusqu'à 2–3 heures d'échéance, mais ne permettent pas de simuler la genèse, la dissipation ou l'évolution des systèmes convectifs [Mandapaka et al., 2012]. À l'inverse, les modèles numériques de prévision (*numerical weather prediction models*) permettent une représentation explicite de ces processus, et peuvent allonger les délais d'anticipation jusqu'à 24–48h [Zanchetta and Coulibaly, 2020]. Cependant, ils sont généralement moins performants à très court terme. Certaines versions à rafraîchissement rapide, comme AROME-NWC en France, ont été spécifiquement développées pour améliorer la prévision immédiate [Auger et al., 2015].

Pour tirer parti des atouts respectifs de ces deux approches, des produits hybrides ont été développés, combinant les extrapolations radar et les sorties des modèles numériques via des techniques de fusion (*blending*, voire figure I.15). Cela permet d'assurer une transition "sans couture" (*seamless*) entre les champs pluviométriques observés et prédits [Poletti et al., 2019, Lovat et al., 2022]. En France, c'est le cas du produit opérationnel PIAF [Moisselin et al., 2019], qui est actuellement utilisé par Météo-France pour la prévision immédiate.

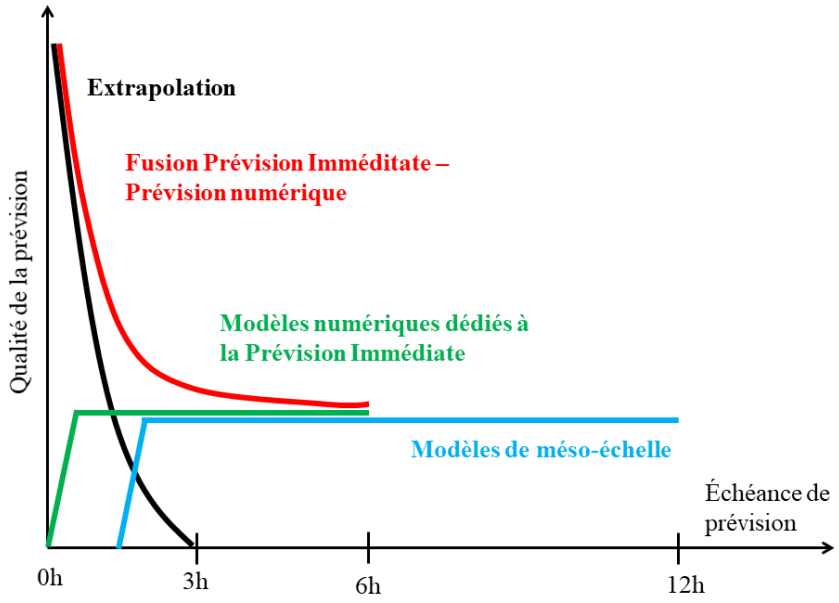


FIGURE I.15 – Qualité des prévisions de pluie en fonction des échéances [Lovat, 2019].

Néanmoins, les modèles numériques de prévision présentent encore des limites importantes, notamment une localisation imprécise des cellules convectives intenses, particulièrement problématique dans le cas des petits bassins versants [Vincendon et al., 2011]. Ces erreurs résultent de plusieurs sources d'incertitude : les conditions initiales et aux limites des modèles, les paramètres physiques des modèles, et les conditions de surface comme la température de la mer ou des sols [Lovat, 2019]. Pour tenir compte de ces incertitudes, l'approche ensembliste consiste à simuler plusieurs scénarios météorologiques, chacun correspondant à une hypothèse plausible sur l'état initial ou les conditions aux limites du système atmosphérique [Bowler et al., 2006, Descamps et al., 2015, Bouttier and Raynaud, 2018]. Plusieurs études, dont celle réalisée par Roux et al. [2020], ont montré la plus-value des approches ensemblistes par rapport aux approches déterministes pour la prévision des crues soudaines.

2.1.2 Les prévisions hydrologiques

Un grand nombre de systèmes d'avertissement aux crues soudaines s'appuient sur des modèles pluie-débit, dont le rôle est de simuler la réponse hydrologique de petits bassins versants, souvent non jaugés. Les

équations gouvernant ces modèles peuvent être de trois types : des relations issues des lois de la physique (modèle à base physique), des relations conceptuelles, qui sont des équations mathématiques mais non issues des lois de la physique (modèle conceptuel), et des relations d'entrées-sorties, faisant par exemple appel à l'intelligence artificielle (modèle "boîte-noire" ou modèle empirique). Les modèles pluie-débit peuvent également être distingués par d'autres critères. Ils peuvent être globaux (le bassin versant est une entité unique), semi-distribués (subdivision en sous-bassins homogènes) ou distribués (l'espace est discrétisé par un maillage : exemple en figure I.16). On peut également parler de modèle événementiel (la simulation s'effectue sur des événements distincts) ou bien de modèle continu (la simulation est faite sur une période temporelle plus longue et continue, ce qui permet de représenter les cycles successifs d'assèchement et d'humidification des sols du bassin versant).

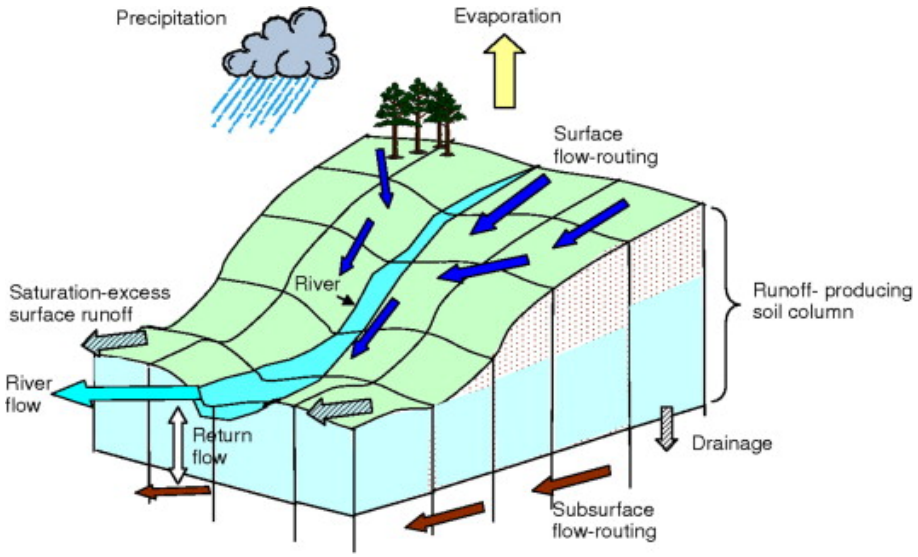


FIGURE I.16 – Illustration d'un modèle hydrologique distribué [Cole and Moore, 2009].

Les modèles sont généralement alimentés par des forçages météorologiques, tels que les précipitations et l'évapotranspiration mesurées et prévues, ainsi que par des informations topographiques, géologiques et pédologiques décrivant les caractéristiques des bassins versants. Le calage des paramètres repose souvent sur des séries temporelles de débits estimés à partir des chroniques de hauteur d'eau enregistrées aux stations hydrométriques (via des courbes de tarage).

Dans le cadre de la prévision des crues soudaines, les modèles hydrologiques utilisés sont souvent distribués ou semi-distribués, car un maillage fin permet de mieux représenter les effets et processus liés à la forte variabilité spatio-temporelle des précipitations. C'est le cas, par exemple, du modèle MARINE développé par Roux et al. [2011], du modèle ISBA-TOPMODEL utilisé dans la thèse de Edouard [2016], ou encore du modèle CINECAR dans celle de Charpentier-Noyer [2022]. Ce choix se retrouve également dans les systèmes opérationnels : le système AFAS développé pour la Nouvelle-Zélande [Cattoën et al., 2022] repose ainsi sur le modèle TopNet [Clark et al., 2008], un modèle physique, continu et semi-distribué. On peut aussi citer le système FLASH aux États-Unis [Gourley et al., 2017], qui intègre notamment le modèle CREST [Wang et al., 2011], ou encore le modèle Grid-to-Grid utilisé au Royaume-Uni [Cole et al., 2016] pour la production des Flood Guidance Statements. Ces deux derniers sont des modèles hybrides, combinant approches physiques et conceptuelles, appliqués à l'échelle de la maille kilométrique, correspondant généralement à la résolution des données de pluie utilisées en entrée.

En France, le système Vigicrues Flash [Javelle et al., 2016, Piotte et al., 2020] s'appuie actuellement sur la plateforme de modélisation hydrologique SMASH (Spatially-distributed Modelling and ASsimilation for Hydrology) [Colleoni et al., 2025]. Cette plateforme met en œuvre des modèles de type "réservoirs"

issus ou dérivés de la famille GR ("Génie Rural", Perrin et al. [2003]). Dans sa version opérationnelle, Vigicrues Flash utilise un modèle de type GR-Distribué, fonctionnant de manière continue, avec une résolution spatiale de 1 km et un pas de temps de 15 minutes. Ce modèle sera présenté plus en détail dans le chapitre II du manuscrit.

2.1.3 La prévision des zones inondées

Contrairement à la prévision hydrologique, qui vise à estimer le débit ou la hauteur d'eau dans les cours d'eau, la prévision des zones inondées cherche à déterminer l'emprise spatiale des inondations, et les hauteurs d'eau associées, en tout point du domaine de modélisation, ce qui est crucial notamment dans le cas de crues débordantes dans lesquelles les cours d'eau envahissent une partie de leur lit majeur.

Établir une cartographie des zones inondées nécessite d'intégrer un bloc de modélisation supplémentaire, qui peut être plus ou moins complexe. Les approches les plus simples reposent uniquement sur la topographie, via des méthodes dites de remplissage de MNT (*bathtub modeling*) qui combinent des formules hydrauliques simples et des descripteurs topographiques tels que la méthode HAND (Height Above Nearest Drainage, Nobre et al., 2011). Bien que rapides et faciles à mettre en œuvre, ces méthodes sont aujourd'hui largement critiquées, et plusieurs travaux ont montré que les approches de modélisation de type plus physique permettent de réduire les biais de modélisation [Hocini et al., 2021, Sanders et al., 2024].

Une deuxième famille de modèles utilisés pour la prévision des zones inondées sont les modèles hydrauliques à base physique, dont les équations différentielles représentant les bilans de masse et de quantité de mouvement dans un écoulement à surface libre. Il s'agit précisément des équations de Navier-Stokes, simplifiées dans différentes versions des équations de Barré de Saint-Venant. La modélisation hydraulique 1D, qui suppose un écoulement unidimensionnel, consiste à résoudre ces équations selon l'axe principal du cours d'eau. Elle a été largement utilisée pour la prévision des crues, avec des avantages notables tels que des temps de calcul relativement courts pour un bon niveau de précision [Le Bihan et al., 2017, Lamichhane and Sharma, 2018]. Toutefois, ces modèles présentent des limites similaires aux approches topographiques : ils reproduisent mal les écoulements complexes sur les plaines inondables, bien qu'ils restent attractifs par leur simplicité et leur rapidité. À l'inverse, des modèles tridimensionnels (3D), bien plus complexes, peuvent être explorés pour dépasser ces limites [Luo et al., 2018], mais leur usage opérationnel reste très limité, en raison de leur complexité et du manque de données de calibration [Bates, 2022]. Le choix du type de modèle hydraulique dépend donc des objectifs visés (échelle de modélisation, niveau de précision attendu, contraintes de temps de calcul) [Laganier et al., 2014]. Dans la pratique, les modèles 2D sont souvent considérés comme un bon compromis, et de nombreuses études ont mis en évidence leur plus-value pour des applications en prévision des inondations. Par exemple, les travaux de Hocini et al. [2021] ont montré la supériorité d'une modélisation hydraulique 2D (Floodos [Davy et al., 2017]) par rapport à la méthode topographique HAND [Nobre et al., 2011] et à la méthode caRtino 1D [Pons et al., 2014], pour la modélisation de crues soudaines sur des bassins méditerranéens français.

Le développement de ces approches permettant de cartographier les zones inondées est soutenu par l'amélioration continue des données d'entrée, en particulier des MNT à haute résolution. Les progrès récents en acquisition LiDAR ont renforcé la qualité et la couverture des données topographiques, avec par exemple 99% du territoire anglais couvert en LiDAR à 1m [{Environment Agency}, 2024], et le projet Lidar HD qui vise une couverture complète de la France à 0,5m [IGN, 2024]. En parallèle, plusieurs travaux portent sur la reconstruction du lit mineur, encore largement absent des MNT disponibles, bien qu'il joue un rôle central dans les dynamiques d'écoulement [Lague and Feldmann, 2020, Pricope and

Bashit, 2023, Frizzle et al., 2024].

À l'échelle globale ou continentale, des simulations hydrauliques 2D ont été réalisées à des résolutions relativement grossières, de l'ordre de 100m. C'est le cas des cartes d'inondation européennes produites par Alfieri et al. [Alfieri et al., 2014, Dottori et al., 2017], basées sur le modèle LISFLOOD [Bates et al., 2010], ou des Global Flood Maps réalisées par Fathom, initialement à 90m [Sampson et al., 2015], puis affinées à 30m dans leur version la plus récente [Wing et al., 2023]. Cependant, plusieurs travaux soulignent que ces résolutions sont peu adaptées à la représentation de crues locales, notamment dans les petits bassins versants ou les zones urbaines [Fleischmann et al., 2019]. Des efforts croissants visent donc à développer des simulations à des résolutions inférieures à 10m, notamment 5m dans le cas de travaux récents sur les crues éclairs [Hocini et al., 2021].

Dans un contexte de prévision temps réel, les modèles hydrauliques 2D se heurtent toutefois à une contrainte majeure : leur coût en temps de calcul. Simuler un événement sur un périmètre étendu peut nécessiter plusieurs heures de calcul, ce qui est difficilement compatible avec les exigences opérationnelles. Pour cette raison, les chaînes de prévision s'appuient aujourd'hui presque systématiquement sur des catalogues de simulations (*flood libraries*) : des bases de données de scénarios pré-calculés pour différentes hypothèses de débit, de période de retour ou de hauteur d'eau à une station donnée [Dottori et al., 2017, Alfieri et al., 2024, Mosimann et al., 2024]. Cette approche, illustrée en figure I.17, permet d'accéder rapidement à une simulation réaliste en fonction des prévisions hydrologiques.

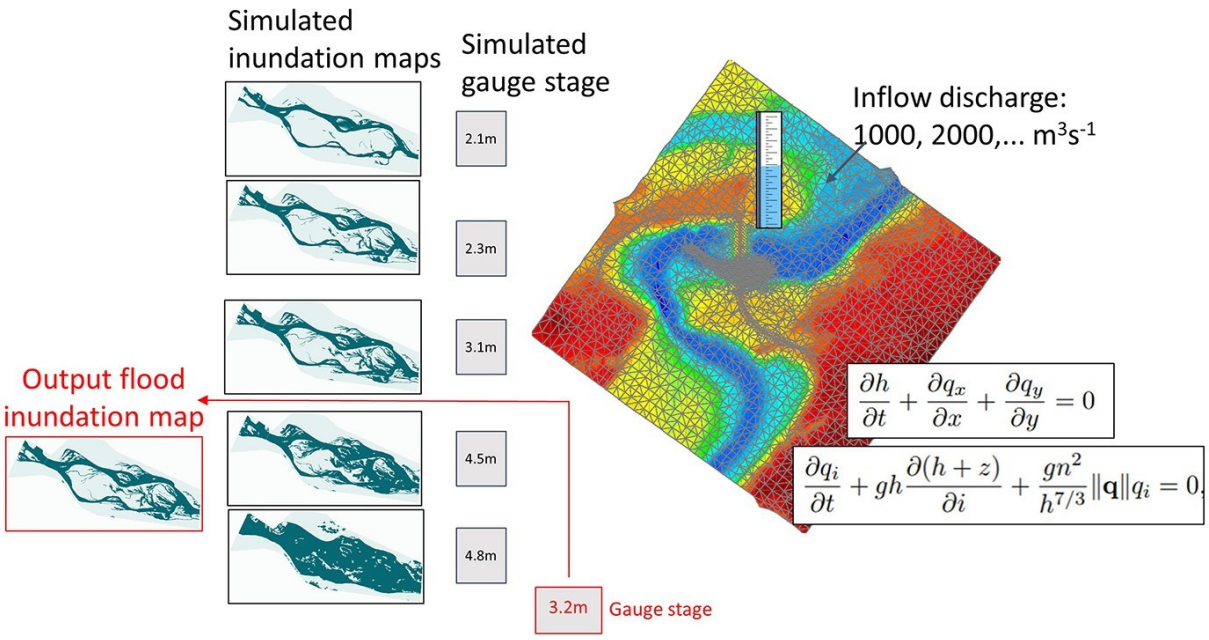


FIGURE I.17 – Illustration d'une approche de prévision des inondations par utilisation d'un catalogue de scénarios, estimés par modélisation hydraulique 2D (équations de Saint-Venant avec hypothèse de stationnarité) [Nevo et al., 2022]

Ces cartes peuvent être produites de manière encore plus rapide en ajoutant des hypothèses simplificatrices au modèle hydraulique 2D, notamment l'hypothèse de régime permanent. Le résultat correspond alors à une photographie statique de la crue supposant que le débit de pointe s'est maintenu sur une longue durée, ce qui peut être pertinent pour estimer un aléa maximal, mais empêche de représenter la dynamique temporelle de l'inondation (temps de montée, de propagation et de décrue). En particulier, cette hypothèse fait abstraction des délais de remplissage des lits majeurs, ce qui peut conduire, dans certains cas, à une surestimation des emprises inondées. Une autre simplification couramment adoptée consiste à négliger les termes inertiels (liés à l'accélération de l'eau) dans l'équation de Saint-Venant, lorsque le nombre de

Froude est relativement faible. Cela permet de réduire considérablement les temps de calcul, mais peut introduire des erreurs, notamment dans les zones où l'écoulement change brusquement de direction ou à proximité d'obstacles, où les effets d'inertie jouent un rôle important.

Enfin, des travaux récents explorent la possibilité de réaliser des simulations hydrauliques "quasi-temps-réel", grâce à l'optimisation des modèles et à l'augmentation des capacités de calcul [Najafi et al., 2024]. Des approches basées sur l'intelligence artificielle commencent à être explorées, toutefois, leur mise en œuvre reste encore limitée dans les bassins non jaugés, en raison du manque de données d'apprentissage [Nevo et al., 2022].

2.1.4 La prévision des impacts

Les approches visant à estimer les impacts associés aux crues dans une chaîne de prévision sont variées. Dans les blocs de modélisation évoqués précédemment - précipitations, débit, zones inondées - les variables sorties des modèles sont globalement homogènes d'un système à l'autre, et ne posent pas question. En revanche, au niveau du bloc de modélisation des impacts, les possibilités sont multiples, tant du point de vue des méthodes de calcul mobilisées que des types de résultats visés.

Tableau I.3 – Exemples de catégorisation des impacts dans trois systèmes d'avertissement [Cole et al., 2016, Schroeder et al., 2016, Ritter et al., 2021]

Flood Guidance System (UK)		Flash Flood Severity Index (USA)		ReAFFINE (Europe)	
Catégorie	Impact	Catégorie	Impact	Catégorie	Impact
Mineur	• 1–100 habitations • 1–2 services communautaires ou réseaux (eau, électricité. . .) • 1–20 bâtiments commerciaux	Mineur	Débordement de rivière/ruisseau, inondation de jardins, sous-sols ou terres agricoles	Faible	1–99 personnes exposées/km²
		Modéré	Inondations de rues et de routes, routes fermées		
Significatif	• 1–100 habitations • > 2 services communautaires ou réseaux (eau, électricité. . .) • > 20 bâtiments commerciaux • 5m de réseau routier ou ferroviaire	Sérieux	Véhicules, maisons et bâtiments inondés, ponts endommagés	Modéré	100–999 personnes exposées/km²
Sévère	> 100 bâtiments résidentiels	Sévère	Véhicules et maisons mobiles emportées	Fort	+1000 personnes exposées/km²
		Catastrophique	Bâtiments et infrastructures submergés / détruits		

Une première approche consiste à combiner directement, par superposition spatiale, les enjeux exposés et les cartes d'emprise inondée prédites, puis à définir des catégories d'intensité des impacts pour faciliter la communication du résultat. Des exemples sont donnés dans le tableau I.3 : premièrement pour le Flood Guidance System appliqué au Royaume-Uni [Cole et al., 2016], qui évalue les impacts à l'échelle de mailles de 1 km^2 ; ensuite pour le Flash Flood Severity Index, conçu pour compléter les Flash Flood Guidances aux États-Unis [Schroeder et al., 2016], qui propose une catégorisation plus qualitative ; et finalement pour la méthode ReAFFINE qui propose, à l'échelle européenne, des prévisions sur la population exposée aux crues soudaines (des détails sur ce calcul sont fournis en figure I.18). Ce tableau

illustre la grande diversité des méthodes de classification des impacts dans les systèmes d'avertissements.

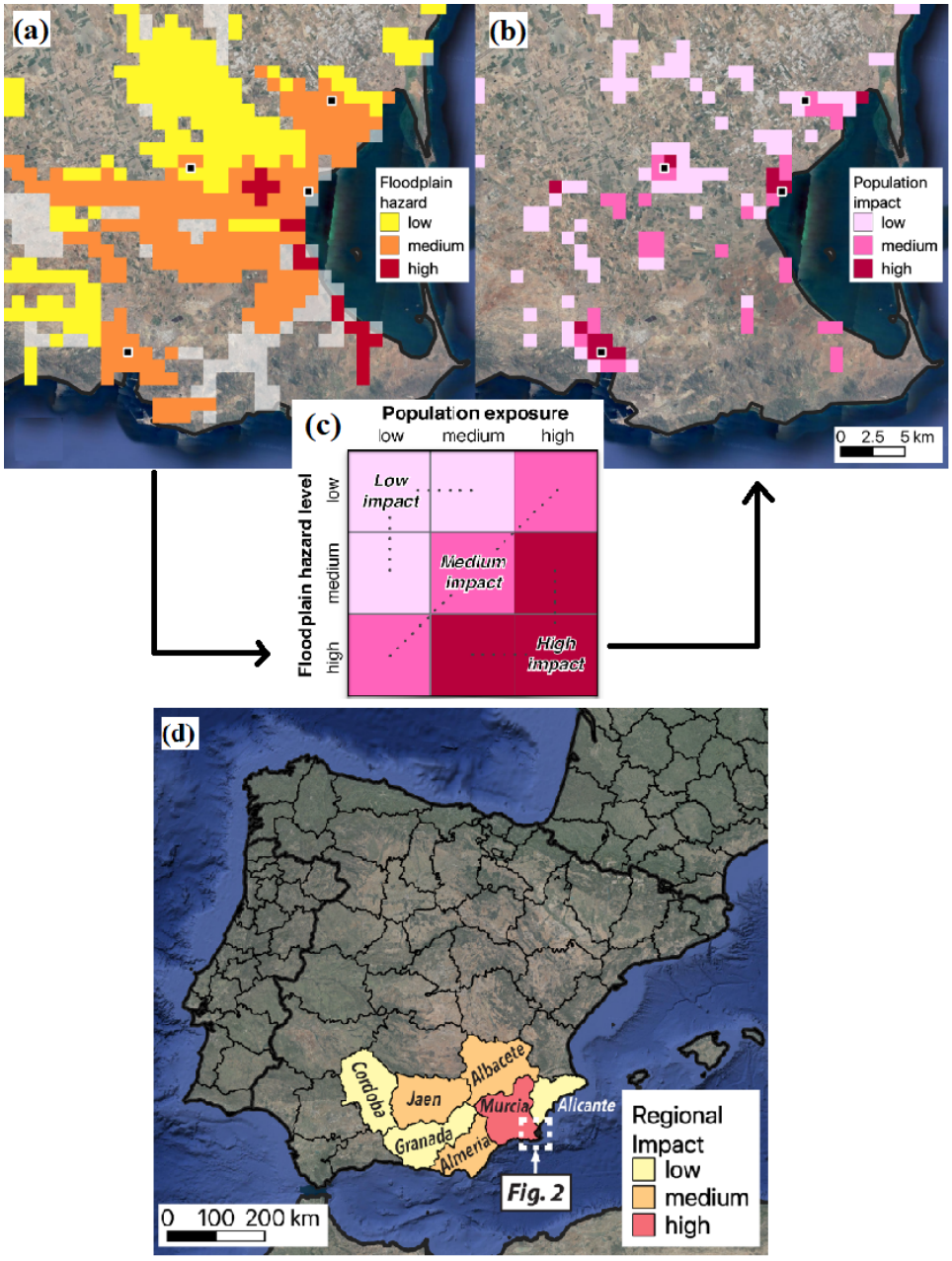


FIGURE I.18 – Méthode d'estimation des impacts dans la chaîne ReAFFINE [Ritter et al., 2021] : (a) affectation d'un niveau d'aléa selon la période de retour sur chaque maille identifiée comme « inondée » ; (b) carte des impacts par maille ; (c) matrice d'impact croisant aléa et exposition ; (d) agrégation régionale.

La notion de vulnérabilité peut également être intégrée au calcul, selon la formule classique vue en début de ce chapitre $Risque = Aléa \times Exposition \times Vulnérabilité$. C'est par exemple la démarche utilisée pour le développement d'un indice de risque dans le cadre du système Vigicrues Flash en France [Saint-Martin et al., 2016], qui repose sur une hiérarchisation des enjeux construite en lien avec les acteurs du territoire (réseaux, bâtiments sensibles, cultures). Ce type d'approche est également utilisé dans des outils comme Flood-PROOFS pour l'Afrique [Alfieri et al., 2024] ou le système du CEMADEN au Brésil [de Moraes, 2023].

D'autre part, plusieurs travaux ont proposé d'estimer des impacts plus spécifiques : nombre de bâtiments affectés [Le Bihan et al., 2017], pertes humaines [Terti et al., 2019a], sécurité des personnes [Silvestro et al., 2019], accessibilité et coupures de routes [Coles et al., 2017, Naulin et al., 2013], pertes économiques [Ritter et al., 2020, Alfieri et al., 2024], ou fonctionnement des services de secours

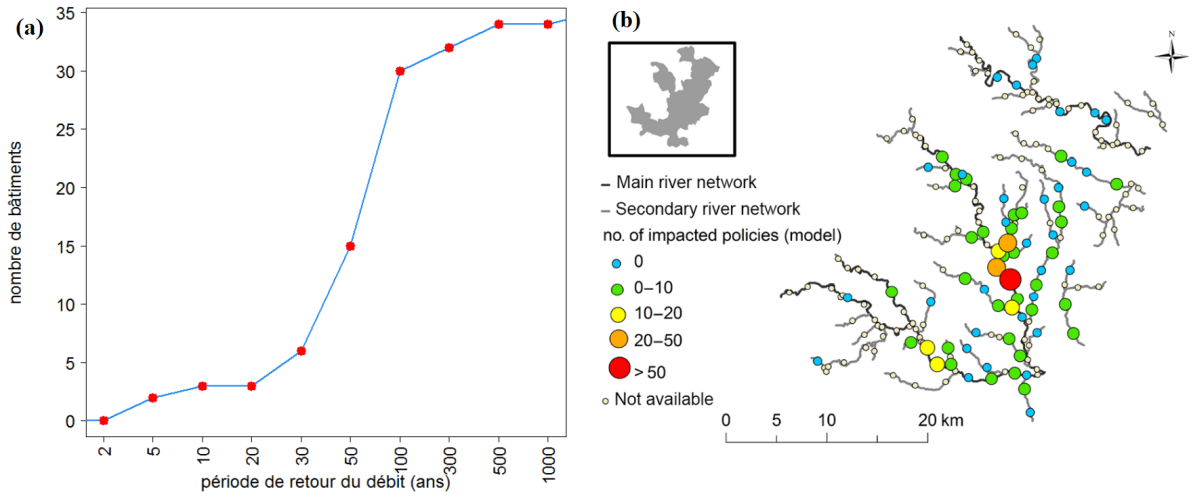


FIGURE I.19 – Méthode de calcul des impacts dans Le Bihan et al. [2017] : (a) relation entre période de retour et nombre de bâtiments affectés ; (b) spatialisation des impacts prédits à l'échelle des tronçons.

[Charpentier-Noyer, 2022].

Il est intéressant de noter que dans certains cas, le calcul d'impacts est réalisé en temps réel, par croisement géospatial des cartes d'aléa prédites et des couches d'enjeux — comme dans les méthodes de Ritter et al. [2021] ou Najafi et al. [2024]. Mais il est également possible de précalculer les impacts en amont, notamment lorsque l'aléa est représenté à partir d'un catalogue de scénarios — comme dans les travaux de Charpentier-Noyer [2022] ou Le Bihan [2016]. La figure I.19 illustre l'approche retenue par Le Bihan et al. [2017], où le nombre de bâtiments inondés est estimé par tronçon de rivière pour un ensemble de scénarios hydrauliques préétablis (périodes de retour de 2 à 1000 ans). Une courbe « période de retour – nombre de bâtiments affectés » est construite par interpolation linéaire (figure I.19a), permettant ensuite de prédire rapidement l'impact associé à une période de retour estimée dans une chaîne de prévision. Ainsi, l'impact est ici estimé sans opération géospatiale, par lecture directe de la courbe en fonction du débit prévu (figure I.19b).

En résumé, la prévision des impacts présente une très grande diversité, tant en termes de résultats visés (types d'impacts considérés, niveau de détail attendu, rendu sous forme de décompte ou de catégorisation), qu'en termes de méthodes mobilisées. Certaines approches reposent sur des comptages directs d'enjeux exposés, d'autres sur des scores intégrant également la notion de vulnérabilité. Le calcul peut être effectué en temps réel ou précalculé à partir de scénarios, et l'intégration dans les chaînes de prévision reste très variable selon les contextes d'application.

2.2 Opportunités et verrous posés par les avertissements basés sur les impacts

2.2.1 Des avertissements basés sur les impacts : un changement de paradigme

Pendant plusieurs décennies, la conception des systèmes d'avertissement s'est fondée presque exclusivement sur la prévision de l'aléa attendu (e.g., intensité de la crue). L'approche dominante consistait à suivre l'évolution de variables physiques (par exemple, le débit ou la hauteur d'un cours d'eau au niveau d'une station, dans le cadre de la prévision des crues), puis à déclencher un avertissement dès qu'un seuil fixé à l'avance était dépassé. Ces avertissements, définis ici comme "Hazard-Based Warnings" (HBW), ont structuré l'architecture de la plupart des dispositifs opérationnels. En France, c'est notamment l'approche retenue pour Vigicrues Flash, qui repose sur la comparaison entre les débits simulés et des seuils de débits correspondant à des périodes de retour prédéfinies (voir figure I.13).

Ces dernières années, un changement de paradigme s'est amorcé avec la montée en puissance des avertissements fondés sur les impacts (Impact-Based Warnings, IBW), soutenus notamment par les recommandations de l'Organisation Météorologique Mondiale [WMO, 2015].

Citation issue de WMO [2015] :

> *In the case of both weather forecasts and warnings, the focus is on what the **weather will be**. These Guidelines advocate evolving from this weather-based paradigm to one that is focused primarily on forecasting impacts. In other words, the focus should evolve to what the **weather will do.***

L'objectif de ces systèmes est d'enrichir l'information transmise en intégrant non seulement la sévérité de l'aléa, mais aussi ses conséquences attendues sur les personnes, les biens ou les infrastructures, en tenant compte des éléments exposés et de leur vulnérabilité. L'approche reste compatible avec une logique par seuils, mais ceux-ci portent désormais sur des indicateurs d'impacts, tels que le nombre de bâtiments endommagés ou de personnes potentiellement affectées.

Une étude menée par Kaltenberger et al. [2020] auprès de 32 services météorologiques et hydrologiques nationaux européens (national meteorological and hydrological services, NMHSs) met en évidence l'évolution des systèmes d'avertissement, passant de dispositifs fondés sur des seuils fixes ou des périodes de retour ("Hazard-Based") vers des systèmes basés sur les impacts (voir figure I.20). Dans cette étude ainsi que dans les recommandations de WMO [2015], une distinction est faite entre les "Impact-based warnings" et les "Impact warnings" : les premiers font une mention claire d'impacts attendus, mais de manière plutôt générique (par exemple : *"De fortes pluies sont attendues. Des inondations localisées pourraient provoquer la fermeture de plusieurs routes secondaires et perturber les déplacements dans la commune"*), tandis que les seconds reposent sur des critères individuels (par exemple : *"En raison des inondations prévues, votre trajet habituel vers l'hôpital risque d'être bloqué demain matin. Prévoir un itinéraire alternatif ou retarder votre déplacement"*, plus proches des "people-centered warnings" évoqués par Potter et al. [2021].

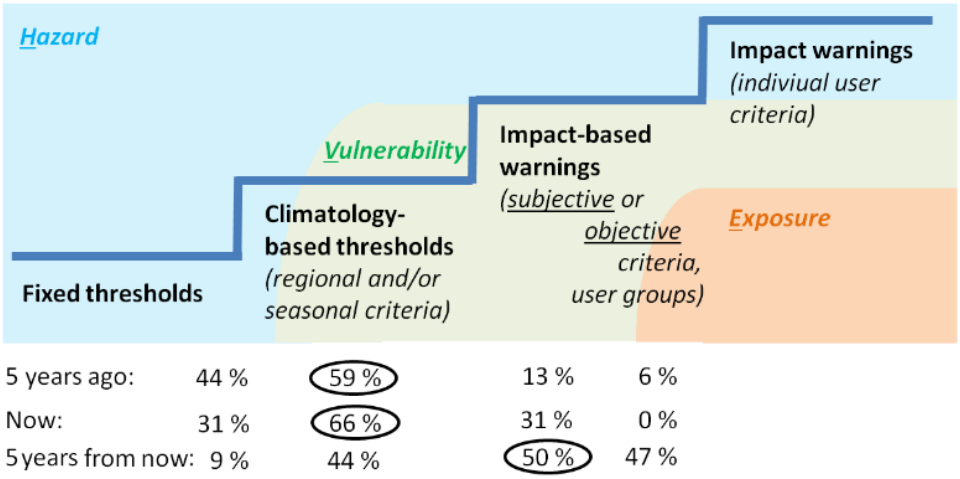


FIGURE I.20 – Évolution des avertissements météorologiques et hydrologiques, intégrant des informations sur l'aléa, la vulnérabilité et/ou l'exposition, dans 32 pays européens il y a 5 ans, aujourd'hui (2018/2019) et dans 5 ans. Les valeurs maximales entourées illustrent l'évolution attendue vers les impacts-based warnings. Extrait de l'étude de Kaltenberger et al. [2020].

2.2.2 Avantages des avertissements basés sur les impacts

Les bénéfices attendus des IBW concernent plusieurs dimensions de la gestion du risque. La première porte sur la communication et la compréhension du risque. En mettant en avant les conséquences concrètes

d'un événement, les IBW permettent de réduire le fossé entre experts et usagers [WMO, 2015, Potter et al., 2021, Yu et al., 2022, Kuller et al., 2021, Meléndez-Landaverde et al., 2024, Potter et al., 2025]. Par exemple, au lieu d'annoncer un débit ou un niveau d'eau attendu, un avertissement pour crue soudaine pourrait indiquer que certains quartiers risquent d'être inondés, avec des habitations, des routes ou des infrastructures critiques touchées.

Un deuxième axe est celui de la gestion opérationnelle des crises. En croisant les aléas prévus avec les enjeux exposés, les IBW permettent d'identifier les zones prioritaires, de hiérarchiser les interventions et de mieux coordonner les secours [Merz et al., 2020, Apel et al., 2022, Le Bihan et al., 2017, Ritter et al., 2022, Saint-Martin et al., 2018, Charpentier-Noyer et al., 2025]. D'autre part, les IBW pourraient moduler les avertissements émis pour les zones à faible enjeu ou peu vulnérables, et ainsi réduire le nombre de fausses alarmes, ce qui permettrait d'augmenter la confiance dans les systèmes d'alerte et de mobiliser plus efficacement les moyens d'intervention [Potter et al., 2021]. Enfin, ils pourraient accélérer la prise de décision en automatisant le croisement entre aléa et enjeux, ce qui est aujourd'hui souvent réalisé manuellement par les services de gestion de crise, avec une part de subjectivité [Ritter et al., 2021]. Plus généralement, l'apport des IBW serait de fournir directement une information utile à la décision [Taylor et al., 2018].

Enfin, un troisième ensemble de bénéfices concerne la planification à long terme et l'intégration multi-aléas. Les IBW invitent à dépasser la logique mono-risque et favorisent l'émergence de systèmes d'avertissement multi-aléas en créant un langage commun basé sur les impacts, qui facilite les échanges entre disciplines traditionnellement cloisonnées [Potter et al., 2025]. De tels systèmes seraient capables de traiter des événements combinés (par exemple, des pluies intenses survenant sur un sol déjà saturé ou fragilisé par une catastrophe récente comme un incendie). Ils incitent également à une meilleure prise en compte des conditions initiales et des effets en cascade, en soulignant la vulnérabilité accrue des territoires successivement touchés par plusieurs événements (comme le village de Saint-Martin-Vésubie dans les Alpes-Maritimes, fortement touché par la tempête Alex en 2020, et de ce fait rendu plus vulnérable aux épisodes pluvieux observés les années suivantes) [WMO, 2018, Ritter et al., 2022, Potter et al., 2025].

2.2.3 Limites et difficultés liées aux avertissements basés sur les impacts

La transition de HBW vers des IBW pose cependant plusieurs difficultés. En apparence, la modélisation des impacts peut sembler moins complexe que celle de l'aléa, car elle cherche à décrire des liens directs de cause à effet, plutôt que des processus physiques complexes. Pourtant, cette étape introduit souvent davantage d'incertitudes [Merz et al., 2020]. D'une part, la modélisation des impacts reste encore peu développée, notamment en raison du manque de données d'observation systématiques sur les dommages. D'autre part, les impacts eux-mêmes dépendent de nombreux facteurs (y compris des comportements humains) qui rendent leur prévision particulièrement délicate. Pour cette raison, l'intégration d'un module d'impact dans une chaîne de prévision peut nécessiter des ressources accrues en matière d'infrastructures de calcul et d'expertise, dont de nombreux NMHSs ne disposent pas (66 % des NMHSs européens en faisaient le constat dans le sondage de Kaltenberger et al., 2020).

La dissémination des IBW soulève également la question des rôles et responsabilités des différentes organisations. D'après le sondage de Potter et al. [2021], les NMHSs ne sont pas toujours perçus comme les entités les plus légitimes pour communiquer des IBW, en raison de leur manque d'expertise concernant la vulnérabilité et l'exposition.

De plus, la plupart des bénéfices supposés des IBW restent à démontrer [Merz et al., 2020]. Certaines études questionnent même l'intérêt réel des IBW par rapport aux HBW. Par exemple, il a été montré que dans certaines situations, lorsque la peur est déjà un facteur de motivation majeur, ajouter des informations

sur les impacts ne change pas les comportements [Weyrich et al., 2020]. Finalement, une autre limite liée aux IBW, qui sera le sujet de la section suivante, concerne le manque de disponibilité de données d'évaluation.

2.2.4 Évaluation des systèmes d'avertissements basés sur les impacts

L'évaluation des avertissements constitue une étape essentielle mais complexe. Le premier défi réside dans la définition d'un « bon » avertissement. Selon Golding [Golding, 2022], un avertissement parfait est « celui qui permet d'éviter toutes les conséquences négatives évitables ». Le passage des HBW aux IBW soulève de nouveaux défis d'évaluation — au cœur notamment du projet High Impact Weather porté par l'OMM [Potter et al., 2025] — comme la capacité à mesurer dans quelle mesure un IBW améliore la compréhension du risque par ses destinataires, par comparaison à un HBW.

Lorsqu'on se focalise sur le composant "prévision" (estimation des conditions physiques futures) de l'avertissement, il est relativement simple d'évaluer la performance des HBW, par exemple en comparant les débits de pointe effectivement observés (ou reconstitués) aux débits simulés par le modèle hydrologique [Javelle et al., 2014]. En revanche, l'évaluation des IBW suppose de disposer de données d'impacts observés, ce qui est d'autant plus délicat que ces systèmes ont précisément pour objectif de réduire les impacts... et donc potentiellement de les rendre invisibles dans les données [Potter et al., 2021, Hemingway and Robbins, 2020]. Ce problème ne se pose pas ici, puisque les IBW étudiés dans ce manuscrit ne sont pas encore opérationnels, et que les observations d'impacts utilisées pour l'évaluation ne sont pas influencées par ces avertissements.

L'une des principales limites à l'évaluation des IBW réside dans le manque de données post-événement systématiques, exhaustives et continues, un constat largement documenté [Potter et al., 2025, Harrison et al., 2020, 2021, 2022a, Kaltenberger et al., 2020, Meléndez-Landaverde and Sempere-Torres, 2022]. Pour y remédier, plusieurs pistes ont été avancées, notamment le développement de collaborations entre services hydrométéorologiques, services de secours, autorités et autres acteurs opérationnels [Harrison et al., 2024, Potter et al., 2025]. Les travaux de Ruin et al. [2014] illustrent par exemple une initiative associant sciences physiques et sociales pour la collecte de données post-événement. D'autres recherches proposent également des méthodes visant à faciliter la collecte de données d'impacts [Saint-Martin et al., 2018, Wyatt et al., 2024]. Même lorsque les données sont disponibles, encore faut-il qu'elles soient cohérentes avec la variable à évaluer [Robbins et al., 2022]. Par exemple, si un IBW prévoit le nombre de bâtiments inondés par commune, il faut que les observations post-événement soient également disponibles à cette échelle, ce qui est rarement le cas. Par ailleurs, les données d'impact sont rarement produites dans une optique d'évaluation, et nécessitent souvent un retraitement, notamment lorsqu'on souhaite croiser plusieurs sources pour obtenir une vision plus exhaustive [Wyatt et al., 2023, Potter et al., 2025].

Les biais présents dans les bases de données d'impacts accessibles publiquement constituent une autre difficulté. Par exemple, un biais de seuil est observé lorsque seules les crues les plus graves sont enregistrées, au détriment des événements à impacts plus modérés [Wyatt et al., 2023]. Ce biais introduit un décalage entre les impacts prévus et observés, rendant l'évaluation plus difficile. Une piste proposée pour limiter ce problème est de diversifier les sources de données mobilisées, afin d'obtenir une représentation plus réaliste des impacts.

L'enquête menée par Kaltenberger et al. [2020] auprès de 32 services hydro-météorologiques nationaux (NMHSs) en Europe révèle que seulement 34% d'entre eux évaluent régulièrement leurs IBW. La majorité (63%) ne dispose pas d'une base de données d'impacts pouvant servir à des fins d'évaluation, et 75% ne collaborent pas avec d'autres organisations (par exemple, services de secours ou autorités) pour la mise en place de telles bases. De plus, 47% des NMHSs n'organisent pas de campagnes de

retour d'expérience post-événement permettant de collecter des données d'impacts. Ces chiffres illustrent clairement les limites mentionnées précédemment.

Dans le cas plus spécifique des crues et crues soudaines, l'évaluation du bloc « impacts » dans les chaînes de prévision reste peu fréquente, en raison du manque de données de dommages fiables [Molinari et al., 2019, Merz et al., 2010, Meyer et al., 2013, Jongman et al., 2012, Molinari et al., 2014], notamment à des échelles spatiales fines [de Moel et al., 2015]. La majorité des études s'appuie sur des observations agrégées, comme les populations affectées ou les pertes économiques à l'échelle régionale ou nationale [Alfieri et al., 2024, Ritter et al., 2021, de Moraes, 2023]. Quelques travaux ont toutefois exploré des évaluations à des échelles plus fines : comté [Terti et al., 2019a], commune [Silvestro et al., 2019, Saint-Martin, 2018], ou tronçon de rivière (environ 2 km) [Le Bihan et al., 2017]. La plupart de ces travaux se concentrent sur un nombre limité d'événements [Alfieri et al., 2024, Ritter et al., 2021, Terti et al., 2019a, Silvestro et al., 2019, Le Bihan et al., 2017], et les initiatives d'évaluation systématique sur de longues chroniques restent rares [de Moraes, 2023, Saint-Martin, 2018]. Quelques bases de données exhaustives adaptées à l'évaluation des systèmes de prévision des crues soudaines existent cependant, comme le jeu de données SHAVE ou la Storm Events Database aux États-Unis, construites à partir de remontées citoyennes [Calianno et al., 2013].

3 Objectifs de la thèse et organisation du manuscrit

3.1 Objectifs de la thèse

3.1.1 Mise en place d'une chaîne de prévision immédiate des impacts des crues soudaines pour l'Arc Méditerranéen

En France, la Vigilance Crues est déployée sur les grands cours d'eau instrumentés. Les niveaux d'avertissement (ou niveaux de vigilance) sont définis en tenant compte conjointement de l'aléa et des enjeux potentiellement affectés. En revanche, ce système ne couvre pas les petits bassins versants très réactifs, qui causent des crues soudaines parfois dévastatrices. Des outils locaux existent pour certaines zones particulièrement exposées mais leur mise en œuvre à plus grande échelle serait trop coûteuse. Le système Vigicrues Flash permet actuellement de pallier en partie ce manque, en émettant des avertissements automatiques offrant un minimum d'anticipation sur un grand nombre de bassins versants non jaugés à l'échelle nationale. Toutefois, il repose sur des seuils de débit pour émettre les avertissements, sans considération des impacts potentiels associés.

Comme vu précédemment, les approches fondées sur les impacts sont aujourd'hui largement encouragées, bien qu'elles impliquent une complexification de la chaîne de prévision. Le premier objectif de cette thèse est de développer un système de prévision basé sur les impacts potentiels des crues soudaines, à partir du système existant Vigicrues Flash, ainsi que de modèles et méthodes développées dans le cadre de différentes thèses précédant ce travail : Le Bihan [2016], Saint-Martin [2018], Jay-Allemand [2020], Hocini [2022], Charpentier-Noyer [2022]. L'échelle visée est régionale : l'ensemble de l'arc méditerranéen français.

Ce premier objectif a donc une visée opérationnelle, mais il pose également de nombreuses questions de recherche, parmi lesquelles :

- Un système HBW utilisant un modèle hydrologique régionalisé unique peut-il couvrir l'ensemble des cours d'eau méditerranéens, y compris ceux au comportement particulier (karst, neige, bassins influencés, etc.), tout en conservant des performances homogènes ?
- Comment connecter les sorties d'un modèle hydrologique maillé à un modèle hydraulique fondé sur une représentation vectorielle des cours d'eau, pour aboutir à une représentation des zones inondées nécessaire dans un système de type IBW ?
- Quel type de modélisation hydraulique peut-on mobiliser pour la représentation des zones inondées, et quel est l'impact de l'amélioration de la représentation des lits mineurs sur les résultats de modélisation de l'inondation ?
- Comment reconstituer la prévision d'inondation et la prévision d'impacts dans la chaîne de type IBW, en se basant sur des scénarios d'inondations pré-calculés ?
- Comment définir des avertissements fondés sur les impacts, à l'image de ceux actuellement fondés sur l'aléa dans Vigicrues Flash ?
- Est-il possible de coupler le module hydrologique de la chaîne d'avertissement avec des prévisions de pluie d'ensemble, afin d'augmenter les délais d'anticipation et de représenter les incertitudes associées aux prévisions ? Quelle est la plus-value d'un tel couplage ?

3.1.2 Évaluation détaillée sur une large fenêtre spatio-temporelle

Au-delà de la mise en œuvre d'une chaîne de prévision basée sur les impacts à l'échelle de l'arc méditerranéen, l'originalité de cette thèse réside dans l'évaluation exhaustive des différents blocs de modélisation constituant la chaîne. Comme évoqué précédemment, les chaînes de prévision sont géné-

ralement évaluées à l'échelle d'un événement donné, alors que les évaluations continues, incluant des événements de moindre ampleur et permettant d'évaluer statistiquement la performance des avertissements (notamment l'équilibre entre les non-détections et les fausses alarmes), restent rares.

L'objectif de la thèse est donc de mener une évaluation complète d'une chaîne de prévision combinant la modélisation des crues, des inondations et de leurs impacts, sur un vaste territoire (environ 20 000 km de linéaire de cours d'eau sur l'arc méditerranéen) et sur une longue période (2010–2022). Ce travail soulève plusieurs enjeux méthodologiques et permet de répondre à différentes questions de recherche :

- Quelles données mobiliser pour évaluer un modèle hydrologique estimant des débits et des périodes de retour sur un réseau en grande partie non jaugé ?
- Quelles sont les performances d'une chaîne de type HBW pour la détection des crue soudaine ?
- L'utilisation de données d'impacts permet-elle d'identifier les limites d'un système HBW, en particulier l'absence de prise en compte de l'exposition des enjeux dans la définition des seuils d'avertissement ?
- Quelles données et méthodes peut-on mobiliser pour évaluer des scénarios d'inondation calculés pour des périodes de retour fixes à échelle régionale ?
- Quels sont les ordres de grandeur des erreurs de hauteur d'eau associées à une modélisation hydraulique simplifiée appliquée à grande échelle ?
- Quelles données d'impacts peuvent être mobilisées pour l'évaluation de l'ensemble de la chaîne, sur une large fenêtre spatio-temporelle ?
- Quelle méthode d'évaluation peut-on appliquer pour évaluer la plus-value des IBW par rapport aux HBW ?
- Les résultats d'évaluation permettent-ils de confirmer cette plus-value en termes de réduction de fausses alarmes et de hiérarchisation des zones impactées ?
- Comment évaluer la plus-value d'un produit de prévisions d'ensemble de pluie par rapport à des produits déterministes, à l'échelle de quelques événements de crues soudaines ?
- Comment aborder la problématique de la propagation des incertitudes au sein de la chaîne de prévision ?

3.2 Organisation du manuscrit

Le manuscrit a été organisé autour des articles publiés ou soumis dans des revues internationales au cours des trois années de thèse. Certains chapitres ou sections reprennent donc directement l'intégralité d'un article rédigé en anglais. Dans ce cas, une synthèse en français, mettant en évidence les points principaux et les résultats clés de l'article, sera systématiquement proposée avant le texte intégral, accompagnée d'un résumé graphique.

La progression du manuscrit suit les différents blocs de la chaîne de prévision basée sur les impacts qui a été mise en œuvre :

- Le chapitre II aborde la question de l'évaluation d'un système d'avertissement aux crues soudaines de type HBW (analogue à Vigicrues Flash) couvrant un réseau essentiellement non jaugé, sur une large fenêtre spatio-temporelle, à partir de jeux de données variés. La question de l'homogénéité spatiale des performances du système est examinée. Ce chapitre a fait l'objet d'une publication dans le *Journal of Flood Risk Management* en mai 2025 [Godet et al., 2025b].
- Le chapitre III s'intéresse aux problématiques d'évaluation et de précision des scénarios d'inondation pouvant être mobilisés pour développer des systèmes d'avertissement aux crues soudaines de type IBW. Ce chapitre a fait l'objet d'une publication dans *Earth System Science Data* en juin 2025

[Godet et al., 2025a].

- Le chapitre IV décrit l'implémentation technique de la chaîne de prévision pour un système de type IBW, et s'intéresse notamment aux méthodes de couplage entre les blocs de modélisation hydrologique et hydraulique, ainsi qu'à la reconstitution de l'emprise inondée à partir de scénarios pré-calculés. Ce travail a fait l'objet d'une publication dans *Hydrology and Earth System Sciences* en mars 2024 [Godet et al., 2024].
- Le chapitre V détaille l'évaluation du système d'avertissement de type IBW sur une large fenêtre spatio-temporelle. Il décrit notamment les données d'impacts observés utilisées, ainsi que la méthode d'évaluation basée sur la comparaison avec un système d'avertissement de type HBW. Le contenu du chapitre a été soumis à l'*International Journal of Disaster Risk Reduction* en octobre 2025.
- Le chapitre VI, qui s'inscrit davantage dans une visée prospective, s'intéresse dans une première partie à la plus-value de prévisions d'ensemble des pluies à courte échéance dans une chaîne de prévision de type HBW, travail ayant fait l'objet d'une publication en novembre 2023 dans *Natural Hazards and Earth System Sciences* [Godet et al., 2023]. La seconde partie s'intéresse l'utilisation de ces mêmes prévisions dans une chaîne de type IBW, en particulier pour analyser la propagation des incertitudes à travers la chaîne.

Chapitre II

Évaluation d'un système d'avertissement aux crues soudaines basé sur l'intensité de l'aléa

Objectifs et organisation du chapitre :

Ce chapitre a pour objectif d'évaluer la performance des modèles hydrologiques pluie-débit régionalisés, utilisés pour la surveillance des petits cours d'eau non jaugés au sein des systèmes d'avertissement aux crues soudaines de type "Hazard-Based Warnings" (HBWs). Il propose une évaluation sur une large fenêtre spatio-temporelle d'une chaîne de modélisation très proche de celle du système opérationnel Vigicrues Flash. La principale difficulté de cette évaluation réside dans la disponibilité très limitée d'observations de débits sur les cours d'eau majoritairement non jaugés couverts par le modèle. Les résultats apportent un éclairage sur les questions suivantes :

- Quelles sont les performances du module hydrologique de la chaîne Vigicrues Flash, notamment par rapport à d'autres systèmes similaires de la littérature ?
- Que révèle une première estimation des délais d'anticipation fournis par un modèle hydrologique forcé par des observations de pluie, dont l'anticipation dépend uniquement des temps de réponse des bassins ?
- Un système HBW utilisant un modèle hydrologique régionalisé unique peut-il couvrir l'ensemble des cours d'eau méditerranéens, y compris ceux au comportement particulier (karst, neige, bassins influencés, etc.) ?
- L'utilisation de données d'impact pour l'évaluation permet-elle d'identifier les limites d'un système HBW, en particulier l'absence de prise en compte de l'exposition des enjeux dans la définition des seuils d'avertissement ?

Ce chapitre reprend l'intégralité d'un article publié en 2025 dans la revue internationale *Journal of Flood Risk Management* [Godet et al., 2025b]. La première section propose une synthèse en français de cet article, et les sections suivantes reproduisent le texte intégral de l'article, rédigé en anglais.

0 Synthèse en français

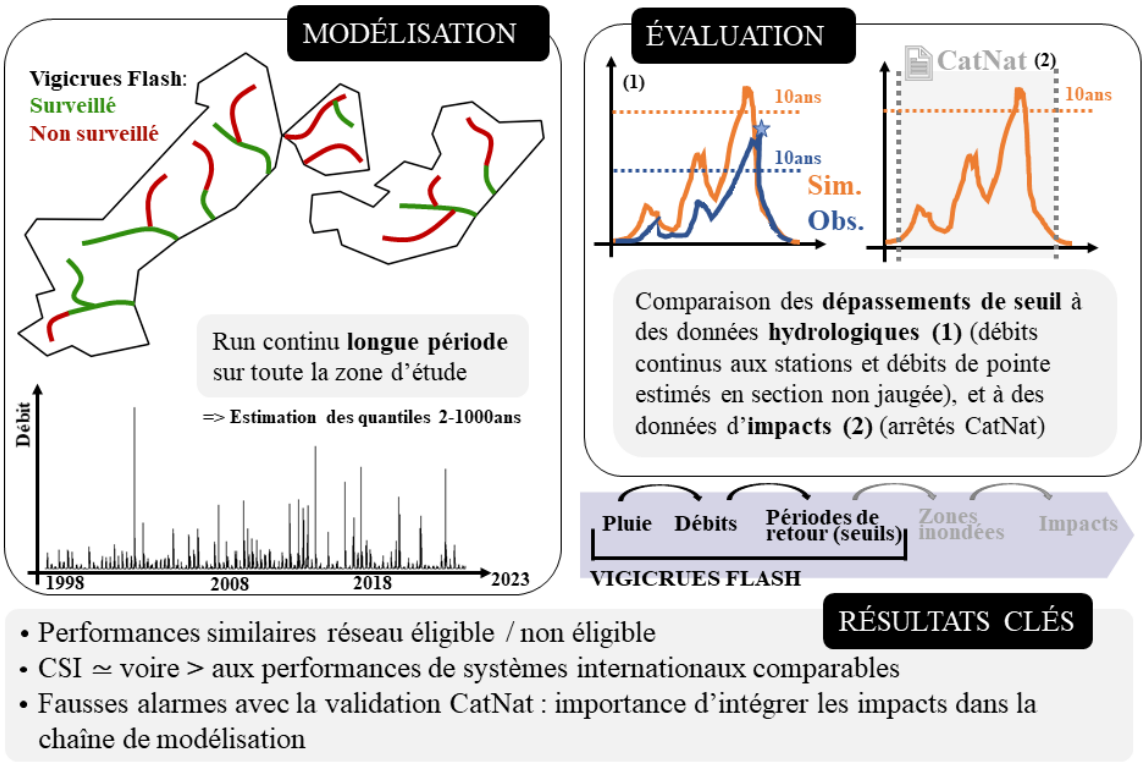


FIGURE II.1 – Résumé graphique de l'article Godet et al. [2025b].

Cet article présente un travail d'évaluation d'une chaîne d'avertissement de type "HBW", très proche de la chaîne opérationnelle Vigicrues flash actuelle. La chaîne évaluée se base en effet sur le modèle SMASH-Lag0 et émet des avertissements à partir des dépassements de seuils de débits. Par rapport à Vigicrues Flash, elle couvre un réseau hydrographique plus étendu (le système opérationnel se limite aujourd'hui à une sélection de cours d'eau) et mobilise une gamme plus large de seuils (périodes de retour de 2 à 100 ans, alors que Vigicrues Flash n'utilise que deux seuils de périodes de retour de 5 et 20 ans). L'évaluation est conduite à l'échelle de l'arc méditerranéen, sur une période continue de 15 ans (2008–2023), ce qui dépasse largement les études menées jusqu'ici sur le système opérationnel Vigicrues Flash (opérationnel depuis 2017). Ce travail a nécessité de lever plusieurs verrous méthodologiques :

- L'estimation de quantiles de crue élevés, notamment au-delà de la période de retour 20 ans, seuil maximal utilisé dans la version opérationnelle de Vigicrues Flash.
- La mobilisation de données hydrologiques pour valider l'approche, à la fois sur le réseau jaugé et non jaugé.
- La comparaison entre dépassements de seuils hydrologiques (approche Vigicrues Flash) et données d'impacts, afin de compléter la validation.
- L'application d'une méthode d'évaluation multi-sources à grande échelle : plus de 10 ans de données sur l'ensemble de l'arc méditerranéen.

L'évaluation repose, dans un premier temps, sur l'intensité des crues, en comparant les dépassements de seuils issus des débits observés (stations de jaugeage ou relevés post-crue effectués pour des crues d'intensité remarquable) et simulés. Elle couvre deux périodes : la période de calibration du modèle (2008–2018) et la période de validation (2018–2023). Elle inclut également deux réseaux hydrographiques : les tronçons actuellement éligibles au service Vigicrues Flash, et ceux qui ont été jusqu'ici été écartés en raison de doutes sur les performances du modèle sur ces tronçons. Les seuils de débits utilisés

pour cette évaluation correspondent à des quantiles associés à des périodes de retour comprises entre 2 et 1000 ans, calculés sur l'ensemble de ce réseau à partir de l'approche SHYREG. Dans un second temps, l'évaluation porte sur les impacts, en comparant les avertissements issus des débits simulés avec les arrêtés CatNat (décrets d'indemnisation des catastrophes naturelles). Contrairement à la validation hydrologique, limitée aux stations de mesure et aux relevés ponctuels, cette évaluation couvre l'ensemble de la zone d'étude. La résolution de cette évaluation est celle des communes de l'arc méditerranéen français ($\simeq 20km^2$), qui correspond à l'échelle de définition des arrêtés CATNAT.

En termes d'évaluation basée sur l'intensité des crues, le système affiche un Critical Success Index (CSI) de 73 % pour les débits de pointe issus des relevés post-crue (dont les estimations ont été généralement faites surtout sur des tronçons très impactés, diminuant donc les chances de fausses alarmes) et de 41 % pour les débits mesurés aux stations. Les délais d'anticipation ont été estimés supérieurs à une heure dans 60% à 70% des cas d'avertissements émis. Il faut également noter que l'évaluation menée n'a pas montré de différence notable entre les périodes de calibration (2008–2018) et de validation (2018–2023), ce qui souligne la robustesse du système. Aucune différence significative de performance n'a été observée entre les tronçons actuellement éligibles au service Vigicrues Flash et ceux non éligibles, ce qui suggère qu'une extension du système à l'ensemble du réseau hydrographique (incluant les bassins au comportement hydrologique particulier) sans perte majeure de performance est possible.

L'évaluation basée sur les impacts, réalisée en comparant les HBWs aux arrêtés CatNat, présente un CSI de 35% pour un seuil d'avertissement en période de retour égal à 20 ans. Bien que ce score soit inférieur à celui de l'évaluation hydrologique, il constitue une première référence pour évaluer la capacité du système à anticiper les impacts réels. La baisse de performance constatée par rapport à l'évaluation hydrologique peut être attribuée en partie à certaines limites des données de validation utilisées, et plus particulièrement à la difficulté de distinguer les inondations par ruissellement des inondations par débordement de cours d'eau dans les arrêtés CatNat.

L'évaluation du système HBW à partir de données d'impacts observés a également mis en lumière de nombreux cas où des avertissements de "crues très fortes" (T>20ans) correspondaient à des fausses alarmes selon l'évaluation à partir de données de débits aux stations, mais se transformaient en "hits" selon l'évaluation à partir des arrêtés CatNat. A l'inverse, de nombreux cas d'avertissements émis pour "crue forte" (T>5ans) correspondaient à des "hits" selon l'évaluation hydrologique, et se transformaient en fausses alarmes selon l'évaluation basée sur les impacts. Ces deux résultats ont montré que l'évaluation d'un système HBW avec des données d'impacts observés peut apporter un éclairage très différent sur les résultats, et souligne bien l'importance d'intégrer une représentation des impacts attendus dans la chaîne de prévision. L'enquête menée auprès des utilisateurs de Vigicrues Flash a confirmé ce constat : les avertissements reçus correspondaient à un événement où le niveau du cours d'eau a évolué mais sans aucun débordement ni dommage dans près de 50% des cas.

En résumé, cette étude montre que le modèle hydrologique utilisé dans Vigicrues Flash permet de fournir des avertissements sur l'ensemble du réseau hydrographique avec un niveau de détection comparable aux résultats de la littérature, en se basant sur une évaluation mobilisant des données hydrologiques. Elle souligne aussi que les bases de données d'impacts représentent une voie prometteuse pour l'évaluation des chaînes d'avertissement, et permettent de mettre en évidence les limites des systèmes de type HBW, dont les avertissements ne traduisent pas toujours des impacts.

1 Introduction

1.1 Flash floods and operational early warning systems worldwide

At a time when extreme weather events are becoming more frequent, flash floods stand out as one of the most recurrent, unpredictable and destructive natural disasters. They represent 85% of all reported damaging flood events and claim 5000 lives per year worldwide [WMO, 2020]. Some single flash floods may induce economic losses exceeding a billion euros [Gaume et al., 2009]. More generally, damages related to flood (including flash-floods) are likely to increase in the near future : according to Dottori et al. [2023], in Europe, without climate change mitigation and adaptation, annual damage from river flooding would rise from €7.6 billion (present conditions) to €44 billion (2100), and the exposed population would increase from around 166,000 to 500,000. To address this threat, operational flash flood early warning systems (FFEWS) have been put into operation, particularly in developed regions, such as the FLASH system in the USA [Gourley et al., 2017], the ERIC-ERICHA system in Europe [Raynaud et al., 2015, Park et al., 2019], the AFAS system in New Zealand [Cattoën et al., 2022] or the system developed by the National Early Warning and Monitoring Centre for Natural Disasters (CEMADEN) in Brazil [de Moraes, 2023]. Progress is also being made in developing countries, e.g. Alfieri et al. [2024] for an operational system in the Greater Horn of Africa, in particular through the Early Warnings for All initiative led jointly by the World Meteorological Organisation (WMO) and the United Nations Office for Disaster Risk Reduction (UNDRR).

These FFEWS often share similar characteristics. While some systems directly compare observed or predicted rainfall with reference thresholds to estimate flood severity, e.g. the Flash Flood Guidance System in the US [Clark et al., 2014] or in Haiti and the Dominican Republic [Shamir et al., 2013], most FFEWS are based on rainfall-runoff models, which generally provide better results [Zanchetta and Coulibaly, 2020]. Spatial and temporal resolutions vary between models, although they are usually in the range of 1km/15min-1h, to match the spatial resolution of the input rainfall products, and to be consistent with the rapid dynamics of flash floods. A combination of three types of models is generally used to forecast river discharges : a water balance model to take into account the vertical processes of infiltration and evapo-transpiration ; a lateral routing model on slopes to represent surface and subsurface transfers ; and finally a routing model within the hydrographic network. Lumped models have gradually been replaced by distributed models, e.g. the Sacramento soil moisture accounting SAC-SMA was mostly replaced by the Hydrology Laboratory Research Distributed Hydrologic Model (HL-RDHM) for the Flash Flood Guidance system in the US [Zanchetta and Coulibaly, 2020]. Furthermore, most models are now based on hybrid conceptual-physical approaches, e.g. the CREST [Wang et al., 2011] and Continuum [Silvestro et al., 2013] models. Finally, most rainfall-runoff models are forced with rainfall forecasts (to improve anticipation), especially ensemble forecasts to represent and communicate uncertainties, which is of great value for decision making [Fundel et al., 2019]. Some of the FFEWS extend the hydrological modelling to flood mapping in order to provide impact-based warnings, as urged by WMO [2015] and Merz et al. [2020]. Following these guidelines, the Flood Guidance Statements in the UK [Cole et al., 2016] are based on a flood risk matrix approach that combines potential impact severity and likelihood of occurrence.

1.2 The flash flood warning system in France

France frequently experiences severe flash floods, particularly in the Mediterranean region [Gaume et al., 2016]. In response, the Ministry of the Environment launched the Vigicrues Flash service in March

2017 [Javelle et al., 2016, Fouchier et al., 2017, Piotte et al., 2020], complementing the Vigicrues flood forecasting service established in 2003. While Vigicrues provides forecasts for the main river network (covering 22,000 km of rivers), it has shown limitations during catastrophic flash flood events in small, ungauged basins outside this network. For instance, the flash floods in the Var department in June 2010 caused €750 million in damages [CCR, 2023a]. These limitations prompted the development of the Vigicrues Flash system. Vigicrues Flash integrates radar-gauge rainfall grids at 1-km resolution into a conceptual distributed rainfall-runoff model, producing peak discharge estimates every 15 minutes across the river network (see section 3.1). These estimates are compared to flood quantiles calculated using the same hydrological model. Since Vigicrues Flash does not currently incorporate rainfall forecasts, its anticipation capabilities are limited, making it more accurately described as a Flash Flood Monitoring System rather than a Flash Flood Early Warning System. It issues nowcasts rather than forecasts ; however, we retain the term "Early Warning System" for simplicity and generality.

This system sends automatic warnings to local authorities that have subscribed to the service, although anyone can visualize the maps of municipalities that might be impacted on https://apic-vigicruesflash.fr (last accessed : 17/07/2024). Two warning thresholds have been defined : "severe flood" (5-y return period) and "very severe flood" (20-y return period). The Vigicrues Flash system is accompanied by the APIC system (Warning for Intense Rainfall at the Municipality scale), which provides maps of characterization of heavy rain intensities on the same web platform, with warnings for "intense rainfall" (10-y return period) and "very intense rainfall" (50-y period). Vigicrues Flash covers about 30 000 km of river network and 11 319 out of the 35 000 municipalities in France (see figure II.2). The eligibility of a municipality to the service is determined based on the criteria summarised in the table II.1. Notably, Table II.1 shows that catchments with concentration times of less than 3 hours are not eligible for the Vigicrues Flash service (representing roughly 38% of all catchments in France). This decision was made at the operational level to maximize the chances of providing sufficient anticipation time for response coordination. Since the model is forced only with rainfall observations and not forecasts, the anticipation time is approximately equal to the catchment's concentration time, although it also depends on the location of heavy rainfall. This anticipation is further reduced by the delay in receiving rainfall data (approximately 10 minutes), while model computations can be considered instantaneous.

Tableau II.1 – Automatic criteria used to determine the eligibility of river reaches for the Vigicrues Flash service (Source : SCHAPI, Caruso et al. [2013])

Type of criterion	**Description**
Hydrology	Drainage area >10km2
	Catchment concentration time >3h
	Share of snowfall in the basin water supply <10%
	Ratio "Annual base flow / biennal daily flood flow" >3
	i.e. (QA in mm/year) * 365 / (QJ2 in mm/24h) >3
Topography / Geography	Mean altitude difference in the watershed >0.5m
	Non-tidal watershed
	Watershed not monitored by the Vigicrues system
	Watershed area outside France <10%
Human influence	Watershed area controlled by dam <5%
Geology	Watershed area in karst zone <5%

Given its recent implementation, the performance of the Vigicrues Flash system has rarely been evaluated so far. However, feedbacks from subscribed users were analysed for warnings sent between April 2017 and April 2019 [Parra, 2019] and between December 2021 and July 2024 (see appendix 1). Globally, the

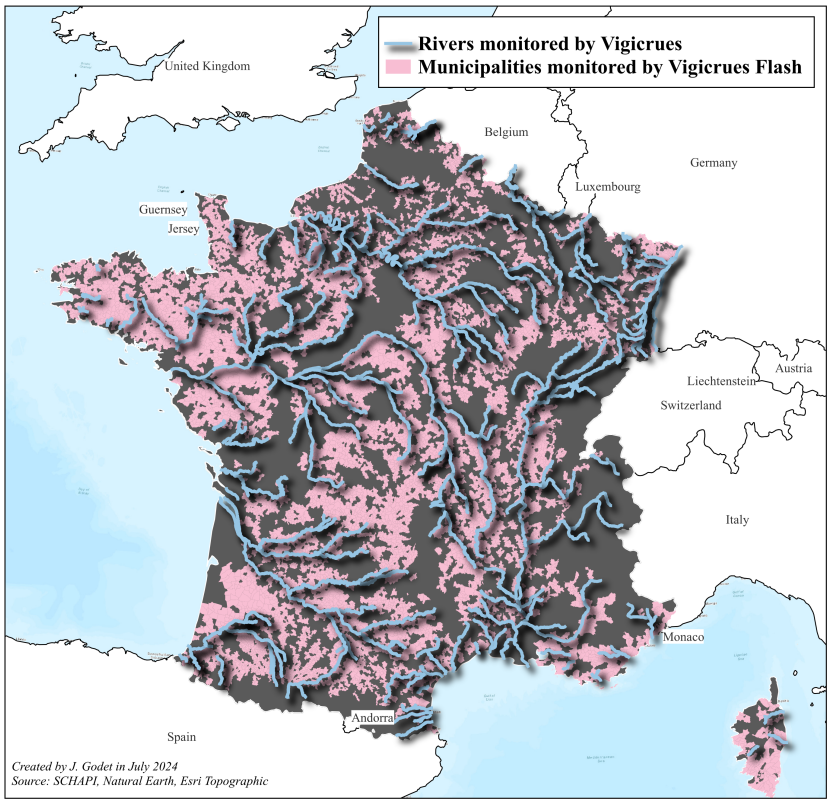


FIGURE II.2 – Municipalities eligible for Vigicrues Flash and river network covered by Vigicrues, July 2024

results show that 2/3 of the respondents found the warnings accurate and useful, and 2/3 of the respondents received a warning before the significant reaction of the river occurred (although this percentage may not reflect the real anticipation of the system because of the confusion between "river reaction" and "flood" for some end-users). The feedback from subscribers also indicated that only 13,2% of warnings sent between 2017 and 2019 resulted in damaging events, and roughly 25% for 2021-2024. Finally, in terms of false alarms, figure B.1e shows that when the warning did not indicate a flood risk (i.e. there was an APIC warning, but no Vigicrues Flash warning) flooding was observed in only 12% of the cases. Beyond this qualitative assessment of the system, flood intensity-based validation was investigated : Javelle et al. [2019] have looked at the performance of Vigicrues Flash during the storm of June 12th 2018 in France, with encouraging results.

1.3 Assessing the performance of flood warnings

The evaluation of the performance of flood warnings is a step that should not be overlooked. Lumbroso [2018] for instance has shown that only three out of the thirty flood early warning systems set up in African countries appear to be really effective in reducing humanitarian impacts. Flood warning systems are often assessed at the event scale [Shamir et al., 2013, Javelle et al., 2019, Chahinian et al., 2023, Thieken et al., 2023]. While analyzing system performance during high-impact events is crucial, we argue that such evaluations should be complemented by more generic statistical assessments over long periods of time. For instance, Clark et al. [2014] have assessed the skill of the US Flash Flood Guidance system over a 4-year period and Cattoën et al. [2022] proposed a validation of the Aotearoa (New Zealand) Flood Awareness System at both the event scale and on a continuous period (2.5 years). Furthermore, a long-time evaluation makes it possible to assess the system's behaviour for small to moderate events. While discharge measurements during flood events are generally used for validating FFEWS, they cannot

always be relied upon : either because flash floods often occur on small ungauged rivers, or because gauging stations can be destroyed or damaged during a flood, or even because the estimated discharges lie outside the reliability range of the river gauging stations' rating curves. For these reasons, it may be relevant to look for additional validation datasets, such as indirect discharge estimates from post-flood surveys.

In this paper, we propose an evaluation of the hydrological model used in Vigicrues Flash. This evaluation first focuses on the capacity of the chain to detect flood threshold exceedances. Specifically, we aim to determine whether the model can distinguish moderate floods (2-year to 10-year return period), severe floods (10-year to 100-year return period), and very severe floods (100-year to 1000-year return period) without generating too many false alarms. Defining discharge thresholds in terms of return periods raises the issue of flood frequency analysis in the context of ungauged catchments, for which we propose an approach based on the French SHYREG statistical interpolation method [Arnaud et al., 2014, Aubert et al., 2014]. In a second step, we aim to assess whether Vigicrues Flash can detect damaging floods by comparing the warnings to legal CatNat decrees recognizing natural disasters. The article is organized as follows : Section 2 details the validation datasets (discharge data and legal CatNat decrees). Section 3 presents the methods (the rainfall-runoff modeling chain used in Vigicrues Flash, the validation method, and the flood frequency analysis method). Section 4 presents the validation results. Limitations and possible improvements are highlighted in the discussion and conclusions section (Section 5).

2 Validation data

2.1 Continuous measurements from hydrometric stations

Mean discharges at the 15 minute time step were estimated at 117 gauging stations from the Hydroportail database (https://hydro.eaufrance.fr/, last accessed : 26/07/2024). This selection of river gauges was validated with the local flood forecasting services to ensure a good quality of data in the range of flood discharges (reliability of rating curve extrapolation). Among this selection, 75 stations are located on river reaches currently monitored by Vigicrues Flash (see figure II.3a). The catchment areas range from 19 km² to 1958 km². The oldest station went into service in 1963. The mean percentages of missing data per year between 1998 and 2023 (simulation period) are summarised in table II.2.

Tableau II.2 – Mean percentages of missing discharge data per year, on the 117 gauging stations selected

1998	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
34%	31%	32%	34%	28%	34%	27%	30%	31%	35%	25%	19%	15%
2011	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**
14%	14%	11%	10%	7%	6%	7%	4%	4%	3%	2%	2%	1%

2.2 Peak discharge estimates from post-event surveys

The river gauge data was complemented with peak discharge data gathered for ungauged rivers within the framework of the Cévennes-Vivarais Mediterranean Hydrometeorological Observatory (ohmcv.osug.fr, last access 2024-11-22) and the HyMeX (Hydrological Cycle in the Mediterranean Experiment, www.hymex.org, last access 2024-11-22) programs [Ducrocq et al., 2019]. These two long-term observation programs organize systematic field campaigns after outstanding floods, to provide estimates of the peak discharges in the ungauged river reaches affected by these events. In total, more than 400 peak discharge estimates were produced for fourteen flood events over the period 2002-2020. [Brunet and Bouvier, 2017, Amponsah

et al., 2018, Lebouc et al., 2019a,b, Payrastre et al., 2019, Brigode et al., 2021, Payrastre et al., 2022]. These estimates, which locations are mapped on figure II.3a, form an original independent dataset which was not previously used to calibrate or validate the rainfall-runoff model of Vigicrues Flash.

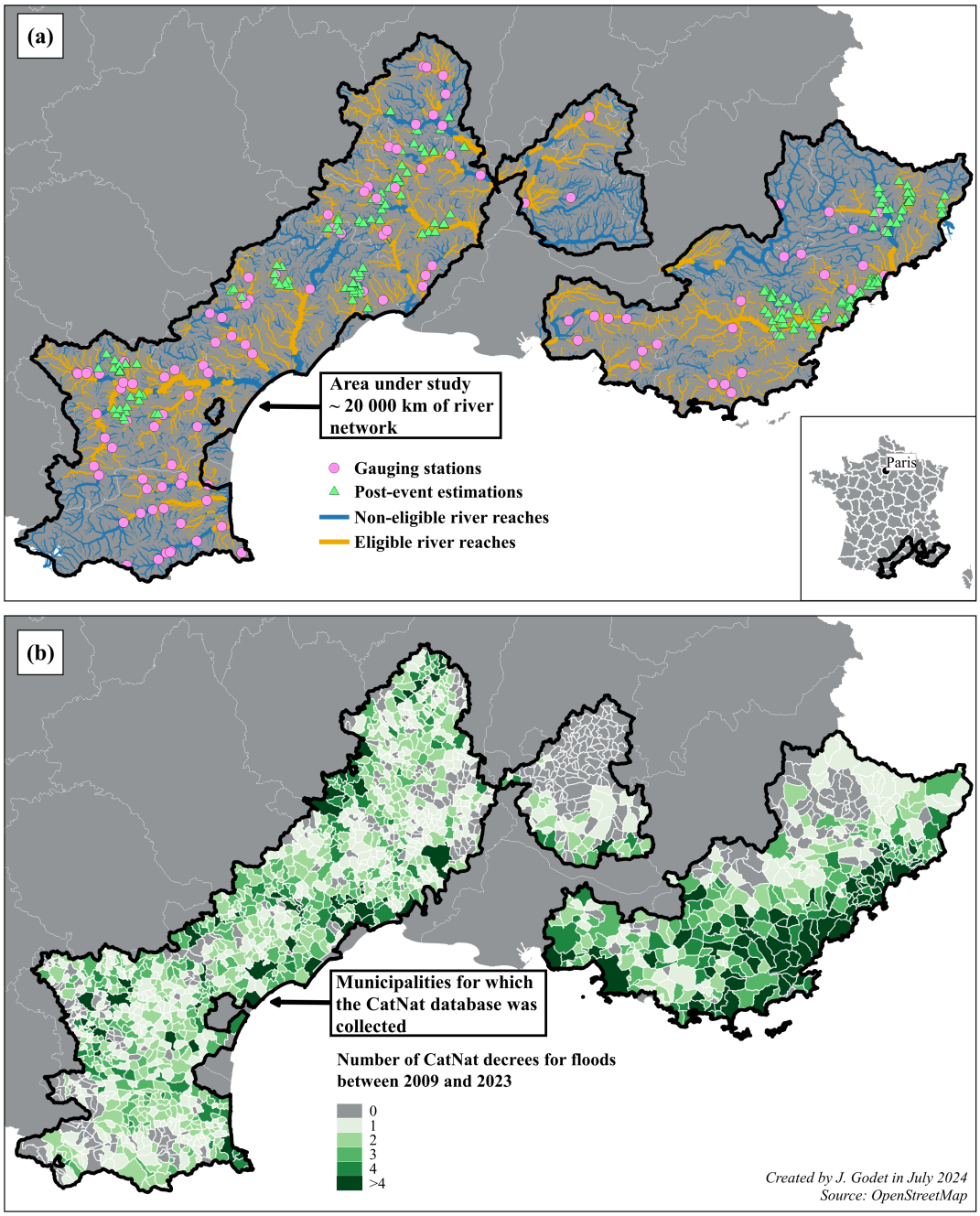


FIGURE II.3 – Maps of the validation datasets : (a) discharge measurements on eligible / ineligible river reaches and (b) CatNat database (sources : SCHAPI 2024, GASPAR, Hydroportail)

2.3 CatNat decrees

In France, "CatNat" decrees (short for "Catastrophes Naturelles," or "Natural Disasters") are official governmental declarations that recognize certain events as natural disasters. These decrees enable affected individuals and communities to receive compensation for damages under their insurance policies. The system operates on the principle of national solidarity, where collective national resources are mobilized to support those impacted by events like floods, landslides, earthquakes, droughts, and others. When a natural disaster occurs, causing significant damage to property and infrastructure, local governments or

municipalities submit a request to the national government for the event to be recognized as a natural disaster. If the event meets predefined criteria regarding its severity and intensity (such as a theoretical return period of at least 10 years for the rainfall or flood event), a decree is issued, officially recognizing the event as a natural disaster. This recognition allows affected individuals to access compensation and other forms of assistance provided by the national government.

The historical records of CatNat decrees with the corresponding affected municipalities since 1982 is available in the GASPAR database. The use of CatNat decrees for flood event inventory has already been investigated by Douvinet [2006], Douvinet and Vinet [2012], who pointed out some limitations : (i) the criterion of the 10-year return period seems flexible, (ii) the decision for the decree is subject to political pressures, (iii) damages induced by river and pluvial floods cannot be distinguished. Nevertheless, the CatNat decrees extracted from the GASPAR database have previously been used for forecast chain validation purposes [Saint-Martin, 2018]. From this database we extracted the 3647 CatNat decrees corresponding to our study area (1985 municipalities located in the French Mediterranean region) between september 2009 and september 2023. They are illustrated in figure II.3b.

Since 2018, the procedure for requesting CatNat decrees has been computerized, using a platform called iCatNat. Although not published in open access, the iCatNat database could be obtained from the Ministry of the Interior for this study. It is much more informative than the GASPAR database, for three main reasons : both accepted and rejected applications are listed, together with a description of the hazard type (A1 : river flooding ; A2 : surface flooding, B : torrential flooding) and the number of declared damaged buildings per municipality. However, some limitations remain : e.g. the number of flooded buildings is approximate, there may be confusion between hazard types : river floods of intermittent river may be classified as surface floods, and the database only goes back to 2018.

3 Methods

3.1 Rainfall-runoff model

As described in Piotte et al. [2020], Godet et al. [2023], the hydrological model used in Vigicrues Flash is a reservoir-based model inspired by the GR ("Génie Rural") family of models [Perrin et al., 2003]. It is a continuous, conceptual, fully distributed gridded model, operating at a 15-min time step with a 1-km² resolution. In this work, we used the same model structure as Vigicrues Flash, via the SMASH (Spatially-distributed Modelling and ASsimilation for Hydrology) hydrological modeling platform [Colleoni and Huynh, 2024], which is the platform currently used to enhance the hydrological modelling for future versions of Vigicrues Flash. The general expression of the model structure is described by :

$$q(i,t) = f([P,E](i,t), [c_p, c_t](i), [h_p, h_t](i,t))$$

Detailed explanations on how the function f is numerically solved can be found in [Jay-Allemand et al., 2020] ; only the concept will be explained here (see figure II.4). In its current operational version, the SMASH model is based on two distributed parameters determined at the pixel level : the capacity c_p of the production reservoir and the capacity c_t of the transfer reservoir. At each pixel i and time-step t, the net evapotranspiration $e_n(i,t)$, the production infiltrating evaporation $e_s(i,t)$, and the net rainfall $p_n(i,t)$ are estimated based on the atmospheric forcings P (rainfall) and E (evapotranspiration). A fraction of $p_n(i,t)$ is used to fill the production reservoir (of level $h_p(i,t)$), and the remaining fraction $p_{r,i}$ is directed towards a transfer reservoir (of level $h_t(i,t)$) that generates the pixel runoff flow $q(i,t)$. To obtain the

total flow rate $Q(i,t)$ at the basin outlet, the runoff discharges $q(t,i)$ from all the pixels of the basin are transferred to the routing module. In the Vigicrues Flash system, in order to better anticipate floods while ingesting only precipitation observations, the chosen routing module is the "lag0" module, i.e. "Instanteaneous Routing" : it means that the discharge $Q(i,t)$ for a cell is the simple sum of the runoff discharges of all the upstream grid cells $q(i,t)$, produced for the same time step t.

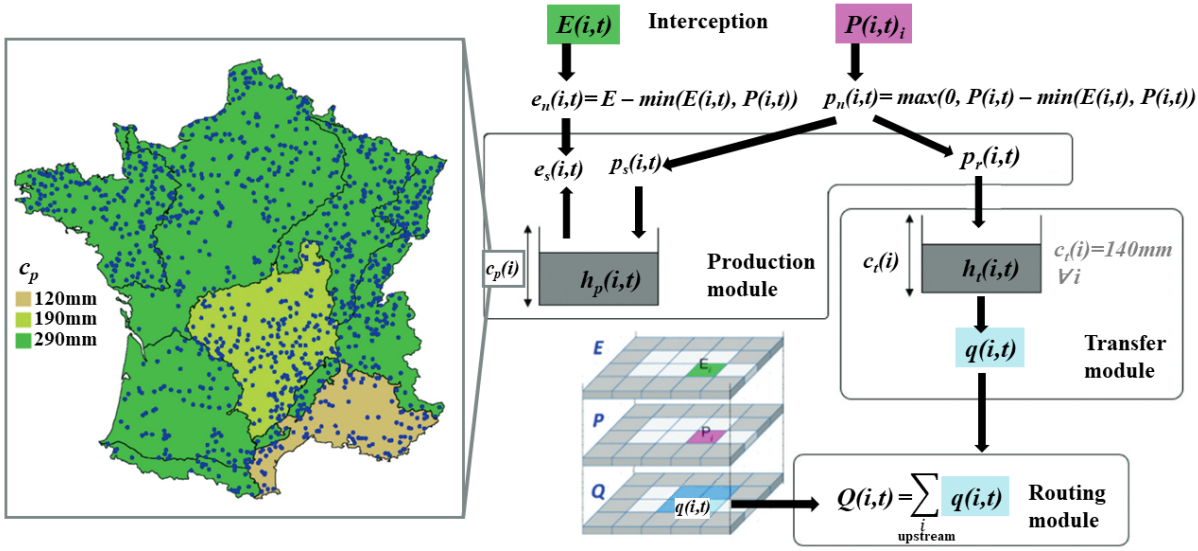


FIGURE II.4 – The simplified hydrological model used in Vigicrues Flash. All parameters are explained in the text. The location of the 1192 calibration basin outlets is shown on the left is shown along with the values of the regionalized parameter c_p. The figure is inspired from Piotte et al. [2020].

The main disadvantage of the lag0 method is that it does not account for flood routing in the hydrographic network and thus does not provide realistic hydrographs (see figure II.5). However, to account for potential bias when estimating real-time peak flows, these peak flows are compared to discharge quantiles estimated from the same model : the 5-year and 20-year discharge quantiles are estimated from long simulations of the SMASH model with rainfall reanalysis data, by fitting of a Gumbel distribution to the maximum annual simulated peak discharges. Although the Gumbel distribution is less suited for extreme values than the GEV, it remains a common choice for low return periods (up to 20 years) and was selected here for consistency with the operational method. However, for larger quantiles, we use the SHYREG method, as explained in Section 3.3. This method is based on the assumption that, even if the discharge values produced by the model are imperfect, the ranking of the flood magnitudes remains reliable. The concept of using discharge exceedance threshold relative to climatological model simulations in order to provide forecasts on the severity of a flood event is also found in other FEWS [Alfieri et al., 2013, Reed et al., 2007].

In this work, simulated flows are produced at a 15-minute time step and 1-km^2 scale using a reanalysis of the ANTILOPE [Laurentin, 2008] QPE (Quantitative Precipitation Estimate) product, which combines observations from the Météo-France radar and rain gauge network. The real-time ANTILOPE QPE product is used in the Vigicrues Flash operational model, however the reanalysis (which includes nearly 40% of additionnal rain gauge data) was used here to reduce the errors introduced by the input rainfall, and to simplify the model (avoiding the need to replicate the detailed operational process, which updates results daily upon receiving the ANTILOPE reanalysis). This choice may result in slightly more optimistic outcomes compared to the actual operational system. The lag0 method generally provides some anticipation even without rainfall forecasts, depending on the typical response time of the catchment and the distri-

bution of the rainfall within the basin (e.g., whether the rain falls in the upstream or downstream part of the catchment). The evapotranspiration corresponds to inter-annual values. The simulation period is split in two : the calibration period 2008-2018, on which the parameters c_p and c_t were previously calibrated, and the validation period 2018-2023. The model outputs are discharge time series at each simulation grid cell. Simulated flows to be evaluated correspond to basin outlets where observation is available, from gauging stations and post-event peak discharge estimates. Methods for allocating these basin outlets to specific hydrological modelling grid cells based on a flow direction grid have been documented by Godet et al. [2024], who advocate a "contour-based" allocation based on similarity of catchment boundaries. Given the limited number of points to be allocated in this study, they were allocated to grid cells based on distance and area proximity, and positions were manually checked to avoid allocation errors.

3.2 Contingency-based validation

In this study, the verification method was based on contingency tables and associated scores, which are often used to qualify the ability of forecasting systems to detect categorical discrete events [Mason, 1982]. Still in the context of flood modelling, similar approaches can also be used to assess the performance of flood mapping methods by comparing observed and simulated flood extents [Hocini et al., 2021]. Contingency tables, for which an example is given in table II.3 for a binary event (with only one threshold and two possible categories), describe the discrete joint distribution of simulation (or forecast) and observation in terms of the frequencies of occurrence for defined categories. In the case of a single threshold, four categories are defined depending on the correspondence between simulation (or forecast) and observation : hit, miss, false alarm, and correct rejection. Based on this table various contingency scores can be computed, such as the Critical Success Index, which measures the overall forecast accuracy according to the number of hits (a), false alarms (b), and misses (c) : $CSI = \frac{a}{a+b+c}$ (see appendix 2).

Tableau II.3 – Content of a 2x2 contingency table for a binary event with two possible outcomes. a, b, c, and d correspond to the number of occurrences in each category.

Event simulated	**Event observed**	
	Yes	**No**
Yes	HIT (a)	FALSE ALARM (b)
No	MISS (c)	CORRECT REJECTION (d)

3.3 Flood intensity-based contingencies

The contingency tables have been defined from multiple discharge thresholds to evaluate the quality of Vigicrues Flash for different flood severity levels : eight discharge quantiles were estimated for return periods ranging from 2 to 1000 years (see details on their calculation in section 3.5). The observed and simulated flood/no-flood occurrences are counted for a defined set of events. When peak discharge estimates from post-event surveys are used as a reference, the event dates are easily defined. When using continuous time series, a hydrograph segmentation method was used to identify individual flood events with their associated observed/simulated peak discharges. The segmentation method, described in the SMASH documentation and in Huynh et al. [2023], uses the rainfall gradient and energy to define the start date of an event and the discharge baseflow for the ending date, as well as a maximum time period. The segmentation was performed using both observed and simulated discharge series with SMASH default parameters, to ensure a comprehensive detection of misses as well as false alarms. For each event and station, it is then evaluated whether the observed and simulated peak discharge exceed each of the eight discharge thresholds (see figure II.5). It should also be noted that in this study a HIT has been

counted even if the threshold crossing occurred later for simulation than for observation. However, we have calculated anticipation times (difference between the moments of the first threshold crossing for obs. and sim. discharges), and negative anticipation times thus enable to identify these late detections.

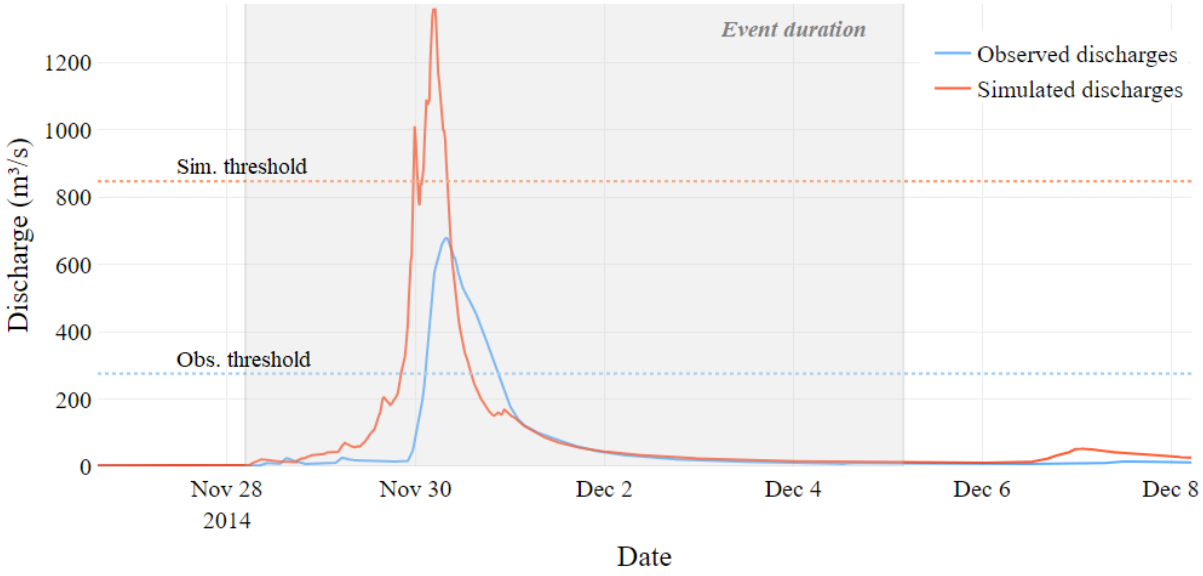


FIGURE II.5 – Illustration of a HIT counting in the validation process, for a given return period threshold.

When the results are aggregated for all stations and all events, the output is a table similar to the one shown in figure II.6. It counts the number of times the observed (resp. simulated) maximum exceeded return period falls within each interval, i.e., $[2y; 5y[$, $[5y; 10y[$, etc. In general, this table is only used as a calculation step for the contingency scores. However, it can also give an idea of the performance of the model by observing "how diagonal" the matrix is. We argue that a measure of the diagonality of a matrix

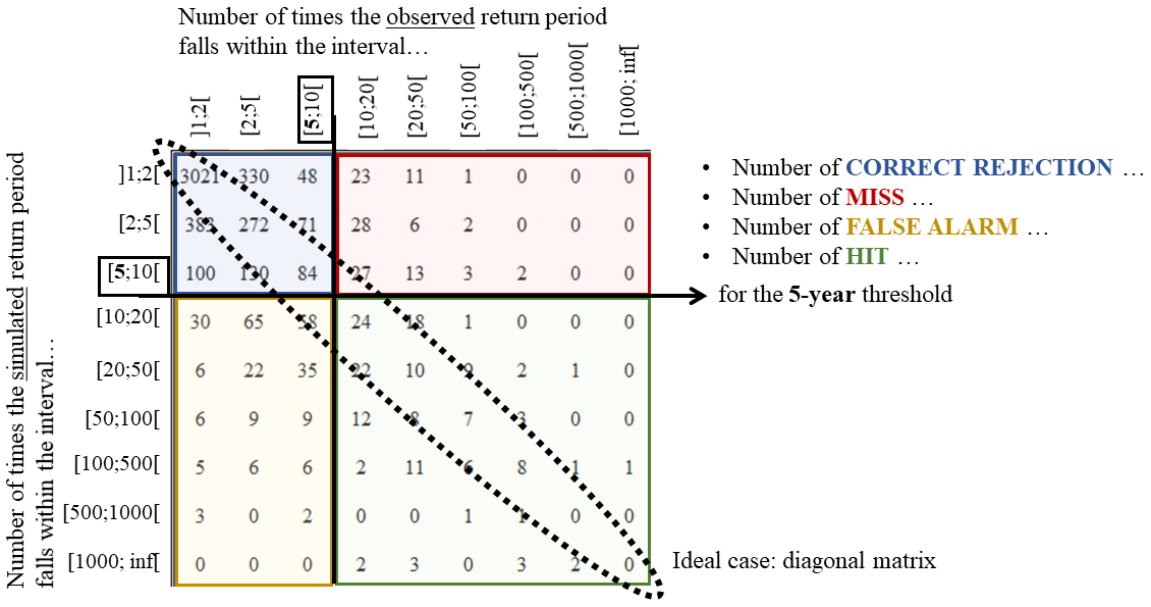


FIGURE II.6 – Illustration of the counting of threshold exceedances for each return period, a step towards the filling of the contingency table.

representing the number of events can be obtained using the Pearson correlation coefficient. Specifically, if X (resp. Y) is a random variable describing the event severity, and thus takes values in [0,8] (0 if the peak discharge return period of the observed or simulated event does not exceed 2 years, and 8 if it exceeds

1000 years), then the matrix of figure II.6 is generated by incrementing the count at position (X, Y). If X and Y are perfectly correlated, the matrix will be diagonal and the Pearson coefficient will be equal to 1. We therefore propose to complement the classical contingency scores with this measure, which provides a good synthesis score (see appendix 2 for more details).

3.4 Impact-based contingencies

The emissions of CatNat decrees were simulated for the period 2009-2023. Following the principle of Vigicrues Flash, the return periods were estimated for each river reach. According to Table II.3, an "Event observed" corresponds to the actual issuance of a CatNat decree in a municipality, while an "Event simulated" occurs when at least one peak discharge return period on one river reach crossing the municipality exceeds the 5-year or 20-year return period (even if these thresholds differ from the 10-year threshold used to release a CatNat decree, we chose here to keep the thresholds used in the operational Vigicrues flash system). This validation process is thus conducted at the municipality level. The duration of observed events, as stated in the CatNat decrees, averages 2 days, whereas a simulated event lasts between the first and last instance a river reach in the municipality exceeds the 2-year threshold (averaging 1.75 days). A HIT is recorded if a CatNat decree is issued and if at least one river reach crossing the municipality exceeds the 5-year or 20-year return period during the overlap of these two periods.

3.5 Determination of rare discharge quantiles based on the SHYREG method

In Vigicrues Flash, the discharge thresholds used to define warning levels correspond to the 5- and 20-year return periods, determined by fitting a Gumbel distribution to the annual peak flows extracted from an extensive chronicle of model simulations. However, for validation purposes, we aimed to assess whether the model could also identify more severe events with return periods exceeding 20 years. Given the 25-year simulation period (1998–2023), limited by the availability of rainfall data, the Gumbel distribution appears unsuitable for estimating flood discharges for higher return periods [Odry and Arnaud, 2017]. Therefore, we opted to calibrate the SHYREG method for each available discharge series to estimate flood quantiles for return periods ranging from 2 to 1,000 years.

3.5.1 The SHYREG approach

The SHYREG (REGionalized SHYpre model) approach is a regional flood frequency analysis method which aim is to simulate regional rainfall and estimate flood quantiles from 2 years to 1000 years of different durations at both gauged and ungauged sites. It combines a regionally calibrated stochastic hourly rainfall generator with a GR-type gridded rainfall-runoff model (grid scale : 1km^2) [Arnaud et al., 2014, Arnaud and Lavabre, 2002]. The rainfall-runoff model is a conceptual model with reservoirs [Edijatno and Michel, 1989, Folton and Lavabre, 2007], simplified to have a unique parameter to calibrate locally : the initial filling rate of the production reservoir, denoted as S0/A. This parameter can be considered as a non-linear runoff coefficient dependent on the hydric condition of the studied area. Assuming that the parameter distribution has already been determined, rainfall events generated by the rainfall generator can then be transformed into flood events by the rainfall-runoff model. By performing a stochastic frequency analysis of simulated flood events over a 10,000-year period [Arnaud et al., 2016], we can obtain flood quantiles from 2 to 1,000 years. This procedure enables to estimate the distribution of quantiles at the calculation unit (grid cell). Quantiles at catchment outlets are estimated by summing upstream grid cell quantiles with a reduction function that considers catchment area and flow quantile

duration, which coefficients have been calibrated for catchment areas ranging from 1 to 10 000 km² in previous studies [Aubert et al., 2014, Aubert, 2012, Fouchier, 2010]. This process has been repeated for a large range of values of S0/A (between 0 and 1).

3.5.2 Calibration on observed discharge series

The calibration of the SHYREG approach involves determining the optimal spatial distribution of the S0/A parameter. Arnaud et al. [2014] developed the SHYREG database by adjusting the S0/A value for each catchment to ensure that the simulated SHYREG flow quantiles closely match the observed flow quantiles for commonly occurring events. This SHYREG database has been used to obtain discharge quantiles associated with streamgauge series and post-event peak discharge estimates.

3.5.3 Calibration on SMASH discharge series

In this study, we recognize that the choice of a lag0 routing module in our SMASH model results in the hydrograph not being accurately reproduced and leads to a tendency to overestimate peak flows. To address this, we systematically calibrated the S0/A parameter based on SMASH-simulated discharge series, ensuring that the SHYREG flow quantiles align closely with the SMASH flow quantiles for return periods of 2, 5, and 10 years, as illustrated in Figure II.7. After completing the calibration, the optimized S0/A parameter distribution was used to simulate flood quantiles for return periods up to 1,000 years at the study sites.

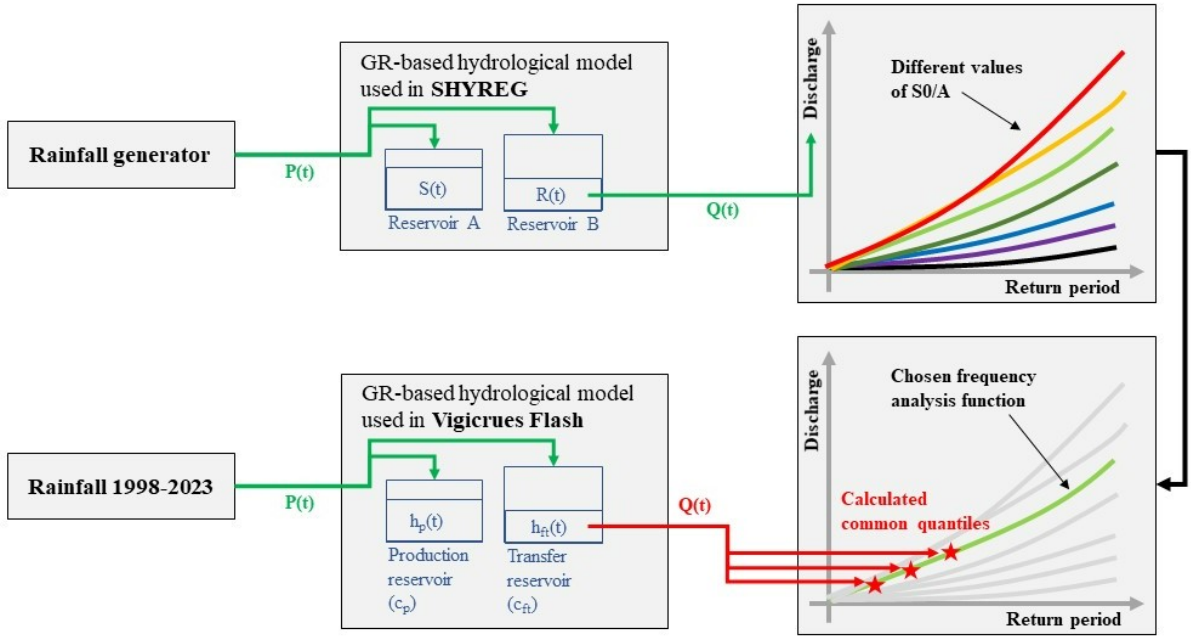


FIGURE II.7 – Simplified illustration of the frequency analysis method (inspired by Arnaud et al. [2014]).

For a deeper understanding of the SHYREG approach, including model calibration, regionalization, and other enhancements, refer to studies such as Ding and Arnaud [2022] and Ding and Arnaud [2024]. Despite its limitations, previous research, including Odry and Arnaud [2017], has demonstrated that SHYREG outperforms the Gumbel fitting approach used in the operational Vigicrues Flash model. It is important to note that, although both the SHYREG and Vigicrues Flash models belong to the GR family of hydrological models, they are not identical, which is a limitation of the proposed approach. The key distinction is that the SHYREG model provides a synthetic and statistical representation of extreme floods, whereas SMASH is designed for real-time flood forecasting. Therefore, SHYREG is event-based

and lumped, while SMASH operates as a continuous and distributed model, making them unsuitable for direct substitution. However, ongoing research is exploring ways to adapt the SHYREG method to the SMASH model.

4 Results

4.1 Flood intensity-based validation

The results of the comparison between the estimated return periods of the simulated and observed peak discharges (either at gauging stations or stemming from the HyMeX post-flood surveys at ungauged sites) are presented in tables II.4, II.5, B.1 and B.2. Several comments can be made :

- Aggregated results across all return periods show higher performance for the validation based on peak discharge (CSI = 73%) compared to the validation at gauging stations (CSI = 41%), primarily due to the reduced number of false alarms. This outcome can be attributed to the fact that post-event peak discharge estimates focus on river sections where significant floods were observed, thereby significantly limiting the number of false alarms. Nevertheless, the CSI scores in both cases are comparable to, or even exceed, those of other models [Zanchetta and Coulibaly, 2020].
- For each score, the results of the first and second lines (validation VS calibration periods) are generally similar, which is satisfying.
- For each score, the results are generally very similar for eligible and non-eligible river reaches (see table 1 for the criteria used to define eligibility), for the low return periods (lower than the 10-years). For more severe floods, the differences become slightly more pronounced, but neither eligible nor non-eligible river reaches systematically perform better (depending on the calibration/validation period, and the data used for validation).
- The CSI scores clearly decrease as the discharge threshold increase, which is consistent with the theory developed by Schaefer [1990], according to which the CSI depends on the frequency of the events to be detected.

The corresponding Pearson correlation coefficients (PCCs) are provided in table II.6. As a reminder, this score provides a measure of "how diagonal" the matrix from figure II.6 is, i.e. how much the obs. and sim. return periods are consistent (i.e. the return period ranges are correlated). Thus, it can be considered as a synthesis score of all the contingency scores presented in tables II.4 and II.5. As in the previous observations, these results do not show a clear difference in the performance of the rainfall-runoff estimates between eligible and non-eligible river reaches. For the evaluation based on gauging stations, the PCCs are quite satisfactory and slightly higher for stations located on eligible river reaches (about 0.7 VS 0.6). However, they decrease for the evaluation based on peak discharge estimates, and the tendancy is reversed : better scores are obtained on non-eligible river reaches.

Finally, the anticipation delays of the Vigicrues Flash system have been estimated by calculating the difference between the moments of the first threshold exceedance for obs. and sim. discharges, using the gauging station data. The values are plotted on figure II.8. The results show that, amongst all "HITs", 20 % threshold exceedance were detected concomitantly or after the observed threshold exceedance, which, in terms of forecasting, would account for a "MISS". In other words, the model anticipates flood peaks 80% of the time without being forced with rainfall forecasts. This is "on the safe side," as incorporating rainfall forecasts (as done in the study by [Godet et al., 2023]) would likely improve anticipation further. However, 60% to 70 % of threshold exceedance were detected at least 1h before the observed events, which is a higher percentage than the one estimated according to the 9 subscribers who provided a feedback on this

question, see figure B.1c in appendix A. Finally, figure II.8a and II.8b being very similar, it seems that the criteria for eligibility do not influence the anticipation delays.

TABLEAU II.4 – Contingency results for the validation based on discharges recorded at gauging stations (bold, top of each line : on validation period 2018-2023 ; bottom of each line : on calibration period 2008-2018 ; orange : rivers eligible for Vigicrues Flash ; blue : rivers non eligible for Vigicrues Flash. The last line indicates the aggregated CSI for eligible / ineligible river reaches and calibration / validation periods.)

	2y		5y		10y		20y		50y		100y		500y		1000y	
hit	**181**	**72**	**79**	**33**	**45**	**11**	**32**	**3**	**15**	**1**	**7**	**1**	**0**	**0**	**0**	**0**
	237	112	100	43	36	10	12	3	2	2	0	1	0	0	0	0
miss	**75**	**29**	**29**	**22**	**15**	**11**	**6**	**3**	**5**	**1**	**2**	**0**	**1**	**0**	**0**	**0**
	103	38	39	24	33	15	13	8	4	2	0	0	0	1	0	1
fa	**120**	**47**	**75**	**29**	**64**	**19**	**42**	**16**	**30**	**10**	**19**	**7**	**6**	**0**	**3**	**0**
	124	77	89	44	54	27	29	15	15	6	8	2	3	0	3	0
cr	**897**	**353**	**1090**	**417**	**1149**	**460**	**1193**	**479**	**1223**	**489**	**1245**	**493**	**1266**	**501**	**1270**	**501**
	1487	534	1723	650	1828	709	1897	735	1930	751	1943	758	1948	760	1948	760
CSI	**0.48**	**0.49**	**0.43**	**0.39**	**0.36**	**0.27**	**0.40**	**0.14**	**0.30**	**0.08**	**0.25**	**0.13**	**0.00**	**NA**	**0.00**	**NA**
	0.51	0.49	0.44	0.39	0.29	0.19	0.22	0.12	0.10	0.20	0.00	0.33	0.00	0.00	0.00	0.00
CSI	0.50		0.42		0.30		0.27		0.21		0.19		0		0	

TABLEAU II.5 – Contingency results for the validation based on estimated peak discharges (bold, top of each line : on validation period 2018-2023 ; bottom of each line : on calibration period 2008-2018 ; orange : rivers eligible for Vigicrues Flash ; blue : rivers non eligible for Vigicrues Flash.The last line indicates the aggregated CSI for eligible / ineligible river reaches and calibration / validation periods.)

	2y		5y		10y		20y		50y		100y		500y		1000y	
hit	**53**	**24**	**53**	**21**	**51**	**18**	**48**	**14**	**36**	**6**	**28**	**4**	**12**	**4**	**3**	**2**
	85	45	79	44	67	43	57	37	42	29	24	21	5	1	4	0
miss	**6**	**0**	**6**	**0**	**7**	**1**	**7**	**0**	**15**	**3**	**21**	**3**	**11**	**0**	**13**	**1**
	1	2	5	2	6	1	6	3	9	5	11	3	7	4	4	2
fa	**0**	**2**	**0**	**3**	**1**	**4**	**1**	**7**	**1**	**8**	**1**	**6**	**9**	**2**	**11**	**2**
	1	0	3	0	10	1	15	5	15	8	23	11	16	9	9	9
cr	**0**	**0**	**0**	**2**	**0**	**3**	**3**	**5**	**7**	**9**	**9**	**13**	**27**	**20**	**32**	**21**
	0	0	0	1	4	2	9	2	21	5	29	12	59	33	70	36
CSI	**0.90**	**0.92**	**0.90**	**0.88**	**0.86**	**0.78**	**0.86**	**0.67**	**0.69**	**0.35**	**0.56**	**0.31**	**0.38**	**0.67**	**0.11**	**0.40**
	0.98	0.96	0.91	0.96	0.81	0.96	0.73	0.82	0.64	0.69	0.41	0.60	0.18	0.07	0.24	0.00
CSI	0.95		0.91		0.85		0.78		0.64		0.48		0.28		0.15	

TABLEAU II.6 – Results of the Pearson correlation coefficient

	Eligible for Vigicrues Flash	Peak discharge estimates	Gauging stations
Calibration	Yes	0.38	0.69
	No	0.56	0.61
Validation	Yes	0.41	0.70
	No	0.70	0.61

4.2 Impact-based validation

The comparison between hypothetical Vigicrues Flash warnings—i.e., the warnings that would have been issued if the service had been operational—and CatNat decrees from 2009 to 2023 is shown in Figure II.9 and table II.7, at the municipality level.

This comparison highlights the challenge of establishing a warning threshold that aligns with CatNat declarations. A 5-year threshold results in too many false alarms, while a 20-year threshold reduces false alarms but doubles the number of missed events. However, aligning warning thresholds with CatNat declarations is inherently difficult, as the declaration process depends not only on hazard intensity but also on exposure and the resulting damages. Significant impacts can occur even during relatively small

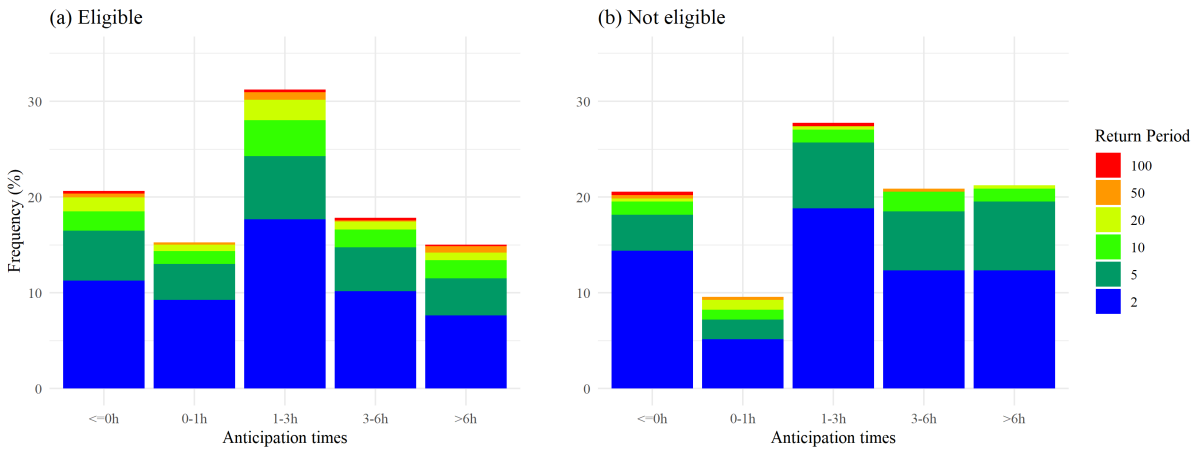


FIGURE II.8 – Histograms of anticipation times for each threshold, for stations located on (a) eligible river reaches and (b) non-eligible river reaches.

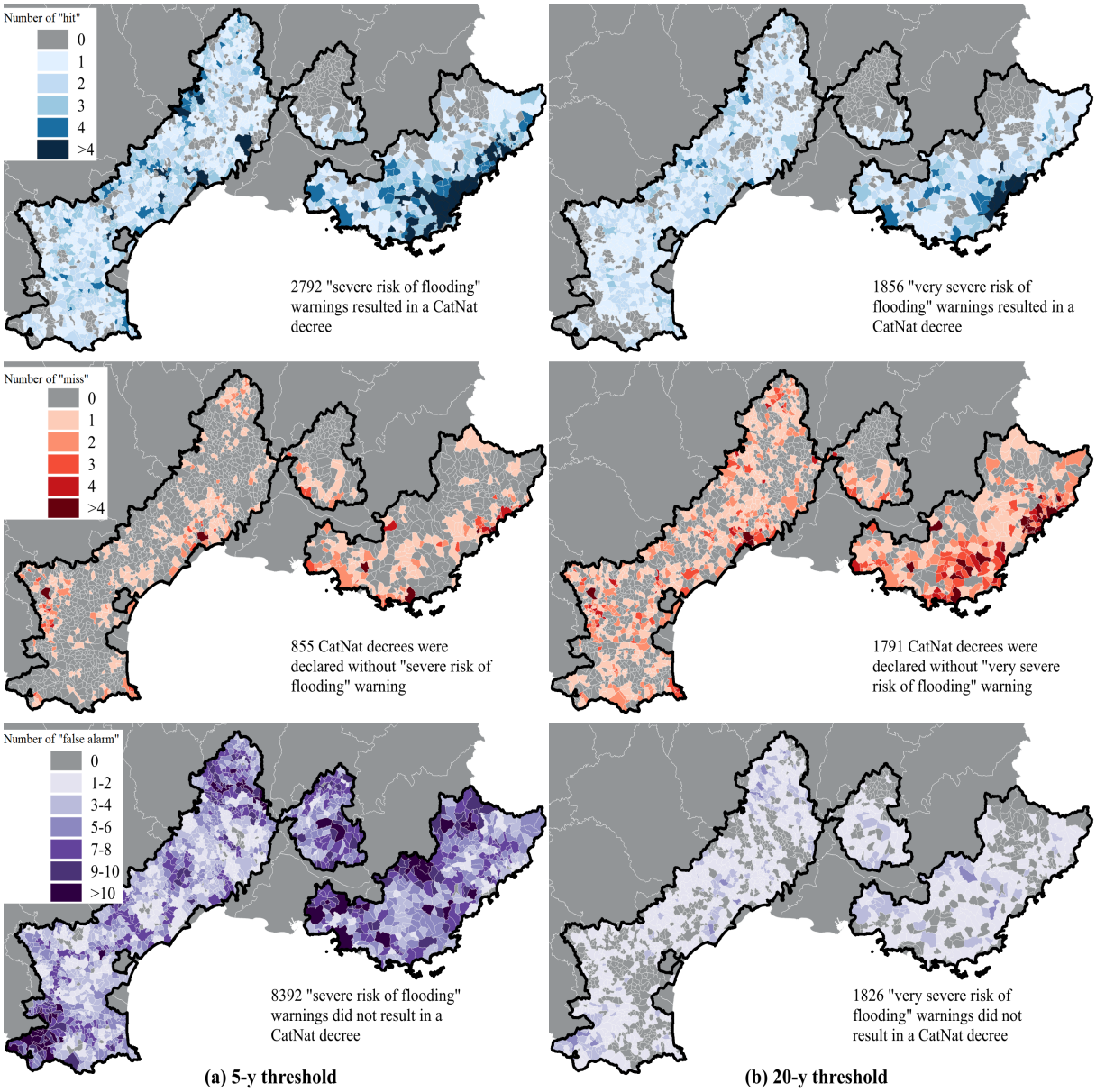


FIGURE II.9 – Results of the validation based on CatNat decrees, for the 5-year and 20-year return period thresholds used in Vigicrues Flash

(a)	5y	20y
hit	2792	1856
miss	855	1791
fa	8392	1826
CSI	0.22	0.35
FAR	0.75	0.50

(b)	5y	20y
hit	255	171
miss	44	128
fa	618	131
CSI	0.28	0.40
FAR	0.71	0.43

TABLEAU II.7 – Aggregated validation results based on CatNat decrees for the 5-year and 20-year return period thresholds used in Vigicrues Flash : (a) for all 1985 municipalities, and (b) for the 109 municipalities equipped with a gauging station. The CSI and FAR scores are detailed in appendix 2.

flood events if exposure is high. It is important to note that the warnings from Vigicrues Flash are not intended to predict the need for a CatNat request by a municipality. This is why the return periods used by Vigicrues Flash do not align with the 10-year criterion in the CatNat decrees. In terms of CSI, the 20-year threshold performs better (35%) than the 5-year threshold (22%), as shown in table II.7a. Although the total number of false alarms for the 20-year threshold may seem high (1,826 false alarms, nearly equal to the number of correctly detected alerts), only 8% of the municipalities studied received more than two false alarms during the 2009–2023 period, which can be considered relatively low.

4.2.1 Analysis regarding hydrological validation

Table II.7b focuses on the 109 out of 1985 municipalities equipped with a studied gauging station. These results can be directly compared to the hydrological-based validation presented in the previous section, as they cover the same locations over an almost identical period. The aggregated CSI scores for gauging stations, shown in Table II.4, indicate a CSI of 42% for the 5-year return period and 27% for the 20-year return period. Notably, these CSI values are similar to those observed in the "CatNat" validation. However, the 5-year return period performed better in accordance with the theory of Schaefer [1990]. For the 20-year return period, the CSI improved in the "CatNat" validation due to a significant decrease in false alarm rates (67% in the "gauging station" validation compared to 43% in the "CatNat" validation). This reduction in false alarms suggests that certain events flagged as false alarms in the hydrological-based validation were, in fact, damaging events when assessed through the impact-based approach. [1]

To focus on the remaining significant false alarms for the 20-year return period in the "CatNat" validation, we isolated the 168 municipalities with more than two false alarms during the 2009-2023 period (8% of all municipalities). We then selected the gauging stations located in these municipalities (13 stations, or 11% of all studied stations) and calculated the "gauging station" False Alarm Ratio (FAR) for the 20-year return period for these stations. The result, 82%, is significantly higher than the FAR aggregated across all stations. This indicates that municipalities with significant false alarms in terms of impact detection also have higher false alarm ratios in terms of flood intensity estimation, which is consistent.

4.2.2 Analysis regarding the user survey

The large number of false alarms suggests that the system reacted in situations where no significant local damage occurred, as the purpose of the CatNat system is to compensate for damage. Figures 1f and

1. ***Complément par rapport au contenu de l'article :*** Pour la période de retour 5 ans, le FAR est de 71% via l'évaluation basée sur les arrêtés CatNat (tableau II.7b), et de 48% via l'évaluation basée sur les débits aux stations (calcul sur le tableau II.4). Cette fois-ci, l'évaluation basée sur les impacts met en lumière des avertissements HBWs considérés comme des "hits" selon l'évaluation hydrologique qui se transforment en "false alarms" selon l'évaluation basée sur les impacts, car une crue de période de retour "modérée" (5-20 ans) ne génère pas systématiquement des dommages. Ces deux résultats montrent donc que l'évaluation d'un système HBW avec des données d'impacts observés apporte un éclairage différent sur les performances du système, en comparaison à une évaluation plus classique mobilisant des données hydrologiques.

1g show that in nearly 50% of cases where a warning was issued, the river reacted but no damage was recorded. This confirms that the occurrence of damage is influenced not only by flood intensity (discharge return period) but also by local vulnerability—an aspect not accounted for by Vigicrues Flash.

4.2.3 Analysis regarding surface flooding in iCatNat

Using additional information from the iCatNat database for the 2018–2023 period, we calculated that 54% of the correctly detected alarms ("hits") for the 5-year return period corresponded to CatNat decrees recognizing damage caused by surface flooding, a hazard not covered by Vigicrues Flash. This could be seen as a "lucky break" or as evidence that the system, with its hazard-based thresholds at the municipal level, accounts for the likelihood that significant water level rises from heavy rainfall often lead to surface flooding. However, we also found that 66% of the "misses" for the 5-year return period were associated with CatNat decrees for surface flooding (compared to 40% of all CatNat decrees), suggesting that excluding surface flooding from the simulation partially explains the high number of missed warnings.

5 Discussion and conclusions

The aim of this paper is to provide an initial quantitative evaluation of the French Vigicrues Flash system, which generates "nowcasts" for small, ungauged river reaches not covered by the main Vigicrues flood forecasting service. Validation was performed using two approaches :

- Flood intensity-based validation, which compared return periods derived from observed discharges (either from gauging stations or post-event surveys) and simulated discharges.
- Impact-based validation, which compared simulated Vigicrues Flash warnings to CatNat decrees.

The flood intensity-based validation spanned two periods : the model calibration period (2008–2018) and the validation period (2018–2023). It also considered two river networks : river reaches currently eligible for the Vigicrues Flash service and those that are not.

5.1 Study limitations

5.1.1 Limitations related to the observation data used as reference

This study faces several limitations related to the observation data used as reference, which are far from perfect. First, while discharge estimates are generally deemed reliable, they rely on (i) rating curves, which are often extrapolated for high discharges, and (ii) peak discharge estimates, which are typically determined months after an event using simplified Manning-Strickler equations, high-water marks, and cross-section measurements. Furthermore, estimating the "observed" return period remains challenging. Here, we utilized SHYREG reference quantiles, commonly used in France for discharge return period estimates. However, these are based on a simplified hydrological model with known limitations [Arnaud et al., 2014]. Lastly, to cover the entire validation region — not just specific points on the river network — we utilized CatNat decrees from the GASPAR database. These decrees, issued at the municipality level, represent areas where insurable damage was reported following a flood event. However, we observed that a significant portion of the CatNat decrees between 2018 and 2023 were related to surface flooding, while Vigicrues Flash only accounts for river flooding. It is also possible that some decrees classified as "riverine flooding" reflect damage on river reaches that are not included in the modeled network, as only catchments larger than 5km^2 are represented. Furthermore, CatNat decisions can be influenced by political factors. It is also important to note that the Vigicrues Flash system was not designed to anticipate

CatNat requests, as reflected in the different threshold return periods used (5 and 20 years in Vigicrues Flash vs. 10 years in the CatNat procedure).

5.1.2 Limitations related to the validation approach

The validation approach also presents certain limitations. The SHYREG method was again employed to determine simulated discharge quantiles, but this came with an additional simplification : SHYREG discharge quantiles were calibrated using SMASH discharge series, despite slight differences in the hydrological models used by both methods. This flood frequency analysis approach is novel and open to debate, but we believe it is not inferior to other methods for estimating flood quantiles associated with simulated discharge series, especially when only short-duration simulated series are available. Another potential criticism of the validation method is that it may overestimate the system's performance by registering a "HIT" even when the simulated peak discharge occurs after the observed peak discharge. As shown in the anticipation time histograms in Figure II.8, this happens in approximately 20% of cases. We argue, however, that the current Vigicrues Flash system is not a real forecasting system but rather a monitoring system, as it does not incorporate precipitation forecasts. For this reason, we refer to Vigicrues Flash products as "warnings" rather than "forecasts" or "early warnings".

5.2 Key results and interpretations

Despite these limitations, this study is a first step toward the assessment of the performance of the Vigicrues Flash system, led seven years after its operationnal implementation. We found globally very encouraging scores in terms of event detection for the validation based on peak discharge estimates, with a CSI of 73% for all return periods combined. Satisfactory results were also obtained for the validation based on discharges recorded at gauging stations : the CSI of 41 %, can be reflected with the CSI=20% result for the evaluation of the Flash Flood Guidance system in the US by in Clark et al. [2014], and with the CSI=38% result found by Gourley et al. [2017] in their study to improve tools for flash flood monitoring in the US. The mean Pearson correlation coefficient (for both calibration and validation periods, and in and outside of the eligible river network) is equal to 58% (64% in Gourley et al. [2017]). It is important to emphasize that the results are consistent between the validation and calibration periods. Additionally, we report a CSI of 35% for the 20-year return period in the impact-based validation, which can serve as a benchmark for future system improvements. The slightly lower performance of the impact-based validation compared to the flood intensity-based validation could be attributed to several factors : difficulties in distinguishing between surface water flooding and river flooding in CatNat decrees, and the potential lack of reported impacts despite certain hydrological thresholds being exceeded, which increases false alarms in the impact-based validation.

It is also important to acknowledge that flood impacts depend not only on flood hazard but also on local exposure and vulnerability conditions, which Vigicrues Flash does not account for. For instance, flood defenses may protect certain river reaches from high-magnitude events, while in other areas, significant impacts might occur even for lower-magnitude events due to high exposure levels or insufficient mitigation measures. This limitation partially explains the discrepancies observed between hydrological and impact-based validation results and highlights the challenge of validating a hazard-based warning system with observed impact data.

This work also aimed at comparing the validation results on rivers monitored by the Vigicrues Flash system, and rivers that are currently not eligible to the Vigicrues Flash service, for hydrological, geological, or geographical reasons (see table II.1). Neither the CSI, nor the Pearson correlation coefficient nor the

anticipation lead times detected a clear difference of system performances between eligible and non-eligible river networks. Therefore, this study reveals that the caution taken in determining the eligibility of rivers to ensure a good performance of the system is probably unjustified, and that the warning system could be extended without too much deterioration in performance.

5.3 Perspectives

Flood monitoring and early warning systems in France are undergoing a substantial transformation. While the Vigicrues service currently provides early warnings (24h ahead) and flood forecasts for large rivers, and the Vigicrues-Flash system provides nowcasts on selected small, ungauged river reaches, the French Ministry for Environment aims to extend the coverage of Vigicrues early warnings to the entire river network. Our results indicate that this expansion appears feasible, even for catchments with a concentration time of less than 3 hours, which is one of the eligibility conditions for Vigicrues Flash. These forecasts are intended to be probabilistic [Belleudy et al., 2024], which could benefit Vigicrues Flash, as demonstrated by Godet et al. [2023]. Additionally, impact-based warnings are being explored to deliver clearer and more actionable information for improved crisis management. In the validation based on CatNat decrees, many false alarms issued by the system could likely be attributed to river reactions that were observed but did not result in physical damage. This highlights the potential benefits that impact-based warning systems could offer. However, impact-based forecasts and warnings present their own challenges, including questions of responsibility (should weather services be accountable for assessing impacts and proposing responses ?), input variables (how can impact data be sourced, especially at local scales ?), as well as cost and feasibility [Potter et al., 2021, 2025]. Unlike hydrological-based warnings that rely primarily on discharge simulation, impact-based approaches require hydraulic computations to assess flooded extents and exposed assets, which can be computationally demanding [Mosimann et al., 2024], although ongoing research is yielding promising results [Najafi et al., 2024]. A faster computation method is to use pre-calculated flood libraries [Mosimann et al., 2024, Alfieri et al., 2024, Dottori et al., 2017, Godet et al., 2025a]. AI-based approaches could also be explored to improve impact prediction ; however, the lack of data makes them difficult to implement for ungauged basins [Nevo et al., 2022].

Furthermore, the issue of validation data will remain a challenge for any future improvements to flood early warning systems. In this study, we used two unconventional datasets : legislative decrees declaring municipalities as natural disaster zones, and discharge data obtained from post-event surveys. Field observations following flood events are invaluable and warrant further development. Although efforts have been made in the French Mediterranean region (e.g., Saint-Martin et al. [2018]), achieving comprehensive data collection remains challenging. Ongoing research is exploring French insurance databases (such as in Le Bihan et al. [2017]) and historical records of fire service interventions, with the results to be presented in a forthcoming paper.

Chapitre III

Évaluation d'une modélisation hydraulique couvrant les petits cours d'eau à une échelle régionale

Objectifs et organisation du chapitre :

Ce chapitre porte sur l'évaluation de scénarios d'inondation pouvant être générés sur les petits cours d'eau méditérranéens et intégrés, sous forme de catalogues préétablis, dans des chaînes d'avertissement de type "Impact-based Warning". L'objectif est de définir un cadre d'évaluation mobilisant différentes sources de données afin d'apprécier la qualité des cartographies de zones inondées obtenues à partir de modèles hydrauliques automatisés, appliqués à l'ensemble des cours d'eau à une échelle régionale. Ce cadre a été appliqué à des catalogues de scénarios couvrant l'arc méditerranéen, établis à une résolution de 5 m pour huit périodes de retour comprises entre 2 et 1000 ans, et construits à partir de différentes données topographiques d'entrée. L'évaluation vise à répondre aux questions suivantes :

- Quels sont les ordres de grandeur des erreurs de hauteur d'eau associées à une modélisation hydraulique simplifiée appliquée à grande échelle pour différents scénarios d'inondation ?
- Quelles sont les principales sources d'erreur ?
- Quel est l'impact du prétraitement des lits mineurs sur les performances de la modélisation hydraulique ?

Ce travail a été valorisé dans un Datapaper publié en 2025 dans la revue internationale *Earth System Science Data* [Godet et al., 2025a]. Cet article décrit les méthodes et résultats de l'évaluation, ainsi que les données d'entrée et de sortie du modèle hydraulique et les données d'évaluation, en accompagnement du jeu de données associé. La première section du chapitre propose une synthèse en français de cet article, et les sections suivantes en reproduisent le texte intégral, rédigé en anglais.

0 Synthèse en français

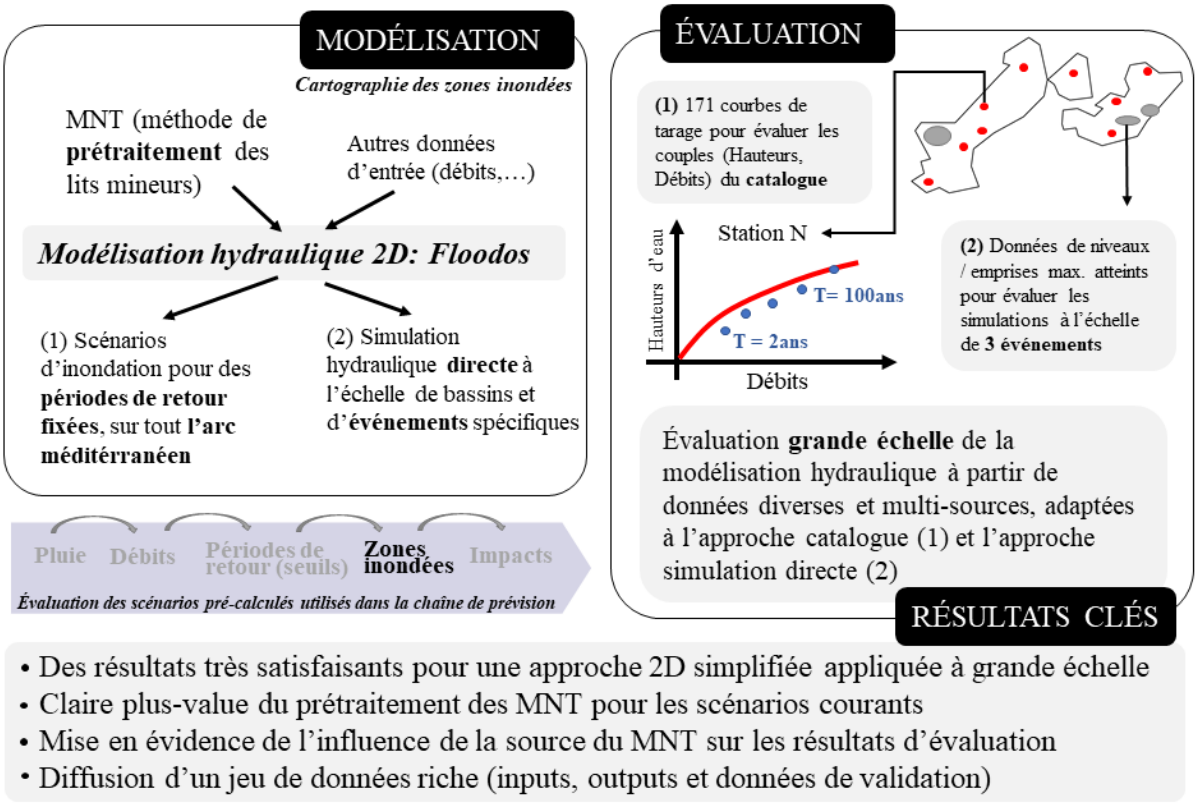


FIGURE III.1 – Résumé graphique de l'article [Godet et al., 2025a].

Cet article propose un cadre d'évaluation et des jeux de données d'entrée/sortie permettant d'évaluer, dans des conditions maîtrisées, des méthodes de cartographies automatisées susceptibles de fournir des scénarios d'inondation sur l'ensemble des cours d'eau de l'arc méditérranéen français (soit 20 000 km de linéaire cours d'eau au total). Deux types de données topographiques d'entrée sont mis à disposition : un MNT brut et un MNT prétraité, destiné à améliorer la représentation des lits mineurs. Ce cadre d'évaluation est appliqué ici à une approche de cartographie basée sur le modèle Floodos [Davy et al., 2017], utilisé en régime permanent, et déjà mobilisé par Nicolle et al. [2021] pour établir des catalogues de scénarios couvrant huit périodes de retour sur l'ensemble de l'arc méditerranéen. Floodos est un modèle résolvant les équations de Saint-Venant en deux dimensions, en négligeant les termes d'inertie. D'après l'étude de Hocini et al. [2021], il a montré de meilleures performances pour la cartographie de trois événements de crues soudaines qu'une approche 1D et qu'une approche dite de « remplissage de MNT ».

Deux approches d'évaluation sont proposées dans l'article : (1) la comparaison de la relation hauteur–débit modélisée avec celles issues des courbes de tarage, et (2) la simulation d'événements réels.

La première évaluation s'appuie sur un ensemble de 171 courbes de tarage collectées sur l'ensemble de l'arc méditerranéen, mises à disposition dans le jeu de données publié. Après un important travail de prétraitement et d'uniformisation, les couples (hauteur d'eau, débit) simulés pour différentes périodes de retour au niveau des 171 stations hydrométriques peuvent être comparés aux courbes de tarage. Les résultats apparaissent très satisfaisants pour une modélisation hydraulique simplifiée à grande échelle : l'erreur médiane (simulation – observation) est de –0,04 m pour le scénario de période de retour 2 ans, et de –0,14 m tous scénarios confondus. Le prétraitement des lits mineurs a nettement amélioré les performances pour le scénario 2 ans (erreur médiane de 0,54 m avec les lits mineurs bruts). Les gammes

d'erreurs (écarts interquartiles) associées à la modélisation utilisant les lits mineurs prétraités se situent entre ±50 cm pour le scénario le plus fréquent et se maintiennent entre ±75 cm jusqu'au scénario de période de retour 50 ans, mais s'élargissent nettement pour les scénarios extrêmes. L'évaluation de ces derniers reste plus délicate car elle s'appuie sur un nombre réduit de données, souvent peu fiables : les courbes de tarage sont généralement extrapolées à partir de jaugeages qui excèdent rarement la période de retour 2 ans.

La deuxième approche d'évaluation s'appuie sur l'étude menée par Hocini et al. [2021], mobilisant des données de Plus Hautes Eaux (PHE) et des emprises observées lors de trois événements de crues soudaines. Cette évaluation présente l'avantage d'examiner à la fois la capacité du modèle à représenter l'emprise inondée maximale et la précision sur les hauteurs d'eau atteintes. Les résultats montrent que les scores CSI, mesurant la précision du modèle dans la délimitation des contours de l'inondation, varient entre 55 % et 85 % pour 90 % des tronçons de rivière évalués, avec une médiane proche de 80 %, ce qui traduit de très bonnes performances. Les erreurs moyennes sur les hauteurs d'eau maximales se situent entre ±30 et ±50 cm, soit un ordre de grandeur cohérent avec les incertitudes généralement associées à ce type d'évaluation [Bates, 2022].

Enfin, ce cadre d'évaluation a permis d'identifier certaines sources d'erreur : fiabilité variable des données de validation, ou encore influence marquée de la qualité du MNT d'entrée, avec des erreurs majeures souvent corrélées à des MNTs produits sans données Lidar.

1 Introduction : context and objectives

Flash floods are a weather-related type of natural hazard, which, in Europe, most frequently occur in the Mediterranean region. Some sub-regions in the north-western Mediterranean seem to experience more frequent extreme discharges [Gaume et al., 2016]. Several of these "hotspots" have been identified in the French Mediterranean area, which has been hit by many disastrous flash floods in the last decade. Recent examples of these events include the 2019 floods in the Var, Alpes-Maritimes, and Vaucluse departments, which caused 13 fatalities and up to EUR 450 million in insured losses, or the 2020 "Storm Alex" event in the Alpes-Maritimes department, which caused 19 fatalities and up to EUR 230 million in insured losses [CCR, 2021]. Although the effects of climate change on the intensity and frequency of flash floods in the French Mediterranean region remain unclear [Tramblay et al., 2019], rapid urbanization and population growth in this area will likely increase the risk of disastrous flash floods. Therefore, the improvement of tools for flood risk management is in high demand, especially since many recent studies emphasize the usefulness of information about potential impacts [Merz et al., 2020, Apel et al., 2022], which requires to develop enhanced flood mapping capacities.

Fortunately, the scientific context is conducive to tackling these issues, as flood mapping is being increasingly improved by recent developments in Digital Terrain Model (DTM) resolution and accuracy—especially thanks to Light Detection and Ranging (LIDAR) technology. For instance, LIDAR has led to a coverage of 99% of England at a 1m spatial resolution [{Environment Agency}, 2024], and the ongoing project LIDAR HD plans to provide DTMs, DEMs (Digital Elevation Models) and DSM (Digital Surface Models) for France at a 0.5m spatial resolution [IGN, 2024]. Additionally, there have been recent advances in the estimation of river channel bathymetry in DTMs, which is unobserved in current products and represents a challenging issue for the future [Lague and Feldmann, 2020, Pricope and Bashit, 2023, Frizzle et al., 2024].

Current flood mapping methods typically rely on two main approaches. The first, commonly used, assumes that an area is flooded if the elevation of the floodplain cell is lower than an assumed or estimated flood water height in the channel, a GIS-based method known as DTM filling. For example, the Height Above Nearest Drainage (HAND) method [Nobre et al., 2011] estimates flood water height based on topographic relationships derived from the terrain and channel structure. This method offers a simple and computationally efficient way to approximate inundation extents. The second approach, which is the focus of this paper, is based on the numerical solution of two-dimensional shallow water equations (SWE), commonly referred to as hydraulic models. While SWE are typically solved in two dimensions, one-dimensional (1D) models have also been used, such as in Le Bihan et al. [2017], Lamichhane and Sharma [2018], due to their shorter computation times and reasonable accuracy. In fact, 1D models share some features with GIS-based methods : while they use a hydraulic model to estimate flood stage along the channel network, they similarly rely on a GIS operation to project water surface elevations onto the terrain to derive inundation extents. However, 1D models have limitations in representing complex floodplain flows. To address this, three-dimensional (3D) models are being explored [Luo et al., 2018], but their real-world application remains limited by complexity and the current lack of calibration and validation data [Bates, 2022]. In short, hydraulic 2D SWE models offer a good compromise between data availability, accuracy, and computation times.

Continental and global-scale 2D flood mapping simulations have been largely performed since the early 2010s, most often at 10^2 m resolutions. Examples include the 100m-European flood maps calculated by Alfieri et al. [2014] and updated by Dottori et al. [2022], and the Global Flood Maps calculated by Fathom, initially at a spatial scale of approximately 90m [Sampson et al., 2015, Andreadis et al., 2022].

However, it can be argued that such large-scale models are not always relevant at local scales [Fleischmann et al., 2019], which are targeted for the representation of small rivers. Consequently, efforts are being made to develop methods applicable at finer resolutions (10^1 m). The latest version of Fathom's global maps [Wing et al., 2023] has been refined to achieve a spatial resolution of approximately 30m, with performance evaluated to be close to local model skill in many cases. Spatial resolutions lower than 10m are especially targeted for areas prone to flash floods, such as the 5m resolution used in Hocini et al. [2021].

Whatever flood mapping method is used, the assessment of its performance generally suffers from a lack of validation data, especially for the simulation of flash floods, which often occur in data-scarce regions. Crucial parameters that need to be assessed in flood mapping studies include flood extents, flood depths, and water velocities, all of which are rarely recorded after a flash flood event for several reasons, including measurement difficulties [Molinari et al., 2019]. Validation methods based on reproducing actual flood events may use observed extensions of the flooded area [Dottori et al., 2017, Hocini et al., 2021, Wing et al., 2023], high water marks [Hocini et al., 2021, Wing et al., 2023], as encouraged by Bösmeier et al. [2022], or flood impacts [Le Bihan et al., 2017, Ritter et al., 2021], and sometimes gauge measurements [Gebrehiwot and Hashemi-Beni, 2022]. Another possible data source for the validation of hydraulic models are the results of simulations performed by experts [Alfieri et al., 2014, Sampson et al., 2015, Albano et al., 2020, Hocini, 2022, Dottori et al., 2022, Wing et al., 2023]. It can consist of historical event simulations or flood hazard maps (e.g., those produced as part of the European Flood Directive). In both cases, the main limitation lies in the input data, which sometimes are unavailable or undocumented. Without access to the original input data used in a study, it becomes challenging to conduct a relevant comparison between models.

The aim of this article is to help fill the gap in validation data by providing a benchmark dataset for flash flood mapping in the French Mediterranean region. Figure III.2 specifies the content of the dataset and the corresponding structure of this paper, which is separated in two parts : flood hazard and flood events. On one hand, we provide inputs for flood hazard calculations in steady state, with an example of results obtained with the Floodos model (section 2.2.1), as well as a large sample of rating curves usable for validating the flood hazard maps (section 2.2.2). On the other hand, we provide inputs for simulating three historical flash flood events, as well as the corresponding Floodos simulation results (section 2.3.1), and the observed flood extents and/or high water levels used for the validation (2.3.2). The provided data can be used to run and evaluate similar simulations (flood hazard maps or specific flood events) with any other hydrodynamic modeling tool. In this paper, we also present the production methods, including the DTM preprocessing method (section 3.1), the Cinecar rainfall-runoff model (section 3.2) and the Floodos method (2D SWE-based approach, without inertia, detailed in section 3.3). Finally, we illustrate the validation results obtained with the Floodos approach (section 4).

2 Data description

This section aims to describe the provided data (see table C.1 for an exhaustive list of the data made available). We first present the areas of the French Mediterranean region covered by the simulations. Then, following the structure of the dataset, we describe successively the data related to flood hazard maps and the data related to historical flood events. In each of these two sections, the data are presented according to their role in hydraulic modelling : input/output or validation data.

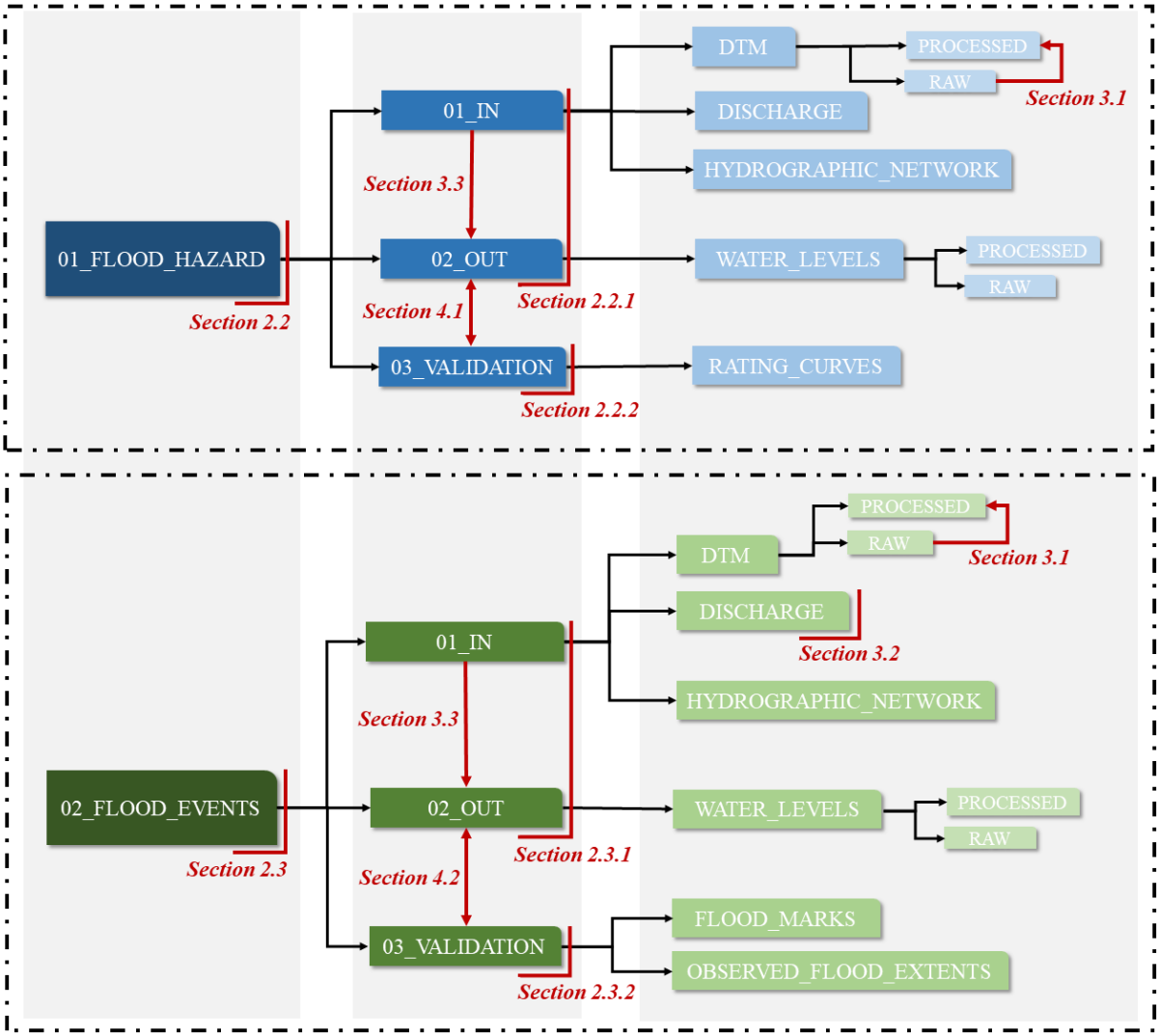


FIGURE III.2 – Dataset structure and organisation, and related sections of the data paper. The signification of each item is summarised in table C.1.

2.1 Study areas

The whole simulation area covers 61 elementary hydrological areas (watersheds grouping together nearly 300 sub-basins of sizes between 20km^2 and 750km^2, 140 km^2 on average) located in the French Mediterranean region, which represent nearly 42,000 km^2 of drained area and 20,000 km of river network (see figure III.3). The Rhône and Durance rivers were excluded from the study area because they are seldom impacted by flash floods and are significantly influenced by artificial structures, such as dams. The area covers a wide range of climatic patterns, hydrological and geological properties, and population densities. It has been subject to numerous flash floods in the last decades, especially the floods in the Argens watershed (15 June 2010), in the Alpes Maritimes department (3 October 2015), and in the Aude watershed (15 October 2018), which are studied in this paper. These floods led to up to EUR 710 million of insured losses and 25 victims for the most devastating event (June 2010 flood). Detailed descriptions of each event with meteorological and hydrological information can be found in Hocini et al. [2021], whose case studies, input and validation data have been integrated in this article.

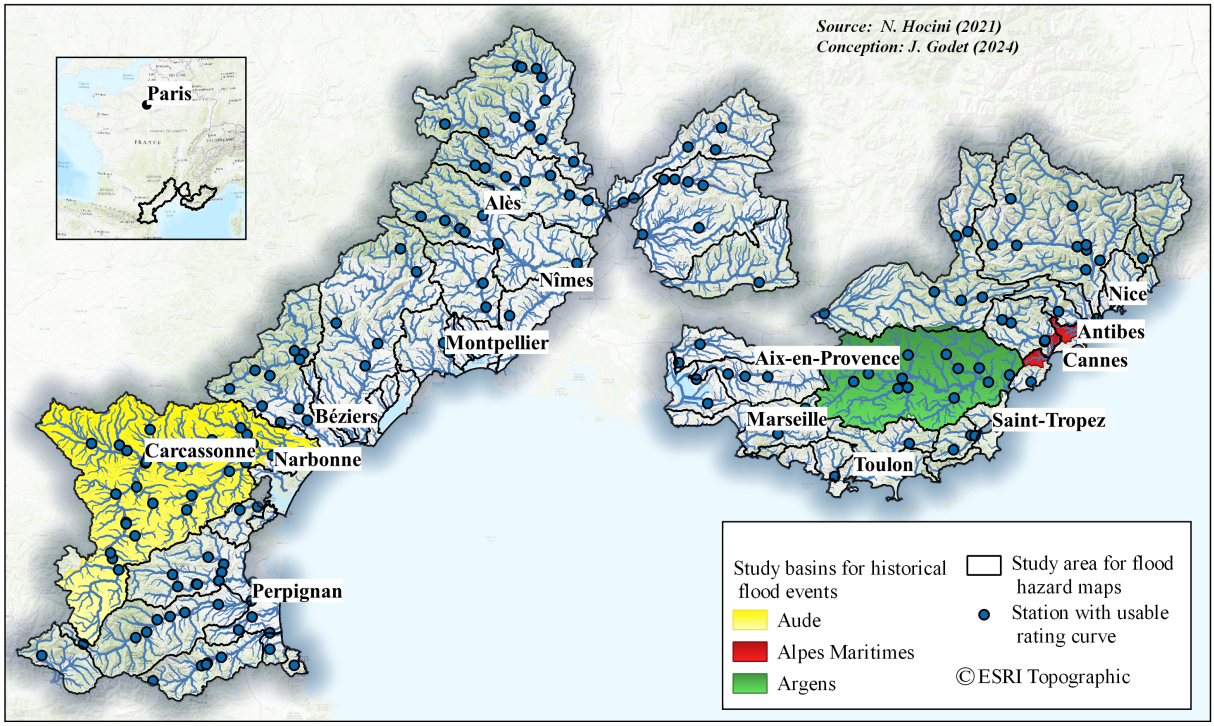


FIGURE III.3 – Study basins for the simulation of historical events (coloured), study areas for flood hazard maps (framed in black) and the selected stations for rating curve data.

2.2 Computation and validation of flood hazard maps

Flood hazard maps corresponding to eight return periods between 2 and 1000 years were generated on the whole simulation area. The 2D Floodos computation results included in the dataset were obtained based on the following input data : the Digital Terrain Models (DTMs), the vector description of the hydrographic network, and the peak discharges for each return period. For the asessment of modeling results, we provide in the dataset rating curves from stream gauges in the study area. Figure III.4 illustrates the computation and validation procedures of flood hazard maps.

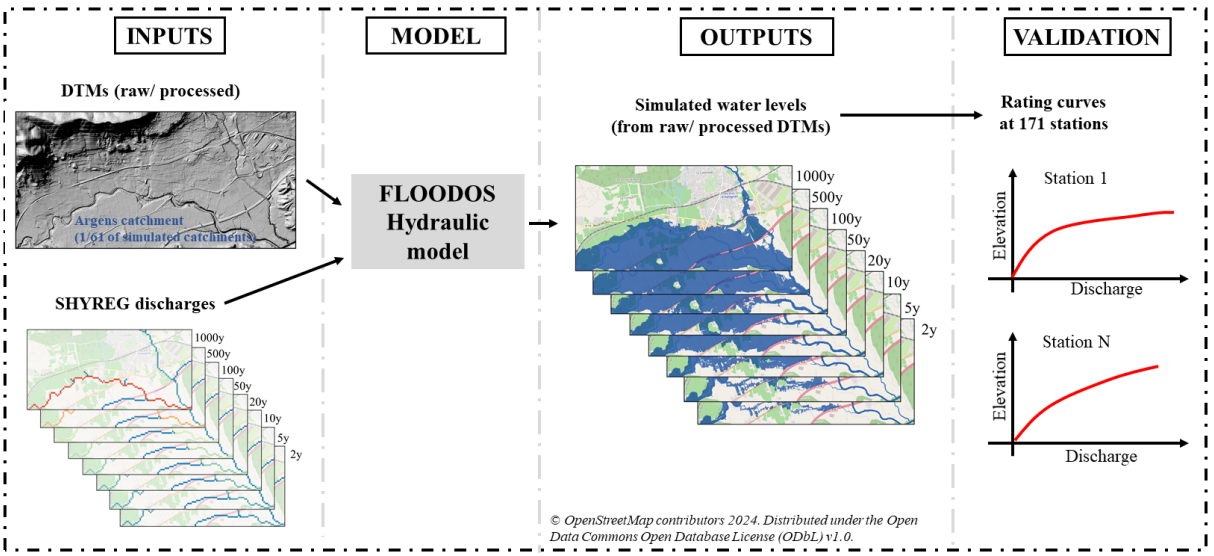


FIGURE III.4 – Illustration of the computation and validation of flood hazard maps

2.2.1 Flood hazard mapping inputs and outputs

The DTM used is the RGE ALTI® produced by IGN (Institut national de l'information géographique et forestière, https://geoservices.ign.fr/rgealti), with a spatial resolution of 1m, that includes the latest lidar measurements available in October 2022 (date of extraction from IGN repository). It must be noted that the entire study area was not covered by a lidar measurement at the time of the calculations (see Figure III.11). This DTM provides average altitude information for each 1m × 1m pixel. It was used as is for the DTM cleaning process (see section 3.1), and it was resampled at the resolution of 5m × 5m for the hydraulic calculations. The 5m DTM directly obtained from the resampling is referred to as "raw", while the 5m DTM that underwent a cleaning process is referred to as "processed".

The hydrographic network provided in the dataset is a raster description of the streambeds, derived from BD TOPAGE®, which is the French reference hydrographic network database, accessible at https://www.data.gouv.fr/fr/datasets/bd-topage-r/ (last accessed 14/03/2025). Based on the provided raster description of the streambeds, two different Manning's friction coefficient values were used in the hydraulic simulations : $n = \frac{1}{10}s.m^{-1/3}$ for the flood plains, and $n = \frac{1}{18}s.m^{-1/3}$ for the streambeds, for all river reaches. These values have been determined with a sensitivity analysis led by Hocini [2022] in the study area.

In order to carry out modelling of flood hazard over the whole computational domain, it is necessary to gather discharge data for defining the inflow boundary conditions of the hydraulic model. This discharge data was extracted from the SHYREG (REGionalized SHYpre model) database, and is included in the dataset. SHYREG is a national database providing synthetic rainfall intensity and peak discharge quantile estimates for the whole French territory [Arnaud et al., 2014, Aubert et al., 2014]. The SHYREG quantiles are derived from stochastic simulations combining an hourly stochastic rainfall generator and a simplified distributed rainfall–runoff model, respectively calibrated against the existing raingauge and streamgauge measurements. The database includes the peak discharge quantiles for 8 return periods of 2, 5, 10, 20, 50, 100, 500, and 1000 years. These discharge quantiles are provided continuously along the hydrographic network, for drainage areas larger than 5 km^2, on a 50 m resolution grid. Therefore, this method can provide discharge information for small ungauged river reaches, which is crucial for evaluating risks related to flash floods. Like all methods, estimations provided by SHYREG are subject to uncertainty, due to major simplifications in the model, concerning, for instance, the influence of hydraulic regulation structures, karstic zones, rain and discharge data uncertainty, etc. Caruso et al. [2013] thus provide confidence levels to qualify the applicability of the method on each stream network in France. Details on how the SHYREG quantiles were incorporated into the Floodos computational domains can be found in section 3.3 and appendix 4.

The flood mapping procedure described in section 3.3, applied to the aforementioned data, yielded 61 × 8 water level rasters (one for each of the 61 sub-domains of figure III.3 and for each of the 8 return periods) at the 5m × 5m resolution. Initial results, derived from the same hydraulic modelling approach but with different input data (i.e. different DTMs that were pre-processed using a much more simplified approach than the one described in section 3.1), showed a large overestimation of the flooded areas for the 2-year return period, mainly attributed to residual defects in representing the geometry of the streambeds in the DTMs [Hocini, 2022]. This justifies the application of a specific DTM processing to improve the representation of the streambeds. Figure C.1 illustrates the impact of this processing procedure on river streambeds. For the sake of comparison and illustration of the efficiency of the DTM preprocessing method, results obtained from both raw and processed DTM are included in the dataset. A quantitative assessment of the results based on stage-discharge rating curves available at gauging stations over the studied river networks is carried out in section 4.1, and shows that the DTM preprocessing method largely

address the overestimation for low return periods.

2.2.2 Validation data : Rating curves

As discussed in the introduction, the assessment of a set of flood hazard maps for different return periods is highly limited by the lack of validation data. The proposed approach leverages the expertise of local hydrometry services, which is summarized in the rating curves available at streamgauge river sections. Such an evaluation was already applied by Le Bihan et al. [2017], who examined gauging data and rating curves from twelve hydrometric stations. It was applied here for a much broader assessment, at a regional scale ; on 171 gauging stations selected as follows. Initially, both gauged (i.e. measured) discharges and rating curves were considered. However, due to the very limited number of measured discharges corresponding to return period exceeding 2 year (only 2% of the total amount of gauged discharges, spread over 35% of the gauging stations), it was decided to focus on the rating curves. The rating curves are also subject to uncertainty, as they are derived from measured discharge data, which themselves carry significant uncertainty [Di Baldassarre and Montanari, 2009]. Notably, tools exist to quantify the uncertainty associated with hydrometric rating curves [Le Coz et al., 2014]. In this study, we argue that they serve as a convenient and independent source of data for evaluation despite their inherent uncertainties, which should be considered when analyzing the results. They were provided by SPC Grand Delta, SPC Méditerranée Ouest, and SPC Méditerranée Est, either through the open-access Hydroportail platform (`https://hydro.eaufrance.fr/`, last accessed : April 7th, 2024), or directly by the data producers through the BAREME tool [Bechon et al., 2013], when the data was not available on the Hydroportail platform (representing roughly 1/3 of the stations). The data provided in the dataset do not correspond to the raw data extracted from Hydroportail/BAREME - they have undergone a procedure to provide uniform (data extracted from Hydroportail are not formatted the same way as data extracted from BAREME), verified (some errors in the data could be detected), suitable (for flood hazard maps verification purposes), and simplified (keeping only the variable of interest) information. The procedure is detailed in appendix 2.

The application of this procedure reduced the initial number of gauging stations of interest from 418 to 171. The 171 selected stations are mapped in Figure III.3, and Table III.1 summarises the number of stations for which the rating curve providing estimated stage-discharge values for each of the eight estimated return periods.

T (years)	2	5	10	20	50	100	500	1000
N	171	164	155	140	112	85	33	20

Tableau III.1 – Number of stations for which the rating curve has stage-discharge values exceeding each return period.

The data provided in the dataset include metadata such as station name, code, location, reference altitude of the water level gauge, etc., but also a "Quality index for high flow data" which is a qualitative score (between 1 and 3, 3 : reliable data, 2 : moderately reliable data, 1 : uncertain data) that was estimated based on discussions with the local forecasting services. It quantifies the precautions that need to be taken when interpreting the gap between the model simulations and the rating curves for high return periods. Furthermore, reliability zones (i.e. water levels between which the data is deemed reliable by the local forecasting service) which were directly extracted from Hydroportail / BAREME are also provided. Finally, despite this careful data selection, multiple errors can remain and uncertainties cannot be avoided, and this is particularly true when working with data as sensitive as the rating curves, which are often

extrapolated from direct gauge measurements below the 2-year return period.

The provided rating curves can take two forms : either as "piecewise linear" curves based on linear interpolation of stage-discharge pairs, or as a "power function" $Q = \alpha \times (H - \beta)^{\gamma}$, where Q is the discharge in m^3/s, H the water stage over the zero reference altitude of the water level gauge, and α, β, γ are parameters to be determined. Rating curves are constructed across a large range of discharge rates, although gaugings may not be available for the entire range. Return periods are not provided by local hydrometric services and are determined independently using the SHYREG method.

2.3 Computation and validation of actual extreme flood events simulations

Hocini et al. [2021] compared three automated methods for flash flood inundation mapping, based on the simulation results of three observed flood events. In this section, we present the data associated with these three events (see Figure III.5 for an overview of each dataset's role in the simulation chain). The input discharges and validation data (observed inundation extents and high water marks) are shared with the study by Hocini et al. (2021). However, the input DTMs and the resulting simulated water levels differ, highlighting the impact of DTM preprocessing on hydraulic results.

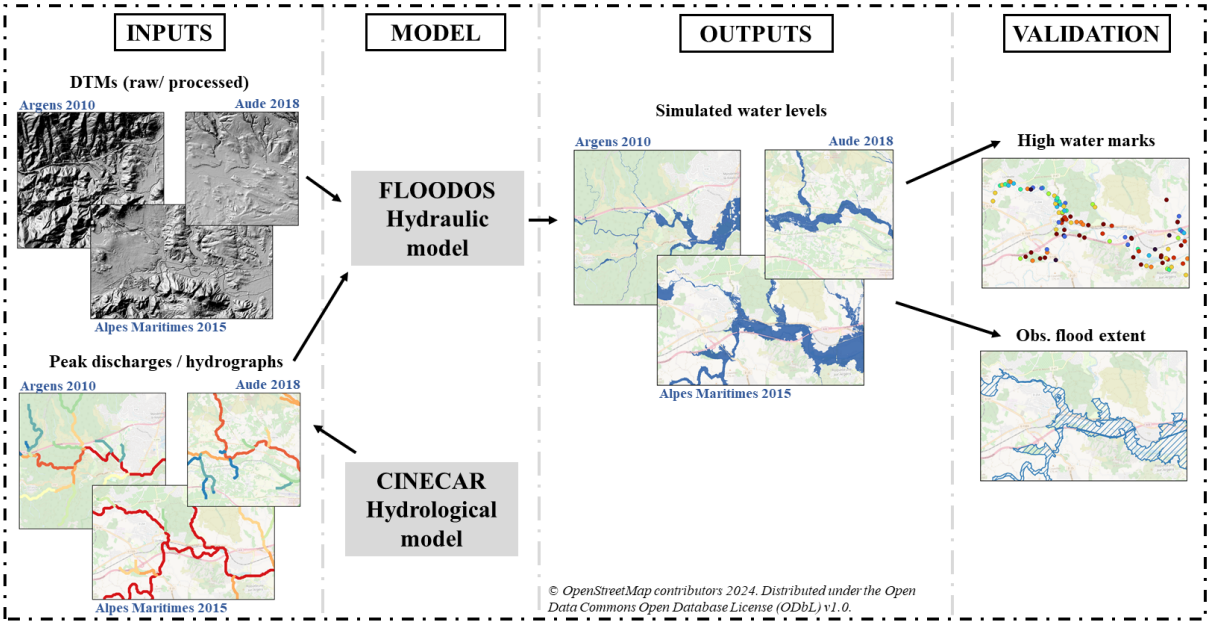


FIGURE III.5 – Illustration of the simulation and validation of extreme flood events

2.3.1 Flood mapping inputs/ outputs

DTMs and hydrographic networks are the same as those described in section 2.2.1, except that the Siagne and Brague watersheds were merged to enable the simulation of the Alpes Maritimes event. Manning's friction coefficients were set to $n = \frac{1}{18} s.m^{-1/3}$ in the streambed and $n = \frac{1}{10} s.m^{-1/3}$ in the flood plains. Computations were performed in steady-state regime, using simulated peak discharges from Cinecar rainfall-runoff model (see section 3.2 for more details). Discharges are considered homogeneous on each river reach, of average length lower than 2km. We provide the full hydrographs at each river reach outlet, as well as the lateral inflow hydrographs, to enable unsteady computations for future users.

The water level maps (flood inundation extents and depths) generated with the Floodos model that are included in the dataset, correspond to the maximum of inundation according to the peak discharge of

the flood, they may thus correspond to different instants of the flood for the various reaches of the river network.

2.3.2 Validation data : Flood marks and flood extents

As mentioned earlier, the validation data used to assess the quality of flood event simulations originate from the study by Hocini et al. [2021]. However, they are properly described and published for the first time in this datapaper. We also repurpose these data to compare the impact of DTM preprocessing on hydraulic results (see Section 4.2). Two types of data are provided : high water marks (HWMs) and reconstituted flood extents from field observations.

HWMs were extracted from the French HWM database (last access : April 19th, 2024). We provide 642 HWMs for the 2010 Argens event, 556 HWMs for the 2015 Alpes Maritimes event, and 1089 HWMs for the 2018 event. To ensure that collected HWMs only correspond to river flooding, HWMs were manually filtered according to their proximity to the modelled river network. While these data should not include large errors as they were systematically checked, they are not immune to occasional problems. For instance, amongst the HWMs collected for the 2010 Argens flood event, systematic errors in altimetric referencing were observed on a 28km-long river reach of the Argens river, running upstream from the north-west of the town of Vidauban. It was particularly obvious since some HWMs were clearly under the natural terrain level, as revealed by a simple comparison with the DTM. All these detected errors concerned HWMs surveyed by the same firm at the same date. While these errors may have been caused by human error, other ones can also be the result of local obstacles affecting water surface elevation, capillary rise of moisture in walls, instrumentation limitations, etc. The detected faulty HWMs were systematically removed from the dataset.

The reconstituted flood extents provided in the dataset are also a combination of maps obtained from several firms of engineering consultants, commissioned by local authorities. Reconstitution methods were generally based on field surveys, inventories and locations of HWMs, photos, videos, and satellite images. Limits can come from local interpolations between field observation points. These maps are available for the Argens 2010 event and the Aude 2018 event, however the Alpes Maritimes 2015 event was characterised by very rapid kinetics in urban environments, thus it was more complicated to conduct the same reconstitution procedure. Local authorities provided a map which are partly the result of hydraulic simulations, thus it was decided not to use it for validation purposes. The flood extent dataset also provides a shapefile of subcatchments under study, which were limited to the river network where the observed flooded areas are available. These subcatchments were calculated based on a 50m flow direction grid and have an average area of $2.5 km^2$. They were designed in order to conduct the evaluation process at a local river reach scale.

3 Description of data prepocessing methods

3.1 DTM preprocessing

As mentioned in section 2.2.1, a specific processing of the DTMs was applied to attenuate or remove several imperfections suceptible to cause large errors in hydraulic computations : imperfect interpolations of DTMs in stream beds leading to noisy cross-sectional and longitudinal bathymetric profiles, missing description of the capacity of covered watercourses and presence of undesirable obstacles to flow in river beds (typically bridges decks incompletely removed). These imperfections may generally lead to overestimate flooded areas particularly for the low (i.e. 2y) return periods.

The DTM processing procedure was based on the work of Kalsron et al. [2019] to "reopen" of covered river reaches, and on the methods of Wimmer et al. [2021] for the delineation of the streambed. They are thoroughly described in a technical report to the French Ministry for the Environment [Nicolle et al., 2023], and shortly presented in appendix 3.

The effects of these DTM corrections on hydraulic computation results were analysed in three main steps. First, hydraulic modeling results samples were visually inspected. An example is provided in figure III.6 for the Brague river. It illustrates several of the targeted imperfections. A better definition of the streambeds' cross-sections clearly limits simulated overflows for the T=2 year floods (two comparisons on the top of the figure) or eliminates the overflow upstream a bridge not eliminated in the original DTM (bottom of the figure). Then, the total number of flooded pixels outside the streambed for the 2-year maps were counted : it should be as limited as possible since it is generally considered that the streambed has the capacity to convey the 2-year flood without overflow. On average on all the stream reaches, the number of inundated pixels outside the streambed for the 2-year return period was reduced by 40% due the DTM processing (minimum decrease of 4%, maximum decrease of 95%, no increase). Finally, validation results for both configurations (raw DTMs and processed DTMs) were compared using the validation data described in Section 2.2.2 (the rating curves). These results demonstrate the statistic relevance of the DTM correction method in river channels, as discussed in Section 4.1.

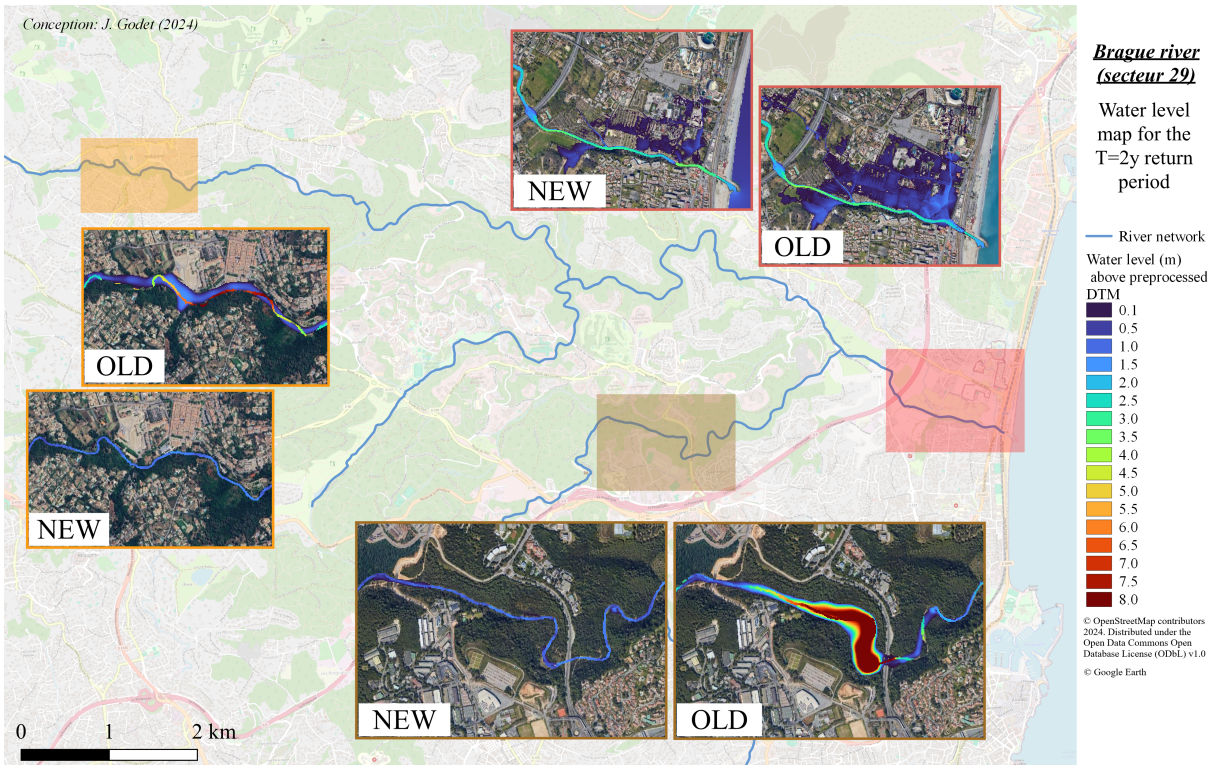


FIGURE III.6 – Comparison of simulation results obtained from raw and processed DTMs for the Brague river and the T=2y return period.

3.2 Hydrograph calculations

The dataset provides flood hydrographs for the three historical flash flood events under study (Argens 2010, Alpes Maritimes 2015, Aude 2018), for each of the considered river reaches. These were calculated using the Cinecar distributed rainfall-runoff model [Gaume et al., 2004, Naulin et al., 2013], calibrated for each flood event based on the available measured and estimated discharges. Cinecar is based on a division of the river network into river reaches (1.5km-long on average in the present cases) that are

each connected to two rectangular slopes, representing the left and right bank subwatersheds. Cinecar simulates the evolution over time of runoff coefficients and volumes on each subwatershed based on the Soil Conservation Service-curve number model (SCS-CN). The computation time step can be adjusted. It has been set to 15 min in this work. To route the effective rainfall through the watershed, the kinematic wave model is used on the hillslopes and in the river network, except for river reaches with slopes lower than 0.6%, for which the Hayami solution for diffusive wave model is used.

For the Argens 2010 and the Alpes Martimes 2015 events, Cinecar was forced with the Antilope J+1 rainfall product [Champeaux et al., 2009], which combines radar and rainfall records. It is an operational product provided by Météo-France the day following the precipitation record date, permitting an incorporation of 40% of additional data from rainfall network compared to the online Antilope rainfall product. This may not always prevent from observing important errors locally (especially in regions scarce in rain gauge networks), which is why the model was forced with an improved rainfall reanalysis for the Aude 2018 event [Caumont et al., 2021]. This product takes advantage of new rainfall data, in particular amateur measurements, and is considered to be more reliable than the initial Antilope J+1 product. The original time step is 1h, thus the rainfall rates were uniformly distributed over the 15min sub time step to feed the Cinecar model. Cinecar was calibrated for each event using all available observations of peak discharges, including a large number of estimations at ungauged sites, gathered within the HyMeX program [Ducrocq et al., 2019]. As Hocini et al. [2021] reported, the resulting relative difference between simulated and observed peak discharges rarely exceeds 20% (see figure III.7).

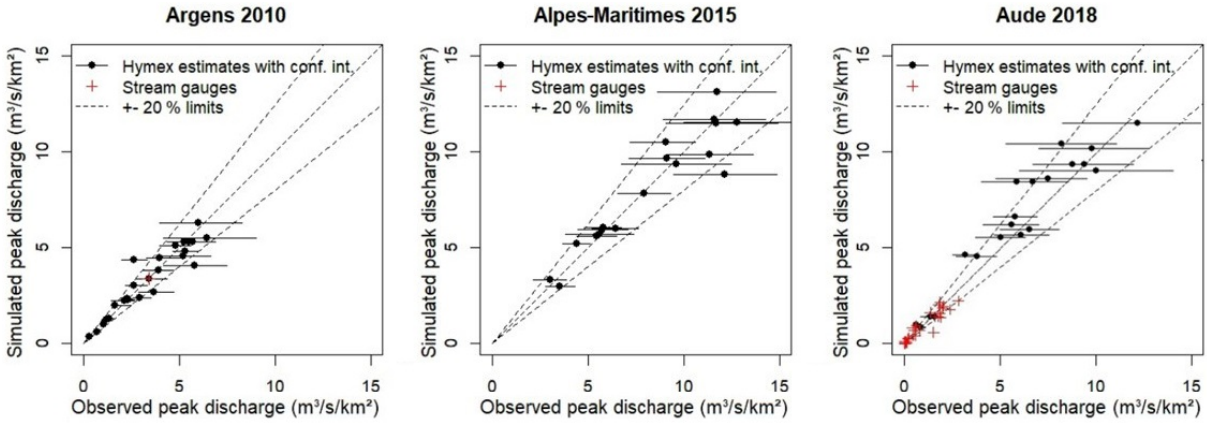


FIGURE III.7 – Observed vs. simulated peak discharges with the Cinecar rainfall–runoff model for the three simulated flood events (adapted from [Hocini et al., 2021]).

3.3 Floodos computations

Floodos is a 2D lagrangian hydrodynamic model solving the 2D SWE equations, neglecting inertia. It routes elementary water volumes (i.e. precipitons) over the topography, the number of precipitons entering the model per computation time step depending on the discharge [Davy et al., 2017]. It is coded in C++ and it is distributed in the form of a Windows executable that can be launched from three input files representing the topography (GSBG file, ".alt"), the discharges to be injected (GSBG file, ".rain") and the parameters to be applied (text file, ".dat"). In this work, the routine including the generation of these files and launching the calculation has been wrapped in R.

Hocini et al. [2021] showed that for steady state simulations, Floodos led to more accurate results for 2D modeling of flooding than a 1D hydrodynamic approach (caRtino 1D) and than a DTM filling approach (HAND/MS). Floodos also has the advantage of being relatively simple to implement, as the calculations are performed directly on the DTM mesh, and of being very fast. However, inertia terms

are neglected in the resolution, which can cause some errors in areas with abrupt changes in direction and/or in the vicinity of obstacles in the flow. Additionally, the model requires careful verification of convergence, which has been automated here. In this work, simulations are conducted independently for each river reach and are performed in a steady-state regime based on flood peak discharges (inflow discharges, the hydrological/hydraulic coupling being described in appendix 4) provided by the SHYREG database (for the flood hazard maps) or by the Cinecar hydrological model (for the historical flash flood events). The steady-state assumption can overestimate inundation extents and depths if the flood wave volume is relatively small compared to the floodplain's storage capacity. However, this assumption is considered reasonable here, as the floodplains are only a few hundred meters wide, and their storage capacities are therefore limited. Additionally, computations based on peak discharges may overestimate backwater effects at confluences due to the assumption that peak discharges occur simultaneously across all river branches at a confluence.

Details on the implementation of the Floodos numerical model over the whole French Mediterranean region is presented in appendix 4, and is derived from the previous works of Hocini [2022], Nicolle et al. [2021].

4 Illustration of the validation methods and of the results obtained with the Floodos model

4.1 Validation of flood hazard maps using hydrometric data (rating curves)

This section illustrates the use of rating curves for the verification of the flood hazard maps. In order to compare the rating curve values to the simulated water levels, the first step consists in extracting from hydraulic computation data the simulated water levels and discharge values at each gauging station location. To get the discharge values, the gauging stations were connected to the nearest SHYREG pixel. Connection based on a distance criterion (150m here) might not be the most appropriate method to determine the corresponding hydrological pixel of a gauging station - Godet et al. [2024] have documented this issue - but as the gauging stations only concern intermediate to large catchments, it was considered that the risk of error was limited, though all allocated positions were manually checked. The allocation gauging station / SHYREG pixel makes it possible to extract, for each gauging station, discharge values $Q_{shy}(T)$ corresponding to each return period T. Then, we retrieve the simulated water altitude $Z(x, y; T) = h(x, y; T) + b(x, y)$ with h the water depth and b the bathymetry elevation, at the original location of the gauging station, i.e. not at the location of the allocated SHYREG pixel. DEM_{alti} is the altitude of the pixel according to the DTM. This process results in eight triplets $\{T, W_{alti}(T), Q_{shy}(T)\}$ (one for each simulated return period) for each gauging station, which can be directly compared to the rating curve on a graph (see figure III.8).

To complete the evaluation of the effect of the DTM preprocessing (see section 3.1) on hydraulic modelling results, simulation results obtained from both raw and processed DTMs can be compared. Examples of such comparisons are provided in figure III.8. As illustrated in these examples, for a significant number of gauging stations there is good or very good agreement between the theoretical rating curves established by the hydrometric services and the simulated water levels, with differences often lower than 0.5m, even though largest errors can also be observed (e.g. the example of the station of Le Gapeau à Hyères). The results do not indicate a systematic overestimation or underestimation of the water levels, even if underestimation increasing with the return period, like in the illustrated example of La Tech at Amélie-les-Bains-Palada, is observed several times. The example of the Arc à Meyreuil or the

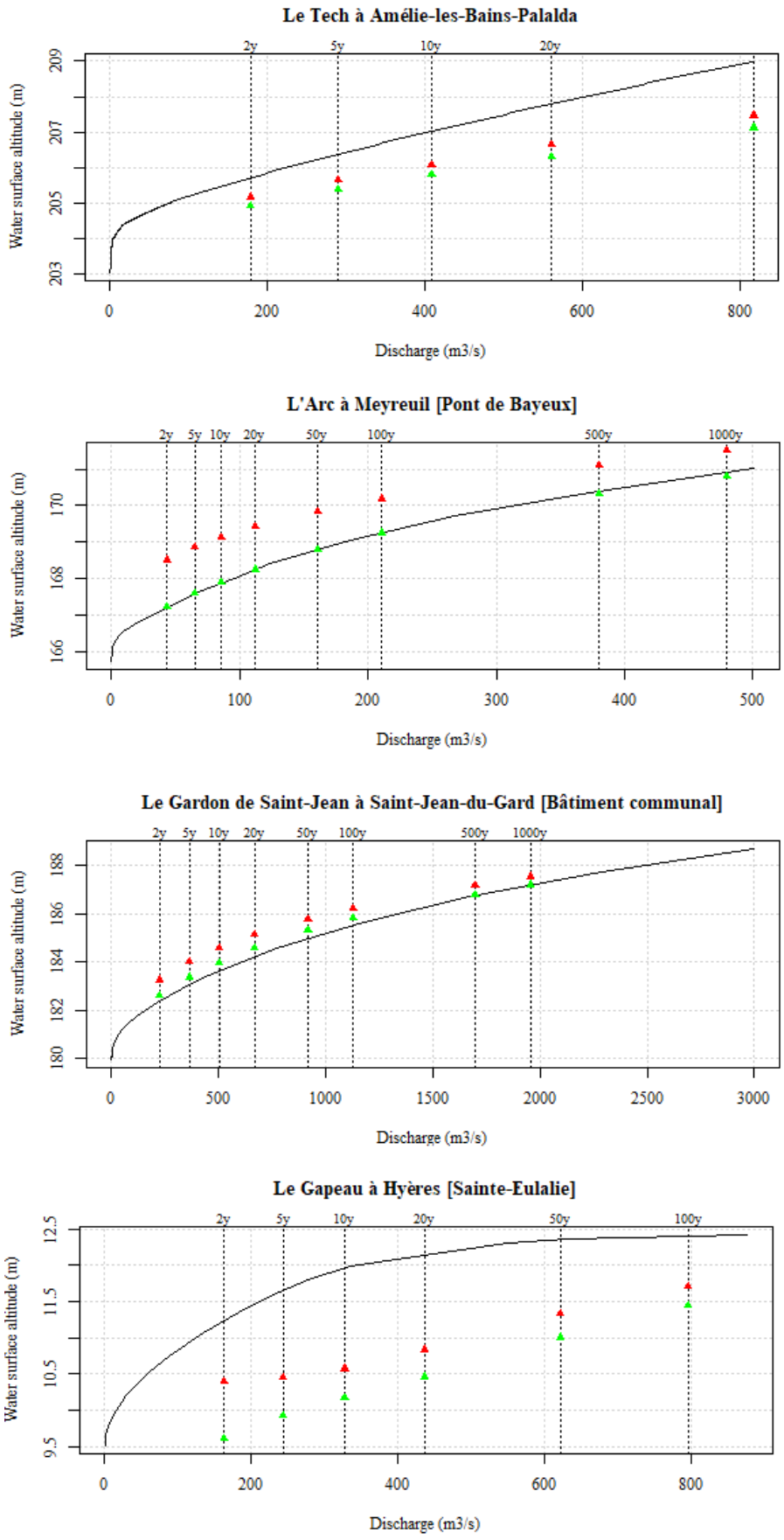


FIGURE III.8 – Examples of comparisons of theoretical rating curves elaborated by local hydrometric services (black line) and stage-discharge relations based on the regionally implemented Floodos model based on the raw (red points) or processed (green points) DTMs.

Gardon at Saint Jean illustrate the expected improvement induced by the DTM preprocessing. however such an improvement is not systematically observed, as illustrated in 2 out of 4 of the presented examples : globally, when the water levels were already underestimated, the DTM pre-processing tend to exacerbate the inaccuracies in the results.

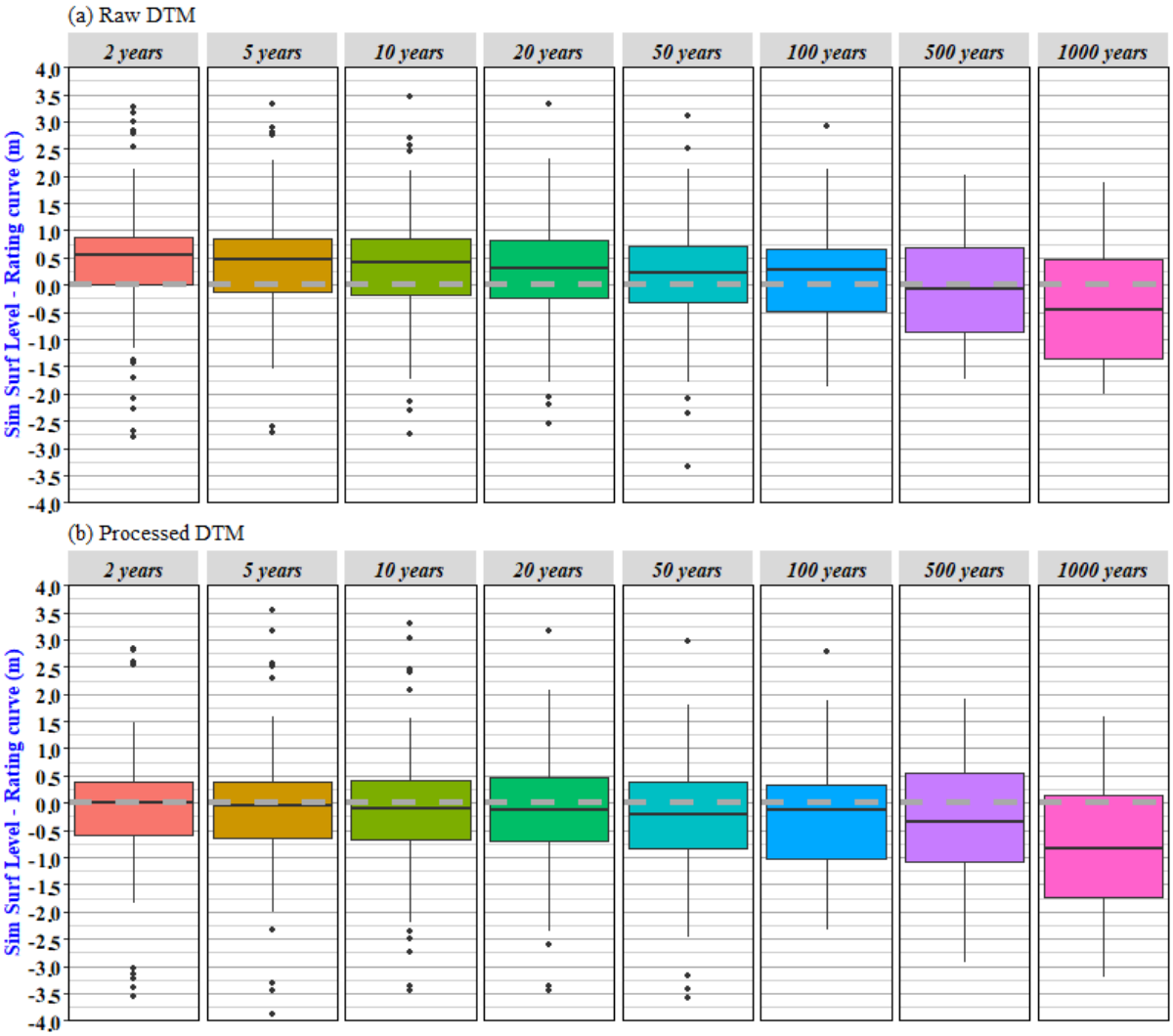


FIGURE III.9 – Boxplot of water altitude differences between the local expert rating curves and the regional 2D simulation results based on (a) the raw DTM and (b) the processed DTM for the eight considered return periods. The box plots represent, respectively, the 5 % and 95 % (whiskers) and the 15 % and 85 % quantiles (boxes).

To provide a more comprehensive evaluation, boxplots of differences in water altitude are presented for each return period in Figure III.9. This confirms that the general effect of the DTM preprocessing is a decrease in the simulated water altitudes. This was expected since the DTM preprocessing procedure described in section 3.1, leads to increase the hydraulic capacity of the streambeds. This impact is noticeable for all return periods and leads to an overall improvement of the agreement between the expert-based rating curves and the simulated water levels for low to medium return periods : the median error for the T=2years is reduced from 0.54m (raw DTM) to -0.04m (processed DTM). The median error (sim. - obs.) for all return periods combined is of 0.38m for the water levels simulated using raw DTMs, and of -0.14m for the water levels simulated using processed DTMs. These results are encouraging given that they are of the same order of magnitude than typical errors observed for the validation of flood events mapping using high water marks (see Hocini et al. [2021] and section 4.2), despite the fact that this verification method is very challenging. It is important to recall that the simulations are based on a

numerical 2D model, implemented at a large regional scale with fixed friction parameters values and no further local adjustment.

The results show that the primary objective of the DTM preprocessing, i.e. reducing excessive overflow for low return periods, is reached. However, they also show that simulated water levels are less consistent with local theoretical rating curves for large return periods, i.e. 500-year and 1000-year. However, the boxplots for the 500-year and 1000-year return periods are only based on 33 and 20 values respectively (table III.1), as few local forecasting services venture into this type of highly uncertain extrapolation of rating curves. As indicated in section 2.2.2, only 2% of directly measured (gauged) discharges exceed the 2-year return period, which does not allow for reliable extrapolations. Even when the rating curves suggest values for the 500-year and 1000-year return periods, this is systematically far outside of the reliability range defined by the local forecasting service. Finally, the spatial distribution of average errors has been explored. This does not reveal any clear spatial trend or concentration of high errors values. This analysis did reveal that large differences (i.e. exceeding 2m) affect generally both the results obtained with the processed and raw DTMs. This indicates that the processing of the DTMs could not significantly reduce the distance between the locally established rating curve and the simulation results for the considered gauging stations, meaning that the source of the difference is probably not related to the limitations of the DTM. As highlighted in Section 3.3, other sources of error may stem from neglecting inertia terms in the resolution and assuming a steady-state regime.

4.2 Validation of flood event inundation maps using observed flood extents and high water marks

Readers are encouraged to refer to Hocini et al. [2021] for a more detailed explanation of the validation method used here, which consisted in comparing the inundation simulation results obtained for three specific events with observed flood extents and high water marks. In this section, we replicate this evaluation for our simulation results, obtained again with the Floodos hydraulic approach but using both an updated version of the raw DTM, and the processed version of this DTM. The results, shown in figure III.10.b for the case of the Argens 2010 flood, align with those in figure III.9, since they indicate a tendency to underestimate water levels during the most intense events (the three considered floods are major ones). This underestimation may be attributed to the uncalibrated Manning friction coefficients used or to the lack of representation of backwater effects from structures, among other factors. Although the influence of solid transport related to erosion and deposition processes was limited during these three events — unlike a more recent event, the Alex storm in October 2020, which severely affected the study area (see Payrastre et al. [2022]) — localized blockages were observed during the Aude 2018 event, as documented by Hocini et al. [2021]. These blockages may have contributed to significant underestimations of water levels and flood extent in the vicinity of the blocked bridges.

Despite these limitations, the results showed a median (mean) altitude difference of -0.47 m (-0.57 m) for the Argens 2010 event, -0.30 m (-0.45 m) for the Aude 2018 event, and -0.31 m (-0.20 m) for the Alpes Maritimes 2015 event. These values fall within the typical error range of ± 0.3–0.5 m for validation based on this type of data, as noted by Bates [2022]. The simulated inundation maps were also in good agreement with the observed flood areas (figures III.10.a and III.10.c), with CSI scores (measuring the overlap between observed and simulated flood extents) ranging from 55% to 85% for 90% of the river reaches studied. The median CSI values, close to 80%, are particularly satisfactory, considering that scores above 65% are generally regarded as indicators of reliable local estimates [Fleischmann et al., 2019].

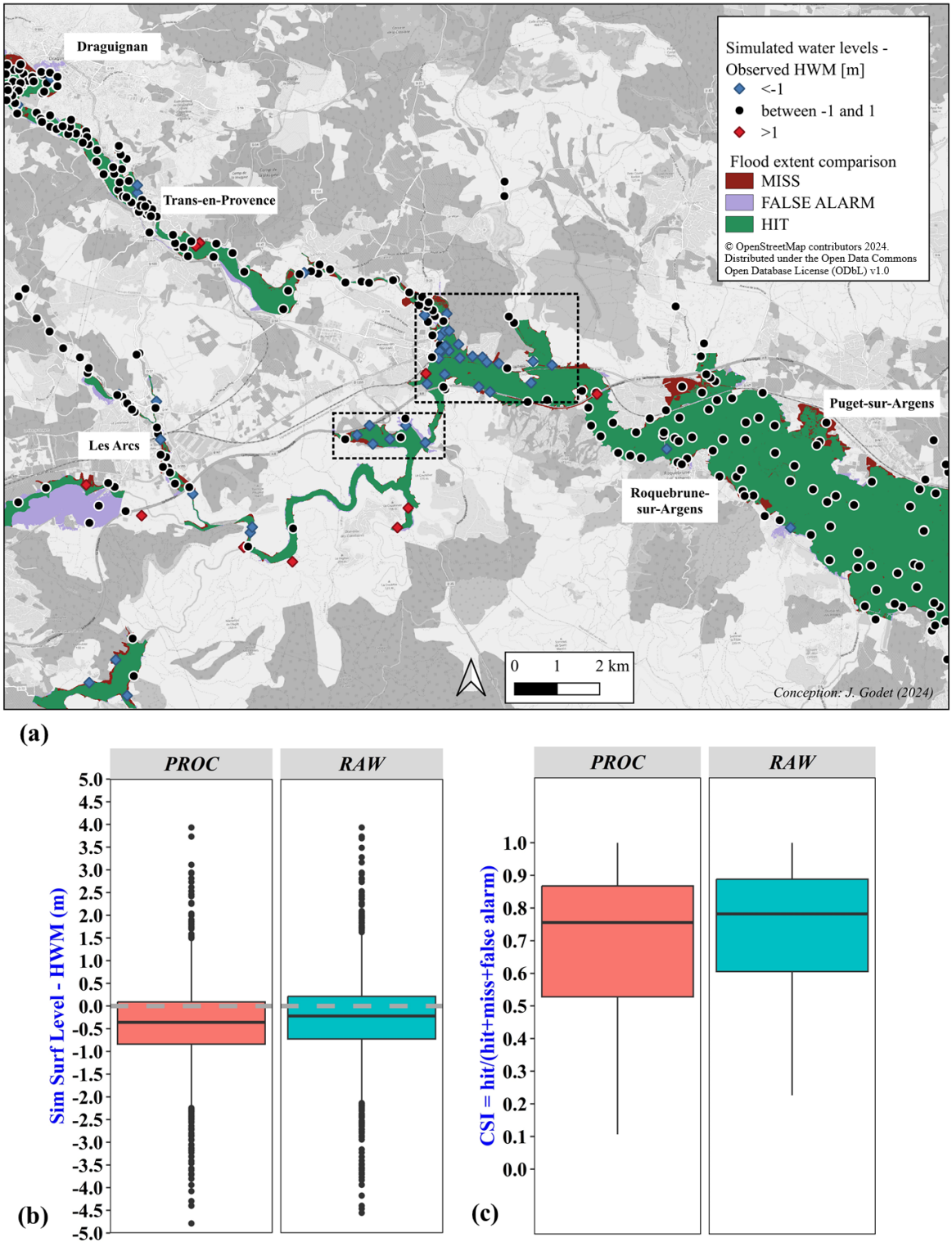


FIGURE III.10 – (a) Comparison of simulated and observed flood areas and water levels for the Argens 2010 event and the processed DTM. The dotted black rectangles refer to error clusters. (b) Comparison of simulated water levels based on the raw and processed DTMs, with observed high water marks (HWM) for the three events. (c) Comparison of simulated and observed flood extents for the Argens 2010 and the Aude 2018 events, using the CSI metric.

5 Discussions

The flood hazard maps, obtained from automated computations implemented on large areas - which inevitably implies errors and uncertainties - resulted in very promising results when compared to the 171 available locally adjusted rating curves. The local comparison of simulated inundation maps with observed flood extents for three specific intense flood events also led to satifactory results, as already demonstrated by Hocini et al. [2021]. However, these results also showed locally large differences between simulated and reference water altitudes (see table III.2 for some examples in the case of flood hazard maps). This reveals some remaining limits which must be underlined.

A first explanation can be found in the DTM source. Notably, the entire study area was not covered by a Lidar DTM at the time of the calculations (Figure III.11). Less accurate DTM production techniques, such as satellite radar and aerial photogrammetry, still predominate in some areas of the region under consideration. Table III.2 indicates that among the 5% of stations with the largest errors, 6 out of 9 were located in areas without Lidar coverage. Figure III.11 shows clear examples where the streambed is poorly defined or even non-existent in the original DTM. The pre-processing technique described in appendix 3 allows the definition of a streambed when it is absent or poorly defined in the raw DTM. However it is not always able to compensate for all the limitations of an initially poorly defined DTM.

Code	Name	**Mean abs. diff. (m)**	**% Lidar**	**% Photo**	**% Radar**
Y004501001	Le Riu de Quérol à Porta	16.04	0	100	0
Y047406001	La Têt à Saint-Féliu-d'Amont	5.28	100	0	0
V603501001	Le Toulourenc à Malaucène	5.14	65	0	35
Y136401001	Le Fresquel à Carcassonne	5.10	100	0	0
Y663501001	La Bévéra à Sospel	5.05	0	0	100
Y040401001	La Têt à Mont-Louis	4.93	0	100	0
V505401001	L'Ardèche à Vallon-Pont-d'Arc	3.95	0	100	0
Y201002001	L'Arre au Vigan	3.71	0	100	0
Y622401001	La Tinée à Saint-Sauveur-sur-Tinée	3.43	0	0	100

TABLEAU III.2 – The 5% of stations where the flood hazard maps showed the largest errors (calculated on the 2-y, 5-y, 10-y, 20-y return periods, with the results obtained from processed DTMs) and the source of the DTM slab they are located on.

As an illustration, the mean absolute errors calculated for the 2-year, 5-year, 10-year, and 20-year flood hazard maps across all stations is 1.03 meters, doubling to 1.95 meters for stations located in areas without Lidar coverage. This clearly demonstrates the advantages of Lidar technology. But even when Lidar measurements are available, achieving a perfect description of the geometry of streambeds on a large scale remains challenging, particularly due to the lack of bathymetric data, as the commonly used lasers do not penetrate the water column. Hocini et al. [2021] have also discussed main error sources affecting their results on the simulation of historical events. They have observed that large errors are spatially clustered (see the rectangles in figure III.10) : they had already pointed out the limitations of the DTMs and particularly the lack of river bathymetry. The resulting underestimation of the streambed's cross-sectional area and hydraulic capacity will then depend on the season of survey and on the morphology of the river beds. Larger underestimations are more likely for winter surveys than for summer surveys. Similarly, ponds in the riverbed will be more frequent in the downstream sections of watercourses, often equipped with engineered weirs or exhibiting a succession of pools and riffles. Ideally, Lidar surveys should be conducted during the Summer when water levels are low. However, the dense vegetation in Summer, which may partially or completely cover the streambed, especially in small rivers, can also pose problems for Lidar measurements.

Remaining errors can arise from a variety of other sources, including particularly geographical or

altimetric referencing errors of sensors and/or HWM. In the case of station Y004501001 on the Riu de Quérol, which shows the largest mean difference between the simulated water altitudes and the rating curve amongst all stations under study (average difference of -16m for the results obtained with the processed DTM), the exact location of the water level sensor can be questioned. The official altitude of the zero of the water level gauge is 1372m NGF on Hydroportail. According to the Hydroportail database, the streamgauge is located under a bridge which deck altitude is 1363m NGF. This clearly is inconsistent. The local slope of the stream is about 5%. The altitude of the zero level of the gauge, if reliable, thus suggests that the sensor should be located several hundred meters upstream this bridge. Exchanges with the local hydrometric service could not confirm either of the two hypotheses : error in the reference altitude or location of the sensor. This example shows that despite a careful problem tracking with the help of local hydrometric services, some errors may remain.

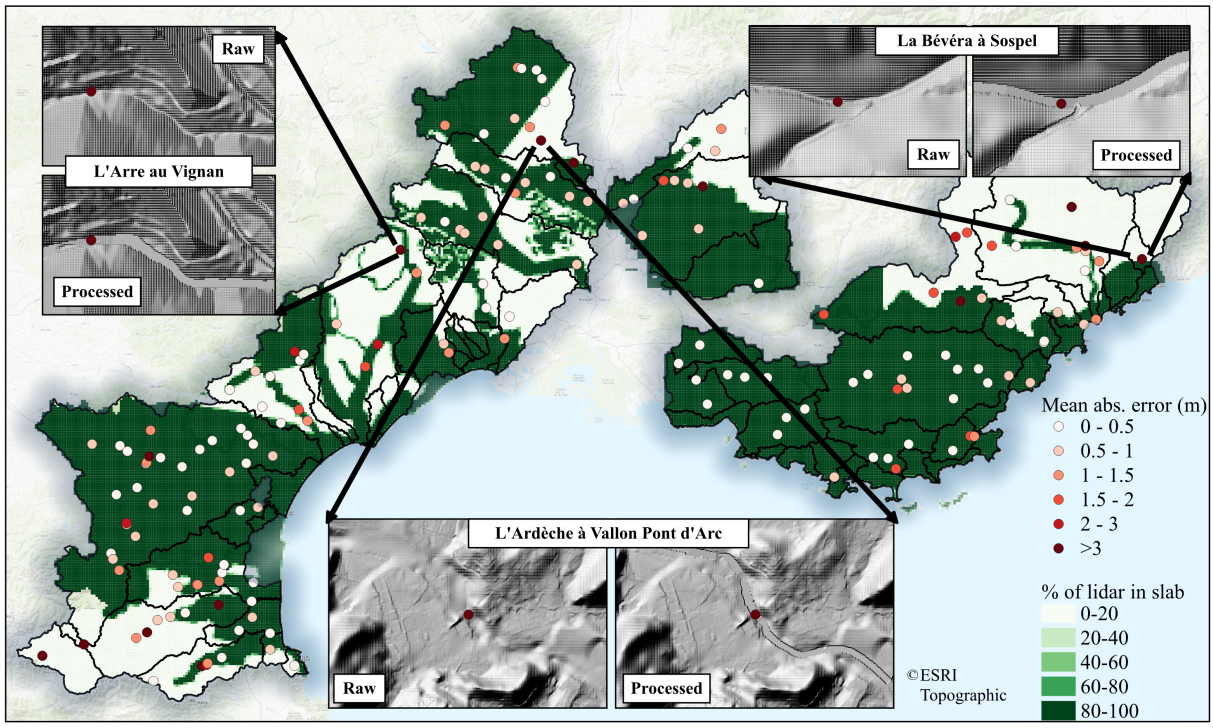


FIGURE III.11 – Portion of the area monitored by the radar technology, combined with the mean absolute errors between the simulated water altitudes and the rating curves (calculated on the 2-y, 5-y, 10-y, 20-y return periods), and examples where the streambeds where very poorly detected

It is worth noting that Figure III.9 also shows that the chosen streambed/ flood plain couple of Manning's friction coefficients (1/18, 1/10 $s.m^{-1/3}$) is adapted to low return periods, but that it leads to underestimate on average the water levels for larger return periods. If we trust the locally established rating curves, this suggests that for extreme flood events, the apparent friction of Mediterranean streambeds is certainly even greater than that adjusted in this application (rather 1/15 than 1/18 $s.m^{-1/3}$). This observation is consistent with the conclusions drawn by Lumbroso and Gaume [2012], who concluded, based on post-event survey data, that Manning roughness coefficients should be much higher than the commonly recommended values (ranging from 1/30 to 1/25 $s.m^{-1/3}$), to provide realistic stage-discharge relations for extreme flash-flood events. Large regional implementations of hydraulic models may also provide or confirm precious information about hydraulic properties of river beds.

6 Conclusions

This article documents a dataset including input and validation data, for the computation of flood hazard maps and inundation maps of historical floods over the French Mediterranean area. Reference computation results are also provided, usingthe Floodos 2D hydraulic model, which solves the shallow water equations without inertia and in steady state regime. The article illustrates both the variety and richness of the proposed evaluation dataset in a region frequently affected by severe flash floods, and the relevance of approaches based on Lidar DTM and 2D numerical hydraulic models for high resolution regional flood mapping.

It was diagnosed that some major errors were linked to the DTM source, nevertheless other errors are probably introduced because of the simplicity of the chosen modelling approaches : simplified reservoir-based hydrological models, hydrodynamic method neglecting inertia terms and carried out independently for each reach, calculations in steady-state regime, monofrequency catalog of scenarios, etc. The proposed dataset provides the opportunity to address these issues and to assess the improvements of more complex approaches in the future. However, the results presented here, show that simple 2D models like Floodos, implemented on a regional scale, already deliver very satisfactory outcomes. These models are already effective for simulating historical flood extents and computing flood levels and hazard maps for reference discharge values. The challenging comparison with locally calibrated rating curves is particularly interesting, with an absolute mean difference of less than 0.5m between simulated levels and the reference rating curve in more than half of the cases, even for hazards of limited intensity (low return periods). This confirms the promising results already obtained by Hocini et al. [2021] based on the comparison with HWM levels for the three historical flash flood events.

Chapitre IV

Implémentation de la chaîne de prévision immédiate basée sur les impacts

Objectifs et organisation du chapitre :

Ce chapitre s'intéresse aux aspects techniques de la construction d'une chaîne d'avertissement de type "Impact-based Warnings" (IBW). Il traite en particulier des modalités d'association entre les variables d'entrée et de sortie des différents modèles mobilisés, ainsi que de la combinaison entre des calculs hydrologiques effectués en temps réel et des calculs hydrauliques préétablis sous forme de catalogues de scénarios d'inondation.

La première problématique abordée concerne la connexion entre un modèle hydrologique reposant sur des mailles de résolution kilométrique et des modélisations hydrauliques, dans lesquelles les débits sont injectés en des points situés le long du réseau hydrographique. Ce travail a fait l'objet d'une publication dans la revue *Hydrology and Earth System Science* [Godet et al., 2024], dont le texte intégral, rédigé en anglais, est reproduit dans la première section du chapitre, après un court paragraphe de synthèse.

Dans un second temps, le chapitre traite des modalités de reconstitution d'un scénario d'inondation, puis du scénario d'impacts associé, à l'échelle du tronçon de cours d'eau, au sein d'une chaîne d'avertissement IBW applicable en temps réel à l'ensemble du réseau hydrographique. Cette seconde partie permet de présenter en détail le fonctionnement et les choix opérés pour la chaîne d'avertissement IBW utilisée dans la suite de la thèse.

1 Connecter l'hydrologie et l'hydraulique : associer des points-rivière aux mailles d'une grille de modélisation hydrologique

1.0 Synthèse en français

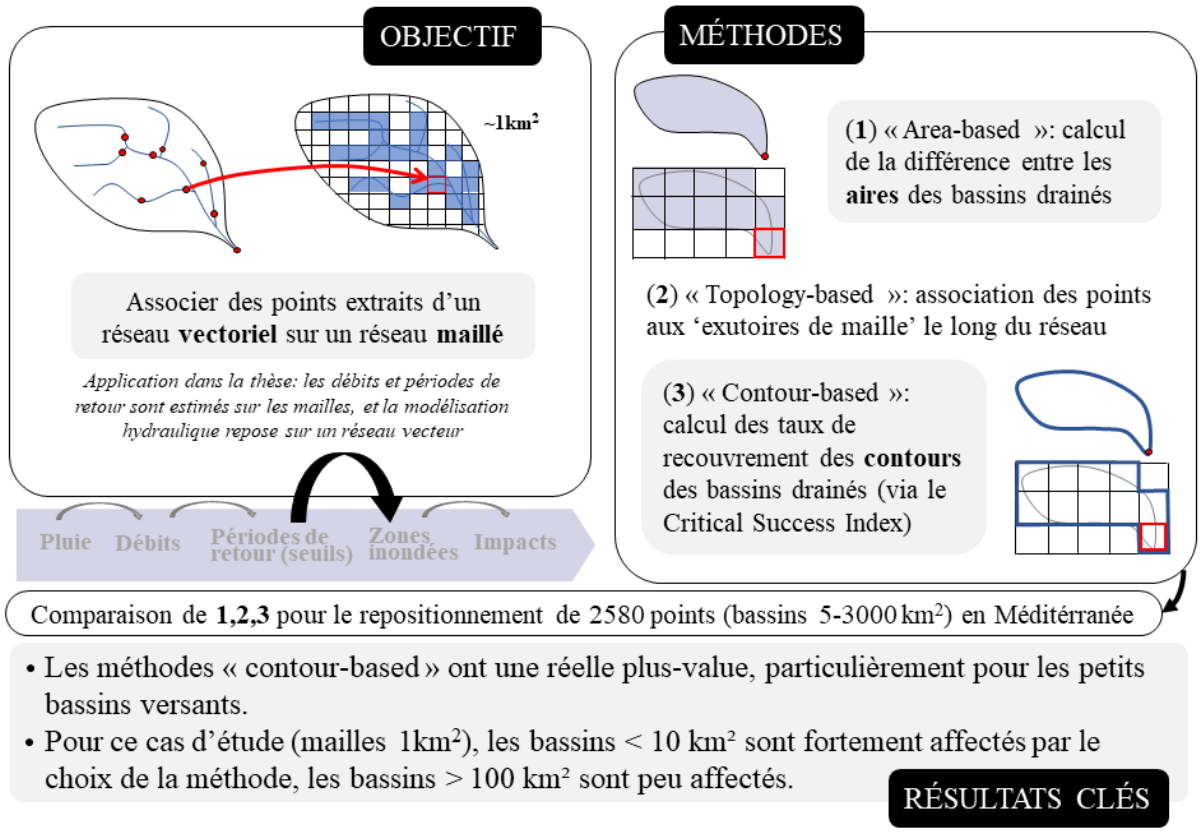


FIGURE IV.1 – Résumé graphique de l'article [Godet et al., 2024].

Cet article s'intéresse à la connexion entre le bloc "hydrologie" et le bloc "hydraulique" d'une chaîne de modélisation de type IBW, en partant du postulat que la résolution des modèles hydrologiques (environ kilométrique) est plus grossière que la résolution des modélisations hydrauliques (de l'ordre de quelques mètres) utilisées pour simuler les inondations. L'étude de cas proposée correspond au contexte de la prévision des crues en France. Le réseau hydrographique utilisé est celui de la BNBV (représentation vectorielle approchée du réseau hydrographique français dès 5km² de surface drainée, issue d'un MNT de résolution 50m). Les points exutoires de ce réseau, pour lesquels la base SHYREG fournit des quantiles de débits associés à des périodes de retour de 2 à 1000 ans, constituent des points d'entrée pertinents pour les modélisations hydrauliques des zones inondées. En revanche, le modèle hydrologique utilisé dans Vigicrues Flash fonctionne sur une grille de résolution kilométrique, et ne peut donc pas directement fournir d'hydrogrammes à l'échelle fine des exutoires BNBV.

Cette problématique d'association se pose également en dehors du cadre des chaînes de prévision IBW : pour caler ou valider un modèle hydrologique maillé, il est inévitable de pouvoir définir quel pixel, ou quelle maille, représente une station hydrométrique. Pour cela, les approches traditionnelles se concentrent généralement sur la simple comparaison des valeurs d'aires drainées amont (différence entre l'aire de référence du bassin versant associé à la station, et l'aire drainée amont du pixel) et/ou sur la distance (entre la station et le pixel). Cependant, ces approches sont de plus en plus critiquées par les hydrologues, qui soulignent les problèmes pouvant se produire au niveau des confluences, lorsque deux petits affluents peuvent avoir des valeurs d'aires drainées amont similaires. De nouvelles approches ont

donc été développées, notamment des méthodes qui visent à comparer les contours des bassins versants plutôt que seulement les aires drainées. Cependant, la plus-value de ces nouvelles approches n'a jamais été réellement évaluée.

Cet article apporte des premiers éléments de réponse en comparant différentes méthodes pour l'association de "points-rivière" (pouvant être des stations hydrométriques) avec des pixels d'une grille hydrologique à la résolution d'1 km^2, sur la région Méditerranée-Est.

Trois approches représentant diverses familles de méthodes ont été appliquées et comparées : des méthodes basées sur la superficie ("area-based"), sur la topologie ("topology-based") et sur le contour des bassins versants ("contour-based"). Ces méthodes sont mises en œuvre pour associer 2580 points-rivières à une grille de modélisation hydrologique de résolution kilométrique. Ces points sont répartis le long de l'ensemble du réseau hydrographique de la région méditerranéenne sud-est de la France, couvrant des bassins versants de taille variant entre 5 et 3000 km^2. La comparaison des trois approches se base sur un score CSI (Critical Success Index), qui quantifie le recouvrement du bassin choisi par la méthode d'association au bassin de référence.

Sur une grille de modélisation hydrologique de résolution 1 km^2, utilisée dans ce travail, les bassins inférieurs à 10 km^2 sont inévitablement représentés de manière imprécise : leurs scores médians de précision (critère de recouvrement CSI) atteignent à peine 60 % avec la meilleure méthode, contre plus de 90 % pour des bassins supérieurs à 100 km^2, quelle que soit l'approche. En revanche, la dispersion des scores CSI s'avère beaucoup plus faible avec la méthode basée sur les contours (premier et dernier quartiles : 0.55 et 0.65) qu'avec celle basée sur les aires (premier et dernier quartiles : 0.05 et 0.65). En résumé :

- Les méthodes "contour-based" surpassent généralement les méthodes "area-based" : C'est particulièrement vrai pour les petits bassins versants, pour lesquels ces méthodes réduisent significativement les erreurs.
- Pour les faibles surfaces drainées, les méthodes "area-based" peuvent associer des bassins qui ne se correspondent en réalité pas du tout, ce qui peut se produire lorsque deux branches différentes du réseau hydrographique sont proches et ont des aires drainées similaires.
- Les différences de performance diminuent à mesure que la taille du bassin versant augmente : Cela explique probablement pourquoi les méthodes basées sur l'aire drainée ont été si largement utilisées dans les études précédentes, où l'accent a souvent été mis sur les grands bassins versants.

1.1 Introduction

In hydrology, rainfall-runoff models' outputs are often compared to observed discharge series at gauging stations for calibration or evaluation purposes. Vector-based hydrological models are adequate to meet these objectives, because it is straightforward to locate a gauging station along the river network. However, when using gridded models, it is necessary to allocate each gauging station to a specific cell in the model grid. In the literature, terms such as "co-registering" [Fekete et al., 2002], "co-referencing" [Döll and Lehner, 2002], and "matching" [Wang et al., 2018] are also used to describe this process. The allocation of specific river points to a coarse-resolution cell can also be necessary when connecting the output of a hydrological distributed model (providing hydrographs or peak discharges on a grid) to a hydraulic model for inundation modelling. As an example, Dottori et al. [2017] developed a European-wide flood risk assessment system, based on the European Flood Awareness System (EFAS, see Thielen et al. 2009, Bartholmes et al. 2009). The discharge output of EFAS is provided on a grid of spatial resolution of $5km$, which needed to be downscaled to a resolution of $100m$ in order to derive flood hazard maps at the pan-European scale. Dottori et al. [2015] opted for a basic method consisting in allocating the $100m$ river pixel to the $5km$ river cell containing this pixel. This approach has limitations, particularly when the two river networks defined at the $100m$ and $5km$ scales do not overlap.

Generally, the allocation of river points to a coarse-resolution grid for hydrological modelling relies on distance and upstream drainage area (UPA) error criteria [Döll and Lehner, 2002, Fekete et al., 2002, Lehner, 2012, Zhao et al., 2017, Sutanudjaja et al., 2018, Wang et al., 2018, Burek et al., 2020, Polcher et al., 2023]. However, this process is also prone to errors, especially near confluences, where points of different branches of the river network may have similar UPAs. As a result, their allocation on the coarse-resolution grid can lead to assigning a point to the wrong hydrological grid cell and corresponding upstream watershed, based on a slightly better UPA fit or a slightly shorter distance (see figure IV.2 for an example).

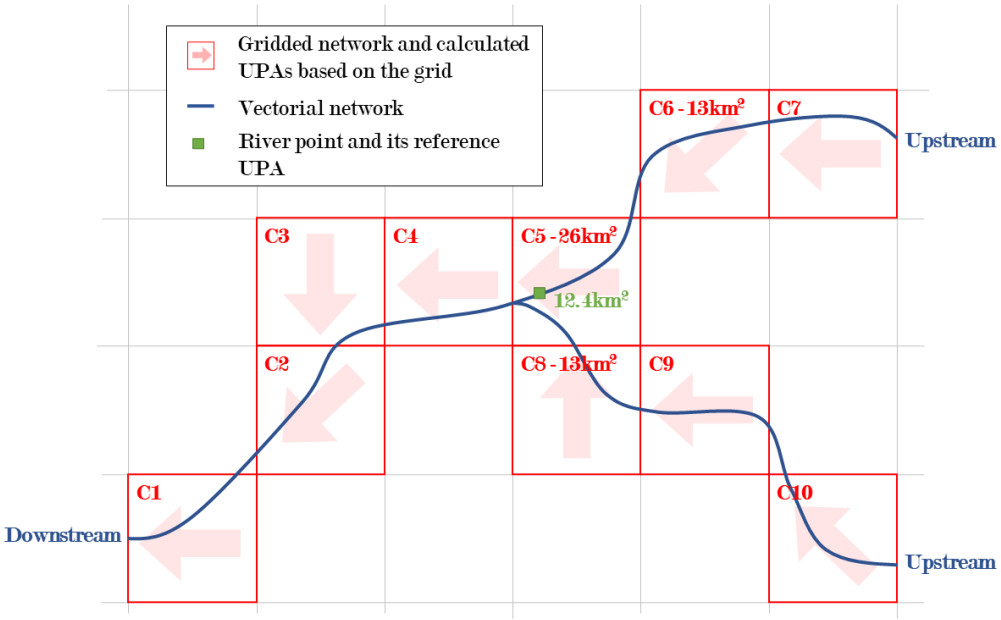


FIGURE IV.2 – Example of allocation process failure when based on distance and UPA error criteria : green river point is allocated to grid cell C8 instead of grid cell C6.

Considering this possible limitation, efforts have been made to propose more effective protocols for allocating river points to hydrological grid cells. For instance, the methods proposed by Burek and Smilovic [2023], Munier and Decharme [2022] combine distance and UPA error criteria with a

comparison between the gauging station's basin boundaries, delineated on the basis of a fine-resolution Digital Elevation Model (DEM), and the basin boundaries of the allocated cell based on the coarse-resolution hydrological grid. In both studies, the similarity between the watershed limits is characterized by the Intersection Over Union index [Rezatofighi et al., 2019].

The idea of comparing basin boundaries had been previously considered, although in a slightly different context, namely the evaluation of hydrological grids obtained from an upscaling algorithm (i.e., transforming a fine-resolution grid into a coarser-resolution grid). Initially, upscaling algorithms were also guided by a comparison between the UPA values of "small pixels" and the corresponding upscaled "large cells" [Reed, 2003, Paz et al., 2006, Yamazaki et al., 2008, Eilander et al., 2021]. Visual inspections were performed to identify obvious inconsistencies between coarse and high-resolution river networks. To reduce these inconsistencies automatically, additional criteria have been proposed to complement the UPA criterion, such as the mean distance between river networks and the percentage within a buffer [Davies and Bell, 2009], the correctness index and the figure of merit [Li and Wong, 2010], and the watershed delineation percentages of consistency [Sousa and Paz, 2017].

In summary, numerous methods are available to achieve the objective of allocating a river point to a coarse-resolution grid cell. However, these methods have been developed in different contexts and have rarely been compared. This study aims at comparing the results obtained from three different types of methods for allocating a large number (2580) of river points to a coarse-resolution hydrological grid ($1km \times 1km$). The first method belongs to the category of area-based methods and employs distance and UPA error criteria. The third method is a contour-based approach. The second method is a topological method based on proximity along the river network. Another unique aspect of this work is that it deals with a detailed river network that includes river points with small drainage areas (minimum of $5km^2$), whereas most previous studies have been limited to the main river networks (catchments larger than $500km^2$ in Dottori et al. [2017] for instance). In this study, $1km \times 1km$ is considered as "coarse" resolution because the hydrological model is intended for the regional scale. However, the same problem could arise for hydrological modelling applied on a continental scale where the resolution will be coarser than 1km (i.e. 5 to 10 km).

The paper is structured as follows : Section 2 presents the three tested allocation methods as well as the validation metrics. Section 3 provides an overview of the case-study. Section 4 compares the results obtained with the three tested allocation methods, and Section 5 highlights the limitations of the study and summarises the key messages.

1.2 Allocation methods and validation metrics

In this section, we describe three methods that allow for the allocation of a river point to a coarse-resolution grid cell. The validation metrics used to compare the methods are described at the end of the section.

1.2.1 Method 1 : area-based method

Area-based methods can be traced back to Döll and Lehner [2002], who proposed allocating river points to the coarse-resolution grid cells containing the points, provided that the relative difference between coarse and reference resolution UPAs did not exceed 5%. This criterion led, in their case, to a manual re-allocation of 35% of the points. In order to automate the allocation procedure, Lehner [2012] proposed to select the grid cell within a 5km radius search area around each river point (see Figure IV.3) with the lowest value of a discrepancy criterion $D = RA + 2R$, where RA stands for the relative difference

between UPAs that should not exceed 50% and R for the distance between the point and the centre of the grid cell. In most other works [Zhao et al., 2017, Sutanudjaja et al., 2018, Wang et al., 2018, Burek et al., 2020], RA is the only selection criterion, the radius of the search area (1-25km, depending on the spatial resolution) and the maximum acceptable RA value for a successful allocation (10-30%) varying between studies. To maximise the proportion of allocated river points and to optimise the computation times, the approach proposed herein proceeds in three possible successive steps. At step 1, the closest grid cell verifying $RA < 10\%$ and $R < 3$ cells is selected, if it exists. If it does not, the maximum RA values is increased to 20% at step 2 and 30% at step 3. It can be noted that the proposed approach combines area and distance criteria. Also, if a maximum difference between UPAs of 30% is a recurrent choice in the literature, e.g [Burek et al., 2020], regardless of the studied model resolutions, the distance criterium $R < 3$ is more study-dependant. In the present study, it appeared after some tests as a good compromise providing accurate results. However when using global-scale hydrological models and coarser grids, the value of R may have to be adjusted.

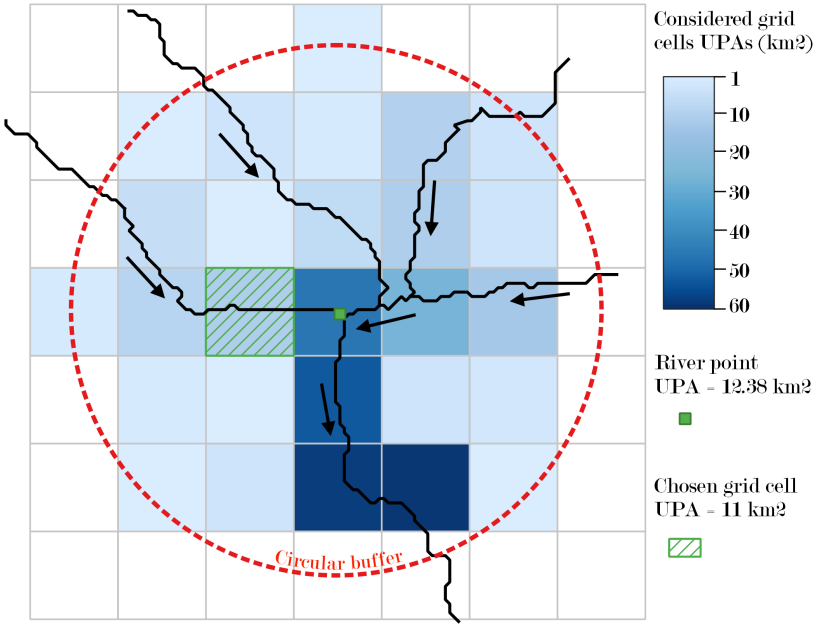


FIGURE IV.3 – Illustration of Method 1 : grid cells candidates for a specific river point. In this situation, the river point is allocated to the green-hatched grid cell.

1.2.2 Method 2 : topology-based method

This second method requires a vector-based river network and the definition of coarse grid cells' *outlet points*. The cells' outlet points are located and selected according to the IHU upscaling method [Eilander et al., 2021], used to generate the coarse-resolution hydrological modelling grid (see section 1.3.2). Each river point can then be connected to the closest upstream or downstream grid cell outlet point and hence allocated to the corresponding grid cell, provided that both points belong to the same river reach, i.e. are not separated by a network confluence (case of point P3 in Figure IV.4). With this method, river points located between two confluences within the same grid cell cannot be allocated.

1.2.3 Method 3 : contour-based method

Several previous works have stressed the importance of considering the consistency of watershed contours for the evaluation or optimisation of upscaling or allocation methods [Davies and Bell, 2009,

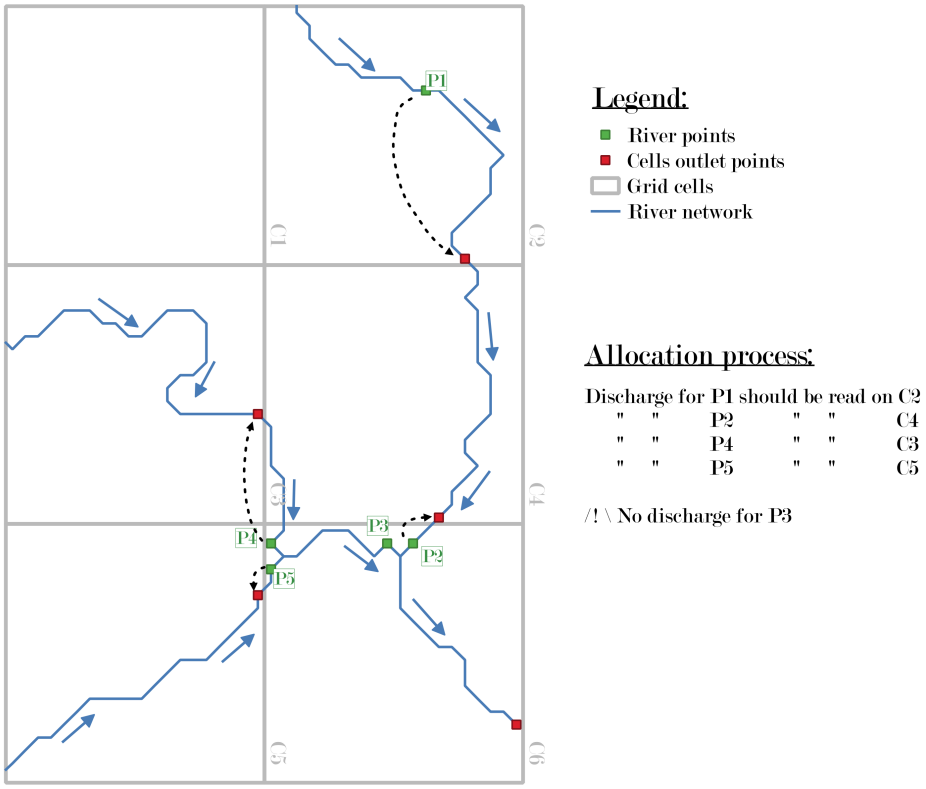


FIGURE IV.4 – Illustration and limits of Method 2 : connection of river points and cells outlet points (black dotted arrows) with an impossibility for point P3, located between two confluences in the same grid cell.

Li and Wong, 2010, Sousa and Paz, 2017]. Method 3 is similar to method 1, except that it is not only based on the comparison between upstream watershed surfaces, but also between upstream watershed contours. While Munier and Decharme [2022], Burek and Smilovic [2023] did propose an allocation criterion based on a combination of an area-based and a contour-based criterion, it is proposed herein to base the selection of the appropriate grid cell for each river point based on a single criterion, namely the critical success index, CSI (see eq. IV.1). The CSI is a standard score to compare surfaces, often used to compare flood inundation models for instance [Fleischmann et al., 2019, Hocini et al., 2021].

$$CSI = \frac{a}{a+b+c} \tag{IV.1}$$

Where a (HIT, see figure IV.5) is the overlapping area between the reference upstream watershed of the considered river point and the upstream watershed of the candidate grid cell, defined on the coarse-resolution grid using the TAUDEM library [Tarboton, 1997]. "b" is the area of the reference watershed not overlapping with the coarse grid watershed (MISS) and conversely, "c" is the area of the coarse grid watershed, not overlapping with the reference watershed (FALSE ALARMS). CSI is equal to 1 in case of a perfect overlap and 0 when there is no overlap at all. It can be noted that the CSI has often been used with alternative denominations in previous studies, such as the Intersection over Union criterion [Munier and Decharme, 2022, Burek and Smilovic, 2023], the Figure of Merit [Li and Wong, 2010], or Fit Metric [Fleischmann et al., 2019].

The grid cell maximising the CSI value is selected among the 9 cells closest to the considered river point. If a minimum CSI value is not reached (i.e. 0.4 for watershed areas under 10 km^2 and 0.6 otherwise), the search area is extended to the $7 \times 7 = 49$ closest grid cells. Tests carried out prior to this choice showed that extending the research area to the $5 \times 5 = 25$ closest grid cells did not allow most of the targeted cases

to be corrected. The CSI thresholds can be adjusted (see section 1.4.4), as well as the surface threshold and the extended research area, since they may depend on the model resolution.

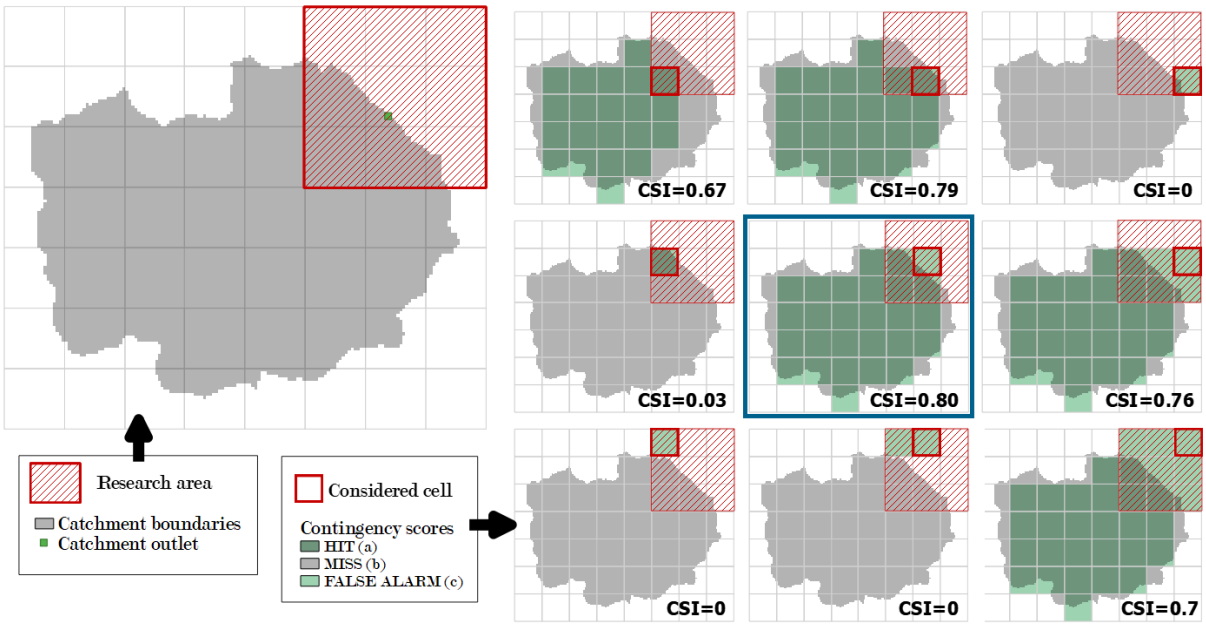


FIGURE IV.5 – Illustration of Method 3 : CSI calculations for the nine candidate cells. In this situation, the river point is allocated to the candidate cell of the blue box scenario, which maximises the CSI (CSI = 0.80).

1.2.4 Evaluation metrics

The allocation procedures aims at relating points of the river network, often corresponding to a gauging station, to coarse grid cells of a distributed hydrological model, with the upstream watershed closest to the actual watershed, or at least closest to the watershed delineated based on the finest available geographical data. Therefore, the CSI appears as the best suited score for the efficiency evaluation of allocation methods and will be used hereafter. The CSI is calculated based on the low resolution catchment contours (reference contours). By construction, method 3 should then lead to the best performances, but at the cost of a higher implementation complexity and significantly larger computation times, as will be illustrated herein. Hence, the main question will be : "how well do methods 1 and 2 compared to method 3 ?"

1.3 Case study and data

The test area covers three departments in the Eastern Mediterranean region of France, with a total surface area exceeding 15,000 km^2 (figure IV.6). Two geographical datasets have been used to implement and compare the allocation methods : shapefiles of the river points to be allocated and their reference catchments boundaries, and a raster describing the coarse-resolution grid specifically designed for the purpuse of the study.

1.3.1 The river points to be allocated

The BNBV (Base Nationale des Bassins Versants) is a French reference GIS layer describing the network of rivers with an upstream catchment area larger than $5km^2$, over the whole territory of France. It was produced by Organde et al. [2013], based on the processing of a hydrologically-validated 50m-resolution

flow direction grid. It includes a vector description of the river reaches and identifies approximately 15,000 points of interest along the river network across mainland France. These points correspond to locations of gauging stations, urban areas, confluences, river mouths. Intermediate outlets were also added to ensure comprehensive coverage of the whole river network. The upstream watershed limits and areas are associated to each BNBV point.

Figure IV.6 displays the BNBV outlets and river reaches in the Eastern Mediterranean region. BNBV outlets located in the étang de Berre area were excluded from our study due to the lack of meaningful flow direction in a lagoon. The region contains 2580 BNBV outlets that will be allocated to a coarse-resolution grid. The vector representation of the upstream catchment limits serves as the reference for the evaluation and for the implementation of Method 3, while the vector description of river reaches is essential for Method 2.

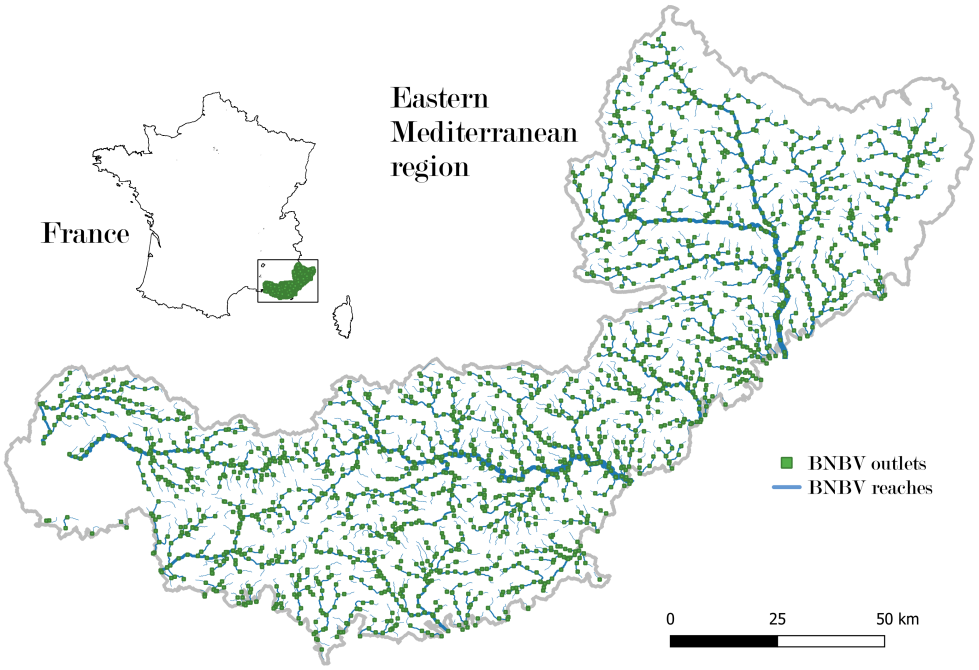


FIGURE IV.6 – The 2580 BNBV outlets on the French Eastern Mediterranean region.

1.3.2 The coarse-resolution hydrological modelling grid

Regional gridded hydrological models are often implemented on 1km resolution grids, aligning with the typical resolution of operational radar-based quantitative precipitation estimates. In this context, the $1km \times 1km$ hydrological modelling grid was herein generated by upscaling the 50m flow direction grid from the BNBV database using the IHU method [Eilander et al., 2021], see figure IV.7 to visualize the upscaling results. The IHU method incorporates principles from previous upscaling methods [Döll and Lehner, 2002, Fekete et al., 2001, Olivera et al., 2002, Paz et al., 2006, Wu et al., 2011] and has demonstrated superior performance compared to the benchmarks methods, e.g. DMM method [Olivera et al., 2002] and EAM method [Yamazaki et al., 2008]. Moreover, the IHU method is the only fully automated and open-source flow direction grid upscaling method known to the authors. After implementing the IHU method, we made minimal manual corrections to the flow direction grid. Only a small number of cells, approximately a dozen out of around 14,000 cells, required adjustments. These manual corrections were primarily made along the zone's borders, particularly near the coastline.

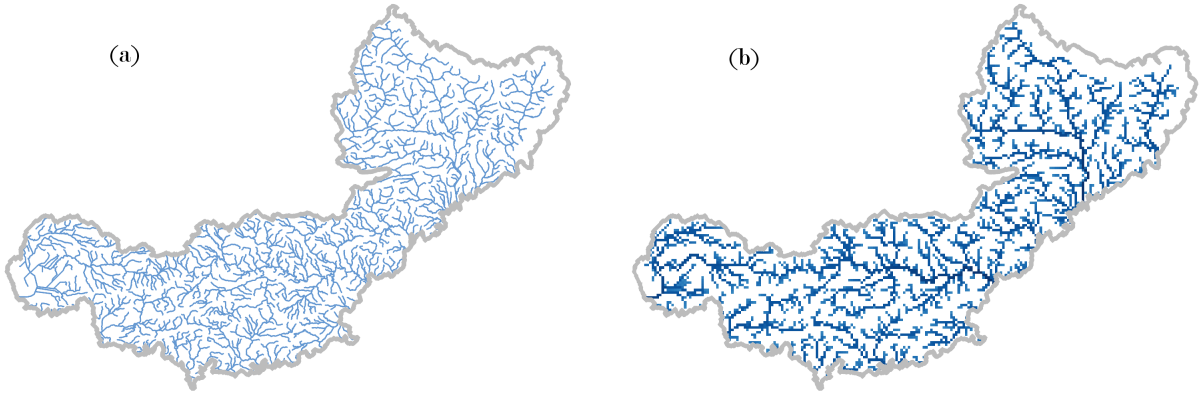


FIGURE IV.7 – Initial $50m$ river network vectorized for $S > 5km^2$ (a) and $1km$ upscaled surface accumulation ($S > 5km^2$) grid (b).

1.4 Results

1.4.1 Advantages and disadvantages of each method

Table IV.1 provides an overview of the general advantages and disadvantages of each method. Method 1 and 2 demonstrate superior performance in terms of computation times. The most time-consuming stage of Method 3 is the computation of the upstream catchment for each candidate cell, and its comparison with the reference catchment. The duration of these processes will naturally increase with the size of the considered case study and the number of points to be allocated. However, the computation times might not be considered as a crucial issue, since the allocation of river points to grid cells need do be performed only once, before hydrological modelling. Method replicability is probably a more important issue for this work. Method 2 is limited in its compatibility as it relies on the IHU upscaling, As a consequence, it makes this method inoperable in contexts where the coarse resolution gridded network does not coincide with the river network (i.e. both networks may come from different data sources). On the other hand, Method 3 requires reference catchment boundaries. Furthermore, the confidence level varies between the methods. Method 1 relies solely on basin area information, which means that near confluences, it may allocate a river point to a neighboring grid cell with a similar basin area but belonging to a different catchment. Method 2 ensures that the river points and chosen grid cell belong to the same river reach, while Method 3 selects the cell with the most similar upstream catchment in terms of basin contour and location.

	Method 1	**Method 2**	**Method 3**
Computation times	≃ Second	≃ Second	≃ Hour
Replicability	Replicable	Requires the definition of *cells outlet points*	Replicable
Additional data	None	River network vector and cell outlet points	Reference basin limits
Confidence level	Variable	High	High
% of allocated outlets	Potentially < 100	Potentially < 100	100

TABLEAU IV.1 – Advantages and disadvantages of each method.

Lastly, Method 3 guarantees the allocation of all river points, covering 100% of the dataset, whereas the other two methods may not achieve complete allocation. For instance, river points similar to P3 in Figure IV.4, i.e. points located between two confluences within the same grid cell, are not allocated with Method 2. These points represent about 2% of the total number of points to be allocated in the considered region. One potential solution could consist in selecting, among the closest upstream or downstream outlets of grids, the one with the nearest usptream watershed area. Another solution would be to allocate

P3 to both upstream cells (C3 and C5 in Figure IV.4), however in this paper we chose not to permit the allocation to several grid cells. In these cases, Method 2 detects that all allocations to a single grid cell will be imperfect, which is why it was not deemed essential to propose an allocation solution for these specific points, for the sake of comparison.

1.4.2 Comparison of allocation performances

In order to compare the quality of allocation among the three methods, we initially examined the CSI statistics. However, it should be noted that while Method 1 and Method 3 successfully allocated all 2580 considered river points (BNBV outlets in the Eastern Mediterranean zone), Method 2 only allocated 2532 points (98%). To ensure a fair comparison, these 48 points were excluded from the analysis. Figure IV.8 displays the histograms of CSI values for each method, along with the corresponding mean and median values.

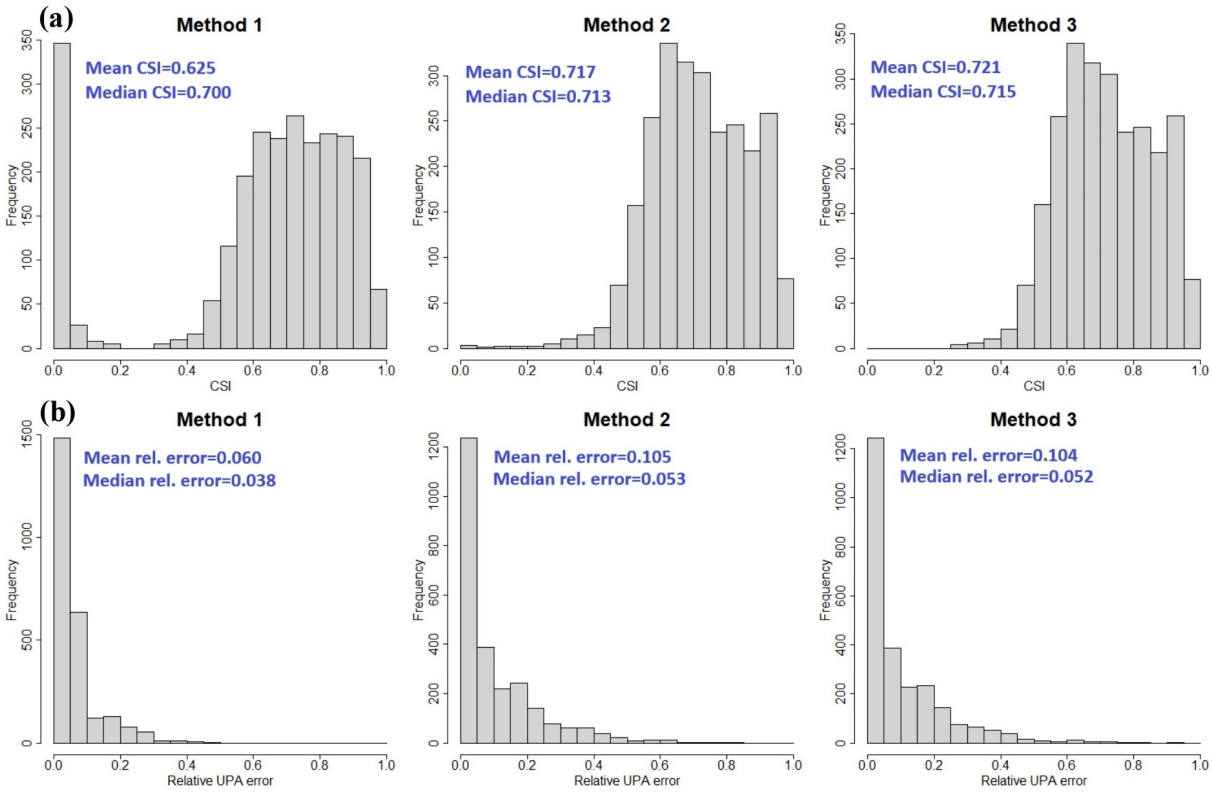


FIGURE IV.8 – Histograms of (a) CSI values and (b) UPA relative error for the three allocation methods.

The results indicate that Method 1 has significantly lower performances compared to Methods 2 and 3. This discrepancy is primarily attributed to a high percentage of river points with very low CSI scores (less than 0.05) in Method 1. These low scores occur when a river point is allocated to a cell solely based on similar basin area, disregarding substantial differences in the shape and location of the upstream catchments (e.g. figure IV.2). In its state, Method 2 cannot allocate 100% of the river outlets, even though it could with a small extension (see previous section). Method 3 was able to allocate the excluded outlets successfully, with a mean CSI of 0.7. Additionally, Figure IV.8 highlights that Method 3 consistently yields a minimum CSI of around 0.25, whereas Method 2 falls below 0.05 (with 12 river points having CSI < 0.25). Figure IV.9 provides a more detailed comparison of CSI scores between Method 2 and Method 3.

It clearly demonstrates the consistent superiority of Method 3 over Method 2. While the CSI differentials are generally small, the green circle in the figure highlights cases where the differential can be

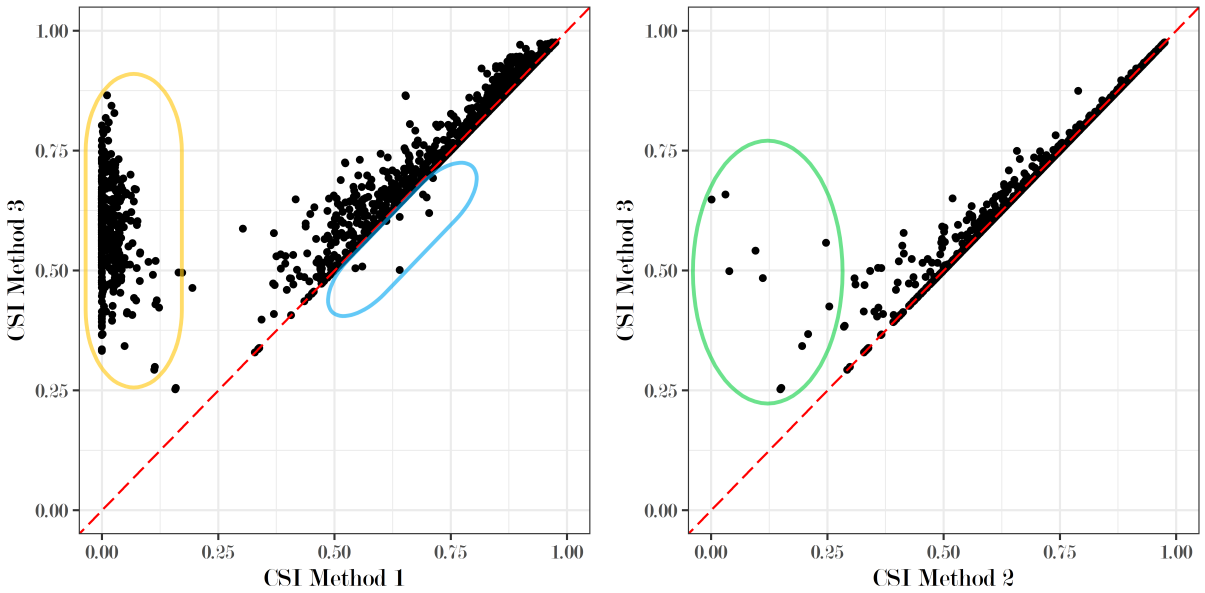


FIGURE IV.9 – Comparison of CSI scores for each outlet, between Methods 3 and 1, and between Methods 3 and 2.

significant. These are characterised by river points with small upstream basin areas (less than $12km^2$) located far from the nearest cell outlet point, resulting in notable differences in UPAs. Figure IV.10 shows an example of these situations, which are often complex configurations where different branches of the network are in close proximity to each other, often within the same grid cell. As a consequence, the upscaled flow directions cannot represent faithfully the river network, and all candidate cells would give a mediocre allocation result. However we argue that Method 3 will give the most correct allocations in these situations. Furthermore, Figure IV.9 reveals that Method 3 effectively corrects the allocation errors made by Method 1 (indicated by orange circles in the figure). However, it also indicates that there are surprisingly cases where Method 1 yields slightly better CSI scores than Method 3 (nine cases circled in blue). This suggests that the minimum CSI values defined in section 1.2.3 for the two possible steps of Method 3 may not be optimal, a topic that will be further discussed in Section 1.4.4.

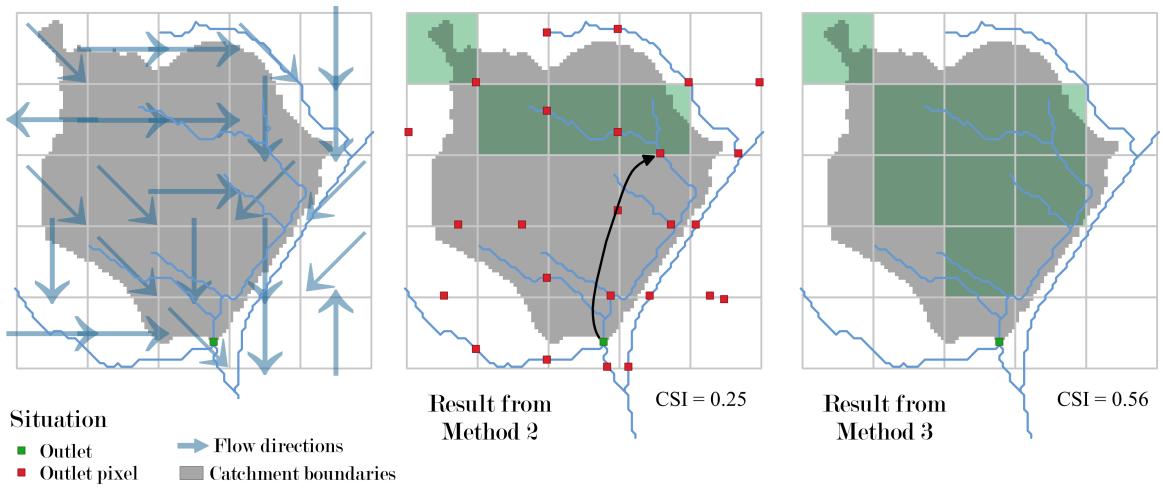


FIGURE IV.10 – An example of high CSI differential between Methods 2 and 3 (basin area $12km^2$). For the sake of comprehension, small ($\leq 5km^2$) tributaries were added to the vectorial network in this figure, explaining why the river outlet can be allocated with Method 2 despite being apparently located between two confluences within the same grid cell.

Finally, the difference in UPAs between each reference catchment and its corresponding coarse

resolution catchment was also calculated, even though we decided not to use this metric for the comparison of the three methods, in agreement with the many hydrologists (e.g. Davies and Bell [2009]) who have pointed out that an evaluation based on UPA comparison is highly uncertain. This thesis is supported by the results presented on figure IV.8b, which show a slightly smaller difference in upstream area for Method 1 than for Methods 2 and 3. Thus, comparing the three allocation methods based on this criterion only would be misleading, because it would not account for all the cases where a river point is allocated to a hydrological modelling cell describing a different upstream catchment, if they have similar UPAs. However, these results show that the relative difference in UPAs remain limited (mostly lower than 15%) for all three methods.

1.4.3 The influence of the upstream basin area

The analyses performed in this study also revealed that basin area plays a significant role in explaining the largest allocation errors. Specifically, among all the river points allocated using Method 1 and having a CSI lower than 0.05, 80% have a basin area smaller than $9km^2$. Similarly, with Method 3, among all the river points with CSI scores lower than 0.6, 100% have a basin area lower than $25km^2$, and 92% have a basin area lower than $10km^2$.

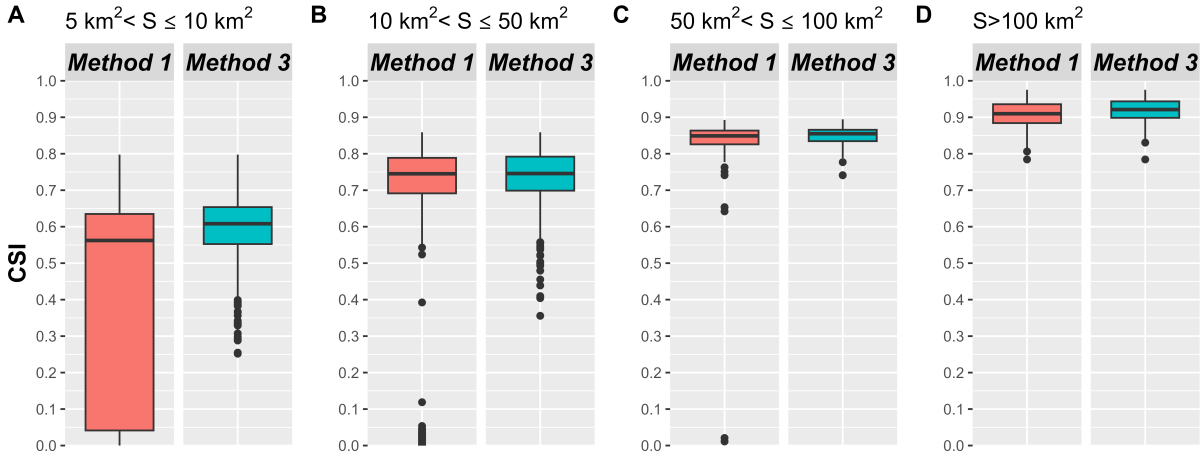


FIGURE IV.11 – Boxplots of CSIs divided into classes of surfaces for Methods 1 and 3.

The boxplot comparisons in Figure IV.11 highlights two important observations. Firstly, Method 3 effectively prevents the largest allocation errors generated by Method 1. Secondly, these allocation errors predominantly occur for small catchment areas. For catchments larger than $100km^2$, CSI scores systematically exceed 0.75 for both methods. Consequently, Method 1 can be considered a reliable method when applied to the main river network, which explains its widespread usage in previous studies : i.e. for almost all the studied watersheds, the minimum catchment size in various previous research works is between 100 and 10, 000 km^2 [Fekete et al., 2002, Döll and Lehner, 2002, Sutanudjaja et al., 2018, Wang et al., 2018, Zhao et al., 2017, Burek et al., 2020, Polcher et al., 2023].

In addition, Figure IV.11 also confirms that, with Method 3, the lowest CSI scores are obtained for small catchments - the minimum CSI value is consistently higher than 0.7 for catchment sizes larger than $50km^2$. The low CSI values reflect the uncertainties affecting the boundaries of small basins defined on a 1km resolution grid, particularly pronounced for narrow watersheds. However, it is important to note that the impact of low CSI scores on the representation of rainfall is relatively minor for small catchments compared to larger ones, because of the limited variability of rainfall at scales of the order of a square kilometre.

1.4.4 Effect of the criterion conditioning Method 3 iteration

While Method 3 appears, without surprise, to be the most reliable method, it has also some shortcomings. One is the limited search zone on the coarse-resolution hydrological grid. The current implementation first considers only the nine surrounding pixels and then extends the search area to forty-nine surrounding pixels under certain conditions. This approach was chosen to reduce the computation time, which is significantly longer for Method 3 compared to the other two methods. The criterion used for the second iteration depends on the catchment size, based on the assumption that the upstream drainage area (UPA) would influence the allocation results.

The second iteration of Method 3, is only activated for 70 outlets out of 2580. Extending the research area in the second iteration improved the allocation for only 10 out of these 70 outlets. However, the increase in CSI for these improved allocations was quite significant, with a median increase of 133%. Moreover, it is obvious on figure IV.9 that Method 3 does not lead to the optimal allocation in some cases : see the 9 cases circled in blue. A visual analysis indicates that the search area composed of the 49 surrounding grid cells would have solved the problem, if the second iteration of Method 3 would have been activated. Increasing the threshold values for this activation to 0.55 (resp. 0.65) for catchment areas lower (resp. higher) than 10 km^2, would solve the problem encountered for these 9 points in the present case study, but at the cost of higher computing times : 353 points processed in the second iteration of method 3, instead of 70 for the initial theshold values.

1.5 Discussion and conclusions

This comparative study of methods allocating river points to coarse grid cells was driven by the shift in approach from area-based methods to contour-based methods. In this work, we compared these two categories of methods and introduced a new approach based on topological proximity.

1.5.1 Performance and scope of each method

The study results revealed that contour-based methods were more relevant and satisfying from a hydrological point of view, although costly in terms of computing time. The introduced topology-based method is a good compromise because it leads to similar allocation quality than the contour-based method. However, it is inoperable when the fine and coarse resolutions river networks come from different data sources, since it requires the definition of "cells outlet points" as well as the vector-based description of the river network. Moreover, it cannot allocate as many points as the contour-based method.

The area-based method generated numerous allocation errors, which the contour-based method was able to address for a significant portion of them. However, upon closer examination, it was observed that the performance gap between both methods was more pronounced for small catchments, while being less significant for larger catchments (with $S > 100km^2$). The area-based methods thus lead to satisfying results if we only consider river points with large UPAs compared to the grid cell resolution. Based on the results obtained, we would recommand a minimum factor of 100 between a river point's UPA and the resolution of the hydrological modelling grid for the application of an area-based method.

The transferability of the results outside the test area and to coarser resolutions (i.e. global hydrological models) is debatable, as there are many uncertainties and non-linearities in the representation of hydrological information at such larger scales and coarser resolutions. However, it is very likely that, with coarser resolution grids, allocation problems will increase and that errors related to area-based methods will impact larger catchments (i.e. larger than $100km^2$). Even if this needs to be verified, the

"contour-based" method will certainly remain more effective at coarser resolutions than the "area-based" method.

1.5.2 Remaining limitations of allocating river points to coarse grid cells

Even if allocation errors are reduced, some low CSI scores remain with contour-based methods, due to the inherent difficulty of representing the boundaries of small basins at a 1km resolution. The situation illustrated in figure IV.10 is an example of configurations where different branches of the river system are crossing the same coarse grid cell, which makes it impossible to correctly represent the river network and basin contour with coarse flow directions. However, it is important to note that these inevitable errors resulting in low CSI scores are generally found on small catchments, and are less problematic than low CSI scores on larger catchments, because of the more limited variability of rainfall at smaller spatial scales.

A possible way to reduce these remaining errors could be to allocate river points to multiple grid cells, by either taking the sum of the upstream cells, or the difference of the downstream cells, even though it could complicate the hydrological modelling at a later stage. Another approach that could circumvent the challenges faced would be to use hydrological models structured based on vectorial objects instead of regular grids. These models preserve the topology of river networks and allow seamless integration of observational data. However, vector-based modeling also introduces its own challenges related to data and computational requirements and the need for accurate input data.

2 Mise en œuvre d'une chaîne de prévision basée sur les impacts

La modélisation hydrologique, présentée en Chapitre II, permet d'estimer les débits et périodes de retour atteints sur une grille de résolution kilométrique, dont les pixels peuvent être associés à chaque exutoire de la BNBV en utilisant l'approche "contour-based" décrite dans la section précédente. La chaîne de simulation étudiée au Chapitre II permet donc d'obtenir, à chaque pas de temps de 15 minutes sur la période d'intérêt (septembre 2009 - septembre 2023), un intervalle de périodes de retour $[T_{min}; T_{max}[$ atteint sur chaque tronçon de la BNBV. Ceci correspond donc au fonctionnement de la chaîne opérationnelle Vigicrues Flash. En revanche, dans ce système, l'estimation des quantiles de crue est limitée à la période de retour 20 ans (seuil d'avertissement maximal), tandis que dans ce travail, les périodes de retour sont estimées jusqu'à 1000 ans, via la méthode SHYREG (Chapitre II, Section 3.5).

Cette section vise à décrire les méthodes mises en place pour passer d'une chaîne d'avertissement de type HBW, à une chaîne d'avertissement de type IBW (voir figure IV.12), en mobilisant notamment des données déjà présentées, telles que le catalogue de scénarios d'inondation pré-établi pour des crues théoriques de périodes de retour fixées comprises entre 2 et 1000 ans (Chapitre III).

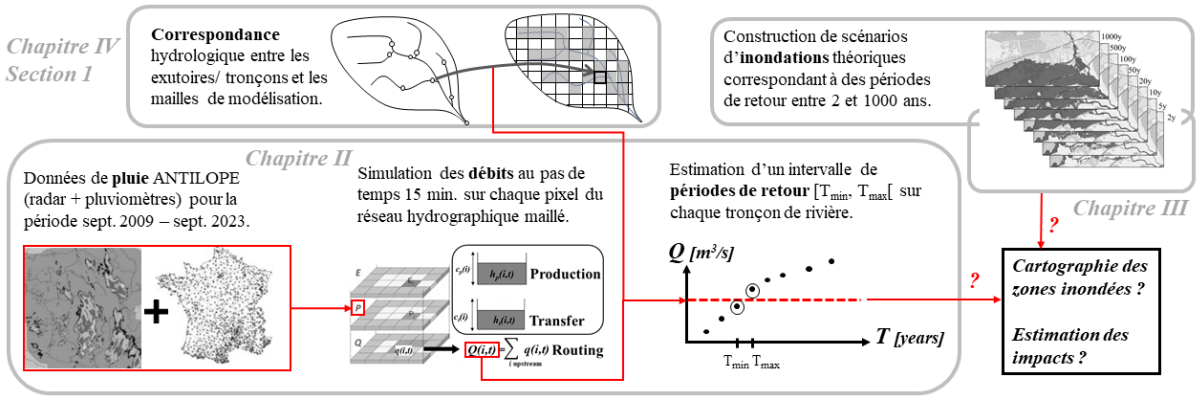


FIGURE IV.12 – État de la construction de la chaîne de modélisation à ce stade du manuscrit.

2.1 Reconstitution de l'emprise inondée

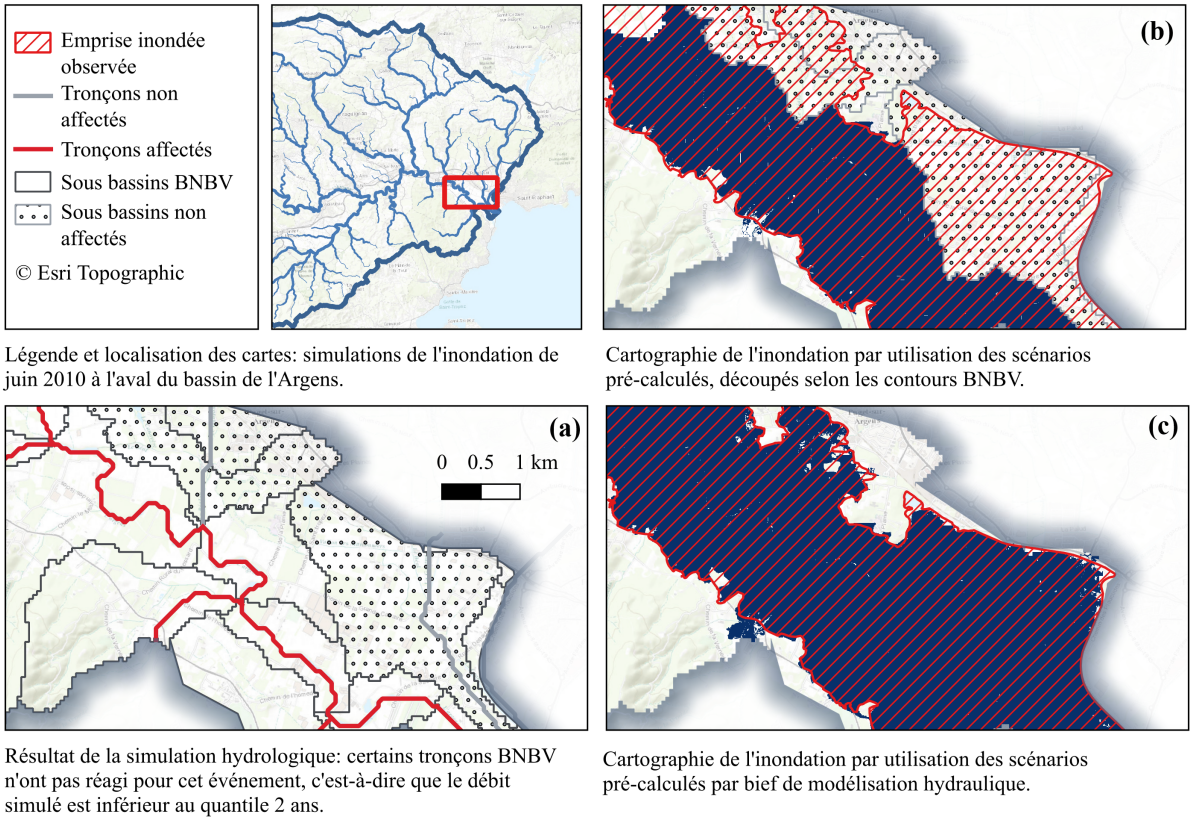
La cartographie des zones inondées dans une chaîne de prévision appliquée à grande échelle repose généralement sur l'utilisation de scénarios d'inondation pré-calculés [Dottori et al., 2017, Alfieri et al., 2024]. La réalisation de calculs hydrauliques en temps réel, bien que potentiellement plus précis, restent aujourd'hui incompatibles avec les contraintes opérationnelles, en raison des temps de calcul élevés pour des simulation hydraulique 2D appliqués à des résolutions fines (<5m). Ceci est d'autant plus valable pour les crues soudaines, qui se forment très rapidement et laissent très peu de temps de préparation aux utilisateurs.

2.1.1 Méthode (1) existante pour la reconstruction de l'emprise inondée.

Dans des travaux menés sur l'arc méditerranéen par Le Bihan [2016], Charpentier-Noyer [2022], la méthode consistait à utiliser le découpage en tronçons de cours d'eau de la BNBV [Organde et al., 2013], sur lesquels étaient estimés les débits et leurs périodes de retour, afin d'établir pour chaque tronçon une relation entre période de retour et emprise inondée.

Une approche analogue a été testée dans un premier temps pour la construction de la chaîne d'avertissement de type IBW appliquée dans cette thèse. Toutefois, l'analyse d'événements réels a révélé d'importantes limites. La figure IV.13 illustre le cas de la crue de juin 2010 à l'aval du bassin de l'Argens.

La carte IV.13a montre les tronçons affectés selon la modélisation hydrologique, c'est-à-dire ceux dont la période de retour estimée dépasse le seuil de 2 ans. Dans la méthode de Dottori et al. [2017], l'inondation est cartographiée en sélectionnant, pour chaque tronçon affecté, le scénario dont la période de retour est immédiatement inférieure à celle estimée. En conséquence, aucun scénario n'est choisi pour les tronçons dont la période de retour est inférieure à 2 ans (tronçons gris). La carte IV.13b présente le résultat obtenu après assemblage des scénarios d'inondation sélectionnés sur chaque tronçon de la BNBV. Comparée à l'emprise réellement observée, la sous-estimation est manifeste.



Légende et localisation des cartes: simulations de l'inondation de juin 2010 à l'aval du bassin de l'Argens.

Cartographie de l'inondation par utilisation des scénarios pré-calculés, découpés selon les contours BNBV.

Résultat de la simulation hydrologique: certains tronçons BNBV n'ont pas réagi pour cet événement, c'est-à-dire que le débit simulé est inférieur au quantile 2 ans.

Cartographie de l'inondation par utilisation des scénarios pré-calculés par bief de modélisation hydraulique.

FIGURE IV.13 – Illustrations de deux méthodes de reconstruction de l'emprise inondée dans la chaîne de modélisation, pour l'événement de juin 2010 sur le bassin de l'Argens.

En effet, certains tronçons considérés comme non affectés apparaissent pourtant inondés, comme le montrent les zones en pointillés de la figure IV.13b. Ces débordements sont très probablement liés à des tronçons voisins, situés sur une zone plate à l'aval du bassin possédant de nombreuses confluences, où l'inondation associée à un tronçon peut très facilement s'étendre au delà des lignes de crête séparant les sous bassins versants associés à chaque tronçon. Dans ce cas, la méthode qui repose sur un découpage des emprises inondées en fonction des contours de chaque sous-bassin versant [Le Bihan, 2016, Charpentier-Noyer, 2022] conduit à une sous estimation marquée. Le problème est d'autant plus critique que les zones concernées, très urbanisées, comptent ici plus de 1100 bâtiments résidentiels, commerciaux, industriels ou agricoles.

2.1.2 Méthode (2) proposée pour la reconstruction de l'emprise inondée

Pour pallier ces limites, il est proposé de ne plus découper les emprises inondées en fonction des contours de sous-bassins versants BNBV (figure IV.14a), qui restent liés à la modélisation hydrologique et ne sont pas adaptés à l'application hydraulique. L'approche développée dans ce travail consiste à se baser directement sur les limites de domaines de modélisation hydraulique qui ont été utilisés pour

construire le catalogue de scénarios d'inondation sur l'arc méditerranéen (cf. Chapitre III). Comme détaillé au Chapitre III (voir annexe 4), le réseau hydrographique de la BD TOPAGE® a été retenu pour la modélisation hydraulique, offrant une représentation plus fine que celle de la BNBV. Il a été découpé en tronçons afin de limiter la taille des domaines de calcul hydraulique. Chaque tronçon a été prolongé de 500m vers l'aval pour réduire l'influence des conditions aux limites, et le domaine de simulation correspondant a été étendu de 2km dans le 4 directions cardinales, de façon à inclure l'ensemble du lit majeur et à permettre un large débordement du cours d'eau le cas échéant.

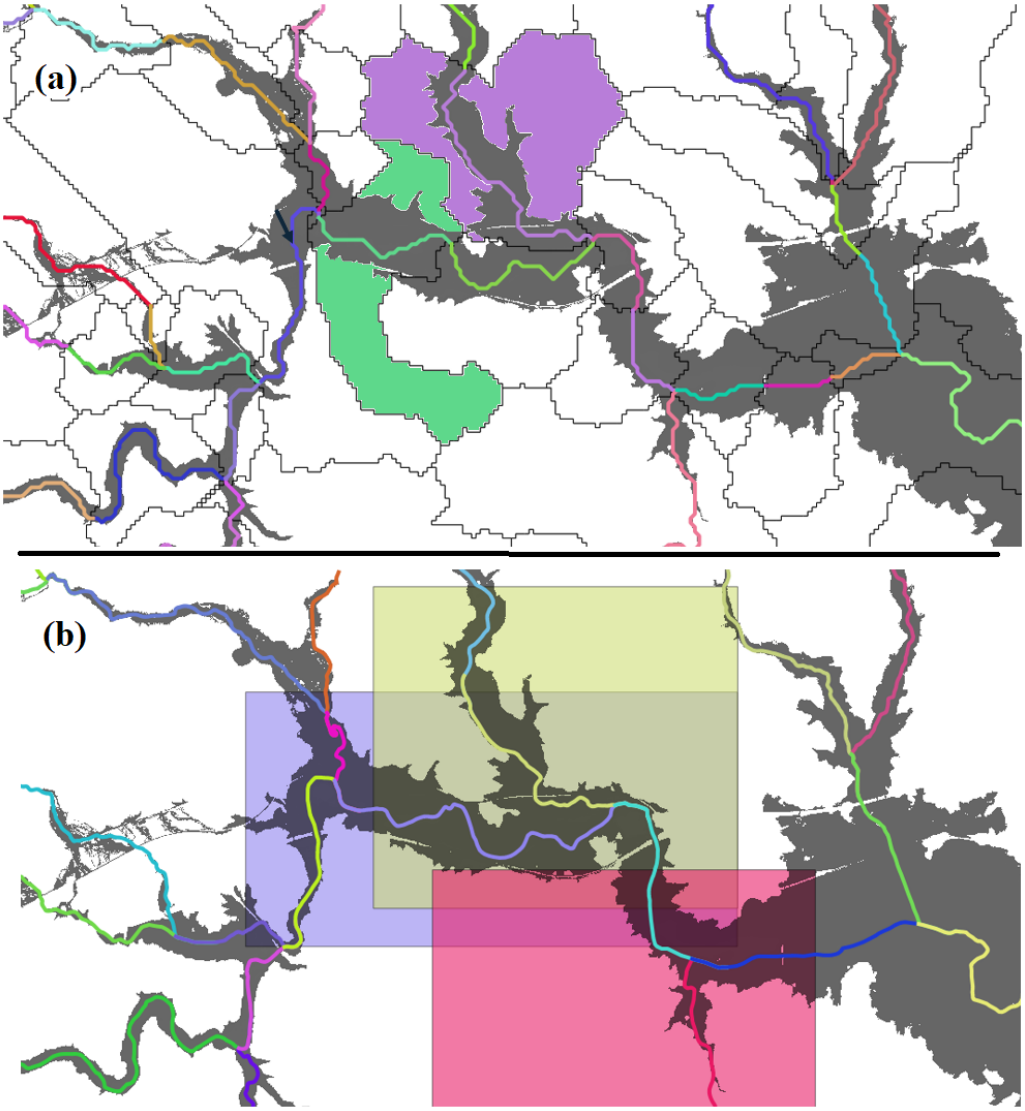


FIGURE IV.14 – Illustrations de deux méthodes de reconstruction de l'emprise inondée : (a) les périodes de retour sont estimées sur le réseau hydrographique de la BNBV et les scénarios d'inondation sont découpés selon les contours du sous-bassin associé à chaque tronçon ; (b) les périodes de retour sont estimées sur le réseau hydrographique utilisé dans la modélisation hydraulique (BD TOPAGE), et les scénarios d'inondation calculés sur chaque tronçon sont superposés sans découpage.

La méthode développée dans ce travail consiste à associer une période de retour à chaque tronçon utilisé pour les calculs hydrauliques [1], puis à sélectionner directement le scénario hydraulique correspondant sur tout le domaine de calcul, sans effectuer de découpage (figure IV.14b). Dans les zones de recouvrement, la hauteur d'eau maximale est retenue. Appliquée à la crue de l'Argens en juin 2010, cette

1. Étape déjà réalisée lors du prétraitement des données d'entrée pour la construction des scénarios d'inondation (voir annexe 4) : les débits SHYREG, disponibles pour des périodes de retour de 2 à 1000 ans sur le réseau BNBV, ont été injectés dans le réseau de la BD TOPAGE pour alimenter le modèle hydraulique.

méthode (figure IV.13c) reproduit beaucoup plus fidèlement l'emprise inondée observée.

2.2 Calculs des impacts

2.2.1 Construction d'une relation Débits–Impacts

L'avantage de la méthode 1 présentée précédemment pour la reconstruction des surfaces inondées est de traiter chaque tronçon indépendamment, puisque les contours de découpe associés à chaque tronçon ne se recouvrent pas. Cette approche permet, comme l'ont proposé Le Bihan et al. [2017], Charpentier-Noyer [2022], de définir des relations Débits (ou périodes de retour) – Impacts pouvant être pré-calculées et appliquées directement dans la chaîne de simulation, sans avoir à reconstituer l'emprise inondée (figure I.19).

La méthode 2 développée dans ce travail évite les limites identifiées pour la méthode 1, mais rend plus complexe l'établissement d'une telle relation. En effet, l'emprise inondée ne dépend plus seulement du tronçon concerné, mais également de ceux situés à proximité.

Ainsi, avec la méthode 2, le calcul anticipé des impacts doit intégrer non seulement l'ensemble des tronçons d'une zone donnée, mais aussi ceux extérieurs dont les domaines de calcul hydraulique s'entrecoupent avec elle. Le terme zone désigne ici l'échelle à laquelle les impacts sont comptabilisés : soit le tronçon de cours d'eau, soit la commune (voir chapitre V). Pour une zone recoupée par N domaines de calcul hydrauliques, il faut considérer toutes les combinaisons possibles de scénarios : absence d'inondation (période de retour < 2ans) ou dépassement des seuils 2, 5, 10, 20, 50, 100, 500 ou 1000 ans. Soit, pour chaque zone, 9^N scénarios d'impacts à pré-calculer.

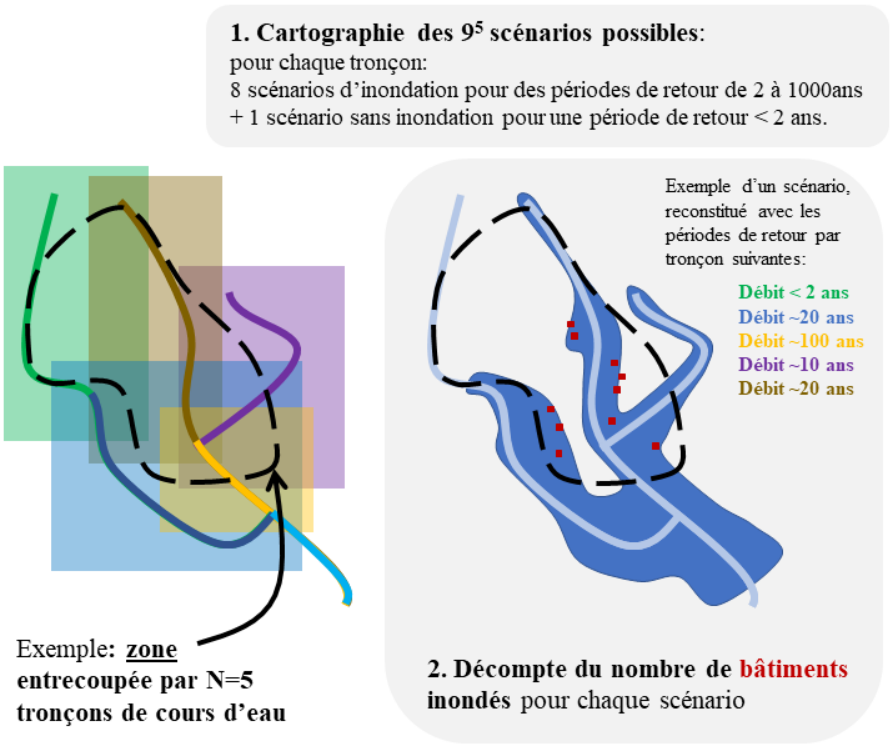


FIGURE IV.15 – Illustration de la construction d'une matrice Débits-Impacts.

2.2.2 Sélection d'une variable représentative des impacts

Jusqu'ici, le terme impacts désignait le dernier maillon de la chaîne de modélisation, sans préciser leur nature. Or, les impacts potentiels sont multiples (voir Chapitre I, section 2.1.4) : dommages économiques

ou humains, coupures de routes, mobilisation des secours, etc. Dans ce travail, le choix a été fait de se concentrer sur les impacts liés au bâti - afin de correspondre au mieux aux données de validation, comme cela sera approfondi dans le Chapitre V. Concrètement, il s'agit de compter, pour chaque zone (tronçon ou commune), le nombre de bâtiments inondés. Finalement, la comparaison de ce nombre à un nombre de bâtiments "seuil" prédéfini permet de définir les avertissements IBW.

La BD TOPO, produite par l'IGN, a été utilisée à cette fin. Cette base vectorielle décrit divers éléments du territoire (administratif, bâti, hydrographie, infrastructures, etc.). Seule la couche bâtiments a été retenue, après un prétraitement détaillé en annexe A, afin de ne conserver que les bâtiments résidentiels, commerciaux, agricoles ou industriels.

2.2.3 Représentation statique du bâti sur une période de plus de dix ans

La version de 2020 de la BD TOPO a été utilisée dans ce travail : https://geoservices.ign.fr/bdtopo#telechargementshpdep2020 [2].

Cette décision n'a pas été prise sans connaissance des limites introduites. En effet, ce travail couvre une période de plus de dix ans, durant laquelle les enjeux présents sur le territoire ont naturellement évolué. Ainsi, la simulation d'un événement s'étant produit en 2010 utilisant les enjeux présents sur le territoire en 2020 est source d'erreurs. Cependant, il faut noter que si l'évolution du nombre d'enjeux présents sur le territoire métropolitain est avérée, elle n'est pas forcément significative [3], d'autant plus en zone inondable, qui constitue l'emprise spatiale maximale de notre modélisation. On peut en effet penser que depuis la loi du 2 février 1995 permettant le financement des Plans de Préventions de Risques, l'augmentation du nombre de PPRi [4] a permis de limiter les nouvelles constructions dans les zones les plus à risque d'inondation.

Pour justifier notre choix de retenir un scénario unique de représentation des enjeux sur le territoire, nous avons tracé la surface de bâtiments en zones inondables par commune au cours du temps. Pour cela, nous avons extrait les données de bâti des BD TOPO 2010, BD TOPO 2014, et BD TOPO 2018. Nous avons compté le nombre de bâtiments de nature "Indifférenciée" (regroupant en grande partie les bâtiments résidentiels) dans l'emprise inondée associée au scénario d'une crue millénale.

Ces résultats, illustrés en figure IV.16, montrent que l'évolution des bâtiments en zone inondable est globalement faible, même si quelques points se démarquent : les trois points entourés en bleu correspondent aux communes de Claira dans les Pyrénées-Orientales, de Lançon-Provence dans les Bouches du Rhône, et de Mauguio dans l'Hérault. Dans ces communes, l'hypothèse de stabilité des enjeux en zone inondable est plus débattable. Néanmoins, nous considérons valable cette hypothèse de manière globale. Par ailleurs, les limites de la BD TOPO présentées en figure IV.17 découragent l'utilisation de BD TOPO anciennes. Pour cette raison et afin de faciliter les calculs, nous avons donc choisi de ne travailler qu'avec la BD TOPO 2020.

2.2.4 Limites liées à l'utilisation de la BD TOPO

Cette section présente quelques limites rencontrées lors de l'utilisation de la BD TOPO dans le cadre des objectifs de la thèse. D'une part, dans certains secteurs de la zone d'étude, les anciennes versions de la BD TOPO présentaient des limites notables quant à la précision de la représentation des bâtiments. La figure IV.17 illustre ces différences, notamment en ce qui concerne la séparation de maisons mitoyennes

2. Ce choix avait été effectué pour être en cohérence avec les données de population de l'INSEE, dont le dernier recensement date de 2020 - qui n'ont, au final, pas été utilisées dans cette thèse, mais dans des travaux connexes [Nicolle et al., 2024]
3. Selon l'INSEE [2023], le parc de logements en France hors Mayotte s'est accru de 1,1 % par an en moyenne.
4. En 1982, moins de 2000 communes étaient couvertes par un PPRi, contre plus de 10 000 en 2022, selon la CCR [2023c].

dans une zone résidentielle, entre les versions 2010 et 2018. Cette limite est particulièrement déterminante dans le cadre de ce travail, où les avertissements basés sur les impacts reposent sur le dépassement d'un seuil de nombre de bâtiments inondés. Ce problème n'a toutefois pas été observé dans la version utilisée dans la thèse (2020).

D'autre part, des vérifications manuelles ont permis d'identifier des cas où la classe des bâtiments n'était pas correctement identifiée, malgré les prétraitements décrits en annexe A. Si de nombreux bâtiments résidentiels sont classés comme "Indifférencié" dans les champs "NATURE", "USAGE1" et "USAGE2", cette catégorie peut également concerner des bâtiments manifestement non résidentiels (par exemple des commerces, des services administratifs, etc.).

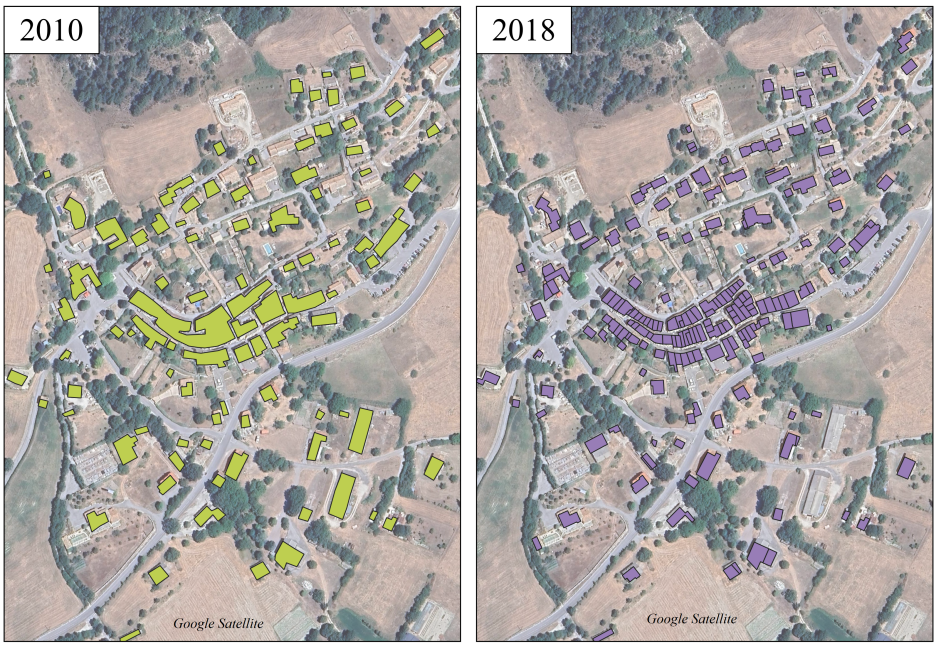


FIGURE IV.17 – Limites de la BD TOPO liées au dénombrement des bâtiments

2.3 Application de la chaîne complète pour le besoin la thèse et transférabilité en contexte opérationnel

La chaîne de modélisation décrite aux paragraphes précédents a été rejouée de septembre 2009 à septembre 2023, depuis les données de pluie en entrée jusqu'au décompte des bâtiments inondés en sortie. Les débits et leurs périodes de retour ont été estimés pour chaque tronçon toutes les 15 minutes, puis la valeur journalière maximale a été retenue. Ainsi, la sortie du modèle correspond à un nombre de bâtiments inondés par jour, par commune ou par tronçon. Ce choix est cohérent avec l'évaluation de la chaîne (Chapitre V), qui a été conduite soit de façon continue au pas de temps journalier, soit à l'échelle d'événements d'au moins une journée.

La chaîne d'avertissement de type IBW proposée est parfaitement transposable en contexte opérationnel. Certes, certaines étapes de prétraitement – notamment la construction des relations Débits–Impacts – sont chronophages, mais elles ne doivent être réalisées qu'une seule fois. En phase opérationnelle, le délai principal provient de la réception des données pluviométriques ANTILOPE TR 15min, produites par Météo-France et transmises au Service Central Vigicrues en environ 8 minutes. Les étapes suivantes – modélisation hydrologique à l'échelle régionale, calcul des périodes de retour, puis estimation des impacts via les relations pré-calculées – sont, elles, quasi instantanées.

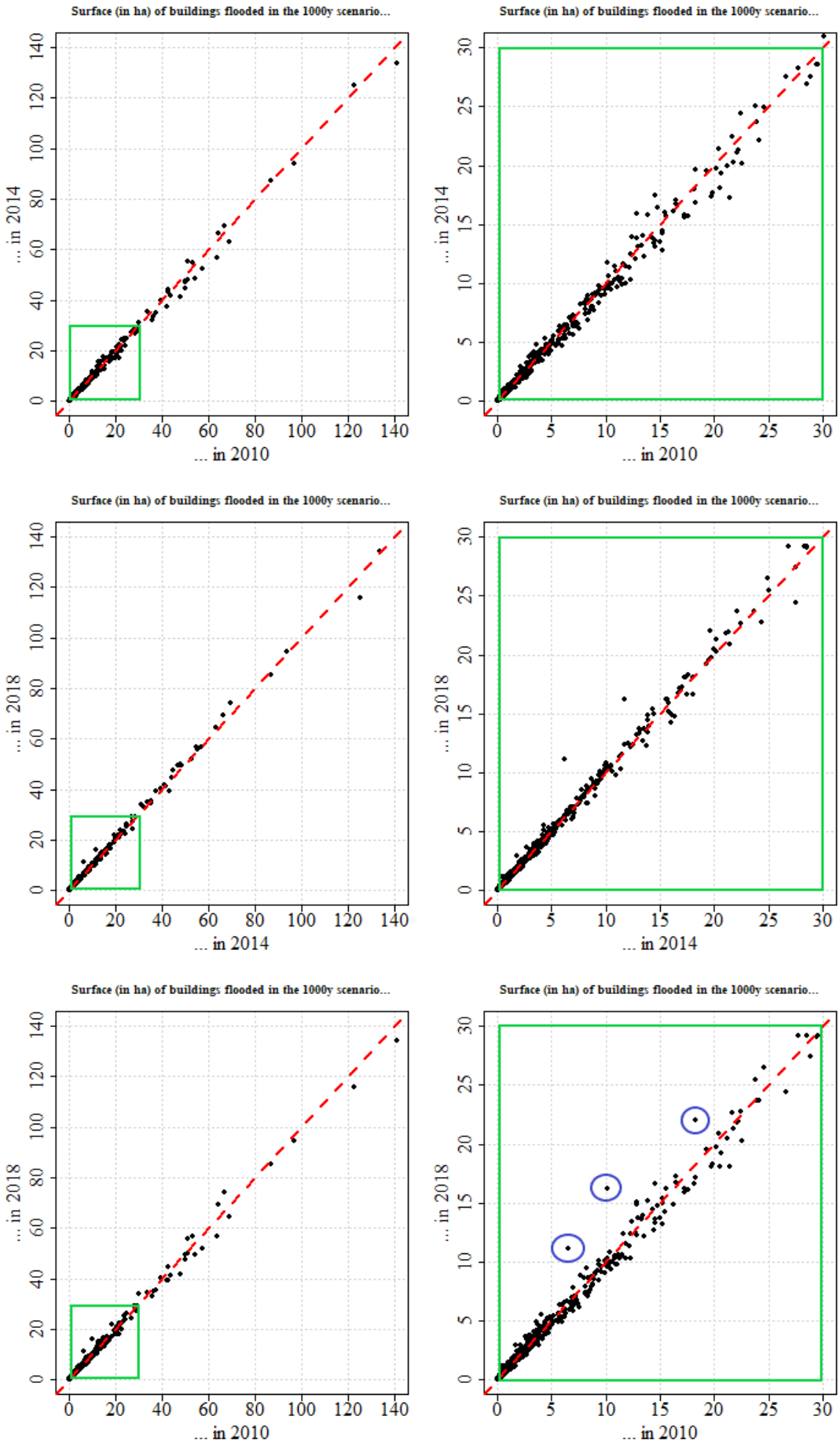


FIGURE IV.16 – Evolution de la surface (en hectares) des bâtiments en zone inondable par une crue millénale, par commune située dans la zone d'étude

3 Discussions et conclusions

Ce chapitre s'est attaché aux aspects techniques de l'implémentation d'une chaîne d'avertissement de type IBW. Il fait suite au Chapitre II, dans lequel la modélisation hydrologique, basée sur le système opérationnel Vigicrues Flash, a été présentée et évaluée. Il prolonge également le Chapitre III, consacré à la modélisation hydraulique et notamment aux scénarios d'inondation pré-calculés qui sont directement mobilisés dans la chaîne de prévision proposée. Dans ce chapitre, trois points techniques majeurs ont été traités.

La connexion entre l'hydrologie (maille) et l'hydraulique (vectoriel)

Le premier point technique abordé concerne l'association des débits de pointe définis sur des mailles de la grille de modélisation hydrologique avec des points du réseau rivière (vectoriel) utilisé pour la modélisation hydraulique. Traditionnellement, cette association repose sur des méthodes comparant les aires des bassins versants amonts aux points étudiés. Plus récemment, des méthodes fondées sur la comparaison des contours des bassins versants amonts ont émergé, car jugées plus robustes, mais leur plus-value n'avait encore jamais été quantifiée. Une comparaison de ces deux approches a donc été proposée et a montré clairement l'intérêt des méthodes basées sur les contours, notamment pour les petits bassins (surfaces inférieures à 10 km^2) particulièrement vulnérables aux crues soudaines. Les résultats de cette étude ont également montré que l'écart de performance entre les deux approches diminue fortement à mesure que la taille du bassin augmente. Dans le cadre des crues soudaines, déclenchées par des cellules pluvieuses très localisées, il reste cependant primordial d'assurer une correspondance point/maille optimale pour ces petits bassins. À la suite de ce travail, la méthode basée sur les contours des bassins a été adoptée [5].

Cette problématique de connexion entre un point sur un réseau rivière (par exemple une station hydrométrique) et une maille d'un modèle maillé est en réalité inhérente à l'utilisation d'un modèle hydrologique à mailles régulières. Elle peut néanmoins être évitée en utilisant un modèle hydrologique dont la structure repose sur une représentation vectorielle du réseau hydrographique. C'est le cas du modèle Cinecar [Gaume et al., 2004, Naulin et al., 2013], brièvement présenté dans le Chapitre III (section 3.2). Comme l'illustre la figure IV.18, le réseau vectoriel y est découpé en tronçons, chacun étant associé à deux versants rectangulaires. Les débits en sortie du modèle peuvent ainsi être directement récupérés aux exutoires de chaque sous-bassin. Des recherches sont actuellement en cours pour tester un couplage entre SMASH (modèle maillé utilisé dans cette thèse) et Cinecar : dans ce schéma, SMASH assurerait la fonction de production et produirait des pluies nettes "pixellisées" qui seraient injectées, en lieu et place des pluies radar, dans le modèle Cinecar afin d'assurer le transfert dans le réseau hydrographique.

La reconstruction des emprises inondées

Le deuxième point technique traité a porté sur la reconstitution de l'emprise inondée à partir de scénarios pré-calculés pour huit périodes de retour (présentés au Chapitre III). L'application d'une première méthode, inspirée de travaux de thèse antérieurs [Le Bihan, 2016, Charpentier-Noyer, 2022], a mis en évidence plusieurs limites sur un cas d'étude : en considérant les tronçons de cours d'eau de manière indépendante, cette méthode présente un risque important de sous-estimation dans les zones plates où l'inondation peut dépasser les lignes de crête associées à un tronçon. La méthode proposée dans ce chapitre repose plutôt sur les domaines de modélisation hydraulique associés à chaque tronçon,

5. Suite à ce travail, la méthode "contour-based" a été implémentée dans la plateforme de modélisation hydrologique SMASH [Colleoni et al., 2024].

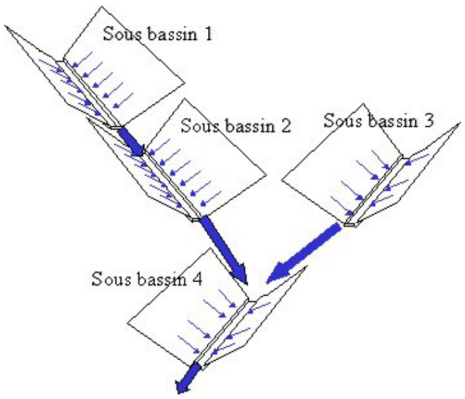


FIGURE IV.18 – Représentation schématique d'un bassin versant dans le modèle hydrologique Cinecar [Le Bihan, 2016].

qui peuvent se recouper, permettant ainsi une représentation plus réaliste, notamment dans les zones de confluence.

Il n'est actuellement pas possible de réaliser des calculs hydrauliques fins à l'échelle de grands bassins versants, ou de grandes zones géographiques regroupant un grand nombre de bassins versants, en temps réel. Pour cette raison, l'utilisation de scénarios d'inondation pré-calculés est une pratique courante. Dans ce travail, huit scénarios ont été utilisés, couvrant des périodes de retour de 2 ans à 1000 ans. À l'échelle de chaque tronçon, le choix appliqué consiste à sélectionner le scénario dont la période de retour est immédiatement inférieure à celle du débit simulé par le modèle hydrologique, comme proposé par Dottori et al. [2017] à l'échelle européenne. Cette approche par scénarios introduit nécessairement des incertitudes. Des recherches sont en cours pour développer des méthodes d'interpolation entre scénarios sur l'arc méditerranéen français.

Le calcul des impacts

Enfin, le troisième point technique abordé concerne le calcul des impacts, qui constitue le dernier bloc de la chaîne de prévision de type IBW appliquée dans cette thèse. Le choix a été fait de se concentrer sur l'exposition du bâti, afin de correspondre au mieux aux données de validation mobilisées dans le Chapitre V. Les données utilisées proviennent de la BD TOPO, qui ont été critiquées et triées (cf. Annexe A). Une méthode a ensuite été proposée pour la construction d'une relation Débits–Impacts, de manière à rendre la chaîne de prévision basée sur les impacts largement compatible avec un contexte opérationnel. Pour répondre à l'objectif fixé dans le Chapitre V, visant à évaluer la plus-value de cette chaîne par rapport au système opérationnel Vigicrues Flash, il a été choisi de produire en sortie un décompte des bâtiments affectés par l'inondation, à l'échelle du tronçon de cours d'eau ou de la commune, et à un pas de temps journalier.

Toutefois, les sorties de la chaîne pourraient être diversifiées afin de répondre aux besoins variés des utilisateurs. Dans ce travail de thèse, nous nous sommes focalisés sur le bâti, mais d'autres enjeux sont tout aussi importants : les services, les réseaux de transport, les réseaux électriques, etc. Ces informations peuvent être représentées sur des cartes distinctes ou bien agrégées sur une carte unique en attribuant des poids à chaque élément, comme proposé par [Saint-Martin et al., 2016].

Chapitre V

Évaluation de la plus-value d'une chaîne de prévision immédiate basée sur les impacts

Objectifs et organisation du chapitre :

Ce chapitre porte sur l'évaluation de la plus-value d'Impact-based Warnings (IBW) par rapport aux "Hazard-based Warnings" (HBW) pour la détection des zones impactées par les crues soudaines. L'évaluation proposée couvre une large fenêtre spatio-temporelle, grâce au rejeu des deux chaînes d'avertissement sur une période continue de 13 ans (2010-2022) et sur près de 20 000 km de linéaire de cours d'eau de l'arc méditerranéen français. La question de la disponibilité et de l'exhaustivité des données d'impacts observés est abordée, au travers de l'étude d'un large éventail de bases de données d'impacts. La comparaison des avertissements émis (IBW et HBW) avec les données d'impacts retenues permet de répondre aux questions suivantes :

- Les IBW permettent-ils de réduire significativement les fausses alarmes par rapport aux HBW ?
- À l'échelle de l'événement, les IBW offrent-ils une meilleure hiérarchisation des zones affectées que les HBW ?
- Les différentes bases de données d'impacts utilisées pour l'évaluation conduisent-elles aux mêmes conclusions ? L'évaluation permet-elle de mettre en évidence les limites associées à chacune de ces bases ?

Ce chapitre reprend l'intégralité d'un article soumis en octobre 2025 dans la revue *International Journal of Disaster Risk Reduction*. La première section propose une synthèse en français de cet article, et les sections suivantes reprennent le texte intégral de l'article, rédigé en anglais.

0 Synthèse en français

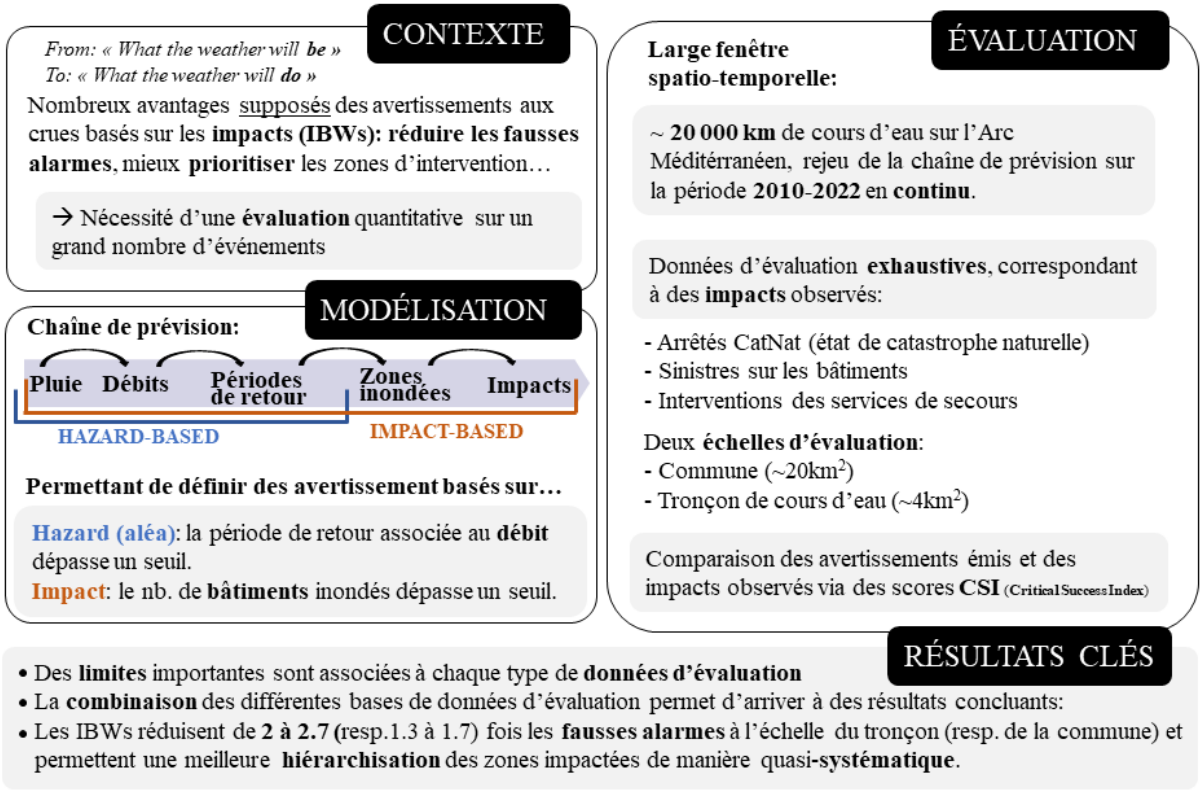


FIGURE V.1 – Résumé graphique de l'article soumis dans la revue *International Journal of Disaster Risk Reduction*.

Cet article s'intéresse à l'évaluation et la quantification de la plus-value apportée par les avertissements basés sur les impacts (IBW), par rapport aux avertissements basés sur l'aléa (HBW). Les avertissements de type HBW étudiés suivent l'approche Vigicrues Flash, qui compare les débits simulés par le modèle hydrologique à des seuils correspondant à des périodes de retour (5 et 20 ans dans la version opérationnelle de Vigicrues Flash). Les avertissements de type IBW reposent, quant à eux, sur le décompte du nombre de bâtiments situés en zone inondée, l'inondation étant estimée à partir de scénarios pré-calculés pour différentes périodes de retour (de 2 à 1000 ans). Dans les deux chaînes d'avertissements, la modélisation se limite aux rivières dont le bassin versant est supérieur à 5 km^2.

Ces avertissements HBW et IBW, définis au pas de temps journalier, ont été comparés à des données d'impacts observés, dont la collecte, l'analyse et le traitement ont constitué un travail conséquent. Pour permettre une évaluation complète, les données d'impacts recherchées devaient être continues dans le temps, exhaustives dans l'espace et disposant de la résolution la plus fine possible. Un large éventail de données a été considéré (y compris des données originales telles que celles issues des réseaux de télécommunication, des réseaux électriques ou encore des réseaux sociaux). Toutefois, les critères mentionnés plus haut ont fortement limité le volume de données mobilisables pour l'évaluation. Trois bases principales ont finalement été retenues :

(i) la liste des arrêtés CatNat déclarés sur les communes de l'arc méditerranéen, disponibles dans la base Gaspar pour l'ensemble de la période d'étude (environ 2010-2022), ainsi que dans la base iCatNat depuis 2018 (offrant davantage de détails) ;

(ii) la liste et la hiérarchie par événement des tronçons de cours d'eau ayant enregistré des sinistres selon les données de la Caisse Centrale de Réassurance (CCR) sur l'ensemble de la zone et de la période

étudiées ;
(iii) le détail des interventions des services de secours liées aux inondations pour deux départements de la zone d'étude (Gard et Alpes-Maritimes) depuis 2012.

Compte tenu de la résolution des données d'impacts collectées, l'évaluation a été menée à deux échelles spatiales : d'une part, l'échelle de la commune (soit 20 km^2 en moyenne), et d'autre part, l'échelle du tronçon de cours d'eau (soit 4 km^2 en moyenne). Les avertissements HBW et IBW émis à ces deux échelles ont été comparés aux impacts réellement observés, sur plus d'une décennie et sur l'ensemble de l'arc méditerranéen, permettant le dénombrement des "hits", "misses", "false alarms" et "correct rejections".

Les résultats ont montré qu'à taux de "hits" équivalents, les IBW réduisent significativement le nombre de fausses alarmes par rapport aux HBW. La comparaison avec les données de sinistralité a mis en évidence que le nombre de fausses alarmes à l'échelle de la commune était divisé par 1,3, et par 2 à l'échelle du tronçon de cours d'eau. La comparaison avec les interventions des services de secours a montré une réduction encore plus marquée : le nombre de fausses alarmes était divisé par 1,7 à l'échelle de la commune et par 2,7 à l'échelle du tronçon de cours d'eau.

En revanche, la comparaison aux arrêtés CatNat n'a pas permis d'aboutir aux mêmes conclusions. L'analyse des données iCatNat (fournissant des précisions sur le type d'inondation justifiant l'arrêté et sur le nombre de bâtiments endommagés) a mis en évidence certaines limites de ces données pour l'évaluation conduite ici. D'une part, des reconnaissances CatNat peuvent être demandées et accordées même en l'absence de dommages importants, le critère reposant alors uniquement sur la période de retour de l'événement, suivant une logique HBW. D'autre part, les arrêtés CatNat contenus dans la base Gaspar ne distinguent pas les phénomènes de "ruissellement", de "crue" ou de "crue torrentielle". Or, la modélisation sous-jacente aux HBW et IBW se concentre uniquement sur le débordement des rivières dont le bassin versant est supérieur à 5 km^2, et ne permet donc pas de représenter les phénomènes locaux de ruissellement. On observe ainsi de nombreux cas où des HBW sont correctement émis face à des arrêtés CatNat "ruissellement" (ces ruissellements étant souvent causés par des pluies intenses, se traduisant par des périodes de retour de débits élevées), alors que les IBW ne peuvent pas refléter les dégâts liés à ces ruissellements non modélisés.

La performance des IBW a été quantifiée via des scores CSI agrégés sur l'ensemble de la zone et de la période d'étude. Face aux données de sinistralité de la CCR, il est intéressant de noter que les IBW performent mieux lorsqu'un critère sur le nombre de bâtiments inondés est combiné à un critère sur la période de retour atteinte. Ce constat est cohérent avec l'existence d'un seuil théorique de période de retour de 10 ans pour qu'un événement soit reconnu CatNat et ouvre droit à indemnisation par les compagnies d'assurance, via la CCR. Le score CSI maximal obtenu est de 30 % à l'échelle de la commune, pour un seuil de période de retour de 20 ans combiné à un nombre de bâtiments en zone inondée supérieur à 10. Ce score peut sembler faible, mais il n'est en réalité pas surprenant au regard d'études antérieures, notamment Moncoulon et al. [2020], qui ont montré qu'une part importante des sinistres (environ 40 %) demeure inexpliquée, même avec des modélisations intégrant à la fois les inondations par débordement de rivière et par ruissellement. Les scores CSI issus de l'évaluation basée sur les interventions des services de secours sont nettement plus faibles, ce qui peut s'expliquer par plusieurs facteurs, notamment la difficulté à isoler les interventions réellement liées à un débordement de cours d'eau, et la méthode de recensement des interventions (une intervention enregistrée dans un secteur pouvant en réalité correspondre à des actions multiples dans d'autres secteurs, rattachées à un même point de départ).

Les données de sinistralité fournies par la CCR à l'échelle du tronçon de cours d'eau détaillent également la hiérarchie des zones impactées (en termes de nombre de sinistres enregistrés) pour chacun

des événements disponibles dans la base sur notre zone et période d'étude. Les avertissements HBW et IBW ont été adaptés à cette logique de classement : les premiers en fonction de la période de retour estimée, les seconds en fonction du nombre de bâtiments en zone inondée. Le score d'évaluation CSI a ainsi été adapté pour mesurer la capacité des modèles à classer les zones les plus impactées. Les résultats ont montré que les prévisions basées sur les impacts permettaient, dans la quasi-totalité des cas, une meilleure hiérarchisation des zones affectées.

En conclusion, ce chapitre a permis de démontrer et de quantifier la plus-value des IBW par rapport aux HBW, tant en matière de détection des zones impactées que de priorisation de ces zones, sur la base d'un cadre d'évaluation particulièrement robuste (large échelle spatio-temporelle). Les résultats ont également mis en évidence les limites propres aux différentes bases de données mobilisées, dans un contexte où la collecte systématique de données d'impacts post-crues demeure insuffisamment exploitée, malgré son rôle essentiel pour l'amélioration des modèles de prévision des crues. Il convient enfin de rappeler que cet article ne traite que d'un aspect des bénéfices des IBW par rapport aux HBW ; une analyse complète devrait également mobiliser les sciences sociales afin d'évaluer les bénéfices liés notamment à une meilleure communication et compréhension des prévisions par les utilisateurs.

1 Introduction

Recently, Europe and North America have been hit by catastrophic flash floods, such as those in València (Spain) in autumn 2024 and in Texas in summer 2025. These events remind us of the devastating impact that floods can have on our societies. They also reinforced a key lesson : timely and reliable flood warnings are essential to saving lives and reducing damage, when authorities act on them through preventive evacuations and other protective measures.

Warnings are traditionally characterized by their content, format, and communication channel [Kuller et al., 2021]. The content refers to the message conveyed by the warning producer to the receiver. This message may include details such as the nature of the oncoming hazard, the start and duration of the event, its location, predicted intensity, and/or anticipated consequences. It may also provide recommendations for actions and information about associated uncertainties. Examples of warning formats include text, graphics, and maps, while communication channels may involve radio, television, or SMS. The forecast, defined as the estimation of future conditions on which the warning is based, forms an integral part of the warning's content. In this study, although we adopt the term "warning" for simplicity and generality, our focus lies specifically on forecasts, or more accurately, nowcasts, as we concentrate on a short-range warning system. Strictly speaking, however, our analyses rely on hindcasts, since we replay forecasting chains over a long period for evaluation purposes, while still using the encompassing term "warnings."

1.1 Hazard-based and impact-based warnings

Originally, warnings relied solely on the characterization of the oncoming hazard, with a recurring approach of using thresholds (on selected variables to estimate the intensity of the forthcoming hazard, e.g., the expected rain / discharge / wind speed / snow depth / etc.) to launch a warning. In this paper, these types of warnings are referred to as hazard-based warnings (HBW). However, increasing efforts are now being directed towards the development of impact-based warnings (IBW). Following the recommendations of the World Meteorological Organization, these warnings aim to communicate not only what the hazard will be, but also what it will do, by explicitly considering the elements at risk, their exposure and vulnerability [WMO, 2015]. Impact-based warnings can use quantitative physical or socio-economic indicators, such as the expected number of damaged buildings or affected people, which keeps the possibility of using a threshold approach to decide the warning level.

Impact modeling generally appears less complex than hazard modeling, as its outputs are typically limited to consequences on objects, areas, and people. However, the resulting uncertainties are often considered more significant [Merz et al., 2020]. This disparity arises from our limited experience with impact modeling, such as the less systematic collection of impact data compared to hazard-related variables such as river water levels, as well as the inherent complexities that impact modeling entails, despite its apparent simplicity. Impacts are shaped by a multitude of factors, including human behavior, which can lead to unpredictable consequences.

Despite the larger uncertainties, impact-based warnings are expected to be more useful than hazard-based warnings for various reasons.

1. Firstly, they could help bridge the communication gap between scientists, warning services, and the target audience [WMO, 2015, Potter et al., 2021, Yu et al., 2022, Kuller et al., 2021, Meléndez-Landaverde et al., 2024, Potter et al., 2025]. For example, instead of providing expected water levels, an impact-based flood warning could highlight that homes, roads, and critical infrastructure in specific neighborhoods could be inundated, making the information more accessible to non-experts.

2. Secondly, they could help identify and prioritize affected areas [Merz et al., 2020, Apel et al., 2022, Le Bihan et al., 2017, Ritter et al., 2022, Saint-Martin et al., 2018]. For example, during a flood, impact-based warnings could identify areas with older infrastructure as a priority for evacuation and emergency assistance.
3. They could also raise awareness of antecedent conditions and cascading hazards, helping decision makers and communities prepare for secondary effects [Potter et al., 2021, Pascal et al., 2006, WMO, 2018]. For example, a flood warning should not only emphasize the immediate risks in the inundated areas, but also account for cumulative and cascading effects, such as the paralysis of transport networks or failures in electricity and drinking water supply systems. These indirect impacts may in turn create evacuation needs beyond the directly flooded zones.
4. They could limit the number of "false alarms", improve trust in warning systems, and dispatch available emergency resources to relevant areas [Potter et al., 2021]. For example, impact-based systems could help ensure that warnings are targeted to areas where there are significant assets or vulnerabilities, avoiding unnecessary alerts in regions with few or no assets at risk.
5. They could help save time and speed up the decision process by automatically combining hazard information with risk elements, which is traditionally performed manually by emergency managers and is prone to subjectivity [Ritter et al., 2021]. More broadly, the added value of IBW would be to facilitate decision making by automatically providing useful information on potential impacts [Taylor et al., 2018].
6. Finally, they could facilitate the transition to multihazard early warning systems by providing a common language based on impact that helps overcome barriers between disciplines that traditionally do not interact [Potter et al., 2025].

However, most of these assumptions remain unverified with robust evidence [Merz et al., 2020]. In fact, some studies even challenge the belief that IBW are inherently more useful than HBW. For example, the findings suggest that when fear is the primary motivator, the additional information provided by IBW may have a limited impact [Weyrich et al., 2020]. In this paper, our objective is to contribute to the growing body of evidence supporting the added value of IBW compared to HBW, particularly in terms of identifying and prioritizing affected areas (item 2) and reducing false alarms (item 4).

1.2 Impact-based warnings in the context of flash floods

In this paper, we focus on warning systems for flash floods. It is important to note that the warning systems for floods and flash floods may differ. For example, in France, the Vigicrues system monitors large rivers with a 24-hour anticipation window, using four warning levels determined based on the expected flood intensity, adjusted by integrating elements at risk. Smaller rivers that are not monitored by Vigicrues and are more prone to rapid flooding can be eligible for the Vigicrues Flash service [Javelle et al., 2016], which will serve as a benchmark in our study. This automated warning service offers limited anticipation, better described as "nowcasting" rather than "forecasting", and determines whether an ongoing flood is "severe" or "very severe" based on the exceedance of flood return periods (5 years or 20 years). The current operational flash-flood warning system in France is therefore entirely hazard-based. Other hazard-based systems for flash flood monitoring and forecasting have been developed, for example, in the US [Gourley et al., 2017], New Zealand [Cattoën et al., 2022], Spain [Corral et al., 2008] Switzerland [Liechti et al., 2013], and Europe with the ERIC-ERICHA system [Raynaud et al., 2015, Park et al., 2019].

The approaches to estimating the impacts of predicted floods and flash floods are diverse. For example, the ReAFFINE method [Ritter et al., 2021], developed from the ERICHA system, calculates socio-economic impacts, such as the affected population and disrupted critical infrastructure. The "Risk

Index" based on France's Vigicrues Flash system [Saint-Martin et al., 2016] evaluates exposure by ranking various assets such as transportation networks, agricultural crops or specialized buildings, based on surveys with stakeholders. This kind of approach that combines the vulnerabilities of different asset classes into a score is also used in tools like Flood-PROOFS for Africa [Alfieri et al., 2024] and the CEMADEN tool for Brazil [de Moraes, 2023]. The flood risk matrix used in the UK [Cole et al., 2016] assigns impact categories (minor, significant, or severe) directly based on the number of inundated assets, such as roads, railways, and residential properties, compared to predefined thresholds. Similarly, the Flash Flood Severity Index, a potential future warning tool for the United States [Schroeder et al., 2016], ranks floods from minor to catastrophic based on impact types (e.g., "homes swept away" indicates a catastrophic flood, while "river overflowing" suggests a minor flood). Other approaches, though less focused on operational contexts, aim to estimate specific impacts. For example, some studies focus on estimating the number of affected buildings on detailed spatial scales [Le Bihan et al., 2017], advancing the prediction of human losses [Terti et al., 2019a] and human safety [Silvestro et al., 2019], or forecasting emergency service operations [Charpentier-Noyer, 2022] and accessibility [Coles et al., 2017] and the location of road closures due to floods [Naulin et al., 2013]. In this work, we have also focused on specific impacts, which can be seen as a very simplified approach but aligns with verification constraints, as will be explained in the following.

1.3 Verification of impact-based warnings

Verification of warnings is a critical but complex process. The first challenge is to define what constitutes a correct warning. According to Golding [Golding, 2022], a "perfect" warning is "one that prevents all [negative] consequences that are avoidable". Transitioning from HBW to IBW introduces additional verification challenges - one of the main research areas of the WMO World Weather Research Program High Impact Weather project [Potter et al., 2025] - such as measuring how IBW improve the understanding of the hazard by recipients. Focusing on the forecast component of the warning, it is relatively straightforward to assess the quality of HBW (e.g., by comparing measured or reconstituted discharges to flash flood warnings based on discharge thresholds [Javelle et al., 2014]).

In contrast, IBW must be compared with impact data, based on ground observations, which is particularly challenging because of the lack of systematic, exhaustive, and continuous post-event impact data collection. This data gap has been widely documented [Potter et al., 2025, Harrison et al., 2020, 2021, 2022a, Kaltenberger et al., 2020, Meléndez-Landaverde and Sempere-Torres, 2022]. Researchers have proposed addressing this issue through partnerships and collaborations between warning services, emergency responders, and other stakeholders [Harrison et al., 2024, Potter et al., 2025], and by developing methods to improve impact data collection [Saint-Martin et al., 2018, Wyatt et al., 2024]. Even when impact data are available, they must align with the variable being evaluated [Robbins et al., 2022]. For instance, if an IBW predicts the number of flooded buildings per municipality, the validation process requires corresponding post-event observations, which may not always be collected or made available. Moreover, impact data are rarely collected specifically for the verification of IBFW systems and often need additional processing, especially when integrating multiple data sources to achieve a comprehensive perspective [Wyatt et al., 2023, Potter et al., 2025].

Biases in publicly available impact data further complicate IBW validation. For example, Threshold Bias arises when databases predominantly include high-severity events, excluding lower-impact occurrences [Wyatt et al., 2023]. This bias affects IBW validation by creating discrepancies between predicted and observed impacts within the geographic domain. To mitigate this problem, Wyatt et al. [2023] recommends the use of various data sources to obtain a more realistic representation of the impacts of natural

hazards.

In the context of floods and flash floods, the impact component of the warning chain is rarely evaluated due to the limited availability of high-quality damage data [Molinari et al., 2019, Merz et al., 2010, Meyer et al., 2013, Jongman et al., 2012, Molinari et al., 2014], particularly at detailed spatial scales [de Moel et al., 2015]. Most studies have evaluated impact predictions using observations on affected populations or aggregated economic damages on regional or national scales [Alfieri et al., 2024, Ritter et al., 2021, de Moraes, 2023]. However, some validations have been performed at finer spatial resolutions, such as county [Terti et al., 2019a], municipality [Silvestro et al., 2019, Saint-Martin, 2018], or river reach scales (approximately 2 km) [Le Bihan et al., 2017]. Furthermore, validation efforts often focus on a limited number of specific events [Alfieri et al., 2024, Ritter et al., 2021, Terti et al., 2019a, Silvestro et al., 2019, Le Bihan et al., 2017], with relatively few initiatives pursuing continuous validation efforts [de Moraes, 2023, Saint-Martin, 2018]. Examples of exhaustive impact databases suitable for the verification of flash flood forecasting tools include the SHAVE (Severe Hazards Analysis & Verification Experiment) dataset and the Storm Events Database in the United States, which are built from personal reports [Calianno et al., 2013].

This study presents a validation of flash flood IBW using various data sources, including legislative decrees attesting to natural disasters in municipalities, insurance claim records, and fire and rescue service operation records. The validation spans a continuous time window from 2010 to 2022, a vast area covering the French Mediterranean regions, and is conducted at high spatial resolutions (municipality and river reach scales). Using these data, we compare the performance of a simple impact-based model (which issues warnings based on thresholds of number of flooded buildings) to the Vigicrues Flash hazard-based model (which issues warnings based on discharge thresholds). Therefore, the originality of this study lies in the quantitative comparison of HBW and IBW based on a continuous, long-term implementation of both models over a very large spatio-temporal domain, allowing a systematic assessment of their added value, which goes beyond existing case-based, short-term or spatially limited studies. In the following, Section 2 describes the impact databases. Section 3 details the impact-based and hazard-based models, along with the validation methods. Section 4 presents and discusses the results.

2 Flood Impact data

2.1 Collecting flash flood impact data in South-Eastern France

This study focuses on the French Southeastern Mediterranean region, which includes several "hotspots" of frequent extreme discharges in Europe [Gaume et al., 2016]. A notable example of a catastrophic flash flood in this area is the 2010 flood in the Var department, which resulted in 25 deaths and caused insured losses of up to €771 million (adjusted to 2023 values) [CCR, 2024]. The study domain spans 61 main river basins in the French Mediterranean region, covering nearly 42,000 km² of drained area and 20,000 km of river network (see figure V.2a). Rivers with drainage areas smaller than 5 km² are excluded from the river network. In addition, the Rhône and Durance rivers are omitted from the study domain because of their low susceptibility to flash floods and their significant regulation by artificial structures, such as dams.

In this study, we compare impact-based warnings and hazard-based warnings issued at two spatial scales : the river reach scale - illustrated in Figure V.2 b (with a mean sub-catchment area of 4 km2) - and the municipality scale - illustrated in Figure V.2 c (with a mean municipality area of 20 km²). The high resolution requirements of this analysis preclude the use of widely recognized international impact databases, such as EM-DAT, as such granularity can only be achieved at the infra national level. Impacts

are typically classified as direct, indirect, tangible, or intangible [Merz et al., 2010]. Here, we focus exclusively on impact data that can be directly compared with simulated direct and tangible impacts, such as damage to buildings or roads. Furthermore, our study is limited to the flash flood hazard affecting rivers with drainage areas greater than 5 km^2. Therefore, the impact data we want to collect should ideally only represent the damage caused by these rivers (thus excluding smaller rivers or surface flooding). This constitutes an important limitation, as in the case of flash floods, surface flooding can be significant and can cause important damage.

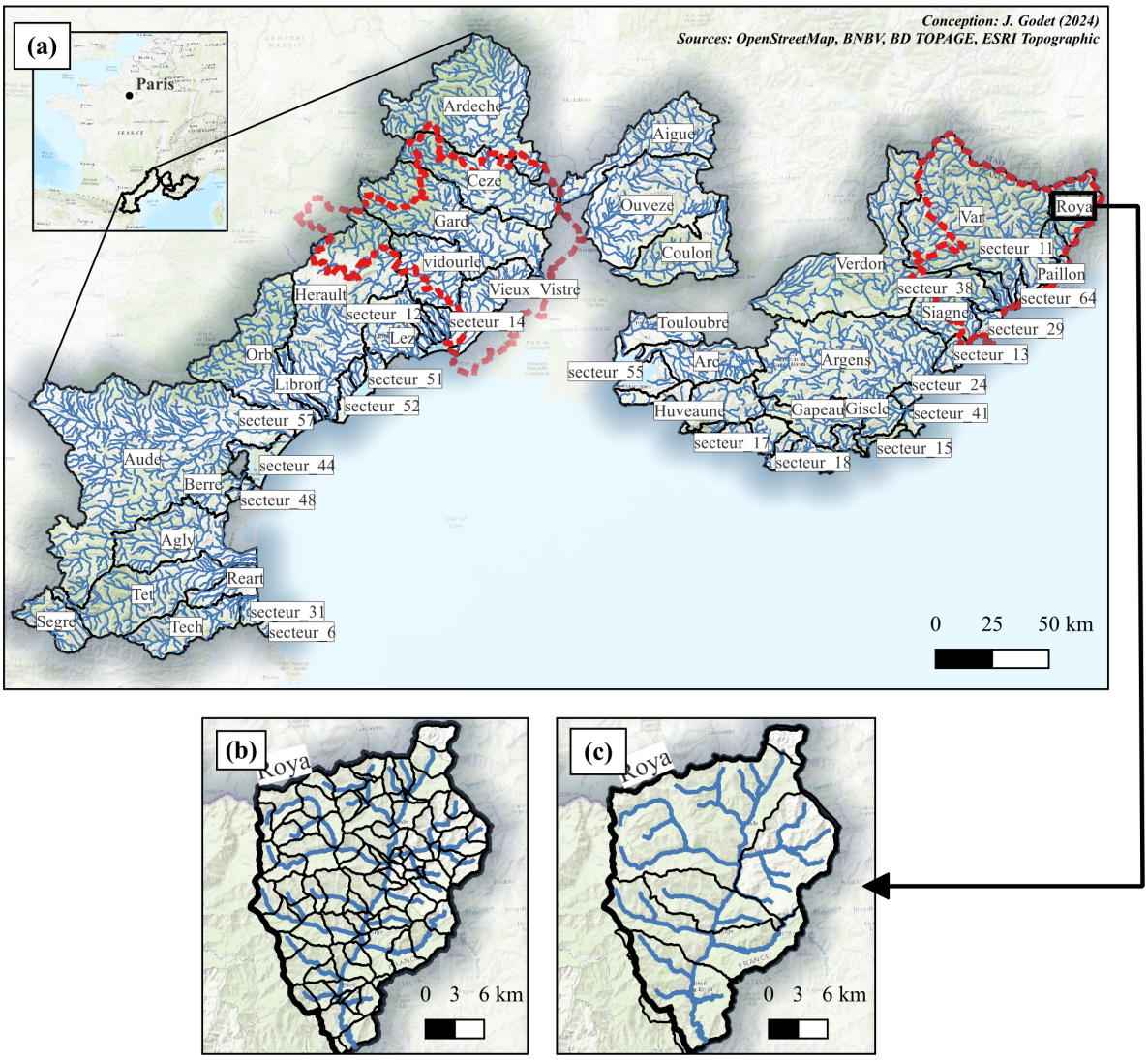


FIGURE V.2 – (a) Presentation of the studied river network and of the main river basins on which the CatNat decrees and the insurance database were collected. The red dotted lines represent the areas on which the fire and rescue service operations records were collected. (b) Illustration of the river reach scale on the Roya catchment. (c) Illustration of the municipality scale on the Roya catchment.

The impact data collection has faced the first challenge of the French administrative mille-feuille, characterized by overlapping and sometimes conflicting responsibilities across administrative and environmental governance levels. Traditional administrative bodies, such as departments, offer relatively stable and historically grounded frameworks for data collection. On the other hand, environmental governance structures, such as river basin authorities, provide a territorially coherent approach but add complexity. This fragmented landscape results in uneven data availability, influenced by local priorities, budgets, and institutional capacities. Consequently, identifying key actors and harmonizing efforts remain

critical challenges.

Although traditional institutional actors are key to impact data collection, exploring beyond these channels can also be valuable. The rise of voluntary geographic information (VGI) and social data, driven by the adoption of smartphones, allows two-way communication between crisis managers and the public. Platforms like Twitter and Waze provide useful tools for detecting and validating flood impacts, particularly in areas overlooked [Lowrie et al., 2022, de Bruijn et al., 2019]. Despite biases such as urban overrepresentation, these data sources complement institutional data sets and improve flood monitoring and crisis management.

Given these requirements and constraints, we identified several categories of flood impact data. On one hand, comprehensive datasets provide continuous coverage of flood impacts and are therefore particularly valuable for validation : they make it possible to check that the warning system is triggered during damaging events, while ensuring that it does not overreact in the absence of impacts. Among these, datasets related to building damages (CatNat decrees, CCR insurance data, and fire and rescue service operations SDIS ; which will be explicited in sections 2.2, 2.3 and 2.4, respectively) are especially suitable, as the number of affected buildings represents a simple but relevant indicator of flood impacts that can be modeled with some simplifications (Table V.1). On the other hand, other impact datasets cover assets such as roads, electricity, or telecommunication networks (Tables D.1 and 1). These datasets could in principle support continuous validation, but are more difficult to exploit here since their indicators are not directly comparable with building-related damages. In addition, some datasets provide detailed descriptions of specific flood events (Tables D.2 and 2), which are highly valuable for event-scale analyses but not adapted to the continuous validation framework adopted in this study.

Based on these considerations, we selected three building-related impact datasets for this study : CatNat decrees, CCR insurance records, and fire and rescue service operation records. In the following subsections, each of these databases is described in detail.

Tableau V.1 – Summary of impact databases used in this study

Database	**Description**	**Access and Coverage**	**Impact type**	**Spatial resolution**
CatNat / iCatNat databases	Municipal declarations of natural disaster states after significant flood events (T >10 years).	Open-access, full spatiotemporal coverage	Building damage	Polygon data (municipality scale)
Insurance records from CCR database	Locations and ranking of river reaches with insured building damage within CatNat declared events, selected under the 1000-y scenario of the studied river network.	Private data, full spatio-temporal coverage	Building damage	Polygon data (river reach scale)
Fire and rescue service operations records	fire and rescue services operations selected under the 1000y scenario of the studied river network	Restricted data, accessed on 22% of study area, Jan. 2012 - Sept. 2023	Reconnaissance, evacuations, pumping operations (mostly near buildings)	Point data

2.2 CatNat decrees

2.2.1 Definition and limitations

In France, CatNat decrees (abbreviations for "Catastrophes Naturelles" or "Natural Disasters") are official declarations issued by the government to recognize specific events as natural disasters. These decrees facilitate compensation for damages under insurance policies, based on a system of national solidarity that mobilizes collective resources to support affected individuals and communities. Natural hazards such as floods, landslides, earthquakes, and droughts can qualify. When significant damage to property or infrastructure occurs, municipalities can request the national government to classify the event as a natural disaster. If the event meets required criteria, such as a theoretical return period of at least 10 years, an official decree is issued. This return period can refer either to rainfall accumulations over durations ranging from one hour to several days, or to peak discharges. This recognition enables those affected to access financial compensation and additional support provided by the government.

The open access database GASPAR provides historical records of CatNat decrees and the corresponding affected municipalities since 1982. Previous studies [Douvinet, 2006, Douvinet and Vinet, 2012] have evaluated the use of these decrees for flood event inventories, highlighting certain limitations : (i) the criterion requiring a 10-year return period can be inconsistently applied, (ii) political influence can affect decisions, and (iii) it is difficult to distinguish between damage caused by river floods and pluvial floods. Despite these drawbacks, CatNat decrees from the GASPAR database have already been utilized to validate forecast chains [Saint-Martin, 2018]. For this study, we extracted all 3,588 flood decrees issued between September 2009 and September 2023 for 1,985 municipalities in the French Mediterranean region. Therefore, the spatial resolution of this database is the municipality scale (figure V.2c).

2.2.2 The iCatNat platform : some extra information to be treated with caution

Since 2018, the CatNat request submission process has been digitalized through the iCatNat platform. Although the iCatNat database is not publicly accessible, an access to it was provided by the Ministry of Interior for the purpose of this study. Compared to the GASPAR database, iCatNat offers more detailed information for three main reasons : it includes both approved and rejected applications, specifies the type of flooding hazard (A1 : river flooding ; A2 : surface flooding ; B : torrential flooding), and records the number of damaged buildings per municipality. This is particularly interesting for our study, which only considers damage related to river flash floods on catchments larger than 5 km^2. However, it also has its limitations : the reported number of damaged buildings is only an estimate, misclassification between hazard types (for example, labeling intermittent river floods as surface floods) can occur, and the database only covers events from 2018 onward.

Analysis of iCatNat flood data for 2018–2023 shows that 40% of the triggering hazards were surface flooding, 10% river or torrential flooding, and 50% mixed events involving surface and river/torrential flooding. However, characterizing these hazards can be problematic, as the initial classification is made by municipal departments who are not hydrometeorological experts, and who may therefore confuse different flood types. Moreover, for multi-hazard events, the return period used to assess eligibility may be based on rainfall or discharge without explicit distinction, which can further blur the classification.

It is important to note that the classification of flood types is primarily intended to refine the CatNat request process and ensure that appropriate experts are assigned to evaluate each case. This classification is not designed to achieve perfect scientific precision. Therefore, we chose to use the CatNat decrees issued between 2009 and 2023, which are summarized in the GASPAR database, where the types of flood are not indicated. The iCatNat database will be used as supplementary data to provide insight into the

gaps and potential biases observed in the validation results based on the CatNat database. An illustration of the CatNat data is provided in Figure V.3a.

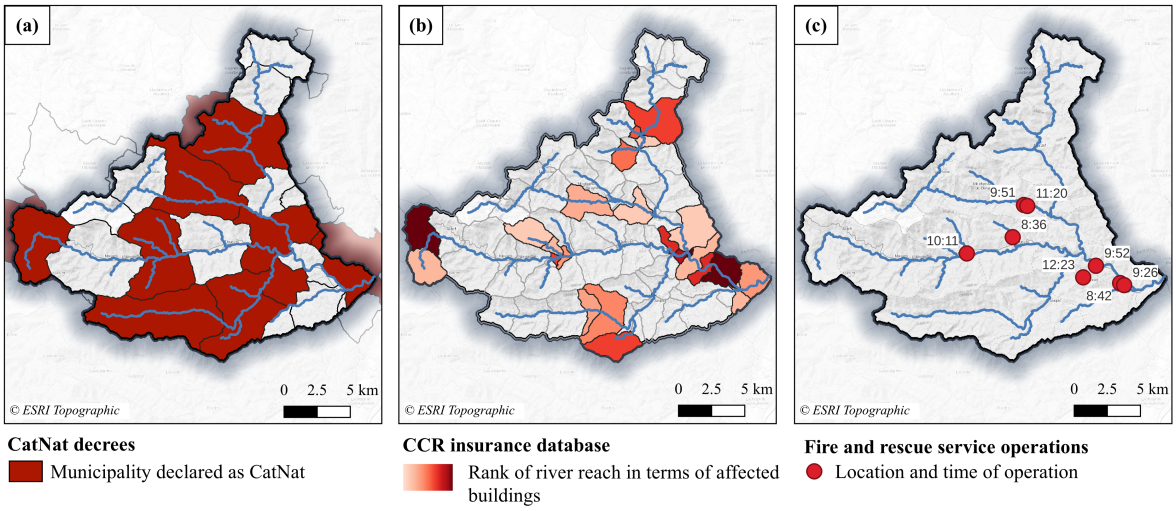


FIGURE V.3 – Illustration of the three impact databases (as originally collected) applied to a hypothetical flood event of the Ardèche River. However, in the verification process, the CCR insurance database was also aggregated at the municipality scale, and the fire and rescue service operations were aggregated at both the river reach and municipality scales.

2.3 Insurance data

2.3.1 Content and limitations of the CCR insurance database

When an event is classified as a CatNat event, a state-owned reinsurance company, CCR ("Caisse Centrale de Réassurance"), provides state-guaranteed coverage to insurance companies. Consequently, CCR has developed an extensive insurance claim database as part of its reinsurance agreements with its clients [Moncoulon et al., 2014]. This database, which covers the entire French territory, is considered reliable for data from 1997 onward and is regarded as the most comprehensive resource on flood-related impacts in France [Le Bihan et al., 2017]. However, it has certain limitations, including : (i) Occasional inaccuracies in geocoding ; (ii) Non-systematic reporting of claims within the database ; (iii) Errors in data entry during the collation process, based on information provided by insurance companies. For example, the address associated with an insurance policy may not always correspond to the location of the affected property, as is often the case with holiday homes.

A study using the same dataset calculated the observed claim-to-policy ratio within the 1000-year flood extent for two major flood events [Le Bihan et al., 2017]. This “1000-year flood extent” corresponds to the area that would be inundated by an event of extremely low probability (return period of 1000 years), and in practice provides a very large envelope of exposure. The calculated ratios, 0.37 and 0.43, were well below 1, which may seem surprising at first glance. This discrepancy may be explained by claims without compensation, the absence of claims, apartments located on the upper floor, or the presence of protective features such as elevated basements or local flood defenses. Another difficulty in interpreting claim data is that it merges damage caused by phenomena other than river flooding, such as small unmodeled watercourses, surface flooding, sewer overflows, or even roof infiltration claims that may inadvertently slip into the database.

As an illustration, 38% of the flood damages reported in the Mediterranean region between 1995 and 2019 were located outside the flooded areas modeled by the CCR flood model [Moncoulon et al.,

2020], even though this model accounts for both river and surface flooding. This result highlights a broader point : whatever the impact model used, large discrepancies must be expected between modeled hazard extents and observed damages. Only a fraction of the properties located in high-exposure zones are actually significantly damaged, and conversely, more than half of the observed losses during major floods are not directly attributable to river overflow.

2.3.2 The insurance data collected for this study

The data used for this study was provided by CCR, and includes data aggregated on the river reach scale (Figure V.2b) to maintain confidentiality. Before this aggregation, CCR filtered the claim data by selecting only those related to residential, commercial, industrial, and agricultural buildings that could be geocoded with the street number. In addition, they included only claims located within the 1,000-year flood extent (i.e. claims that may correspond to river flooding). This 1000-year flood map is described by Godet et al. [2025a]. This approach minimized the likelihood of comparing our warnings with flood impacts caused by unmodeled hazards, such as surface flooding or floods from rivers with drainage areas smaller than 5 km^2. However, some residual instances of such impacts may still occur within the 1000-year flood extent.

For each event, CCR counted the number of claims per river reach and provided the ranking of river reaches based on the observed damage (number of claims), as shown in Figure V.3b. This data preparation method ensured both confidentiality and alignment with our objectives : focusing primarily on river flooding and enabling the evaluation of our warnings to detect and classify affected areas.

CCR extracted data based on a continuous time window from January 2010 to December 2022, resulting in 402 events distributed in the 61 basins shown in Figure V.2a. An illustration of the CCR insurance data is provided in Figure V.3b. To address potential errors in the insurance database, such as incorrect reporting of policy addresses, event dates, or cases where the policyholder's address does not match the location of the damage, we implemented an additional verification step. This involved cross-referencing the CCR insurance database with the CatNat database described in the previous section. Since all claims in the CCR database should correspond to a recognized CatNat event, this filtering process allowed us to identify and exclude discrepancies. As a result, 20% of the events were removed, highlighting the importance of considering these limitations.In particular, the excluded events were highly localized, with 75% involving only a single river reach and a maximum of four river reaches affected. Such cases may reflect isolated errors (e.g. geolocation inaccuracies, misreported addresses such as holiday homes) rather than actual flood events, which could explain why they are not identified in the CatNat database.

2.4 Fire and rescue service operations records

2.4.1 Presentation of the role of fire and rescue services in flood management and resulting datasets

In France, fire and rescue services are responsible for providing emergency assistance to population facing natural hazards. They generally adopt a phased approach to managing flood events, adapting their actions to the event's progression. Before the flood, efforts focus on organizing preventive evacuations and pre-positioning resources. During the flood, the priority shifts to immediate rescues and securing affected areas, with specialized teams deployed to assist populations and assess the situation. After the flood, recovery operations take precedence, including pumping water, clearing debris, and restoring safe conditions in affected areas. This structured response ensures effective management of each flood phase.

In general, fire and rescue service operations are archived in departmental databases, which typically include timestamps and locations for each intervention. Statistical reports on fire and rescue services activities are compiled annually at the national level; in 2022 alone, 4,968,500 interventions were recorded. Among these, 30,663 were related to flooding, according to the French Ministry of Interior and Overseas Territories. However, these national reports conceal the disparities in the way such interventions are documented. The lack of standardization in this process poses a significant challenge, as impacts must be recorded as rigorously and homogeneously as possible. However, there is a clear trend toward national standardization, which should help overcome these current limitations.

To our knowledge, fire and rescue service operations records have never been used for flood warning verification. However, they have been highlighted as promising datasets in several studies [Saint-Martin, 2018, Le Bihan, 2016], have inspired flood impact modeling [Charpentier-Noyer, 2022], and similar data have been used in other countries for purposes such as analyzing the relationship between rainfall events that trigger flash floods and the impacts of these floods as recorded by fire and rescue service operations [Papagiannaki et al., 2015].

2.4.2 Description of collected fire and rescue service operations records, benefits and limitations

Records of fire and rescue service operations are primarily collected at the departmental level. The study area presented in Figure V.2a covers 14 different departmental fire and rescue services (SDIS), all of which were systematically contacted. A high response rate was observed, with 10 of 14 services responding to emails or calls. Among these, three provided operational databases, six expressed the potential for future collaboration, and one rejected the request. Among the three databases collected, one was excluded because it only provided geocoded operations to the municipality center. The analysis focuses on the remaining two databases, which include precise coordinates for each operation. These correspond to the Gard department (on the left side of the map) and the Alpes-Maritimes department (on the right side of the map), referred to as "SDIS30" and "SDIS06," respectively, with SDIS standing for "departmental fire and rescue service" and "30" and "06" representing the department codes.

The SDIS06 database, which dates back to December 2011, includes operations labeled "Flooding of land / public roadway," "Flooding of premises due to adverse weather," "Flooding of roadway due to adverse weather", "Flood assessment following adverse weather" and "Important flooding" in its comprehensive operation records. After removing entries with significant missing data (for example, missing coordinates), 4,180 operations remained. Operations with keywords such as "watering," "leak," "hydrant," "tap," "pipeline," or "sprinkler" were excluded, reducing the database to 2,991 operations. However, despite these precautions, some records may still represent floodings caused by non-meteorological events. To address this, a final filtering step was applied : Only operations located within the 1000-year flood extent were retained (as for the CCR insurance data). This resulted in a final dataset of 741 operations.

The SDIS30 database covers operations since May 2012, with an initial count of 5,574 records labeled as "Non-urgent actions following a weather event" or "Rescues or securing operations following a weather event." These labels may cover operations caused by weather hazards other than floods. To mitigate this issue, the same filtering step was applied for the SDIS06 database, retaining only operations within the 1000-year flood extent. This resulted in a final dataset of 1,693 operations. The SDIS30 and SDIS06 databases were aggregated on the river reach scale after selection to maintain confidentiality.

Although the operation types differ between these two databases, they share several features. Both include precise timestamps (accurate to the minute) for when the call was received, when the rescue team departed, and when the operation ended. They also detail the types of resources and vehicles deployed. This information is particularly valuable, as one of the main limitations of fire and rescue

service operation records is that some interventions can be masked. For instance, a single operation in the records can correspond to deployments across multiple areas, particularly when the coordinates correspond to an "advanced command post," where strategic decisions are made directly on the ground. Such cases can often be identified by the large number of vehicles assigned to a single operation.

Despite uncertainties about the origin of some operations (e.g., whether they were truly flood-related, and specifically related to rivers with drainage areas larger than 5 km^2), as well as the potential masking of multiple interventions under a single record, this dataset was deemed valuable. These limitations are consistent in both hazard-based and impact-based approaches, ensuring that they do not bias the comparison. Unlike the CatNat and CCR databases, fire and rescue service records are not subject to the threshold bias associated with the 10-year return period, which makes them particularly useful for analyzing smaller flood events. In addition, their precise spatial and temporal information uniquely allows for validation of the modeled timing of events. Finally, while some interventions occur outside buildings (e.g., rescuing individuals from vehicles), we adopted the simplifying assumption that interventions occur where people are located, and people are mainly located in buildings. This assumption is supported by our data, where most interventions during flood events were building-related (e.g., reconnaissance, evacuations, pumping). An illustration of the fire and rescue service databases is provided in Figure V.3c.

3 Nowcasting chains and validation methods

In this section, we present the flash flood warning services under comparison. The first is the hazard-based (HB) warning service currently operational in France, which will be referred to as the HB model. This model serves as the benchmark for assessing the added value of a simple impact-based (IB) warning service, referred to as the IB model. The "nowcasts" produced by these two models were simulated over a continuous time window from September 2009 to September 2023. Figure V.4 provides an illustration of the two modeling chains.

3.1 Hazard-based nowcasting chain

The HB model that serves as the benchmark in this study is the Vigicrues flash monitoring system for flash floods in France. As it has been thoroughly described and assessed in a previous study [Godet et al., 2025b], only a brief summary is provided here. Vigicrues Flash uses a reservoir-based rainfall-runoff model developed by INRAE, that estimates flood severity (in terms of return periods) every 15 minutes [Piotte et al., 2020]. The hydrological model is highly simplified, relying on a structure with two reservoirs (production and transfer) and an "instantaneous" routing scheme. It is fully distributed and operates at a resolution of 1 km^2. Currently, Vigicrues Flash is forced with observed radar rainfall estimates rather than rainfall forecasts. Consequently, the anticipation time cannot exceed the concentration time of the catchment, and highly depends on the location of heavy rainfall. For this reason, the term "nowcasts" is used in this study instead of "forecasts". Furthermore, in this work, the model is forced with the ANTILOPE reanalysis rainfall product [Laurentin, 2008], which could yield slightly more optimistic results compared to the operational system, where real-time ANTILOPE data are utilized.

In Vigicrues Flash, discharge outputs from the hydrological model are converted into return periods using a Gumbel fitting method applied to long discharge series generated by the model. This is considered sufficient for estimating discharge quantiles for low return periods (Vigicrues Flash does not estimate beyond a 20-year return period). However, since this study considers return periods ranging from 2 to 1000 years, we applied the SHYREG method [Arnaud et al., 2014] for larger quantiles (see details in [Godet et al., 2025b]).

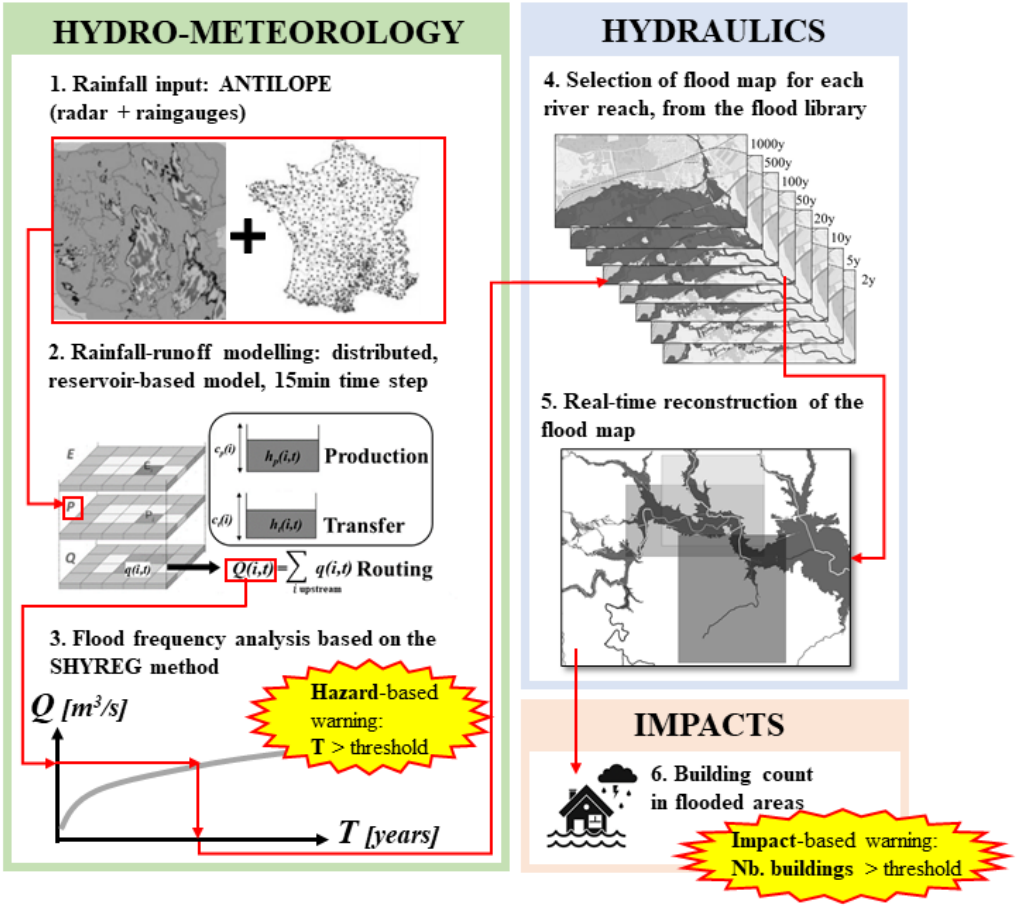


FIGURE V.4 – Illustration of the structure of the HB and IB nowcasting chains under comparison

Currently, only a selected river network is eligible for Vigicrues Flash. However, the results of [Godet et al., 2025b] indicated that extending the service to the entire river network did not lead to a systematic decrease in performance. Based on these findings, we applied the model to the full river network presented in Figure V.2. Since the hydrological model used in Vigicrues Flash is gridded, discharge outputs and return periods are computed for each pixel in the study area. To integrate these results with the vectorized river network shown in Figure V.2, points from the river network are assigned to the closest grid pixel of the hydrological model based on the similarity of their upstream catchment contours (see [Godet et al., 2024] for a detailed description of the method).

Finally, in the Vigicrues Flash service, results are aggregated at the municipality scale : a warning is issued if the discharge return period for any river reach within the municipality exceeds a predefined threshold. Our HB model replicates this approach, but also evaluates the results directly on the river reach scale.

3.2 Impact-based nowcasting chain

To modify the previously described HB model into an IB model, the first challenge encountered is flood mapping. Impact-based approaches require hydraulic calculations to assess the extent of floods and exposed assets, which can be computationally demanding [Mosimann et al., 2024]. Although fast and simple methods such as "bathtub modeling" have been used in the past, they are increasingly discouraged [Sanders et al., 2024], and physics-based approaches are now recommended - such as hydraulic models solving the shallow water equations (SWE) [Hocini et al., 2021] - especially as ongoing research shows promising results with near-real-time hydrodynamic modeling [Najafi et al., 2024]. However, for large

study areas such as the one considered in this work, direct hydraulic calculations remain infeasible, and flood libraries are traditionally used in such cases [Le Bihan et al., 2017, Dottori et al., 2017, Alfieri et al., 2024, Mosimann et al., 2024]. AI-based approaches are also being explored for flood mapping, but their implementation in ungauged basins remains challenging due to data scarcity [Nevo et al., 2022].

In this study, we adopt the traditional flood library approach. The flood library used in this study is documented in [Godet et al., 2025a] and consists of eight flood maps, each corresponding to a single frequency scenario for return periods of 2, 5, 10, 20, 50, 100, 500, and 1000 years, computed using a 2D SWE model at 5m resolution. The real-time reconstruction of flood extents based on the flood library treats each river reach separately. For each river reach, the return period is determined from the HB model, and the hydraulic results from the closest exceeded return period are retrieved from the flood library. This consists of a raster of water level values computed for the given river reach, with a slight extension beyond its limits (2 km on each side) to account for potential backwater effects at confluences. In areas where flood extents from different river reaches overlap, the maximum water level value is selected. This process can be pre-calculated for all possible scenarios, allowing immediate application in real-time forecasting. In an operational context, it would produce a flood map every 15 minutes. In this study, however, flood maps were aggregated at the daily scale by selecting, for each river reach, the maximum simulated discharge of the day.

From the previous stage, we obtained flood maps (including water levels and water velocities) for each day when the 2-year return period discharge was crossed anywhere in the computation domain. This opens up numerous possibilities for impact modeling. As discussed in the Introduction, flood impact modeling encompasses a wide range of approaches, including the estimation of flooded buildings [Le Bihan et al., 2017], economic damage and affected populations [Ritter et al., 2021, Dottori et al., 2017, Alfieri et al., 2024], human losses [Terti et al., 2019a], emergency service operations [Charpentier-Noyer et al., 2025], and road closures [Naulin et al., 2013]. When the corresponding observed impact data are available, these indicators can be directly compared to observations to validate forecasts. However, in some cases, the output of the impact model is synthesized into impact scores or categorical impact levels.

In this study, we have adopted a highly simplified approach to impact modeling by counting the number of buildings in flooded areas. Despite its simplicity, we argue that this flood impact indicator aligns well with the three comprehensive impact databases we collected - CatNat decrees, insurance records, and fire and rescue service operation records. For CatNat decrees and insurance records, which primarily focus on building damage, this comparison is straightforward. Although fire and rescue service operation records do not always refer to buildings - since interventions can also involve rescuing people from vehicles, for example - our analysis of available samples indicated that the majority of interventions recorded during flood events were related to buildings. This simplification enables the evaluation of a single flood impact model using three different validation data sets.

The building shapefile was extracted from BD-TOPO (`https://geoservices.ign.fr/bdtopo`, last accessed 10/02/2025) and pre-processed to group residential, agricultural, industrial and commercial buildings. We selected the shapefile corresponding to the year 2020 to ensure consistency with the Digital Elevation Model used for hydraulic modeling. Some verifications confirmed that the number of buildings within the 1,000-year flood extent remained relatively stable throughout the simulation period. The BD-TOPO building shapefile was rasterized to match the 5 m spatial resolution of the flood maps. A building was considered flooded if at least one of its "building pixels" overlapped a "flood pixel" where the water level exceeded 15 cm, roughly corresponding to a typical door threshold. The number of flooded buildings was then counted within each river reach to match the spatial resolution of the insurance records data and within each municipality to align with the CatNat data. This aggregation also ensured confidentiality in

the processing of the fire and rescue service operations data.

3.3 Validation approach

The comparison between the aforementioned models and the validation datasets allows us to assess the models' ability to (1) detect impacted areas and (2) rank impacted areas, as detailed in this section. This assessment is designed for users, such as emergency managers, whose objective is to determine where to allocate resources for effective crisis management and how to prioritize the distribution of limited resources.

3.3.1 Assessing the ability of the models to identify impacted areas

In this study, the nowcasts provided by the models are limited to the deterministic detection of threshold crossings. The contingency table is a common validation method when using a single threshold, and is illustrated in Table V.2. This validation based on contingency tables has been used here to assess how the different models, based on hazards, impacts and hybrid models, perform in detecting (and anticipating) damaged areas from a binary perspective : "impacted" / "not impacted".

A wide range of thresholds have been tested in this study for each model. For the HB model, available return periods range from 2 to 1000 years, and for the IB model, we focus in particular on thresholds greater than one flooded building. This choice is motivated by several factors : buildings located within flooded areas may experience limited or no damage due to local protective measures, and even when damage does occur, it may not result in an insurance claim, for instance, due to high deductibles or other reasons that discourage reporting.

Tableau V.2 – Illustration of a contingency table built for the evaluation of a nowcasting chain based on impact data. a, b, c, and d correspond to the number of occurrences in each category.

Warning triggered	**Impact observed**	
	Yes	**No**
Yes	HIT (a)	FALSE ALARM (b)
No	MISS (c)	CORRECT REJECTION (d)

The contingency tables were built in the following manner : if a warning was triggered in the same area where an impact was observed, within a +/- 2-day window, the HIT counter was incremented by one. In contrast, if an impact was recorded in an area where no warning was issued within the +/- 2-day window, a MISS was counted. Similarly, a FALSE ALARM was recorded when a warning was issued without a corresponding observed impact.

In this validation exercise, the "Area" refers to two different spatial resolutions : the municipality scale (Figure V.2c) or the river reach scale (Figure V.2b). Evaluating results at both scales allows for a comparative assessment of model performance at different resolutions. The results are expected to be more reliable on the municipal scale, as the river reach scale imposes stricter requirements and demands greater precision in impact localization.

"Impact" refers to the selected validation datasets : A municipality where a CatNat decree was issued for a flood event, defined by its start and end dates ; A river reach or municipality where at least one insurance claim - selected under the 1000-year scenario of our flood model — was reported in the CCR insurance records during a flood event defined by start and end dates ; A river reach or municipality where at least one fire and rescue service operation — selected under the 1000-year scenario of our flood model — was recorded in the fire and rescue service database, with a precisely defined operation time.

Based on the contingency tables, traditional contingency scores were calculated, such as the CSI (Critical Success Index) score : $CSI = \frac{a}{a+b+c}$, which ranges from 0 to 1, with a value of 1 indicating a perfect forecast. POD (Probability Of Detection or Hit rate) : $POD = \frac{a}{a+c}$, it focuses on the set of observed events and equals 0 when all observed events are missed, and 1 when all observed events are correctly forecasted. The FAR (False Alarm Ratio) : $FAR = \frac{b}{a+b}$, it focuses on the set of issued warnings and equals 0 when no false alarms are issued, and 1 when all issued warnings correspond to false alarms. An illustration of the validation process is provided in figure V.5a.

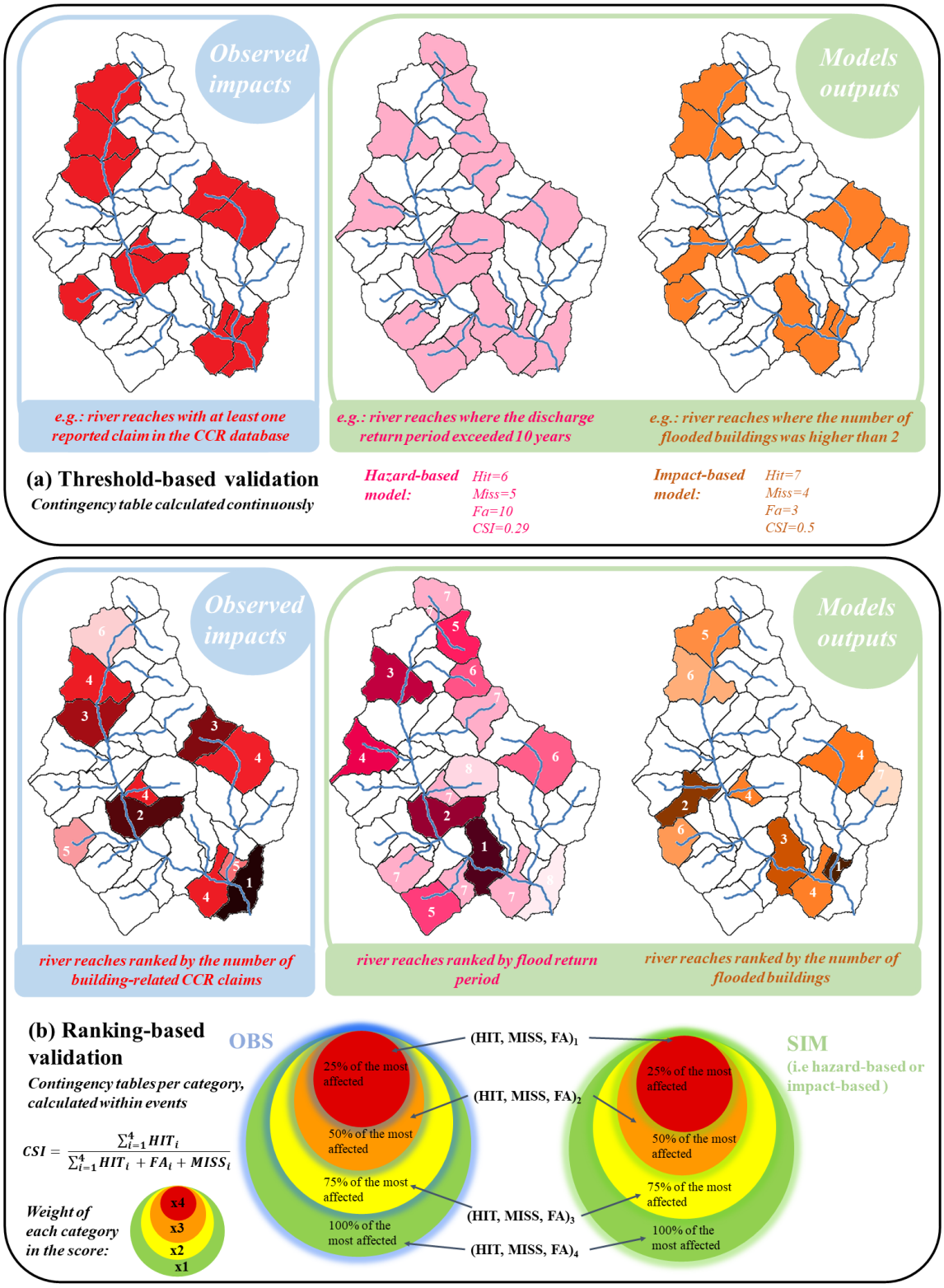


FIGURE V.5 – Illustration of (a) the threshold-based validation and (b) the ranking-based validation

3.3.2 Assessing the ability of the models to rank impacted areas

Although all three selected validation databases allow for threshold-based validation, the CCR insurance records originally provided ranked lists of affected areas based on the numbers of reported claims for each event. This observed ranking, available on the river reach scale, can be reproduced using the different models : in the HB model, areas are ranked by flood return period ; in the FB model, by the percentage of flooded area within the 1000-year flood scenario ; and in the IB model, by the number of flooded buildings. This ranking-based validation is particularly valuable, as it extends beyond assessing the models' ability to detect impacted areas, offering insight into their ability to rank them appropriately. This distinction is critical in a real-world context, where emergency resources are limited and prioritization is essential for effective crisis management.

An illustration of this ranking-based validation process is provided in Figure V.5b. To assess the ability of the models to rank impacted areas, four contingency tables are constructed. They correspond to different percentile-based subsets of the most affected river reaches, defined according to the model-specific impact indicator. Specifically, the 25%, 50%, and 75% most affected reaches correspond to those whose simulated impact values belong to the upper 25%, 50%, and 75% of the distribution, respectively, while the 100% category includes all impacted reaches. The global agreement between observed and simulated rankings, is then summarized based on an aggregated CSI score which sums these contingency tables. This score applies different weights to each category : the 25% most affected river reaches are weighted x4, the next category 25 to 50% is weighted x3, category 50 to 75% is weighted x2, and category 75 to 100% is weighted x1. This approach ensures greater emphasis on the most severely affected areas. However, this score does not fully account for near misses, that is, cases where the observed and predicted categories are adjacent rather than identical, which potentially makes the validation overly stringent. This limitation could be addressed in future work by adopting alternative scoring methods, such as the Gerrity score [Gerrity, 1992], which better accounts for the ordinal nature of the categories, although with a more complex formulation and a lower degree of comparability with the metrics used so far.

Figure V.9 shows examples of the CSI scores obtained with the proposed metric, based on two real flood events. In the case of the October 2015 flood in the Siagne catchment (Figure V.9a), the IB model achieves a CSI of 0.63, reflecting a fairly good alignment between the impact rankings observed and predicted. Although this value may seem a bit low given the visual similarity between the two maps, the aim here is less to achieve a perfect score and more to highlight the added value of the IB model compared to the HB approach. In that sense, the relative difference between the two models is what matters most.

This ranking-based validation using the CCR insurance database can only be performed on the river reach scale, as the rankings cannot be meaningfully aggregated at the municipality level. It is also not applicable to the CatNat database, which provides only binary impact information. Although in theory fire and rescue service operation records could be used to rank river reaches based on the number of recorded operations, this approach was ultimately discarded. As discussed previously in the limitations of the dataset, a single recorded operation can correspond to deployments in multiple areas, making it unreliable for direct ranking comparisons.

4 Results and discussions

This section presents the results of comparing the hindcasts produced by the nowcasting chains (sections 3.1 and 3.2) with the flood impact data (section 2) across all municipalities and river reaches in the covered Mediterranean region (figure V.2). Sections 4.1, 4.2, 4.3, 4.4, and 4.5 present and discuss the results of the threshold-based validation approach (section 3.3.1) spanning the continuous 2010–2022

period, using all three validation datasets : CatNat decrees, CCR insurance claims, and fire and rescue service operations. Section 4.6 presents and discusses the results obtained at the event scale for both the threshold- and ranking-based validation approaches, using only the insurance claims dataset.

4.1 Comparison between hazard-based and impact-based models : significant reduction of false alarms

Figure V.6 presents the results obtained from the threshold-based validation of the HB and IB nowcasting chains, using the insurance claim impact data. This figure illustrates the evolution of CSI scores with warning thresholds (based on return periods for the Hazard-based (HB) model or on the number of flooded buildings for the Impact-based (IB) model). The CSI curves exhibit a 'hill' shape due to a trade-off effect : as the warning threshold increases, the number of hits decreases (while the number of misses, although not shown, symmetrically increases), but the number of false alarms decreases as well. The CSI reaches its maximum where this trade-off is balanced. Although a higher maximum CSI is achieved with the IB model compared to the HB model, it is difficult to draw broader conclusions from this figure alone. In particular, the threshold that optimizes the CSI depends on the model, the spatial scale, and the validation data set considered.

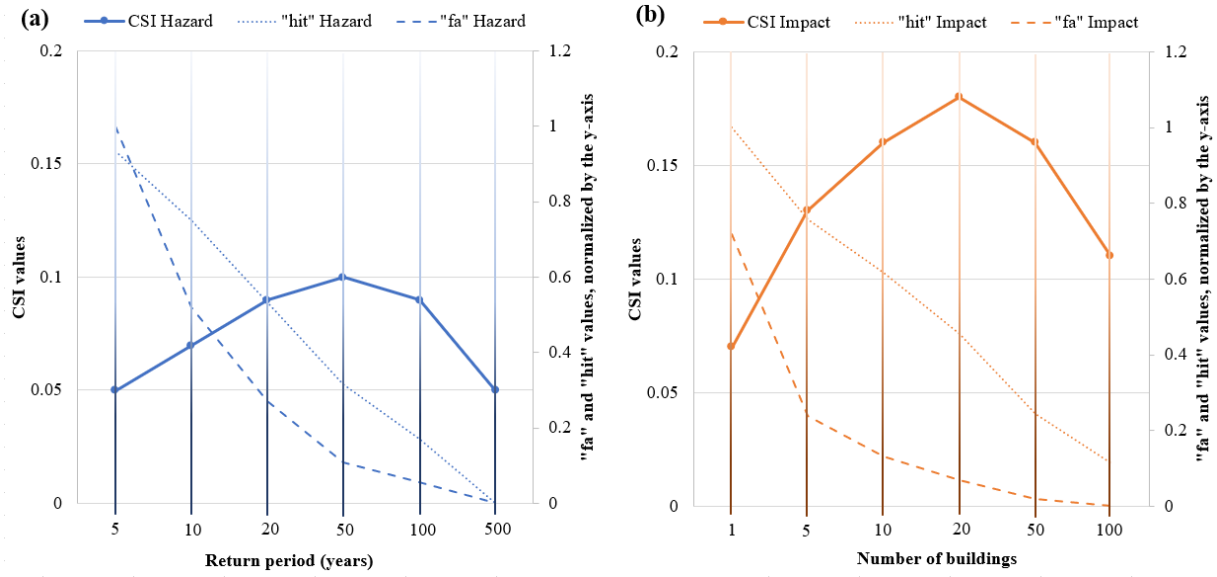


FIGURE V.6 – Results of the threshold-based validation using insurance claim data at the river reach scale, showing the evolution of CSI, false alarms, and hits with increasing warning thresholds for (a) the hazard-based model and (b) the impact-based model.

To facilitate the comparison between the different nowcasting chains, Table V.3 illustrates the scores obtained with thresholds leading to a similar hit rate (or Probability of Detection, POD), for both HB and IB models. This table helps to quantify the reduction in false alarms when switching from an HB to an IB model, under a common POD target of 60 to 70%, which is commonly considered indicative of satisfactory detection performance in binary flood forecasting evaluations. However, this POD target could not be reached when validating against fire and rescue service data, a limitation that will be discussed in more detail hereafter.

The validation results presented in Table V.3 for the insurance data and the fire and rescue service data clearly indicate that IB models significantly reduce the number of false alarms, reinforcing the idea that HB models are less effective in identifying impacted areas. This result is expected, as flood impacts depend not only on flood hazard but also on local exposure and vulnerability conditions, which HB

Tableau V.3 – Validation results for thresholds (Thres.) providing a similar Probability of Detection (POD) = hit/(hit + miss) across all models, at (a) the municipality scale and (b) the river reach scale. For the Hazard-Based model, thresholds are expressed as return periods (years), while for the Impact-Based models, thresholds correspond to the number of flooded buildings (b.).

Database	Model	(a) Municipality scale							(b) River reach scale						
		hit	miss	fa	FAR	POD	CSI	Thresh.	hit	miss	fa	FAR	POD	CSI	Thresh.
CATNAT	HB	2454	1134	4381	0.64	0.68	0.31	10y							
	IB	2173	1415	10122	0.82	0.61	0.16	1 b.							
INSURANCE	HB	693	446	3022	0.81	0.61	0.17	10y	1350	658	17626	0.93	0.67	0.07	10y
	IB	694	445	2381	0.77	0.61	0.20	20 b.	1363	645	8496	0.86	0.68	0.13	5 b.
FIRE SERVICE	HB	171	447	1906	0.92	0.28	0.07	5y	170	611	5616	0.97	0.22	0.03	5y
	IB	170	448	1067	0.86	0.28	0.10	5 b.	165	616	1886	0.92	0.21	0.06	5 b.

models do not account for.

In addition, the discrepancy in false alarms appears to be amplified on the river reach scale. HB warnings result in 1.3 to 1.7 times more false alarms on the municipal scale, but this ratio increases to 2 to 2.7 on the river reach scale. This can be attributed to the ability of the IB approach to detect areas without exposed assets, to avoid false alarms. This added value is limited at larger spatial scales, such as municipalities (mean area 20 km2), where the likelihood that exposed assets are present somewhere within the area is higher compared to the river reach scale (mean subcatchment area 4 km^2).

4.2 Discrepancy between CatNat data and the IB model

The results obtained from the CatNat validation data in table V.3 do not show the same tendency of reduction of false alarms with the IB approach. First, we observe that the IB hit rate cannot exceed 61%, as this score was obtained using the lowest impact-based threshold, that is, 1 building in a flooded area. This means that for almost 40% of CatNat decrees - which are theoretically only requested when damage to buildings occurs - our simulation chain did not identify any flooded building. This conclusion remains even if a building is considered flooded under less than 15cm of water (which was the threshold chosen for an approximate correspondence with a typical door threshold).

The limited performance of the IB approach for the detection of CatNat decrees can be attributed to the fact that the impacts resulting from surface flooding - a hazard that we do not account for - are not distinguished from the river flooding impacts in the GASPAR database. This distinction can be crucial, as a study led by CCR in the French Mediterranean region revealed that the proportion of the impacts due to "surface flooding" (in terms of economic damage) is 60 % within all CatNat flood events between 1995 and 2019 [Moncoulon et al., 2020]. Furthermore, the analysis of the iCatNat database, which contains additional information on the type of flooding hazard, revealed that 40% of all CatNat requests between 2018 and 2023 in the study area were exclusively triggered by surface flooding phenomena. This could explain the fact that the CatNat decrees are not appropriate to validate our IB approach, which focuses only on river flooding for catchments greater than 5 km^2. In contrast, a Hazard-based model emitting warnings based on flow return periods is more consistent with CatNat decrees data, since surface flooding impacts are directly caused by heavy rainfall patterns that are reflected in the flows return periods.

Figure V.7 provides an example of these limitations, for the September 2020 flood event in the Gardon catchment. In this case, the validation based on CatNat decrees yield better performance for the HB model than for the IB model. Using additional information from the iCatNat database (i.e., the number of affected buildings declared by the mayor and the type of flood hazard), it is observed that all CatNat decrees involving "river flooding", "torrential flooding", or both are detected by both the HB

and IB models. However, among the five CatNat decrees detected only by the HB approach (highlighted with red hatching in the figure), we find that two were declared for surface flooding only and three for a combination of surface flooding, river flooding, and torrential flooding. This may explain why the IB approach did not identify damage : it was not designed to account for surface water flooding or for the impacts generated by sediment transport. Another important limitation emerges. For four out these five municipalities, the CatNat request submitted by the mayor did not report any damage to buildings. This observation contradicts the assumption underlying our use of CatNat decrees as a proxy for flood-related impacts, namely that such decrees systematically reflect actual damage within the municipality. In four of these cases, it therefore suggests that the IB model was correct in not issuing a warning, as no building damage was apparently observed. These findings can be related to the limitations briefly discussed in Section 2.2.1 : while CatNat decrees are in principle based on objective criteria, the recognition process may sometimes reflect precautionary requests from municipalities, which means that a return period of 10 years does not always correspond to significant observed damage.

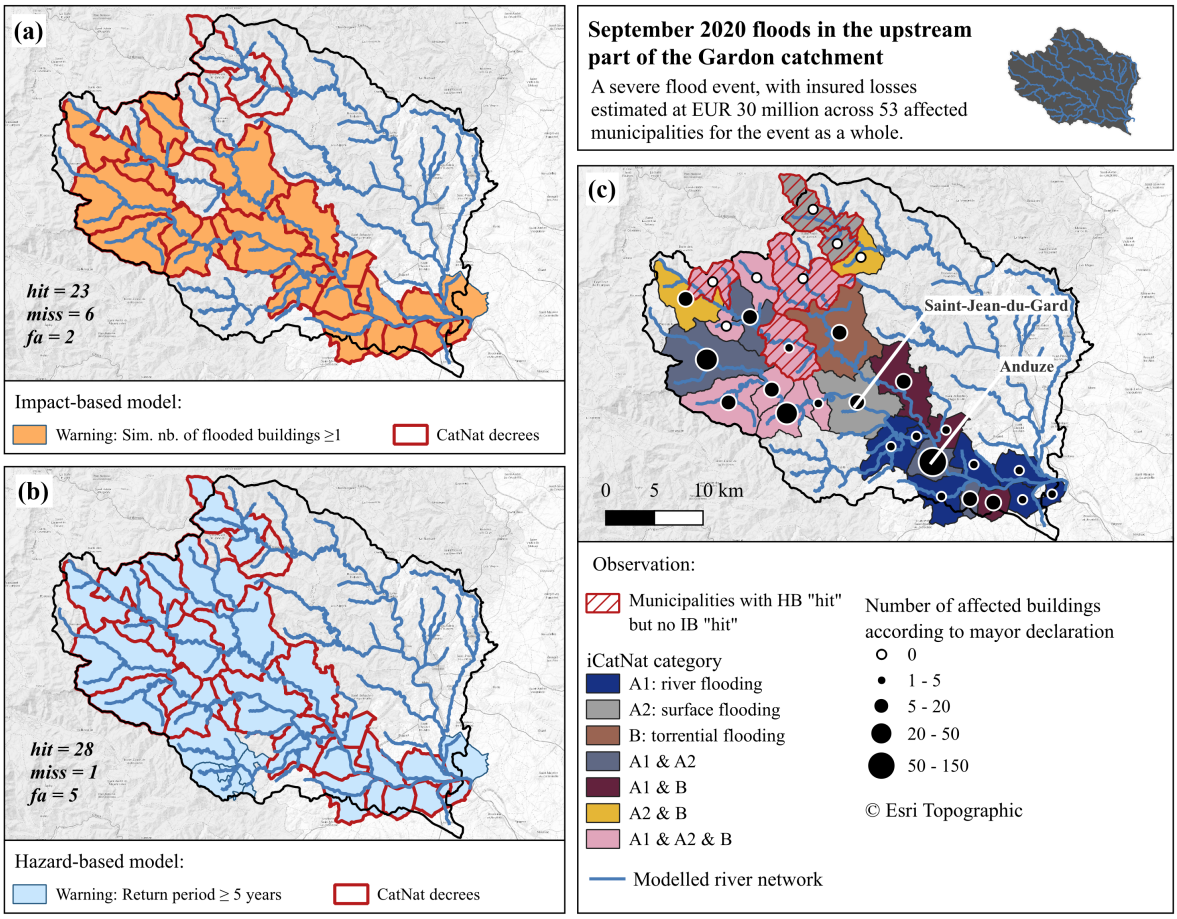


FIGURE V.7 – Illustration of the discrepancy between CatNat data and the IB approach, based on the September 2020 flood event in the Gardon catchment : (a) Results of the IBW for a 1 building threshold ; (b) Results of the HBW for a 5-year threshold ; and (c) Types of flooding and reported impacts from the iCatNat database.

4.3 Low hit rates for the fire and rescue services operations

Another striking result from Table V.3 is that even at low thresholds, the validation based on fire and rescue service operations yields very low hit rates. This is observed for both HB and IB approaches. The thresholds shown in the table, namely, a 5-year return period for the HB model and a 5-building

threshold for the IB model, were selected because they result in nearly identical hit rates. However, even when using the lowest possible thresholds (that is, a 2-year return period or a 1-building threshold), the hit rates remain below 0.30.

One explanation lies in the fundamental difference between the fire and rescue service operation data and the insurance claims and CatNat datasets : unlike the latter, the fire and rescue service data are not limited to significant or officially recognized events. In theory, only events with a return period of at least 10 years should be recorded in the CatNat or insurance data. In contrast, the analysis of the fire and rescue service operations collected for the Gard and Alpes-Maritimes departments (Figure V.2) shows that 60% occurred outside of CatNat-declared events. More strikingly, over 90% of the operations not detected by the warning chains and thus classified as 'misses', correspond to interventions that did not take place during any CatNat event. In the cases where detailed descriptions of the interventions were available, we found that many of these 'misses' cases were associated with events outside the scope of our nowcasting chains, for example, basement flooding, sewer overflows after heavy rain, multiple interventions in different areas that are recorded within a single point, reconnaissance missions with no reported damage, or water infiltration due to poor building maintenance. This highlights a key limitation of using this database for model validation : it captures a broader and more diverse range of impacts, many of which are either unpredictable or irrelevant in the context of our warning chains.

However, despite these limitations, the fire and rescue services data remains highly valuable in the context of our study. Since our goal is to compare warnings issued by hazard-based versus impact-based approaches, and particularly to quantify the possible reduction in false alarms with the IB warnings on long periods of time, the richness of the fire and rescue service operations data still appears interesting for this purpose, as illustrated in table V.3 .

4.4 Flood-based and hybrid models

Although the objective of this study is to quantify the added value of the Impact-based chain compared to the Hazard-based chain, we introduce intermediate models for discussion purposes.

First, a Flood-Based (FB) Model : This model outputs the real-time percentage of the forecasted flooded area within the 1,000-year flood envelop, for each river reach or municipality. Comparing this model with the HB and IB models allows us to quantify the added value of the hydraulic modeling stage and to ensure that the expected benefits of the IB model truly originate from asset description, rather than only from the additional information provided by hydraulic modeling. Although this assumption may seem intuitive, it still requires verification.

Then, we introduce Hybrid Models : These models integrate elements from both the HB and IB approaches. Although HBW are activated when the discharge return period exceeds a predefined threshold, and IBW are activated when the number of affected buildings exceeds a set threshold, hybrid models combine both criteria. For instance, a hybrid HB-IB warning could be based on the simultaneous exceedance of both thresholds, such as : "The flow return period exceeds the 10-year threshold AND the number of affected buildings is greater than 5."

Table V.4 presents the performance of all nowcasting chains (HB, IB, FB, and hybrid HB&IB) based on the Critical Success Index (CSI). For each approach, the optimal threshold was selected, as illustrated in Figure V.6. Due to the limitations identified in Section 4.2 concerning the unsuitability of the CatNat database for evaluating the IB model, the comparison focuses solely on validation results based on the insurance and fire and rescue service datasets.

These new results first confirm some conclusions drawn from Table V.3 : the added value of the IB model is greater on the river reach scale than on the municipality scale. They also bring several

new learnings. First, the Flood Based (FB) model, which triggers warnings based solely on the extent of the flood, performs significantly worse than the other models. This confirms the added value of incorporating exposure information when moving from HB to IB approaches. Then, for the insurance data, the best performance is achieved by combining hazard-based and impact-based thresholds. This approach significantly reduces the number of false alarms without severely impacting the number of hits. This is consistent with the fact that insurance data are only available for CatNat events, which are theoretically declared when return periods exceed 10 years. In contrast, for fire and rescue service data, combining hazard- and impact-based thresholds does not improve performance. This aligns with the operational logic of fire and rescue service mobilization, which is not conditioned by the hazard return period.

4.5 Model uncertainties and data limitations

If we examine the results in absolute terms—rather than relatively by comparing different models, as done in the previous discussions—the CSI scores remain low. For results based on fire and rescue service data, the limitations discussed in section 4.3 help account for these values. For results based on insurance data, the highest CSI score obtained is 0.30, at the municipality scale, for a Hybrid HB & IB model issuing a warning when the discharge return period exceeds 20 years on a river within the municipality and at least 10 buildings are counted within the expected flood extent (table V.4). Even with this strict warning threshold, the number of false alarms (735) exceeds the number of correctly detected events (553), while the number of missed events also remains substantial (586).

To put these numbers into perspective, all the uncertainties propagating through the nowcasting chains must be considered. Rainfall input data may contain inaccuracies ; the hydrological model used is simplified and uncalibrated, introducing additional uncertainties ; the flood extent representation relies on a limited number of pre-calculated scenarios, themselves produced with a simplified hydraulic model at regional scale ; and the impact model is again highly simplified, defining a flood impact as any area where more than a given number of buildings are exposed to more than 15 cm of water. Yet we are well aware that this does not accurately reproduce real processes of damage generation. As mentioned earlier, the literature has shown that claims ratios within simulated flooded areas are relatively low ($\simeq$ 40%) [Le Bihan et al., 2017], which helps explain the large number of false alarms.

In addition to model-related uncertainties, the limitations of the evaluation datasets—already discus-

Tableau V.4 – Validation results for thresholds (Thres.) providing the best Critical Sucess Index (CSI) = hit/(hit + miss + fa) across all models, at (a) the municipality scale and (b) the river reach scale. For the Hazard-Based model, thresholds are expressed as return periods (years), while for the Impact-Based models, they correspond to the number of flooded buildings (b.), and for the flood-based model, to a percentage of the surface of the 1000-year flood extent.

Database	Model	(a) Municipality scale							(b) River reach scale						
		hit	miss	fa	FAR	POD	CSI	Thresh.	hit	miss	fa	FAR	POD	CSI	Thresh.
CCR	HB	499	640	1329	0.73	0.44	0.20	50y	646	1362	4291	0.87	0.32	0.10	50y
	HB & FB	300	839	683	0.69	0.26	0.16	50y & 15%	612	1396	3720	0.86	0.31	0.11	50y & 20%
	FB	265	874	1716	0.87	0.23	0.09	50%	586	1422	6194	0.91	0.29	0.07	70%
	HB & IB	553	586	735	0.57	0.49	0.30	20y & 10 b.	836	1172	1671	0.67	0.42	0.23	10y & 10 b.
	IB	694	445	2381	0.77	0.61	0.20	20 b.	870	1138	2940	0.77	0.43	0.18	20 b.
FIRE SERVICE	HB	139	479	1063	0.88	0.23	0.08	10y	129	652	2831	0.96	0.17	0.04	10y
	HB & FB	133	486	984	0.88	0.21	0.08	10y & 10%	102	670	1642	0.94	0.13	0.04	10y & 40%
	FB	76	543	591	0.89	0.12	0.06	50%	89	683	2050	0.96	0.12	0.03	60%
	HB & IB	117	501	311	0.73	0.19	0.13	10y & 10 b.	112	666	609	0.84	0.14	0.08	5y & 10 b.
	IB	144	474	512	0.78	0.23	0.13	20 b.	125	656	733	0.85	0.16	0.08	20 b.

sed in detail—must also be taken into account. Despite our efforts, the "observed impact" does not always match the simulated impact, which only accounts for river flooding impacts in catchments larger than 5 km^2. Some of the "misses" may be explained by evaluation data that still includes impacts from surface or local flooding. Moreover, previous studies [Moncoulon et al., 2020] have shown that a substantial share of flood-related building damage remains unexplained ($\simeq$ 40%), even when using advanced models that combine river and surface flooding. This also contributes to the relatively high number of misses.

More broadly, our findings highlight several key recommendations for future impact data collection and management in the context of IBW evaluation. First, multiple and complementary impact data sources should be jointly leveraged, as relying on a single dataset may lead to biased conclusions regarding the added value of IBW. In particular, our results show that using the CatNat database alone would have resulted in a different assessment of IBW performance. Second, the spatial resolution of impact observations is an important factor : higher-resolution datasets lead to a greater apparent added value of IBW. Third, caution is required when comparing absolute verification scores across datasets. Although the same IB model is applied to all impact observation types, the resulting CSI values are not directly comparable due to intrinsic limitations and reporting practices specific to each dataset, as illustrated by the low CSI values obtained when validating against fire and rescue service interventions. Consequently, in this study, we focus on comparing HB and IB model performance within a given dataset, rather than across different data sources.

Finally, the validation approach applied here is particularly demanding, requiring both spatial and temporal accuracy. It is performed continuously over 13 years of data, with high spatial resolution (average municipality area : 20 km^2 ; average river reach : 4 km^2) across a large study area of 42,000 km^2. The CSI metric itself is strict in its definition, as it does not account for near misses—for instance, when a warning is issued for a river reach adjacent to the one actually impacted. Only a slight tolerance of two days on event dates was introduced to account for uncertainties in the timing of observed events.

4.6 Validation results at the event scale

4.6.1 Performance of HB and IB models at detecting impacted areas

Figure V.8a shows the distributions of CSI scores calculated separately for each event (instead of aggregated CSI scores over the 2010-2022 period presented in table V.4). In this representation, each point corresponds to a specific flood event (approximately 300 events in total). The scatterplots compare the CSIs obtained from the best-performing hazard-based (HB) and impact-based (IB) models, as defined in Table V.4, using the insurance dataset at the river reach scale. This event-based representation clearly shows that the IB model generally outperforms the HB model : the vast majority of points lie above the 1 :1 line. The next section will explore this further by incorporating additional information available in the insurance dataset : that is, the ranking of affected river reaches based on the number of insurance claims reported.

4.6.2 Performance of HB and IB models performance at ranking impacted areas

This last section focuses on the ranking-based validation approach presented in section 3.3 and Figure V.5. Figure V.8b shows the ranking performance of both the HB and IB models, calculated individually for each of the ~ 300 flood events recorded between 2010 and 2022 in the 61 catchments in the study area. Each point represents one event and the CSI scores compare in this case the ability of each model to reproduce the observed ranking of affected river reaches according to insurance claim data.

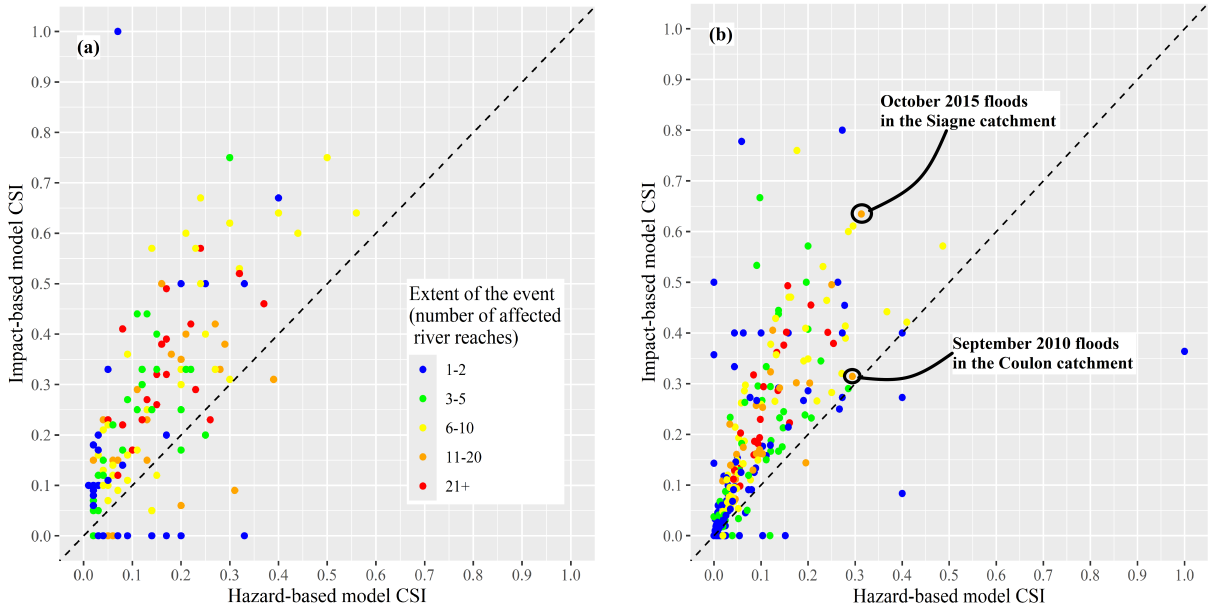


FIGURE V.8 – Comparison of the (a) Threshold-based and (b) Ranking-based validation results at the event scale based on the insurance claims dataset, for the HB and IB warning approaches. Each point represents a flood event.

The figure clearly demonstrates that the IB model consistently outperforms the HB model. Most points lie above the 1 :1 line, indicating a higher CSI for the IB model. In the rare cases where the HB model performs better, the events tend to be very localized, typically involving only one or two affected reaches (blue points). In these cases of localized events, the differences of performance between both approaches are more subject to random fluctuations.

These results finally confirm that the IB warning approach, which prioritizes areas based on the simulated number of flooded buildings, offers a significantly better ability to reproduce the observed spatial distribution of impacts than the HB approach, which ranks areas based only on flood return periods. This suggests that incorporating vulnerability and exposure elements into early warning models should significantly improve their relevance for emergency response.

Cartographic illustrations are provided in Figure V.9 for two specific events. For the October 2015 event in the Siagne catchment (Figure V.9a), the ranking predicted by the IB model closely matches the observed classification, resulting in an CSI of 0.63. In contrast, the HB model shows significant discrepancies, particularly in the western part of the catchment, and yields a much lower CSI of 0.31. Conversely, Figure V.9b illustrates the case of a very localized event—the September 2010 flood in the Coulon basin - where both models provide only moderate performance in reproducing the observed impact ranking. The IB and HB models produce similar CSI scores (0.31 and 0.29, respectively), suggesting a limited, though not negligible, ability to capture the spatial distribution of damage in this case.

5 Conclusions

This study provides new quantitative evidence of the added value of impact-based warnings (IBW) for flash flood monitoring. Using a large-scale, multisource validation framework covering the French Mediterranean region from 2010 to 2022, we demonstrated that IBW substantially reduce false alarms compared to hazard-based warnings (HBW) on a river reach scale, by a factor of 2 to 3. We also showed that IBW outperform HBW in their ability to identify and prioritize impacted areas. These findings highlight the potential of IBW to better support decision-making during flood events.

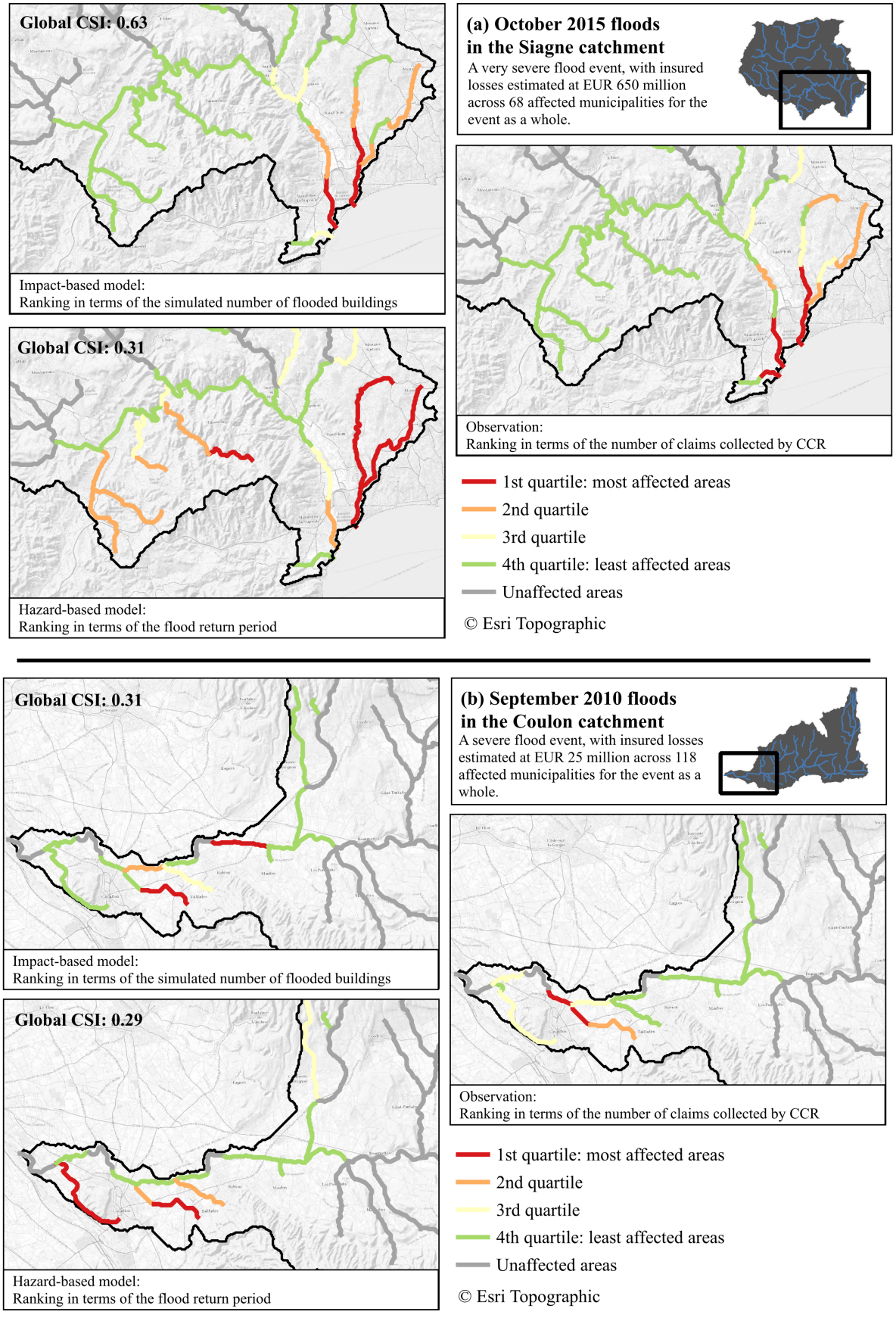


FIGURE V.9 – Illustration of ranking-based validation results for two selected flood events : (a) October 2015 event in the Siagne catchment, (b) September 2010 flood event in the Coulon catchment.

At the same time, the evaluation revealed important limitations, particularly regarding the quality and consistency of available validation datasets with the nowcasting chains used to issue warnings. CatNat decrees, while offering broad coverage, are subject to severity thresholds, political influence, and inconsistencies in hazard attribution, all of which may skew validation results. Insurance data, in the specific French context studied here, are likewise threshold-dependent and merge diverse flooding processes (e.g., surface and river flooding), while also reflecting a complex claim process. Fire and rescue service operation records, although more detailed and less constrained by severity thresholds, present challenges due to heterogeneity and the inclusion of out-of-scope impacts linked to localized flooding or unclear causes. Beyond quantifying IBW performance, this study therefore underscores the critical need for systematic and harmonized post-event impact data collection. Appendix D points to promising but underused datasets, particularly those related to road closures, which represent both a major source of disruption and a modeling challenge. Future research should investigate how such datasets could be integrated into IBW validation frameworks.

The IBW assessed here relied on simple deterministic thresholds of flooded buildings. However, impact modeling and warning generation could benefit from alternative approaches. Recent work suggests that warning thresholds should be adjusted to reflect people's tolerance levels for false alarms, in order to improve warning effectiveness [Zhang et al., 2024]. Similarly, shifting from deterministic to ensemble-based nowcasting could better represent uncertainty and enhance confidence in early warnings, a direction supported by previous research on the Vigicrues Flash system [Godet et al., 2023]. This would also include replacing precipitation observations by forecasts, which will be important to account for in the evaluation process.

Finally, while our results focus on measurable gains in false alarm reduction and spatial prioritization, other potential benefits of IBW remain to be systematically evaluated. These include improved communication of forecasts, as several studies suggest that HBW are not always well understood by citizens, e.g., [Amengual et al., 2024]; a direct comparison of the public understanding of IBW and HBW therefore appears crucial, with a first attempt proposed by Potter et al. [2018]. Additional benefits of IBW may arise from stronger cross-sector collaboration, which could be verified by assessing the added value of IBW co-developed through interactions between forecasters and end-users, compared to HBW developed within meteorological or hydrological services alone [Potter et al., 2021]. Finally, it would be interesting to verify that IBW facilitate multihazard approaches by enabling the quantification of the added value of a unique IBW, compared to a HBW for each hazard, which in the case of flash flooding can include strong winds, surface flooding, torrential flooding or landslide. Assessing these advantages requires interdisciplinary work bridging hydrology, impact modeling, crisis management, and social sciences.

Finally, while our results focus on measurable gains in false alarm reduction and spatial prioritization, other potential benefits of IBW remain to be systematically evaluated. These include improved communication of forecasts, stronger cross-sector collaboration, and support for multihazard approaches. Assessing these advantages requires interdisciplinary work bridging hydrology, impact modeling, crisis management, and social sciences.

In conclusion, this study contributes both methodological insight and empirical evidence of the operational potential of IBW in flash flood contexts. It also highlights critical areas for improvement and future research, supporting the ongoing transition toward more impact-oriented and people-centered flood warning systems.

Chapitre VI

Incertitudes : prévisions d'ensemble de pluie et propagation des erreurs

Objectifs et organisation du chapitre :

Dans les chapitres précédents, l'approche retenue pour la chaîne de prévision immédiate était entièrement déterministe. Ce chapitre, qui vise principalement à explorer de nouvelles pistes de recherche et des perspectives, s'intéresse à la valeur ajoutée des prévisions d'ensemble de pluie en entrée de la chaîne.

La première partie étudie la plus-value d'un produit de prévisions d'ensemble des pluies à horizon +3h (PIAF-ensemble), par rapport à des prévisions déterministes, dans une chaîne de type Hazard-based Warning, proche du système Vigicrues Flash. L'objectif est de définir un cadre d'évaluation cohérent avec le principe HBW fondé sur les dépassements de seuils de débits, et de l'appliquer à huit événements de crues soudaines, afin d'analyser le comportement de PIAF-ensemble sur plusieurs événements significatifs. Ce travail, publié dans la revue *Natural Hazards and Earth System Sciences* [Godet et al., 2023], est repris dans la première section du chapitre (texte intégral en anglais), précédé d'un court paragraphe de synthèse.

La deuxième partie s'intéresse à l'apport des prévisions d'ensemble dans une chaîne de type Impact-based Warning. Cette thématique a été initiée lors d'un séjour de recherche de deux mois à Earth Sciences New Zealand, effectué pendant la thèse. Deux objectifs principaux sont poursuivis : d'une part, illustrer les changements et problématiques induits par le passage d'une approche déterministe à une approche ensembliste dans une chaîne de type IBW ; d'autre part, fournir des premiers éléments, à la fois méthodologiques et relatifs aux résultats, concernant la propagation des incertitudes à travers les différents blocs de la chaîne IBW.

1 Plus-value de prévisions d'ensemble de pluies dans une chaîne de type HBW

1.0 Synthèse en français

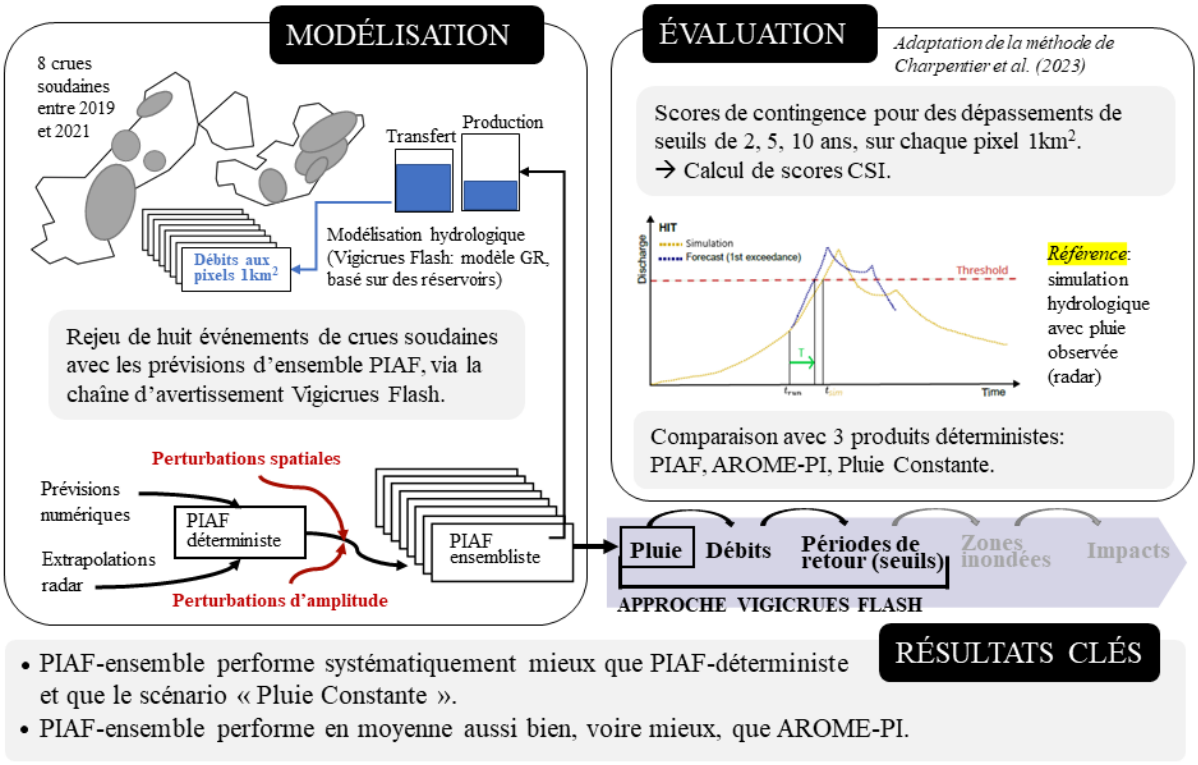


FIGURE VI.1 – Résumé graphique de l'article Godet et al. [2023].

L'objectif de ce travail est d'évaluer la plus-value apportée par des prévisions d'ensemble de pluie dans la chaîne d'avertissement de type HBW (Vigicrues Flash). Les prévisions de pluies étudiées sont une première version ensembliste du système de prévision immédiate PIAF, développée dans le cadre du stage de Walcarius [2021]. PIAF (Prévision Immédiate Agrégée Fusionnée) [Moisselin et al., 2019] est un produit dédié à la prévision immédiate (horizons 0-3h), qui repose sur la fusion d'images de lames d'eau radar extrapolées et d'images simulées par le modèle numérique AROME-PI [Auger et al., 2015].

Cet ensemble de prévisions de pluie est composé de 17 membres. Ces derniers sont injectés en amont du modèle hydrologique similaire à la chaîne opérationnelle Vigicrues Flash décrite au Chapitre II. Huit événements de crues soudaines ont été rejoués par le modèle. Les débits ont été calculés sur chaque pixel de 1 km^2 (une différence notable par rapport au reste du manuscrit, où les débits sont évalués à l'échelle du tronçon de cours d'eau) et comparés à des quantiles de débits établis à partir d'un rejeu climatique du modèle et d'un ajustement Gumbel, afin de définir l'émission des avertissements aux crues soudaines.

L'évaluation de ces avertissements repose sur des scores de contingence. Une simulation de référence est construite en forçant le modèle hydrologique avec des observations de pluie. Les dépassements de seuil de cette simulation sont ensuite comparés à ceux issus des prévisions. Des "hits", "misses" et "false alarms" sont comptabilisés selon une méthode adaptée de Charpentier-Noyer et al. [2023], avec une adaptation concernant la sélection des échéances de prévisions (méthode de stratification, Bellier et al., 2017).

Les tableaux de contingence obtenus peuvent être cartographiés pour visualiser géographiquement les résultats de détection sur chaque événement, ou agrégés en sommant les résultats sur chaque pixel et

en calculant un CSI (Critical Success Index) : $\frac{hit}{hit+miss+false\,alarm}$. Les scores CSI issus des différents déciles de PIAF-ensemble ont été comparés à ceux de trois produits déterministes : AROME-PI, un scénario “naïf” de pluie constante, et PIAF déterministe.

À l’échelle de chaque événement, le score CSI agrégé sur l’ensemble des mailles est supérieur pour l’approche PIAF-ensemble par rapport aux approches à pluie constante et PIAF-déterministe (figure VI.10). Ce constat se vérifie systématiquement pour le décile 50 % de l’ensemble, et généralement pour une gamme élargie de déciles. Toutefois, dans certains cas, AROME-PI obtient de meilleures performances ; un résultat déjà observé où AROME-PI surpassait ponctuellement PIAF-déterministe [Lovat et al., 2022]. En revanche, lorsque les résultats sont moyennés sur l’ensemble des événements, les déciles intermédiaires (40–50 %) de PIAF-ensemble présentent des scores équivalents, voire supérieurs, à ceux d’AROME-PI (figure VI.11).

D’autre part, une analyse des délais d’anticipation des dépassements de seuils de débits a montré que la plus-value de PIAF et de PIAF-ensemble par rapport à AROME-PI concernait majoritairement les échéances de prévision très courtes (entre 0 et 1 h), ce qui correspond précisément à la gamme temporelle couverte par les extrapolations radars, ajoutées dans ces produits mais absentes de l’approche AROME-PI.

Ce travail met donc en évidence la plus-value du produit PIAF-ensemble dans une chaîne d’avertissement de type HBW analogue à Vigicrues Flash. Cette évaluation portant sur un échantillon de 8 événements, elle devra être complétée par des analyses statistiques réalisées sur de longues chroniques d’événements et de non événements.

1.1 Introduction

The year 2022, particularly the summer season, was marked by several deadly and catastrophic flash floods in Pakistan, Kentucky (USA), Iran, Sierra Leone, Bangladesh, Australia, and unfortunately many other countries. Very few parts of the world seem to be spared from flash floods. According to the World Meteorological Organization [WMO, 2020], floods are the deadliest natural hazards, and flash floods account for 85 % of the flooding events and have the highest mortality rate within the category (5000 victims annually). In France, the Mediterranean region is particularly prone to severe flash floods. Even though an intensification of extreme rainfall events in response to anthropogenic influence was diagnosed [Ribes et al., 2019], the consequences of climate change on flash floods remain unclear in this region, particularly because of the compensating effect of the expected decrease in soil moisture [Tramblay et al., 2019]. However, the increase in the vulnerability to these episodes may lead to an increase in the global risk associated with flash floods in the future years.

In this context, developing flash flood forecasting is of crucial interest to limit the death toll and optimize the emergency response. Several operational flash flood warning systems have recently been developed worldwide, and they generally have similar features. The observed or forecasted rainfall can be directly compared to reference thresholds to estimate the flash flood likelihood. This is the case for instance in the Flash Flood Guidance system in the US [Clark et al., 2014] or the ERIC–ERICHA system in Europe [Raynaud et al., 2015, Corral et al., 2019]. Rainfall data can also be used as input to highly distributed hydrological models, which may bring additional information about the intensity and temporal dynamics of the floods and may be particularly interesting for decision-making [Zanchetta and Coulibaly, 2020]. The FLASH system in the USA [Gourley et al., 2017] and the Vigicrues Flash service in France [Javelle et al., 2016, Piotte et al., 2020] follow this second principle. The operational systems using hydrological models are still often based on radar quantitative precipitation estimates (QPEs), without involving quantitative precipitation forecasts (QPFs). This choice not only increases the quality of detection and limits the risks of false alarms but also highly limits the anticipation that cannot exceed the (limited) response times of the small catchments where flash floods do occur.

The development of convection-permitting numerical weather prediction (NWP) models has paved the way for the use of QPFs as input to flash flood warning systems, with the objective of extending anticipation lead times up to 24-48 h [Collier, 2007, Hapuarachchi et al., 2011, Zanchetta and Coulibaly, 2020]. Convection-permitting models offer an interesting capacity to describe heavy precipitation events and offer space and time resolutions which are suited to the hydrological models used in flash flood warning systems. However, the current QPF products still show spatial and temporal uncertainties in the description of intense rainfall cells that may significantly exceed the typical scales of small river basins [Roberts and Lean, 2008, Clark et al., 2016, Armon et al., 2020]. This may highly limit the capacity to issue relevant flash flood warnings, without appropriate strategies to represent or reduce uncertainties [Silvestro et al., 2011, Vincendon et al., 2011, Furnari et al., 2020]. Even if ensemble approaches have been widely used as input to flash flood forecasting chains [Vié et al., 2012, Davolio et al., 2013, 2015, Hally et al., 2015, Nuissier et al., 2016, Amengual et al., 2017, 2021, Furnari et al., 2020, Sayama et al., 2020], uncertainties in QPFs can still hardly be reduced for lead times exceeding 6–8 h, even with enhanced assimilation schemes in NWP models [Davolio et al., 2017, Lagasio et al., 2019].

Efficient flash flood forecasting strategies can also be developed for short lead times (< 6 h, i.e. the nowcasting range), with a high update frequency (typically 5 min to 1 h between two runs of forecasts) to regularly benefit from the last available observations [Lovat et al., 2022]. For such applications, the QPF products can be derived either from adapted versions of convection-permitting NWP models [Auger et al., 2015, Benjamin et al., 2016] or by extrapolating the last radar observations [Berenguer et al.,

2011, Silvestro and Rebora, 2012, Imhoff et al., 2022]. Simple Lagrangian radar extrapolations can easily outperform NWP models for lead times up to 2–3 h [Mandapaka et al., 2012] ; however they are not suited to larger lead times because they cannot reproduce the physical changes occurring in the atmosphere. For that reason, up-to-date short-range QPF approaches now combine both information sources through blending techniques to offer a seamless transition between observed and forecasted rainfall fields [Poletti et al., 2019, Lovat et al., 2022, Scheufele et al., 2014]. However, despite all these efforts to create seamless short-range QPFs products, the forecast uncertainties still remain significant and need to be quantified through ensemble approaches [Bowler et al., 2006, Seed et al., 2013, Descamps et al., 2015, Osinski and Bouttier, 2018, Bouttier and Raynaud, 2018].

The objective of this paper is to assess the potential of a new seamless short-range ensemble QPF product called PIAF-EPS ("PIAF" meaning Prévision Immédiate Agrégée Fusionnée and "EPS" meaning ensemble prediction system) and recently developed by Météo-France for flash flood forecasting purposes. This ensemble aims to represent very shortrange forecast uncertainties. It can be frequently updated at a very small numerical cost, in order to keep it consistent with the latest nowcasting data based on radar images. The aim is to confirm the benefits of using such an ensemble seamless product as input to flash flood nowcasting chains, compared to other short-range deterministic QPF products from which PIAF-EPS is directly derived. For this purpose, a reanalysis of eight flash flood events observed in the French Mediterranean region between 2019 and 2021 is proposed, using a similar hydrological-modelling chain as the one implemented in the French Vigicrues Flash operational flash flood monitoring system. Since the selected flash floods mainly occurred on small rivers, the proposed evaluation framework not only focuses on a couple of gauged outlets but also offers comprehensive coverage of the small rivers hit by the studied rainfall events. This is achieved by comparing the hydrological forecasts obtained using QPFs with simulated discharges (i.e. based on QPEs) at each pixel of the hydrological-model grid (1 km resolution) and following a methodology adapted from [Charpentier-Noyer et al., 2023].

The paper is organized as follows : Sect. 1.2 describes the hydrometeorological forecasting chains compared in the study ; Sect. 1.3 provides details about the case studies used for the evaluation as well as the chosen verification method ; and, finally, Sect. 1.4 presents and discusses the verification results.

1.2 The short-range hydrometeorological forecasting chains

1.2.1 General structure of the chains

The forecasting chains applied in this study are directly inspired by the French Vigicrues Flash operational flash flood monitoring service [Javelle et al., 2016, Piotte et al., 2020]. They are presented in Fig. VI.2. The chains evaluate the severity of the floods by comparing the simulated and forecasted hydrographs to reference discharge quantiles. These hydrological data are obtained using a fully distributed rainfall runoff model, detailed in Sect. 1.2.4. This hydrological model is forced with the PANTHERE (Projet Aramis Nouvelles Technologies en Hydrométéorologie Extension et Renouvellement) rainfall QPEs, derived from a network of about 30 radars over mainland France and its vicinity. [Tabary et al., 2013]. As mentioned in the Introduction, the choice of using radar QPEs without QPFs would tend to limit the false alarms emitted by the chain but would also drastically limit its anticipation capacity.

In this paper, we thus combined QPEs with different QPFs products as input of the chain, with the objective of increasing the current anticipation levels. The common time step for QPE–QPF and the hydrological model is 15 min. All the QPFs mentioned are available up to 6 h forecast lengths, but their refresh times depend on the considered product. For the present study, we decided to consider QPFs only for 0–3 h forecast ranges and with a common refresh time of 1 h. The QPF products include the

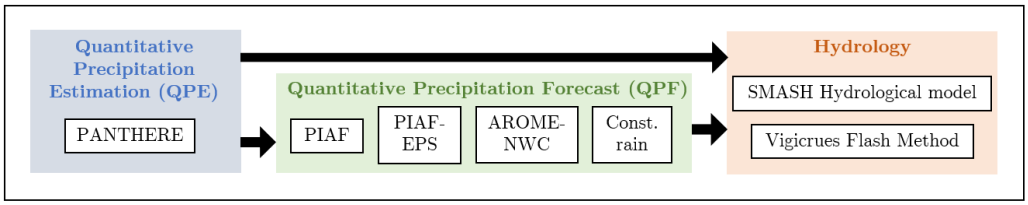


FIGURE VI.2 – General structure of the forecasting chains

new PIAF-EPS ensemble product, and three deterministic products are used as a reference. Two of these reference QPFs are directly involved in the generation of the PIAF-EPS ensemble (see Fig. VI.3), i.e. the deterministic version of PIAF, and the AROME-NWC numerical weather prediction model (AROME, Applications de la Recherche à l'Opérationnel à Méso-Echelle ; NWC, nowcasting). The third reference QPF corresponds to a naive constant-rain scenario.

The next sections present each of the components involved in the forecasting chains applied in this study.

1.2.2 The three deterministic QPFs : AROME-NWC, deterministic PIAF, and naive constant-rain scenario

The first QPF product used as input to the chain corresponds to the AROME-NWC system documented in [Auger et al., 2015]. It is a rapid-refresh version of the AROME convection-permitting numerical weather prediction system. It is updated every hour by a 3D-Var (3D variational) data assimilation system with a 10 min observation cutoff (i.e. the initial state of each forecast is prepared using observations collected up to 10 min after its validity time), from which 6 h forecasts are produced at 1.3 km resolution with a 20 min delivery time. Each 3D-Var analysis updates the model state by multivariately blending tens of thousands of observations from various meteorological networks (radar winds and reflectivities, satellite radiances, GPS data, in situ surface and aircraft reports, etc.). More information about the AROMENWC 3D-Var can be found in [Auger et al., 2015].

The second QPF product involved is the deterministic rain nowcasting system called PIAF (Prévision Immédiate Agrégée Fusionnée in French) [Moisselin et al., 2019]. Each PIAF forecast blends rainfall fields of radar QPF products and AROME-NWC numerical predictions as explained hereafter. The radar QPF product is derived from the PANTHERE radar QPEs. Rainfall accumulations are estimated every 5 min at 1 km resolution and extrapolated in time using an optical-flow technique that maintains the apparent motion of reflectivity from recent radar images. The blending of radar extrapolations and AROME-NWC follows the equation PIAF = $\alpha\times$ radar QPF + (1-α)$\times$ Arome-NWC, where α is a forecast range-dependent weighting factor. At short forecast ranges, α is equal to 1 so that PIAF is equivalent to the extrapolated radar QPF, which tends to be better than AROMENWC. At longer forecast ranges, typically beyond 1 to 2 h, α smoothly decreases towards 0 so that the PIAF converges to the latest available AROME-NWC precipitation forecast, which consistently outperforms radar QPF at longer forecast ranges. The speed at which α decreases is case-dependent : it is determined by a simple online machine learning procedure [Auer et al., 2002, Devaine et al., 2013] that minimizes the average forecasting errors over the past 6 h, as measured by a Gerrity score over large subdomains. In a nutshell, this algorithm produces a smooth transition (as a function of forecast range) between the latest available radar extrapolation and AROME-NWC forecast ; compared to climatologically optimal weights, this transition occurs earlier if AROME-NWC performed better than average during the 6 preceding hours (relative to radar extrapolation). Evaluations of the deterministic PIAF precipitation forecasts [Moisselin et al., 2019, Lovat et al., 2022]

indicate that they statistically outperform both radar QPF products and AROME-NWC forecasts for forecast ranges between 0 and 3 h.

Finally, we considered a third "naive" QPF scenario, corresponding to a constant future rain. Despite its very simplistic principle, this scenario may give valuable information since flash floods are often caused by quasi-stationary storm systems [Gaume et al., 2009].

1.2.3 The new PIAF-EPS ensemble QPF product

PIAF-EPS is a new experimental short-range ensemble rainfall product, which is built by adding perturbations to the deterministic PIAF nowcast. The ensemble generation is original and inspired by previously proposed stochastic nowcasting schemes, e.g. those of Bowler et al. [2006], Seed et al. [2013]. The perturbation tuning parameters have been kept to a minimum, in order to facilitate future operational deployment and maintenance of the proposed system. The perturbation technique is an adaptation to nowcasting ranges of the "pertDpepi" method used by Peredo et al. [2021], Charpentier-Noyer et al. [2023]. It is illustrated in Fig. VI.3. Each PIAF forecast (available every 5 min) is used to generate 16 perturbed members using equiprobable perturbations of the precipitation field : spatial perturbations and amplitude perturbations.

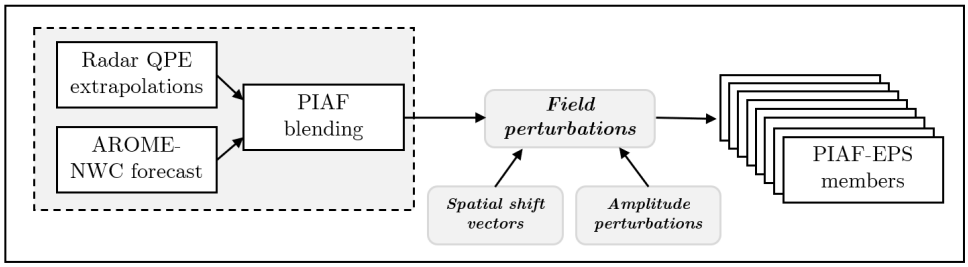


FIGURE VI.3 – Illustration of the generation of the PIAF-EPS ensembles

The 16 spatial perturbations are pseudorandom shifts that approximate (together with the unperturbed forecast) a 17member, isotropic Gaussian sample in the 2D space. The shifting vectors are computed following the recommendations and dataset of Wang et al. [2019], which include a Dirac mixture algorithm involving the Cramér–von Mises method. It is a deterministic 2D distribution that is on average a better approximation of a Gaussian than a Monte Carlo sample, given the small ensemble size. The vector directions are constant in time for each ensemble member. The vector amplitudes are scaled as a function of lead time so that the amplitude of the spatial shifts grows linearly from 0 to 30 km over 3 h, after which it is kept constant. This setting was based on a visual examination of spatial prediction errors for a set of high-impact precipitation events (independent of the ones used for the evaluations in this study).

The 16 amplitude perturbations are multiplications of 2D patterns by the spatially shifted fields. Each pattern is an independent realization of a 2D random field that has Gaussian autocorrelations in space and a serial time autocorrelation from a clipped AR(1) autoregressive process. The autocorrelation scales are set to approximately 40 km and 6 h, respectively. Thus, the amplitude perturbations are independent between members, and they slowly evolve in time. The standard deviation of the perturbation amplitude grows linearly in time for the first forecast hour, after which it is kept constant ; it has been tuned to produce reliable average standard deviations of the precipitation spread (as measured by the spread–skill ratio of the whole ensemble) over a large forecast tuning sample (1 month, independent of the cases evaluated in this study). Likewise, a small bias correction (amplification of the highest precipitation intensities) of the forecasts with respect to precipitation observations has been applied using the same tuning sample. An example of the perturbations is given in Fig. VI.4.

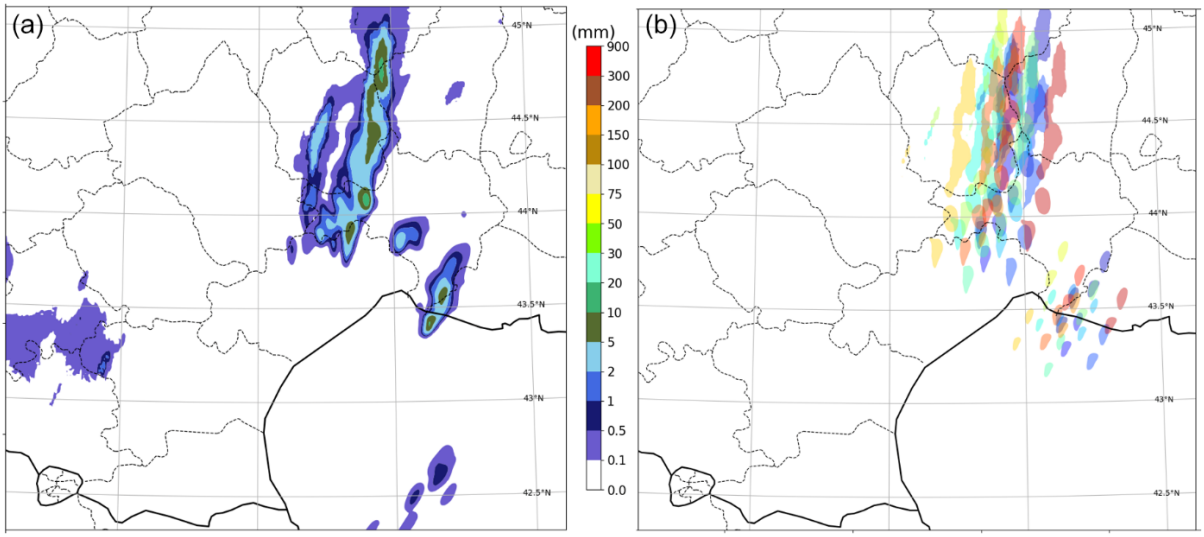


FIGURE VI.4 – Example of PIAF-EPS ensemble forecast perturbations. (a) Deterministic PIAF forecast of 15 min rainfall accumulation (forecast start : 19 September 2020 at 06 :00 UTC, forecast range : 2 h). This is used as member 0 of the ensemble. (b) Same field in members 1 to 16; the shading represents rainfall areas above 5 mm, with one colour for each member.

The unperturbed PIAF forecast is used as a 17th ensemble member, which makes the ensemble slightly nonequiprobable but minimizes the risk of corrupting a good deterministic forecast by applying ensemble perturbations that are too large. The justification is that, in a few high-impact cases, experience shows that intense Mediterranean precipitation can be predicted quite precisely by numerical models thanks to the influence of local orographic features. Further improvement to our (purely statistical) ensemble generation technique would be needed to automatically reduce the perturbation amplitudes in such cases, which is left for a future study.

1.2.4 The SMASH hydrological model and the Vigicrues Flash method

The rainfall-runoff part of the forecasting chains is based on SMASH (Spatially distributed Modelling and ASsimilation for Hydrology models). SMASH is a highly distributed, continuous, and conceptual hydrological model developed at INRAE (Institut national de la recherche pour l'agriculture, l'alimentation et l'environnement) and Hydris Hydrologie [Jay-Allemand et al., 2020]. The general principle of the model is presented in Fig. fig :smash. SMASH is inspired by the GR (Génie Rural) reservoir-based family of models [Perrin et al., 2003]. For each pixel of the territory, the model includes a production reservoir (capacity c_p); a transfer reservoir (capacity c_{tr}); and an adapted cell-to-cell routing model, represented by a routing reservoir (capacity c_r).

The version of SMASH used in this study is the one that is currently operational in the Vigicrues Flash system. This version is working on a 1 km grid, at a 15 min time resolution. It is a "lag-0" version, which means that there is no cell-to-cell routing scheme (or, in other words, that the routing velocity is infinite) : the discharge on a cell is the sum of the instantaneous discharges of all the upstream cells. This method does not provide realistic hydrographs, but this is not considered a problem, since the warning thresholds are defined based on a "climatological" run of the same model (see the next paragraph).

According to the Vigicrues Flash method, the forecasted hydrographs obtained with SMASH are compared with reference discharge quantiles corresponding to defined return periods. These reference values are obtained by running the SMASH model for a long and continuous period and by adjusting a Gumbel distribution to the corresponding annual maximum series. For this study, a 15-year-long

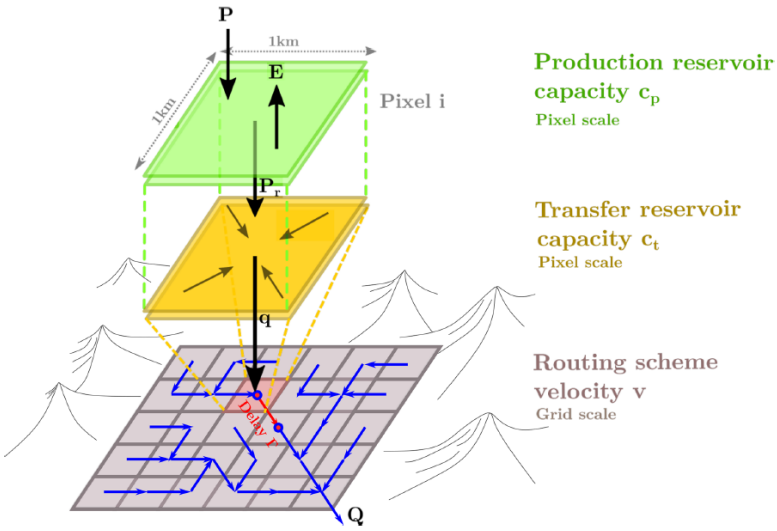


FIGURE VI.5 – General outlines of SMASH [Jay-Allemand, 2020]. P represents the local rainfall over one cell ; E is the potential evapotranspiration ; P_r is the effective rainfall ; q is the elementary discharge ; and Q is the total routed discharge

simulation period was used, which is the longest period that can be simulated based on an homogeneous PANTHERE QPE product. Discharge quantiles of T = 2, 5, and 10 years were obtained for each 1 km2 pixel of the studied area (see Sect. 1.3.1). Figure VI.6b illustrates the discharge quantiles obtained for the return period of T = 2 years. In the west zone, the effect of relief in the Cévennes mountainous area is clearly distinguishable, logically resulting in higher rainfall amounts and higher flood quantiles. However, the results appear less consistent in the east zone, firstly because several very intense events occurred in the 2006–2021 simulation time window (sensibility to sampling) and secondly because the quality of the radar rainfall is questionable in this area. Indeed, a V-shaped band can be clearly observed, which is probably the result of bad calibration of the Collobrières radar. However this does not alter the methodology and results proposed in this study, since only simulated (and not observed) discharges are used to assess the quality of the forecast results.

1.3 Case studies and verification method

1.3.1 Study area and selected events

The south of France, particularly the Mediterranean region, has experienced a large number of catastrophic flash flood events this last decade, both in terms of economical damage and casualties. The study has been focused on the most recent events that hit this area, since the ensemble PIAF-EPS forecasts can be released only from February 2019 (it would be labour intensive to process older cases because of technical constraints in the archiving system, and they would be less and less relevant to current operational forecasting systems because the AROME and PIAF systems are frequently upgraded, typically once a year). Eight heavy precipitation events which occurred between 2019 and 2021 in the southeastern region of France (see Fig. VI.6a) were selected. Figure VI.7 shows the maps of rainfall accumulations for each event, and Table VI.1 provides additional information including the duration, the maximum rainfall accumulation (spatial maximum), and the intensity and geographical extent of the hydrological responses simulated by the SMASH model.

The rainfall accumulations presented in Fig. VI.7 show that the selected events have very different features. Some events, such as events A and E, show a wide spread of rainfall. For these events, the larger rainfall accumulations appear homogeneous over areas covering one or several departments. The

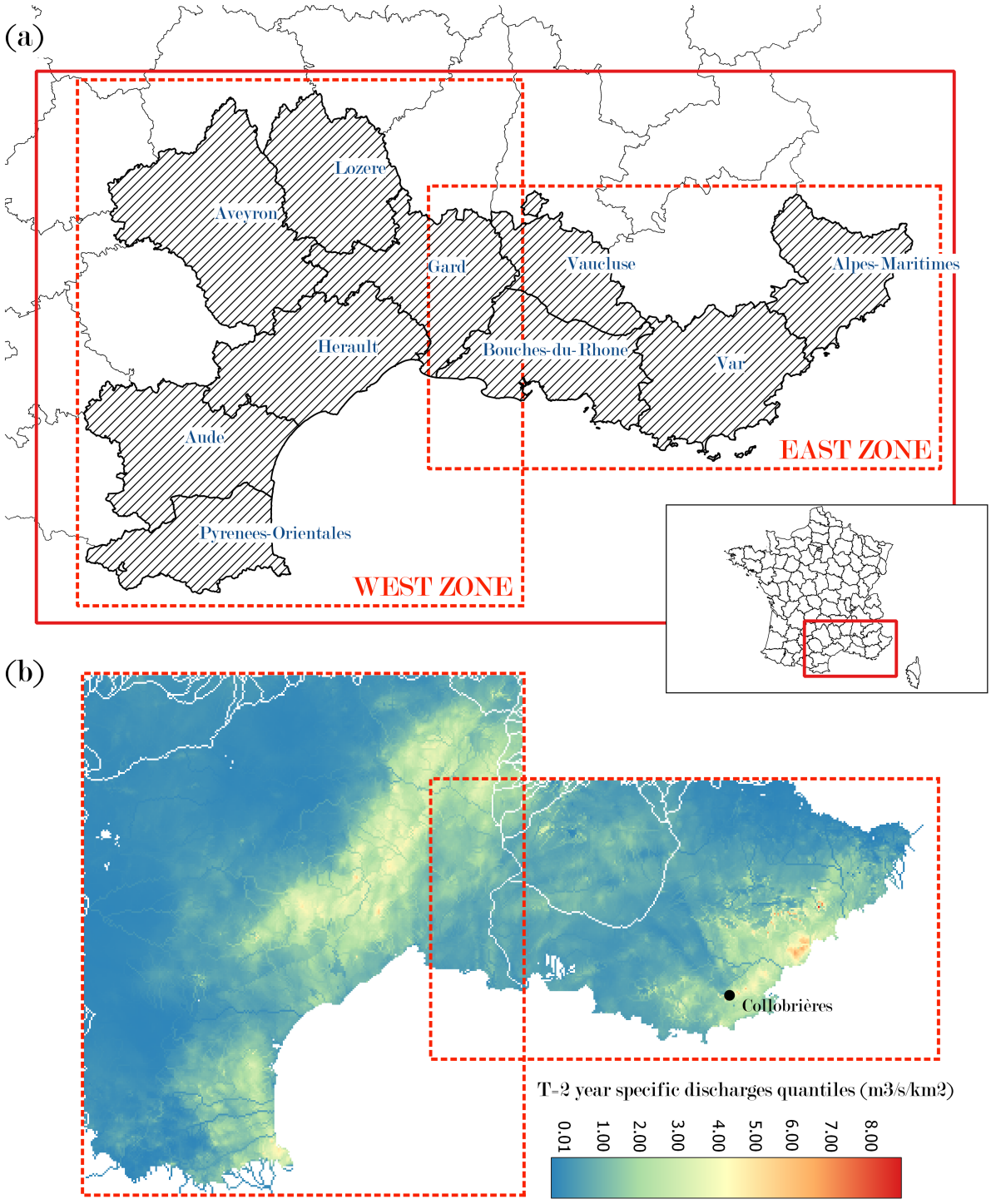


FIGURE VI.6 – Study area and French departments affected by the events (a) and T = 2years specific discharge quantiles estimated on the study area based on a 15-year SMASH simulation (b)

other events are much more localized and have a larger variability in rainfall accumulations. Some of them show locally very intense rainfall cells (events C, D, and F). The rainfall accumulation map for the October 2021 event shows that two separated zones were affected by the heavy rains, and the study of the QPE over time revealed that both zones were not affected at the same time : the heavy rainfall hit the department of Lozère first, on 3 October, and then the departments of Bouches-du-Rhône and Var, on 4 October. It was therefore decided to separate this event into two distinct events, labelled D and H.

For most of the eight selected events, the larger hydrological responses occurred in small ungauged catchments. However, for events A, B, and C, post-event studies could estimate the maximum peak

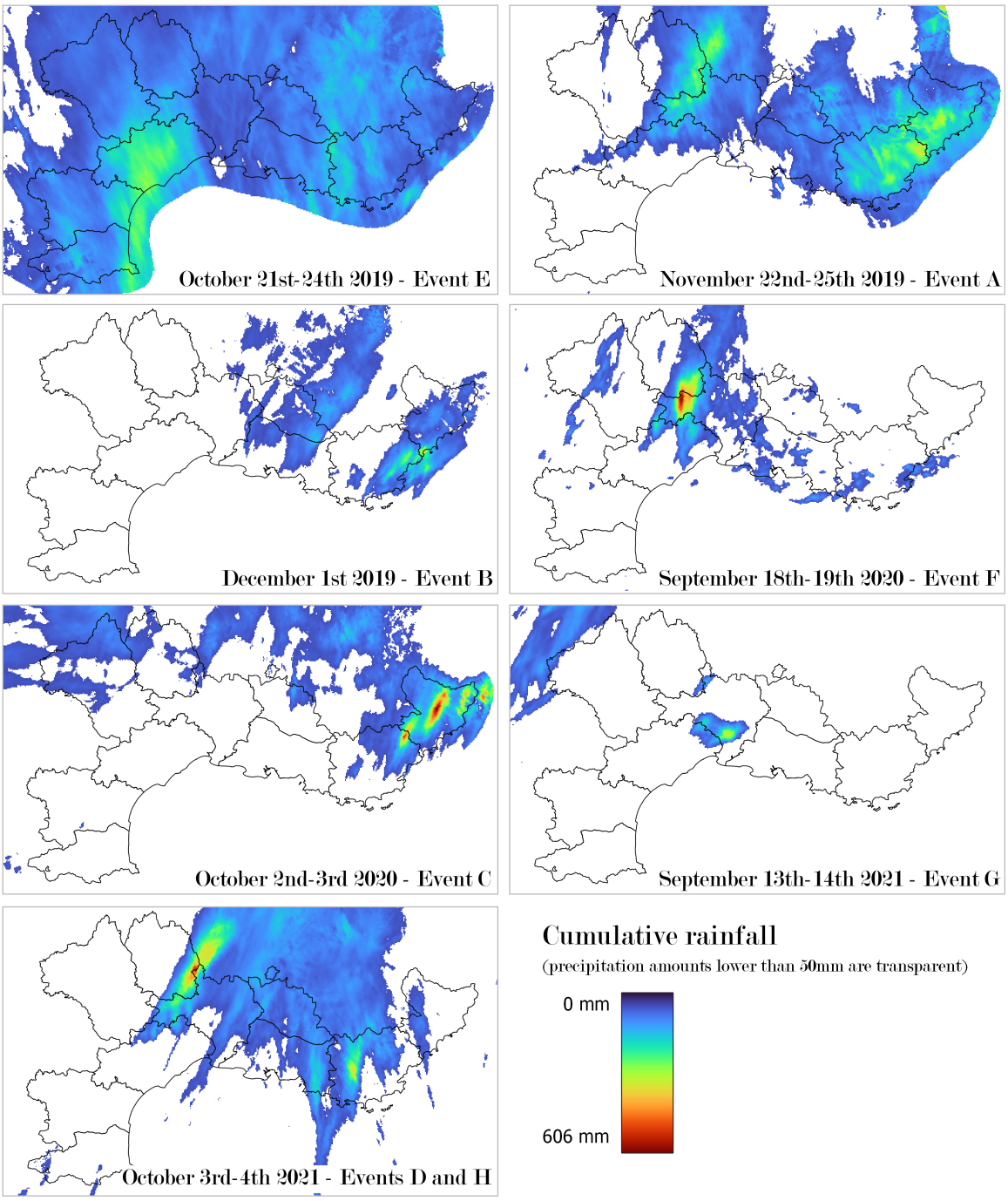


FIGURE VI.7 – Maps of rainfall accumulations for each of the selected events. These maps were drawn using the ANTILOPE QPE [Champeaux et al., 2009], i.e. the best reanalysed QPE merging radar estimations and rain gauge observations.

discharge values [Lebouc and Payrastre, 2020, Brigode et al., 2021, Payrastre et al., 2022]. For events A and B, it was estimated that peak discharges locally reached a $5 - 15\ m^3 s^{-1} km^{-2}$ range in the small basins hit by the larger rainfall accumulations. These two events are hydrologically interesting because they happened close in time and in the same area : the maximum peak discharges were probably observed during event B because of larger soil saturation and higher rainfall intensities observed on short time steps [Brigode et al., 2021]. For event C (Storm Alex) which hit the same department (Alpes-Maritimes), Payrastre et al. [2022] estimated that despite significantly higher rainfall accumulations, the peak discharges were globally similar to those observed during the November–December 2019 flash flood events, except on some upstream basins where the estimated peak discharges reached values in

Duration	Max. rainfall	$\frac{S_{Q \geq Q_{T=2y}}}{S_{Zone}} \times 100$	$\frac{S_{Q \geq Q_{T=5}}}{S_{Zone}} \times 100$	$\frac{S_{Q \geq Q_{T=10y}}}{S_{Zone}} \times 100$	Zone	Label
22-25/11/2019	380 mm	22.06	7.92	3.86	East	A
01/12/2019	360 mm	11.92	6.53	4.77	East	B
02-03/10/2020	600 mm	2.42	1.28	0.97	East	C
03-04/10/2021	370 mm	4.31	2.00	1.33	East	D
21-24/10/2019	300 mm	11.03	6.91	5.39	West	E
18-20/09/2020	590 mm	1.35	0.77	0.53	West	F
13-14/09/2021	330 mm	0.36	0.27	0.25	West	G
03-04/10/2021	540 mm	3.94	1.76	1.38	West	H

TABLEAU VI.1 – Description of the flash flood events : date, maximum rainfall accumulation (spatial maximum), percentages of the study area where the SMASH simulated peak discharge exceeds the 2-year, 5-year and 10-year discharge thresholds, affected zone (by reference to figure VI.6) and assigned label

the $15 - 20\ m^3 s^{-1} km^{-2}$ range. Considering these discharge values, events B and C are among the most intense floods which have been observed in the departments of Var and Alpes Maritimes.

Because of the limited information about the actual intensity and location of the flood responses during the eight selected events, it is difficult to more thoroughly compare the characteristics of the flood events. The comparison of simulation results obtained with the SMASH model can, nevertheless, bring additional information. Table VI.1 gives the percentage of the surface where several discharge thresholds (2-, 5-, and 10-year return periods) were exceeded by the reference simulation. It refers to Fig. E.1, which gives an idea of the extent of the flood responses exceeding these thresholds for each event. Again, we can see here that the considered flood events show very different characteristics in terms of spatial extent : from very localized events (F and G) up to more generalized flood responses (events A, B and E), according to the SMASH simulations.

1.3.2 Verification method

As mentioned in the Introduction, the objective is to assess the benefits of forcing the flash flood nowcasting chain with the PIAF-EPS forecasts. The whole river network of the study area, including small ungauged rivers, should be considered. As a consequence, the verification process has been applied at each 1 km^2 pixel of the SMASH model to provide as detailed an evaluation as possible. Since no discharge observation is generally available at this 1 km^2 scale, the discharge simulated by the SMASH hydrological model forced with the observed PANTHERE QPE was used as a reference. This reference also allows for ignoring the errors due to the hydrological model, the performance of which have not been assessed in this study. Therefore, in the following, Q_{sim} is assigned to the discharge calculated by the SMASH model forced with PANTHERE QPE, while Q_{for} is assigned to the discharge forecasts obtained by forcing the model with the different rainfall forecasts.

The verification process aims to evaluate if the exceedances by Q_{sim} of discharge thresholds Q_t, are well anticipated by Q_{for}, i.e. when rainfall forecasts are used as input of the chain, instead of the PANTHERE QPEs. Since Q_{for} may correspond either to ensemble forecasts (case of PIAF-EPS QPF) or to deterministic forecasts (case of other QPFs), we selected verification scores that can be applied on both deterministic and probabilistic forecasts.

The verification is based on the filling of contingency tables, which are commonly used for assessing the ability of forecasting systems to detect binary events [Mason, 1982]. The contingency tables are filled by comparing the reference (Q_{sim}) and forecasted (Q_{for}) discharges to a threshold (Q_t), resulting

in the four outcomes presented in table VI.2. Then, the probability of detection ($POD = \frac{a}{a+c}$) and the probability of false detection ($POFD = \frac{b}{b+d}$) can be calculated. In the case of ensemble forecasts, they can be plotted for different forecast probabilities to create a ROC curve [Mason, 1982].

	$Q_{for} \geq Q_t$	$Q_{for} < Q_t$
$Q_{sim} \geq Q_t$	HIT (a)	MISS (c)
$Q_{sim} < Q_t$	FALSE ALARM (b)	CORRECT REJECTION (d)

Tableau VI.2 – Content of a contingency table. a, b, c, and d correspond to the number of forecasts in each category.

Contingency tables are usually filled by combining a continuous temporal sequence of forecasts, at one single site and for one unique lead time. However, we followed here the principle proposed by Charpentier-Noyer et al. [2023] to build the contingency tables by aggregating the forecasts issued during the most critical phase of the event (i.e. the forecasts issued during the flood rising limb, just before the threshold exceedance by Q_{sim}, independently of the lead time). The detailed methodology is explained in appendix E, section 2.1. It includes a selection process of the forecasts considered to fill the contingency table, called stratification. Some adaptations have been introduced in this process in order to create a forecast-based stratification, rather than an observation-based stratification, following the recommendations of Bellier et al. [2017]. The details are provided in appendix E, section 2.2.

Here we considered the forecasts obtained for each pixel of the SMASH model to build the contingency tables. Three discharge thresholds were considered at each pixel : $Q_t = \{Q_{T=2y}, Q_{T=5y}, Q_{T=10y}\}$ (see section 1.2.4 and appendix E, section 1). The contingency tables obtained for each threshold may be visualised directly on a map (see figure VI.8), or summarized using synthetic scores such as the POD or POFD. However, the POFD score is sensitive to the extent of the verification area, which directly determines the number of correct rejections (see figure VI.8). The choice of the verification area was already identified as an important issue by Charpentier-Noyer et al. [2023] who suggested to pay a particular attention to the choice of the "HFA" (Hydrological Focus Area). For that reason, we chose to summary the contingency tables based on the Critical Success Index ($CSI = \frac{a}{a+b+c}$) instead of POD/POFDs. The CSI score does not take into account the correct rejections, thus is much less sensitive to the choice of the verification spatial window. This choice avoided the issue of defining an appropriate verification zone for each of the considered events.

It is important to mention that although the verification method is based on classical statistical scores, it cannot characterize the performances of the QPFs for long temporal series. We are only assessing here the ability of the different QPFs to correctly predict several specific events of high intensity. The obtained results, even if providing interesting information, cannot be extrapolated to future events because of the too limited number of events considered in this study.

1.4 Results and discussion

1.4.1 Maps of contingency tables : presentation on a result sample

As explained in section 1.3.2, the contingency tables filled for each event can be represented on maps. One map can be extracted for each discharge threshold $Q_t = \{Q_{T=2y}, Q_{T=5y}, Q_{T=10y}\}$, and each QPF product. In the case of the PIAF-EPS ensemble forecast, one map is obtained for each percentile of the forecast ensemble.

These maps allow for observing and comparingthe performance of the forecasts. For example, figure VI.8 shows the maps obtained in the case of event F and a threshold $Q_t = Q_{T=2y}$. For this event, the

constant rain and the AROME-NWC forecasts showed poor performance, with the first one correctly emitting warnings on the affected region but emitting many false alarms elsewhere, and the other one missing most of the area affected by the event. The CSI scores summarizing the content of the contingency tables (see section 1.3.2) are 0.12 and 0.13 respectively for these two forecasts. The improvements observed for the deterministic PIAF forecast (CSI=0.21) confirm that, in this case, the blending of radar QPFs and AROME-NWC was effective. Moreover, the 60% percentile of the PIAF-EPS forecast, which has the highest CSI score among the other percentiles, shows even better results than the PIAF forecast (CSI=0.27), by notably reducing the area affected by false alarms.

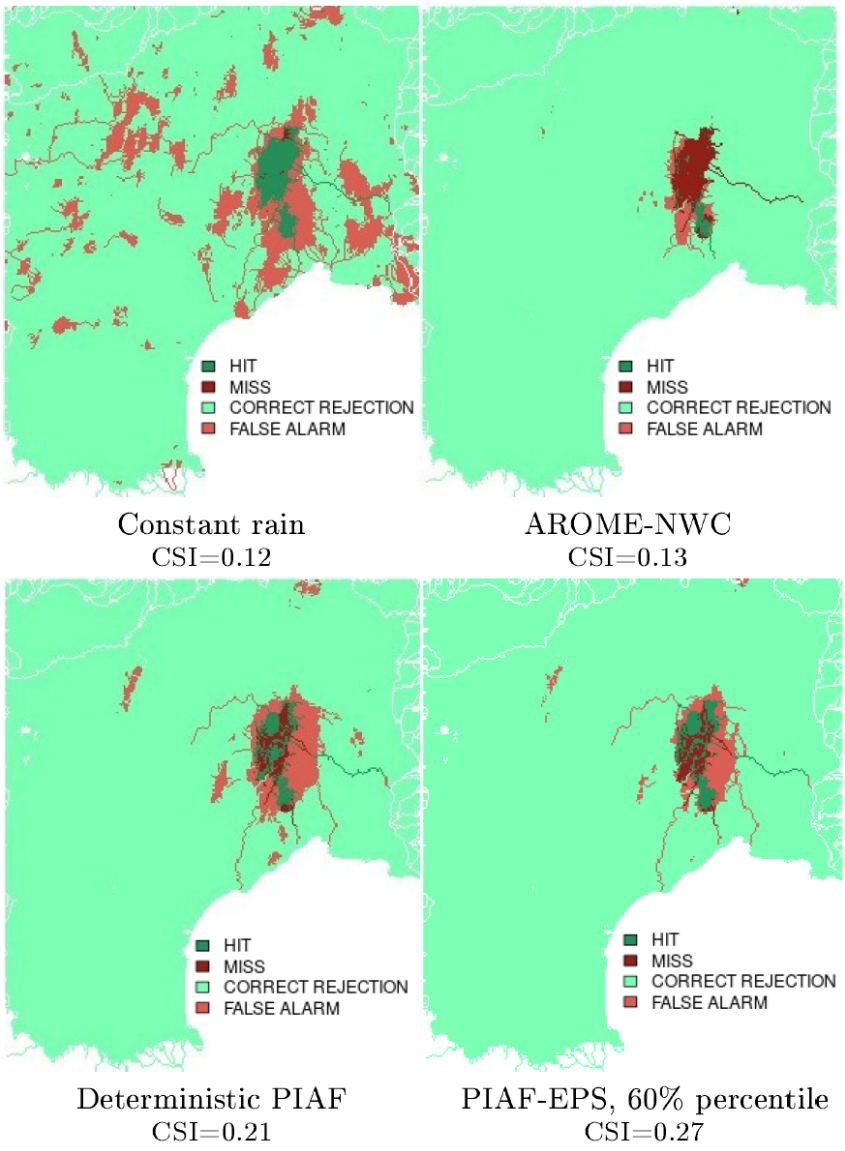


FIGURE VI.8 – Maps of contingency tables for each forecast product, for the T=2-year threshold (event F).

Figure VI.9 shows the maps of the contingency tables obtained for another single event (event B) and for the different PIAF-EPS percentiles. This figure illustrates the evolution of the forecast performance depending on the considered percentile of the ensemble forecast. Logically, for low percentiles, the number of correct detections is very small and much lower than the number of missed warnings, resulting in low values of CSI. For intermediate percentiles, the number of correct detections increases with respect to the number of missed warnings and false alarms, which results in an increase of the CSI values. However, for larger percentiles, the number of false alarms increases up to the point that it outweighs the increase

of correct detections, resulting in a decrease of the CSI scores.

This evolution of detection performance depending on the percentile of the ensemble forecasts directly explains the expected "hill" shape of the CSI curves presented on figure VI.10. The "best" ensemble percentiles can be identified by selecting the maximum on these curves. In the case of event B, figures VI.9 and VI.10 show an optimal HIT/MISS/FALSE ALARM balance for percentiles around 50-60%.

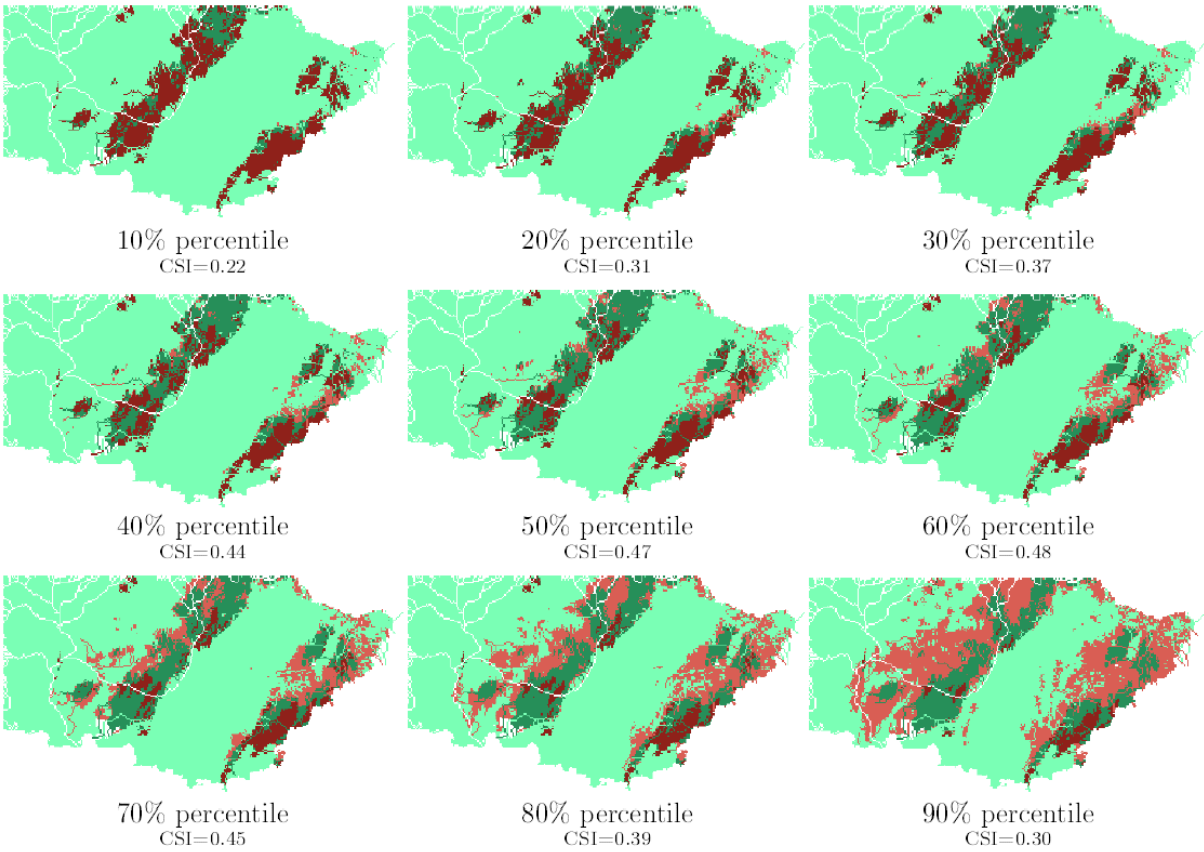


FIGURE VI.9 – Map results of contingency tables for the T=5-year threshold (event B).

1.4.2 Analysis of CSI scores at the event scale

Before comparing the different forecast approaches, two generic observations can be made on figure VI.10. First, the performance of all forecasts tends to decrease as the discharge thresholds increase. This is in agreement with the theory developed by Schaefer [1990], according to which the CSI is biased by the frequency of the forecast event. Typically, the rarer an event is, the higher its return period is, and the lower is the CSI. Here it leads to a CSI decrease of 0.1 to 0.2, when comparing the 2-year and 10-year discharge thresholds. Second, the CSI curves for the PIAF-EPS forecasts do not always have the expected "hill" shape with a maximum at intermediate percentiles. Especially, for events D, E, G and H, the maximum CSI is reached at low percentiles. It means that, for those events, PIAF-EPS had a tendency to over-estimate the discharge probabilities. This is consistent with the rank diagrams plotted for each event in appendix E, section 3 (figure E.3), which generally show a slight positive discharge bias for the same events (D,E,G,H). A larger bias can even be observed for the forecast discharges exceeding the 2-year return period threshold, even if in this specific case the bias can be increased by the stratification process (see appendix E, section 3).

Furthermore, figure VI.10 shows that the intermediate percentiles of PIAF-EPS (i.e. 40% to 60%)

outperform almost systematically the naive forecast approach (constant rain), and the deterministic PIAF forecast. This confirms that adding spatial and amplitude perturbations to the PIAF QPFs, to obtain the PIAF-EPS ensemble QPF product, resulted in better performances of the flash-flood forecasts, at least for the 8 intense flash-floods considered in this study.

The results are more mixed concerning AROME-NWC. For 5 out of 8 events, AROME-NWC shows equivalent or lower CSI values than the deterministic PIAF, and it is outperformed by the PIAF-EPS forecasts in these cases, at least for intermediate percentiles. Conversely, there are 3 events for which AROME-NWC leads to significantly higher CSI values than deterministic PIAF (events E, G and H). For two of these three events (events E and G), PIAF-EPS largely compensates the poor performance of deterministic PIAF, leading to CSI values that are similar to AROME-NWC (events E and G). Lovat et al. [2022] have already shown that, depending on the lead time, AROME-NWC can outperform deterministic PIAF forecasts. The relationship between the CSI values and the lead times can hardly be investigated in this work, as the verification method applied looks at all the forecast runs emitted in a specific time window, regardless of their respective lead times. An explanation for the poor performance of deterministic PIAF on these events could be a sudden stationarization of the rain cells. In such a situation, the Lagrangian radar QPE extrapolation becomes a very poor rain predictor, because it cannot account for rapid changes in the speed of high precipitation areas. The results obtained here suggest that PIAF-EPS can at least partly handle the inherent uncertainty of these situations where the blending with radar QPE extrapolations limits the quality of the deterministic PIAF forecast.

1.4.3 Averaged CSI scores

Averaged CSI scores were also calculated in order to provide a more synthetic view of the forecast performance, by aggregating them over all studied events. There are two possible ways of calculating global CSI scores :

$CSI_1 = \frac{1}{N_{evt}} \sum_{n=1}^{N_{evt}} \frac{a_n}{a_n+b_n+c_n}$ and $CSI_2 = \frac{\sum_{n=1}^{N_{evt}} a_n}{\sum_{n=1}^{N_{evt}} a_n+b_n+c_n}$

Since the studied events have very different spatial extents (see figure E.1 and table VI.1), we chose to use the CSI_1 formula where the averaging implies that all the events have the same weight. CSI_2 would have given much more relative weight to the large-scale events.

The global CSI curves obtained from CSI_1 are presented in figure VI.11 and indicate that globally, PIAF-EPS 30-60% percentiles outperform all the deterministic reference forecasts. This identification of "best" percentiles can be useful for end-users [WMO, 2012], particularly if they remain relatively stable depending the considered events. In this paper, we assess the value of these "best" percentiles in a slightly overoptimistic way, since we use all events to derive the optimum CSI values (i.e. it is not an "out of sample" optimization), because our sample is very small. This methodological weakness does not invalidate our conclusions, however, since the CSI curves are rather stable from case to case in our sample. The global CSI_1 scores also confirm that AROME-NWC globally performs better than PIAF for the eight studied events. However, this conclusion is largely influenced by events G and H, for which the CSI differences between AROME-NWC and deterministic PIAF are the largest.

Finally, the averaged CSI values obtained with the intermediate percentiles (40%-50%) of PIAF-EPS are in the 0.3-0.4 range, depending on the return period of the discharge threshold considered. The CSI scores can reach up to 0.6 for some specific events. These CSI values may appear relatively low in the perspective of operational decision making. However, the real added value of these forecasts for decision making can only be evaluated by considering the balance between the gains associated with the hits, and the costs related to false alarms. Moreover, other studies dealing with flash-flood nowcasting found

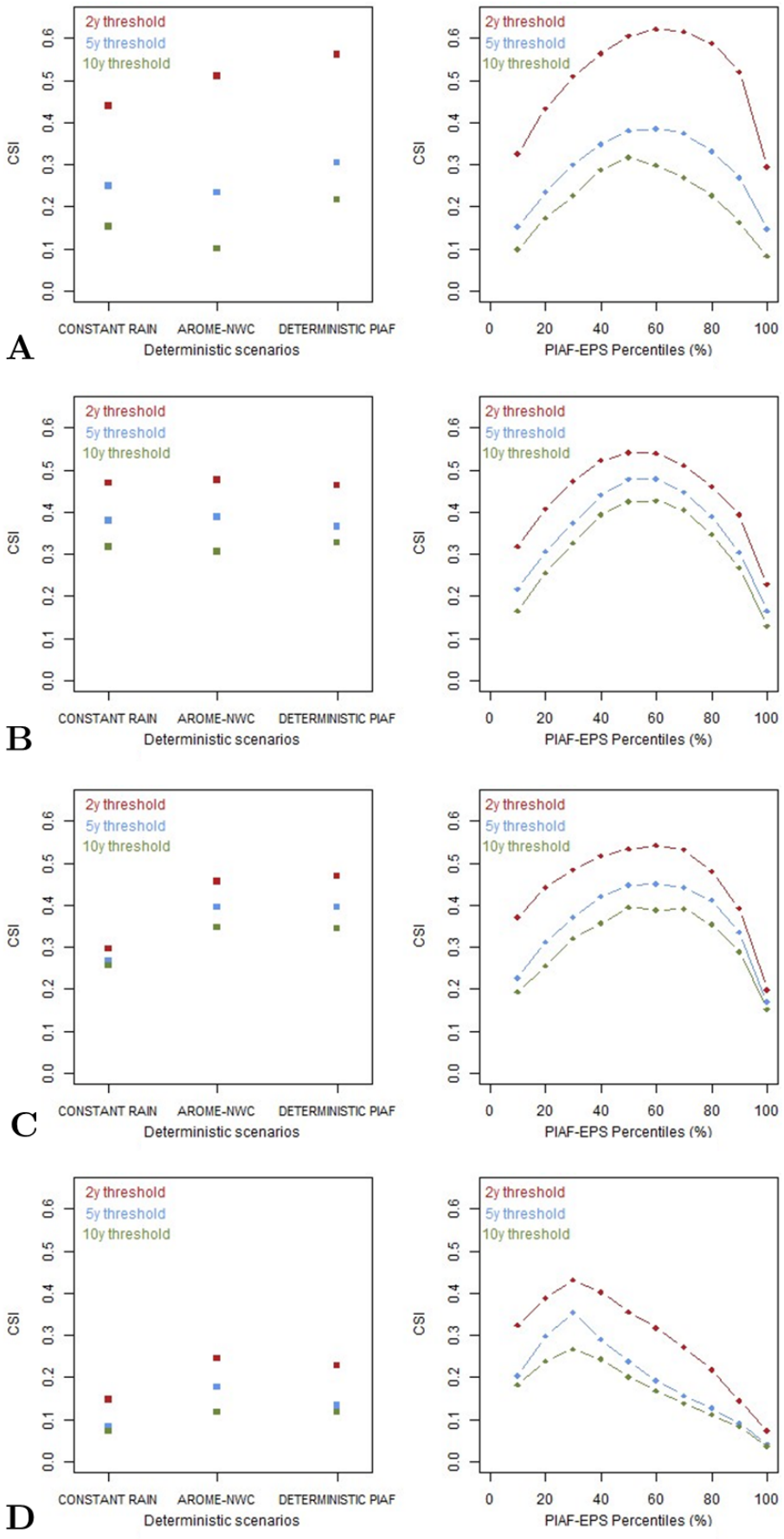


FIGURE VI.10 – CSI curves of each event for the various forecast products and thresholds.

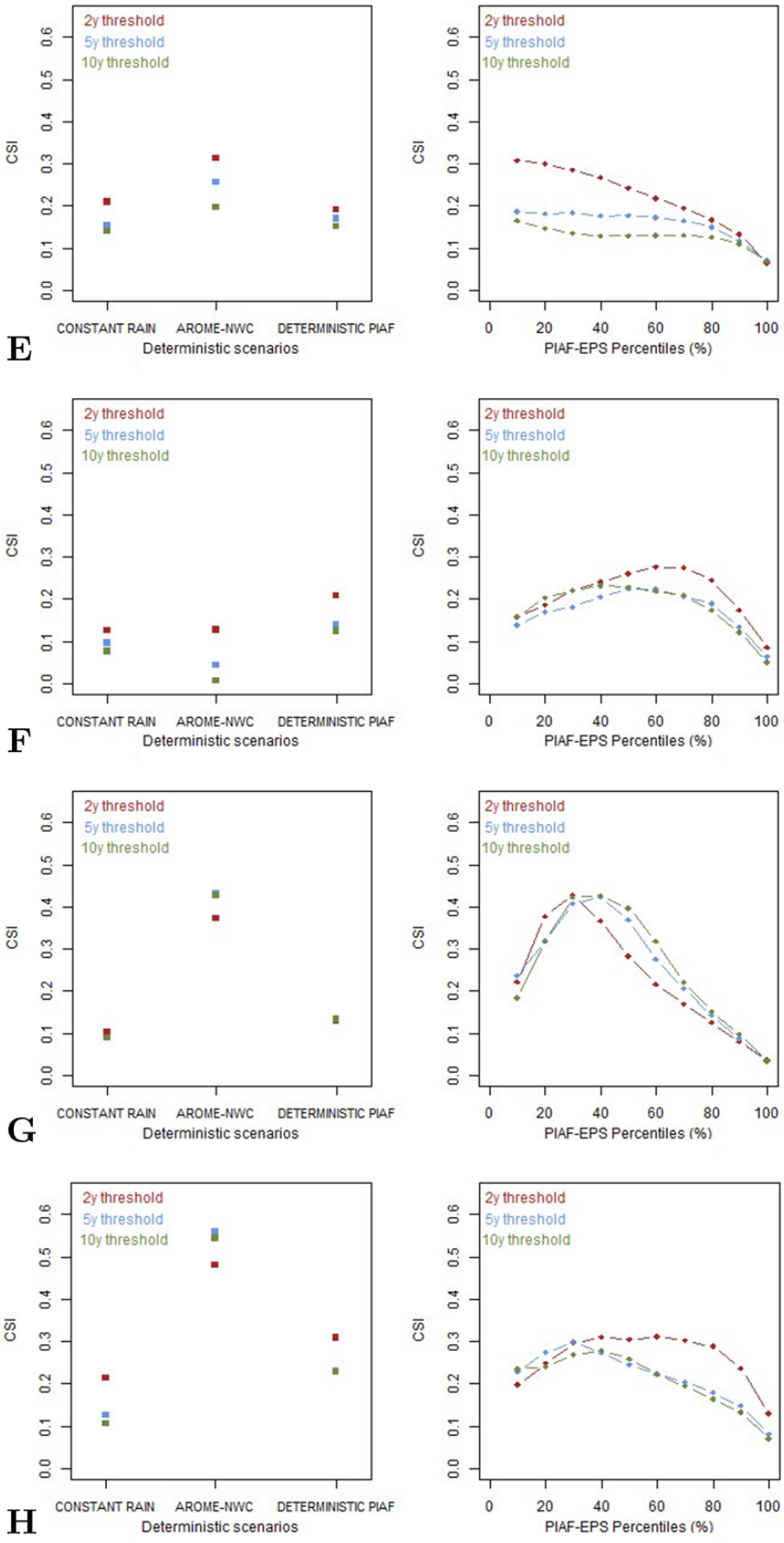


FIGURE VI.10 – CSI curves of each event for the various forecast products and thresholds.

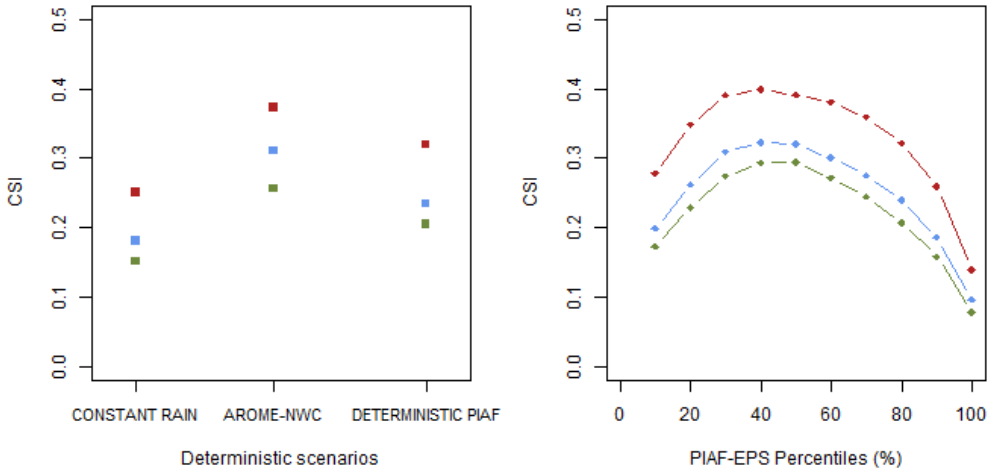


FIGURE VI.11 – Averaged CSI scores calculated for the eight flood events.

similar CSI values of 0.20 [Clark et al., 2014] and 0.38 [Gourley et al., 2017], even though these CSI values were obtained in very different contexts (using observations over the whole U.S.), and thus cannot be directly compared with the values of this paper.

1.4.4 Anticipation lead times

The CSI scores presented above assess the ability of the rainfall products to predict discharge threshold exceedances, regardless of anticipation. However, maximizing the anticipation times of good forecasts is another desirable property in an operational forecasting context. An estimation of the anticipation times associated with the HITs in contingency tables was proposed by Charpentier-Noyer et al. [2023] by computing the difference $t_{sim} - t_{run}$, where t_{sim} represents the first threshold exceedance by the reference simulation, and t_{run} corresponds to the starting time of the first forecast that identifies this threshold exceedance event. Anticipation times for each QPF (50th percentile for PIAF-EPS, as intermediate percentiles where identified as optimal in the previous section) and for the 5-year threshold are presented in figure VI.12. Firstly, the results show that the anticipation times can reach up to 6 hours. This is due to the choice of counting a HIT when t_{sim} falls within the interval $]t_{run}; t_{run} + T + 3h]$ (see appendix E, section 2.2), T being the forecast range ($0 < T \leq 3$ hours). Anticipation times exceeding the forecast length of 3 hours, even if helpful in anticipating threshold exceedances, result from unrealistic forecasts where the threshold crossing is forecasted too early. It is thus logical to observe that the constant rain scenario has the highest number of anticipation times exceeding 3 hours, and it is rather satisfying to note that PIAF-EPS has the least occurrences in this anticipation range. The comparison of histograms in the 0-3h range of anticipation times confirms that PIAF and PIAF-EPS yield a larger number of HITs globally. Additionally, it shows that this increase of HITs is primarily obtained in the 0-2h range of anticipation times compared to AROME-NWC. This logic is clear as it corresponds to the forecast range where radar extrapolations are involved in building PIAF and PIAF-EPS. Furthermore, it suggests that PIAF-EPS brings additional HITs mainly in the 0-1h range of anticipation times when compared to PIAF. However, drawing systematic conclusions is complicated, as we are only examining one ensemble percentile here, and we are considering all events, while important differences may exist within each event.

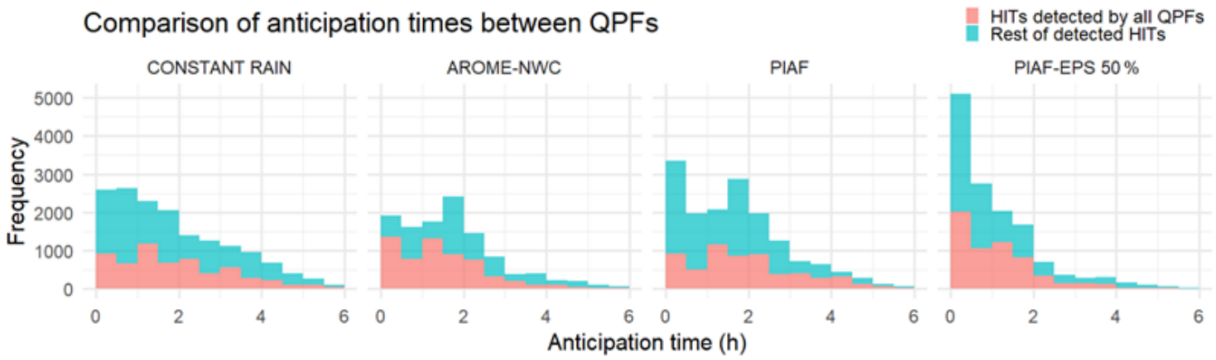


FIGURE VI.12 – Anticipation times aggregated for all events and for each QPF (50th percentile for PIAF-EPS) for the 5-year threshold

1.5 Conclusions

The development of efficient tools and methods for flash-flood forecasting is of crucial importance to limit the often catastrophic consequences of flood hazards. The objective verification of newly developed forecasting methods and products is a key step before their integration into operational forecasting suites. In the current study, the potential of the experimental PIAF-EPS short-range ensemble rainfall product for flash flood forecasting purposes has been assessed. Eight heavy precipitation events that occurred between 2019 and 2021 in the South-Eastern region of France were reanalysed using a hydrological forecasting suite similar to the one that is currently operational in the French national flash flood warning system, Vigicrues Flash. An original verification process, directly derived from Charpentier-Noyer et al. [2023], was performed on each $1 km^2$ pixel of the area. This allowed us to plot maps to precisely visualize the forecasts performance, and to summarize it as CSI scores.

The hydrological forecasts based on PIAF-EPS have been compared to those obtained with deterministic PIAF and AROME-NWC rainfall forecasts, since PIAF-EPS is directly obtained from these two deterministic products. A naive constant rainfall scenario was also used as a reference. The results showed that PIAF-EPS systematically outperformed the constant rainfall and the deterministic PIAF forecasts. As indicated in previous studies, it was also observed that PIAF does not always outperform AROME-NWC, because the forecast quality depends on the lead time and on the performance of radar QPE extrapolations. Over the eight events considered in this study, it was observed that the PIAF-EPS performance is generally similar to, or better than AROME-NWC.

In a nutshell, the results obtained confirm the added value of using the PIAF-EPS products for anticipating flash floods in the Mediterranean area. We argue that statistical scores such as the CSI provide valuable indications of performance despite not being applied on long data series, but on only eight particularly intense flash-floods. Indeed, when assessing the performance of such a new forecasting product, it is essential to carefully check its behaviour on some high-impact events, as a complement to more generic statistical evaluations. The results presented here should nevertheless be complemented with more robust statistical evaluations over longer periods of time and on a larger number of high precipitation events, bringing a more generic overview of the quality of the forecast ensembles.

2 Mobilisation de prévisions d'ensemble de pluies en entrée d'une chaîne de type IBW

La section précédente s'est focalisée sur l'évaluation de la plus-value de prévisions d'ensemble en entrée d'une chaîne d'avertissement basée sur l'aléa (HBW). Dans cette section, les prévisions d'ensemble sont utilisées pour alimenter la chaîne d'avertissement basée sur les impacts (IBW), de façon à illustrer la propagation des incertitudes de prévision de pluie sur les impacts estimés.

La première partie de cette section illustre la production de prévisions d'impacts probabilistes à l'échelle d'un événement significatif - les inondations du 1er décembre 2019 sur le bassin du Riou de l'Argentière - et ouvre la discussion sur les problématiques liées à la représentation des prévisions d'ensemble dans une chaîne de type IBW. La seconde partie propose des premiers éléments de méthode et de discussion sur la propagation des incertitudes entre les différents blocs de modélisation de la chaîne IBW, toujours à l'échelle de ce même événement. Ces deux thématiques ont été initiées lors d'un séjour de recherche effectué à Earth Sciences New Zealand, à partir d'un cas d'étude néo-zélandais. Le travail réalisé durant ce séjour, présenté en annexe F, sera régulièrement évoqué à titre de comparaison et de complément.

2.1 Mise en œuvre de la chaîne IBW à l'échelle de l'événement du 1er décembre 2019

Cette section porte sur l'épisode pluvieux du 1er décembre 2019 dans les Alpes-Maritimes, identifié comme l'événement B dans la section 1.3. Selon les retours d'expérience réalisés pour cet épisode [Brigode et al., 2021, Lebouc and Payrastre, 2020], il se caractérise par des intensités très élevées : plus de 100 mm de pluie en moins de trois heures et plus de 270 mm en 48 h. Les débits spécifiques ont localement atteint 10 à 15 $m^3/s/km^2$, ce qui place cet épisode parmi les plus marquants de la décennie 2010-2020, à un niveau comparable à celui des crues d'octobre 2015 dans la même région.

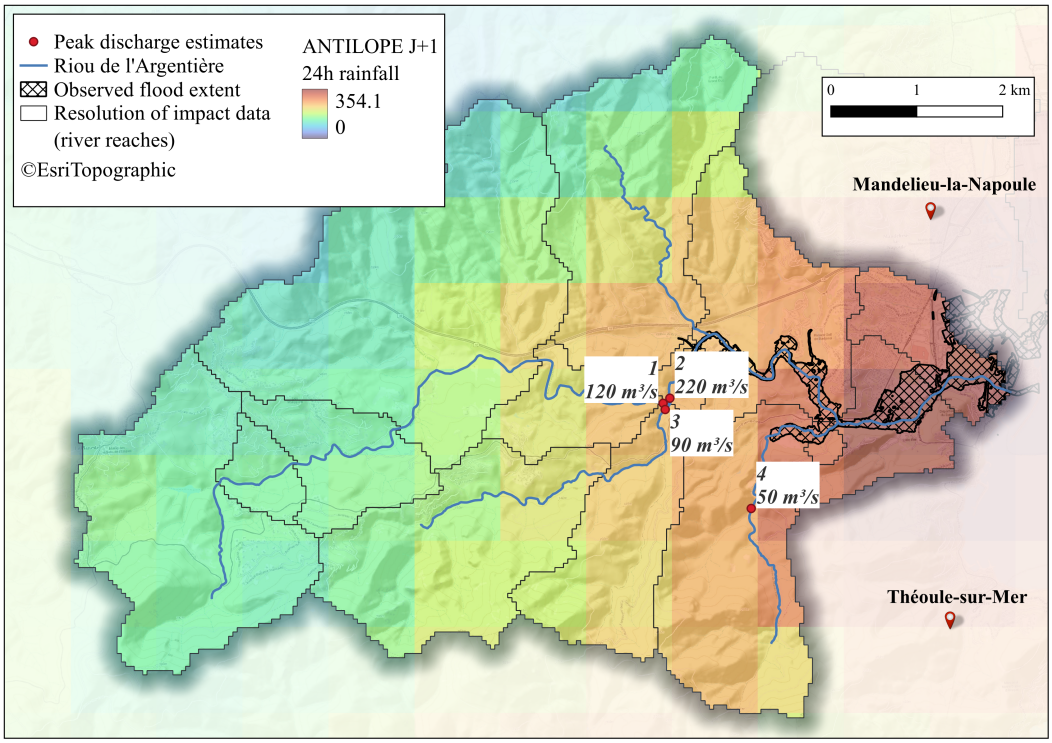


FIGURE VI.13 – Présentation de l'événement du 1er décembre 2019 sur le bassin du Riou de l'Argentière ($50km^2$) et des données de validation associées.

Cet épisode est survenu une semaine après un premier événement pluvieux intense (22-25 novembre 2019, événement A dans la section 1.3). Le bilan cumulé de ces deux événements est lourd : 13 décès, environ 400 M€ de dommages assurés (en EUR 2024), plus de 2000 interventions des services de secours, 170 hélitreuillages et 1600 personnes évacuées, d'après la CCR.

La proximité temporelle des deux événements rend difficile leur distinction dans les données collectées sur le terrain (par ex. les débits de pointe ou les données d'impacts), ce qui complique l'évaluation des prévision émises lors de ces deux événements. Pour contourner ce problème, l'analyse qui suit se concentre sur le bassin du Riou de l'Argentière (figure VI.13), pour lequel les rapports de Lebouc and Payrastre [2020] permettent de justifier que les cumuls maximaux de pluie, les débits de pointe, l'extension maximale de l'inondation et les impacts associés correspondent à l'événement du 1er décembre 2019.

Cet événement a été rejoué par la chaîne de prévision immédiate basée sur les impacts (chapitre V), mais en utilisant cette fois les prévisions d'ensemble PIAF (section 1.2) en entrée, plutôt que des observations de pluie. Cela permet d'accroître le délai d'anticipation (jusqu'à trois heures) et de générer non pas une prévision unique, mais plusieurs scénarios (17 membres au sein de l'ensemble). Les membres de l'ensemble peuvent être utilisés pour calculer des quantiles sur les variables de sortie, ou encore pour obtenir des probabilités de dépassement de seuils (de débits, de hauteur d'eau, ou d'impacts).

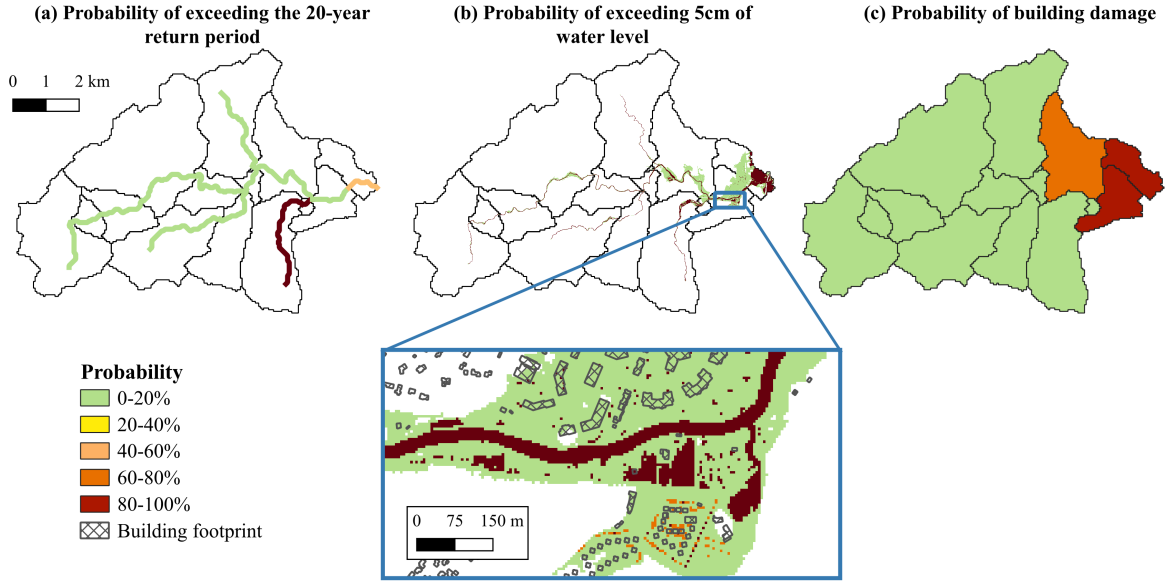


FIGURE VI.14 – Exemples de résultats disponibles à différents niveaux de la chaîne de prévision probabiliste pour le run lancé le 1er décembre à 17h : (a) prévision hydrologique : probabilité que le quantile de débit 20 ans soit dépassé sur le tronçon, (b) prévision hydraulique : probabilité que le seuil de hauteur d'eau 5cm soit dépassé sur le pixel, (c) prévision d'impacts : probabilité que le seuil de 1 bâtiment inondé soit dépassé sur le sous-bassin versant.

La figure VI.14 illustre plusieurs sorties possibles de cette chaîne de prévision : cartes de probabilité de dépassement d'un seuil de période de retour par le débit de pointe, d'un seuil de hauteur d'eau, ou d'un seuil de nombre de bâtiments impactés. Ces résultats soulignent la diversité des représentations possibles et la richesse de l'information apportée par la chaine d'avertissement IBW. Le premier constat prolonge la problématique abordée au chapitre précédent : les cartes VI.14a et VI.14c, l'une basée sur l'intensité de l'aléa et l'autre sur les impacts attendus, ne mettent pas en évidence les mêmes zones sensibles. À l'échelle du bassin, la carte d'inondation (probabilité de dépasser 5 cm d'eau) paraît moins pertinente : les secteurs les plus touchés ressortent mal et l'information est difficile à agréger. En revanche, elle peut s'avérer utile à une échelle plus fine, lorsque l'attention se porte sur des zones particulières.

2.2 Analyse de la propagation des incertitudes de prévision de pluie au sein de la chaîne IBW

Dans ce travail de thèse, la prévision des impacts d'inondation repose sur le chaînage de plusieurs blocs de modélisation (voir Chapitre IV). Ces blocs ont été évalués séparément (bloc hydrologique au Chapitre II, bloc hydraulique au Chapitre III) ainsi qu'en sortie de la chaîne de prévision immédiate des impacts (Chapitre V). Cependant, au-delà de la vision modulaire, la chaîne forme un système complexe où les incertitudes se propagent d'un bloc à l'autre. La propagation des incertitudes dans une chaîne de prévision des impacts d'inondation ne semble pas avoir été abordée à ce jour dans la littérature.

2.2.1 Application dans des travaux antérieurs

Titley et al. [2024] proposent une première approche en analysant les corrélations entre erreurs des différents blocs d'une chaîne de prévision appliquée à l'ouragan Iota (Amérique centrale, novembre 2020). Leur chaîne s'arrêtait toutefois au calcul des débits, sans aller jusqu'aux impacts. La figure VI.15 illustre la méthodologie utilisée : des scores d'erreurs sont calculés à différents stades de la chaîne de prévision probabiliste des crues (trajectoire du cyclone, intensité du cyclone, précipitations et débits prévus), pour chaque membre de l'ensemble des pluies (51 membres, horizon de 10 jours). Ces scores sont ensuite comparés deux-à-deux afin d'examiner les relations entre les incertitudes associées aux différents blocs de la chaîne. Les nuages de points obtenus et les coefficients de corrélation calculés permettent de visualiser et de quantifier la manière dont les incertitudes d'un bloc peuvent se propager dans un autre. Les résultats de cette étude ont notamment mis en évidence une forte corrélation entre les incertitudes liées aux précipitations (erreur moyenne sur le cumul total) et celles associées aux débits (erreur évaluée par un score KGE), malgré une dispersion importante.

L'objectif de l'analyse proposée ici n'est donc pas d'évaluer les sorties de la chaînes de prévision probabilistes en se basant sur les critères classiques d'évaluation des prévisions d'ensemble - e.g., le CRPS [Hersbach, 2000], le Brier score [Brier, 1950], ou le diagramme de rang [Hamill, 2001]. Ces scores ont en effet peu de sens lorsqu'il sont appliqués sur un événement unique.

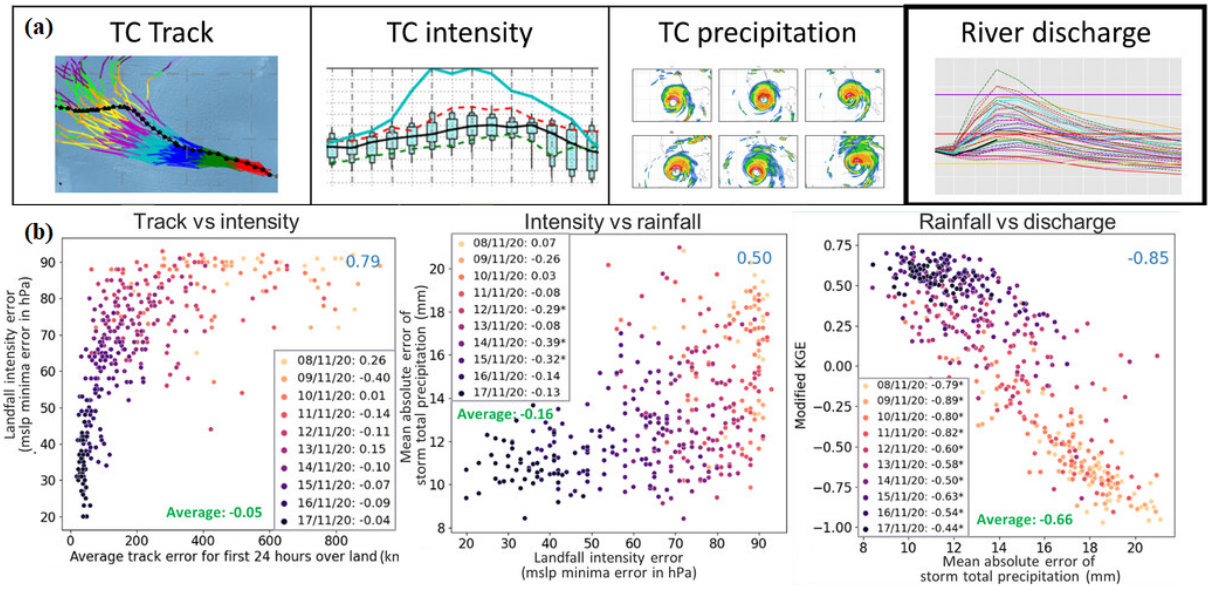


FIGURE VI.15 – Figures extraites de Titley et al. [2024]. (a) Composants de la chaîne de prévision étudiée : trajectoire et intensité du cyclone, précipitation et débits. (b) Relations entre les variables d'erreurs des composants de la chaîne de prévision.

Le séjour effectué en Nouvelle-Zélande au cours de la thèse a permis d'explorer le potentiel de l'approche de Titley et al. [2024] pour examiner la propagation des erreurs dans une chaîne de prévision de type IBW. Pour cela, une chaîne IBW pluie-débit-inondation-impact a été appliquée sur le cas des inondations de juillet 2021 dans la ville de Westport (annexe F). La chaîne IBW de ce cas d'étude est très similaire, dans sa structure, à celle développée dans le cadre de cette thèse, mais les échelles d'application diffèrent fortement : dans le cas français, l'échelle considérée est celle de l'ensemble de l'arc méditerranéen, avec une modélisation hydraulique régionale et simplifiée ; dans le cas néo-zélandais, la modélisation hydraulique avait été développée spécifiquement à l'échelle de la ville de Westport, et les événements de calibration incluaient notamment l'inondation de juillet 2021. Des scores d'erreurs avaient été calculés pour chaque membre de l'ensemble de pluies considéré (18 membres, horizon de 5 jours), et pour chaque bloc de modélisation de la chaîne.

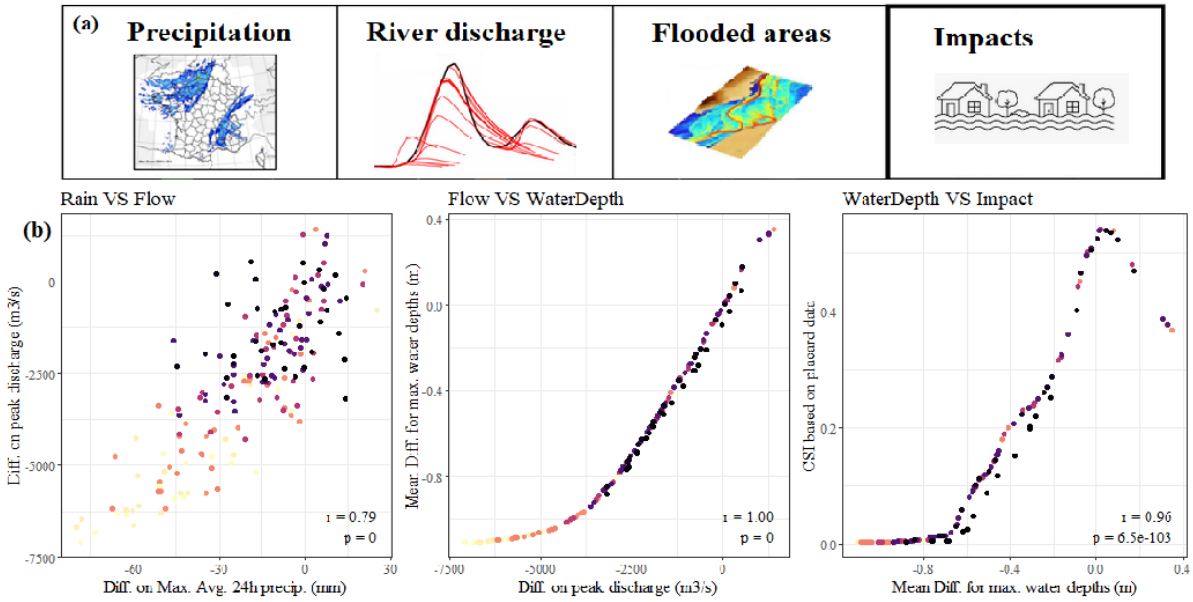


FIGURE VI.16 – (a) Composants de la chaîne de prévision IBW. (b) Relations entre les variables d'erreurs des composants de la chaîne IBW appliquée en Nouvelle-Zélande (annexe F).

Les résultats présentés en figure VI.16 ont apporté un éclairage intéressant sur la comparaison entre approches HBW et IBW : les blocs de modélisation hydraulique et de représentation des impacts semblent introduire très peu d'incertitudes supplémentaires, ce qui constitue un nouvel argument en faveur du développement d'approches de type IBW. L'objectif des sections suivantes est donc de réitérer cette analyse sur le cas des inondations du 1er décembre 2019 sur le bassin du Riou de l'Argentière, et d'examiner si des résultats similaires peuvent être obtenus, malgré des contextes différents.

2.2.2 Méthodes et données d'évaluation sur le cas d'étude du 1er décembre 2019

Les données d'évaluation disponibles pour chacun des blocs de la chaîne IBW étant différentes entre les cas d'études néo-zélandais et français, les scores d'erreur ont dû être adaptés. Cette section présente ainsi les modalités de calcul des erreurs associées à chaque entrée et sortie des blocs de modélisation, depuis la pluie jusqu'aux impacts, appliquées à l'événement de décembre 2019 sur le bassin du Riou de l'Argentière.

2.2.2.1 Erreurs sur les pluies. Les données radar ANTILOPE J+1 [Laurentin, 2008] servent de référence pour la pluviométrie. Afin de correspondre aux prévisions PIAF (échéance maximale : 3 h), des

cumuls sur trois heures sont calculés pour le 1er décembre 2019 à 17 h, 18 h, 19 h et 20 h, pour chacun des 17 membres de l'ensemble. Ces cumuls sont calculés sur les bassins versants amont des points 1 à 4 de la figure VI.13, correspondant aux localisations de débits de pointe estimés lors de la campagne HyMeX [Ducrocq et al., 2019, Lebouc and Payrastre, 2020].

2.2.2.2 Erreurs sur les débits. Ces débits de pointe de référence sont comparés aux débits maximaux simulés par les runs de prévision lancés aux mêmes heures. Cette comparaison n'est toutefois pas directe : comme détaillé au Chapitre II, section 3.1, le module hydrologique surestime les pics crues. Pour corriger ce biais, un rejeu du modèle sur une longue chronique a permis de relier les débits simulés à des périodes de retour (via les quantiles de crue calés sur les simulations, section 3.5.3). Ces périodes de retour sont ensuite converties en débits plus réalistes grâce aux quantiles calés sur les débits observés en station (section 3.5.2).

2.2.2.3 Erreurs sur les emprises inondées. Chaque prévision fournit une emprise inondée et des hauteurs d'eau maximales sur les trois heures suivantes. En l'absence de repères de crue pour cet événement (contrairement aux cas étudiés au Chapitre III, section 2.3.2), les contours d'inondation observés provenant des relevés terrain coordonnés par la DDTM 06 (figure VI.13) ont été utilisés. Le score d'erreur retenu est le Critical Success Index, calculé entre les emprise prédites et observées. Ce score est calculé sur les tronçons dont le débit est connu (tronçons 2 et 4 sur la figure VI.13), de façon à permettre la comparaison avec les erreurs sur les débits. Il est également calculé globalement sur la zone d'étude (en comparant avec le débit à l'exutoire du bassin, approché en cumulant les débits aux points 2 et 4, considérés comme les principaux contributeurs des débordements avals), notamment pour permettre la comparaison avec les scores d'erreur sur les impacts.

2.2.2.4 Erreurs sur les impacts. L'évaluation du bloc impacts de la chaîne est particulièrement délicate (voir Chapitre V). Plusieurs sources d'information existent pour cet événement [Dias, 2024] : interventions du SDIS, données VISOV, photothèque de la DDTM 06, arrêtés CatNat par commune et sinistralité CCR par tronçon de cours d'eau. Les premières sources, bien que précises spatialement, ne sont pas exhaustives. La base iCatNat, qui fournit un décompte de bâtiments impactés par commune, n'est pas adaptée non plus car le bassin du Riou de l'Argentière recoupe plusieurs communes également affectées par d'autres cours d'eau (Fréjus, Mandelieu-la-Napoule). L'évaluation s'appuie donc sur les données CCR. La liste des tronçons impactés (i.e. au moins un bâtiment est exposé à plus de 15 cm d'eau selon le modèle) est comparée à celle issue de la CCR, et un "CSI-impacts" est calculé sur l'ensemble des 14 tronçons du bassin. Compte tenu de la résolution des données, ce calcul est réalisé à l'échelle du bassin entier et comparé au "CSI-inondation" basé sur l'emprise inondée observée (hachures noires, figure VI.13).

2.2.3 Résultats sur le cas d'étude du 1er décembre 2019

Les scores décrits dans la section précédente ont été calculés et les relations entre les variables d'erreurs des différents composants de la chaîne de prévision sont présentés en figure VI.17.

2.2.3.1 Erreurs pluie/débit. Le premier graphique de la figure VI.17 illustre les incertitudes liées à la modélisation pluie-débit. Celles-ci sont importantes, mais restent cohérentes avec l'utilisation d'un modèle pluie-débit appliqué à l'échelle régionale, sans calibration spécifique pour le bassin du Riou de l'Argentière. La corrélation entre les erreurs sur le débit de pointe et celles sur la pluie (pluie moyenne

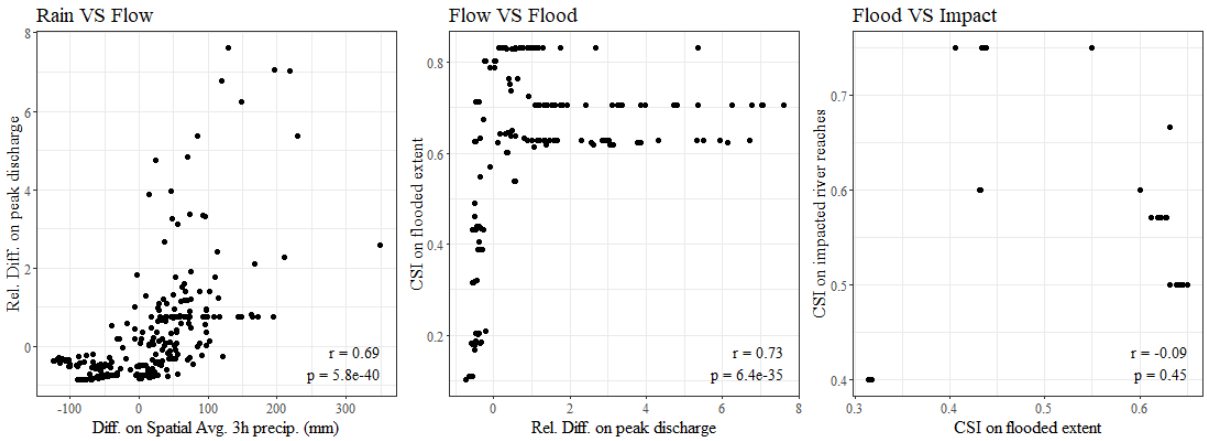


FIGURE VI.17 – Relations entre les variables d'erreurs des composants de la chaîne de prévision. L'erreur Pluie correspond à la différence (Sim–Obs) du cumul sur 3 h (pluie moyennée sur le bassin). L'erreur Débit correspond à la différence relative du débit de pointe aux points étudiés. L'erreur Inondation est un CSI basé sur l'emprise inondée observée. L'erreur Impact est un CSI à l'échelle du tronçon de cours d'eau. Un point correspond à un membre de l'ensemble et un run de prévision.

sur le bassin, cumul sur 3 h) est toutefois marquée, avec un coefficient de Spearman $r \simeq 0.7$. Cette valeur est plus faible que celles rapportées par d'autres études : $r \simeq 0.85$ chez Titley et al. [2024], obtenus en comparant le KGE [Gupta et al., 2009] appliqué aux prévisions de débits du modèle global GloFAS, aux erreurs moyennes sur les cumuls de pluie ; et $r \simeq 0.79$ dans le cas d'étude néo-zélandais (annexe F, section 4), en comparant les écarts absolus entre débits de pointe simulés (modèle hydrologique TopNet Clark et al., 2008, utilisé dans le système national de prévision [Cattoën et al., 2022]) et observés, aux erreurs moyennes de pluies cumulées sur 24 h. Les indicateurs d'erreur diffèrent donc légèrement entre ces cas, et une différence notable du cas français est que les débits ne sont pas simulés directement par le modèle hydrologique de la chaîne de prévision, mais reconstitués à partir de la période de retour estimée, ce qui introduit une source supplémentaire d'incertitude.

Il faut également rappeler que si le débit de pointe est sensible à la pluviométrie, d'autres facteurs que le cumul moyen sur 3h influent sur ce débit, notamment le cumul total sur l'événement, la position du pic d'intensité dans l'événement, la répartition spatiale des pluies, les antécédents pluviométriques et le niveau de saturation initial des sols, etc.

2.2.3.2 Erreurs débit/inondation. Le deuxième graphique de la figure VI.17 illustre la relation entre les incertitudes sur les débits et celles sur l'étendue de l'inondation. On observe que les scores de CSI liés à l'inondation peuvent être faibles, voire très faibles, lorsque les débits sont sous-estimés, et dans ce cas la corrélation est forte. En revanche, ils se maintiennent au-dessus de 0.60 (valeur considérée comme une bonne performance pour un modèle d'inondation) lorsque les débits sont correctement estimés ou sur-estimés. Les scores de CSI présentent un effet de plafonnement à trois valeurs : 0.63, 0.71 et 0.83. Ces seuils correspondent aux différentes zones de calcul décrites en section 2.2.1.2 : 0.63 pour le tronçon situé à l'aval de la confluence, 0.71 pour l'affluent en rive gauche, et 0.83 pour l'affluent en rive droite. Ces valeurs maximales sont atteintes pour une période de retour de $T = 1000$ ans sur chacun de ces tronçons, soit le scénario le plus extrême de la bibliothèque disponible. Ce phénomène reflète donc une situation où l'inondation a vraisemblablement atteint les limites du lit majeur, ce qui explique que les scores ne diminuent pas même lorsque le débit est fortement surestimé.

L'analyse menée sur le cas d'étude néo-zélandais montre au contraire une corrélation parfaite entre les incertitudes sur les débits et celles sur les inondations. Ceci est principalement dû au fait que le

modèle hydraulique utilisé a été spécifiquement développé pour la ville de Westport, et calibré sur deux événements, dont les inondations de Juillet 2021. En France la modélisation hydraulique est appliquée à l'échelle régionale de l'arc Méditerranéen, sans calibration locale des paramètres du modèle. D'autre part, dans le contexte néo-zélandais, des hauteurs d'eau de référence sont disponibles sur l'ensemble de la zone d'étude, alors qu'en France, faute de repères de crue pour l'événement étudié, l'analyse ne peut s'appuyer que sur le couple (débit, étendue inondée) plutôt que (débit, hauteur d'eau), ce qui modifie la nature de la relation observée. Par ailleurs, bien que la chaîne de prévision néo-zélandaise repose elle aussi sur des scénarios d'inondation pré-calculés, ceux-ci sont beaucoup plus nombreux que les huit scénarios utilisés pour le cas français, et une interpolation entre scénarios est mise en œuvre.

2.2.3.3 Erreurs inondation/impact. Finalement, le troisième graphique de la figure VI.17 illustre les incertitudes associées à la modélisation inondation/impacts. Le CSI calculé sur les impacts y présente une gamme de valeurs très étroite (résumée dans le tableau VI.3). Cela s'explique par deux facteurs principaux : d'une part, l'événement étudié est relativement localisé ; d'autre part, la résolution des données d'impacts disponibles est assez grossière. Concrètement, seuls 4 tronçons de cours d'eau présentent des sinistres enregistrés dans la base de la Caisse Centrale de Réassurance, tandis que l'emprise inondée du scénario 1000 ans recoupe du bâti sur 7 tronçons (soit le nombre maximal de tronçons pouvant être impactés selon les prévisions). Le nombre restreint de tronçons concernés limite donc fortement l'éventail des valeurs possibles pour le CSI.

Un autre résultat marquant est l'absence totale de corrélation entre les erreurs sur l'inondation et celles sur les impacts. Pour mieux comprendre ce phénomène, la figure VI.18 examine le scénario présent dans tous les cas où le CSI-impact atteignait 0.75 alors que le CSI-inondation restait inférieur à 0.45 (points situés en haut à gauche du troisième graphique de la figure VI.17). Dans cet exemple, le faible recoupement entre les emprises inondées observée et prédite conduit logiquement à un CSI-inondation bas. Pourtant, la prévision identifie des bâtiments situés en zone inondée au sud de la zone (cercle vert), en dehors de l'emprise inondée observée. Or, le tronçon correspondant est répertorié comme sinistré dans la base CCR : il est donc comptabilisé comme un "HIT". Cette situation illustre comment un faible CSI-inondation peut coexister avec un CSI-impact élevé, en raison de l'échelle d'évaluation imposée (ici, le tronçon de cours d'eau).

CSI	0.40	0.50	0.57	0.60	0.67	0.75
Fréquence	0.15	0.10	0.43	0.22	0.01	0.09

Tableau VI.3 – Répartition des valeurs de CSI-impact obtenues (calculées sur les tronçons de cours d'eau affectés) pour les 17 membres de l'ensemble et les 4 cycles de prévision sélectionnés (soit sur $17 \times 4 = 68$ valeurs).

L'analyse menée sur le cas néo-zélandais montre au contraire une corrélation très forte entre incertitudes d'inondation et d'impact. Toutefois, les résultats ne sont que partiellement comparables. Comme indiqué dans le paragraphe précédent, la modélisation hydraulique diffère significativement entre les deux applications. De plus, les indicateurs d'erreur ne sont pas les mêmes : dans le cas néo-zélandais, l'erreur moyenne sur les hauteurs d'eau est utilisée pour l'inondation, tandis que les impacts sont évalués via un CSI calculé directement à l'échelle du bâtiment (grâce à la base de données des "placards" de sécurité apposés sur les bâtiments après une catastrophe naturelle).

2.2.3.4 Bilan En résumé, les résultats obtenus sur le cas français ne montrent pas des résultats aussi clairs que pour le cas néo-zélandais (où la chaîne de prévision est appliquée à l'échelle d'une

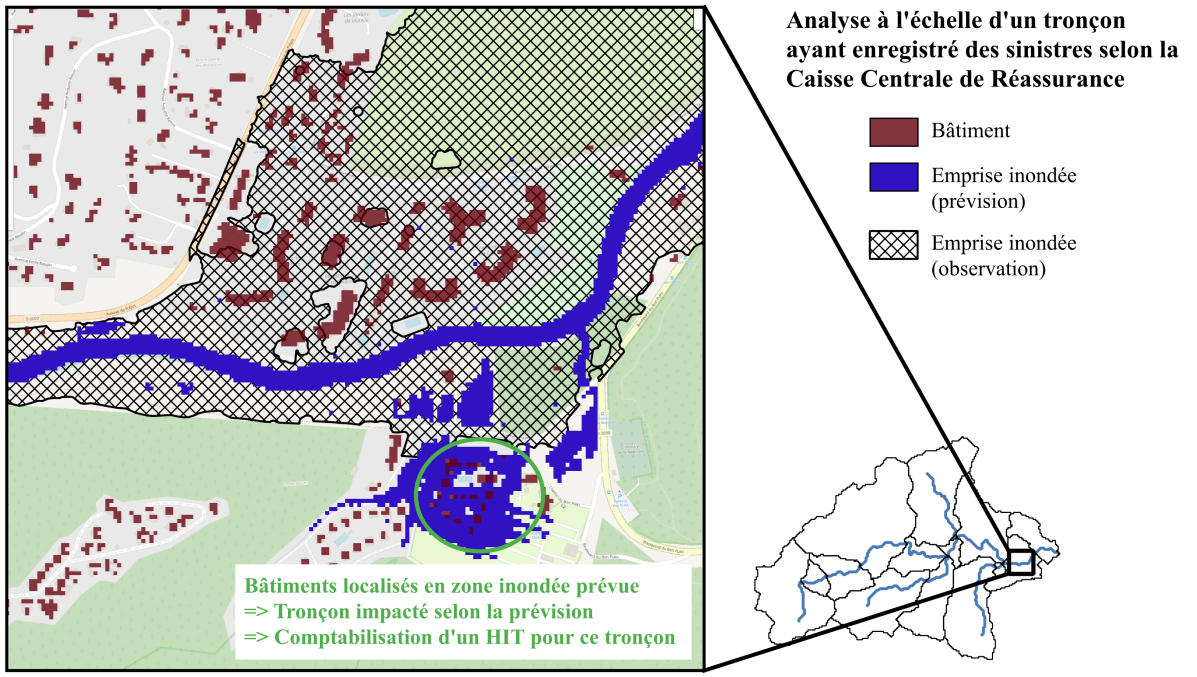


FIGURE VI.18 – Illustration d'un tronçon pour lequel une mauvaise prévision de l'étendue de l'inondation peut se traduire en une prévision d'impact correcte.

ville et la modélisation hydraulique est calibrée sur l'événement étudié, et où les jeux de données d'évaluation disponibles sont plus riches). Le cas français apporte toutefois un éclairage complémentaire avec l'utilisation d'un modèle hydraulique non calibré et appliqué à une échelle régionale pour le calcul d'un plus faible nombre de scénarios pré-calculés, qui sont utilisés dans la chaîne de prévision sans interpolation. Il met également en évidence l'influence des scores d'erreur retenus (ou imposés par les données d'évaluation disponibles) sur l'analyse des relations entre incertitudes des variables en entrée et en sortie, ainsi que l'importance de l'échelle d'évaluation, en particulier pour le bloc de modélisation des impacts.

3 Conclusions et perspectives

L'objectif de ce chapitre était d'explorer des perspectives sur l'implémentation d'approches ensemblistes dans une chaîne de prévision immédiate basée sur les impacts.

Apport des prévisions d'ensemble dans une chaîne de prévision immédiate des crues soudaines

La première partie de ce chapitre a évalué la valeur ajoutée de prévisions d'ensemble de pluie en entrée d'un système d'avertissement de type HBW. L'ensemble de 17 membres du produit PIAF, fournissant des prévisions jusqu'à trois heures d'échéance, a été injecté dans une chaîne très similaire au système opérationnel Vigicrues Flash. Des tableaux de contingence permettant d'évaluer la qualité d'anticipation des dépassements de seuils de débits ont été construits à partir d'une méthodologie adaptée de Charpentier-Noyer et al. [2023] et ont permis de comparer PIAF-ensemble à sa version déterministe, au produit AROME-PI [Auger et al., 2015], ainsi qu'à un scénario naïf de pluie constante. Les résultats montrent que les prévisions d'ensemble surpassent les produits déterministes. Cette évaluation, réalisée sur huit événements de crues soudaines, gagnerait toutefois à être prolongée sur des chroniques plus longues, et en mobilisant d'autres critères d'évaluation (e.g., le CRPS [Hersbach, 2000] ou le Brier score [Brier, 1950])

D'autre part, il serait possible d'adapter l'évaluation hydrologique proposée ici afin de mesurer la plus-value des prévisions d'ensemble sous l'angle des impacts. Contrairement à l'approche de Charpentier-Noyer et al. [2025], qui repose sur des données synthétiques et non sur des observations (simulation de déplacement d'équipes de secours selon différentes prévisions probabilistes de pluies), la base de données constituée durant cette thèse permet aujourd'hui d'envisager une telle analyse avec de réelles données observées (cf. Chapitre V). Si cette évaluation n'a pas été conduite dans le cadre de la thèse, c'est pour plusieurs raisons : (i) le programme de travail était déjà dense ; (ii) les archives de prévisions d'ensemble disponibles restent encore relativement récentes et peu nombreuses pour réaliser une analyse robuste. Néanmoins, les développements menés ici posent les bases méthodologiques nécessaires pour conduire ce type d'évaluation à l'avenir.

Analyse de la propagation des incertitudes dans une chaîne IBW

La deuxième partie de ce chapitre a étudié l'intérêt de prévisions d'ensemble de pluie pour un système d'avertissement de type IBW. La méthode consistait à définir des scores adaptés à chaque variable d'entrée/sortie des blocs de modélisation, pertinents vis-à-vis des données de validation disponibles. La méthode, initialement définie et appliquée à l'inondation de Westport en 2021 en Nouvelle-Zélande (dans le cadre d'un séjour de recherche, cf annexe F), avait donné des résultats intéressants sur cette étude de cas. Toutefois, la même analyse menée sur l'événement de 2019 n'a pas conduit aux mêmes conclusions, soulignant d'une part les différences entre une approche développée et calibrée à l'échelle d'une ville et d'un événement, vis-à-vis d'une approche régionale non calibrée, et d'autre part l'influence des scores d'erreurs choisis, et de la disponibilité et de la résolution des données de validation.

Ces premières applications d'une méthodologie inspirée de Titley et al. [2024] ouvrent la voie à de futures analyses de propagation des incertitudes dans les chaînes de prévision des impacts d'inondations. Celles-ci pourraient être généralisées à un plus grand nombre d'événements, en mobilisant les jeux de données exploités dans les chapitres II, III et V : respectivement sur les débits observés, sur les niveaux d'eau maximaux atteints et sur les impacts constatés. Les inondations de l'Aude en octobre 2018 pourraient

être un cas intéressant : de nombreuses données de débits, hauteurs maximales sont disponibles, et des données d'impacts sur le bâti à échelle plus fine que le tronçon pourraient être exploitées.

Représentation et communication de prévisions d'impacts probabilistes

L'application de la chaîne de type IBW utilisant des prévisions d'ensemble de pluie sur l'événement du 1er décembre 2019 (figure VI.14) a donné un premier aperçu du large éventail de représentations possibles pour ce type de prévisions. Mais la communication de telles prévisions demeure un défi majeur : un message clair pour un prévisionniste peut être source de confusion pour les utilisateurs (trop d'informations, interprétation contre-intuitive, ou au contraire simplification excessive) [Jean et al., 2023]. Dans le cadre du séjour de recherche en Nouvelle-Zélande, un questionnaire a été élaboré et diffusé pour sonder les attentes et besoins des utilisateurs vis-à-vis de ces représentations probabilistes basées sur les impacts. Le contenu des questions, les types de représentations proposées et les réponses recueillies sont détaillés dans l'Annexe F, section 7. Ces questions relatives à la représentation et à l'agrégation des prévisions d'ensemble basées sur les impacts pourraient être explorées en soumettant le questionnaire développé en Nouvelle-Zélande à un public français, voire international : e.g. au Canada et au Royaume-Uni, dans lesquels des contacts ont déjà été établis.

Une approche régulièrement citée dans la littérature concernant la représentation de prévisions probabilistes d'impacts est celle utilisée par le Flood Forecasting Center pour l'Angleterre et le Pays de Galles. Une matrice de risque, ou "flood risk matrix" [Cole et al., 2016], illustrée en figure VI.19, combinant probabilité du phénomène et impacts potentiels, est utilisée pour agréger les prévisions à l'échelle de chaque collectivité locale (voir figure VI.20). Cette représentation a été proposée dans le questionnaire élaboré en Nouvelle-Zélande, cependant les participants semblaient être partagés entre cette approche et un prototype dans lequel les scénarios min., médian, et max. sont fournis à la place d'une représentation probabiliste (Figure F.25).

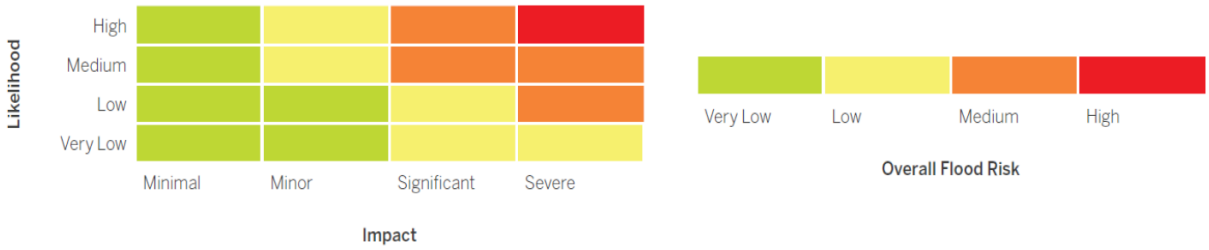


FIGURE VI.19 – "Flood risk matrix" utilisée dans les Flood Guidance Statements en Angleterre et au Pays de Galles (source : Flood Forecasting Center/ MetOffice).

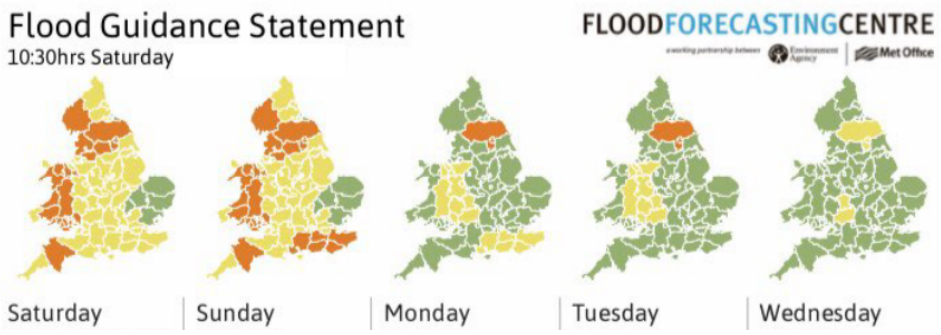


FIGURE VI.20 – Exemple de prévisions utilisant l'approche "flood risk matrix" en Angleterre et au Pays de Galles (source : Flood Forecasting Center/ MetOffice).

Chapitre VII

Conclusions et perspectives

Objectifs et organisation du chapitre :

Ce chapitre présente les conclusions de l'ensemble des travaux conduits au cours de la thèse. La première partie reprend les objectifs et les questions de recherches formulés en introduction, et résume les approches adoptées ainsi que les enseignements obtenus pour chacun d'eux. La seconde partie ouvre des perspectives de recherche liées au sujet, qui pourraient nourrir des travaux futurs.

1 Enseignements tirés en lien avec les objectifs de la thèse

Cette vise à mettre en œuvre et évaluer une chaîne de prévision immédiate des crues soudaines fondée sur les impacts, applicable à l'échelle régionale sur l'arc méditerranéen. Elle repose sur un enchaînement cohérent de choix méthodologiques, depuis la modélisation pluie-débits jusqu'à la définition d'avertissements fondés sur les impacts, avec l'objectif de proposer des solutions opérationnelles, compatibles avec les contraintes de couverture spatiale, de temps de calcul et de disponibilité des données.

Les travaux menés montrent qu'il est possible de s'appuyer sur un modèle hydrologique régionalisé unique pour couvrir l'ensemble des cours d'eau méditerranéens, y compris ceux au comportement complexe, tout en conservant des performances satisfaisantes. Le couplage de ce modèle avec des scénarios d'inondation pré-calculés, issus d'une modélisation hydraulique volontairement simplifiée, permet de représenter de manière réaliste les zones inondées à grande échelle. Cette approche constitue un compromis assumé entre précision et faisabilité opérationnelle.

La transition vers des avertissements fondés sur les impacts constitue un apport central de la thèse. En reliant directement les débits simulés aux enjeux exposés, cette approche permet de dépasser les limites des avertissements fondés uniquement sur l'aléa, en intégrant explicitement l'exposition des territoires. L'évaluation conduite sur une large fenêtre spatio-temporelle montre ainsi que les avertissements fondés sur les impacts réduisent significativement le nombre de fausses alarmes et améliorent la hiérarchisation des zones réellement affectées lors des événements.

Enfin, l'intégration de prévisions de pluie d'ensemble ouvre des perspectives importantes pour l'augmentation des délais d'anticipation et la prise en compte des incertitudes liées à la prévision météorologique. Les résultats obtenus soulignent l'intérêt d'une approche probabiliste. L'ensemble de ces travaux contribue ainsi à poser les bases méthodologiques d'une chaîne de prévision immédiate fondée sur les impacts, évaluée de manière rigoureuse et pensée dès sa conception pour une utilisation opérationnelle.

La suite de cette partie reprend les objectifs et les questions de recherches formulés en introduction, et résume les approches adoptées ainsi que les enseignements obtenus pour chacun d'eux.

1.1 Mise en place d'une chaîne de prévision immédiate des impacts des crues soudaines pour l'Arc Méditerranéen

Dans un premier temps, ce travail de thèse s'est concentré sur la mise en place d'une chaîne de prévision immédiate constituée de plusieurs blocs de modélisation : pluie, débits, inondations et impacts. Bien que cet objectif ait une finalité opérationnelle et s'appuie largement sur des outils existants, il soulevait de nombreuses questions de recherche :

Un système HBW (Hazard-Based Warning) utilisant un modèle hydrologique régionalisé unique peut-il couvrir l'ensemble des cours d'eau méditerranéens, y compris ceux au comportement particulier, tout en conservant des performances homogènes ? Le modèle pluie-débit utilisé est basé sur celui actuellement opérationnel dans Vigicrues Flash. Il s'agit d'un modèle hydrologique distribué (mailles de 1 km^2), fonctionnant de manière continue au pas de temps de 15 minutes. Son principe repose sur deux réservoirs (production et transfert) dont les paramètres sont estimés de manière simplifiée sur de larges zones. Le travail présenté au Chapitre II a permis de montrer que (i) le modèle pouvait être appliqué à l'ensemble de la zone d'étude sans perte notable de performance par rapport à l'approche de Vigicrues Flash, qui se limite aux cours d'eau au comportement simple, (ii) il était possible d'obtenir au moins une heure d'anticipation pour 60 à 70 % des événements détectés – d'après l'analyse menée sur

le réseau jaugé – malgré l'utilisation de données de pluie observées et non de prévisions en entrée du modèle, et (iii) le calage de la méthode SHYREG sur les débits simulés offrait un moyen d'estimer des quantiles de débits jusqu'à la période de retour de 1000 ans, permettant de corriger le biais du modèle hydrologique lié à l'absence de module de routage réaliste.

Comment connecter les sorties d'un modèle hydrologique maillé à un modèle hydraulique fondé sur une représentation vectorielle des cours d'eau, pour aboutir à une représentation des zones inondées nécessaire dans un système de type IBW ? La mise en place de la chaîne de prévision complète a ensuite nécessité de connecter les sorties du modèle pluie-débit, estimées sur une grille hydrologique de mailles carrées kilométriques, à des scénarios d'inondation établis sur un réseau vectoriel de tronçons de rivière. Cette étape d'appariement entre mailles hydrologiques et biefs de cours d'eau, incontournable avec des modèles hydrologiques maillés, a été explorée en comparant l'approche classique - basée sur la correspondance des aires drainées amont - à une méthode plus récente prenant aussi en compte les limites des bassins versants. Le Chapitre IV (Partie 1) a montré la nette plus-value de cette seconde méthode, en particulier pour les petits bassins, pour lesquels les erreurs d'appariement peuvent être fréquentes, et qui sont aussi les plus vulnérables aux crues soudaines. Elle a donc été retenue pour le couplage hydrologie / hydraulique.

Quel type de modélisation hydraulique peut-on mobiliser pour la représentation des zones inondées, et quel est l'impact de l'amélioration de la représentation des lits mineurs sur les résultats de modélisation de l'inondation ? La représentation des zones inondées dans la chaîne de prévision s'est appuyée sur des scénarios d'inondation pré-calculés pour huit périodes de retour comprises entre 2 et 1000 ans (Chapitre III). Ces scénarios ont été produits à partir d'un modèle hydraulique simplifié : équations de Saint-Venant 2D résolues en régime stationnaire et en négligeant les termes inertiels, sans calage des paramètres de rugosité. Les résultats se sont révélés satisfaisants, voire très satisfaisants pour une approche appliquée à l'échelle régionale, notamment lorsque les données topographiques utilisées étaient améliorées (données Lidar prétraitées pour assurer une meilleure continuité hydraulique des lits mineurs).

Comment reconstituer la prévision d'inondation et la prévision d'impacts dans la chaîne de type IBW, en se basant sur des scénarios d'inondations pré-calculés ? L'approche retenue pour reconstituer l'emprise inondée prévue (Chapitre IV, Partie 2) a été améliorée par rapport aux travaux antérieurs, de façon à mieux représenter l'inondation au niveau des confluences. L'approche consiste toujours à sélectionner le scénario dont la période de retour est la plus proche (immédiatement inférieure) à celle estimée par le modèle. En revanche, les scénarios d'inondation associés à chaque tronçon ne sont plus découpés en fonction des limites de bassin associé à chaque tronçon, mais superposées. Les résultats obtenus ont montré une nette amélioration de la représentation des limites d'inondation dans les secteurs de plaine.

Comment définir des avertissements fondés sur les impacts, à l'image de ceux actuellement fondés sur l'aléa dans Vigicrues Flash ? Les impacts ont été évalués en comptant le nombre de bâtiments exposés (issus de la BD TOPO), à l'échelle des tronçons de cours d'eau et des communes. Cette estimation a été réalisée une seule fois pour l'ensemble des scénarios (y compris dans les zones de recouvrement des domaines de calculs hydrauliques), ce qui permet d'obtenir une relation directe Débits–Impacts, particulièrement adaptée à une utilisation opérationnelle, en raison de l'absence de calculs à effectuer

en temps réel. Les IBWs ont ensuite été définis en comparant le nombre de bâtiments inondés à un certain seuil (Chapitre V), dans une logique analogue à celle de Vigicrues Flash qui définit ses HBWs en comparant le débit simulé à un seuil prédéfini.

Est-il possible de coupler le module hydrologique de la chaîne d'avertissement avec des prévisions de pluie d'ensemble, afin d'augmenter les délais d'anticipation et de représenter les incertitudes associées aux prévisions ? Quelle est la plus-value d'un tel couplage ? Le Chapitre VI a étudié le couplage du système Vigicrues Flash avec des prévisions de pluie probabilistes. Les résultats, testés sur plusieurs événements significatifs, ont montré leur plus-value par rapport aux produits déterministes. Cette étude ouvre des perspectives intéressantes pour la construction d'une chaîne de prévision complète fondée sur des prévisions de pluie et non pas simplement des pluies mesurées, et intégrant une représentation probabiliste des impacts attendus.

1.2 Évaluation détaillée sur une large fenêtre spatio-temporelle

L'objectif in fine de ce travail était d'analyser, sur une vaste échelle spatiale et temporelle, la performance d'une chaîne de prévision de type IBW intégrant successivement la modélisation hydrologique et la représentation des zones inondées et des impacts. L'évaluation proposée dans cette thèse a ainsi couvert plus de 20 000 km de linéaire de cours d'eau de l'arc méditerranéen, sur plus d'une décennie (2010–2022), et a permis de traiter différentes questions de recherche :

Quelles données mobiliser pour évaluer un modèle hydrologique estimant des débits et des périodes de retour sur un réseau en grande partie non jaugé ? Dans un premier temps, la capacité de la première partie de la chaîne (modèle de type HBW, similaire à Vigicrues Flash) à détecter des crues soudaines a été évaluée (Chapitre II). Puisque ce modèle est appliqué sur un réseau en grande partie non jaugé, l'évaluation s'est appuyée sur un jeu de données élargi, allant au-delà des débits observés aux stations hydrométriques. Celui-ci comprend des estimations de débits de pointe sur de petits cours d'eau, obtenues lors des relevés de terrain post-crue organisés au sein de l'OHM-CV (Observatoire Hydrométéorologique Méditerranéen Cévennes Vivarais), ainsi que des arrêtés CatNat recensant les communes impactées par des inondations dans la zone d'étude.

Quelles sont les performances d'une chaîne de type HBW pour la détection des crue soudaine ? L'utilisation de données d'impacts permet-elle d'identifier les limites d'un système HBW, en particulier l'absence de prise en compte de l'exposition des enjeux dans la définition des seuils d'avertissement ? L'évaluation de la chaîne de type HBW à partir de données hydrologiques, présentée au Chapitre II, a montré des performances comparables à d'autres systèmes de la littérature (scores CSI d'environ 40 % pour l'évaluation continue aux stations). L'évaluation continue basée sur des données d'impacts a présenté un score CSI de 35 % et a mis en évidence l'importance de la prise en compte des enjeux dans la définition des seuils d'avertissement.

Quelles données et méthodes peut-on mobiliser pour évaluer des scénarios d'inondation calculés pour des périodes de retour fixes à échelle régionale ? La modélisation des zones inondées, nécessaire pour la transition des HBWs vers les IBWs, a été évaluée sur l'ensemble de la zone d'étude (Chapitre III). Un jeu de données composé de 171 courbes de tarage a été constitué et utilisé pour comparer les couples (hauteur, débit) correspondant aux différents scénarios hydrauliques aux valeurs de référence estimées à partir des courbes de tarage. Les courbes de tarage permettent une évaluation à grande échelle

du modèle, et intègrent à la fois des observations pour les débits fréquents et l'expertise hydraulique locale des services hydrométriques pour les crues exceptionnelles.

Quels sont les ordres de grandeur des erreurs de hauteur d'eau associées à une modélisation hydraulique simplifiée appliquée à grande échelle ? Les résultats apparaissent très satisfaisants pour une modélisation hydraulique simplifiée appliquée à grande échelle : l'erreur médiane (simulation – observation) est de –0,04 m pour le scénario de période de retour 2 ans, et de –0,14 m tous scénarios confondus. Les écarts interquantiles se situent entre ±50 cm pour le scénario le plus fréquent et se maintiennent entre ±75 cm jusqu'au scénario de période de retour 50 ans. L'évaluation a également mis en évidence plusieurs sources d'erreurs, dont la fiabilité variable des données de validation et l'absence de données Lidar dans certaines dalles des MNT utilisés.

Quelles données d'impacts peuvent être mobilisées pour l'évaluation de l'ensemble de la chaîne, sur une large fenêtre spatio-temporelle ? Pour permettre une évaluation complète de la chaîne de type IBW, les données de validation recherchées doivent correspondre à un recensement d'impacts observés continu dans le temps, exhaustif dans l'espace et disposant de la résolution la plus fine possible. Un large éventail de données a été analysé au Chapitre V (y compris des sources originales comme celles issues des réseaux de télécommunication, électriques ou encore des réseaux sociaux). Toutefois, les trois critères pré-cités ont fortement limité les sources de données réellement mobilisables. Trois bases de données ont finalement été retenues : la liste des arrêtés CatNat déclarés sur les communes de l'arc méditerranéen (disponibles dans la base Gaspar pour la période 2010–2022), la liste et la hiérarchie par événement des tronçons de cours d'eau sinistrés selon la Caisse Centrale de Réassurance (CCR) sur l'ensemble de la zone et de la période étudiées, et enfin le détail des interventions des services de secours liées aux inondations pour deux départements (Gard et Alpes-Maritimes) depuis 2012.

Quelle méthode d'évaluation peut-on appliquer pour évaluer la plus-value des IBWs par rapport aux HBWs ? Compte tenu de la résolution des données d'impacts collectées, l'évaluation a été menée à deux échelles spatiales (Chapitre V) : la commune ($\simeq$ 20 km^2 en moyenne) et le tronçon de cours d'eau ($\simeq$ 4 km^2). Les avertissements HBWs et IBWs émis à ces deux échelles - selon le rejeu des chaînes de prévision sur plus d'une décennie et sur l'ensemble de l'arc méditerranéen - ont été comparés aux impacts observés au pas de temps journalier, permettant le dénombrement des "hits", "misses", "false alarms" et "correct rejections". De plus, les avertissements HBWs et IBWs ont été adaptés pour fournir un classement des tronçons impactés lors d'un événement, selon la période de retour estimée pour le premier cas, et selon le nombre de bâtiments impactés pour le second. Un score basé sur le CSI a été proposé pour comparer ces classements aux hiérarchies des zones impactées d'après les observations (nombres de sinistres enregistrés par la CCR).

Les résultats d'évaluation permettent-ils de confirmer cette plus-value en termes de réduction de fausses alarmes et de hiérarchisation des zones impactées ? Les résultats sur la période 2010–2022 ont montré qu'à taux de "hits" équivalents, les IBWs réduisent significativement le nombre de fausses alertes par rapport aux HBWs : le nombre de fausses alertes était divisé par 1,3 à 1.7 à l'échelle communale et par 2 à 2.7 à l'échelle des tronçons. L'évaluation à l'échelle événementielle a par ailleurs montré que les prévisions basées sur les impacts permettaient, dans la quasi-totalité des cas, une meilleure hiérarchisation des zones affectées.

Comment évaluer la plus-value d'un produit de prévisions d'ensemble de pluie par rapport à des produits déterministes, à l'échelle de quelques événements de crues soudaines ? Le dernier chapitre de la thèse s'intéresse, dans un premier temps, au couplage de la chaîne de type HBW avec des prévisions d'ensemble de pluie (version ensembliste du produit PIAF - prévision agrégée fusionnée - de Météo France). La plus-value de celles-ci par rapport à trois produits déterministes a été évaluée en rejouant les avertissements émis - basés sur des dépassements de quantiles de débits - pour chaque décile de l'ensemble et pour chaque scénario déterministe, puis en les comparant aux avertissements émis pour un scénario de référence correspondant aux simulations pluie-débit avec pluie observée. Cette méthode, qui a été adaptée par rapport à celle initialement proposée par Charpentier-Noyer et al. [2023], ne mobilise pas de données d'observation mais permet de calculer des scores d'erreur à l'échelle de chaque pixel de 1 km² du domaine. Les "hits", "misses", "false alarms" et "correct rejections" sont ensuite cartographiés pour chaque événement, puis agrégés via des scores CSI calculés à l'échelle événementielle et moyennés. Cette méthode a montré que les déciles intermédiaires (40–50 %) des prévisions d'ensemble obtenues présentaient des scores supérieurs aux trois scénarios déterministes considérés.

Comment aborder la problématique de la propagation des incertitudes au sein de la chaîne de prévision ? Le produit de prévision d'ensemble de pluie a également été utilisé en entrée de la chaîne de type IBW. À l'échelle d'un événement bien documenté, des scores d'erreur adaptés à chaque variable de sortie des différents blocs de la chaîne ont pu être calculés pour chaque membre de l'ensemble et pour différentes échéances de prévision. Cette méthode permet de suivre l'évolution des incertitudes tout au long de la chaîne. Ce travail a fourni des premiers éléments d'analyse pour le cas des inondations du 1er décembre 2019 dans les Alpes Maritimes, et ouvre la voie vers une évaluation plus large sur plusieurs événements sur lesquels des données d'observation sont disponibles (débits de pointe, marques de plus hautes eaux, etc.).

2 Perspectives

2.1 Prévision des pluies : allonger les délais d'anticipation

Ce travail s'est majoritairement concentré sur les blocs débits-inondation-impacts de la chaîne de prévision, qui était alimentée par des observations de pluie. Cette méthode permet de réduire nettement les incertitudes liées aux données pluviométriques d'entrée, mais limite significativement les délais d'anticipation. Un couplage de la chaîne implémentée dans la thèse avec des prévisions d'ensemble de pluie à courte échéance (horizon de 3h pour le produit de prévision d'ensemble PIAF) a été testé et a révélé des premiers résultats encourageants quant à la plus-value de ces prévisions d'ensemble par rapport à des scénarios déterministes, à l'échelle de huit événements de crues soudaines.

Cependant, l'amélioration continue des modèles numériques de prévision du temps ouvre la voie vers des prévisions allant jusqu'à 24h ou 48h d'échéance, avec une approche ensembliste pour la représentation des incertitudes associées [Furnari et al., 2020, Amengual et al., 2021]. C'est notamment l'approche adoptée en Grande-Bretagne pour la prévision opérationnelle des inondations, avec des horizons allant jusqu'à quatre jours dans les "Flood Guidance Statements" [Maybee et al., 2024] et jusqu'à deux jours dans les "Rapid Flood Guidance Statements" [Pilling et al., 2025]. En France, le projet Vigicrues 2030 vise à étendre la vigilance crues à 24 h, actuellement limitée aux grands cours d'eau instrumentés, à l'ensemble du territoire métropolitain. Le cadre d'évaluation développé dans cette thèse, ainsi que celui proposé par Charpentier-Noyer et al. [2025], permettra d'analyser l'utilité effective de ces nouveaux produits selon le type de décision visé et la précision spatiale désirée. En effet, si l'allongement des échéances de prévision constitue un progrès important, il s'accompagne d'une incertitude croissante : les prévisions à 24 h pourront contribuer à la mise en vigilance sur de larges territoires, sans pour autant remplacer les prévisions immédiates indispensables à la gestion de crise lors du déclenchement des crues soudaines.

2.2 Prévision des débits : améliorer les modèles hydrologiques

Les premiers résultats obtenus sur l'analyse de la propagation des incertitudes dans la chaîne de prévision, à partir des cas néo-zélandais et français, encouragent à améliorer la connaissance des relations pluie–débits. Le modèle hydrologique utilisé dans cette thèse repose sur une structure simplifiée à deux réservoirs (production et transfert), avec des paramètres uniformes à l'échelle de grandes régions, et sans module réaliste de routage des écoulements. Les travaux récents de Huynh [2025] se sont concentrés sur l'amélioration de modèles hydrologiques adaptés à la prévision opérationnelle des crues rapides en Méditerranée française, en mobilisant notamment des approches basées sur l'intelligence artificielle. Ils ont notamment mis en évidence le potentiel de l'intégration de réseaux de neurones pour la régionalisation des paramètres, dans le cadre de structures d'opérateurs hydrologiques et de routage plus complexes, permettant une meilleure représentation des flux et processus de la transformation pluie–débit.

Ces avancées sont intégrées dans l'outil SMASH [Colleoni et al., 2025], qui repose sur une représentation « pixelisée » du réseau hydrographique. De futurs travaux devraient également explorer la plus-value d'un couplage de SMASH avec un module de routage basé sur une représentation vectorielle du réseau, ce qui faciliterait le lien avec les approches hydrauliques dans une chaîne de prévision de type IBW. Améliorer le réalisme du module de routage des débits simulés vers l'exutoire des bassins versants permettrait également d'éviter le passage par les périodes de retour lors de la connexion entre les prévisions hydrologiques et les scénarios hydrauliques, qui est une source d'incertitude importante dans l'approche proposée ici.

2.3 Prévision des zones inondées : complexifier et interpoler

L'évaluation des scénarios d'inondation utilisés dans la chaîne de prévision de type IBW (Chapitre III) a mis en évidence l'importance de la précision des MNTs en entrée du modèle hydraulique, en particulier l'impact de l'absence de données Lidar dans certaines dalles. L'ensemble de la zone étudiée est désormais couvert par le Lidar HD (voir https://geoservices.ign.fr/lidarhd); une mise à jour des scénarios d'inondation serait donc possible et profitable. Par ailleurs, la simplicité de l'approche hydraulique mobilisée - hypothèse de régime permanent et négligence des termes inertiels - pourrait être remise en question à l'avenir, compte tenu des progrès rapides des modèles optimisant les schémas de résolution et leur parallélisation pour permettre des simulations hydrauliques plus complexes à très fines résolutions [Costabile et al., 2023]. Toutefois, de telles méthodes ne semblent pas encore compatibles avec une application en temps réel à l'échelle régionale. L'approche fondée sur des scénarios préétablis demeure donc, à ce jour, la plus adaptée pour éviter des calculs trop lourds. Il serait néanmoins intéressant d'étudier dans de futurs travaux la plus-value de techniques d'interpolation entre les scénarios disponibles. Cette interpolation reste complexe, notamment au niveau des confluences, mais le développement récent de modèles de substitution basés sur l'intelligence artificielle ouvre la voie à des perspectives prometteuses [Kabir et al., 2020, Wang et al., 2024].

2.4 Prévision des impacts : mieux répondre aux besoins des utilisateurs

Pour dépasser le modèle d'impact très simple consistant à localiser et compter le nombre de bâtiments exposés à une inondation prévue, il serait intéressant d'intégrer d'autres types de variables pour représenter les impacts, par exemple les routes et voies ferrées, les réseaux d'électricité ou encore de communication. Ces variables sont également plus directement reliées à des sources de données d'impacts observés (cf. Annexe D), et donc probablement plus proches des réalités de certains gestionnaires de crise. De plus, une perspective prometteuse pour mieux représenter les impacts engendrés par une inondation réside dans l'utilisation de modèles multi-agents, qui permettent notamment de simuler les processus complexes d'interaction entre différents acteurs (citoyens, services de secours, gestionnaires de crise, etc.) et leurs réactions en cas de crise [Yang et al., 2018, Li et al., 2019, Novoa et al., 2023]. Ces approches sont intéressantes pour préparer la gestion de crise et tester la robustesse de plans d'actions, mais aussi pour évaluer l'utilité des prévisions [Charpentier-Noyer et al., 2025]

Plus généralement, des progrès sont possibles afin que les outils développés répondent davantage aux besoins des utilisateurs, en leur permettant de prendre des décisions plus efficaces en situation de crise. Cela illustre l'évolution du paradigme sous-tendant cette thèse - *from "What the weather will be" to "What the weather will do"* [WMO, 2015] - vers un nouvel objectif : *"What the users should do"*. Autrement dit, il s'agit d'évoluer des HBWs et des IBWs vers des Action-Based Warnings (ABWs). Pour y parvenir, il est notamment nécessaire de poursuivre les recherches sur la compréhension des processus de décision, par exemple à travers l'utilisation de jeux sérieux [Terti et al., 2019b].

2.5 Chaîne de prévision complète : mieux comprendre, mieux évaluer, mieux communiquer

Au-delà des améliorations portant individuellement sur chacun des blocs de modélisation de la chaîne de prévision, la compréhension de la propagation des informations - et notamment des incertitudes - au sein de cette chaîne doit être approfondie. Les premiers travaux sur ce sujet, présentés au Chapitre VI et inspirés par Titley et al. [2024], méritent d'être étendus à un plus grand nombre de cas d'étude. À l'échelle d'événements spécifiques, il serait par ailleurs pertinent d'exploiter la précision temporelle et

géographique de certaines sources de données d'impacts observés, en particulier les interventions des services de secours, qui sont précisément géolocalisées et horodatées.

De plus, de nombreuses initiatives récentes de collecte, d'uniformisation et de mise à disposition des données d'impacts laissent penser que le contexte devient de plus en plus favorable à des évaluations sur de longues périodes temporelles, à l'image de celle menée dans cette thèse. On peut notamment citer la forte mobilisation des services d'incendie et de secours dans la démarche de collecte des données présentées au Chapitre V, ainsi que leur volonté d'harmoniser le recensement des interventions [Dias, 2024]. On peut également souligner que l'historique des événements climatiques en France, disponible en libre accès sur le site de la CCR, a été récemment mis à jour et intègre désormais des cartes de coûts à l'échelle communale pour chaque inondation passée.

Enfin, il convient de rappeler que l'amélioration des chaînes de prévision des inondations ne sera véritablement efficace que si l'information est transmise de manière claire et adaptée aux utilisateurs. De futurs travaux, qui pourront s'appuyer notamment sur le questionnaire présenté en Annexe F, devront explorer plus en profondeur les enjeux liés au contenu des avertissements : leur accessibilité, leur clarté, leur compréhensibilité, mais aussi leur capacité à être inclusifs et adaptés à la diversité des utilisateurs.

Bibliographie

Raffaele Albano, Caterina Samela, Iulia Craciun, Salvatore Manfreda, Jan Adamowski, Aurelia Sole, Ake Sivertun, and Alexandru Ozunu. Large Scale Flood Risk Mapping in Data Scarce Environments : An Application for Romania. *Water*, 12(6) : 1834, June 2020. ISSN 2073-4441. doi : 10.3390/w12061834. URL https://www.mdpi.com/2073-4441/12/6/1834. Number : 6 Publisher : Multidisciplinary Digital Publishing Institute.

L. Alfieri, P. Burek, E. Dutra, B. Krzeminski, D. Muraro, J. Thielen, and F. Pappenberger. GloFAS - global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences*, 17(3) :1161–1175, 2013. ISSN 1027-5606. doi : 10.5194/hess-17-1161-2013. URL https://hess.copernicus.org/articles/17/1161/2013/. Publisher : Copernicus GmbH.

Lorenzo Alfieri, Peter Salamon, Alessandra Bianchi, Jeffrey Neal, Paul Bates, and Luc Feyen. Advances in pan-European flood hazard mapping. *Hydrological Processes*, 28(13) :4067–4077, 2014. ISSN 1099-1085. doi : 10.1002/hyp.9947. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.9947. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/hyp.9947.

Lorenzo Alfieri, Andrea Libertino, Lorenzo Campo, Francesco Dottori, Simone Gabellani, Tatiana Ghizzoni, Alessandro Masoero, Lauro Rossi, Roberto Rudari, Nicola Testa, Eva Trasforini, Ahmed Amdihun, Jully Ouma, Luca Rossi, Yves Tramblay, Huan Wu, and Marco Massabò. Impact-based flood forecasting in the Greater Horn of Africa. *Natural Hazards and Earth System Sciences*, 24(1) :199–224, 2024. ISSN 1561-8633. doi : 10.5194/nhess-24-199-2024. URL https://nhess.copernicus.org/articles/24/199/2024/. Publisher : Copernicus GmbH.

A. Amengual, D. S. Carrió, G. Ravazzani, and V. Homar. A Comparison of Ensemble Strategies for Flash Flood Forecasting : The 12 October 2007 Case Study in Valencia, Spain. *Journal of Hydrometeorology*, 18(4) :1143–1166, April 2017. ISSN 1525-7541, 1525-755X. doi : 10.1175/JHM-D-16-0281.1. URL https://journals.ametsoc.org/view/journals/hydr/18/4/jhm-d-16-0281_1.xml. Publisher : American Meteorological Society Section : Journal of Hydrometeorology.

A. Amengual, A. Hermoso, D. S. Carrió, and V. Homar. The Sequence of Heavy Precipitation and Flash Flooding of 12 and 13 September 2019 in Eastern Spain. Part II : A Hydrometeorological Predictability Analysis Based on Convection-Permitting Ensemble Strategies. *Journal of Hydrometeorology*, 22(8) :2153–2177, August 2021. ISSN 1525-7541, 1525-755X. doi : 10.1175/JHM-D-20-0181.1. URL https://journals.ametsoc.org/view/journals/hydr/22/8/JHM-D-20-0181.1.xml. Publisher : American Meteorological Society Section : Journal of Hydrometeorology.

Arnau Amengual, Romu Romero, María Carmen Llasat, Alejandro Hermoso, and Montserrat Llasat-Botija. Hydrometeorological controls of and social response to the 22 October 2019 catastrophic flash flood in Catalonia, north-eastern Spain. *Natural Hazards and Earth System Sciences*, 24(6) :2215–2242, July 2024. ISSN 1561-8633. doi : 10.5194/nhess-24-2215-2024. URL https://nhess.copernicus.org/articles/24/2215/2024/. Publisher : Copernicus GmbH.

William Amponsah, Pierre-Alain Ayral, Brice Boudevillain, Christophe Bouvier, Isabelle Braud, Pascal Brunet, Guy Delrieu, Jean-François Didon-Lescot, Eric Gaume, Laurent Lebouc, Lorenzo Marchi, Francesco Marra, Efrat Morin, Guillaume Nord, Olivier Payrastre, Davide Zoccatelli, and Marco Borga. Integrated high-resolution dataset of high-intensity European and Mediterranean flash floods. *Earth System Science Data*, 10(4) :1783–1794, October 2018. ISSN 1866-3508. doi : 10.5194/essd-10-1783-2018. URL https://essd.copernicus.org/articles/10/1783/2018/. Publisher : Copernicus GmbH.

Konstantinos M. Andreadis, Oliver E. J. Wing, Emma Colven, Colin J. Gleason, Paul D. Bates, and Casey M. Brown. Urbanizing the floodplain : global changes of imperviousness in flood-prone areas. *Environmental Research Letters*, 17(10) :104024, September 2022. ISSN 1748-9326. doi : 10.1088/1748-9326/ac9197. URL https://dx.doi.org/10.1088/1748-9326/ac9197. Publisher : IOP Publishing.

Heiko Apel, Sergiy Vorogushyn, and Bruno Merz. Brief communication : Impact forecasting could substantially improve the emergency management of deadly floods : case study July 2021 floods in Germany. *Natural Hazards and Earth System Sciences*, 22(9) :3005–3014, September 2022. ISSN 1561-8633. doi : 10.5194/nhess-22-3005-2022. URL https://nhess.copernicus.org/articles/22/3005/2022/. Publisher : Copernicus GmbH.

Moshe Armon, Francesco Marra, Yehouda Enzel, Dorita Rostkier-Edelstein, and Efrat Morin. Radar-based characterisation of heavy precipitation in the eastern Mediterranean and its representation in a convection-permitting model. *Hydrology and Earth System Sciences*, 24(3) :1227–1249, March 2020. ISSN 1027-5606. doi : 10.5194/hess-24-1227-2020. URL https://hess.copernicus.org/articles/24/1227/2020/. Publisher : Copernicus GmbH.

P. Arnaud, Yann Y. Aubert, D. Organde, P. Cantet, C. Fouchier, and N. Folton. Estimation de l'aléa hydrométéorologique par une méthode par simulation : la méthode SHYREG : présentation – performances – bases de données. *La Houille Blanche - Revue internationale de l'eau*, 2 :20, 2014. doi : 10.1051/lhb/2014012. URL https://hal.inrae.fr/hal-02573913.

Patrick Arnaud and Jacques Lavabre. Coupled rainfall model and discharge model for flood frequency estimation. *Water Resources Research*, 38(6) :11–1–11–11, 2002. ISSN 1944-7973. doi : 10.1029/2001WR000474. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2001WR000474. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2001WR000474.

Patrick Arnaud, Philippe Cantet, and Yoann Aubert. Relevance of an at-site flood frequency analysis method for extreme events based on stochastic simulation of hourly rainfall. *Hydrological Sciences Journal*, 61(1) :36–49, 2016. ISSN 0262-6667. doi : 10.1080/02626667.2014.965174. URL https://doi.org/10.1080/02626667.2014.965174. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/02626667.2014.965174.

Yann Aubert. *Estimation des valeurs extrêmes de débit par la méthode Shyreg : Réflexions sur l'équifinalité dans la modélisation de la transformation pluie en débit.* thesis, Doctorat de l'UPMC, 2012. URL https://hal.inrae.fr/tel-02606132. Pages : 318.

Yoann Aubert, Patrick Arnaud, Pierre Ribstein, and Jean-Alain Fine. The SHYREG flow method—application to 1605 basins in metropolitan France. *Hydrological Sciences Journal*, 59(5) :993–1005, May 2014. ISSN 0262-6667. doi : 10.1080/02626667.2014.902061. URL https://doi.org/10.1080/02626667.2014.902061. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/02626667.2014.902061.

Peter Auer, Nicolò Cesa-Bianchi, and Claudio Gentile. Adaptive and Self-Confident On-Line Learning Algorithms. *Journal of Computer and System Sciences*, 64(1) :48–75, February 2002. ISSN 0022-0000. doi : 10.1006/jcss.2001.1795. URL https://www.sciencedirect.com/science/article/pii/S0022000001917957.

Ludovic Auger, Olivier Dupont, Susanna Hagelin, Pierre Brousseau, and Pascal Brovelli. AROME–NWC : a new nowcasting tool based on an operational mesoscale forecasting system. *Quarterly Journal of the Royal Meteorological Society*, 141 (690) :1603–1611, 2015. ISSN 1477-870X. doi : 10.1002/qj.2463. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.2463. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2463.

J. C. Bartholmes, J. Thielen, M. H. Ramos, and S. Gentilini. The european flood alert system EFAS – Part 2 : Statistical skill assessment of probabilistic and deterministic operational forecasts. *Hydrology and Earth System Sciences*, 13(2) :141–153, February 2009. ISSN 1027-5606. doi : 10.5194/hess-13-141-2009. URL https://hess.copernicus.org/articles/13/141/2009/. Publisher : Copernicus GmbH.

Paul D. Bates. Flood Inundation Prediction. *Annual Review of Fluid Mechanics*, 54(Volume 54, 2022) :287–315, January 2022. ISSN 0066-4189, 1545-4479. doi : 10.1146/annurev-fluid-030121-113138. URL https://www.annualreviews.org/content/journals/10.1146/annurev-fluid-030121-113138. Publisher : Annual Reviews.

Paul D. Bates, Matthew S. Horritt, and Timothy J. Fewtrell. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology*, 387(1) :33–45, June 2010. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2010.03.027. URL https://www.sciencedirect.com/science/article/pii/S0022169410001538.

Pierre-Marie Bechon, Jérôme Le Coz, Isabelle Leleu, and Benjamin Renard. Des outils du réseau Etat en hydrométrie et leur ouverture aux notions d'incertitude. *Journées de l'hydraulique*, 35(1) :1–7, 2013. URL https://www.persee.fr/doc/jhydr_0000-0001_2013_act_35_1_1311. Publisher : Persée - Portail des revues scientifiques en SHS.

Anne Belleudy, Renaud Marty, Etienne Le Pape, Didier Narbais-Jaureguy, and Félicien Zuber. Vers la généralisation de la prévision hydrologique probabiliste au sein du réseau Vigicrues : estimation, évaluation et communication. *LHB*, 0(0) : 2374079, 2024. ISSN null. doi : 10.1080/27678490.2024.2374079. URL https://doi.org/10.1080/27678490.2024.2374079. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/27678490.2024.2374079.

Joseph Bellier, Isabella Zin, and Guillaume Bontron. Sample Stratification in Verification of Ensemble Forecasts of Continuous Scalar Variables : Potential Benefits and Pitfalls. *Monthly Weather Review*, 145 :3529–3544, September 2017. ISSN 0027-0644. doi : 10.1175/MWR-D-16-0487.1. URL https://ui.adsabs.harvard.edu/abs/2017MWRv..145.3529B. ADS Bibcode : 2017MWRv..145.3529B.

Stanley G. Benjamin, Stephen S. Weygandt, John M. Brown, Ming Hu, Curtis R. Alexander, Tatiana G. Smirnova, Joseph B. Olson, Eric P. James, David C. Dowell, Georg A. Grell, Haidao Lin, Steven E. Peckham, Tracy Lorraine Smith, William R. Moninger, Jaymes S. Kenyon, and Geoffrey S. Manikin. A North American Hourly Assimilation and Model Forecast Cycle : The Rapid Refresh. *Monthly Weather Review*, 144(4) :1669–1694, April 2016. ISSN 1520-0493, 0027-0644. doi : 10.1175/MWR-D-15-0242.1. URL https://journals.ametsoc.org/view/journals/mwre/144/4/mwr-d-15-0242.1.xml. Publisher : American Meteorological Society Section : Monthly Weather Review.

Marc Berenguer, Daniel Sempere-Torres, and Geoffrey G. S. Pegram. SBMcast – An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation. *Journal of Hydrology*, 404(3) :226–240, July 2011. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2011.04.033. URL https://www.sciencedirect.com/science/article/pii/S0022169411002940.

David Bourguignon. *Événements et territoires - le coût des inondations en France : analyses spatio-temporelles des dommages assurés*. These de doctorat, Montpellier 3, November 2014. URL https://theses.fr/2014MON30042.

François Bouttier and Laure Raynaud. Clustering and selection of boundary conditions for limited-area ensemble prediction. *Quarterly Journal of the Royal Meteorological Society*, 144(717) :2381–2391, 2018. ISSN 1477-870X. doi : 10.1002/qj.3304. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.3304. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3304.

Neill E. Bowler, Clive E. Pierce, and Alan W. Seed. STEPS : A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quarterly Journal of the Royal Meteorological Society*, 132(620) :2127–2155, 2006. ISSN 1477-870X. doi : 10.1256/qj.04.100. URL https://onlinelibrary.wiley.com/doi/abs/10.1256/qj.04.100. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1256/qj.04.100.

Glenn W. Brier. VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY. *Monthly Weather Review*, 78(1) :1–3, January 1950. ISSN 1520-0493, 0027-0644. doi : 10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/mwre/78/1/1520-0493_1950_078_0001_vofeit_2_0_co_2.xml. Publisher : American Meteorological Society Section : Monthly Weather Review.

Pierre Brigode, Sarah Vigoureux, Olivier Delestre, Pierre Nicolle, Olivier Payrastre, Raphaëlle Dreyfus, Stan Nomis, and Leslie Salvan. Inondations sur la Côte d'Azur : bilan hydro-météorologique des épisodes de 2015 et 2019. *LHB*, 107(1) :1–14, 2021. ISSN null. doi : 10.1080/27678490.2021.1976600. URL https://doi.org/10.1080/27678490.2021.1976600. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/27678490.2021.1976600.

Pascal Brunet and Christophe Bouvier. Retour d'expérience sur la crue du 12 septembre 2015 à Lodève (Hérault, France) : influence du karst sur les débits de pointe de crue. *La Houille Blanche*, 103(3) :39–46, 2017. ISSN 0018-6368. doi : 10.1051/lhb/2017020. URL https://doi.org/10.1051/lhb/2017020. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1051/lhb/2017020.

Peter Burek and Mikhail Smilovic. The use of GRDC gauging stations for calibrating large-scale hydrological models. *Earth System Science Data*, 15(12) :5617–5629, December 2023. ISSN 1866-3508. doi : 10.5194/essd-15-5617-2023. URL https://essd.copernicus.org/articles/15/5617/2023/. Publisher : Copernicus GmbH.

Peter Burek, Yusuke Satoh, Taher Kahil, Ting Tang, Peter Greve, Mikhail Smilovic, Luca Guillaumot, Fang Zhao, and Yoshihide Wada. Development of the Community Water Model (CWatM v1.04) – a high-resolution hydrological model for global and regional assessment of integrated water resources management. *Geoscientific Model Development*, 13(7) :3267–3298, July 2020. ISSN 1991-959X. doi : 10.5194/gmd-13-3267-2020. URL https://gmd.copernicus.org/articles/13/3267/2020/. Publisher : Copernicus GmbH.

Mike Bush, Tom Allen, Caroline Bain, Ian Boutle, John Edwards, Anke Finnenkoetter, Charmaine Franklin, Kirsty Hanley, Humphrey Lean, Adrian Lock, James Manners, Marion Mittermaier, Cyril Morcrette, Rachel North, Jon Petch, Chris Short, Simon Vosper, David Walters, Stuart Webster, Mark Weeks, Jonathan Wilkinson, Nigel Wood, and Mohamed Zerroukat. The first Met Office Unified Model–JULES Regional Atmosphere and Land configuration, RAL1. *Geoscientific Model Development*, 13(4) :1999–2029, April 2020. ISSN 1991-959X. doi : 10.5194/gmd-13-1999-2020. URL https://gmd.copernicus.org/articles/13/1999/2020/. Publisher : Copernicus GmbH.

Annette Sophie Bösmeier, Iso Himmelsbach, and Stefan Seeger. Reliability of flood marks and practical relevance for flood hazard assessment in southwestern Germany. *Natural Hazards and Earth System Sciences*, 22(9) :2963–2979, September 2022. ISSN 1561-8633. doi : 10.5194/nhess-22-2963-2022. URL https://nhess.copernicus.org/articles/22/2963/2022/. Publisher : Copernicus GmbH.

Martin Calianno, Isabelle Ruin, and Jonathan J. Gourley. Supplementing flash flood reports with impact classifications. *Journal of Hydrology*, 477 :1–16, January 2013. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2012.09.036. URL https://www.sciencedirect.com/science/article/pii/S0022169412008396.

G. Candille and O. Talagrand. Evaluation of probabilistic prediction systems for a scalar variable. *Quarterly Journal of the Royal Meteorological Society*, 131(609) :2131–2150, 2005. ISSN 1477-870X. doi : 10.1256/qj.04.71. URL https://onlinelibrary.wiley.com/doi/abs/10.1256/qj.04.71. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1256/qj.04.71.

Anthony Caruso, Alix Guillot, and Patrick Arnaud. Notice sur les indices de confiance de laméthode Shyreg-Débit - Définitions et calculs. Technical report, IRSTEA, 2013.

Céline Cattoën, Jono Conway, Nava Fedaeff, Daniel Lagrava, Paula Blackett, Kelsey Montgomery, Ude Shankar, Trevor Carey-Smith, Stuart Moore, Andrea Mari, Tilmann Steinmetz, and Sam Dean. A national flood awareness system for ungauged catchments in complex topography : The case of development, communication and evaluation in New Zealand. *Journal of Flood Risk Management*, n/a(n/a) :e12864, 2022. ISSN 1753-318X. doi : 10.1111/jfr3.12864. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jfr3.12864. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1111/jfr3.12864.

Olivier Caumont, Marc Mandement, François Bouttier, Judith Eeckman, Cindy Lebeaupin Brossier, Alexane Lovat, Olivier Nuissier, and Olivier Laurantin. The heavy precipitation event of 14–15 October 2018 in the Aude catchment : a meteorological study based on operational numerical weather prediction systems and standard and personal observations. *Natural Hazards and Earth System Sciences*, 21(3) :1135–1157, March 2021. ISSN 1561-8633. doi : 10.5194/nhess-21-1135-2021. URL https://nhess.copernicus.org/articles/21/1135/2021/. Publisher : Copernicus GmbH.

CCR. Les catastrophes naturelles en France - Bilan 1982-2018. Technical report, CCR, 2019.

CCR. Les catastrophes naturelles en France : bilan 1982-2020. Technical report, CCR, 2021. URL https://side.developpement-durable.gouv.fr/ACCIDR/doc/SYRACUSE/795441/les-catastrophes-naturelles-en-france-bilan-1982-2020.

CCR. Les catastrophes naturelles en France : bilan 1982-2022. Technical report, CCR, 2023a.

CCR. Conséquences du changement climatique sur le coût des catastrophes naturelles en France à l'horizon 2050. Technical report, CCR, 2023b. URL https://catastrophes-naturelles.ccr.fr/-/cons%C3%A9quences-du-changement-climatique-sur-le-co%C3%BBt-des-catastrophes-naturelles-en-france-%C3%A0-horizon-2050.

CCR. Efficacité des plans de prévention des risques d'inondation sur le coût des dommages assurés : estimation des impacts passés, présents et futurs. Technical report, 2023c.

CCR. Les catastrophes naturelles en France : bilan 1982-2023. Technical report, CCR, 2024. URL https://www.ccr.fr/-/les-catastrophes-naturelles-en-france-de-1982-%C3%A0-2023.

Nanée Chahinian, Matias Alcoba, Ndji dit Jacques Dembélé, Fréderic Cazenave, and Christophe Bouvier. Evaluation of an early flood warning system in Bamako (Mali) : Lessons learned from the flood of May 2019. *Journal of Flood Risk Management*, 16(3) :e12878, 2023. ISSN 1753-318X. doi : 10.1111/jfr3.12878. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jfr3.12878. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1111/jfr3.12878.

Jean-Louis Champeaux, Pascale Dupuy, Olivier Laurantin, Isabelle Soulan, Pierre Tabary, and Jean-Michel Soubeyroux. Les mesures de précipitations et l'estimation des lames d'eau à Météo-France : état de l'art et perspectives. *La Houille Blanche*, 95 (5) :28–34, October 2009. ISSN 0018-6368. doi : 10.1051/lhb/2009052. URL https://doi.org/10.1051/lhb/2009052. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1051/lhb/2009052.

Maryse Charpentier-Noyer. *Evaluation de prévisions hydrométéorologiques d'ensemble des crues soudaines à partir de données sur les impacts*. phdthesis, Nantes Université, December 2022. URL https://theses.hal.science/tel-04341104.

Maryse Charpentier-Noyer, Daniela Peredo, Axelle Fleury, Hugo Marchal, François Bouttier, Eric Gaume, Pierre Nicolle, Olivier Payrastre, and Maria-Helena Ramos. A methodological framework for the evaluation of short-range flash-flood hydrometeorological forecasts at the event scale. *Natural Hazards and Earth System Sciences*, 23(6) :2001–2029, June 2023. ISSN 1561-8633. doi : 10.5194/nhess-23-2001-2023. URL https://nhess.copernicus.org/articles/23/2001/2023/. Publisher : Copernicus GmbH.

Maryse Charpentier-Noyer, Olivier Payrastre, Eric Gaume, Pierre Nicolle, François Bouttier, Axelle Fleury, and Hugo Marchal. An agent-based modeling of rescue operations for the evaluation of short-range flash flood forecasts. *Journal of Hydrology*, page 133048, March 2025. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2025.133048. URL https://www.sciencedirect.com/science/article/pii/S0022169425003865.

Martyn P. Clark, David E. Rupp, Ross A. Woods, Xiaogu Zheng, Richard P. Ibbitt, Andrew G. Slater, Jochen Schmidt, and Michael J. Uddstrom. Hydrological data assimilation with the ensemble Kalman filter : Use of streamflow observations to update states in a distributed hydrological model. *Advances in Water Resources*, 31(10) :1309–1324, October 2008. ISSN 0309-1708. doi : 10.1016/j.advwatres.2008.06.005. URL https://www.sciencedirect.com/science/article/pii/S0309170808001012.

Peter Clark, Nigel Roberts, Humphrey Lean, Susan P. Ballard, and Cristina Charlton-Perez. Convection-permitting models : a step-change in rainfall forecasting. *Meteorological Applications*, 23(2) :165–181, 2016. ISSN 1469-8080. doi : 10.1002/met.1538. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/met.1538. _eprint : https ://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/met.1538.

Robert A. Clark, Jonathan J. Gourley, Zachary L. Flamig, Yang Hong, and Edward Clark. CONUS-Wide Evaluation of National Weather Service Flash Flood Guidance Products. *Weather and Forecasting*, 29(2) :377–392, 2014. ISSN 1520-0434, 0882-8156. doi : 10.1175/WAF-D-12-00124.1. URL https://journals.ametsoc.org/view/journals/wefo/29/2/waf-d-12-00124_1.xml. Publisher : American Meteorological Society Section : Weather and Forecasting.

Steven J. Cole and Robert J. Moore. Distributed hydrological modelling using weather radar in gauged and ungauged basins. *Advances in Water Resources*, 32(7) :1107–1120, July 2009. ISSN 0309-1708. doi : 10.1016/j.advwatres.2009.01.006. URL https://www.sciencedirect.com/science/article/pii/S0309170809000177.

Steven J. Cole, Robert J. Moore, Steven C. Wells, and Paul S. Mattingley. Real-time forecasts of flood hazard and impact : some UK experiences. *E3S Web of Conferences*, 7 :18015, 2016. ISSN 2267-1242. doi : 10.1051/e3sconf/20160718015. URL https://www.e3s-conferences.org/articles/e3sconf/abs/2016/02/e3sconf_flood2016_18015/e3sconf_flood2016_18015.html. Publisher : EDP Sciences.

Daniel Coles, Dapeng Yu, Robert L. Wilby, Daniel Green, and Zara Herring. Beyond 'flood hotspots' : Modelling emergency service accessibility during flooding in York, UK. *Journal of Hydrology*, 546 :419–436, March 2017. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2016.12.013. URL https://www.sciencedirect.com/science/article/pii/S0022169416308022.

François Colleoni and Ngo Nghi Truyen Huynh. SMASH, 2024. URL https://zenodo.org/records/10469390.

François Colleoni, Ngo Nghi Truyen Huynh, Renard Benjamin, Thomas De Fournas, Apolline El Baz, and Pierre-André Garambois. SMASH : Version 1.0.0-rc2.21, April 2024. URL https://zenodo.org/records/10991739.

François Colleoni, Ngo Nghi Truyen Huynh, Pierre-André Garambois, Maxime Jay-Allemand, Didier Organde, Benjamin Renard, Thomas De Fournas, Apolline El Baz, Julie Demargne, and Pierre Javelle. SMASH v1.0 : A Differentiable and Regionalizable High-Resolution Hydrological Modeling and Data Assimilation Framework. *EGUsphere*, pages 1–36, March 2025. doi : 10.5194/egusphere-2025-690. URL https://egusphere.copernicus.org/preprints/2025/egusphere-2025-690/. Publisher : Copernicus GmbH.

C. G. Collier. Flash flood forecasting : What are the limits of predictability ? *Quarterly Journal of the Royal Meteorological Society*, 133(622) :3–23, January 2007. ISSN 0035-9009. doi : 10.1002/qj.29. URL https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.29. Publisher : John Wiley & Sons, Ltd.

C. Corral, D. Velasco, D. Forcadell, D. empere Torres, and E. Velasco. Advances in radar-based flood warning systems. The EHIMI system and the experience in the Besòs flash-flood pilot basin. In *Flood Risk Management : Research and Practice*. CRC Press, 2008. ISBN 978-0-429-20712-9. Num Pages : 1.

Carles Corral, Marc Berenguer, Daniel Sempere-Torres, Laura Poletti, Francesco Silvestro, and Nicola Rebora. Comparison of two early warning systems for regional flash flood hazard forecasting. *Journal of Hydrology*, 572 :603–619, May 2019. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2019.03.026. URL https://www.sciencedirect.com/science/article/pii/S0022169419302665.

Pierfranco Costabile, Carmelina Costanzo, John Kalogiros, and Vasilis Bellos. Toward Street-Level Nowcasting of Flash Floods Impacts Based on HPC Hydrodynamic Modeling at the Watershed Scale and High-Resolution Weather Radar Data. *Water Resources Research*, 59(10) :e2023WR034599, 2023. ISSN 1944-7973. doi : 10.1029/2023WR034599. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2023WR034599. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2023WR034599.

CRED. 2023 Disasters in Numbers : A Significant Year of Disaster Impact. Technical report, CRED, 2024. URL https://civil-protection-knowledge-network.europa.eu/media/2023-disasters-numbers-significant-year-disaster-impact.

HELEN N. Davies and Victoria Bell. Assessment of methods for extracting low-resolution river networks from high-resolution digital data. *Hydrological Sciences Journal*, 54(1) :17–28, February 2009. ISSN 0262-6667. doi : 10.1623/hysj.54.1.17. URL https://doi.org/10.1623/hysj.54.1.17. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1623/hysj.54.1.17.

S. Davolio, M. M. Miglietta, T. Diomede, C. Marsigli, and A. Montani. A flood episode in northern Italy : multi-model and single-model mesoscale meteorological ensembles for hydrological predictions. *Hydrology and Earth System Sciences*, 17 (6) :2107–2120, June 2013. ISSN 1027-5606. doi : 10.5194/hess-17-2107-2013. URL https://hess.copernicus.org/articles/17/2107/2013/. Publisher : Copernicus GmbH.

Silvio Davolio, Francesco Silvestro, and Piero Malguzzi. Effects of Increasing Horizontal Resolution in a Convection-Permitting Model on Flood Forecasting : The 2011 Dramatic Events in Liguria, Italy. *Journal of Hydrometeorology*, 16(4) :1843–1856, August 2015. ISSN 1525-7541, 1525-755X. doi : 10.1175/JHM-D-14-0094.1. URL https://journals.ametsoc.org/view/journals/hydr/16/4/jhm-d-14-0094_1.xml. Publisher : American Meteorological Society Section : Journal of Hydrometeorology.

Silvio Davolio, Francesco Silvestro, and Thomas Gastaldo. Impact of Rainfall Assimilation on High-Resolution Hydrometeorological Forecasts over Liguria, Italy. *Journal of Hydrometeorology*, 18(10) :2659–2680, October 2017. ISSN 1525-7541, 1525-755X. doi : 10.1175/JHM-D-17-0073.1. URL https://journals.ametsoc.org/view/journals/hydr/18/10/jhm-d-17-0073_1.xml. Publisher : American Meteorological Society Section : Journal of Hydrometeorology.

Philippe Davy, Thomas Croissant, and Dimitri Lague. A precipiton method to calculate river hydrodynamics, with applications to flood prediction, landscape evolution models, and braiding instabilities. *Journal of Geophysical Research : Earth Surface*, 122(8) :1491–1512, 2017. ISSN 2169-9011. doi : 10.1002/2016JF004156. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/2016JF004156. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/2016JF004156.

Jens A. de Bruijn, Hans de Moel, Brenden Jongman, Marleen C. de Ruiter, Jurjen Wagemaker, and Jeroen C. J. H. Aerts. A global database of historic and real-time flood events based on social media. *Scientific Data*, 6(1) :311, December 2019. ISSN 2052-4463. doi : 10.1038/s41597-019-0326-9. URL https://www.nature.com/articles/s41597-019-0326-9. Publisher : Nature Publishing Group.

H. de Moel, B. Jongman, H. Kreibich, B. Merz, E. Penning-Rowsell, and P. J. Ward. Flood risk assessments at different spatial scales. *Mitigation and Adaptation Strategies for Global Change*, 20(6) :865–890, August 2015. ISSN 1573-1596. doi : 10.1007/s11027-015-9654-z. URL https://doi.org/10.1007/s11027-015-9654-z.

Osvaldo Luiz Leal de Moraes. An impact-based forecast system developed for hydrometeorological hazards. *International Journal of Disaster Risk Reduction*, 93 :103803, 2023. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2023.103803. URL https://www.sciencedirect.com/science/article/pii/S2212420923002832.

Marleen C. de Ruiter, Jens A. de Bruijn, Johanna Englhardt, James E. Daniell, Hans de Moel, and Philip J. Ward. The Asynergies of Structural Disaster Risk Reduction Measures : Comparing Floods and Earthquakes. *Earth's Future*, 9(1) :e2020EF001531, 2021. ISSN 2328-4277. doi : 10.1029/2020EF001531. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2020EF001531. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2020EF001531.

L. Descamps, C. Labadie, A. Joly, E. Bazile, P. Arbogast, and P. Cébron. PEARP, the Météo-France short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 141(690) :1671–1685, 2015. ISSN 1477-870X. doi : 10.1002/qj.2469. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.2469. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2469.

Marie Devaine, Pierre Gaillard, Yannig Goude, and Gilles Stoltz. Forecasting electricity consumption by aggregating specialized experts. *Machine Learning*, 90(2) :231–260, February 2013. ISSN 1573-0565. doi : 10.1007/s10994-012-5314-7. URL https://doi.org/10.1007/s10994-012-5314-7.

G. Di Baldassarre and A. Montanari. Uncertainty in river discharge observations : a quantitative analysis. *Hydrology and Earth System Sciences*, 13(6) :913–921, June 2009. ISSN 1027-5606. doi : 10.5194/hess-13-913-2009. URL https://hess.copernicus.org/articles/13/913/2009/. Publisher : Copernicus GmbH.

Thomas Dias. *Collecte, critique et utilisation de données multisources pour l'évaluation d'une modélisation d'impacts de crues soudaines dans l'arc méditerranéen*. PhD thesis, Université de Strasbourg, 2024.

Qifan Ding and Patrick Arnaud. Taking Account of Seasonality in a Regional Flood Frequency Estimation Approach Based on Event Simulations. *Water*, 14(9) :1376, 2022. ISSN 2073-4441. doi : 10.3390/w14091376. URL https://www.mdpi.com/2073-4441/14/9/1376. Number : 9 Publisher : Multidisciplinary Digital Publishing Institute.

Qifan Ding and Patrick Arnaud. Using supervised machine learning for regional hydrological hazard estimation in metropolitan France. *Journal of Hydrology : Regional Studies*, 54 :101872, 2024. ISSN 2214-5818. doi : 10.1016/j.ejrh.2024.101872. URL https://www.sciencedirect.com/science/article/pii/S2214581824002209.

Francesco Dottori, Milan Kalas, Peter Salamon, Alessandra Bianchi, J Thielen Del Pozo, and Luc Feyen. A near real-time procedure for flood hazard mapping and risk assessment in Europe. 2015. URL http://89.31.100.18/~iahrpapers/85343.pdf.

Francesco Dottori, Milan Kalas, Peter Salamon, Alessandra Bianchi, Lorenzo Alfieri, and Luc Feyen. An operational procedure for rapid flood risk assessment in Europe. *Natural Hazards and Earth System Sciences*, 17(7) :1111–1126, 2017. ISSN 1561-8633. doi : 10.5194/nhess-17-1111-2017. URL https://nhess.copernicus.org/articles/17/1111/2017/. Publisher : Copernicus GmbH.

Francesco Dottori, Lorenzo Alfieri, Alessandra Bianchi, Jon Skoien, and Peter Salamon. A new dataset of river flood hazard maps for Europe and the Mediterranean Basin. *Earth System Science Data*, 14(4) :1549–1569, April 2022. ISSN 1866-3508. doi : 10.5194/essd-14-1549-2022. URL https://essd.copernicus.org/articles/14/1549/2022/. Publisher : Copernicus GmbH.

Francesco Dottori, Lorenzo Mentaschi, Alessandra Bianchi, Lorenzo Alfieri, and Luc Feyen. Cost-effective adaptation strategies to rising river flood risk in Europe. *Nature Climate Change*, 13(2) :196–202, 2023. ISSN 1758-6798. doi : 10.1038/s41558-022-01540-0. URL https://www.nature.com/articles/s41558-022-01540-0. Number : 2 Publisher : Nature Publishing Group.

Johnny Douvinet. Intérêts et limites des données CatNat pour un inventaire des inondations. L'exemple des crues rapides liées à de violents orages (Bassin parisien, Nord de la France). *Norois*, (201) :17–30, 2006. ISSN 0029-182X, 1760-8546. doi : 10.4000/norois.1733. URL http://journals.openedition.org/norois/1733.

Johnny Douvinet and Freddy Vinet. La carte des arrêtés CatNat pour les inondations : limites et améliorations possibles. *Mappemonde*, (107), 2012.

Véronique Ducrocq, Brice Boudevillain, Christophe Bouvier, Isabelle Braud, Nadia Fourrie, Cindy Lebeaupin-Brossier, Pierre Javelle, Olivier Nuissier, Olivier Payrastre, Hélène Roux, Isabelle Ruin, and Béatrice Vincendon. Le programme HYMEX – Connaissances et prévision des pluies intenses et crues rapides en région méditerranéenne. *La Houille Blanche*, 105(3-4) : 5–12, 2019. ISSN 0018-6368. doi : 10.1051/lhb/2019048. URL https://doi.org/10.1051/lhb/2019048. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1051/lhb/2019048.

Petra Döll and Bernhard Lehner. Validation of a new global 30-min drainage direction map. *Journal of Hydrology*, 258(1) : 214–231, February 2002. ISSN 0022-1694. doi : 10.1016/S0022-1694(01)00565-0. URL https://www.sciencedirect.com/science/article/pii/S0022169401005650.

Edijatno and Claude Michel. Un modèle pluie-débit journalier à trois paramètres. *La Houille Blanche*, 75(2) :113–122, 1989. ISSN 0018-6368. doi : 10.1051/lhb/1989007. URL https://doi.org/10.1051/lhb/1989007. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1051/lhb/1989007.

Simon Edouard. *Prévision d'ensemble des crues rapides méditerranéennes*. phdthesis, Université Paul Sabatier - Toulouse III, December 2016. URL https://theses.hal.science/tel-01761961.

Dirk Eilander, Willem van Verseveld, Dai Yamazaki, Albrecht Weerts, Hessel C. Winsemius, and Philip J. Ward. A hydrography upscaling method for scale-invariant parametrization of distributed hydrological models. *Hydrology and Earth System Sciences*, 25(9) :5287–5313, September 2021. ISSN 1027-5606. doi : 10.5194/hess-25-5287-2021. URL https://hess.copernicus.org/articles/25/5287/2021/. Publisher : Copernicus GmbH.

{Environment Agency}. LIDAR Composite Digital Surface Model (DSM) - 1m, June 2024. URL https://www.data.gov.uk/dataset/cf3f1137-c12b-44a1-a835-e80fe4a60b92/lidar-composite-digital-surface-model-dsm-1m.

Katrin Erdlenbruch and Serge Garcia. Choix résidentiels : quelles informations concernant les inondations sont prises en compte par les ménages ? Technical Report 28, INRAE, 2022.

Guilhem Fabre. La catastrophe hydrologique éclair de Nîmes (3 octobre 1988) (The flash flood disaster at Nîmes on 3 octobre 1988). *Bulletin de l'Association de Géographes Français*, 67(2) :113–122, 1990. doi : 10.3406/bagf.1990.1522. URL https://www.persee.fr/doc/bagf_0004-5322_1990_num_67_2_1522. Company : Persée - Portail des revues scientifiques en SHS Distributor : Persée - Portail des revues scientifiques en SHS Institution : Persée - Portail des revues scientifiques en SHS Label : Persée - Portail des revues scientifiques en SHS Publisher : Association de Géographes Français.

Balázs M. Fekete, Charles J. Vörösmarty, and Richard B. Lammers. Scaling gridded river networks for macroscale hydrology : Development, analysis, and control of error. *Water Resources Research*, 37(7) :1955–1967, 2001. ISSN 1944-7973. doi : 10.1029/2001WR900024. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2001WR900024. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2001WR900024.

Balázs M. Fekete, Charles J. Vörösmarty, and Wolfgang Grabs. High-resolution fields of global runoff combining observed river discharge and simulated water balances. *Global Biogeochemical Cycles*, 16(3) :15–1–15–10, 2002. ISSN 1944-9224. doi : 10.1029/1999GB001254. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/1999GB001254. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/1999GB001254.

Ayan Fleischmann, Rodrigo Paiva, and Walter Collischonn. Can regional to continental river hydrodynamic models be locally relevant ? A cross-scale comparison. *Journal of Hydrology X*, 3 :100027, April 2019. ISSN 2589-9155. doi : 10.1016/j.hydroa.2019.100027. URL https://www.sciencedirect.com/science/article/pii/S2589915519300112.

Nathalie Folton and Jacques Lavabre. Approche par modélisation PLUIE-DEBIT pour la connaissance régionale de la ressource en eau : application à la moitié du territoire français. *La Houille Blanche*, 93(3) :64–70, 2007. ISSN 0018-6368. doi : 10.1051/lhb:2007037. URL https://doi.org/10.1051/lhb:2007037. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1051/lhb :2007037.

Catherine Fouchier. *Développement d'une méthodologie pour la connaissance régionale des crues*. These de doctorat, Montpellier 2, November 2010. URL https://theses.fr/2010MON20226.

Catherine Fouchier, Clotilde Saint Martin, Pierre Javelle, P. Meriaux, D. Organde, and J. Demargne. Mises en oeuvre opérationnelles de la méthode AIGA pour anticiper les crues sur les cours d'eau non surveillés. *Sciences Eaux & Territoires*, (23) : 48–55, 2017. doi : 10.14758/SET-REVUE.2017.23.09. URL https://hal.archives-ouvertes.fr/hal-01538553. Publisher : INRAE.

Catherine Frizzle, Mélanie Trudel, Sylvie Daniel, Antoine Pruneau, and Juzer Noman. LiDAR topo-bathymetry for riverbed elevation assessment : A review of approaches and performance for hydrodynamic modelling of flood plains. *Earth Surface Processes and Landforms*, n/a(n/a) :1–16, 2024. ISSN 1096-9837. doi : 10.1002/esp.5808. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/esp.5808. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/esp.5808.

Kenji Fujiki and Olivier Finance. Exposition et vulnérabilité sociale des villes françaises au risque inondation : une analyse spatiotemporelle à fine échelle (1999-2017). *Cybergeo : European Journal of Geography*, June 2022. ISSN 1278-3366. doi : 10.4000/cybergeo.39179. URL https://journals.openedition.org/cybergeo/39179#tocto2n5. Publisher : CNRS-UMR Géographie-cités 8504.

Vanessa Fundel, Nadine Fleischhut, Stefan Herzog, Martin Göber, and Renate Hagedorn. Promoting the use of probabilistic weather forecasts through a dialogue between scientists, developers, and end-users. *Quarterly Journal of the Royal Meteorological Society*, 145, 2019. doi : 10.1002/qj.3482.

Luca Furnari, Giuseppe Mendicino, and Alfonso Senatore. Hydrometeorological Ensemble Forecast of a Highly Localized Convective Event in the Mediterranean. *Water*, 12(6) :1545, June 2020. ISSN 2073-4441. doi : 10.3390/w12061545. URL https://www.mdpi.com/2073-4441/12/6/1545. Publisher : Multidisciplinary Digital Publishing Institute.

E. Gaume and M. Borga. Etat des connaissances récentes acquises sur les crues éclair en Europe : bilan du projet de recherches européen HYDRATE (2006-2010). *La Houille Blanche*, (2) :24–30, April 2013. ISSN 0018-6368, 1958-5551. doi : 10.1051/lhb/2013012. URL https://www.shf-lhb.org/articles/lhb/abs/2013/02/lhb2013012/lhb2013012.html. Number : 2 Publisher : Société Hydrotechnique de France.

Eric Gaume, Marc Livet, Michel Desbordes, and J. P Villeneuve. Hydrological analysis of the river Aude, France, flash flood on 12 and 13 November 1999. *Journal of Hydrology*, 286(1) :135–154, January 2004. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2003.09.015. URL https://www.sciencedirect.com/science/article/pii/S0022169403003780.

Eric Gaume, Valerie Bain, Pietro Bernardara, Olivier Newinger, Mihai Barbuc, Allen Bateman, Lotta Blaškovičová, Günter Blöschl, Marco Borga, Alexandru Dumitrescu, Ioannis Daliakopoulos, Joachim Garcia, Anisoara Irimescu, Silvia Kohnova, Aristeidis Koutroulis, Lorenzo Marchi, Simona Matreata, Vicente Medina, Emanuele Preciso, Daniel Sempere-Torres, Gheorghe Stancalie, Jan Szolgay, Ioannis Tsanis, David Velasco, and Alberto Viglione. A compilation of data on European flash floods. *Journal of Hydrology*, 367(1) :70–78, 2009. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2008.12.028. URL https://www.sciencedirect.com/science/article/pii/S0022169409000079.

Eric Gaume, Marco Borga, Maria Carmen Llasat, Said Maouche, Michel Lang, and Michalis Diakakis. Sub-chapter 1.3.4. Mediterranean extreme floods and flash floods. In Jean-Paul Moatti and Stéphane Thiébault, editors, *The Mediterranean region under climate change : A scientific update*, Synthèses, pages 133–144. IRD Éditions, Marseille, 2016. ISBN 978-2-7099-2220-3. doi : 10.4000/books.irdeditions.23181. URL https://books.openedition.org/irdeditions/23181. Code : The Mediterranean region under climate change : A scientific update.

Asmamaw Gebrehiwot and Leila Hashemi-Beni. 3D Inundation Mapping : A Comparison Between Deep Learning Image Classification and Geomorphic Flood Index Approaches. *Frontiers in Remote Sensing*, 3, 2022. ISSN 2673-6187. URL https://www.frontiersin.org/articles/10.3389/frsen.2022.868104.

Konstantine P. Georgakakos. On the Design of National, Real-Time Warning Systems with Capability for Site-Specific, Flash-Flood Forecasts. *Bulletin of the American Meteorological Society*, 67(10) :1233–1239, October 1986. ISSN 0003-0007, 1520-0477. doi : 10.1175/1520-0477(1986)067<1233:OTDONR>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/bams/67/10/1520-0477_1986_067_1233_otdonr_2_0_co_2.xml. Publisher : American Meteorological Society Section : Bulletin of the American Meteorological Society.

Joseph P. Gerrity. A Note on Gandin and Murphy's Equitable Skill Score. *Monthly Weather Review*, 120(11) :2709–2712, November 1992. ISSN 1520-0493, 0027-0644. doi : 10.1175/1520-0493(1992)120<2709:ANOGAM>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/mwre/120/11/1520-0493_1992_120_2709_anogam_2_0_co_2.xml. Publisher : American Meteorological Society Section : Monthly Weather Review.

Juliette Godet, Olivier Payrastre, Pierre Javelle, and François Bouttier. Assessing the ability of a new seamless short-range ensemble rainfall product to anticipate flash floods in the French Mediterranean area. *Natural Hazards and Earth System Sciences*, 23(11) :3355–3377, 2023. ISSN 1561-8633. doi : 10.5194/nhess-23-3355-2023. URL https://nhess.copernicus.org/articles/23/3355/2023/. Publisher : Copernicus GmbH.

Juliette Godet, Eric Gaume, Pierre Javelle, Pierre Nicolle, and Olivier Payrastre. Technical note : Comparing three different methods for allocating river points to coarse-resolution hydrological modelling grid cells. *Hydrology and Earth System Sciences*, 28(6) :1403–1413, 2024. ISSN 1027-5606. doi : 10.5194/hess-28-1403-2024. URL https://hess.copernicus.org/articles/28/1403/2024/. Publisher : Copernicus GmbH.

Juliette Godet, Pierre Nicolle, Nabil Hocini, Eric Gaume, Philippe Davy, Frederic Pons, Pierre Javelle, Pierre-André Garambois, Dimitri Lague, and Olivier Payrastre. Benchmark dataset for hydraulic simulations of flash floods in the French Mediterranean region. *Earth System Science Data*, 17(6) :2963–2983, June 2025a. ISSN 1866-3508. doi : 10.5194/essd-17-2963-2025. URL https://essd.copernicus.org/articles/17/2963/2025/. Publisher : Copernicus GmbH.

Juliette Godet, Olivier Payrastre, Qifan Ding, Julie Demargne, Eric Gaume, Pierre Nicolle, Anne Belleudy, and Pierre Javelle. Evaluating the French Flash Flood Warning System Using Hydrological and Impact Data in Southeastern France. *Journal of Flood Risk Management*, 18(2) :e70053, 2025b. ISSN 1753-318X. doi : 10.1111/jfr3.70053. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jfr3.70053. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1111/jfr3.70053.

Brian Golding, editor. *Towards the "Perfect" Weather Warning : Bridging Disciplinary Gaps through Partnership and Communication*. Springer Nature, 2022. doi : 10.1007/978-3-030-98989-7. URL https://library.oapen.org/handle/20.500.12657/57367. Accepted : 2022-07-13T12 :27 :53Z.

Lixia Gong, Liqiang An, Mingzhong Liu, and Jingfa Zhang. Road damage detection from high-resolution RS image. In *2012 IEEE International Geoscience and Remote Sensing Symposium*, pages 990–993, Munich, Germany, July 2012. IEEE. doi : 10.1109/IGARSS.2012.6351235. URL http://ieeexplore.ieee.org/document/6351235/.

Jonathan J. Gourley, Zachary L. Flamig, Humberto Vergara, Pierre-Emmanuel Kirstetter, Robert A. Clark, Elizabeth Argyle, Ami Arthur, Steven Martinaitis, Galateia Terti, Jessica M. Erlingis, Yang Hong, and Kenneth W. Howard. The FLASH Project : Improving the Tools for Flash Flood Monitoring and Prediction across the United States. *Bulletin of the American Meteorological Society*, 98(2) :361–372, 2017. ISSN 0003-0007, 1520-0477. doi : 10.1175/BAMS-D-15-00247.1. URL https://journals.ametsoc.org/view/journals/bams/98/2/bams-d-15-00247.1.xml. Publisher : American Meteorological Society Section : Bulletin of the American Meteorological Society.

Gaby Joanne Gründemann, Nick van de Giesen, Lukas Brunner, and Ruud van der Ent. Rarest rainfall events will see the greatest relative increase in magnitude under future climate change. *Communications Earth & Environment*, 3(1) :1–9, October 2022. ISSN 2662-4435. doi : 10.1038/s43247-022-00558-8. URL https://www.nature.com/articles/s43247-022-00558-8. Publisher : Nature Publishing Group.

Hoshin V. Gupta, Harald Kling, Koray K. Yilmaz, and Guillermo F. Martinez. Decomposition of the mean squared error and NSE performance criteria : Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1) :80–91, 2009. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2009.08.003. URL https://www.sciencedirect.com/science/article/pii/S0022169409004843.

A. Hally, O. Caumont, L. Garrote, E. Richard, A. Weerts, F. Delogu, E. Fiori, N. Rebora, A. Parodi, A. Mihalović, M. Ivković, L. Dekić, W. van Verseveld, O. Nuissier, V. Ducrocq, D. D'Agostino, A. Galizia, E. Danovaro, and A. Clematis. Hydrometeorological multi-model ensemble simulations of the 4 November 2011 flash flood event in Genoa, Italy, in the framework of the DRIHM project. *Natural Hazards and Earth System Sciences*, 15(3) :537–555, March 2015. ISSN 1561-8633. doi : 10.5194/nhess-15-537-2015. URL https://nhess.copernicus.org/articles/15/537/2015/. Publisher : Copernicus GmbH.

Thomas M. Hamill. Interpretation of Rank Histograms for Verifying Ensemble Forecasts. *Monthly Weather Review*, 129 (3) :550–560, March 2001. ISSN 1520-0493, 0027-0644. doi : 10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/mwre/129/3/1520-0493_2001_129_0550_iorhfv_2.0.co_2.xml. Publisher : American Meteorological Society Section : Monthly Weather Review.

H. a. P. Hapuarachchi, Q. J. Wang, and T. C. Pagano. A review of advances in flash flood forecasting. *Hydrological Processes*, 25(18) :2771–2784, 2011. ISSN 1099-1085. doi : 10.1002/hyp.8040. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.8040. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/hyp.8040.

Sara Harrison, Sally Potter, Raj Prasanna, Emma Doyle, and David Johnston. Volunteered Geographic Information for people-centred severe weather early warning : A literature review. *Australasian Journal of Disaster and Trauma Studies*, 24(1) : 3–21, 2020.

Sara E. Harrison, Sally H. Potter, Raj Prasanna, Emma E. H. Doyle, and David Johnston. 'Where oh where is the data¿ : Identifying data sources for hydrometeorological impact forecasts and warnings in Aotearoa New Zealand. *International Journal of Disaster Risk Reduction*, 66 :102619, December 2021. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2021.102619. URL https://www.sciencedirect.com/science/article/pii/S221242092100580X.

Sara E. Harrison, Sally H. Potter, Raj Prasanna, Emma E. H. Doyle, and David Johnston. Identifying the Impact-Related Data Uses and Gaps for Hydrometeorological Impact Forecasts and Warnings. *Weather, Climate, and Society*, 14(1) :155–176, January 2022a. ISSN 1948-8327, 1948-8335. doi : 10.1175/WCAS-D-21-0093.1. URL https://journals.ametsoc.org/view/journals/wcas/14/1/WCAS-D-21-0093.1.xml. Publisher : American Meteorological Society Section : Weather, Climate, and Society.

Sara E. Harrison, Sally H. Potter, Raj Prasanna, Emma E. H. Doyle, and David Johnston. 'Sharing is caring' : A socio-technical analysis of the sharing and governing of hydrometeorological hazard, impact, vulnerability, and exposure data in Aotearoa New Zealand. *Progress in Disaster Science*, 13 :100213, January 2022b. ISSN 2590-0617. doi : 10.1016/j.pdisas.2021.100213. URL https://www.sciencedirect.com/science/article/pii/S2590061721000739.

Sara E. Harrison, Sally H. Potter, Raj Prasanna, Emma E. H. Doyle, and David Johnston. Nurturing partnerships to support data access for impact forecasts and warnings : Theoretical integration and synthesis. *International Journal of Disaster Risk Reduction*, 105 :104395, April 2024. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2024.104395. URL https://www.sciencedirect.com/science/article/pii/S2212420924001572.

Rebecca Hemingway and Joanne Robbins. Developing a hazard-impact model to support impact-based forecasts and warnings : The Vehicle OverTurning (VOT) Model. *Meteorological Applications*, 27(1) :e1819, 2020. ISSN 1469-8080. doi : 10.1002/met.1819. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/met.1819. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/met.1819.

Hans Hersbach. Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. *Weather and Forecasting*, 15(5) :559–570, October 2000. ISSN 1520-0434, 0882-8156. doi : 10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/wefo/15/5/1520-0434_2000_015_0559_dotcrp_2_0_co_2.xml. Publisher : American Meteorological Society Section : Weather and Forecasting.

Nabil Hocini. *Evaluation de méthodes automatisées de cartographie des zones inondables adaptées à la prévision des crues soudaines*. These de doctorat, Nantes Université, March 2022. URL https://www.theses.fr/2022NANU4015.

Nabil Hocini, Olivier Payrastre, François Bourgin, Eric Gaume, Philippe Davy, Dimitri Lague, Lea Poinsignon, and Frederic Pons. Performance of automated methods for flash flood inundation mapping : a comparison of a digital terrain model (DTM) filling and two hydrodynamic methods. *Hydrology and Earth System Sciences*, 25(6) :2979–2995, June 2021. ISSN 1027-5606. doi : 10.5194/hess-25-2979-2021. URL https://hess.copernicus.org/articles/25/2979/2021/. Publisher : Copernicus GmbH.

Ngo Nghi Truyen Huynh. *Hybrid Physics-AI Approaches Integrating Neural Networks into Distributed Hydrological Models for Regional Flood Simulation*. phdthesis, Aix-Marseile Université, June 2025. URL https://hal.science/tel-05151541.

Ngo Nghi Truyen Huynh, Pierre-André Garambois, François Colleoni, and Pierre Javelle. Signatures-and-sensitivity-based multi-criteria variational calibration for distributed hydrological modeling applied to Mediterranean floods. *Journal of Hydrology*, 625 :129992, October 2023. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2023.129992. URL https://www.sciencedirect.com/science/article/pii/S0022169423009344.

IGN. LiDAR HD, 2024. URL https://www.data.gouv.fr/fr/datasets/lidar-hd/.

R. O. Imhoff, C. C. Brauer, K. J. van Heeringen, R. Uijlenhoet, and A. H. Weerts. Large-Sample Evaluation of Radar Rainfall Nowcasting for Flood Early Warning. *Water Resources Research*, 58(3) :e2021WR031591, 2022. ISSN 1944-7973. doi : 10.1029/2021WR031591. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2021WR031591. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2021WR031591.

INSEE. 37,8 millions de logements en France au 1er janvier 2023, October 2023. URL https://www.insee.fr/fr/statistiques/7700305.

Pierre Javelle, Julie Demargne, Dimitri Defrance, Jean Pansu, and Patrick Arnaud. Evaluating flash-flood warnings at ungauged locations using post-event surveys : a case study with the AIGA warning system. *Hydrological Sciences Journal*, 59(7) : 1390–1402, July 2014. ISSN 0262-6667. doi : 10.1080/02626667.2014.923970. URL https://doi.org/10.1080/02626667.2014.923970. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/02626667.2014.923970.

Pierre Javelle, Didier Organde, Julie Demargne, Clotilde Saint-Martin, Céline de Saint-Aubin, Léa Garandeau, and Bruno Janet. Setting up a French national flash flood warning system for ungauged catchments based on the AIGA method. *E3S Web of Conferences*, 7 :18010, 2016. ISSN 2267-1242. doi : 10.1051/e3sconf/20160718010. URL https://www.e3s-conferences.org/articles/e3sconf/abs/2016/02/e3sconf_flood2016_18010/e3sconf_flood2016_18010.html. Publisher : EDP Sciences.

Pierre Javelle, Clotilde Saint Martin, Léa Garandeau, and Bruno Janet. Flash flood warnings : Recent achievements in France with the national Vigicrues Flash system | UNDRR. *Contributing Paper to GAR 2019*, 2019. URL http://www.undrr.org/publication/flash-flood-warnings-recent-achievements-france-national-vigicrues-flash-system.

Maxime Jay-Allemand. *Estimation variationnelle des paramètres d'un modèle hydrologique distribué*. These de doctorat, Aix-Marseille, December 2020. URL https://www.theses.fr/2020AIXM0400.

Maxime Jay-Allemand, Pierre Javelle, Igor Gejadze, Patrick Arnaud, Pierre-Olivier Malaterre, Jean-Alain Fine, and Didier Organde. On the potential of variational calibration for a fully distributed hydrological model : application on a Mediterranean catchment. *Hydrology and Earth System Sciences*, 24(11) :5519–5538, November 2020. ISSN 1027-5606. doi : 10.5194/hess-24-5519-2020. URL https://hess.copernicus.org/articles/24/5519/2020/. Publisher : Copernicus GmbH.

Valérie Jean, Marie-Amélie Boucher, Anissa Frini, and Dominic Roussel. Uncertainty in three dimensions : the challenges of communicating probabilistic flood forecast maps. *Hydrology and Earth System Sciences*, 27(18) :3351–3373, September 2023. ISSN 1027-5606. doi : 10.5194/hess-27-3351-2023. URL https://hess.copernicus.org/articles/27/3351/2023/. Publisher : Copernicus GmbH.

Danielle Johnson, Sarah Harrison, Céline Cattoën, Jordan Luttrell, and Paula Blackett. 'That still haunts me a little bit' : Decision-makers and information providers' experiences of recurring flood events. *International Journal of Disaster Risk Reduction*, 118 :105216, February 2025. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2025.105216. URL https://www.sciencedirect.com/science/article/pii/S2212420925000408.

B. Jongman, H. Kreibich, H. Apel, J. I. Barredo, P. D. Bates, L. Feyen, A. Gericke, J. Neal, J. C. J. H. Aerts, and P. J. Ward. Comparative flood damage model assessment : towards a European approach. *Natural Hazards and Earth System Sciences*, 12(12) :3733–3752, December 2012. ISSN 1561-8633. doi : 10.5194/nhess-12-3733-2012. URL https://nhess.copernicus.org/articles/12/3733/2012/. Publisher : Copernicus GmbH.

Syed Kabir, Sandhya Patidar, Xilin Xia, Qiuhua Liang, Jeffrey Neal, and Gareth Pender. A deep convolutional neural network model for rapid prediction of fluvial flood inundation. *Journal of Hydrology*, 590 :125481, November 2020. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2020.125481. URL https://www.sciencedirect.com/science/article/pii/S0022169420309410.

Jérémy Kalsron, Laurence Jolivet, and Olivier Payrastre. *Construction de MNE adaptés à la simulation d'inondations*. Thèse de master, Université de Lyon, 2019. URL https://documentation.ensg.eu/index.php?lvl=notice_display&id=99151.

Rainer Kaltenberger, Andreas Schaffhauser, and Michael Staudinger. "What the weather will do" – results of a survey on impact-oriented and impact-based warnings in European NMHSs. In *Advances in Science and Research*, volume 17, pages 29–38. Copernicus GmbH, May 2020. doi : 10.5194/asr-17-29-2020. URL https://asr.copernicus.org/articles/17/29/2020/. ISSN : 1992-0628.

Martijn Kuller, Kevin Schoenholzer, and Judit Lienert. Creating effective flood warnings : A framework from a critical review. *Journal of Hydrology*, 602 :126708, November 2021. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2021.126708. URL https://www.sciencedirect.com/science/article/pii/S0022169421007587.

O. Laganier, P. A. Ayral, D. Salze, and S. Sauvagnargues. A coupling of hydrologic and hydraulic models appropriate for the fast floods of the Gardon River basin (France). *Natural Hazards and Earth System Sciences*, 14(11) :2899–2920, November 2014. ISSN 1561-8633. doi : 10.5194/nhess-14-2899-2014. URL https://nhess.copernicus.org/articles/14/2899/2014/nhess-14-2899-2014.html. Publisher : Copernicus GmbH.

Martina Lagasio, Francesco Silvestro, Lorenzo Campo, and Antonio Parodi. Predictive Capability of a High-Resolution Hydrometeorological Forecasting Framework Coupling WRF Cycling 3DVAR and Continuum. *Journal of Hydrometeorology*, 20(7) :1307–1337, July 2019. ISSN 1525-7541, 1525-755X. doi : 10.1175/JHM-D-18-0219.1. URL https://journals.ametsoc.org/view/journals/hydr/20/7/jhm-d-18-0219_1.xml. Publisher : American Meteorological Society Section : Journal of Hydrometeorology.

Dimitri Lague and Baptiste Feldmann. Chapter 2 - Topo-bathymetric airborne LiDAR for fluvial-geomorphology analysis. In Paolo Tarolli and Simon M. Mudd, editors, *Developments in Earth Surface Processes*, volume 23 of *Remote Sensing of Geomorphology*, pages 25–54. Elsevier, January 2020. doi : 10.1016/B978-0-444-64177-9.00002-3. URL https://www.sciencedirect.com/science/article/pii/B9780444641779000023.

Niraj Lamichhane and Suresh Sharma. Effect of input data in hydraulic modeling for flood warning systems. *Hydrological Sciences Journal*, 63(6) :938–956, April 2018. ISSN 0262-6667. doi : 10.1080/02626667.2018.1464166. URL https://doi.org/10.1080/02626667.2018.1464166. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/02626667.2018.1464166.

Laurentin. Hourly rainfall analysis merging radar and rain gauge data. pages 2–8, Grenoble, France, 2008.

Guillaume Le Bihan. *Modèles hydrologiques régionaux pour la prévision distribuée des crues rapides : vers une estimation des impacts et des dommages potentiels*. These de doctorat, Ecole centrale de Nantes, October 2016. URL https://www.theses.fr/2016ECDN0011.

Guillaume Le Bihan, O. Payrastre, Eric Gaume, David Moncoulon, and Frédéric Pons. The challenge of forecasting impacts of flash floods : Test of a simplified hydraulic approach and validation based on insurance claim data. *Hydrology and Earth System Sciences*, 21 :5911–5928, November 2017. doi : 10.5194/hess-21-5911-2017.

J. Le Coz, B. Renard, L. Bonnifait, F. Branger, and R. Le Boursicaud. Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves : A Bayesian approach. *Journal of Hydrology*, 509 :573–587, February 2014. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2013.11.016. URL https://www.sciencedirect.com/science/article/pii/S0022169413008329.

Laurent Lebouc and Olivier Payrastre. Reconstitution des débits de pointe des crues des 23-24 novembre et 1er décembre 2019 dans les départements du Var et les Alpes-Maritimes. Research Report, IFSTTAR - Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux, January 2020. URL https://hal.archives-ouvertes.fr/hal-02933695.

Laurent Lebouc, Olivier Payrastre, and François Bourgin. Reconstitution des débits de pointe des crues du 15 octobre 2018 dans le bassin de l'Aude. Convention DGPR-Ifsttar 2018 n°2201132931 du 22 Mai 2018 - Action 7 appui au SCHAPI. Research Report, IFSTTAR - Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux, 2019a. URL https://hal.science/hal-02110612.

Laurent Lebouc, Olivier Payrastre, and François Bourgin. Reconstitution des débits de pointe des crues du 9 août 2018 dans les bassins de l'Ardèche et la Cèze. Convention DGPR-Ifsttar 2018 n°2201132931 du 22 Mai 2018 - Action 7 appui au SCHAPI. Research Report, IFSTTAR - Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux, 2019b. URL https://hal.science/hal-02110597.

Bernhard Lehner. Derivation of watershed boundaries for GRDC gauging stations based on the HydroSHEDS drainage network - Technical Report prepared for the GRDC, Global Runoff Data Centre. Technical report, Koblenz, Germany, 2012.

Frédéric Leone, Nancy Meschinet de Richemond, and Vinet Freddy. *Aléas naturels et gestion des risques*. Presses Universitaires de France, July 2021. URL https://hal.science/hal-03279213.

Jing Li and David W. S. Wong. Effects of DEM sources on hydrologic applications. *Computers, Environment and Urban Systems*, 34(3) :251–261, May 2010. ISSN 0198-9715. doi : 10.1016/j.compenvurbsys.2009.11.002. URL https://www.sciencedirect.com/science/article/pii/S0198971509000933.

Yi Li, Bisong Hu, Dong Zhang, Jianhua Gong, Yiquan Song, and Jun Sun. Flood evacuation simulations using cellular automata and multiagent systems -a human-environment relationship perspective. *International Journal of Geographical Information Science*, 33(11) :2241–2258, November 2019. ISSN 1365-8816. doi : 10.1080/13658816.2019.1622015. URL https://doi.org/10.1080/13658816.2019.1622015. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/13658816.2019.1622015.

K. Liechti, L. Panziera, U. Germann, and M. Zappa. The potential of radar-based ensemble forecasts for flash-flood early warning in the southern Swiss Alps. *Hydrology and Earth System Sciences*, 17(10) :3853–3869, October 2013. ISSN 1027-5606. doi : 10.5194/hess-17-3853-2013. URL https://hess.copernicus.org/articles/17/3853/2013/. Publisher : Copernicus GmbH.

M. C. Llasat, M. Llasat-Botija, M. A. Prat, F. Porcú, C. Price, A. Mugnai, K. Lagouvardos, V. Kotroni, D. Katsanos, S. Michaelides, Y. Yair, K. Savvidou, and K. Nicolaides. High-impact floods and flash floods in Mediterranean countries : the FLASH preliminary database. *Advances in Geosciences*, 23 :47–55, March 2010. ISSN 1680-7340. doi : 10.5194/adgeo-23-47-2010. URL https://adgeo.copernicus.org/articles/23/47/2010/adgeo-23-47-2010.html. Conference Name : 10th EGU Plinius Conference on Mediterranean Storms (2008) - 10th Plinius Conference on Mediterranean Storms, Nicosia, Cyprus, 22&ndash ;24 September 2008 Publisher : Copernicus GmbH.

Alexane Lovat. *Prévision à très courte échéance des crues rapides méditerranéennes*. These de doctorat, Toulouse, INPT, November 2019. URL https://www.theses.fr/2019INPT0105.

Alexane Lovat, Béatrice Vincendon, and Véronique Ducrocq. Hydrometeorological evaluation of two nowcasting systems for Mediterranean heavy precipitation events with operational considerations. *Hydrology and Earth System Sciences*, 26(10) : 2697–2714, May 2022. ISSN 1027-5606. doi : 10.5194/hess-26-2697-2022. URL https://hess.copernicus.org/articles/26/2697/2022/. Publisher : Copernicus GmbH.

Chris Lowrie, Andrew Kruczkiewicz, Shanna N. McClain, Miriam Nielsen, and Simon J. Mason. Evaluating the usefulness of VGI from Waze for the reporting of flash floods. *Scientific Reports*, 12(1) :5268, March 2022. ISSN 2045-2322. doi : 10.1038/s41598-022-08751-7. URL https://www.nature.com/articles/s41598-022-08751-7. Number : 1 Publisher : Nature Publishing Group.

Darren Lumbroso. How can policy makers in sub-Saharan Africa make early warning systems more effective ? The case of Uganda. *International Journal of Disaster Risk Reduction*, 27 :530–540, 2018. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2017.11.017. URL https://www.sciencedirect.com/science/article/pii/S2212420917303412.

Darren Lumbroso and Eric Gaume. Reducing the uncertainty in indirect estimates of extreme flash flood discharges. *Journal of Hydrology*, 414-415 :16–30, January 2012. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2011.08.048. URL https://www.sciencedirect.com/science/article/pii/S0022169411006044.

Hao Luo, Dimitrios K. Fytanidis, Arthur R. Schmidt, and Marcelo H. García. Comparative 1D and 3D numerical investigation of open-channel junction flows and energy losses. *Advances in Water Resources*, 117 :120–139, July 2018. ISSN 0309-1708. doi : 10.1016/j.advwatres.2018.05.012. URL https://www.sciencedirect.com/science/article/pii/S0309170817306954.

Pradeep V. Mandapaka, Urs Germann, Luca Panziera, and Alessandro Hering. Can Lagrangian Extrapolation of Radar Fields Be Used for Precipitation Nowcasting over Complex Alpine Orography ? *Weather and Forecasting*, 27(1) :28–49, February 2012. ISSN 1520-0434, 0882-8156. doi : 10.1175/WAF-D-11-00050.1. URL https://journals.ametsoc.org/view/journals/wefo/27/1/waf-d-11-00050_1.xml. Publisher : American Meteorological Society Section : Weather and Forecasting.

L. Marchi, M. Borga, E. Preciso, and E. Gaume. Characterisation of selected extreme flash floods in Europe and implications for flood risk management. *Journal of Hydrology*, 394(1) :118–133, November 2010. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2010.07.017. URL https://www.sciencedirect.com/science/article/pii/S0022169410004427.

Ian Mason. A Model for Assessment of Weather Forecasts. *Australian Meteorological Magazine*, 30(4) :291–303, 1982.

Valérie Masson-Delmotte, Panmao Zhai, Anna Pirani, Sarah L. Connors, Clotilde Péan, Sophie Berger, Nada Caud, Yang Chen, Leah Goldfarb, Melissa I. Gomis, Mengtian Huang, Katherine Leitzell, Elisabeth Lonnoy, J. B. Robin Matthews, Thomas K. Maycock, Tim Waterfield, Özge Yelekçi, Rong Yu, and Botao Zhou, editors. *Climate Change 2021 : The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021. doi : 10.1017/9781009157896.

Nikolaos Mastrantonas, Pedro Herrera-Lormendez, Linus Magnusson, Florian Pappenberger, and Jörg Matschullat. Extreme precipitation events in the Mediterranean : Spatiotemporal characteristics and connection to large-scale atmospheric flow patterns. *International Journal of Climatology*, 41(4) :2710–2728, 2021. ISSN 1097-0088. doi : 10.1002/joc.6985. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.6985. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.6985.

Ben Maybee, Cathryn E. Birch, Steven J. Böing, Thomas Willis, Linda Speight, Aurore N. Porson, Charlie Pilling, Kay L. Shelton, and Mark A. Trigg. FOREWARNS : development and multifaceted verification of enhanced regional-scale surface water flood forecasts. *Natural Hazards and Earth System Sciences*, 24(4) :1415–1436, April 2024. ISSN 1561-8633.

doi : 10.5194/nhess-24-1415-2024. URL https://nhess.copernicus.org/articles/24/1415/2024/. Publisher : Copernicus GmbH.

Maurizio Mazzoleni, Francesco Dottori, Hannah L. Cloke, and Giuliano Di Baldassarre. Deciphering human influence on annual maximum flood extent at the global level. *Communications Earth & Environment*, 3(1) :1–10, November 2022. ISSN 2662-4435. doi : 10.1038/s43247-022-00598-0. URL https://www.nature.com/articles/s43247-022-00598-0. Publisher : Nature Publishing Group.

Erika Meléndez-Landaverde, Daniel Sempere-Torres, and Shinju Park. Towards Impact-Based Communication During Climate Emergencies. In *Responding to Extreme Weather Events*, pages 119–140. John Wiley & Sons, Ltd, 2024. ISBN 978-1-119-74137-4. doi : 10.1002/9781119741374.ch6. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119741374.ch6. Section : 6 _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119741374.ch6.

Erika R. Meléndez-Landaverde and Daniel Sempere-Torres. Design and evaluation of a community and impact-based site-specific early warning system (SS-EWS) : The SS-EWS framework. *Journal of Flood Risk Management*, 18(1) :e12860, 2022. ISSN 1753-318X. doi : 10.1111/jfr3.12860. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jfr3.12860. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1111/jfr3.12860.

B. Merz, H. Kreibich, R. Schwarze, and A. Thieken. Review article "Assessment of economic flood damage". *Natural Hazards and Earth System Sciences*, 10(8) :1697–1724, August 2010. ISSN 1561-8633. doi : 10.5194/nhess-10-1697-2010. URL https://nhess.copernicus.org/articles/10/1697/2010/. Publisher : Copernicus GmbH.

Bruno Merz, Christian Kuhlicke, Michael Kunz, Massimiliano Pittore, Andrey Babeyko, David N. Bresch, Daniela I. V. Domeisen, Frauke Feser, Inga Koszalka, Heidi Kreibich, Florian Pantillon, Stefano Parolai, Joaquim G. Pinto, Heinz Jürgen Punge, Eleonora Rivalta, Kai Schröter, Karen Strehlow, Ralf Weisse, and Andreas Wurpts. Impact Forecasting to Support Emergency Management of Natural Hazards. *Reviews of Geophysics*, 58(4) :e2020RG000704, 2020. ISSN 1944-9208. doi : 10.1029/2020RG000704. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2020RG000704. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2020RG000704.

V. Meyer, N. Becker, V. Markantonis, R. Schwarze, J. C. J. M. van den Bergh, L. M. Bouwer, P. Bubeck, P. Ciavola, E. Genovese, C. Green, S. Hallegatte, H. Kreibich, Q. Lequeux, I. Logar, E. Papyrakis, C. Pfurtscheller, J. Poussin, V. Przyluski, A. H. Thieken, and C. Viavattene. Review article : Assessing the costs of natural hazards – state of the art and knowledge gaps. *Natural Hazards and Earth System Sciences*, 13(5) :1351–1373, May 2013. ISSN 1561-8633. doi : 10.5194/nhess-13-1351-2013. URL https://nhess.copernicus.org/articles/13/1351/2013/. Publisher : Copernicus GmbH.

Jean-Marc Moisselin, Philippe Cau, Céline Jauffret, I. Bouissières, and R. Tzanos. Seamless approach for precipitations within the 0-3 hours forecast-interval. April 2019. URL https://repositorio.aemet.es/handle/20.500.11765/10588. Accepted : 2019-05-22T10 :44 :08Z Publisher : Agencia Estatal de Meteorología.

D. Molinari, S. Menoni, G. T. Aronica, F. Ballio, N. Berni, C. Pandolfo, M. Stelluti, and G. Minucci. Ex post damage assessment : an Italian experience. *Natural Hazards and Earth System Sciences*, 14(4) :901–916, April 2014. ISSN 1561-8633. doi : 10.5194/nhess-14-901-2014. URL https://nhess.copernicus.org/articles/14/901/2014/. Publisher : Copernicus GmbH.

Daniela Molinari, Karin M. De Bruijn, Jesica T. Castillo-Rodríguez, Giuseppe T. Aronica, and Laurens M. Bouwer. Validation of flood risk models : Current practice and possible improvements. *International Journal of Disaster Risk Reduction*, 33 : 441–448, February 2019. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2018.10.022. URL https://www.sciencedirect.com/science/article/pii/S221242091830596X.

D. Moncoulon, D. Labat, J. Ardon, E. Leblois, T. Onfroy, C. Poulard, S. Aji, A. Rémy, and A. Quantin. Analysis of the French insurance market exposure to floods : a stochastic model combining river overflow and surface runoff. *Natural Hazards and Earth System Sciences*, 14(9) :2469–2485, September 2014. ISSN 1561-8633. doi : 10.5194/nhess-14-2469-2014. URL https://nhess.copernicus.org/articles/14/2469/2014/. Publisher : Copernicus GmbH.

David Moncoulon, Roxane Marchal, and Thomas Onfroy. Les risques liés au ruissellement, l'apport des données de sinistralité. *La Houille Blanche*, 106(6) :56–65, December 2020. ISSN 0018-6368. doi : 10.1051/lhb/2020058. URL https://doi.org/10.1051/lhb/2020058. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1051/lhb/2020058.

Sandrine Morel. Thousands protest in Valencia as distrust in Spanish government grows over flood response. November 2024. URL https://www.lemonde.fr/en/international/article/2024/11/10/thousands-protest-in-valencia-as-distrust-in-spanish-government-grows-over-flood-response_6732317_4.html.

Markus Mosimann, Martina Kauzlaric, Simon Schick, Olivia Martius, and Andreas Paul Zischg. Evaluation of surrogate flood models for the use in impact-based flood warning systems at national scale. *Environmental Modelling & Software*, 173 : 105936, February 2024. ISSN 1364-8152. doi : 10.1016/j.envsoft.2023.105936. URL https://www.sciencedirect.com/science/article/pii/S1364815223003225.

MRN. La Lettre d'information de la Mission risques naturels. *Lettre MRN*, (19), 2012.

MunichRe. Record thunderstorm losses and deadly earthquakes : the natural disasters of 2023, January 2024. URL https://www.munichre.com/en/company/media-relations/media-information-and-corporate-news/media-information/2024/natural-disaster-figures-2023.html.

Simon Munier and Bertrand Decharme. River network and hydro-geomorphological parameters at 1/12 resolution for global hydrological and climate studies. *Earth System Science Data*, 14(5) :2239–2258, May 2022. ISSN 1866-3508. doi : 10.5194/essd-14-2239-2022. URL https://essd.copernicus.org/articles/14/2239/2022/. Publisher : Copernicus GmbH.

Husain Najafi, Pallav Kumar Shrestha, Oldrich Rakovec, Heiko Apel, Sergiy Vorogushyn, Rohini Kumar, Stephan Thober, Bruno Merz, and Luis Samaniego. High-resolution impact-based early warning system for riverine flooding. *Nature Communications*, 15(1) :3726, May 2024. ISSN 2041-1723. doi : 10.1038/s41467-024-48065-y. URL https://www.nature.com/articles/s41467-024-48065-y. Publisher : Nature Publishing Group.

J. P. Naulin, O. Payrastre, and E. Gaume. Spatially distributed flood forecasting in flash flood prone areas : Application to road network supervision in Southern France. *Journal of Hydrology*, 486 :88–99, April 2013. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2013.01.044. URL https://www.sciencedirect.com/science/article/pii/S0022169413001017.

Sella Nevo, Efrat Morin, Adi Gerzi Rosenthal, Asher Metzger, Chen Barshai, Dana Weitzner, Dafi Voloshin, Frederik Kratzert, Gal Elidan, Gideon Dror, Gregory Begelman, Grey Nearing, Guy Shalev, Hila Noga, Ira Shavitt, Liora Yuklea, Moriah Royz, Niv Giladi, Nofar Peled Levi, Ofir Reich, Oren Gilon, Ronnie Maor, Shahar Timnat, Tal Shechter, Vladimir Anisimov, Yotam Gigi, Yuval Levin, Zach Moshe, Zvika Ben-Haim, Avinatan Hassidim, and Yossi Matias. Flood forecasting with machine learning models in an operational framework. *Hydrology and Earth System Sciences*, 26(15) :4013–4032, August 2022. ISSN 1027-5606. doi : 10.5194/hess-26-4013-2022. URL https://hess.copernicus.org/articles/26/4013/2022/. Publisher : Copernicus GmbH.

Pierre Nicolle, Olivier Payrastre, and Nabil Hocini. Production d'un catalogue d'emprises inondées mono-fréquence sur les petits cours d'eau de l'arc méditerranéen (rapport d'étape). Action 8.1. Research Report, IFSTTAR - Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux, January 2021. URL https://hal.archives-ouvertes.fr/hal-03327095.

Pierre Nicolle, Olivier Payrastre, and Nabil Hocini. Méthode de prétraitement des MNT pour l'amélioration des calculs automatisés de zones inondables. Research Report, Université Gustave Eiffel ; DGPR (Direction Générale de la Prévention des Risques - Ministère de la Transition Ecologique et Solidaire), May 2023. URL https://hal.science/hal-04159337.

Pierre Nicolle, Olivier Payrastre, Juliette Godet, and Thomas Dias. Construction d'échelles de gravité des inondations sur le réseau hydrographique détaillé de l'arc méditerranéen. Technical Report Convention DGPR-Université Gustave Eiffel 2023 - Action 8.1, Université Gustave Eiffel, June 2024. URL https://hal.science/hal-04980845.

NIWA. The largest flood flow ever measured, 2021. URL https://niwa.co.nz/news/largest-flood-flow-ever-measured.

A. D. Nobre, L. A. Cuartas, M. Hodnett, C. D. Rennó, G. Rodrigues, A. Silveira, M. Waterloo, and S. Saleska. Height Above the Nearest Drainage – a hydrologically relevant new terrain model. *Journal of Hydrology*, 404(1) :13–29, June 2011. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2011.03.051. URL https://www.sciencedirect.com/science/article/pii/S0022169411002599.

Diego Novoa, Julian David Reyes-Silva, Björn Helm, and Peter Krebs. Development and Demonstration of an Interactive Tool in an Agent-Based Model for Assessing Pluvial Urban Flooding. *Water*, 15(4) :696, January 2023. ISSN 2073-4441. doi : 10.3390/w15040696. URL https://www.mdpi.com/2073-4441/15/4/696. Publisher : Multidisciplinary Digital Publishing Institute.

O. Nuissier, C. Marsigli, B. Vincendon, A. Hally, F. Bouttier, A. Montani, and T. Paccagnella. Evaluation of two convection-permitting ensemble systems in the HyMeX Special Observation Period (SOP1) framework. *Quarterly Journal of the Royal Meteorological Society*, 142(S1) :404–418, 2016. ISSN 1477-870X. doi : 10.1002/qj.2859. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.2859. _eprint : https ://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2859.

Jean Odry and Patrick Arnaud. Comparison of Flood Frequency Analysis Methods for Ungauged Catchments in France. *Geosciences*, 7(3) :88, September 2017. ISSN 2076-3263. doi : 10.3390/geosciences7030088. URL https://www.mdpi.com/2076-3263/7/3/88. Number : 3 Publisher : Multidisciplinary Digital Publishing Institute.

Francisco Olivera, Mary S. Lear, James S. Famiglietti, and Kwabena Asante. Extracting low-resolution river networks from high-resolution digital elevation models. *Water Resources Research*, 38(11) :13–1–13–8, 2002. ISSN 1944-7973. doi : 10.1029/2001WR000726. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2001WR000726. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2001WR000726.

Didier Organde, Pierre Javelle, Constantin Ardilouze, and Rémi Lamblin. Base nationale des bassins versants du SCHAPI. Technical report, 2013.

Robert Osinski and François Bouttier. Short-range probabilistic forecasting of convective risks for aviation based on a lagged-average-forecast ensemble approach. *Meteorological Applications*, 25(1) :105–118, 2018. ISSN 1469-8080. doi : 10.1002/met.1674. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/met.1674. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/met.1674.

K. Papagiannaki, K. Lagouvardos, V. Kotroni, and A. Bezes. Flash flood occurrence and relation to the rainfall hazard in a highly urbanized area. *Natural Hazards and Earth System Sciences*, 15(8) :1859–1871, August 2015. ISSN 1561-8633. doi : 10.5194/nhess-15-1859-2015. URL https://nhess.copernicus.org/articles/15/1859/2015/nhess-15-1859-2015.html. Publisher : Copernicus GmbH.

Shinju Park, Marc Berenguer, and Daniel Sempere-Torres. Long-term analysis of gauge-adjusted radar rainfall accumulations at European scale. *Journal of Hydrology*, 573 :768–777, June 2019. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2019.03.093. URL https://www.sciencedirect.com/science/article/pii/S0022169419303221.

Célia Parra. *Vers des pistes d'améliorations du service Vigicrues Flash*. Mémoire de master, Institut national universitaire Jean-François Champollion, September 2019.

Mathilde Pascal, Karine Laaidi, Martine Ledrans, Elsa Baffert, Céline Caserio-Schönemann, Alain Le Tertre, Jacques Manach, Sylvia Medina, Jérémie Rudant, and Pascal Empereur-Bissonnet. France's heat health watch warning system. *International Journal of Biometeorology*, 50(3) :144–153, January 2006. ISSN 1432-1254. doi : 10.1007/s00484-005-0003-x. URL https://doi.org/10.1007/s00484-005-0003-x.

Ryan Paulik, Nick Horspool, Richard Woods, Nick Griffiths, Tim Beale, Christina Magill, Alec Wild, Benjamin Popovich, Glenn Walbran, and Russel Garlick. RiskScape : a flexible multi-hazard risk modelling engine. *Natural Hazards*, 119(2) : 1073–1090, November 2023. ISSN 1573-0840. doi : 10.1007/s11069-022-05593-4. URL https://doi.org/10.1007/s11069-022-05593-4.

Ryan Paulik, Alec Wild, Conrad Zorn, Liam Wotherspoon, and Shaun Williams. Evaluation of residential building damage for the July 2021 flood in Westport, New Zealand. *Geoscience Letters*, 11(1) :15, March 2024. ISSN 2196-4092. doi : 10.1186/s40562-024-00323-z. URL https://doi.org/10.1186/s40562-024-00323-z.

Olivier Payrastre and Laurent Lebouc. Evaluation de la sensibilité des réseaux routiers aux inondations : application de la méthode RoadIno dans le département du Gard. Technical report, Université Gustave Eiffel, 2019.

Olivier Payrastre, Eric Gaume, Pierre Javelle, Bruno Janet, Patrick Fourmigué, Philippe Lefort, André Martin, Brice Boudevillain, Pascal Brunet, Guy Delrieu, Lorenzo Marchi, Yoann Aubert, Elisabeth Dautrey, Laurence Durand, Michel Lang, Laurent Boissier, Johnny Douvinet, and Claude Martin. Analyse hydrologique de la crue-éclair catastrophique du 15 juin 2010 dans

la région de Draguignan. *La Houille Blanche*, 105(3-4) :140–148, 2019. ISSN 0018-6368. doi : 10.1051/lhb/2019057. URL https://doi.org/10.1051/lhb/2019057. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1051/lhb/2019057.

Olivier Payrastre, Pierre Nicolle, Laurent Bonnifait, Pierre Brigode, Paul Astagneau, Amaury Baise, Arnaud Belleville, Nedjima Bouamara, François Bourgin, Pascal Breil, Pascal Brunet, Arnaud Cerbelaud, Fanny Courapied, Lise Devreux, Raphaëlle Dreyfus, Eric Gaume, Stan Nomis, Julie Poggio, Frédéric Pons, Yassine Rabab, and Damien Sevrez. Tempête Alex du 2 octobre 2020 dans les Alpes-Maritimes : une contribution de la communauté scientifique à l'estimation des débits de pointe des crues. *LHB Hydroscience Journal*, page 2082891, June 2022. doi : 10.1080/27678490.2022.2082891. URL https://hal.inrae.fr/hal-03710678.

Adriano Rolim Paz, Walter Collischonn, and André Luiz Lopes da Silveira. Improvements in large-scale drainage networks derived from digital elevation models. *Water Resources Research*, 42(8), 2006. ISSN 1944-7973. doi : 10.1029/2005WR004544. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2005WR004544. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2005WR004544.

Daniela Peredo, Maria-Helena Ramos, Hugo Marchal, and François Bouttier. Challenges of event-based evaluation of flash floods : example of the October 2018 flood event in the Aude catchment in France., 2021. URL https://events.ecmwf.int/event/222/contributions/2255/attachments/1291/2358/Hydrological-WS-Peredo.pdf.

Charles Perrin, Claude Michel, and Vazken Andréassian. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279(1) :275–289, August 2003. ISSN 0022-1694. doi : 10.1016/S0022-1694(03)00225-7. URL https://www.sciencedirect.com/science/article/pii/S0022169403002257.

Charlie Pilling, Adrian Wynn, Neil Armstrong, Russel Turner, Julia Perez, Chris Lattimore, and Catherine Birch. Rapid Flood Guidance for flash floods - an operational service for England and Wales in summer 2025. In *EGU25-14408*. Copernicus Meetings, March 2025. doi : 10.5194/egusphere-egu25-14408. URL https://meetingorganizer.copernicus.org/EGU25/EGU25-14408.html. Conference Name : EGU25.

Olivier Piotte, Thibaut Montmerle, Catherine Fouchier, Anne Belleudy, Lea Garandeau, Bruno Janet, Celine Jauffret, Julie Demargne, and Didier Organde. Les évolutions du service d'avertissement sur les pluies intenses et les crues soudaines en France. *La Houille Blanche*, (6) :75–84, December 2020. ISSN 0018-6368, 1958-5551. doi : 10.1051/lhb/2020055. URL https://www.shf-lhb.org/articles/lhb/abs/2020/06/lhb200095/lhb200095.html. Number : 6 Publisher : EDP Sciences.

Guillaume Pla, Julien Crippa, Abdelatif Djerboua, Octavian Dobricean, Frédéric Dongar, Allan Eugene, and Mireille Raymond. ESPADA : un outil pour la gestion en temps réel des crues éclairs urbaines en pleine modernisation. *La Houille Blanche*, (3-4) :57–66, October 2019. ISSN 0018-6368, 1958-5551. doi : 10.1051/lhb/2019027. URL https://www.shf-lhb.org/articles/lhb/abs/2019/03/lhb180028/lhb180028.html. Number : 3-4 Publisher : EDP Sciences.

Jan Polcher, Anthony Schrapffer, Eliott Dupont, Lucia Rinchiuso, Xudong Zhou, Olivier Boucher, Emmanuel Mouche, Catherine Ottlé, and Jérôme Servonnat. Hydrological modelling on atmospheric grids : using graphs of sub-grid elements to transport energy and water. *Geoscientific Model Development*, 16(9) :2583–2606, May 2023. ISSN 1991-959X. doi : 10.5194/gmd-16-2583-2023. URL https://gmd.copernicus.org/articles/16/2583/2023/. Publisher : Copernicus GmbH.

Maria Laura Poletti, Francesco Silvestro, Silvio Davolio, Flavio Pignone, and Nicola Rebora. Using nowcasting technique and data assimilation in a meteorological model to improve very short range hydrological forecasts. *Hydrology and Earth System Sciences*, 23(9) :3823–3841, September 2019. ISSN 1027-5606. doi : 10.5194/hess-23-3823-2019. URL https://hess.copernicus.org/articles/23/3823/2019/. Publisher : Copernicus GmbH.

Nils Poncet, Philippe Lucas-Picher, Yves Tramblay, Guillaume Thirel, Humberto Vergara, Jonathan Gourley, and Antoinette Alias. Does a convection-permitting regional climate model bring new perspectives on the projection of Mediterranean floods ? *Natural Hazards and Earth System Sciences*, 24(4) :1163–1183, April 2024. ISSN 1561-8633. doi : 10.5194/nhess-24-1163-2024. URL https://nhess.copernicus.org/articles/24/1163/2024/. Publisher : Copernicus GmbH.

F. Pons, C. Laroche, P. Fourmigue, and M. Alquier. Cartographie des surfaces inondables extrêmes pour la directive inondation : cas de la Nartuby. *La Houille Blanche*, (2) :34–41, April 2014. ISSN 0018-6368, 1958-5551. doi : 10.1051/lhb/2014014. URL https://www.shf-lhb.org/articles/lhb/abs/2014/02/lhb2014014/lhb2014014.html. Number : 2 Publisher : Société Hydrotechnique de France.

Sally Potter, Sara Harrison, and Peter Kreft. The Benefits and Challenges of Implementing Impact-Based Severe Weather Warning Systems : Perspectives of Weather, Flood, and Emergency Management Personnel. *Weather, Climate, and Society*, 13(2) :303–314, April 2021. ISSN 1948-8327, 1948-8335. doi : 10.1175/WCAS-D-20-0110.1. URL https://journals.ametsoc.org/view/journals/wcas/13/2/WCAS-D-20-0110.1.xml. Publisher : American Meteorological Society Section : Weather, Climate, and Society.

Sally H. Potter, Peter V. Kreft, Petar Milojev, Chris Noble, Burrell Montz, Amandine Dhellemmes, Richard J. Woods, and Sarah Gauden-Ing. The influence of impact-based severe weather warnings on risk perceptions and intended protective actions. *International Journal of Disaster Risk Reduction*, 30 :34–43, September 2018. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2018.03.031. URL https://www.sciencedirect.com/science/article/pii/S2212420918304011.

Sally H. Potter, Thomas Kox, Brian Mills, Andrea Taylor, Joanne Robbins, Carolina Cerrudo, Faye Wyatt, Sara Harrison, Brian Golding, Will Lang, Adam J. L. Harris, Rainer Kaltenberger, Stefan Kienberger, Harold Brooks, and Andrew Tupper. Research gaps and challenges for impact-based forecasts and warnings : Results of international workshops for High Impact Weather in 2022. *International Journal of Disaster Risk Reduction*, 118 :105234, February 2025. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2025.105234. URL https://www.sciencedirect.com/science/article/pii/S2212420925000585.

Narcisa G. Pricope and Md Salman Bashit. Emerging trends in topobathymetric LiDAR technology and mapping. *International Journal of Remote Sensing*, 44(24) :7706–7731, December 2023. ISSN 0143-1161. doi : 10.1080/01431161.2023.2287564. URL https://doi.org/10.1080/01431161.2023.2287564. Publisher : Taylor & Francis _eprint : https ://-doi.org/10.1080/01431161.2023.2287564.

D. Raynaud, J. Thielen, P. Salamon, P. Burek, S. Anquetin, and L. Alfieri. A dynamic runoff co-efficient to improve flash flood early warning in Europe : evaluation on the 2013 central European floods in Germany. *Meteorological Applications*, 22(3) : 410–418, 2015. ISSN 1469-8080. doi : 10.1002/met.1469. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/met.1469. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/met.1469.

Seann Reed, John Schaake, and Ziya Zhang. A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations. *Journal of Hydrology*, 337(3) :402–420, April 2007. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2007.02.015. URL https://www.sciencedirect.com/science/article/pii/S0022169407000923.

Seann M. Reed. Deriving flow directions for coarse-resolution (1–4 km) gridded hydrologic modeling. *Water Resources Research*, 39(9), 2003. ISSN 1944-7973. doi : 10.1029/2003WR001989. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2003WR001989. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2003WR001989.

Hamid Rezatofighi, Nathan Tsoi, JunYoung Gwak, Amir Sadeghian, Ian Reid, and Silvio Savarese. Generalized Intersection Over Union : A Metric and a Loss for Bounding Box Regression. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 658–666, June 2019. doi : 10.1109/CVPR.2019.00075. ISSN : 2575-7075.

Aurélien Ribes, Soulivanh Thao, Robert Vautard, Brigitte Dubuisson, Samuel Somot, Jeanne Colin, Serge Planton, and Jean-Michel Soubeyroux. Observed increase in extreme daily rainfall in the French Mediterranean. *Climate Dynamics*, 52(1) : 1095–1114, January 2019. ISSN 1432-0894. doi : 10.1007/s00382-018-4179-2. URL https://doi.org/10.1007/s00382-018-4179-2.

David S. Richardson, Hannah L. Cloke, and Florian Pappenberger. Evaluation of the Consistency of ECMWF Ensemble Forecasts. *Geophysical Research Letters*, 47(11) :e2020GL087934, 2020. ISSN 1944-8007. doi : 10.1029/2020GL087934. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL087934. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL087934.

Josias Ritter, Marc Berenguer, Carles Corral, Shinju Park, and Daniel Sempere-Torres. ReAFFIRM : Real-time Assessment of Flash Flood Impacts – a Regional high-resolution Method. *Environment International*, 136 :105375, March 2020. ISSN 0160-4120. doi : 10.1016/j.envint.2019.105375. URL https://www.sciencedirect.com/science/article/pii/S0160412019314485.

Josias Ritter, Marc Berenguer, Shinju Park, and Daniel Sempere-Torres. Real-time assessment of flash flood impacts at pan-European scale : The ReAFFINE method. *Journal of Hydrology*, 603 :127022, December 2021. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2021.127022. URL https://www.sciencedirect.com/science/article/pii/S0022169421010726.

Josias Ritter, Marc Berenguer, Francesco Dottori, Milan Kalas, and Daniel Sempere-Torres. Compound flood impact forecasting : integrating fluvial and flash flood impact assessments into a unified system. *Hydrology and Earth System Sciences*, 26(3) : 689–709, February 2022. ISSN 1027-5606. doi : 10.5194/hess-26-689-2022. URL https://hess.copernicus.org/articles/26/689/2022/. Publisher : Copernicus GmbH.

Joanne Robbins, Isabelle Ruin, Brian Golding, Rutger Dankers, John Nairn, and Sarah Millington. Connecting Hazard and Impact : A Partnership between Physical and Human Science. In Brian Golding, editor, *Towards the "Perfect" Weather Warning : Bridging Disciplinary Gaps through Partnership and Communication*, pages 115–147. Springer International Publishing, Cham, 2022. ISBN 978-3-030-98989-7. doi : 10.1007/978-3-030-98989-7_5. URL https://doi.org/10.1007/978-3-030-98989-7_5.

Nigel M. Roberts and Humphrey W. Lean. Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Monthly Weather Review*, 136(1) :78–97, January 2008. ISSN 1520-0493, 0027-0644. doi : 10.1175/2007MWR2123.1. URL https://journals.ametsoc.org/view/journals/mwre/136/1/2007mwr2123.1.xml. Publisher : American Meteorological Society Section : Monthly Weather Review.

H. Roux, D. Labat, P.-A. Garambois, M.-M. Maubourguet, J. Chorda, and D. Dartus. A physically-based parsimonious hydrological model for flash floods in Mediterranean catchments. *Natural Hazards and Earth System Sciences*, 11(9) : 2567–2582, September 2011. ISSN 1561-8633. doi : 10.5194/nhess-11-2567-2011. URL https://nhess.copernicus.org/articles/11/2567/2011/. Publisher : Copernicus GmbH.

Hélène Roux, Arnau Amengual, Romu Romero, Ernest Bladé, and Marcos Sanz-Ramos. Evaluation of two hydrometeorological ensemble strategies for flash-flood forecasting over a catchment of the eastern Pyrenees. *Natural Hazards and Earth System Sciences*, 20(2) :425–450, February 2020. ISSN 1561-8633. doi : 10.5194/nhess-20-425-2020. URL https://nhess.copernicus.org/articles/20/425/2020/. Publisher : Copernicus GmbH.

Isabelle Ruin, Jean-Dominique Creutin, Sandrine Anquetin, Eve Gruntfest, and Céline Lutoff. Human vulnerability to flash floods : Addressing physical exposure and behavioural questions. In *Flood Risk Management : Research and Practice*. CRC Press, 2008. ISBN 978-0-429-20712-9. Num Pages : 1.

Isabelle Ruin, Céline Lutoff, Brice Boudevillain, Jean-Dominique Creutin, S. Anquetin, M. Bertran Rojo, L. Boissier, L. Bonnifait, M. Borga, L. Colbeau-Justin, L. Creton-Cazanave, G. Delrieu, J. Douvinet, E. Gaume, E. Gruntfest, J.-P. Naulin, O. Payrastre, and O. Vannier. Social and Hydrological Responses to Extreme Precipitations : An Interdisciplinary Strategy for Postflood Investigation. *Weather, Climate, and Society*, 6(1) :135–153, January 2014. ISSN 1948-8327, 1948-8335. doi : 10.1175/WCAS-D-13-00009.1. URL https://journals.ametsoc.org/view/journals/wcas/6/1/wcas-d-13-00009_1.xml. Publisher : American Meteorological Society Section : Weather, Climate, and Society.

Clotilde Saint-Martin. *Prise en compte des vulnérabilités territoriales dans l'avertissement des crues rapides : vers une amélioration de la méthode AIGA*. These de doctorat, Montpellier 3, 2018. URL https://www.theses.fr/2018MON30061.

Clotilde Saint-Martin, Catherine Fouchier, Pierre Javelle, Johnny Douvinet, and Freddy Vinet. Assessing the exposure to floods to estimate the risk of flood-related damage in French Mediterranean basins. *E3S Web of Conferences*, 7 :04013, 2016. ISSN 2267-1242. doi : 10.1051/e3sconf/20160704013. URL https://www.e3s-conferences.org/articles/e3sconf/abs/2016/02/e3sconf_flood2016_04013/e3sconf_flood2016_04013.html. Publisher : EDP Sciences.

Clotilde Saint-Martin, Pierre Javelle, and Freddy Vinet. DamaGIS : a multisource geodatabase for collection of flood-related damage data. *Earth System Science Data*, 10(2) :1019–1029, June 2018. ISSN 1866-3508. doi : 10.5194/essd-10-1019-2018. URL https://essd.copernicus.org/articles/10/1019/2018/. Publisher : Copernicus GmbH.

Christopher C. Sampson, Andrew M. Smith, Paul D. Bates, Jeffrey C. Neal, Lorenzo Alfieri, and Jim E. Freer. A high-resolution global flood hazard model. *Water Resources Research*, 51(9) :7358–7381, 2015. ISSN 1944-7973. doi : 10.1002/2015WR016954. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR016954. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/2015WR016954.

Brett F. Sanders, Oliver E. J. Wing, and Paul D. Bates. Flooding is Not Like Filling a Bath. *Earth's Future*, 12(12) : e2024EF005164, 2024. ISSN 2328-4277. doi : 10.1029/2024EF005164. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2024EF005164. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2024EF005164.

Takahiro Sayama, Masafumi Yamada, Yoshito Sugawara, and Dai Yamazaki. Ensemble flash flood predictions using a high-resolution nationwide distributed rainfall-runoff model : case study of the heavy rain event of July 2018 and Typhoon Hagibis in 2019. *Progress in Earth and Planetary Science*, 7(1) :75, December 2020. ISSN 2197-4284. doi : 10.1186/s40645-020-00391-7. URL https://doi.org/10.1186/s40645-020-00391-7.

Joseph T. Schaefer. The Critical Success Index as an Indicator of Warning Skill. *Weather and Forecasting*, 5(4) :570–575, December 1990. ISSN 1520-0434, 0882-8156. doi : 10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/wefo/5/4/1520-0434_1990_005_0570_tcsiaa_2_0_co_2.xml. Publisher : American Meteorological Society Section : Weather and Forecasting.

Katrin Scheufele, Kirstin Kober, George C. Craig, and Christian Keil. Combining probabilistic precipitation forecasts from a nowcasting technique with a time-lagged ensemble. *Meteorological Applications*, 21(2) :230–240, 2014. ISSN 1469-8080. doi : 10.1002/met.1381. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/met.1381. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/met.1381.

Amanda J. Schroeder, Jonathan J. Gourley, Jill Hardy, Jen J. Henderson, Pradipta Parhi, Vahid Rahmani, Kimberly A. Reed, Russ S. Schumacher, Brianne K. Smith, and Matthew J. Taraldsen. The development of a flash flood severity index. *Journal of Hydrology*, 541 :523–532, October 2016. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2016.04.005. URL https://www.sciencedirect.com/science/article/pii/S002216941630186X.

SDES. Chiffres clés des risques naturels. Technical report, Commissariat général au développement durable, Service des données et études statistiques, Paris, February 2024.

Alan W. Seed, Clive E. Pierce, and Katie Norman. Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme. *Water Resources Research*, 49(10) :6624–6641, 2013. ISSN 1944-7973. doi : 10.1002/wrcr.20536. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20536. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/wrcr.20536.

E. Shamir, K. P. Georgakakos, C. Spencer, T. M. Modrick, M. J. Murphy, and R. Jubach. Evaluation of real-time flash flood forecasts for Haiti during the passage of Hurricane Tomas, November 4–6, 2010. *Natural Hazards*, 67(2) :459–482, 2013. ISSN 1573-0840. doi : 10.1007/s11069-013-0573-6. URL https://doi.org/10.1007/s11069-013-0573-6.

Ashish Sharma, Conrad Wasko, and Dennis P. Lettenmaier. If Precipitation Extremes Are Increasing, Why Aren't Floods ? *Water Resources Research*, 54(11) :8545–8551, 2018. ISSN 1944-7973. doi : 10.1029/2018WR023749. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2018WR023749. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2018WR023749.

F. Silvestro and N. Rebora. Operational verification of a framework for the probabilistic nowcasting of river discharge in small and medium size basins. *Natural Hazards and Earth System Sciences*, 12(3) :763–776, March 2012. ISSN 1561-8633. doi : 10.5194/nhess-12-763-2012. URL https://nhess.copernicus.org/articles/12/763/2012/. Publisher : Copernicus GmbH.

F. Silvestro, S. Gabellani, F. Delogu, R. Rudari, and G. Boni. Exploiting remote sensing land surface temperature in distributed hydrological modelling : the example of the Continuum model. *Hydrology and Earth System Sciences*, 17(1) :39–62, 2013. ISSN 1027-5606. doi : 10.5194/hess-17-39-2013. URL https://hess.copernicus.org/articles/17/39/2013/. Publisher : Copernicus GmbH.

F. Silvestro, L. Rossi, L. Campo, A. Parodi, E. Fiori, R. Rudari, and L. Ferraris. Impact-based flash-flood forecasting system : Sensitivity to high resolution numerical weather prediction systems and soil moisture. *Journal of Hydrology*, 572 :388–402, May 2019. ISSN 0022-1694. doi : 10.1016/j.jhydrol.2019.02.055. URL https://www.sciencedirect.com/science/article/pii/S0022169419302124.

Francesco Silvestro, Nicola Rebora, and Luca Ferraris. Quantitative Flood Forecasting on Small- and Medium-Sized Basins : A Probabilistic Approach for Operational Purposes. *Journal of Hydrometeorology*, 12(6) :1432–1446, December 2011. ISSN 1525-7541, 1525-755X. doi : 10.1175/JHM-D-10-05022.1. URL https://journals.ametsoc.org/view/journals/hydr/12/6/jhm-d-10-05022_1.xml. Publisher : American Meteorological Society Section : Journal of Hydrometeorology.

Thâmara Martins Ismael Sousa and Adriano Rolim Paz. How to evaluate the quality of coarse-resolution DEM-derived drainage networks. *Hydrological Processes*, 31(19) :3379–3395, 2017. ISSN 1099-1085. doi : 10.1002/hyp.11262. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.11262. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/hyp.11262.

Edwin H. Sutanudjaja, Rens van Beek, Niko Wanders, Yoshihide Wada, Joyce H. C. Bosmans, Niels Drost, Ruud J. van der Ent, Inge E. M. de Graaf, Jannis M. Hoch, Kor de Jong, Derek Karssenberg, Patricia López López, Stefanie Peßenteiner, Oliver Schmitz, Menno W. Straatsma, Ekkamol Vannametee, Dominik Wisser, and Marc F. P. Bierkens. PCR-GLOBWB 2 : a 5 ;arcmin global hydrological and water resources model. *Geoscientific Model Development*, 11(6) :2429–2453, June 2018. ISSN 1991-959X. doi : 10.5194/gmd-11-2429-2018. URL https://gmd.copernicus.org/articles/11/2429/2018/. Publisher : Copernicus GmbH.

Pierre Tabary, Clotilde Augros, Jean-Louis Champeaux, Jean-Luc Chèze, Dominique Faure, Daniel Idziorek, Richard Lorandel, Bernard Urban, and Valérie Vogt. Le réseau et les produits radars de Météo-France. *La Météorologie*, 8ème série :15–27, November 2013. doi : 10.4267/2042/52050.

David G. Tarboton. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33(2) :309–319, 1997. ISSN 1944-7973. doi : 10.1029/96WR03137. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/96WR03137. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/96WR03137.

Andrea Louise Taylor, Thomas Kox, and David Johnston. Communicating high impact weather : Improving warnings and decision making processes. *International Journal of Disaster Risk Reduction*, 30 :1–4, September 2018. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2018.04.002. URL https://www.sciencedirect.com/science/article/pii/S2212420918304242.

Galateia Terti, Isabelle Ruin, Sandrine Anquetin, and Jonathan Gourley. Dynamic vulnerability factors for impact-based flash flood prediction. *Natural Hazards : Journal of the International Society for the Prevention and Mitigation of Natural Hazards*, 79(3) :1481–1497, 2015. URL https://link.springer.com/article/10.1007/s11069-015-1910-8. Publisher : Springer & International Society for the Prevention and Mitigation of Natural Hazards.

Galateia Terti, Isabelle Ruin, Jonathan J. Gourley, Pierre Kirstetter, Zachary Flamig, Juliette Blanchet, Ami Arthur, and Sandrine Anquetin. Toward Probabilistic Prediction of Flash Flood Human Impacts. *Risk Analysis*, 39(1) :140–161, January 2019a. ISSN 0272-4332. doi : 10.1111/risa.12921. URL https://onlinelibrary.wiley.com/doi/10.1111/risa.12921. Place : Hoboken Publisher : Wiley WOS :000454958400010.

Galateia Terti, Isabelle Ruin, Milan Kalas, Ilona Láng, Arnau Cangròs i Alonso, Tommaso Sabbatini, and Valerio Lorini. ANYCaRE : a role-playing game to investigate crisis decision-making and communication challenges in weather-related hazards. *Natural Hazards and Earth System Sciences*, 19(3) :507–533, March 2019b. ISSN 1561-8633. doi : 10.5194/nhess-19-507-2019. URL https://nhess.copernicus.org/articles/19/507/2019/. Publisher : Copernicus GmbH.

Annegret H. Thieken, Philip Bubeck, Anna Heidenreich, Jennifer von Keyserlingk, Lisa Dillenardt, and Antje Otto. Performance of the flood warning system in Germany in July 2021 – insights from affected residents. *Natural Hazards and Earth System Sciences*, 23(2) :973–990, 2023. ISSN 1561-8633. doi : 10.5194/nhess-23-973-2023. URL https://nhess.copernicus.org/articles/23/973/2023/. Publisher : Copernicus GmbH.

J. Thielen, J. Bartholmes, M.-H. Ramos, and A. de Roo. The European Flood Alert System – Part 1 : Concept and development. *Hydrology and Earth System Sciences*, 13(2) :125–140, February 2009. ISSN 1027-5606. doi : 10.5194/hess-13-125-2009. URL https://hess.copernicus.org/articles/13/125/2009/hess-13-125-2009.html. Publisher : Copernicus GmbH.

H. A. Titley, H. L. Cloke, E. M. Stephens, F. Pappenberger, and E. Zsoter. Using Ensembles to Analyze Predictability Links in the Tropical Cyclone Flood Forecast Chain. *Journal of Hydrometeorology*, 25(1) :191–206, January 2024. ISSN 1525-7541, 1525-755X. doi : 10.1175/JHM-D-23-0022.1. URL https://journals.ametsoc.org/view/journals/hydr/25/1/JHM-D-23-0022.1.xml. Publisher : American Meteorological Society Section : Journal of Hydrometeorology.

Yves Tramblay and Samuel Somot. Future evolution of extreme precipitation in the Mediterranean. *Climatic Change*, 151 (2) :289–302, November 2018. ISSN 1573-1480. doi : 10.1007/s10584-018-2300-5. URL https://doi.org/10.1007/s10584-018-2300-5.

Yves Tramblay, Louise Mimeau, Luc Neppel, Freddy Vinet, and Eric Sauquet. Detection and attribution of flood trends in Mediterranean basins. *Hydrology and Earth System Sciences*, 23(11) :4419–4431, October 2019. ISSN 1027-5606. doi : 10.5194/hess-23-4419-2019. URL https://hess.copernicus.org/articles/23/4419/2019/. Publisher : Copernicus GmbH.

Yves Tramblay, Aurélien Ribes, Samuel Somot, Luc Neppel, Philippe Lucas-Picher, Freddy Vinet, and Eric Sauquet. Impacts du changement climatique sur les pluies intenses et les crues en Méditerranée. *LHB*, 107(1) :1–5, January 2021. ISSN null. doi : 10.1080/00186368.2021.1912971. URL https://doi.org/10.1080/00186368.2021.1912971. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/00186368.2021.1912971.

Yves Tramblay, Guillaume Thirel, Laurent Strohmenger, Guillaume Evin, Lola Corre, Louis Heraut, and Eric Sauquet. Evolution of flood generating processes under climate change in France. *EGUsphere*, pages 1–26, April 2025. doi : 10.5194/egusphere-2025-1635. URL https://egusphere.copernicus.org/preprints/2025/egusphere-2025-1635/. Publisher : Copernicus GmbH.

UNDRR. Understanding and breaking the cycle of risk, November 2019. URL https://www.undrr.org/building-risk-knowledge/understanding-risk.

UNDRR. Hazard definition & classification review. Technical report, United Nations Office for Disaster Risk Reduction, 2020.

UNDRR. Disaster Risk Reduction Terminology, March 2023. URL https://www.undrr.org/drr-glossary/terminology.

UNICEF. Weather-related disasters led to 43.1 million displacements of children over six years, October 2023. URL https://www.unicef.org/press-releases/weather-related-disasters-led-431-million-displacements-children-over-six-years.

P.-A. Versini, E. Gaume, and H. Andrieu. Application of a distributed hydrological model to the design of a road inundation warning system for flash flood prone areas. *Natural Hazards and Earth System Sciences*, 10(4) :805–817, April 2010. ISSN 1561-8633. doi : 10.5194/nhess-10-805-2010. URL https://nhess.copernicus.org/articles/10/805/2010/. Publisher : Copernicus GmbH.

Sergio M. Vicente-Serrano, Yves Tramblay, Fergus Reig, José C. González-Hidalgo, Santiago Beguería, Michele Brunetti, Ksenija Cindrić Kalin, Leonardo Patalen, Aleksandra Kržič, Piero Lionello, Miguel M. Lima, Ricardo M. Trigo, Ahmed M. El-Kenawy, Ali Eddenjal, Murat Türkes, Aristeidis Koutroulis, Veronica Manara, Maurizio Maugeri, Wafae Badi, Shifa Mathbout, Renato Bertalanič, Lilia Bocheva, Ismail Dabanli, Alexandru Dumitrescu, Brigitte Dubuisson, Salah Sahabi-Abed, Fayez Abdulla, Abbas Fayad, Sabina Hodzic, Mirjana Ivanov, Ivan Radevski, Dhais Peña-Angulo, Jorge Lorenzo-Lacruz, Fernando Domínguez-Castro, Luis Gimeno-Sotelo, Ricardo García-Herrera, Magí Franquesa, Amar Halifa-Marín, Maria Adell-Michavila, Ivan Noguera, David Barriopedro, Jose M. Garrido-Perez, Cesar Azorin-Molina, Miguel Andres-Martin, Luis Gimeno, Raquel Nieto, Maria Carmen Llasat, Yannis Markonis, Rabeb Selmi, Soumaya Ben Rached, Slavica Radovanović, Jean-Michel Soubeyroux, Aurélien Ribes, Mohamed Elmehdi Saidi, Siham Bataineh, El Mahdi El Khalki, Sayed Robaa, Amina Boucetta, Karam Alsafadi, Nikos Mamassis, Safwan Mohammed, Beatriz Fernández-Duque, Sorin Cheval, Sara Moutia, Aleksandra Stevkov, Silvana Stevkova, M. Yolanda Luna, and Vera Potopová. High temporal variability not trend dominates Mediterranean precipitation. *Nature*, 639(8055) :658–666, March 2025. ISSN 1476-4687. doi : 10.1038/s41586-024-08576-6. URL https://www.nature.com/articles/s41586-024-08576-6. Publisher : Nature Publishing Group.

Béatrice Vincendon, Véronique Ducrocq, Olivier Nuissier, and Benoît Vié. Perturbation of convection-permitting NWP forecasts for flash-flood ensemble forecasting. *Natural Hazards and Earth System Sciences*, 11 :1529–1544, 2011. doi : 10.5194/nhess-11-1529-2011. URL https://hal-meteofrance.archives-ouvertes.fr/meteo-00595253. Publisher : Copernicus Publ. / European Geosciences Union.

Freddy Vinet, Jean-Philippe Cherel, Karine Weiss, Margaux Lewandowski, and Laurent Boissier. Flood related mortality in the French Mediterranean region (1980–2020). *LHB*, 108(1) :2097022, December 2022. ISSN null. doi : 10.1080/27678490.2022.2097022. URL https://doi.org/10.1080/27678490.2022.2097022. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/27678490.2022.2097022.

B. Vié, G. Molinié, O. Nuissier, B. Vincendon, V. Ducrocq, F. Bouttier, and E. Richard. Hydro-meteorological evaluation of a convection-permitting ensemble prediction system for Mediterranean heavy precipitating events. *Natural Hazards*

and Earth System Sciences, 12(8) :2631–2645, August 2012. ISSN 1561-8633. doi : 10.5194/nhess-12-2631-2012. URL https://nhess.copernicus.org/articles/12/2631/2012/. Publisher : Copernicus GmbH.

Kévin Walcarius. *Prévision immédiate probabiliste de précipitations adaptée à la prévision de crues*. These de master, Ecole Centrale de Lyon, August 2021.

Dantong Wang, Paul Stapor, and Jan Hasenauer. Dirac mixture distributions for the approximation of mixed effects models**We acknowledge financial support by the China Scholarship Council (201706060200) and the German Ministry for Education and Research (BMBF) in the FitMultiCell project (grant number 031L0159A). *IFAC-PapersOnLine*, 52(26) :200–206, January 2019. ISSN 2405-8963. doi : 10.1016/j.ifacol.2019.12.258. URL https://www.sciencedirect.com/science/article/pii/S2405896319321433.

Fuxing Wang, Jan Polcher, Philippe Peylin, and Vladislav Bastrikov. Assimilation of river discharge in a land surface model to improve estimates of the continental water cycles. *Hydrology and Earth System Sciences*, 22(7) :3863–3882, July 2018. ISSN 1027-5606. doi : 10.5194/hess-22-3863-2018. URL https://hess.copernicus.org/articles/22/3863/2018/. Publisher : Copernicus GmbH.

Jiahu Wang, Yang Hong, Li Li, Jonathan J. Gourley, Sadiq I. Khan, Koray K. Yilmaz, Robert F. Adler, Frederick S. Policelli, Shahid Habib, Daniel Irwn, Ashutosh S. Limaye, Tesfaye Korme, and Lawrence Okello. The coupled routing and excess storage (CREST) distributed hydrological model. *Hydrological Sciences Journal*, 56(1) :84–98, 2011. ISSN 0262-6667. doi : 10.1080/02626667.2010.543087. URL https://doi.org/10.1080/02626667.2010.543087. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/02626667.2010.543087.

Xinying Wang, Miao Xiao, Yi Liu, Jun Guo, Yangyang Qin, and Yunkang Zhang. A rapid and efficient method for flash flood simulation based on deep learning. *Engineering Applications of Computational Fluid Mechanics*, 18(1) :2407016, December 2024. ISSN 1994-2060. doi : 10.1080/19942060.2024.2407016. URL https://doi.org/10.1080/19942060.2024.2407016. Publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/19942060.2024.2407016.

Ida K. Westerberg, Giuliano Di Baldassarre, Keith J. Beven, Gemma Coxon, and Tobias Krueger. Perceptual models of uncertainty for socio-hydrological systems : a flood risk change example. *Hydrological Sciences Journal*, 62(11), August 2017. ISSN 0262-6667. Publisher : Taylor & Francis.

Philippe Weyrich, Anna Scolobig, Florian Walther, and Anthony Patt. Responses to severe weather warnings and affective decision-making. *Natural Hazards and Earth System Sciences*, 20(10) :2811–2821, October 2020. ISSN 1561-8633. doi : 10.5194/nhess-20-2811-2020. URL https://nhess.copernicus.org/articles/20/2811/2020/. Publisher : Copernicus GmbH.

Michael H. Wimmer, Markus Hollaus, Günter Blöschl, Andreas Buttinger-Kreuzhuber, Jürgen Komma, Jürgen Waser, and Norbert Pfeifer. Processing of nationwide topographic data for ensuring consistent river network representation. *Journal of Hydrology X*, 13 :100106, December 2021. ISSN 2589-9155. doi : 10.1016/j.hydroa.2021.100106. URL https://www.sciencedirect.com/science/article/pii/S2589915521000341.

Oliver Wing, Niall Quinn, Pete Uhe, James Savage, Chris Sampson, Nans Addor, Natalie Lord, Tom Collings, Simbi Hatchard, Jannis Hoch, Andy Smith, Anthony Cooper, Joe Bates, Hamish Wilkinson, Sam Himsworth, Izzy Probyn, Ivan Haigh, Jeff Neal, and Paul Bates. A 30 m resolution global fluvial&ndash ;pluvial&ndash ;coastal flood inundation model for any climate scenario. Copernicus Meetings, February 2023. doi : 10.5194/egusphere-egu23-6383. URL https://meetingorganizer.copernicus.org/EGU23/EGU23-6383.html. Conference Name : EGU23.

WMO. *Guidelines on Ensemble Prediction Systems and Forecasting*. World Meteorological Organization, 2012. ISBN 978-92-63-11091-6. URL https://library.wmo.int/index.php?lvl=notice_display&id=12962#.Y_N6jq3jKUk. Google-Books-ID : lwOUoAEACAAJ.

WMO. Guidelines on Multi-hazard Impact-based Forecast and Warning Services. Technical report, World Meteorological Organization, 2015. URL https://www.preventionweb.net/publication/wmo-guidelines-multi-hazard-impact-based-forecast-and-warning-services.

WMO. Multi-hazard Early Warning Systems : A Checklist. Technical report, World Meteorological Organization, 2018. URL https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://community.wmo.int/sites/default/files/EWS_Checklist_0.pdf%3FcNCI0mkEKe2FY7nJrkC09atRuoqTCyfc&ved=2ahUKEwji1YCz0fiKAxXlUqQEHejfJTAQFnoECA4QAQ&usg=AOvVaw3N8zLBjo8tztJcoSBUPHAT.

WMO. Climate and water (2020) - Floods, 2020. URL https://wmo.int/about-us/world-meteorological-day/wmd-2020/floods.

Huan Wu, John S. Kimball, Nate Mantua, and Jack Stanford. Automated upscaling of river networks for macroscale hydrological modeling. *Water Resources Research*, 47(3), 2011. ISSN 1944-7973. doi : 10.1029/2009WR008871. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2009WR008871. _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1029/2009WR008871.

Faye Wyatt, Joanne Robbins, and Rebecca Beckett. Investigating bias in impact observation sources and implications for impact-based forecast evaluation. *International Journal of Disaster Risk Reduction*, 90 :103639, May 2023. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2023.103639. URL https://www.sciencedirect.com/science/article/pii/S221242092300119X.

Faye Wyatt, Joanne Robbins, and Simon Eaton. Implementing a routine and standard approach for the automatic collection of socio-economic impact observations for impact-based forecasting and warning. *International Journal of Disaster Risk Reduction*, 110 :104608, August 2024. ISSN 2212-4209. doi : 10.1016/j.ijdrr.2024.104608. URL https://www.sciencedirect.com/science/article/pii/S2212420924003704.

D. Yamazaki, Yuji MASUTOMI, Taikan OKI, and Shinjiro KANAE. An Improved Upscaling Method to Construct a Global River Map. Beijing, 2008.

Liang Emlyn Yang, Jürgen Scheffran, Diana Süsser, Richard Dawson, and Yongqin David Chen. Assessment of Flood Losses with Household Responses : Agent-Based Simulation in an Urban Catchment Area. *Environmental Modeling & Assessment*, 23(4) :369–388, August 2018. ISSN 1573-2967. doi : 10.1007/s10666-018-9597-3. URL https://doi.org/10.1007/s10666-018-9597-3.

Jixin Yu, Jinping Liu, Ji-Won Baek, Clarence Fong, and Michael Fu. Impact-based Forecasting for Improving the Capacity of Typhoon-related Disaster Risk Reduction in Typhoon Committee Region. *Tropical Cyclone Research and Review*, September 2022. ISSN 2225-6032. doi : 10.1016/j.tcrr.2022.09.003. URL https://www.sciencedirect.com/science/article/pii/S2225603222000248.

Andre D. L. Zanchetta and Paulin Coulibaly. Recent Advances in Real-Time Pluvial Flash Flood Forecasting. *Water*, 12(2) :570, 2020. ISSN 2073-4441. doi : 10.3390/w12020570. URL https://www.mdpi.com/2073-4441/12/2/570. Number : 2 Publisher : Multidisciplinary Digital Publishing Institute.

Ruikang Zhang, Dedi Liu, Lihua Xiong, Jie Chen, Hua Chen, and Jiabo Yin. Determining the threshold of issuing flash flood warnings based on people's response process simulation. *Hydrology and Earth System Sciences*, 28(23) :5229–5247, December 2024. ISSN 1027-5606. doi : 10.5194/hess-28-5229-2024. URL https://hess.copernicus.org/articles/28/5229/2024/. Publisher : Copernicus GmbH.

Fang Zhao, Ted I. E. Veldkamp, Katja Frieler, Jacob Schewe, Sebastian Ostberg, Sven Willner, Bernhard Schauberger, Simon N. Gosling, Hannes Müller Schmied, Felix T. Portmann, Guoyong Leng, Maoyi Huang, Xingcai Liu, Qiuhong Tang, Naota Hanasaki, Hester Biemans, Dieter Gerten, Yusuke Satoh, Yadu Pokhrel, Tobias Stacke, Philippe Ciais, Jinfeng Chang, Agnes Ducharne, Matthieu Guimberteau, Yoshihide Wada, Hyungjun Kim, and Dai Yamazaki. The critical role of the routing scheme in simulating peak river discharge in global hydrological models. *Environmental Research Letters*, 12(7) :075003, June 2017. ISSN 1748-9326. doi : 10.1088/1748-9326/aa7250. URL https://dx.doi.org/10.1088/1748-9326/aa7250. Publisher : IOP Publishing.

Annexes

Annexe A

Utilisation de la BD TOPO

1 Représentation des bâtiments résidentiels

La couche "BATIMENTS" de la BD TOPO 2020 n'indique pas clairement les bâtiments résidentiels. En effet, s'il existe des champs permettant de décrire l'usage d'un bâtiment, avec des valeurs possibles parmi "Agricole", "Annexe", "Commercial et services", "Indifférencié", "Industriel", "Religieux", "Résidentiel" et "Sportif", il ne serait pas pertinent de faire une sélection sur la valeur "Résidentiel", car de nombreux exemples montrent que des bâtiments classés en "Indifférencié" correspondent bel et bien à des habitations. Pour cette raison, nous avons appliqué la sélection suivante (inspirée de Fujiki and Finance [2022]) :

Listing A.1 – Requête SQL pour la sélection des batiments résidentiels dans la BD TOPO 2020

```
SELECT *
FROM bat
WHERE (
    (bat.NATURE = 'Indifferenciee'
    AND
    (bat.USAGE1 IN ('Indifferencie', 'Residentiel')
    OR
    bat.USAGE2 IN ('Indifferencie', 'Residentiel')))
    OR
    (bat.NATURE <> 'Indifferenciee'
    AND
    (bat.USAGE1 = 'Residentiel'
    OR
    bat.USAGE2 = 'Residentiel'))
    AND
    bat.ETAT = 'En service'
    AND
    bat.IN_ZONE_ACTI = FALSE
    AND
    bat.HAUTEUR >= 1.8
    AND
    bat.AREA >= 20
);
```

Les champs NATURE, USAGE1 et USAGE2 sont présents dans la couche "BATIMENT" (nommée ici "bat") de la BD TOPO 2020. En revanche, les champs IN_ZONE_ACTI et AREA ont du être

manuellement ajoutés. La signification de chaque champ est décrite ci-dessous :

- NATURE : Attribut permettant de distinguer différents types de bâtiments selon leur architecture.
- USAGE1 : Usage du bâtiment ou d'une partie du bâtiment.
- USAGE2 : Autre usage d'un bâtiment de fonction mixte.
- IN_ZONE_ACTI : Booléen définissant si le bâtiment se situe en zone d'activité (couche ZONE_D_ACTIVITE_OU_ de la BD TOPO 2020).
- AREA : Surface au sol du bâtiment, en mètre carrés.

2 Représentation des bâtiments agricoles, commerciaux et industriels

Listing A.2 – Requête SQL pour la sélection des batiments agricoles commerciaux et industriels dans la BD TOPO 2020

```sql
SELECT *
FROM bat
WHERE (
    (bat.NATURE = 'Industriel, agricole ou commercial'
    AND
    (bat.USAGE1 IN ('Indifferencie', 'Agricole', 'Commercial et services',
        'Industriel')
    OR
    bat.USAGE2 IN ('Indifferencie', 'Agricole', 'Commercial et services',
        'Industriel')))
    OR
    (bat.NATURE <> 'Industriel, agricole ou commercial'
    AND
    (bat.USAGE1 IN ('Agricole', 'Commercial et services', 'Industriel')
    OR
    bat.USAGE2 IN ('Agricole', 'Commercial et services', 'Industriel')))
    AND
    bat.ETAT = 'En service'
);
```

Annexe B

Annexe de l'article Godet et al. [2025b]

1 Feedbacks from Vigicrues Flash users

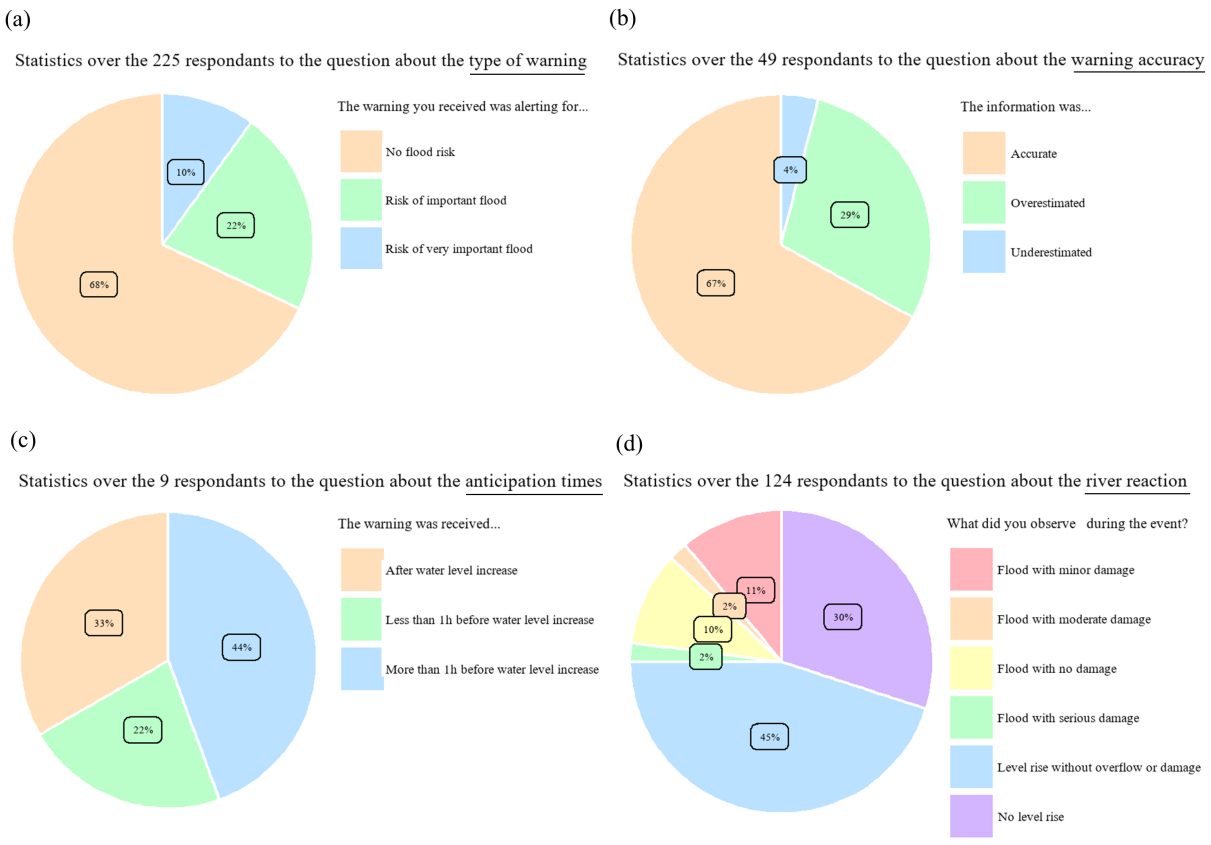


FIGURE B.1 – Pie charts of feedbacks from users who received a warning from Vigicrues Flash between 2021 and 2024

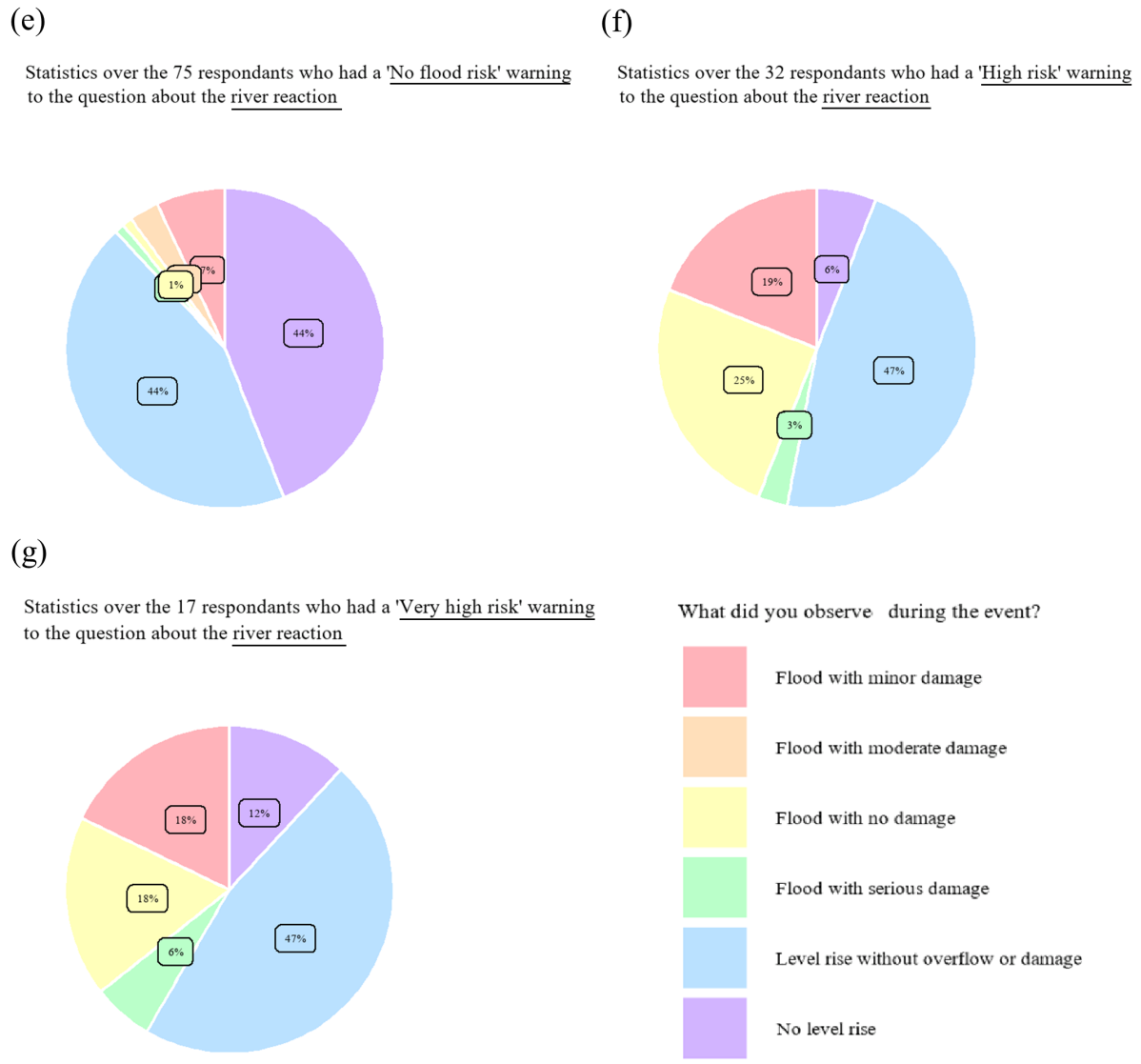


FIGURE B.1 – Pie charts of feedbacks from users who received a warning from Vigicrues Flash between 2021 and 2024

2 Contingency scores

2.1 Classical scores

— The critical sucess index, which looks at missed events and false alerts (without differentiating between these two types of error), and whose optimal score is 1, is defined as follows :

$$CSI = \frac{a}{a+b+c}$$

— The hit rate or probability of detection, which looks at all the events observed, and which is 0 when all the observed events are missed, and is 1 when all the observed events are detected, is defined as follows :

$$POD = \frac{a}{a+c}$$

— The false alarm rate or probability of false detection, which looks at all the non-events observed, and is 0 when there are no false alarms, and is 1 when there is a false alarm issued for each non-event

observed, is defined as follows :

$$POFD = \frac{b}{b+d}$$

— The frequency bias, which approaches 0 when the model tends to miss too many alerts, which increases when the number of false alerts increases, and whose optimal score is 1, is defined as follows :

$$FB = \frac{a+b}{a+c}$$

— The success ratio (respectively the false alarm ratio), which looks at all the alerts issued and is 1 (respectively 0) when there are no false alarms, and is 0 (respectively 1) when all the alerts issued are false alarms, is defined as follows :

$$SR = \frac{a}{a+b}, \quad FAR = \frac{b}{a+b}$$

2.2 The Pearson correlation coefficient

Given a pair of random variables (X, Y), the Pearson correlation coefficient can be expressed as :

$$R = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - (E(X))^2}\sqrt{E(Y^2) - (E(Y))^2}}$$

In our application, where the Pearson correlation coefficient is used as a measure of the diagonality of the square matrix M (of size n), it can be expressed as follows :

$$R = \frac{s_M w - u_x u_y}{\sqrt{s_M v_x - {u_x}^2}\sqrt{s_M v_y - {u_y}^2}}$$

Where :

— j is the $n-$long vector of ones.

— $r_1 = (1, 2, 3, ..., n)$

— $r_2 = (1^2, 2^2, 3^2, ..., n^2)$

— $s_M = jMj^T$

— $u_x = r_1 M j^T$

— $u_y = jM{r_1}^T$

— $v_x = r_2 M j^T$

— $v_y = jM{r_2}^T$

— $w = r_1 M {r_1}^T$

3 Supplementary contingency results for the hydrological validation

Tableau B.1 – Supplementary Table to Table 4 - Contingency results for the validation based on discharges recorded at gauging stations (bold, top of each line : on validation period 2018-2023 ; bottom of each line : on calibration period 2008-2018 ; orange : rivers eligible for Vigicrues Flash ; blue : rivers non eligible for Vigicrues Flash)

	2y		5y		10y		20y		50y		100y		500y		1000y	
POD	**0.71**	**0.71**	**0.73**	**0.60**	**0.75**	**0.50**	**0.84**	**0.50**	**0.75**	**0.50**	**0.78**	**1.00**	**0.00**	**NA**	**NA**	**NA**
	0.70	0.75	0.72	0.64	0.52	0.40	0.48	0.27	0.33	0.50	NA	1.00	NA	0.00	NA	0.00
POFD	**0.12**	**0.12**	**0.06**	**0.07**	**0.05**	**0.04**	**0.03**	**0.03**	**0.02**	**0.02**	**0.02**	**0.01**	**0.00**	**0.00**	**0.00**	**0.00**
	0.08	0.13	0.05	0.06	0.03	0.04	0.02	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00
FB	**1.18**	**1.18**	**1.43**	**1.13**	**1.82**	**1.36**	**1.95**	**3.17**	**2.25**	**5.50**	**2.89**	**8.00**	**6.00**	**NA**	**Inf**	**NA**
	1.06	1.26	1.36	1.30	1.30	1.48	1.64	1.64	2.83	2.00	Inf	3.00	Inf	0.00	Inf	0.00
SR	**0.60**	**0.61**	**0.51**	**0.53**	**0.41**	**0.37**	**0.43**	**0.16**	**0.33**	**0.09**	**0.27**	**0.13**	**0.00**	**NA**	**0.00**	**NA**
	0.66	0.59	0.53	0.49	0.40	0.27	0.29	0.17	0.12	0.25	0.00	0.33	0.00	NA	0.00	NA
FAR	**0.40**	**0.39**	**0.49**	**0.47**	**0.59**	**0.63**	**0.57**	**0.84**	**0.67**	**0.91**	**0.73**	**0.88**	**1.00**	**NA**	**1.00**	**NA**
	0.34	0.41	0.47	0.51	0.60	0.73	0.71	0.83	0.88	0.75	1.00	0.67	1.00	NA	1.00	NA

Tableau B.2 – Supplementary Table to Table 5 - Contingency results for the validation based on estimated peak discharges (bold, top of each line : on validation period 2018-2023 ; bottom of each line : on calibration period 2008-2018 ; orange : rivers eligible for Vigicrues Flash ; blue : rivers non eligible for Vigicrues Flash)

	2y		5y		10y		20y		50y		100y		500y		1000y	
POD	**0.90**	**1.00**	**0.90**	**1.00**	**0.88**	**0.95**	**0.87**	**1.00**	**0.71**	**0.67**	**0.57**	**0.57**	**0.52**	**1.00**	**0.19**	**0.67**
	0.99	0.96	0.94	0.96	0.92	0.98	0.90	0.93	0.82	0.85	0.69	0.88	0.42	0.20	0.50	0.00
POFD	**NA**	**1.00**	**NA**	**0.60**	**1.00**	**0.57**	**0.25**	**0.58**	**0.13**	**0.47**	**0.10**	**0.32**	**0.25**	**0.09**	**0.26**	**0.09**
	1.00	NA	1.00	0.00	0.71	0.33	0.63	0.71	0.42	0.62	0.44	0.48	0.21	0.21	0.11	0.20
FB	**0.90**	**1.08**	**0.90**	**1.14**	**0.90**	**1.16**	**0.89**	**1.50**	**0.73**	**1.56**	**0.59**	**1.43**	**0.91**	**1.50**	**0.88**	**1.33**
	1.00	0.96	0.98	0.96	1.05	1.00	1.14	1.05	1.12	1.09	1.34	1.33	1.75	2.00	1.63	4.50
SR	**1.00**	**0.92**	**1.00**	**0.88**	**0.98**	**0.82**	**0.98**	**0.67**	**0.97**	**0.43**	**0.97**	**0.40**	**0.57**	**0.67**	**0.21**	**0.50**
	0.99	1.00	0.96	1.00	0.87	0.98	0.79	0.88	0.74	0.78	0.51	0.66	0.24	0.10	0.31	0.00
FAR	**0.00**	**0.08**	**0.00**	**0.13**	**0.02**	**0.18**	**0.02**	**0.33**	**0.02**	**0.57**	**0.03**	**0.60**	**0.43**	**0.33**	**0.79**	**0.50**
	0.01	0.00	0.04	0.00	0.13	0.02	0.21	0.12	0.26	0.22	0.49	0.34	0.76	0.90	0.69	1.00

Annexe C

Annexe de l'article Godet et al. [2025a]

1 Contents of the associated dataset

Table C.1 provides a quick access guide to the dataset. Details on the attributes, units, sources, data size, formats, etc., can be found in the readme.txt files provided with each data item, in the dataset.

Data type	Data path	Signification	Section
Vector	02_FLOOD_EVENTS / 01_INPUTS / DISCHARGE / RIVER_NETWORK	River network on which the CINECAR hydrological modelling was performed, for the simulation of the three historical flood events.	2.3.1
	02_FLOOD_EVENTS / 03_VALIDATION / FLOOD_MARKS	High Water Marks collected for the three historical flood events.	2.3.2
	02_FLOOD_EVENTS / 03_VALIDATION / OBSERVED_FLOOD_EXTENTS / ASSESSMENT_AREAS	Subcatchments affected by the Argens 2010 and Aude 2018 events, adapted to the observed flood extents.	2.3.2
	02_FLOOD_EVENTS / 03_VALIDATION / OBSERVED_FLOOD_EXTENTS / OBS_FLOOD_SHP	Observed flood extents for the Argens 2010 and Aude 2018 events.	2.3.2
Raster	01_FLOOD_HAZARD / 01_INPUTS / DTM / RAW	5m DTMs resulting from a sampling of the 1m DTMs provided by IGN, for each area on which a flood hazard map is derived.	2.2.1
	01_FLOOD_HAZARD / 01_INPUTS / DTM / PROCESSED	5m DTMs that underwent the treatments described in section 3.1, for each area on which a flood hazard map is derived.	2.2.1, 3.1
	02_FLOOD_EVENTS / 01_INPUTS / DTM / RAW	5m DTMs resulting from a sampling of the 1m DTMs provided by IGN, for each catchment on which a historical flood event is simulated.	2.3.1
	02_FLOOD_EVENTS / 01_INPUTS / DTM / PROCESSED	5m DTMs that underwent the treatments described in section 3.1, for each catchment on which a historical flood event is simulated.	2.3.1 3.1
	01_FLOOD_HAZARD / 01_INPUTS / HYDROGRAPHIC_NETWORK	Raster description of the minor bed for each area on which a flood hazard map is derived.	2.2.1
	02_FLOOD_EVENTS / 01_INPUTS / HYDROGRAPHIC_NETWORK	Raster description of the minor bed for each catchment on which a historical flood event is simulated.	2.3.1
	01_FLOOD_HAZARD / 01_INPUTS / DISCHARGE	Input SHYREG discharge quantiles for the simulation of flood hazard maps.	2.2.1
	01_FLOOD_HAZARD / 02_OUTPUTS / WATER_LEVELS / RAW	Floodos flood hazard maps (water levels) obtained using raw DTMs.	2.2.1
	01_FLOOD_HAZARD / 02_OUTPUTS / WATER_LEVELS / PROCESSED	Floodos flood hazard maps (water levels) obtained using processed DTMs.	2.2.1
	02_FLOOD_EVENTS / 02_OUTPUTS / WATER_LEVELS / RAW	Floodos flood event maps (water levels) obtained using raw DTMs.	2.3.1
	02_FLOOD_EVENTS / 02_OUTPUTS / WATER_LEVELS / PROCESSED	Floodos flood event maps (water levels) obtained using processed DTMs.	2.3.1
Table	01_FLOOD_HAZARD / 03_VALIDATION / RATING_CURVES	Rating curve data for 171 gauging stations.	2.2.2
	02_FLOOD_EVENTS / 01_INPUTS / DISCHARGE / HYDROGRAPH_TABLES / ... / hydrographs	Simulated discharge values at each 15 min time steps for the three historical events, on each river reach of the river network.	2.3.1
	02_FLOOD_EVENTS / 01_INPUTS / DISCHARGE / HYDROGRAPH_TABLES / ... / lateral_hydrographs	Subcatchment lateral inflow values at each 15 min time steps for the three historical events, on each river reach of the river network.	2.3.1

TABLEAU C.1 – Contents of the dataset

2 Procedure applied on the rating curve data

The data provided in the dataset do not correspond to the raw data extracted from Hydroportail/-BAREME - they have undergone a procedure to provide uniform (data extracted from Hydroportail are not formatted the same way as data extracted from BAREME), verified (some errors in the data could be detected), suitable (for flood hazard maps verification purposes), and simplified (keeping only the variable of interest) information. The procedure is as follows :

— Concerning the locations : stations outside the study area (Figure III.3) were removed. Stations that could not be connected to a pixel of the stream network (distance criterion = 100m) were also removed. Stations that could not be connected to a SHYREG pixel (distance criterion = 150m) were removed. Remaining stations were manually relocated, and the positions were checked with the local hydrometric services. This manual verification can be time-consuming but is crucial, especially for stations located near confluences, which can be attributed to the wrong river reaches. It is important to note that the coordinates of a station do not always correspond to the location of the measurement device, but sometimes to the location of the recording or remote connecting box, that can be located several hectometres from the surveyed river cross-section.

— Concerning the values : stations without a zero reference altitude of the water level gauge were removed, and remaining values were checked : obvious altimetric errors were corrected, and less obvious but detectable errors were discussed with the local hydrometric services. Stations without any recent rating curve (date threshold = 2010) were removed. Rating curves without any point higher than the 2-year return period were removed. In case of several available rating curves for the same gauging station, only the most recent curve was considered.

3 Brief description of the DTM preprocessing method

The DTM preprocessing procedures carried out in this study are very similar to those described by Wimmer et al. [2021], they are organised here into four main stages :

1. Correcting the river axes. To achieve this, orthogonal cross sections were extracted every 1m on the original river network (extracted from BD TOPAGE®, see section 2.2.1), resulting in a cross-sectional profile function $h(d)$ where d is the signed distance orthogonal to the river axis. Then, six weight functions accounting for different geometric and hydraulic criteria (such as the cross section form, water flow properties, etc.) are defined and combined into a global weight function w_{final}. If there exists a $d_{max} > 0$ verifying $w_{final}(d_{max}) = max(w_{final})$, then this d_{max} defines the new location of the corrected river axis, for the considered cross section. These treatments only concern the hydrographic network, and thus keep the DTMs unchanged.

2. Detecting the riverbanks. In order to do this, we work on the same cross sections as before, and we translate the expected shape of the upper edge of embankment (strong negative curvature due to the flattening above the riverbanks) into a mathematical requirement (high first and small second derivative of the cross-sectional profile). Compared to Wimmer et al. [2021], an additional criterion (threshold on the break of slope) was used to better represent V-shaped small rivers.

3. Validating the riverbanks detection. While the detection of riverbanks resulted in satisfying delineations of the streambeds for a high proportion of cross sections, some erratic behaviours have been observed, especially in case of highly vegetated riverbanks. Several criteria were used to decide whether to eliminate a cross section : intersection with water bodies, several intersections with

another river, intersection with bridges, buildings, sports grounds, roads, railways, inconsistency with reference river width classes, detected river width as large as the initial cross section.

4. Reshaping the streambed. First, the positions of the right and left banks of the streambed were smoothed with a moving average procedure implemented on the cross-sectional coordinates of the 5 closest upstream and downstream cross-sections. Then, the longitudinal profile of the minimal altitudes of all cross-sections was checked and local maximums were replaced by linearly interpolated values to produce a river bed with negative longitudinal profile slopes. Finally, in the DTMs, the minimal altitudes of the each cross sections were applied to the whole width of the detected proportion of the stream bed located between the river banks.

Figure C.1 illustrates the results on a 2 km-long reach of the Orb river, upstream of the city of Béziers. In this case, the errors in the raw DTMs, identifiable by the triangular shapes in the streambed, were probably caused by interpolation errors (presumably in the absence of lidar points).

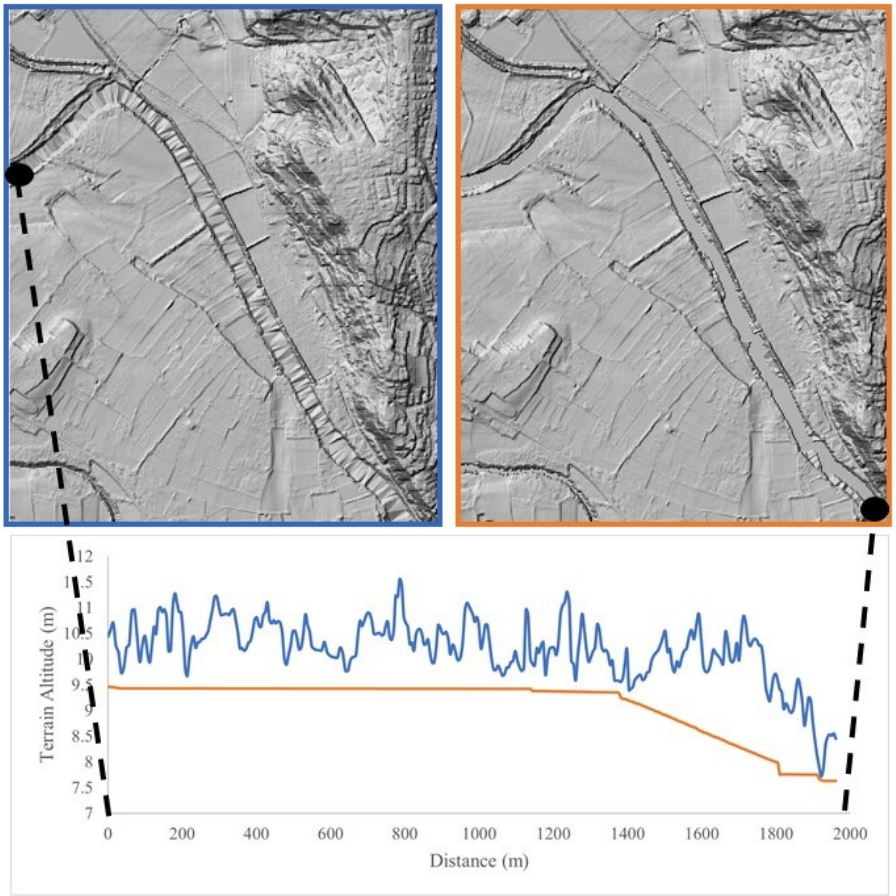


FIGURE C.1 – Illustration of the impact of the DTM processing on the Orb river (near the city of Béziers) streambed cross-sectional and longitudinal profiles : original (left) and processed (right) DTMs and corresponding original (blue) and processed (orange) longitudinal profiles.

4 Details on the implementation of Floodos in this study

Constraints linked to the amount of data to be handled first led to divide the domain into 61 subdomains (see figure III.3), and then to divide the river network of each subdomain into river reaches. This division was initially made at the confluences, and river reaches longer than 10km were divided again, the objective being to create calculation domains with a number of cells not exceeding 25×10^6. In the end, the average length of a river reach is about 5km. Each river reach was extended by 500m downstream, in order to limit

the influence of the dowstream boundary condition on the simulation results. The extent of the domain was also extended by 2km on each side, in order to reduce inconsistencies at confluences (in particular due to possible backwater influences), and to ensure that the computation domain limit is located sufficiently far downstream of the considered river reach. The mesh underwent other treatments, such as the addition of artificial borders upstream the computation domains to force the flow downstream the river reach, and the change (decrease) of elevation values for cells in the Mediterranean sea, to prevent uncontrolled backwater effects at the coastline . These treatments led to the creation of the topography ".alt" file.

The discharges ".rain" file is a grid which values define the discharges to be injected at each injection point. These injection points, corresponding to 5m resolution pixels, need to be correctly located (i.e., on the correct minor bed, which can be tricky near confluences), even though the input SHYREG or CINECAR discharges are originally given at the 50m resolution. Several methods are possible to address this issue, see [Godet et al., 2024]. In this work, points 50m apart were drawn along the hydrographic network, and SHYREG discharges were extracted at these points. The first discharge injection is made on the most upstream point of the river reach, and corresponds to the total discharge circulating in the river reach. Downstream of that point, all injections are increments.

The main challenge concerns the adjustment of the parameters controlingthe convergence of computations. The elementary volume corresponding to each precipiton should be carefully fixed to ensure convergence and computation speed. Davy et al. [2017] have defined a maximum possible value for the volume of the precipiton V^p ensuring the convergence :

$$V^p = 0.75 \times S_0 \times \Delta x^3 \tag{C.1}$$

Where S_0 is the water surface slope. Several tests aiming at finding a compromise between fast calculations and correct convergence led to the following solution, adapted from [Hocini, 2022] : during the initialisation phase of each simulation, the precipiton volume will be progressively reduced to $0.0625 m^3$, which respects the condition of equation C.1 as long as S_0 is lower than 0.07%. The duration of the initialisation phase, as well as the total duration of the simulation are indicated in table C.2, and are defined in Calculation Time Units (CTU). 1 CTU refers to the injection of a fixed number of precipitons. The convergence has been verified on the last 25 CTU of the simulation. The water level was considered as stabilised if the variation was lower than 1mm between iterations.

Parameter	Signification	Value
CTU_{coeff}	Average number of precipitons brought during 1 CTU on one injection point	40
$step$	Number of steps in the initialisation phase	4
$N_{CTU,init}$	Length, in CTU, of the initialisation phase	10
V^P_{start}	Initial volume of precipiton in m3	0.25
V^P_{end}	Final volume of precipiton in m3	0.0625
N_{CTU}	Number of CTU for each simulation	100
$Cond$	Max. authorized error for difference between water level and mean water level during the last 25 CTU, in meters	0.001

Tableau C.2 – Tables of parameters related to the convergence of the model

Calculations were parallelised on a 20 cores and 128 GB RAM cluster and took roughly two months.

Annexe D

Autres sources de données d'impacts observés pour l'évaluation

The three databases used in this study offer continuous coverage of specific flood impacts and are promising validation datasets. However, our impact data collection led to other data that can be very informative even if they do not completely fullfill the requirements of our study. Our validation approach focuses on geolocated data sets, that is, sources where the spatial distribution of impacts can be precisely mapped. However, flood impacts can also be documented in other sources, such as post-event reports (often referred to as 'lessons learned' or 'after-action reports') produced by government agencies or infrastructure operators, summary fact sheets compiled by institutions such as Météo-France, regional authorities, or water management agencies, as well as press archives [Bourguignon, 2014]. Although these sources fall outside the scope of our validation framework, they provide important contextual and qualitative information on past flood events.

1 Exhaustive databases focusing on other types of impacts

A major advantage of the three databases used in this study is their spatio-temporal exhaustiveness, meaning that they cover a continuous time period across the entire study area. Another advantage is that the impacts they represent can be relatively easily modeled through simplifications. As detailed and justified in Section 3, we compared the aforementioned impact observations to simple estimations of flooded buildings, given that CatNat and CCR data primarily relate to damage to buildings and most fire service operations also relate to flooded buildings. In table D.1 and in the text below, we introduce alternative impacts databases that share the same advantage of exhaustiveness but require additional complexity during modeling.

Tableau D.1 – Summary of other exhaustive impact databases

Database	**Description**	**Access and Coverage**	**Impact type**	**Spatial resolution**
The National Database of Bodily Injury Accidents	Accidents occuring on flooded public roads resulting in bodily injuries	Open-access, full spatio-temporal coverage	Human impact (accidents)	Polygon data (municipality scale) and/or Point data

Flood-related mortality database	Detailed database of flood-related fatalities, including circumstances, demographics, and location	Access upon request, full spatio-temporal coverage	Human impact (deaths)	Point data (uncertainties reaching several kilometers)
Road closures	Database of road closures due to flooding, including geolocated segments, closure duration, and potential causes	Restricted data, accessed on 30% of the study area	Road network	Line data (affected road segment)
Telecommunication network data	Records of telecommunication activity, including emergency calls or technical exchanges between devices and antenna	Private data, only samples for 2021 events were accessed	Telecommunication network	Point data (locations of antenna)
Electricity network data	Database of power outages and affected substations caused by floods, with timestamps and affected customer numbers.	Private data, only a sample for an event in 2019 was accessed	Electricity network	Point data (locations of affected posts)

1.1 The National Database of Bodily Injury Accidents

The National Database of Accidents of Bodily Injury provides information on the impact of floods on road safety. This open access database, available on the French Data Gouv platform, records accidents from 2005 to 2022 that meet three criteria : occurring on a public road, involving at least one vehicle, and resulting in bodily injury requiring care. The database includes detailed variables in four categories : characteristics, locations, vehicles, and users. For our study, we focussed on variables such as date, time, weather conditions and road surface conditions (e.g., flooded). After applying strict filters to identify flood-related accidents, only 43 cases were found in the 14 departments covered in our study. However, using the database as a validation data set is challenging, as accidents and casualties are not only caused by flooded roads but also depend on unpredictable human behavior. However, studies [Vinet et al., 2022, Lowrie et al., 2022] indicate that 35–50% of flood-related deaths involve vehicles, highlighting the urgent need to address risky behaviors, such as trying to cross flooded roads.

1.2 Flood-related mortality database

The Vict-In flood-related mortality database [Vinet et al., 2022], spans the 1980-2020 period and documents 328 flood-related deaths, with a high level of precision and extensive historical depth. For our research, we used a subset of data covering 2010-2020, which includes 129 deaths in our study area. The database offers detailed information on each fatality, such as age, gender, nationality, residency status, and occupation, along with the cause of death (e.g. drowning, accident, heart attack) and circumstances (e.g., at home, outdoors, in a vehicle). Additional data includes risk behaviors (for example, ignoring warnings, crossing barriers), housing details (e.g., floor level, type), and vehicle specifics (e.g., type, exit attempts). Environmental data is also included, such as river basin, water height, rainfall, and flood or

weather warnings at various stages of the event. Despite its richness, the database has limitations. The location of fatalities can be imprecise, with uncertainties that can reach several kilometers. Furthermore, the availability of detailed information is inconsistent; for example, the exact time of death is known for only 141 of 328 victims, although age is available for 318. Finally, similar to the previous database, using a flood-related mortality database as a validation data set presents challenges, as flood mortality is influenced by a multitude of factors [Terti et al., 2019a].

1.3 Road closures

Data on road closures during floods have been the focus of numerous studies, either for damage assessment [Gong et al., 2012] or impact prediction [Versini et al., 2010, Naulin et al., 2013]. Although real-time road conditions in France are accessible via platforms such as Inforoutes, historical data is less readily available, requiring collaboration with road management entities such as departmental councils, regional authorities, highway operators, or even platforms like Waze for more recent flood events. As part of previous projects and through new partnerships, we recovered data sets on flood-related road closures from various departments. Although these data sets vary in structure and detail, they typically include geolocated data on flood-induced road closures, such as road names, affected lengths, and, in some cases, the river responsible for the flooding. Data sets are often highly detailed, relying on departmental road reference systems with markers such as kilometer posts and precise coordinates to locate events. However, the level of temporal detail can vary : some datasets provide only daily timestamps, while others offer hourly granularity alongside additional variables, such as reopening times and qualitative descriptions of traffic conditions (e.g., alternating traffic, complete closure). These inconsistencies, while notable, do not diminish the utility of the datasets. The combination of location, temporal tracking, and continuity in data collection makes them invaluable for analyzing the spatio-temporal evolution of flood impacts on roads and improving flash flood impact forecasting. Methods exist to assess whether a road will be flooded using forecasted water levels [Payrastre and Lebouc, 2019], but introduce additional complexity compared to simpler impact-based models that focus on counting affected buildings in flooded areas.

1.4 Telecommunication network data

The use of telecommunication data to detect critical situations and implement customized warning systems is a promising and actively explored approach. For example, ANR DISCRET project used Orange data, a leading telecommunications operator in France, to identify abnormal network activity during events such as the Notre Dame fire in 2019, demonstrating a correlation between critical situations and an increase in activity in cell phone networks. In collaboration with Orange, we accessed sample data sets for several flood events that occurred in our study area in 2021. These data sets include various types of data, such as emergency calls and technical exchanges between devices and antennas. Analyzing telecommunication network activity during floods is complex, as network usage is shaped by a multitude of unpredictable factors, including variation in population density, emergency communications, and infrastructure disruptions. This makes it difficult to use as a validation data set. However, Figure D.1, which presents the activity of the telecommunication network during five flood events, reveals interesting patterns. Peaks in activity or sudden drops, blackouts, most likely caused by failures in the cell phone network, often align with extreme flood events. For instance, the September 2021 event in the Gard (where the maximum simulated return period reached at least 1000 years on some river reaches) shows a significant blackout. Intermediate events, such as those in August and November 2021 (T=20 years for the maximum return period), show moderate reactions (a small peak or a short blackout), while

smaller events (May and September 2021, T=5 to 10 years) do not appear to impact network activity. These correlations suggest that telecommunication data could serve as a valuable real-time indicator of flood severity and human response, offering potential applications for flood risk management and early detection and warning systems.

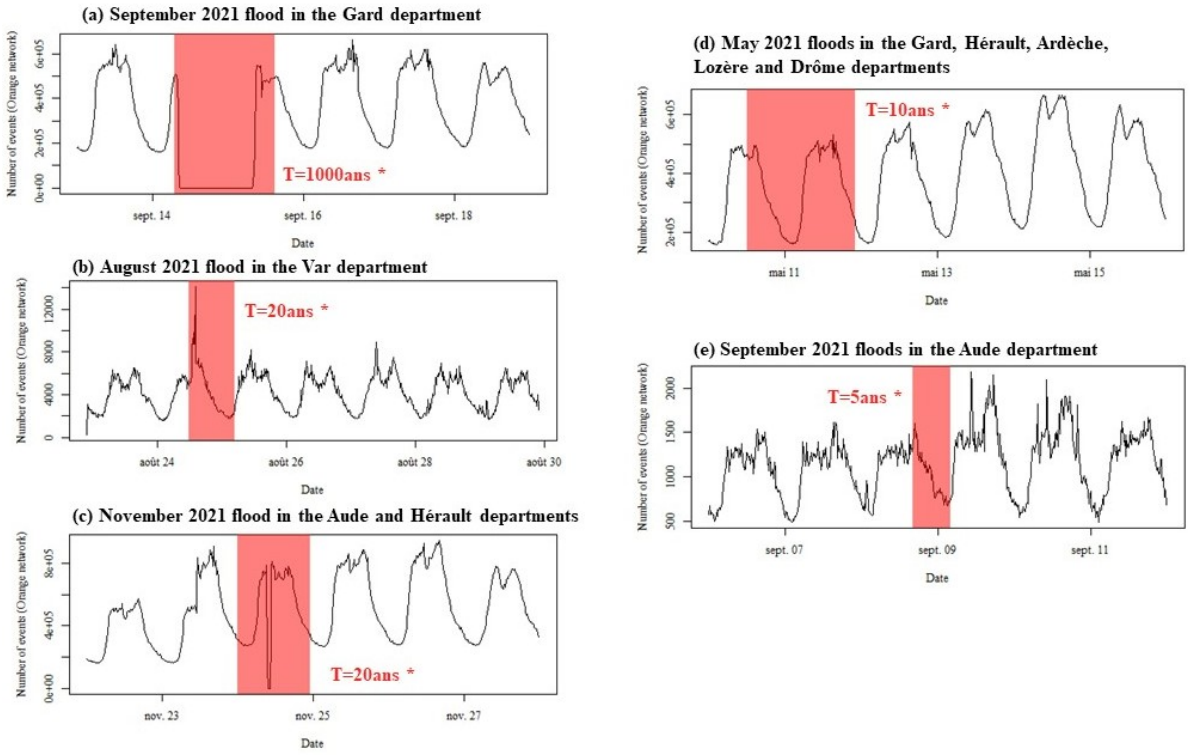


FIGURE D.1 – Telecommunication network activity during five flood events in 2021. The red rectangles indicate the duration of each flood event. The maximum return periods estimated on the affected river reaches are also shown (*return periods were estimated using the implemented warning chain).

1.5 Electricity network data

Enedis manages 95% of France's electricity distribution network, which is highly vulnerable to floods and serves as a key indicator of their impact. The company maintains a database that records the status of the network, allowing queries on power outages classified as "incidents," particularly relevant in flood events. Network management involves securing infrastructure and people from electrical hazards, especially at medium voltage / low voltage (HTA / BT) substations - small facilities in buildings or along roadways that can sustain severe flood damage, leading to material deterioration and electrocution risks. Each substation supplies a subsector where the number of affected users can be determined, a crucial metric in Enedis' database. We access data from two flood events in the French Mediterranean region (November 23-24 and December 1, 2019), including precise outage timestamps, customer impact numbers, and the locations of both preemptively shut and flood-damaged substations. However, the database does not indicate whether preemptively shut substations were later flooded, and even when not, such outages still caused service disruptions. Another limitation is the absence of data on the duration of the outage. Despite these constraints, the dataset is highly valuable for flood forecast verification due to its systematic archiving with precise times and locations. Furthermore, an open-access national dataset cataloging public HTA/BT substations and underground networks makes it possible to simulate the corresponding flood impact and compare it with the collected observations.

2 Detailed impact databases focusing on specific events

In some cases, impact data are recorded or collected only for specific events. This often occurs when field surbeys are conducted after major flood events, which are not systematically organized due to resource limitations. In addition to their lack of continuous temporal coverage, these databases typically do not provide exhaustive spatial coverage of observed impacts (e.g., field campaigns may focus on a particular affected area). However, a key advantage of these databases is that they generally offer more detailed information on the observed impacts. In tables D.2 and in the text below, we provide some insight into the databases we identified for validation of flood forecasts in the French Mediterranean area.

TABLEAU D.2 – Summary of collected impact databases focusing on specific events

Database	**Description**	**Access and Coverage**	**Impact type**	**Spatial precision**
DamaGIS	Manually compiled geodatabase of flood damages based on social media, local authorities, and field surveys	Open-access, 22 events in the study area	Multiple types (agriculture, buildings, road networks. . .)	Point data
Murex	Online platform documenting flood damages from post-flood damage assessment	Open-access, 17 events	Multiple types (Buildings, networks, vehicles...)	Point data
The High Water Marks Database	Records of flood water levels measured in past events, including geolocated high-water marks on buildings and structures.	Open-access, focus on large flood events	High Water Marks (which is not always related to a damage)	Point data
Visov	Database of flood impact reports from social media monitoring by volunteers	Access upon request, 6 events	Multiple types	Point data
MIRV database	Post-flood damage assessment focused on road infrastructure, created after Storm Alex (2020) for reconstruction planning.	Restricted data, 1 event	Road network	Point data
Flood extents	Reconstructed flood extents based on field surveys, satellite images, and high-water marks, used to estimate affected assets	Access upon request, 3 events	Assets in flood extent	Point data

2.1 DamaGIS

The DamaGIS database, developed by Saint-Martin [Saint-Martin et al., 2018], compiles flood impact data for 23 events (2011–2017), addressing gaps in existing data sets, such as poor spatio-temporal resolution, limited damage details, and incomplete event coverage. After adapting it to our study area, 22 events and 628 damage records were retained. Data sources include social networks (64% of entries), media, local authorities, public administrations, field surveys, and online platforms. Each entry is georeferenced, sourced and described, but the dataset remains incomplete due to its manual compilation without institutional support. DamaGIS is structured as a geodatabase with three vector classes (watersheds, flood events, and damages) and categorizes over 14 asset types and 38 subtypes. It also includes a four-level severity index, ranging from minor disruption to complete destruction.

2.2 Murex

The Murex database is an online platform designed to facilitate post-flood damage assessments for local authorities and stakeholders in crisis management in France. The database includes 17 events that fall within the spatial and temporal scope of this study. The Murex data are georeferenced and classified into damage reports (50.4% of all recorded impacts), high-water marks (40%), flood extent zones (0. 1%) and "other" (9. 5%). As a standardized platform, Murex provides additional subcategories, such as "Buildings," "River," "Underground network," "Vehicle" and "Transportation route" for damage reports. It also includes severity levels, such as "Cleaning necessary," "Excavated verges," "Surface layer swept away" and "Base and surface layers swept away" for transport infrastructure damage. However, it is important to note that all entries are manually recorded by field surveyors, making human error and inconsistencies unavoidable. A key limitation of Murex is that it relies on post-flood surveys, which are often incomplete and lack event timestamps. Despite this, its precise geolocation and classification of observed impacts make it a valuable resource for detailed flood forecast verification.

2.3 The High Water Marks Database

The High-Water Mark (HWM) database compiles flood level indicators recorded throughout France. These high-water marks, often engraved on buildings or marked with plaques, serve as historical references for the extent of past floods. This data set shares information with Murex, which includes both water levels and a wider range of impact data. However, the HWM database exclusively focuses on water levels, making it less directly useful for damage assessments, as not all high-water marks correspond to actual damage. Although some HWMs are recorded on buildings with visible damage, as shown in the accompanying photographs, others are measured on flood stains in trees, which can hardly be considered damage. Despite this limitation, the HWM database provides valuable complementary information that is often missing from other impact datasets : the classification of the nature of the flood as river flooding or surface flooding, which is particularly useful for validating flood forecasts that do not account for surface flooding.

2.4 Visov

Organizations are increasingly collecting real-time impact data during crises to improve temporal accuracy and capture information that might otherwise be missed post-event. The VISOV association (Volontaires Internationaux en Soutien Opérationnel Virtuel) follows this approach by leveraging digital volunteerism to monitor crises through social media. Founded in 2014, VISOV operates beyond France

and covers various emergencies, including floods, cyclones, industrial accidents, and emergency service failures. Most of its activations involve intense rain and flooding, often triggered by orange-level weather alerts. During such events, VISOV continuously monitors social networks, sharing critical information between volunteers and crisis management partners. For this study, we accessed VISOV data on six flood events : Aude, Hérault, and Tarn (Oct. 15, 2018, 53 entities), Aude, Hérault, and Pyrénées-Orientales (Oct. 22-23, 2019, 41 entities), Southeastern France (Nov. 23, 2019, 230 entities ; Dec. 1, 2019, 62 entities), Cévennes (Sept. 19, 2020, 36 entities), and Storm Alex (Oct. 2, 2020, 39 entities). All reported impacts are georeferenced, and most include timestamps. However, the data set is highly heterogeneous in impact descriptions. Some entries contain no details ; While they likely correspond to flood impacts shared on social media, their type and severity remain unknown. Others provide only a link to the original post, making them vulnerable to data loss due to deleted accounts or platform policy changes. In some cases, volunteers manually add descriptions that vary widely, such as "Power shutdown," "Flooded street," "Fallen tree on the road," or "Collapsed electricity pole." Furthermore, its reliance on human input affects both data completeness and accuracy : only impacts shared on social networks are recorded, leading to inevitable gaps in spatiotemporal coverage. In addition, volunteers manually geolocate posts, images, and videos, introducing potential inaccuracies.

2.5 An example of a unique and rich After-Event assessment : the MIRV database

The severity of certain events requires extensive post-event assessments to evaluate damages and incorporate lessons into future land use planning. The October 2020 storm Alex, which caused unprecedented destruction in the Alpes-Maritimes (France) and northern Italy, led to a large-scale impact assessment. To support reconstruction, specific missions were established, including the Interministerial Mission for the Reconstruction of Valleys (MIRV), which focused on the Roya Valley, one of the hardest hit areas. MIRV documented structural damage, called "disorders," along departmental roads. The database created for reconstruction efforts includes the precise locations of the affected structures, their level of damage, and their restoration status ("ongoing studies," "completed," or "in progress"). Although it lacks timestamps for individual damages, it provides valuable qualitative insights. Of the 237 recorded structures, 75 were destroyed and 161 damaged, classified as roads and supports (116), retaining walls (67), road surfaces (26), bridges (16) and cliff stabilization works (12). The strength of the database lies in its precise location, categorization, and impact assessment, highlighting the need for a similar detailed data collection for other flood events, not just those of exceptional magnitude.

2.6 Getting impact data from reconstructed water extents

During high-intensity flood events, observed flood extents can be reconstructed using field surveys, HWM inventories, photos, videos, and satellite images. These data sets are valuable for forecast verification, primarily at the hydraulic modeling stage, i.e., before impact modeling, according to the principle of comparing comparable data. However, when impact observations are not available, impacts can be estimated by overlaying the assets layer with the flood extent layer. For this study, we accessed flood extent data for three flood events within our study area : June 2010 in the Argens catchment, Oct. 2018 in the Aude department, and Nov.–Dec. 2019 in the Alpes-Maritimes department. Although these data sets can be useful for events with reduced impact data, they have several limitations. In addition to uncertainties in determining impacts based on flood extent (e.g., a building within a flooded area may not be damaged due to unmodeled flood defenses), the flood extent data itself may contain errors, particularly from local interpolations between field observation points.

Annexe E

Annexe de l'article Godet et al. [2023]

1 Hydrological presentation of the studied events

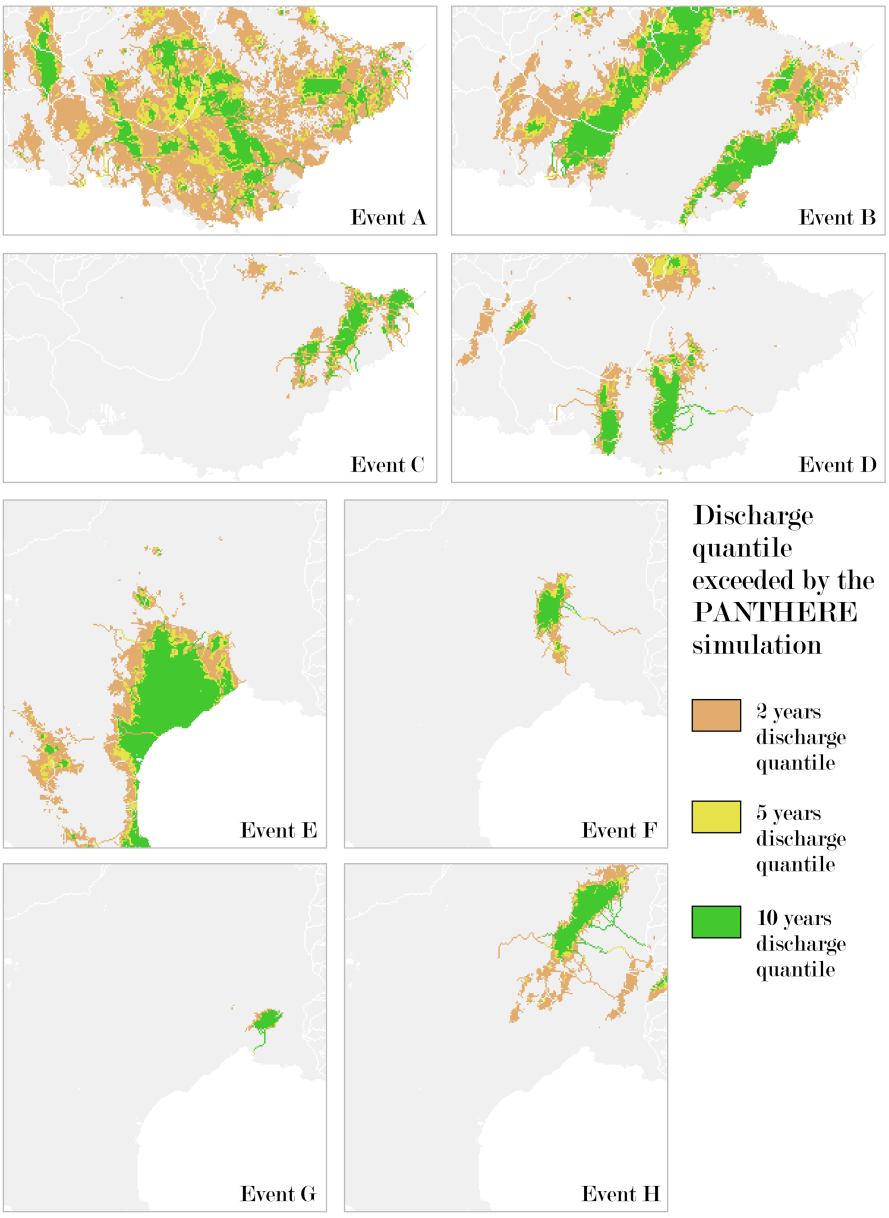


FIGURE E.1 – Discharge quantiles exceeded by the PANTHERE simulation Q_{sim} during each event

2 Methodology for filling the contingency tables

The contingency tables are filled for each pixel of the zones. For an ensemble forecast product, each percentile is considered separately, to be treated as a deterministic product. For each pixel, we look at the hydrograph $Q_{sim}(t)$ simulated by SMASH with the PANTHERE QPE as input, and at the hydrographs $Q_{for}(t)$ forecasted by SMASH when forced with a QPF product (constant rain scenario, AROME-NWC, PIAF or PIAF-EPS percentiles). Let t_{run} be the forecast start time, let T be the forecast range ($0 < T \leq 3h$), and let Q_t be the considered discharge threshold.

2.1 Detailed method of Charpentier-Noyer et al. [2023] : observation-based stratification

The method, applied on each spatial entity and for each forecast probability, is as follows :

— If there exists t such that $Q_{sim}(t) \geq Q_t$, then the date t_{sim} corresponding to the first threshold exceedance by Q_{sim} is selected. A sample S consisting of all the pairs (t_{run}, T) such that $t_{run} < t_{sim} \leq t_{run} + T$ is constructed.

  — If there exists $(t_{run}, T) \in S$ such that $Q_{for}(t_{run}, T) \geq Q_t$, a HIT is counted. The t_{run} corresponding to the first threshold exceedance is selected to calculate the anticipation time : $t_{sim} - t_{run}$.

  — If $Q_{for}(t_{run}, T) < Q_t \quad \forall (t_{run}, T) \in S$, a MISS is counted.

— If $Q_{sim}(t) < Q_t \quad \forall t$, the date t_{sim} corresponding to the global maximum of Q_{sim} is selected and a sample S consisting of all the pairs (t_{run}, T) such that $t_{run} < t_{sim} \leq t_{run} + T$ is constructed.

  — If there is $(t_{run}, T) \in S$ such that $Q_{for}(t_{run}, T) \geq Q_t$, a FALSE ALARM is counted.

  — If $Q_{for}(t_{run}, T) < Q_t \quad \forall (t_{run}, T) \in E$, a CORRECT REJECTION is counted.

2.2 Adapted method : forecast-based stratification

The detailed process used to build the contingency table is as follows :

— If there exists (t_{run}, T) such that $Q_{for}(t_{run}, T) \geq Q_t$, the pair (t_{run}, T) corresponding to the first threshold exceedance by Q_{for} is selected.

  — If $Q_{sim}(t) < Q_t \quad \forall t$, a FALSE ALARM is counted. See figure E.2a.

  — If there is t such that $Q_{sim}(t) \geq Q_t$, the first threshold exceedance occurring at t_{sim} :

    — If $t_{sim} \in]t_{run}; t_{run} + T + 3h]$, a HIT is counted. The anticipation time is equal to $t_{sim} - t_{run}$. The time interval $]t_{run}; t_{run} + T + 3h]$, and particularly the $+3h$ part, was chosen in order to create a tolerant window for the HIT counting. It seemed coherent to use a time tolerance equal to the maximum lead time. See figure E.2b.

    — If $t_{sim} > t_{run} + T + 3h$, a FALSE ALARM is counted. See figure E.2c.

    — If $t_{sim} \leq t_{run}$, a MISS is counted. See figure E.2d.

— If $Q_{for}(t_{run}, T) < Q_t \quad \forall (t_{run}, T)$:

  — If $Q_{sim}(t) < Q_t \quad \forall t$, a CORRECT REJECTION is counted. See figure E.2e.

  — If there exists t such that $Q_{sim}(t) \geq Q_t$, the first threshold exceedance occurring at t_{sim}, a MISS is counted. See figure E.2f.

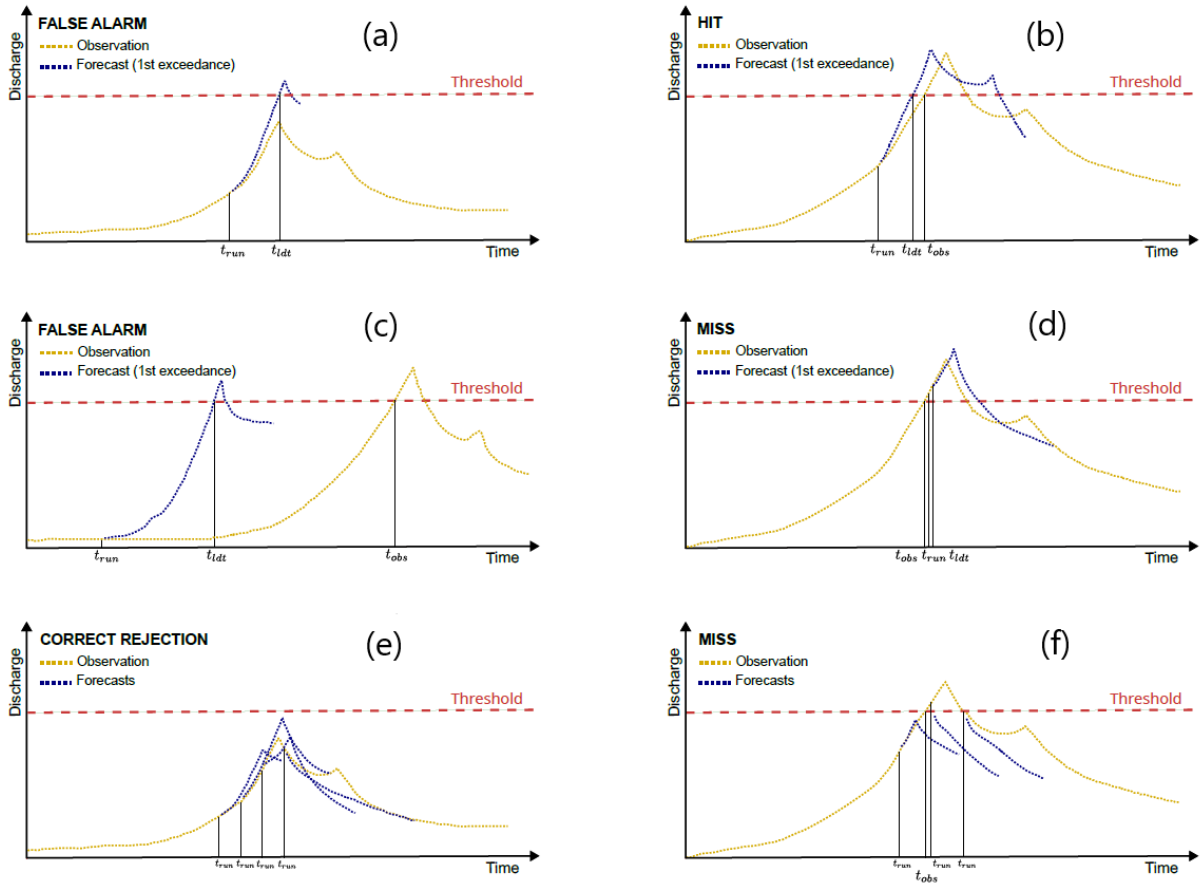


FIGURE E.2 – Six possible cases in the new methodology (inspired by Charpentier-Noyer et al., 2023) : (a) forecasted threshold exceedance not present in the simulated hydrograph (False Alarm), (b) threshold exceedance correctly forecasted (Hit), (c) threshold exceedance anticipated but anticipation largely exceeding the forecast lead time (false alarm), (d) threshold exceedance detected by one forecast, but right after the simulation (miss), (e) absence of threshold exceedance both in the simulation the forecasts (correct rejection), (f) threshold exceedance undetected by all the forecasts (miss)

3 Rank diagrams

Rank diagrams, also called Talagrand diagrams [Candille and Talagrand, 2005], are one of the most common tools for assessing the reliability of meteorological ensemble forecasts. The general idea of this tool is to count the number of times the observation value is included in a given interval of the ensemble forecast quantiles. As a consequence, if the observation value is often close to low quantiles, it means that the forecast model has a tendency to overestimate. On the other hand, if the observation is more often close to high quantiles, then the model tends to underestimate. Logically, a perfect diagram would be perfectly flat, which would mean that the observation is uniformly distributed among the ensemble forecast quantiles. However this never happens in reality. The rank diagram is useful to quickly detect biases in an ensemble forecast. It can detect positive or negative biases, but also under and over dispersion of the ensemble forecasts. Traditionally, the rank diagram is applied to rainfall ensemble forecasts. However, in this study it was applied to discharge ensemble forecasts, at each pixel of the SMASH model computation grid. The rank diagrams presented here combine all the forecasts issued during each event, and are computed for a fixed lead time (one hour).

In order to distinguish the roles of high and low discharges in the rank diagram form, it was decided to build separate rank diagrams for each category. However it is necessary to take precautions concerning the criteria that distinguish those categories. Indeed, Bellier et al. [2017] have shown that a sample stratification based on the observations can introduce bias. A sample stratification based on forecasts is recommended in most of the cases. Therefore, the following categories were chosen :

— $Q_{med} \leq \frac{1}{2} Q_{T=2y}$ for low discharges.

— $\frac{1}{2} Q_{T=2y} < Q_{med} \leq Q_{T=2y}$ for medium discharges.

— $Q_{med} > Q_{T=2y}$ for high discharges.

Where Q_{med} is the median discharge of the hydrological ensemble forecasts and $Q_{T=2y}$ is the 2-year return period quantile, according to the historical run of the SMASH model.

Note that even if based on forecast discharges, this stratification can still cause bias : when only the areas and time steps with high forecast discharges are considered, the overall probability that the considered forecasts exceed the observed discharges tends logically to be higher, and conversely. However, this stratification effect does not affect the global rank diagrams.

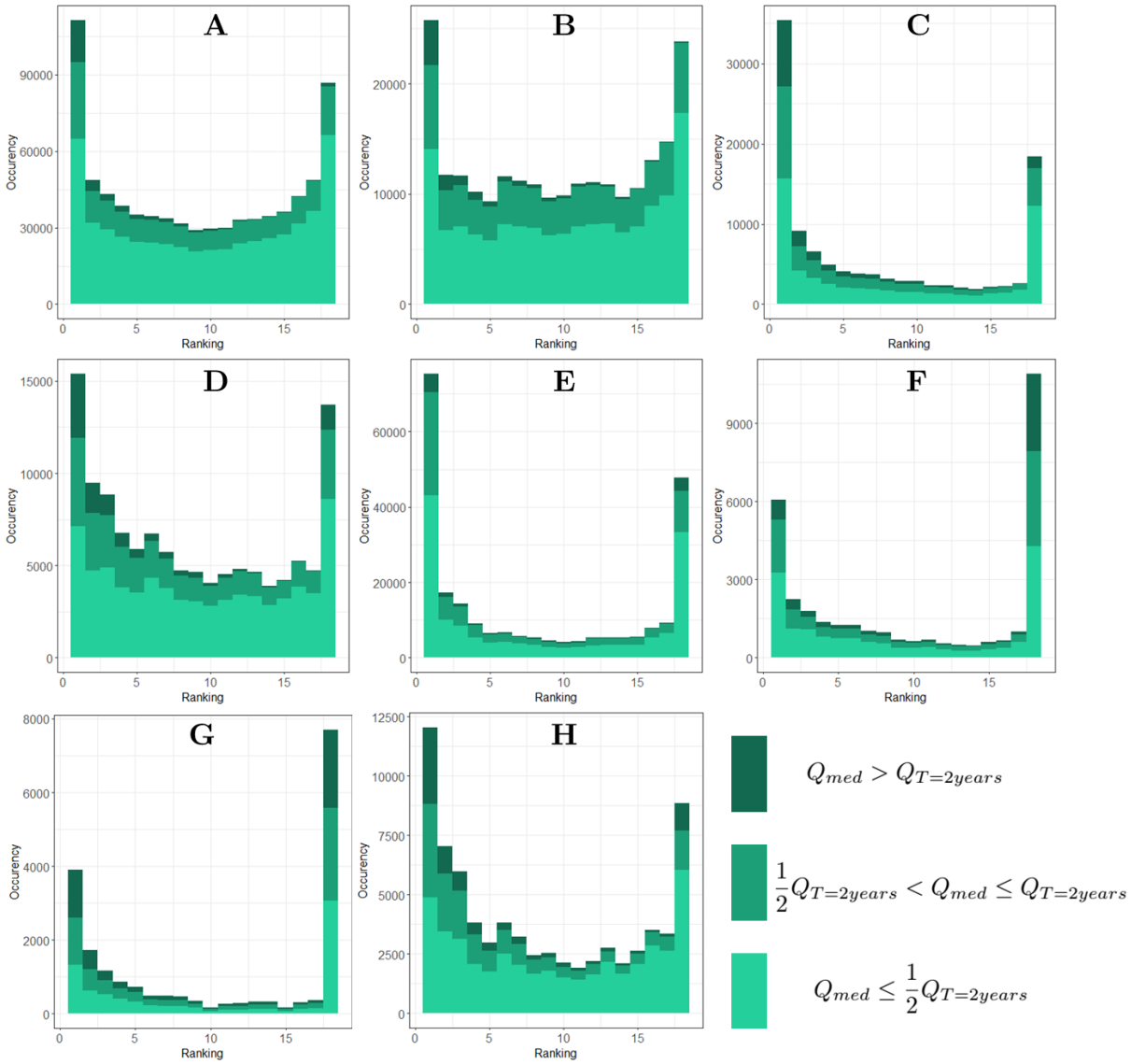


FIGURE E.3 – Rank diagrams obtained for each event, 1hour-lead-time

4 Maps of contingency tables for each event

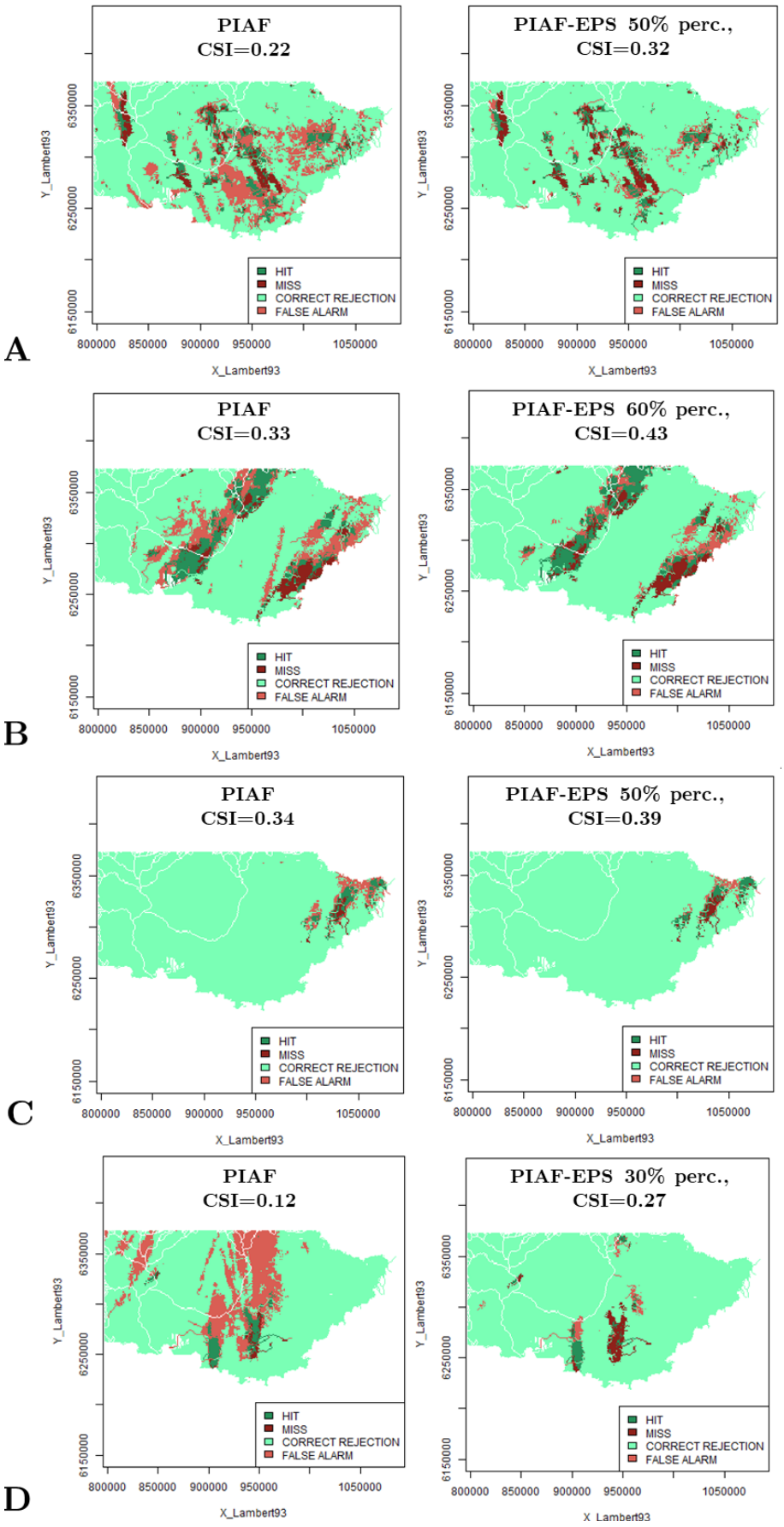


FIGURE E.4 – Best deterministic forecast *(left)* and best PIAF-EPS percentile *(right)* in terms of CSI, for each event, $Q_t = Q_{T=10y}$

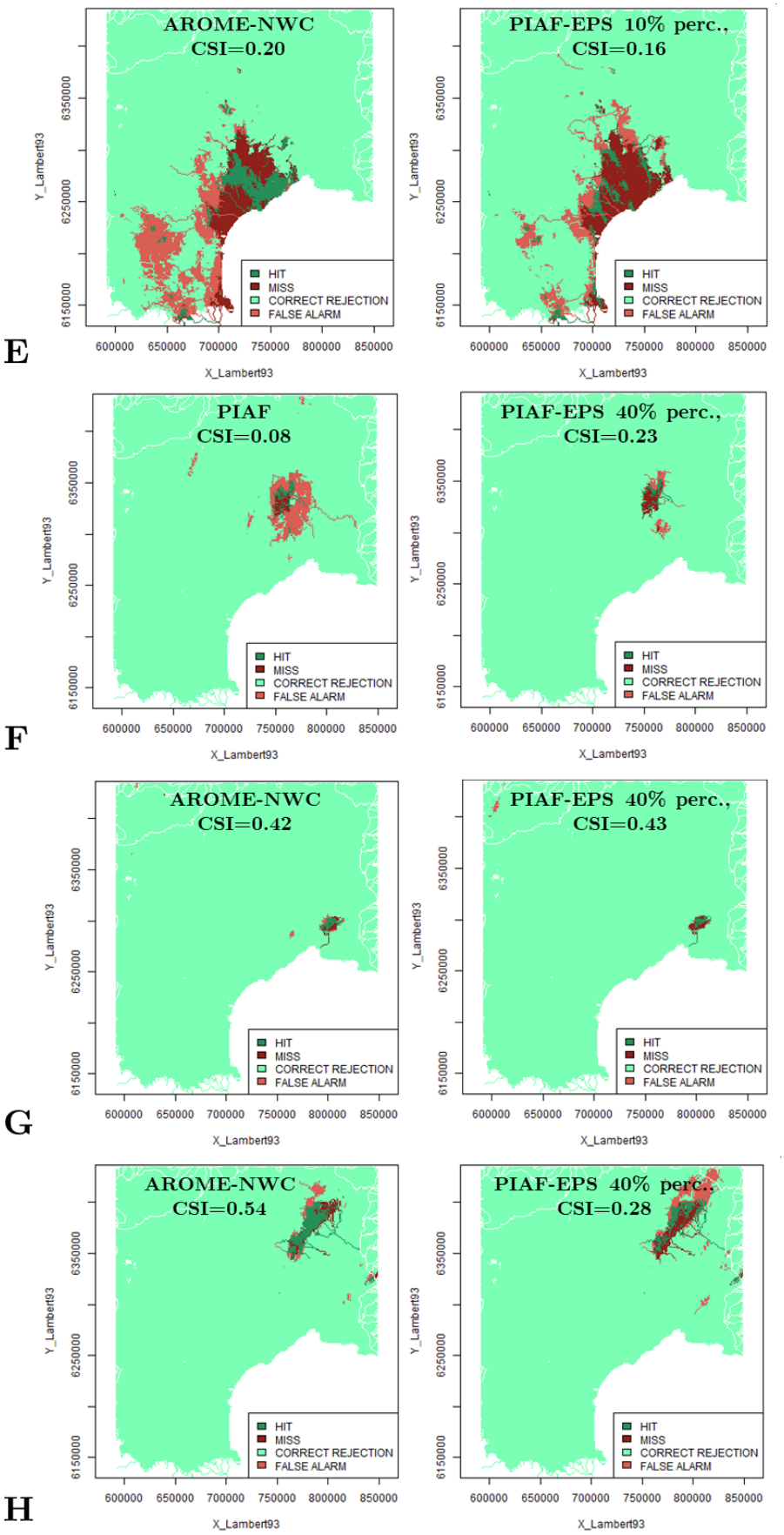


FIGURE E.4 – Best deterministic forecast *(left)* and best PIAF-EPS percentile *(right)* in terms of CSI, for each event, $Q_t = Q_{T=10y}$

Annexe F

Séjour de recherche en Nouvelle-Zélande

Dans le cadre de cette thèse, un séjour de recherche de deux mois a été réalisé au sein de NIWA en Nouvelle-Zélande (National Institute of Water and Atmospheric Research, devenu Earth Sciences New Zealand au 1er Juillet 2025). Ce séjour visait à confronter les approches développées pour la prévision des impacts des crues soudaines à l'échelle de l'arc méditerranéen français, en se basant sur le système opérationnel Vigicrues Flash, à des méthodes similaires mises en œuvre à l'échelle d'une ville en Nouvelle-Zélande. L'objectif était d'explorer la mise en place de prévisions d'impacts dans un contexte géographique et opérationnel différent, en mobilisant les outils de NIWA et les bases de données locales. Le rapport rédigé en anglais à l'issue de ce séjour est directement repris dans la suite de cette annexe.

1 Study case and forecast data

At the beginning of the internship, several flood events were considered as potential case studies, including :

- A series of flood events that affected the town of Westport in 2021 and 2022. This included the July 2021 flood (which caused severe damage), the February 2022 flood (which caused light damage), and the August 2022 flood (which caused no significant damage). NIWA had previously undertaken a project in Westport aimed at implementing a real-time flood model capable of producing inundation maps. As a result, reforecasts of inundation are available for all three events.
- Cyclone Gabrielle in February 2023, a major extreme weather event that struck the North Island of Aotearoa New Zealand, with an estimated total cost of NZD 14.5 billion.

However, due to time constraints and the limited availability of relevant data, the internship ultimately focused on the July 2021 Westport flood event. The following sections present an overview of the event and describe the data provided by NIWA for this study.

1.1 July 2021 Westport flood event

Brief description of the event, *extracted from Johnson et al. [2025]*
"The July 2021 flood in Westport lasted 72 hours from 15th to 18th July. The event was characterized by heavy rainfall and a rapid rise in river levels that caused the Buller River to breach its banks. This caused significant flooding across the Buller District and necessitated the evacuation of 826 properties and over 2,064 people in Westport. Bridges were closed, cutting access to Westport. The flood caused substantial damage, with 71 homes deemed unsafe for habitation and 388 homes requiring assessment for damage. A total of NZ$97.2 million in insurance claims was paid to Westport residents."

Meteorological drivers, *extracted from Cattoën et al. [2022]*
"A pulse of tropical air originating over the Indian Ocean combined with an upper-level trough and front delivered a humid and gusty northerly flow to the central and northern South Island. Within 72 h, rainfall stations in the Southern Alps recorded 697 mm at Ivory Glacier, 413 mm at Arthurs Pass"

Hydrological reaction, *extracted from NIWA [2021] and Johnson et al. [2025] and Cattoën et al. [2022]*

- *"The water level at Te Kuha peaked at 12.8 metres at 1.35pm on Saturday, July 17, several hours ahead of the peak flooding in Westport. In the 1970 flood the flood peak was 11.9 metres. This was the biggest flood recorded since the Te Kuha monitoring site was established in 1963."*
- *"We can confidently say that the recent flood flows in the Buller River are the highest of any river in New Zealand since 1926."*
- *"The predicted peak flow of the probabilistic forecasts (median of 8,564 m3/s on 17th July) was close to the actual peak flow (8,890 m3/s)."*
- *"As bridges closed, the NIWA field team measured 7640 m3/s (for a catchment of 6350 km2 and 454 m^3/s mean flow)—the largest direct streamflow measurement in NZ"*



FIGURE F.1 – Westport, July 2021. ©NZ Defence Force

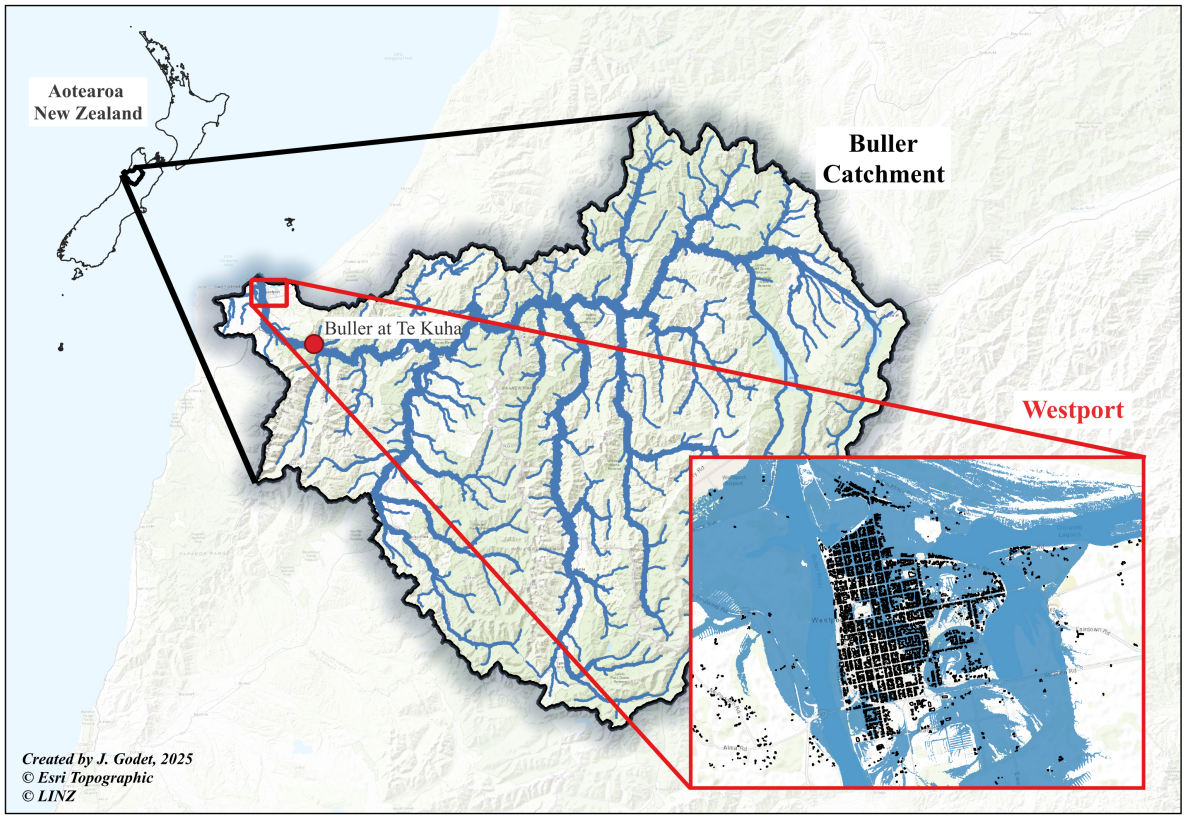


FIGURE F.2 – Presentation of the Buller catchment and Westport, with July 2021 reconstructed flood inundation map (pixels with water depths > 5cm are coloured in blue)

Comparison with Mediterranean Flash Floods

The July 2021 flood event in Westport is markedly different from the Mediterranean flash floods studied in the context of this PhD in France.

— The duration differs : this was not a flash flood with a rapid rise and recession within 24 hours, but rather a prolonged event that lasted several days.

— The flood mechanisms also differ : although this event was primarily riverine, coastal processes play a significant role in Westport's flood dynamics.

— However, some meteorological drivers are comparable : the Buller catchment features an even greater elevation range than typical Mediterranean catchments, and both regions are primarily fed by orographic rainfall.

— The exceptional impacts recorded during this event resonate with the high-impact floods frequently observed in the French Mediterranean region.

1.2 Rainfall Forecasts

In this study we used the NZens ensemble rainfall product, which consists of 18 members with a run length of 5 days. Forecasts are issued twice daily (at the time of the 2021 event; however more recent versions support up to four runs per day). The gridded rainfall resolution is 4.4 km. NZens is based on the RAL1-M regional numerical weather prediction model developed in the UK [Bush et al., 2020].

NZens is a dynamical, convective-scale ensemble model and differs from NZCSM (New Zealand Convective Scale Model), which is used in NIWA's operational flood forecasting system, AFAS [Cattoën et al., 2022]. Unlike NZens, NZCSM is a deterministic model that has been statistically perturbed in terms of rainfall, soil moisture, and baseflow to account for uncertainty in streamflow forecasts.

Rainfall Radar : France vs. New Zealand

Rainfall forecasts used for flood prediction in New Zealand rely on global or regional numerical models. Indeed, radar coverage differs significantly between France and New Zealand. In New Zealand—particularly on the West Coast of the South Island—gaps in radar coverage make it challenging to obtain reliable rainfall observations, especially at high spatial and temporal resolutions.

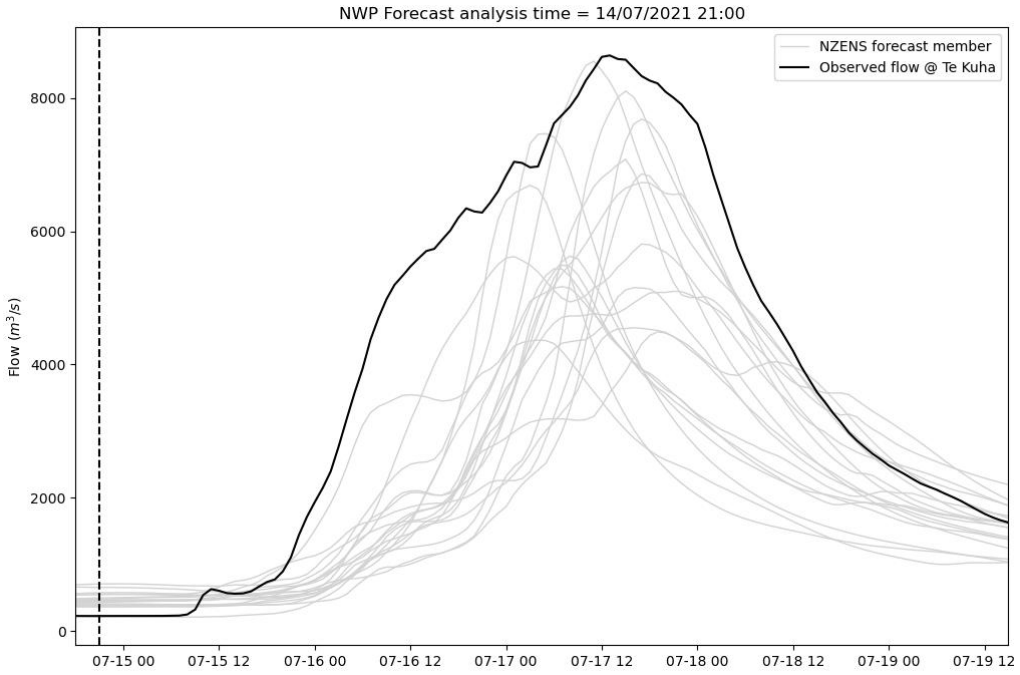
Lead Times – Nowcasting vs. Forecasting

As previously mentioned, NZens forecasts have a run length of 5 days. This contrasts sharply with the ensemble rainfall forecasts used in Vigicrues Flash in earlier work [Godet et al., 2023], where the maximum lead time was just 3 hours. In this context, the term "forecasting" is more appropriate, while "nowcasting" better describes the approach used in France.

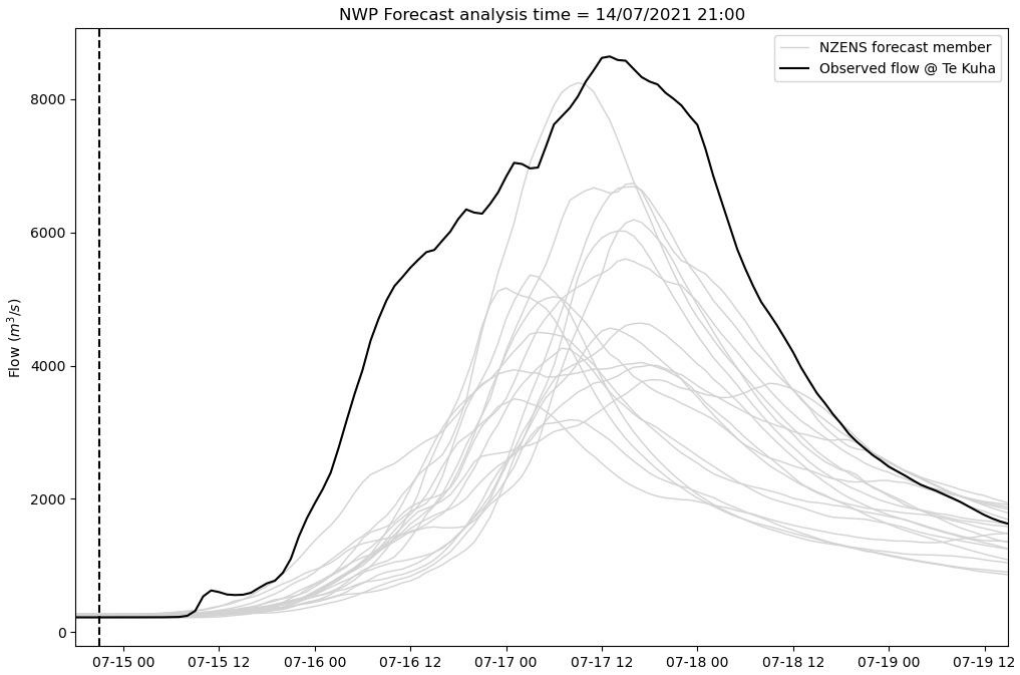
For the validation of the rainfall forecasts, catchment-averaged rainfall was computed for each time step using the "best" gridded rainfall product, which is based on observations. This variable, referred to as *aprecip*, will be used in the validation process as described in later sections.

1.3 Flow Forecasts

Flow forecasts were produced using the TopNet hydrological model [Clark et al., 2008]. TopNet is a continuous, semi-distributed, physically-based hydrological model that operates on an hourly time step. It combines a water balance model applied to each subcatchment with a streamflow routing component using the kinematic wave model. TopNet is used operationally within NIWA's national flood forecasting system, AFAS, as described in [Cattoën et al., 2022]. It has been developed internally by NIWA and is not currently open source.



(a) Without river data assimilation



(b) With river data assimilation

FIGURE F.3 – Different versions of provided forecasts - with and without river data assimilation

TopNet vs. CINECAR

At first glance, TopNet may appear similar to the CINECAR model developed by the Université Gustave Eiffel team [Gaume et al., 2004, Naulin et al., 2013]. However, CINECAR is originally an event-based model, although recent research has begun to explore coupling its vectorial routing module with the continuous SMASH production module [Colleoni and Huynh, 2024].

Initially, flow forecasts were generated using an older version of TopNet, as it was immediately available. Updated forecasts using the newer Strahler 2 version of TopNet were obtained two weeks after the start of the internship. Two versions were provided : with and without river data assimilation

(see Figure F.3). Both versions significantly underestimated river discharge, though the underestimation was less pronounced in the version without data assimilation. Therefore, this version is the focus of the remainder of the analysis.

We focus on the flow forecasts at the Buller River at the Te Kuha station, which is the closest hydrometric station to the town of Westport. Using observed discharge data, the TopNet model was calibrated on the July 2021 event, allowing for the reconstruction of an observed hydrograph. This hydrograph, shown in Figure F.3, serves as a reference and will be used for validation purposes in the subsequent sections of this report.

Similar validation data for France

This compares with the hydrographs reconstructed with Cinecar for the Aude2018, AlpesMaritimes2015 and Argens2010 events that are provided in Godet et al. [2025a].

1.4 Inundation forecasts

1.4.1 Flood library

As mentioned earlier, NIWA has developed a flood model for Westport. It is based on a flood library of approximately one hundred scenarios, each representing a unique combination of two parameters considered to be the main drivers of flooding in the area :

- Peak discharge at the Te Kuha station.
- Peak sea level.

All the scenarios in this 2D parameter space are shown in Figure F.4.

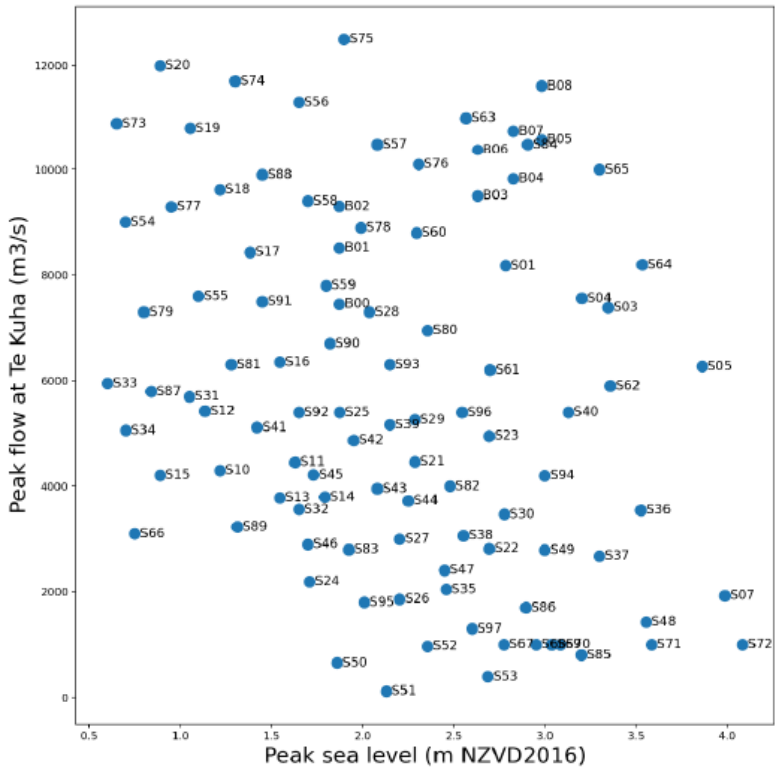


FIGURE F.4 – Precomputed flood library ©C. Cattoën-Gilbert

For research purposes, the NIWA team investigated how to construct a reduced flood library under resource constraints—e.g., if only 15 scenarios could be implemented. They compared different selection

methods, including an "Holistic" approach, a “Maximum Dissimilarity / Uniform” approach, and a “Random” approach. A scientific paper detailing this work will be submitted soon by the NIWA team.

Toward a Mediterranean-Equivalent Flood Library

One of the approaches explored in the Westport study is a “Design Library” method, consisting of 15 scenarios under different climate change conditions (present, RCP 4.5, 6.0, and 8.5), various flow return periods at Te Kuha (20, 50, and 100 years), and different peak sea levels. However, since the July 2021 flood was primarily driven by river discharge, and to enable comparison with the French case study, it is possible to consider only the 1D peak flow dimension. This results in a simplified library of 9 scenarios, summarized in Table F.1.

	RP = 20y	RP = 50y	RP = 100y
present conditions	7451	8514	9304
RCP 4.5		9493	10383
RCP 6		9818	10738
RCP 8.5		10564	11582

TABLEAU F.1 – Flow values (m^3/s) of Te Kuha in the 1D Design flood library

In the French flood library, eight monofrequency scenarios were used, corresponding to the 2-, 5-, 10-, 20-, 50-, 100-, 500-, and 1000-year return periods of peak discharge.

However, to remain consistent with the model delivered to the Westport stakeholders, the results presented in this report are based on the full 100-scenario flood library, with flood maps interpolated using the TIN method. This point will be revisited in Section 7.

Hydraulics Computations

For both the Mediterranean and Westport flood libraries, hydraulic simulations were carried out using a 2D model based on the Saint Venant shallow water equations. However, while the Westport model was run in unsteady mode, using full hydrographs to capture the temporal dynamics of the flood, the Mediterranean simulations were conducted under steady-state assumptions, using only the peak discharge associated with each return period scenario.

1.4.2 Interpolation method

In order to reconstruct flood extents in real time based on the flood library, NIWA has explored several interpolation methods :

- Nearest-neighbour : selecting the closest scenario in the 2D parameter space without interpolation.
- TIN (Triangular Irregular Network) : a 2D linear interpolation method that divides the standardized parameter space into triangles. The forecast lies within one of these triangles, and interpolation is performed based on the values at the three vertices.
- AI-based interpolation : using a Random Forest algorithm.

According to comparison results, which will be discussed in NIWA’s forthcoming publication, the spread of differences between methods narrows as the size of the flood library increases. Furthermore, the AI-based method does not significantly outperform the simpler interpolation approaches.

Comparison with the flood map reconstruction method used in France

In the PhD work focused on Mediterranean catchments, flood maps are reconstructed in real time by selecting the closest scenario for each subcatchment—i.e., no interpolation.

Given that the average size of a Mediterranean subcatchment in this study is 4.2 km^2, which is comparable to the size of Westport's city center, the corresponding method would be equivalent to a 1D nearest-neighbour approach, relying solely on the peak discharge at Te Kuha.

1.4.3 Reference flood map

For the July 2021 Westport event, the NIWA team has reconstructed a water level flood map based on collected high water marks, local knowledge and hydraulics simulations. It is shown in figure F.5.

Similar validation data for France

This compares with the reconstructed flood extents and collected High Water Marks for the Aude2018, AlpesMaritimes2015 and Argens2010 events that are provided in Godet et al. [2025a].

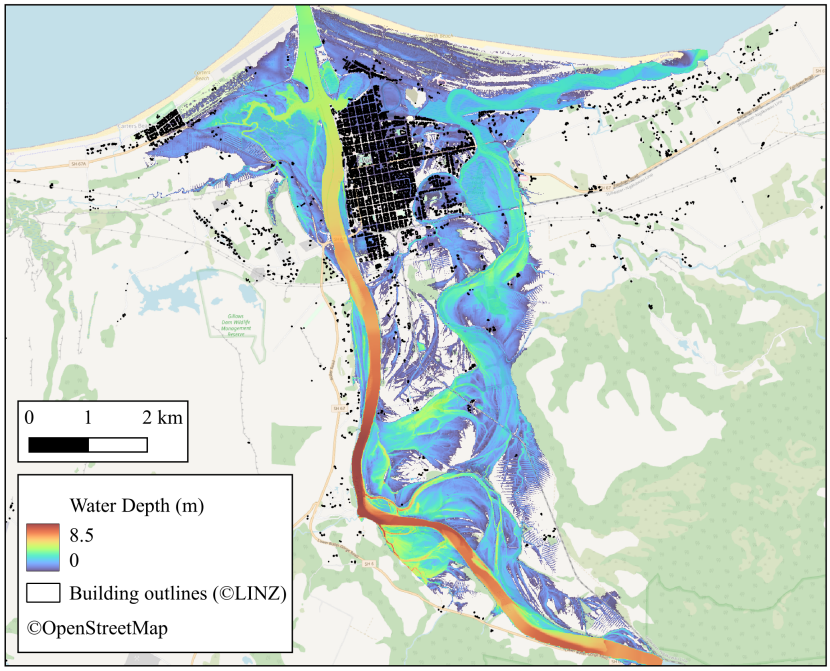


FIGURE F.5 – Reconstructed flood map

2 Collecting impact observations

The objective of this work is to conduct a block-by-block validation of the flood forecasting chain. The previous sections presented the simulation outputs and validation data available for the existing components of the chain : rainfall, streamflow, and inundation. The next step is to integrate an impact component, which poses the significant challenge of obtaining reliable impact observations [Harrison et al., 2021, 2022a,b, 2024]. This section outlines the various datasets that were considered for this purpose.

From broad-scale to high-resolution

In France, the objective was to collect flood impact data over a broad spatio-temporal scale—covering the Mediterranean region over a continuous 10-year period. While a similar approach would be ideal for New Zealand, particularly to support future research, the present study realistically focuses on the July 2021 flood in Westport. This means that the analysis is limited to a single event and a single city. However, this narrower scope allows for a more detailed investigation than what was done in France, particularly in terms of spatial resolution.

2.1 Number of claims from ICNZ

The Insurance Council of New Zealand (ICNZ) has been documenting the economic costs of natural disaster events since 1968. These records detail the costs borne by the insurance industry in settling claims for damage resulting from natural hazards such as floods, storms, wind, and earthquakes. They can be found at https://www.icnz.org.nz/industry/cost-of-natural-disasters/.

The information provided by ICNZ for the "West Coast Flooding" event, which occurred between July 16 and July 19, 2021, is presented in Table F.2. Given that the majority of the impacts were concentrated in the town of Westport, it is reasonable to assume that the reconstructed flood map shown in Figure F.5 encompasses the full extent of the insured losses reported in Table F.2.

Inflation-adjusted cost (2023) : 110.2 million NZD

Claim Type	**Number of Claims**	**Insured Losses (NZD)**
House & Contents	1,368	87,921,014
Commercial	182	5,882,377
Marine	2	51,823
Motor	307	3,130,708
Crops	0	0
Other	5	239,763
Total	**1,864**	**97,225,685**

TABLEAU F.2 – Number of claims and insured losses by claim type for the Westport July 2021 flood.

As of today, the ICNZ database includes 229 events. However, no information was found on the criteria used to include an event in the database (e.g., cost threshold).

French datasets with comparable content and resolution

Similar data sources available in France include :

- For detailed claim counts : the iCatNat database, which provides estimates of affected buildings at the municipality level.
- For insured loss estimates : the event library of the Caisse Centrale de Réassurance (CCR), the public reinsurer offering state-backed coverage for natural disasters. Their portal (https://catastrophes-naturelles.ccr.fr/les-evenements) provides cost ranges for events with insured losses exceeding €10 million.

Similar data, although not publicly available, was provided by CCR specifically for the purposes of this PhD. It consists of a ranking of affected river reaches (with an average size of 4km², thus

offering a higher resolution than municipality- or event-scale data) based on the number of insurance claims per flood event. However, CCR did not provide the exact number of claims nor the insured loss amounts.

2.2 Placard data

In New Zealand, residential properties affected by a natural disaster are assessed using a placard system managed under the Building Act. This system classifies building usability based on observed damage, with three categories : white placards (safe to occupy), yellow placards (restricted use), and red placards (unsafe to enter).

Two datasets were provided by the West Coast Regional Council :

- One corresponding to the initial inspections carried out by Urban Search and Rescue (USAR), which involved surveying all dwellings—including both affected and unaffected ones—and assigning a damage category (see Table F.3).
- The other corresponding to the final placard assessments conducted by MBIE (Ministry of Business, Innovation and Employment). These assessments focused on the subset of buildings identified by USAR as affected. In addition to placard assignment, MBIE assessors also recorded a damage category (see Table F.3).

While both datasets initially contained a large number of attributes (which differed between the two), for the purpose of this study, we retained the damage category, as it was the most relevant variable and the only one common to both datasets. This enabled us to merge them into a single dataset, summarized in Table F.3 and spatially represented in Figure F.6.

damage	source	building_id	DamageNb
NodamageUnaffected	USAR	2513599	0
Slightdamage	USAR	2513425	1
Moderatedamage	USAR	2513602	2
Severedamage	USAR	2513595	3
TotaldamageDestroyed	USAR	2513358	3
light	MBIE	2513596	1
moderate	MBIE	2513597	2
heavy	MBIE	2513604	3

TABLEAU F.3 – Illustration of the placard dataset and of all possible cases in the merging process, resulting in the "DamageNb" column.

Figure F.6 shows that not all buildings in Westport are included in the placard dataset. In some cases, this is because the dataset focuses specifically on residential buildings, and some excluded structures are non-residential. However, manual checks using Google Maps suggest that some excluded buildings do appear to be residential. This indicates that the placard dataset is not fully exhaustive. As a precaution, the analysis comparing simulated impacts to this validation dataset will be restricted to the subset of Westport buildings that are covered by the placard data.

Figure F.6 also illustrates what the dataset looks like when aggregated at a “block” scale (manually defined areas regrouping, on average, 15 buildings). In this example, aggregation was performed by retaining the maximum damage severity observed within each block. Alternatively, aggregation could be based on summing the DamageNb column from Table F.3 within each block. This would produce a Damage Score that captures both the severity and the number of affected buildings, providing a richer (but more complex) representation of the impact.

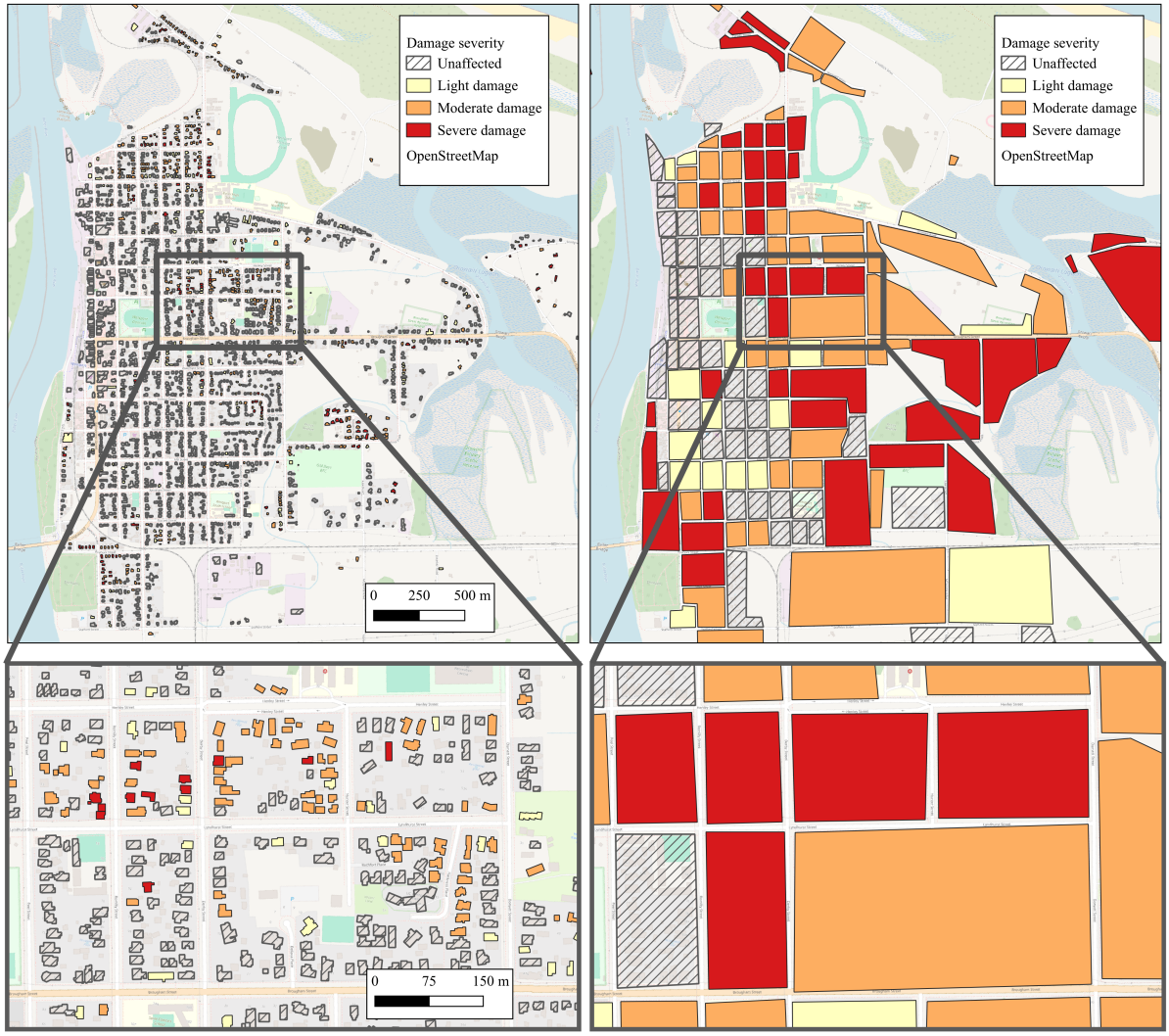


FIGURE F.6 – Damage severity at the building and block scales according to the placard data for 2021 July Westport event.

Such aggregation is useful, as expecting a flood forecasting chain to deliver accurate damage predictions at the building level is extremely demanding—and arguably unrealistic. This limitation will be further discussed later in the report.

French datasets with comparable content and resolution

In France, no dataset exists that documents residential building damage with a similar level of geolocation accuracy, exhaustiveness, and damage detail. The closest available proxy is the database of fire service operations, which is highly detailed (time-stamped, precisely geolocated). However, it has several limitations :

- Not all recorded operations are related to building damage.
- The database does not include damage assessments.
- A single operation recorded at one location may actually involve multiple displacements and multiple buildings.

2.3 Other leads

Validation results based on the two aforementioned impact datasets only will be presented in this report. This section presents other datasets that were collected (but not used for various reasons) and datasets that were considered but could not have been collected.

2.3.1 Building impact survey

On August 10–12, 2021, 24 days after the July flood event, NIWA scientists conducted on-site residential building damage assessments on 247 buildings (see Figure F.7). Their objective was to collect impact data to develop an AI-based model for damage prediction [Paulik et al., 2024]. To achieve this, they gathered information on flood hazard characteristics (e.g., water depth, flow velocity, debris) and building attributes (e.g., dwelling type, structural frame, floor height, foundation type, construction period, number of storeys, wall cladding).

This dataset was collected early during the internship timeline and was initially used to assess whether the forecasting chain could anticipate the buildings that were damaged as recorded in the dataset. However, its main limitation was its lack of exhaustiveness, which made it impossible to conduct a rigorous evaluation that would not only verify correct detection of impacts but also account for false alarms (i.e., predicted impacts where none actually occurred). As a result, this dataset was abandoned in favour of the placard dataset.

French datasets with comparable content and resolution

Several datasets sharing the same benefits (high level of detail, precise geolocation) and drawbacks (lack of exhaustiveness) have been collected for various flood events in the French Mediterranean region :

- **DamaGIS** : Manually compiled geodatabase of flood damages based on social media, local authorities, and field surveys.
- **Murex** : Online platform documenting flood damages through post-flood damage assessments.
- **Visov** : Database of flood impact reports collected through social media monitoring by volunteers.
- **Geolocated pictures** : Photo library capturing impacts such as damaged buildings, flooded vehicles, and devastated landscapes.

However, one difference is that these datasets do not focus exclusively on building damage, but also cover other types of impacts.

2.3.2 Natural Hazard Commission

The Natural Hazards Portal, and particularly its claims map, is a highly valuable data source for natural hazard impacts. It provides data on insurance claims for residential property damage (buildings and land) caused by natural hazards in Aotearoa New Zealand. These claims are lodged with the Natural Hazards Commission (NHC), which provides insurance cover for homeowners whose properties are damaged by a natural hazard, provided their home insurance policy includes fire insurance (which most policies do).

However, unlike hazards such as landslides, earthquakes, tsunamis, and volcanic activity, the Natural Hazards Commission's cover for storms and floods is limited to certain areas of land under and around the

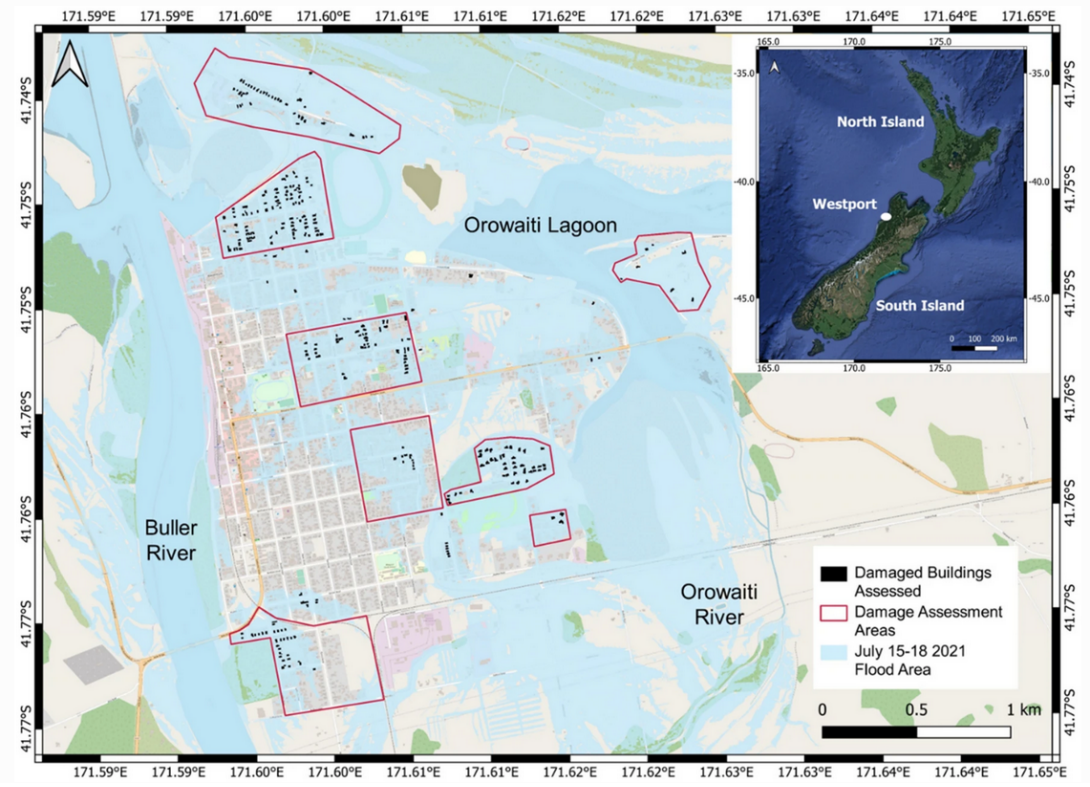


FIGURE F.7 – Residential building damage assessment locations in Westport for the 15–18 July 2021 flood event [Paulik et al., 2024].

insured home — land cover only. Any damage to the home itself must be covered through the homeowner's private insurance policy.

Specifically, in the case of floods :

- NHC provides some cover for damage to land :
  - under the home,
  - under insured separate buildings and structures,
  - within 8 metres of the home or insured separate buildings and structures,
  - under or supporting the main accessway to the home (e.g., a driveway), up to 60 metres from the home.
- NHC provides some cover for bridges and culverts :
  - located on the land within 8 metres of the home or insured structures, or
  - located under or supporting the main accessway, up to 60 metres from the home.
- NHC provides some cover for retaining walls :
  - that support or protect the home, insured structures, or insured land, and
  - are located within 60 metres of the home.

Because the Natural Hazards Commission (NHC) only provides cover for land damage in the case of floods, this database — although rich and exhaustive — is particularly tricky to use. On the one hand, it could be appropriate for validating forecasted building impacts, since in many cases, flooded properties also experience eligible land damage (e.g., debris deposits), which results in a claim. However, this is not systematic : a building can be affected without the associated land qualifying for cover, and homeowners may also be unaware of their eligibility and therefore not file a claim.

Use of insurance claim data for impact model validation

In France, we have been able to collect insurance claim data related to building damage. We have then based our validation on the assumption that for buildings where the water depth exceeded a certain threshold, there would be a claim, although this is not necessarily true, and for various reasons (e.g. deductible effect, precautionary measures, floor heights...).

The NHC database documents claims related to land damage only. Land damage is less directly correlated with water depth and would require much stronger assumptions for modelling. Nevertheless, it remains interesting to compare with impact forecasts, and the richness and exhaustiveness of this data source — available across all of New Zealand, over a long and continuous time scale — make it particularly valuable.

Late data

A bulk export of all flood-related claim data was requested from the NHC on March 13th. The contact point for this request was Geoffrey Spurr, Principal Loss Modelling Analyst at NHC. The data was received on May 7th, after the conclusion of the internship. It was shared for a specific use within the context of the current project and PhD research ; therefore, NIWA will not be permitted to reuse it beyond the scope of this work. However, this initial collaboration has helped establish a clearer understanding of the data's potential value for flood modelling and may facilitate faster access in future requests.

A preliminary analysis of the dataset revealed that for the Westport 2021 floods, only five land claims were recorded in the NHC database—compared to 649 damaged buildings identified in the placard dataset.

2.3.3 NZ Historic Weather Events Catalogue

In the search for flood impact data on a large spatio-temporal scale (as conducted during the PhD work in France), the New Zealand Historic Weather Events Catalogue, developed by NIWA and available at https://hwe.niwa.co.nz/, was considered. This catalogue compiles major weather events that caused significant damage or casualties over the past 200 years, from the 1800s to the present. Information has been gathered from a wide range of sources, including newspaper articles, journals, books, and databases contributed by various organisations and individuals. For each event, the catalogue records the affected regions, the types of hazards involved, and the resulting impacts.

Below is a summary of the information found in the catalogue for the July 2021 flood event :

Impacts of the July 2021 flood in the Buller district *from https://hwe.niwa.co.nz/event/July_2021_South_Island_Flooding*

- The Buller District Council declared a State of Emergency, expecting that houses would be flooded.
- The Buller River was running at a 50-year flood level.
- 1000 people spent the night in evacuation centres in the Buller district.
- A Temporary Accommodation Service was set up to help those whose homes were damaged by the flooding and severe weather across the West Coast and Marlborough.
- The government pledged $300,000 to help the Buller district.
- About 1000 stock may have drowned in the flood event.

- Water conservation orders were issued for residents in Buller. They were asked to use water sparingly.
- The West Coast Civil Defence asked residents in particular streets to voluntarily evecuate. An evacuation centre was set up at Westport South School.
- The NZ Army helped to evacuate 200 Westport residents.
- Eighty-nine Westport homes were red-stickered, meaning they had suffered severe or irreparable damage, and 400 were yellow-stickered, meaning they had suffered moderate damage.

Although the Historic Weather Events Catalogue provides valuable information, it was not used in this study. First, it does not offer precisely geolocated data, which is essential for validating forecasted flood impacts at a fine spatial scale. Second, although aggregated data can still be useful — see section 3.1 — the catalogue does not provide appropriate quantitative data for the purpose of our study. The catalogue may still be useful for broader statistical analyses over long time periods, given its extensive temporal coverage. However, as it only records events associated with significant damage or casualties, threshold bias [Wyatt et al., 2023] will be a limit for systematic impact validation.

2.3.4 Extreme Weather Event database over Aotearoa New Zealand

The Aotearoa New Zealand (ANZ) Extreme Weather Events (EWE) database [**?**] provides a comprehensive record of extreme weather events across ANZ, covering the period from March 1996 to December 2021. For each event, a score is assigned to each affected region, and this score is based on two components :

— The hazard type, labelled as **R** for broad-scale heavy rain events (over areas $> 1000km$), typically driven by synoptic features like frontal bands ; **C** for convective events, primarily thunderstorms, either embedded in frontal bands or isolated ; **W** for broad-scale strong wind events ; **S** for broad-scale heavy snow events ; **T** for extreme temperature events, which may be highly localized.
— An "Extreme likelihood" rating (1 = high, 2 = moderate, 3 = low), assigned based on expert judgment and subjective evaluation.

The information in the EWE database was primarily sourced from MetService and from the aforementioned NZ Historic Weather Events Catalogue, supplemented by media reports and, where available, insured loss data from the Insurance Council of New Zealand (ICNZ).

The main limitation of the EWE database in the context of our work is that it focuses primarily on meteorological rather than hydrological events. Furthermore, as illustrated in table F.4, the information recorded in the database is generic and impacts are not recorded in a structured way or always captured.

Attribute	Value
Unique Identifier	2021.52
Start Date	2021-07-15
Event Title	July 2021 Buller RED warning and extreme flooding
Source	MetService
Northland	
Auckland	
Waikato	3R
Bay of Plenty	3R
Gisborne	3R
Hawke's Bay	
Taranaki	3R ;3W
Manawatu-Whanganui	3R
Wellington	3R ;2W
Tasman/Nelson	1R
Marlborough	1R ;3W
West Coast	1R
Canterbury	2R
Otago	
Southland	
Chatham Islands	
Comments	Widespread severe weather (Thu 15 to Sun 18) ; with Buller hit hardest by heavy rain and extreme flooding ; but also impacting Marlborough ; with thousands evacuated in Westport and Wairau River/Valley. Major road closures. Very heavy rain also in Westland ; Canterbury ; Tasman ; Tararua Range ; Taranaki ; Bay of Plenty. Severe gales also impacted central NZ ; with wind damage in Wellington/Wairarapa ; power outages in Taranaki.
URL	https://hwe.niwa.co.nz/event/July_2021_South_Island_Flooding

TABLEAU F.4 – Details of the July 2021 Buller RED warning and extreme flooding event from the Extreme Weather Event database.

2.3.5 Leads with no results

- **FENZ** : Fire and Emergency New Zealand was approached through Hamish McEwen and Craig MacAlpine to obtain operational records, similar to those collected in the South of France. However, no response was received.
- **NZ Police** : The New Zealand Police were contacted for the same purpose, through James Gunning, who informed us that no relevant data were available for sharing.
- **GEMA** : The Geographic Emergency Management Aotearoa (GEMA) is a GIS community that facilitates collaboration and coordination during disaster response efforts and played an important role during the Cyclone Gabrielle response in 2023. After discussions with members of this community (notably Kate Burns), it was confirmed that GEMA does not retain any operational datasets.
- **Private insurance companies** : As previously discussed, flood-related house damage is typically covered by private insurance companies. After discussions with Professor Ilan Noy (Chair in the Economics of Disasters and Climate Change), who has spent nearly a decade unsuccessfully attempting to obtain private insurance claim data, it was concluded that pursuing this lead would not be worthwhile.

3 Impact modelling

3.1 RiskScape

RiskScape [Paulik et al., 2023] is an open-source software developed by NIWA to assess the impacts of multiple hazards on buildings. In practical terms, it can integrate a variety of input layers and spatially combine them to produce impact estimates.

In our application, RiskScape was used as the impact component of the impact-based flood forecasting chain. It takes as input the layer of buildings representing the June 2021 situation in Westport (extracted from ©LINZ), combines it with the hazard layers (i.e., the flood maps produced for each forecast member), and calculates the flood depth each building is exposed to.

Comparison with the impact component used in France

In the flash flood nowcasting chain implemented in the Mediterranean region, a similar process is used : the water depth is calculated for each building. The method of calculation is also comparable : for each building polygon, the maximum water depth value among all intersecting flood map pixels is retained.

Based on the AI-based physical damage prediction model described in Paulik et al. [2024], RiskScape also provides a damage ratio (i.e., cost to repair divided by cost to replace) for each building in Westport. It is important to note, however, that this damage model was trained on the same flood event that we are reforecasting, although using a different observation dataset (see Section 3.3.1).

Water depth threshold for building damage

In the damage model implemented for Westport [Paulik et al., 2024], the damage ratio is zero if the flood depth is less than 15 cm. It is reassuring to note that this 15 cm threshold was also adopted in the PhD study for considering a building as damaged, corresponding approximately to a typical door threshold. However, in the PhD, the impact-based model was designed to issue warnings based on a threshold number of buildings affected within each spatial entity (river reach or municipality).

In concrete terms, RiskScape was applied in this study by using the following command (Figure F.8) :

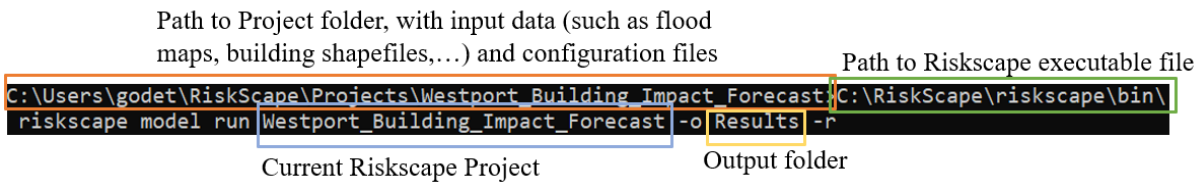


FIGURE F.8 – Riskscape command.

This system command can be directly integrated into the forecasting routine, allowing for the sequential calculation of flow, inundation, and impacts. Figure F.9 shows the structure of the RiskScape project folder, which contains all necessary input data and configuration files. The folder was provided as-is for the purposes of this study and required minimal adaptation (e.g., updating file paths in "2021_Westport_Ensemble.csv", "project.ini", "bookmarks.ini", or in "2021_Westport_Flood_Impacts_18032025.py").

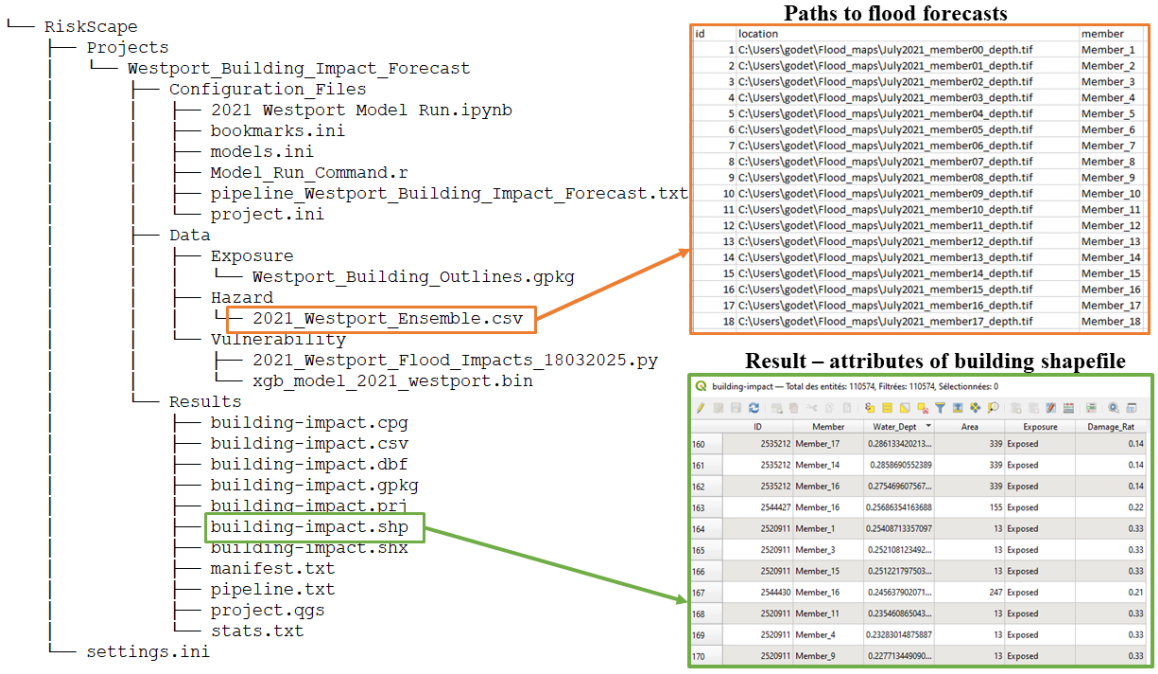


FIGURE F.9 – Riskscape folder.

Perspectives on the use of Riskscape in a flood forecasting routine

In the forecasting chain implemented in this study, RiskScape is used as the impact model, estimating the water depth to which each building is exposed. This represents a basic application of the software—an operation that could alternatively be performed through simple raster comparison, as was done in the French case study during the PhD. However, RiskScape has been optimised for fast execution and includes well-considered methods for determining which water depth value to assign to each building (e.g., pixel under centroid, mean or maximum of all pixel values within the building outlines). While its use in this routine does not introduce substantial added value from a methodological standpoint, it does save time and effort, leverages locally developed tools, and offers potential for broader applications.

3.2 Binary impact detection

The first step of impact-based modelling applied here focuses on building damage and produces a binary output : whether a building is affected or not. Based on the impact outputs provided by RiskScape, two approaches were considered :

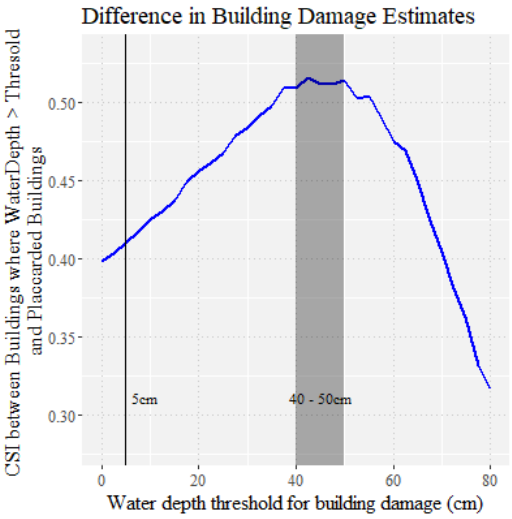
- Using the water depth each building is exposed to and applying a threshold on flood depth.
- Using the damage ratio and applying a threshold on the damage ratio value.

Comparison with the impact component used in France

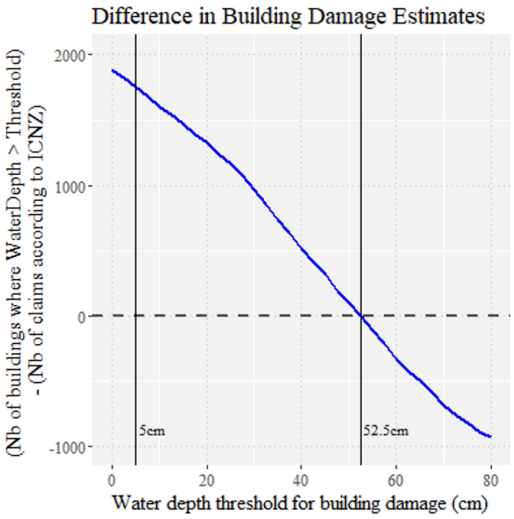
In the flash flood nowcasting chain implemented in the Mediterranean region, a water depth threshold of 15 cm was applied to each building. This value was chosen because it roughly corresponds to the typical height of a door threshold. However, no rigorous sensitivity analysis could be conducted at the time. In contrast, the Westport case study and the available datasets offer the opportunity to explore the choice of threshold more systematically.

To determine the most suitable parameter and associated threshold for binary impact forecasting, the "observed" flood map presented in Section 2.4.3 and Figure F.5 was used to compare the "best" simulated impacts (i.e., impacts derived from the best flood map simulation) against observed impacts described in Sections 3.1 (insurance claims from ICNZ) and 3.2 (placard data). Specifically :

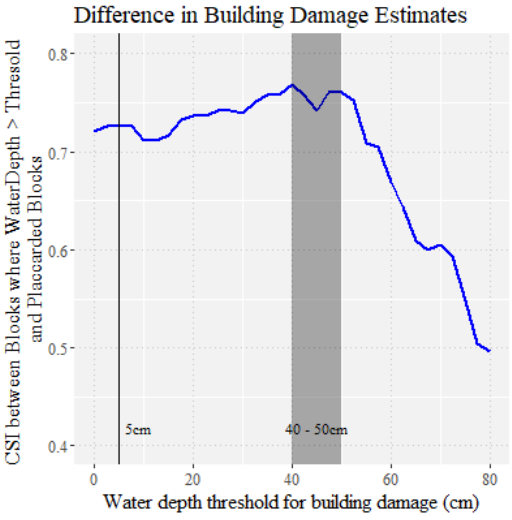
- A CSI score was calculated by comparing the buildings (or blocks) classified as affected according to the placard data with those classified as affected in the simulations, based on thresholds applied to water depth or damage ratio. Results for thresholds ranging from 0 to 80 cm (for water depth) and from 0 to 1 (for damage ratio) are presented in Figures F.10a, F.10c, and F.10d.
- The difference between the number of buildings identified as affected in the simulations and the number of insurance claims reported by ICNZ was calculated. Results for water depth thresholds between 0 and 80 cm are shown in Figure F.10b.



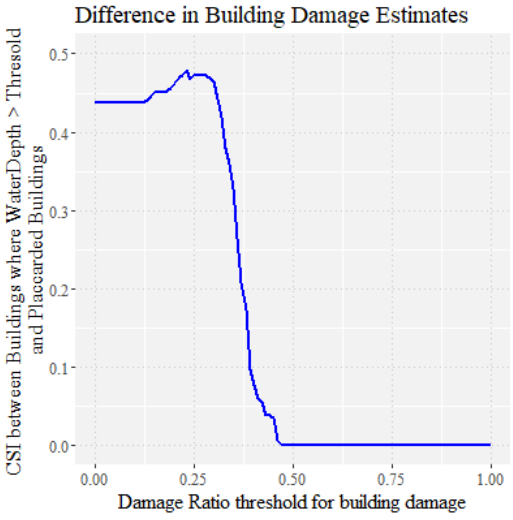
(a) Water threshold - Placard data - Building scale



(b) Water threshold - Number of claims from ICNZ - Building scale



(c) Water threshold - Placard data - Block scale



(d) Damage ratio - Placard data - Building scale

FIGURE F.10 – Sensitivty analysis to the selected parameters (i.e. water depth or damage ratio), based on selected validation data (i.e. placard data or number of claims from ICNZ), at the selected spatial scale (i.e. building scale or block scale).

From this sensitivity analysis, we retained the following conclusions :

- Using the damage ratio instead of the water depth does not lead to better results. Therefore, the water depth was selected as the parameter to consider a building as affected by flooding.

- Across both building and block scales, and for both observation datasets (placard data and insurance claims from ICNZ), a water depth threshold of 50 cm appears to offer the best compromise between over- and under-detection.
- Even when using the "best" flood map, the CSI score does not exceed 0.55 at the building scale and 0.75 at the block scale. Consequently, the validation scores obtained with the flood forecasting chain are expected to be limited by these values. Several explanations can be advanced :
  - Data and model limitations : The "best" flood map remains an approximation rather than a true observed flood map ; observed impacts may not perfectly reflect actual ground truth ; and the impact model based on a simple water depth threshold inevitably simplifies the complexity of flood damage processes.
  - Findings from other studies : Previous research has shown that a significant proportion of flood damage is diffuse, unexplained by flood modelling alone, and linked to complex runoff processes (see below).

A French feedback on low impact detection scores

— Le Bihan et al. [2017] compared insurance claims related to two flash flood events in the Mediterranean region to different flood scenarios. They found that a significant proportion of claims were located outside the estimated 1000-year flood envelope. They suggested that these claims could have been caused by unmodelled flood mechanisms such as runoff accumulation in low points, sewer saturation, or cellar flooding due to rising groundwater levels.

— Moncoulon et al. [2020] performed a broader analysis by comparing all flood claims recorded in the Caisse Centrale de Réassurance database to flood hazard maps for the French Mediterranean region. These maps accounted for both riverine and surface flooding. However, they observed that almost 40% of reported damages were located outside all mapped flood-prone areas, even when considering runoff modelling in upstream areas. They noted that runoff-related damage does not follow hydrographic patterns and that, while it is driven by water height and speed, its paths are much more diffuse and unpredictable than those of river floods. Improved knowledge of drainage and stormwater evacuation network operations could enhance understanding of exposure to runoff.

3.3 Multi-category damage detection

The placard dataset, which includes information on observed damage severity, offers an opportunity to assess whether flood forecasts can go beyond binary outputs (affected / unaffected) and provide insight into damage severity. This section investigates whether such multi-category damage detection can be achieved without significantly compromising forecast accuracy.

3.3.1 At the building scale

To enable multi-category damage detection at the building scale, three water depth thresholds (T_1, T_2, T_3) are defined to match the damage categories present in the placard dataset :

— **Unaffected :** $0\,\text{cm} \leq WD < T_1$

— **Light damage :** $T_1 \leq WD < T_2$

— **Moderate damage :** $T_2 \leq WD < T_3$

— **Heavy damage :** $T_3 \leq WD$

Based on the results from Section 4.2, T_1 is set to 50 cm. The remaining thresholds T_2 and T_3 are determined through a sensitivity analysis using the reference flood map and the placard data at the building scale. For each tested threshold combination, a CSI score is computed for each category, and the values that maximize the total CSI sum are retained. The final thresholds are :

— **Unaffected :** $0\,\text{cm} \leq WD < 50\,\text{cm}$
— **Light damage :** $50\,\text{cm} \leq WD < 65\,\text{cm}$
— **Moderate damage :** $65\,\text{cm} \leq WD < 110\,\text{cm}$
— **Heavy damage :** $110\,\text{cm} \leq WD$

3.3.2 At the block scale

As explained in Section 3.2 and illustrated in Figure F.6, aggregating damage severity raises methodological questions. A straightforward option is to retain the maximum damage category observed within each block. However, this approach does not consider the number of damaged assets within the block.

As an alternative, we define a continuous *Impact Score* by associating each damage category with a numeric value, as shown in Table F.3 :

$$\text{DamageNb} = \begin{cases} 0 & \text{Unaffected} \\ 1 & \text{Light damage} \\ 2 & \text{Moderate damage} \\ 3 & \text{Heavy damage} \end{cases}$$

For a block containing N buildings, the aggregated score is calculated as :

$$\text{ImpactScore}_{\text{block}} = \sum_{i=1}^{N} \text{DamageNb}_i$$

This approach results in a continuous variable, which takes values between 0 and 87 when applied to the observed placard dataset.

Similarities with the French Study Case

This approach is consistent with the impact simulations used to compare modelled impacts with the rankings provided by the Caisse Centrale de Réassurance (CCR). Specifically, CCR supplied ranked lists of impacted areas based on the number of insurance claims for numerous flood events in the French Mediterranean region. Corresponding simulated rankings were produced by ordering spatial entities according to the number of buildings affected in the simulations (i.e., buildings exposed to water depths exceeding 15 cm).

4 Cascading uncertainties

As illustrated by the structure of this report (Section 2), the impact-based flood forecasting chain consists of a sequence of interconnected modelling components. While evaluating the performance of each component independently is important, it is essential to also consider the forecasting system as a whole. A comprehensive understanding of how uncertainties propagate through the chain — and how the different components interact — remains a key challenge in the literature, particularly in the context of impact-based forecasts.

Titley et al. [2024] analysed the relationships between track, precipitation, and river discharge skill within a flood forecast chain, for one flood event (Hurricane Iota, which hit Central America in November 2020). Building on this, the present study extends the analysis by incorporating the flood mapping and impact modelling components, offering a broader view of the cascading uncertainties throughout the entire forecast chain.

Expanding the scope of forecast evaluation in the PhD

In the PhD work focusing on the French Mediterranean region, each component of the flood forecasting chain was rigorously evaluated across a wide spatio-temporal scale — including continuous hydrological validation over more than 10 years for the flow component, comparisons between observed and simulated flow–stage relationships at nearly 200 hydrometric stations for the flood library component, and continuous comparison of simulated warnings to insurance claim data over more than 10 years for the impact component. However, the propagation of uncertainty through the forecasting chain was never explicitly analysed at the event scale. The New Zealand case study offers an opportunity to explore this dimension in greater detail.

4.1 Forecast performance through the chain

Using the forecast and validation data described in the previous sections, errors were calculated for each block of the flood forecasting chain. Boxplots representing the ensemble spread for each forecast cycle are presented in Figure F.11. Specifically :

- For rainfall forecasts, the spatially averaged catchment rainfall over the preceding 48 hours was calculated at each time step. The maximum value was then compared to the best available proxy for 48-hour maximum rainfall, and the difference between the two was computed. Results are shown in Figure F.11a. Additional indicators were also tested, including 24-hour maximum rainfall, total accumulated rainfall, and Nash-Sutcliffe Efficiency (NSE) scores for 3-hour and 6-hour rainfall accumulations. These results are presented in Figure F.17.
- For flow forecasts, the Kling-Gupta Efficiency (KGE) was computed at each time step for the forecasted flows at the Te Kuha station, and compared to the KGE calculated for the best proxy flow. These errors are illustrated in Figure F.11b. In addition, Figure F.17 presents results based on an alternative performance metric : the difference in peak flow values.
- For flood forecasts, forecasted flood water levels were compared to observed water levels (based on the flood map in Figure F.5), using the Root Mean Square Error (RMSE) computed over all 5 m $\times$ 5 m pixels. Results are shown in Figure F.11c.
- For impact forecasts, affected blocks identified in the placard dataset were compared to simulated affected blocks (defined as blocks with at least one building exposed to water depths exceeding 50 cm). This binary comparison enabled the calculation of the Critical Success Index (CSI), with

results presented in Figure F.13d. Additional CSI results at the building scale are presented in Figure F.13a, and boxplots comparing model outputs to the total number of claims reported by ICNZ are shown in Figure F.13c.

For presentation purposes, all boxplots show absolute errors, and CSI (resp. KGE) values are converted to 1 − CSI (resp. 1 − KGE) to allow for a consistent visual interpretation of decreasing error trends.

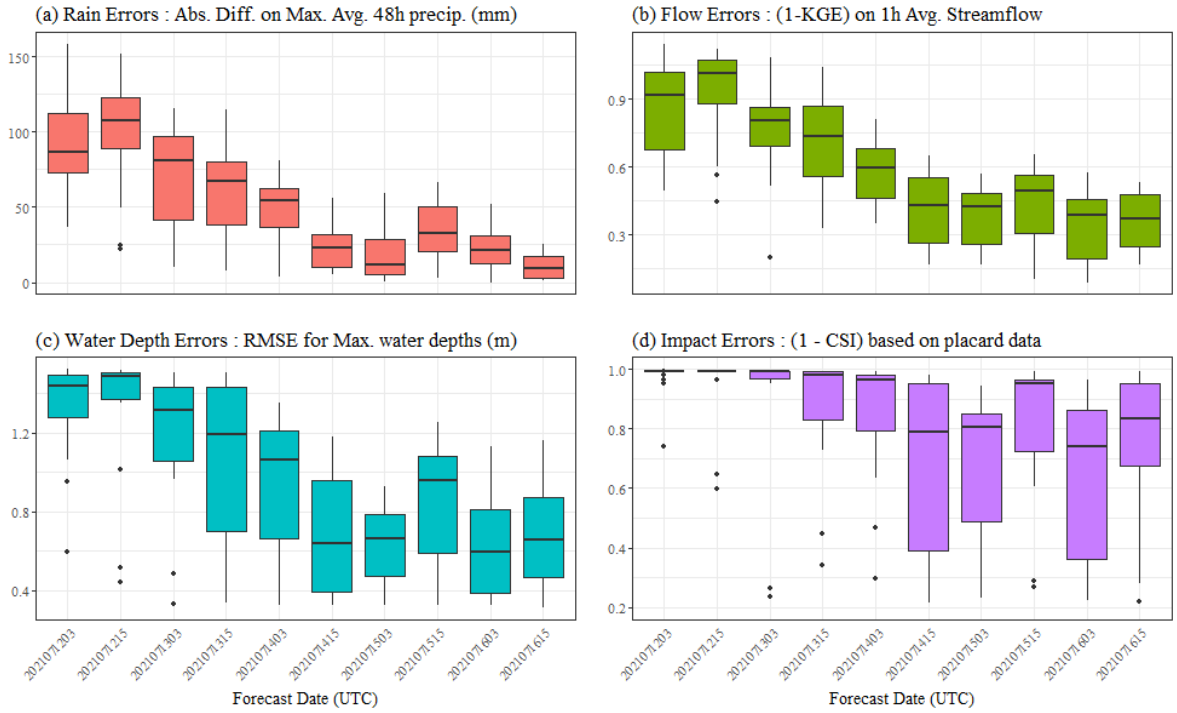
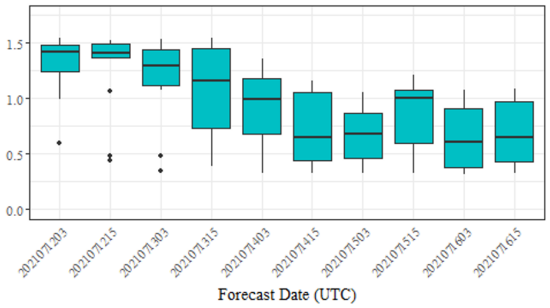


FIGURE F.11 – Boxplots comparing forecast performance across the ensemble members for each forecast cycle

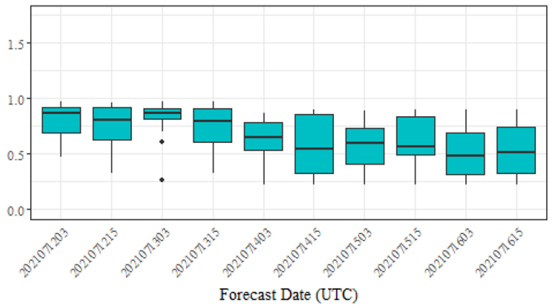
Below are some comments on these results :

- The rainfall forecasts consistently improve up until 2021071415UTC, approximately three days before the event. An unexpected drop in skill is then observed in the 2021071503UTC forecast cycle, leading to a performance degradation that propagates through the entire forecasting chain. Between the final two forecast runs (2021071603 and 2021071615), the rainfall skill improves again ; however, this improvement does not translate across the rest of the chain. The skill between the last two forecasts does not change much for the flow and flood components (this is confirmed in figure F.14), but it significantly deteriorates at the impact level. This non-linear response highlights the complex propagation of uncertainty and the sensitivity of downstream components to modest upstream changes.
- From 2021071415UTC onward, the performance of streamflow forecasts stabilises, with the median KGE converging toward 0.6 (1 − KGE) ≈ 0.4, a value generally considered indicative of “good” predictive skill. However, the event was still markedly underestimated, as shown by the hydrographs in Figure F.3 and the correlation plots in Figure F.17, with most rainfall and discharge errors (Sim. – Obs.) being negative.
- Apart from the skill degradation observed at 2021071515UTC—triggered by a deterioration in rainfall forecast skill—the median RMSE for water depths remains around 60 cm in the last forecast cycles. This error must be interpreted in light of the expected uncertainty in the “observed” water depths used for comparison, which can reach ±30–50 cm according to Bates [2022]. Moreover,

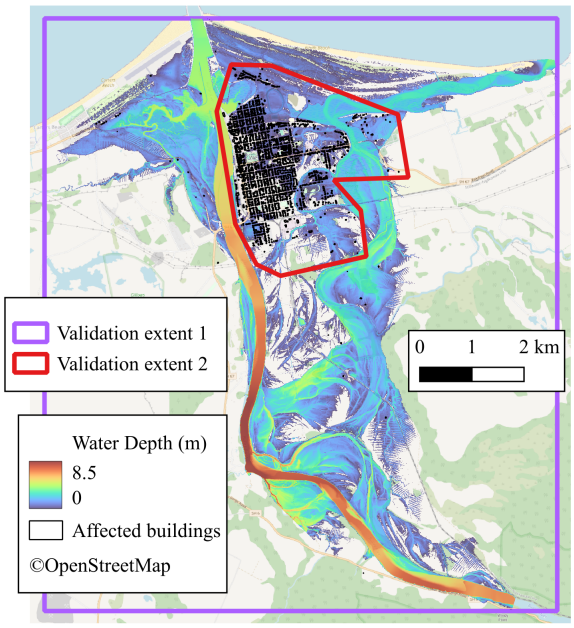
the initial RMSE values were computed over the entire simulation domain, which includes areas far beyond the urban perimeter of Westport (see purple rectangle in Figure F.12c). When the validation domain is restricted to the impacted area only (see red rectangle), the RMSE improves, dropping closer to 50 cm in the final forecasts (see figure F.12b). Nonetheless, even a 50 cm median error significantly affects the impact model's performance, especially given that the threshold for a building to be considered "affected" in the simulation is also set at 50 cm, as detailed in Section 4.2.



(a) RMSE (m) from validation extent 1



(b) RMSE (m) from validation extent 2



(c) Extents on which the flood validation was performed

FIGURE F.12 – Comparison between simulated and observed maximum water depths during the event, evaluated over two different flood extents.

- At the impact level, the skill (translated by the CSI score) is limited and does not exceed 0.8 for the best-performing member, as explained in section 4.2. Even for the best performing forecast cycle, the median CSI is low (0.26).

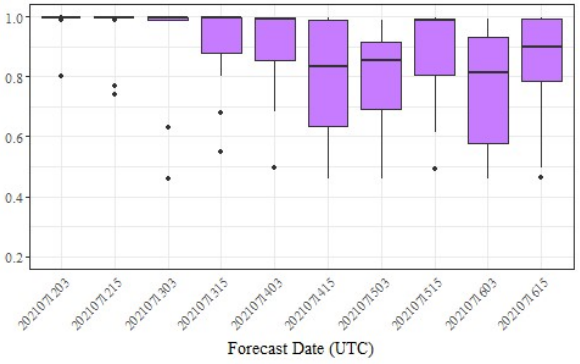
Comparison with impact scores obtained in France

Some explanations for the low impact detection scores have been provided in Section 4.2. Additional insights can be drawn by comparing these results to those obtained from the French case study. Over the last five forecasts, the median CSI values for the impact component at the block scale in the Westport case study (Figure F.11d) were 0.21, 0.20, 0.05, 0.26, and 0.17, respectively, yielding an average of 0.18.

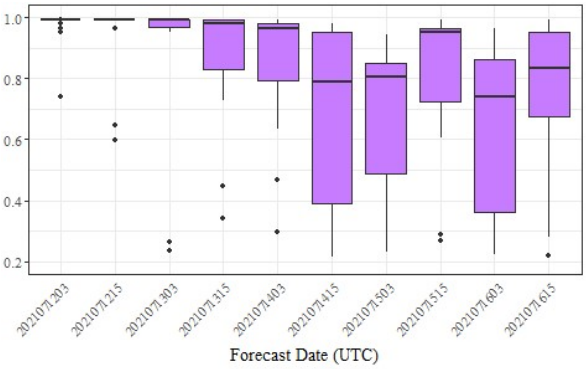
In France, for the continuous validation of the impact-based nowcasting chain at the river reach

scale—using insurance claims as the reference—the obtained CSI was 0.17.

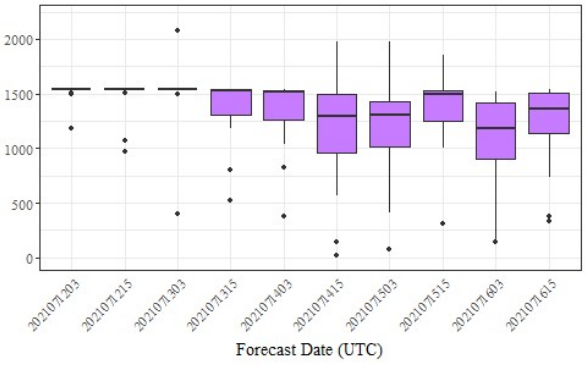
Although the differences in validation window (event-scale vs. continuous), spatial scale (0.03 km^2 for blocks vs. 4 km^2 for river reaches), domain (city scale vs. regional scale), and forecast range (multi-day forecasts vs. nowcasts) make direct comparison difficult, the structure of the forecasting chains and the validation method are broadly similar. Despite these differences, the similarity in CSI values is noteworthy.



(a) CSI based on placard data at the building scale



(b) CSI based on placard data at the block scale



(c) Abs. Diff based on ICNZ number of claims

FIGURE F.13 – Boxplots of impact errors using different validation datasets

4.2 Run-to-run jumps in forecast performance

As in Titley et al. [2024], we calculate the divergence score as defined by Richardson et al. [2020] :

$$\text{Divergence} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|x_i - y_j| - \frac{1}{2M^2}\sum_{i=1}^{M}\sum_{j=1}^{M}|x_i - x_j| - \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}|y_i - y_j|$$

where x_i and y_j are the error values for ensemble members i and j in two successive forecast cycles x and y, with $M = N = 18$ the number of members in each forecast.

This divergence score quantifies the change in ensemble forecast performance from one run to the next. By comparing divergence scores across the rainfall, flow, flood, and impact components of the forecasting chain, the aim is to assess how closely aligned the run-to-run shifts in performance are across the different components.

The results, presented in figure F.14, show that the largest change across all variables occurs between the 2021071415 and 2021071503 runs, with a particularly pronounced peak in rainfall divergence. This marks a clear transition in the forecast system, where the rainfall ensemble undergoes both a substantial shift in median values and a reduction in spread (Fig. F.13). The matching peaks in flow, water depth, and

impact divergence indicate that this change in rainfall forecast propagated through the entire forecasting chain.

By contrast, other forecast transitions such as 2021071515 to 2021071603 show low divergence for rainfall, but a more marked divergence for water depth and impact forecasts. This suggests that even when the rainfall forecast remains structurally stable, changes in downstream components of the forecasting chain can still occur.

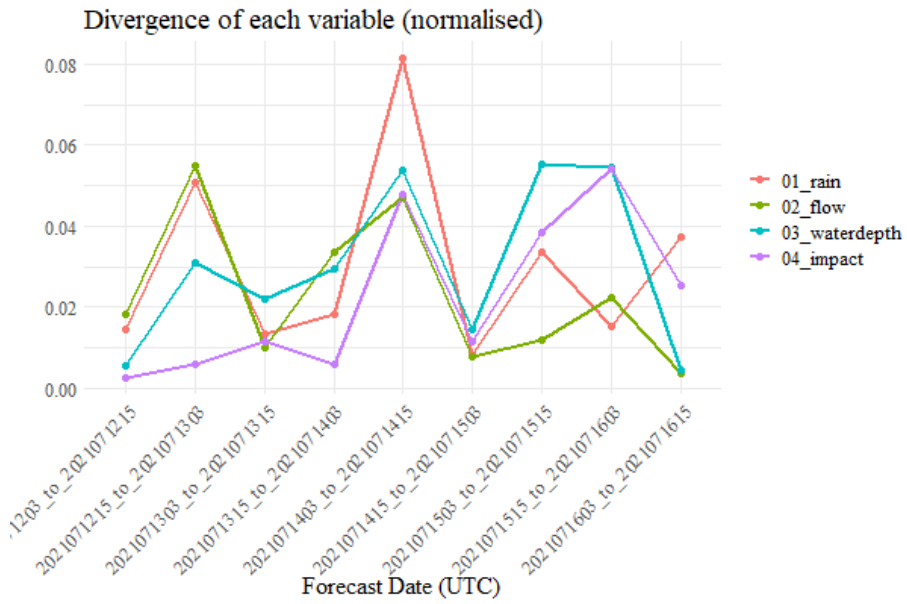


FIGURE F.14 – Divergence score of all error metrics, normalized by the y-axis range

4.3 Correlation of Errors

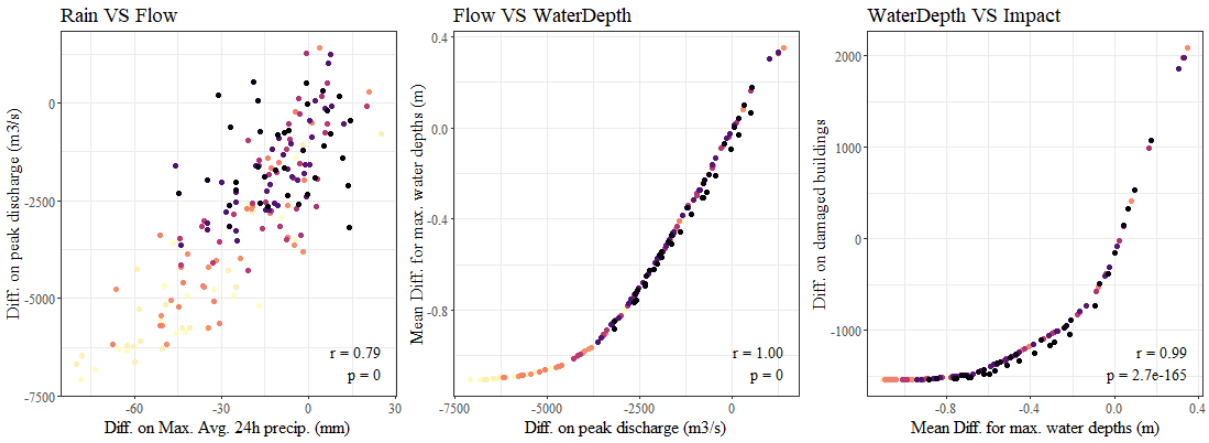
Another way to represent the cascading uncertainties throughout the flood forecasting chain is through error correlation graphs, as proposed by Titley et al. [2024]. Rainfall–flow error indicators are shown in Figure F.17, while correlations between flow and inundation, and between inundation and impacts, are presented in Figure F.15. The results indicate that :

- There are very strong correlations between flow and inundation errors, and between inundation and impact errors, indicating tight coupling across components of the modelling chain.
- In particular, the perfect correlation between peak discharge error and water depth error (Spearman $r = 1$) suggests that the 2021 Westport flood was primarily driven by peak flows, with limited influence from flood volume or tidal/sea-level components.

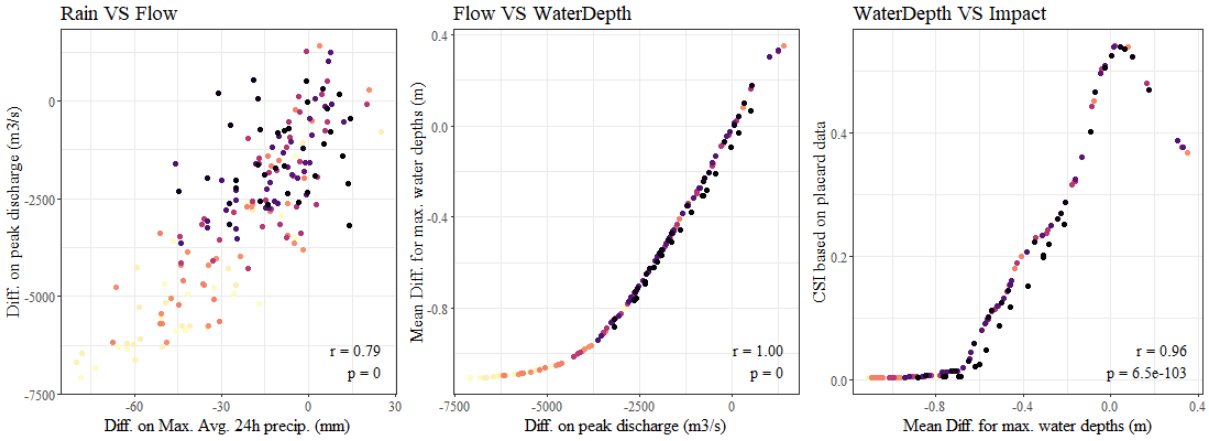
> **Implications for Hydraulic Modelling**
>
> Compared to the Mediterranean case study, these results imply that a steady-state hydraulic modelling approach, as used operationally in France, would likely have provided satisfactory results for this specific flood event.

  The correlation appears linear over a wide range of flow errors ($[-2500, 0]$ m^3/s), but the present dataset does not allow investigation of model behaviour under flow overestimation scenarios.

- The correlation between water depth and impact prediction performance is also strong. The Spearman correlation between water depth error and impact error is slightly lower when validation is



(a) Errors in each model component : Rainfall error is the difference (Sim - Obs) in the temporal maximum of spatially averaged 24-hour catchment precipitation. Flow error is the difference in peak discharge. Inundation error is the mean difference in maximum water depths. Impact error corresponds to the difference in the number of affected buildings, using the ICNZ total claims as a reference.



(b) Same as (a), except for the impact component, which is validated at the building scale using placard data and the CSI as the evaluation metric.

FIGURE F.15 – Scatterplots of errors in each model component ; each dot represents one ensemble member. Members are colored by forecast run (darker = more recent). Spearman correlation coefficients (r) and p-values are reported in each panel.

performed at the building scale using placard data ($r = 0.96$, Figure F.15b), compared to event-scale validation using total claims from the ICNZ dataset ($r = 0.99$, Figure F.15a). This small difference likely reflects the increased complexity and noise introduced at finer spatial scales, where model inaccuracies, local vulnerability factors, and observational uncertainties become more prominent.

Figure F.15b shows that achieving a Critical Success Index (CSI) above 0.5 for the prediction of impacted buildings typically requires the mean water depth error to remain within approximately $[-0.05\,\text{m}, 0.1\,\text{m}]$. This range is slightly asymmetric, potentially indicating greater tolerance for overestimation than underestimation. However, due to the limited number of ensemble members achieving high CSI values, further analysis would be needed to confirm this pattern.

To investigate these relationships at the building scale, Figure F.16 shows the distribution of HIT, MISS, and FALSE ALARM classifications based on the last five forecast cycles across all ensemble members. This complements the domain-scale analysis by highlighting the variability of impact predictions at individual building locations :

– The FALSE ALARM distribution peaks near zero error, underscoring the sensitivity of threshold-based impact models : even small overestimations around the 50cm threshold can

lead to false positive predictions.

- The MISS distribution peaks around -0.5m, indicating that in many false negative cases, the model underestimated water depth by about 50cm. A secondary peak near 0m suggests that even minor underestimations can lead to missed impacts if they occur near the threshold.
- Unlike Figure F.15b, which aggregates errors across the domain, this figure shows that all three outcomes (HIT, MISS, FALSE ALARM) can still occur at near-zero water depth error. This highlights the influence of local uncertainties and the limitations of binary classification. Nonetheless, the peak of the HIT distribution at 0m confirms that accurate water depth predictions generally yield correct impact classification.

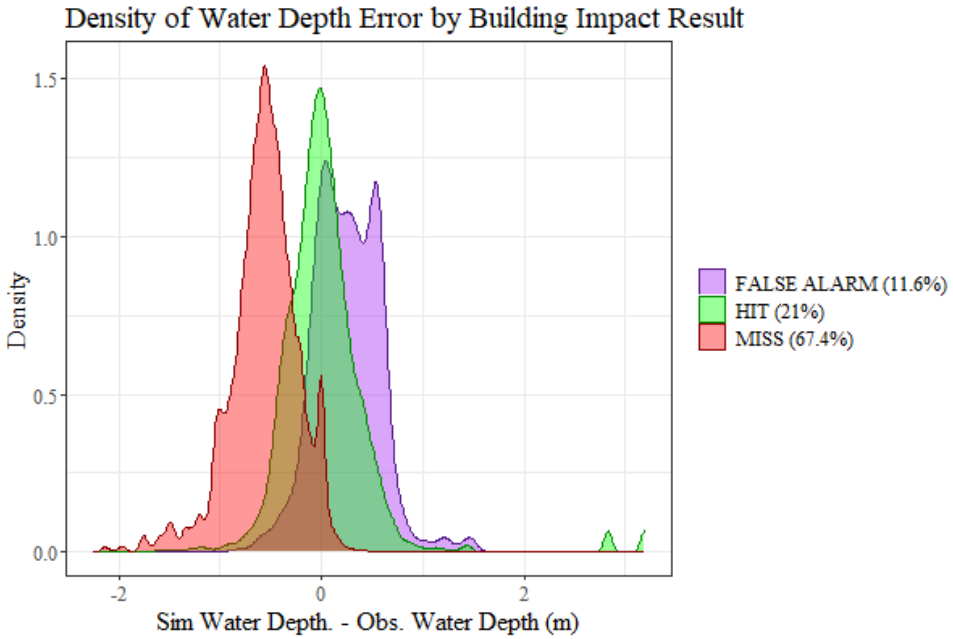
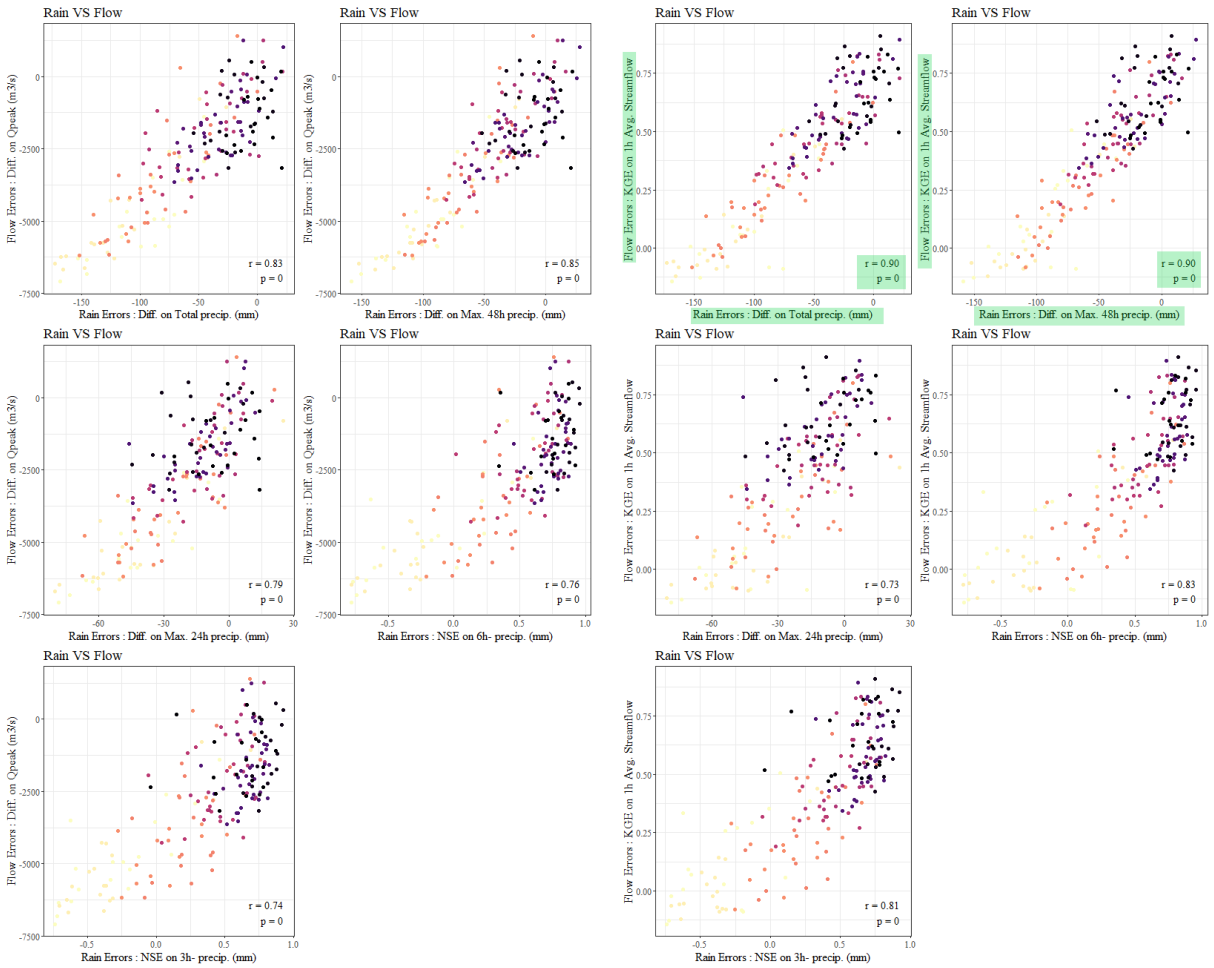


FIGURE F.16 – Density of Water Depth Error by Building Impact Result. The x-axis shows the water depth error at each building location, calculated over the last five forecast cycles and across all ensemble members. Each building is assigned a contingency category (HIT, MISS, or FALSE ALARM) based on whether the simulated water depth exceeds 50cm, and whether the building was recorded as damaged in the placard dataset.

- As shown in Figure F.17, the strongest correlation between rainfall and streamflow occurs when the rainfall error is defined as the difference in total or maximum 48-hour catchment precipitation, and the streamflow performance is measured using KGE. This highlights the importance of accurate cumulative rainfall estimation for runoff modelling. In contrast, rainfall metrics based on short-term dynamics (e.g. 3-hour or 6-hour NSE) are less strongly correlated with flow error, suggesting that runoff generation in this case is more sensitive to total rainfall volume than to sub-daily temporal distribution. Furthermore, the higher correlations obtained with KGE (compared to peak discharge error) confirm that rainfall accuracy influences the full hydrograph shape, not just flood peak values.



(a) Rainfall–flow error correlations using peak discharge as the flow metric, and various rainfall indicators.

(b) Rainfall–flow error correlations using KGE as the streamflow metric, and various rainfall indicators.

FIGURE F.17 – Comparison of rainfall–flow error correlations using different metrics for both rainfall and flow components.

5 Comparing quantile-based approaches

Until now, we have propagated the information through the different components of the forecasting chain using individual ensemble members. However, another possibility is to use probability information such as minimum, maximum, or quantiles. With this second method, it is interesting to study how the final result (i.e., the impact forecast) varies when the probability is calculated at varies locations of the forecasting chain. To illustrate this, let's schematize the ensemble flood forecasting chain that simulates successively the discharge at Te Kuha (Q), the inundation on the simulation domain ($Inundation$) and the impacts in terms of building damage ($Impact$). The quantiles of the ensemble (i.e. 50% quantile, 90% quantile...) can be computed at different locations :

$$\begin{cases} Impact(Inundation(quantile(Q))) \\ Impact(quantile(Inundation(Q))) \\ quantile(Impact(Inundation(Q))) \end{cases}$$

It is interesting to assess how the impact forecast varies depending on how and where the probabilities are calculated, as it helps understanding how probability information and error propagate through the cascade of forecasts. In order to do that, we have considered a lagged-ensemble using the 5 last forecast cycles, to be able to build a $18 \times 5 = 90$ members ensemble which is of more appropriate size to calculate quantile values.

The results, in figure F.18 show very few differences in the final results for the considered probabilities : 50%, 75%, 85%, 90%, 95% and max. This indicates that for this event, the forecasting chain is quite stable and monotonous : high flow values will result into high water depths which in turn yield significant impact. Then it also means that in this case, the simplest approach : $Impact(Inundation(quantile(Q)))$ would be sufficient. In operational terms, this is quite important since it would siginficantly reduce computation times.

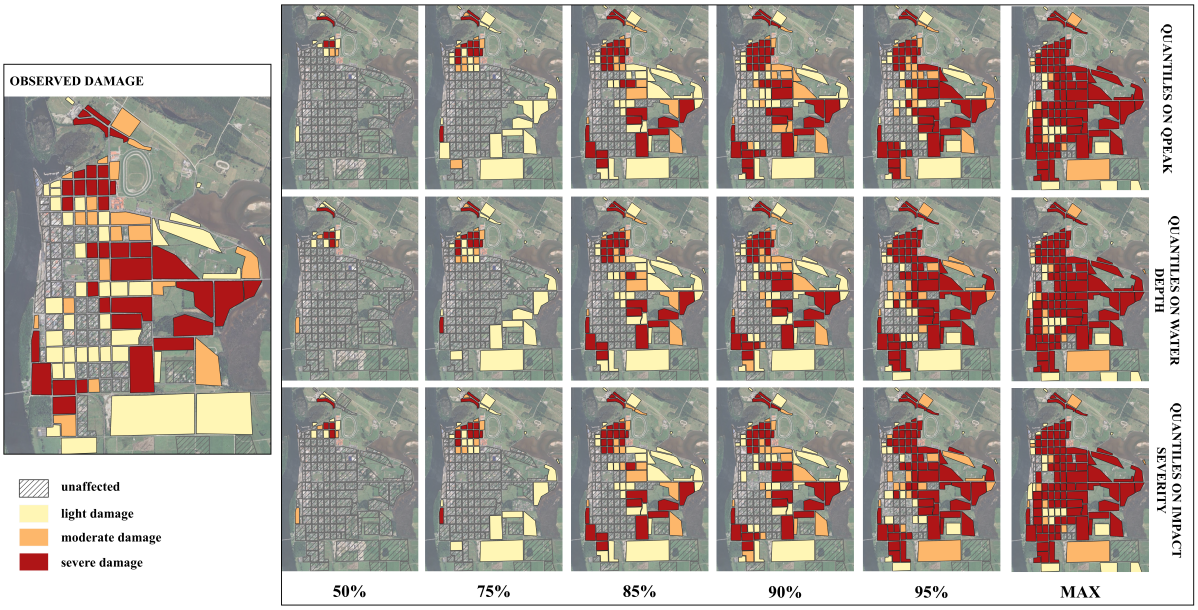


FIGURE F.18 – Impact (damage severity aggregated at the block scale) observed based on the placard dataset (left), and forecasted for different probabilities, with probabilities calculated on peak flows at Te Kuha (top), on the water depth at the pixel-scale (middle), and on the ranking of affected blocks (bottom).

6 Flood library and interpolation methods

In order to assess the impact of flood library selection and interpolation method on forecast performance—and to ensure comparability with the French case studies—the following approaches were tested :

- **Flood library selection (15 scenarios) :**
  - *Uniform library :* Scenarios are selected from the full set of ≃100 simulations using a maximum dissimilarity algorithm to ensure uniform coverage of the 2D parameter space defined by standardised peak sea level and standardised peak flow.
  - *Design library :* Scenarios are selected to represent typical fluvial and coastal return periods, initial fluvial climate change simulations (limited RCPs), sea level rise increments, additional RCP-based fluvial scenarios, and frequent low-impact coastal events.
- **Interpolation methods :**
  - *1D interpolation :* Linear interpolation within the 1D parameter space defined by peak 2-hour flow.
  - *1D nearest neighbour :* Selection of the scenario with the closest peak 2-hour flow, i.e., no interpolation.

Since interpolation is applied in a 1D space defined by flow at Te Kuha, only scenarios with distinct flow values will result in a distinct flood map. For each library, these flow values are presented in table F.5.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Design Library	7451	8514	9304	9493	9818	10564	10383	10738
Uniform Library	8514	11582	8175	3753	11944	5252	5945	5399

	9	**10**	**11**	**12**	**13**	**14**	**15**	**Obs.**
Design Library	11581	1000	1000	1000	1000	1000	1000	**8632**
Uniform Library	117	8994	8169	3096	1000	1000	12492	

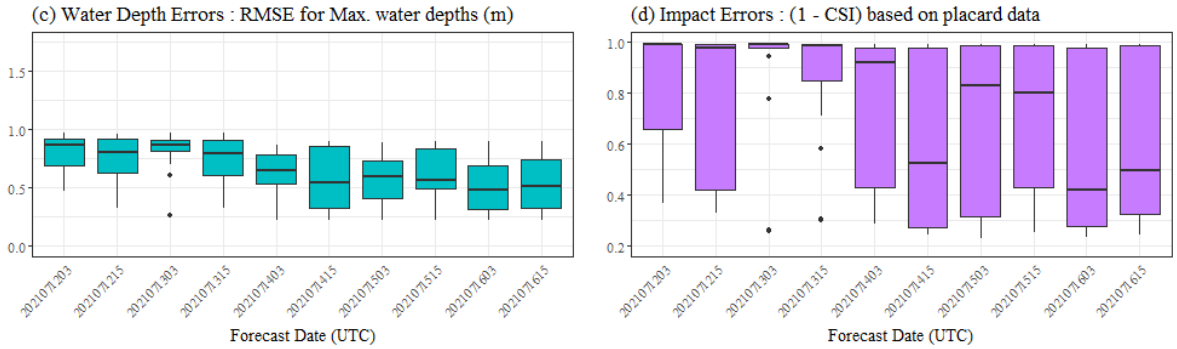
Tableau F.5 – Flow values at Te Kuha (m^3/s) for the 15 scenarios of each flood library considered.

The performance of the Inundation and Impact components for each configuration is illustrated with the boxplots presented in figure F.19.

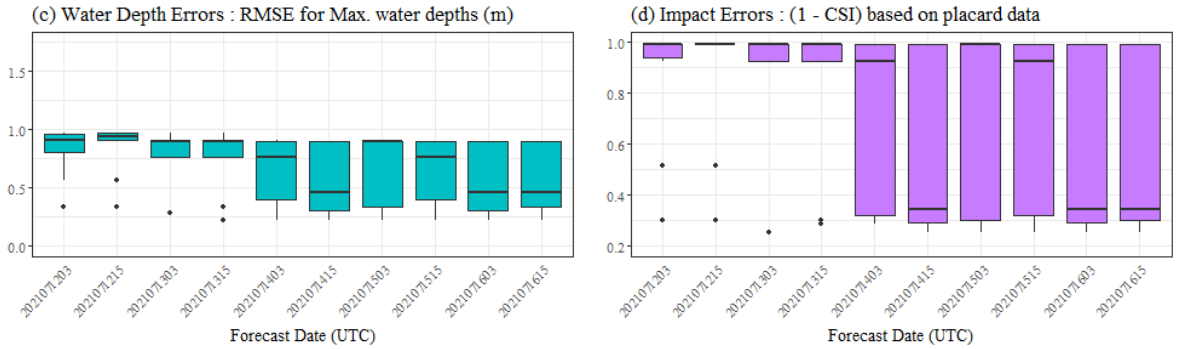
The results show a significantly larger spread and higher errors for the Uniform library. As seen in Table F.5, this can be attributed to the greater representation of lower flow values within the Uniform library—seven scenarios have flow values below 8000 m^3/s, compared to only two in the Design library. Since the forecasts substantially underestimated the peak flow (see Figure F.3a) and rarely reached the reference value of 8632 m^3/s, the Uniform library provided more low-flow scenarios, which ultimately led to larger errors.

In this case, the Design library performs better, but this appears to be mostly coincidental. Had the event been less extreme, the Design library would likely have poorly captured it.

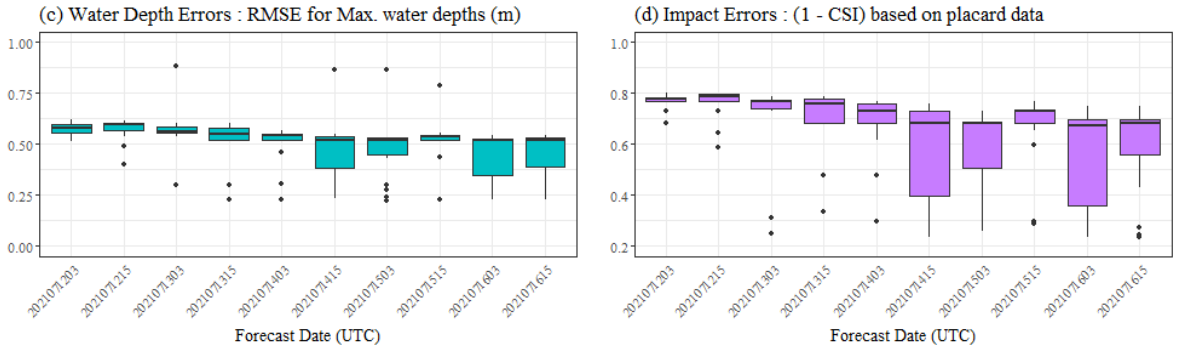
Regarding interpolation methods, linear interpolation generally yields slightly better or comparable results to nearest-neighbor, especially for inundation errors. However, in some cases—such as early forecasts using the Design library—nearest-neighbor interpolation performs similarly or even better.



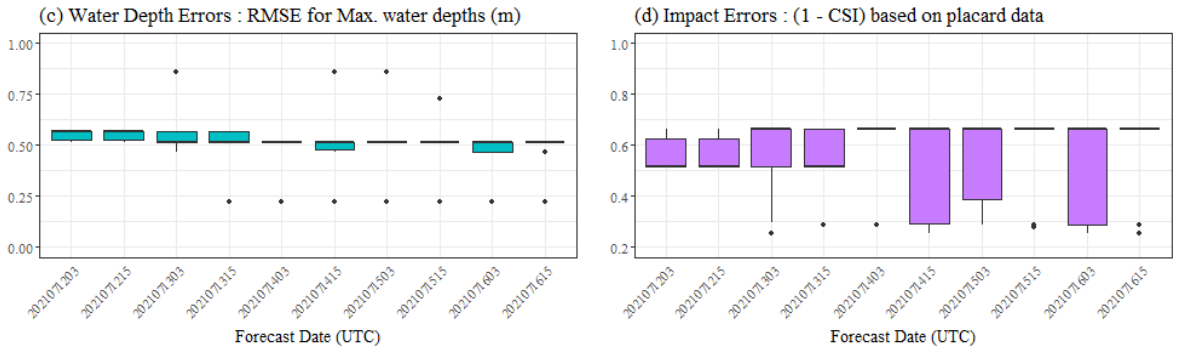
(a) Uniform library, 1D linear interpolation method



(b) Uniform library, 1D nearest neighbor method



(c) Design library, 1D linear interpolation method



(d) Design library, 1D nearest neighbor method

FIGURE F.19 – Comparison of the performance of the Inundation and Impact component for the different flood libraries and interpolation methods tested.

7 Representing impact-based flood forecasts : user insights

Another important aspect of ensemble impact-based flood forecasts is their communication. The communication of ensemble forecasts is a well-known challenge and has been increasingly addressed in recent years. However, most of the existing literature focuses on rainfall and/or flow forecasts. The first study to examine more downstream components of the forecasting chain, specifically inundation forecasts, is Jean et al. [2023]. Key findings from that study are summarised below. Four prototype visualisations of ensemble inundation forecasts were presented to respondents from various user groups :

- Visualisation tools should be adapted to the specific needs of users.
- The amount of information displayed in a single graph should be limited.
- Users interpret information with inherent cognitive biases.
- Probability intervals are generally recommended for communication.
- Across all interview groups, forecasts were mainly considered useful for anticipating potential consequences by comparing to familiar reference events.
- Including both hydrographs and reference thresholds was seen as beneficial.
- Prototype 2—which presents low, median, and high scenarios without explicit probability values—was the most preferred.
- Prototype 4—which asks users to input a threshold and then shows the probability of exceedance—was not well received.
- The concept of probability remains widely misunderstood.

To the best of our knowledge, no study has yet focused on the communication of ensemble impact forecasts, which introduces additional specific challenges. To explore this gap, a survey was designed to investigate preferences regarding the content and visualisation of impact forecasts. It was distributed during a Webinar organised by the Resilient River Community (available at https://www.resilientrivers.nz/webinar-recordings). Two poll questions were displayed live during the Webinar (37 responses), and a more detailed survey was provided afterward (24 responses). Each survey question and the corresponding responses are presented below.

7.1 Questions and answers from the survey

7.1.1 Profile of respondents

Survey on flood warnings

This survey aims to better understand content and visualization needs for ensemble impact-based flood warnings. Your feedback is invaluable.

The estimated time to complete the survey is **5–10 minutes**. It contains **8 questions**.

* Indique une question obligatoire

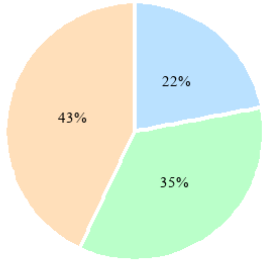
About you – In the context of flood warnings, which role best describes your primary position? *

- ◯ Member of the public (information receiver)
- ◯ Emergency manager or stakeholder (information receiver)
- ◯ Flood forecaster or flood expert (information provider)

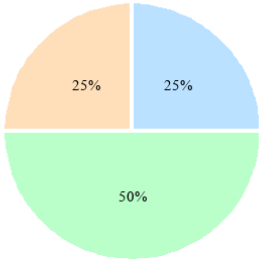
(Optional) Please specify your job title or role:

Votre réponse

(a) Poll responses



(b) Survey responses



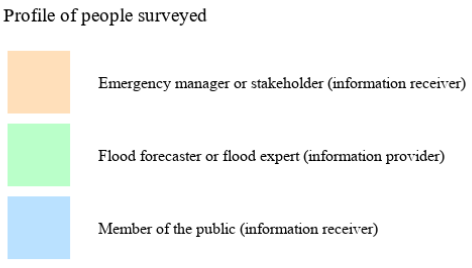
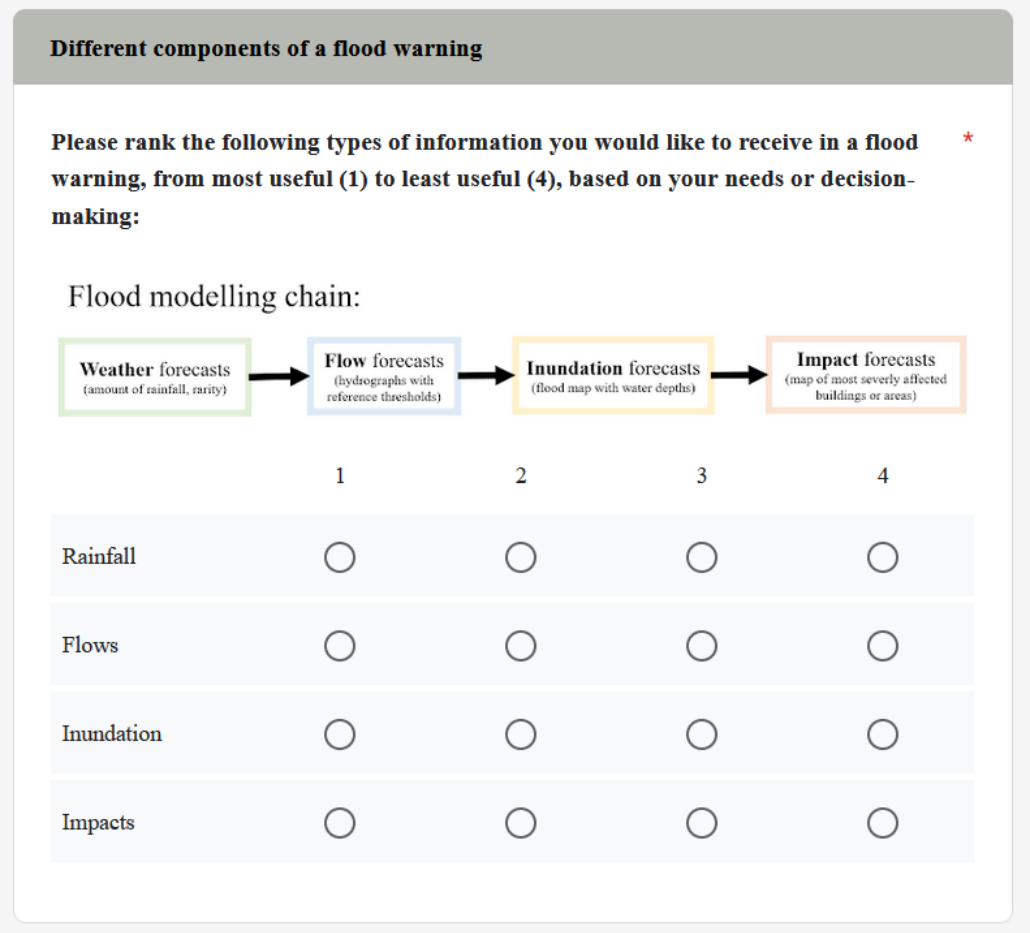


FIGURE F.20 – First question of the survey and results obtained - profile of respondents

7.1.2 Different components of a flood warning



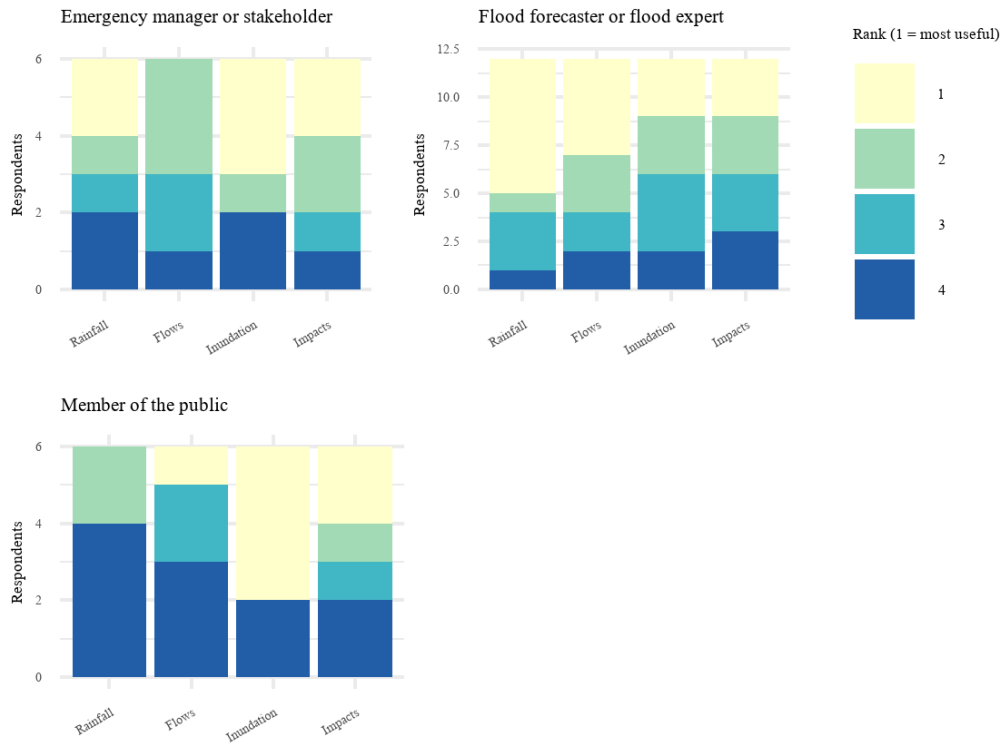
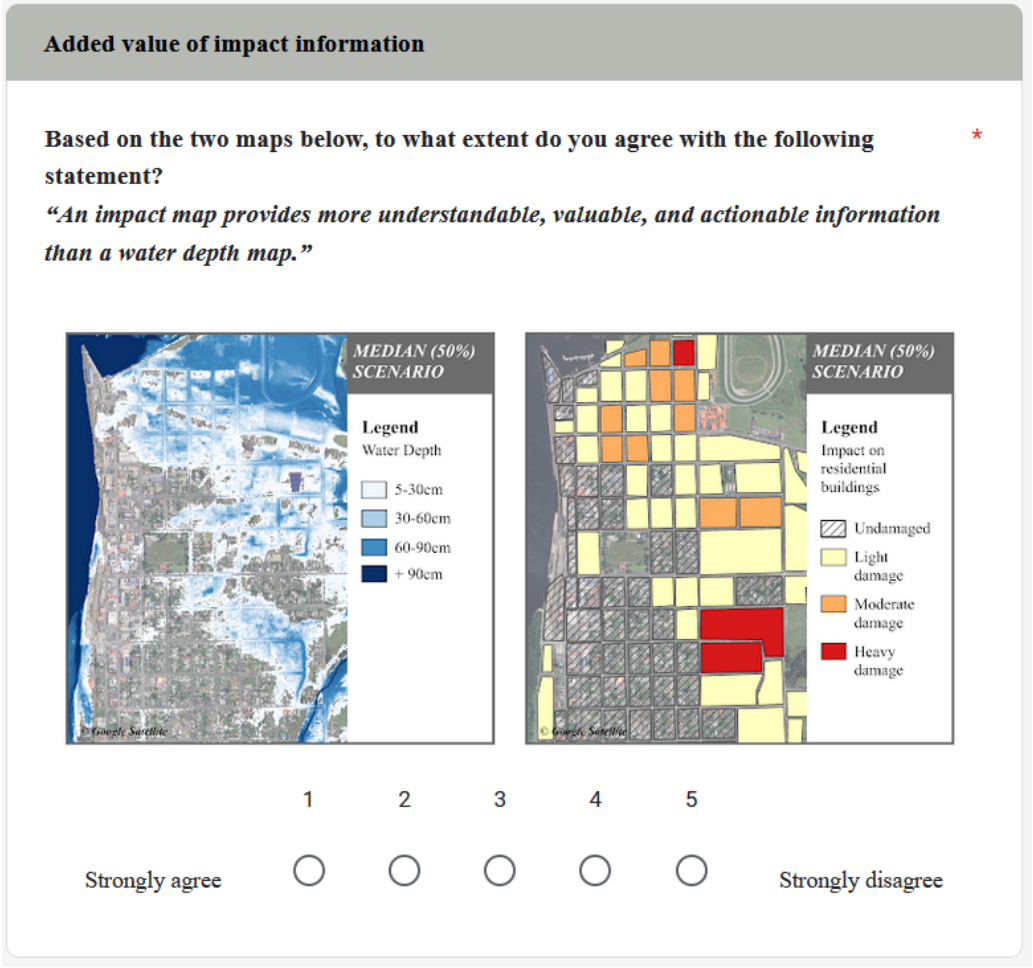


FIGURE F.21 – Second question of the survey and results obtained - different components of a flood warning

7.1.3 Added value of impact information



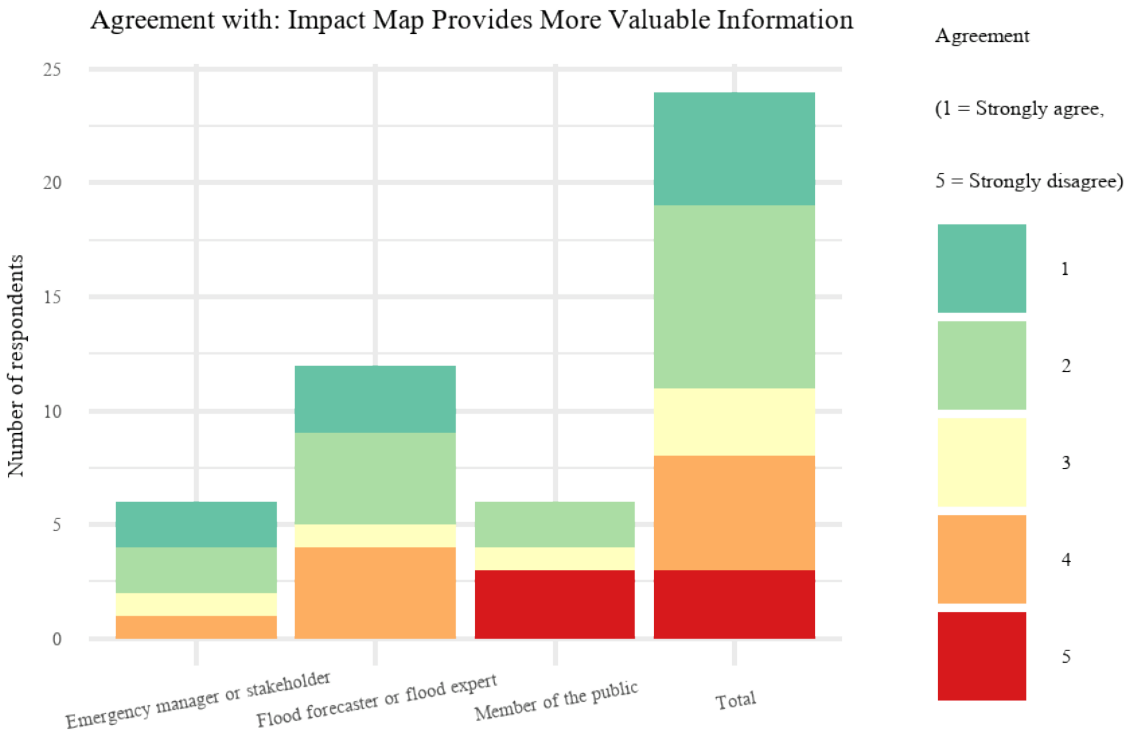


FIGURE F.22 – Third question of the survey and results obtained - added value of impact information

7.1.4 Trade-off between accuracy and familiarity

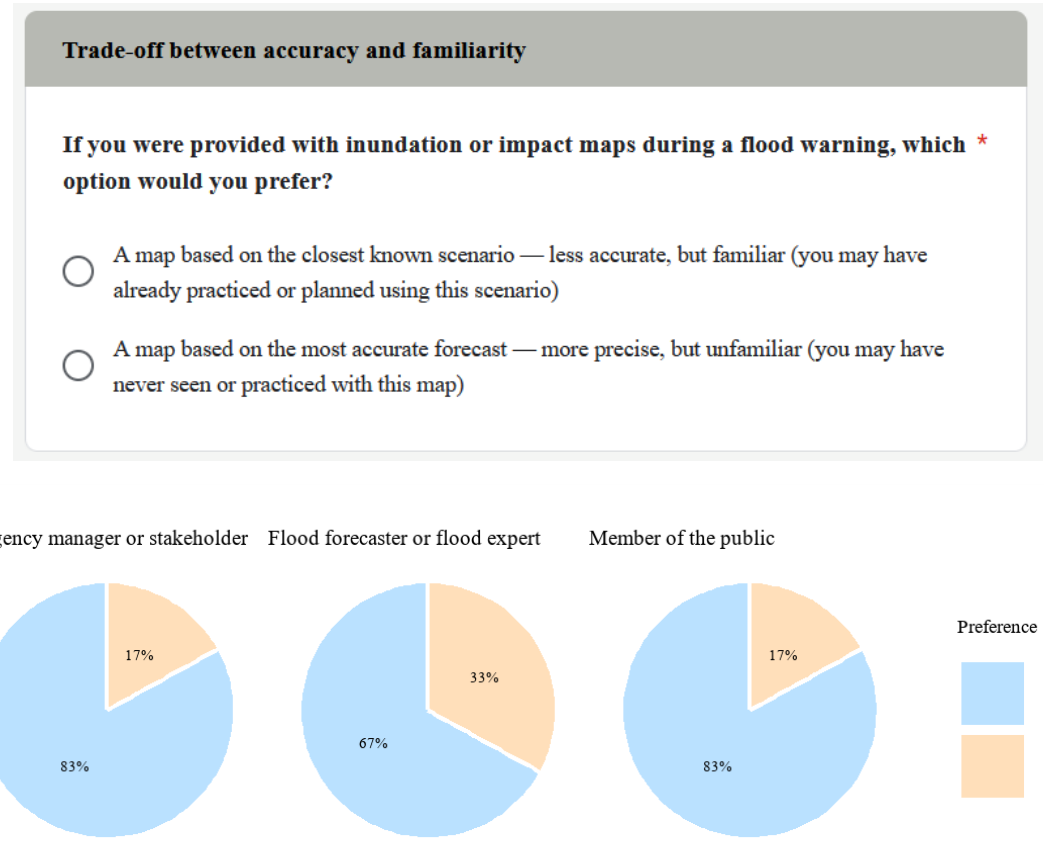
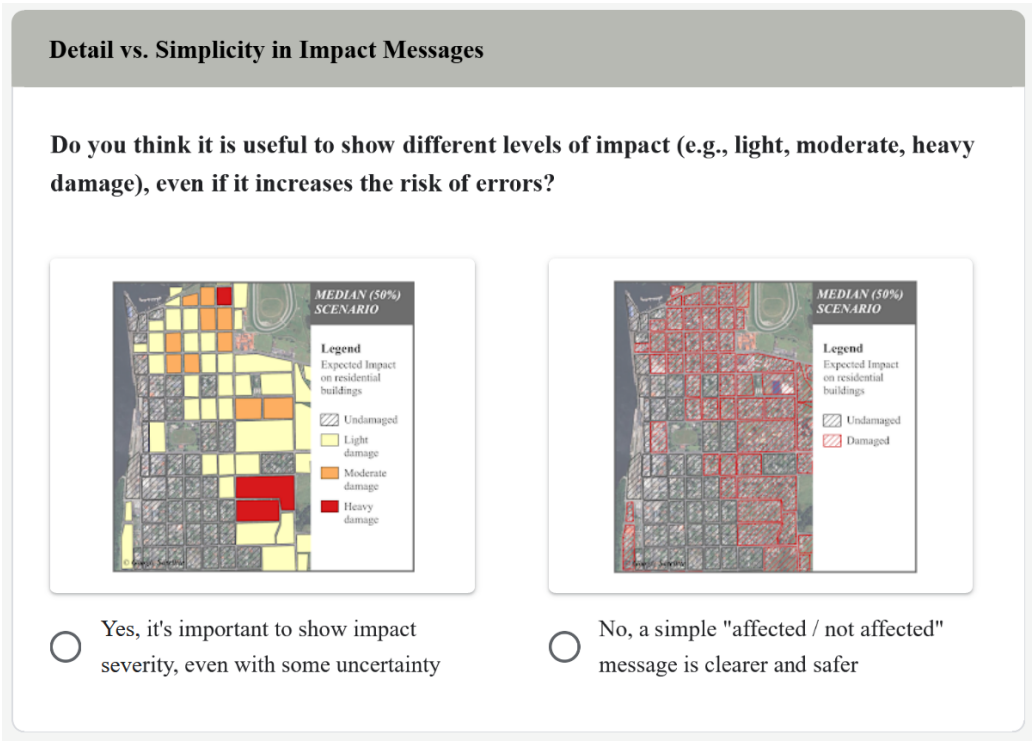


FIGURE F.23 – Fourth question of the survey and results obtained - trade-off between accuracy and familiarity

7.1.5 Detail vs. Simplicity in Impact Messages



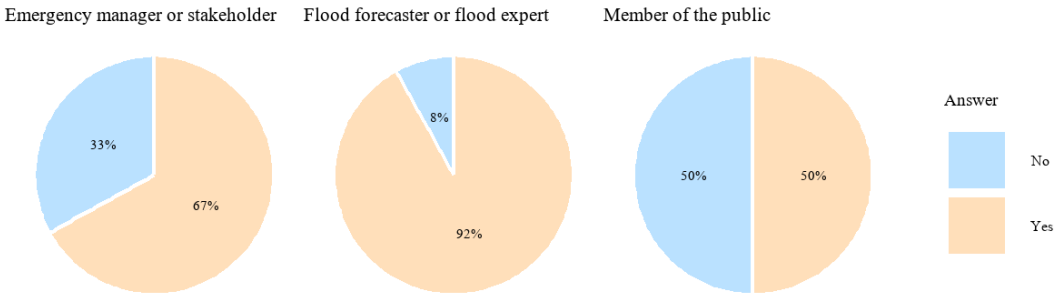


FIGURE F.24 – Fifth question of the survey and results obtained - Detail vs. Simplicity in Impact Messages

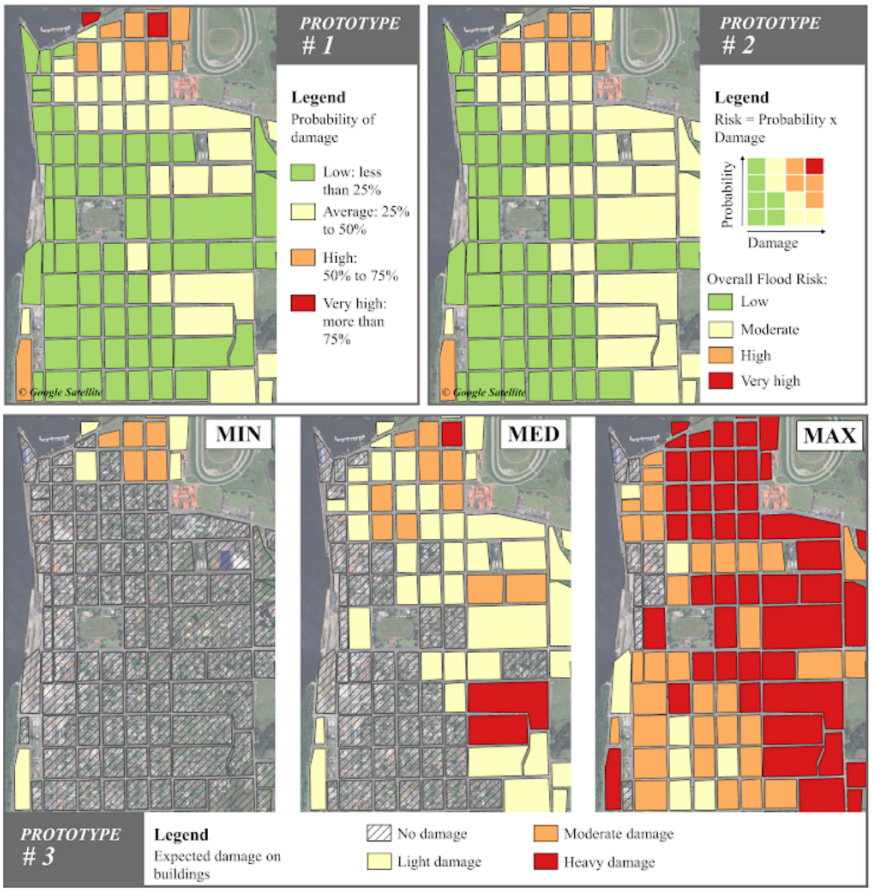
7.1.6 Representation of ensemble-based impact forecasts

Representation of ensemble-based impact forecasts

Ensemble forecasts, which offer multiple equally probable scenarios, are considered more valuable than **deterministic forecasts**,
as they better account for uncertainty. However, this added value often
comes at the cost of increased complexity in communication.

Below, we present different prototypes for communicating **ensemble-based impact forecasts**.

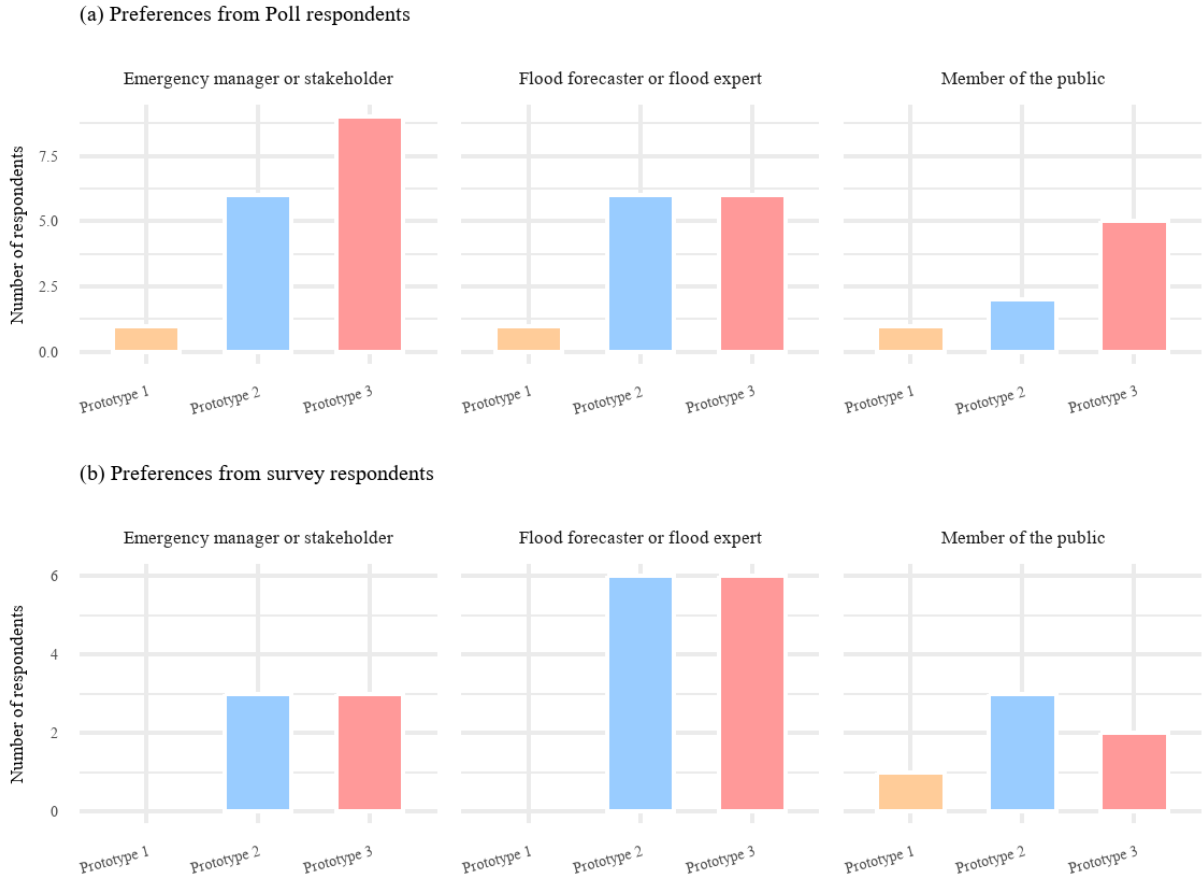
Based on the maps below, which prototype do you prefer for understanding and using impact information during a flood warning? *



- ○ Prototype 1: Shows the probability of any damage occurring to a building — the information is binary (damaged or not), without detail on severity.
- ○ Prototype 2: Shows the overall risk, combining both the probability and the severity of expected damage.
- ○ Prototype 3: Shows the expected damage severity for each block under three different scenarios — minimum, median, and maximum.

(Optional): Please explain why you chose this prototype.

Votre réponse



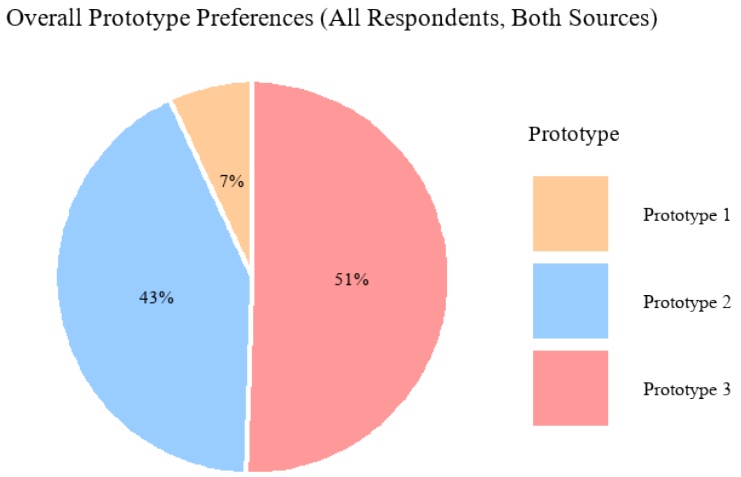
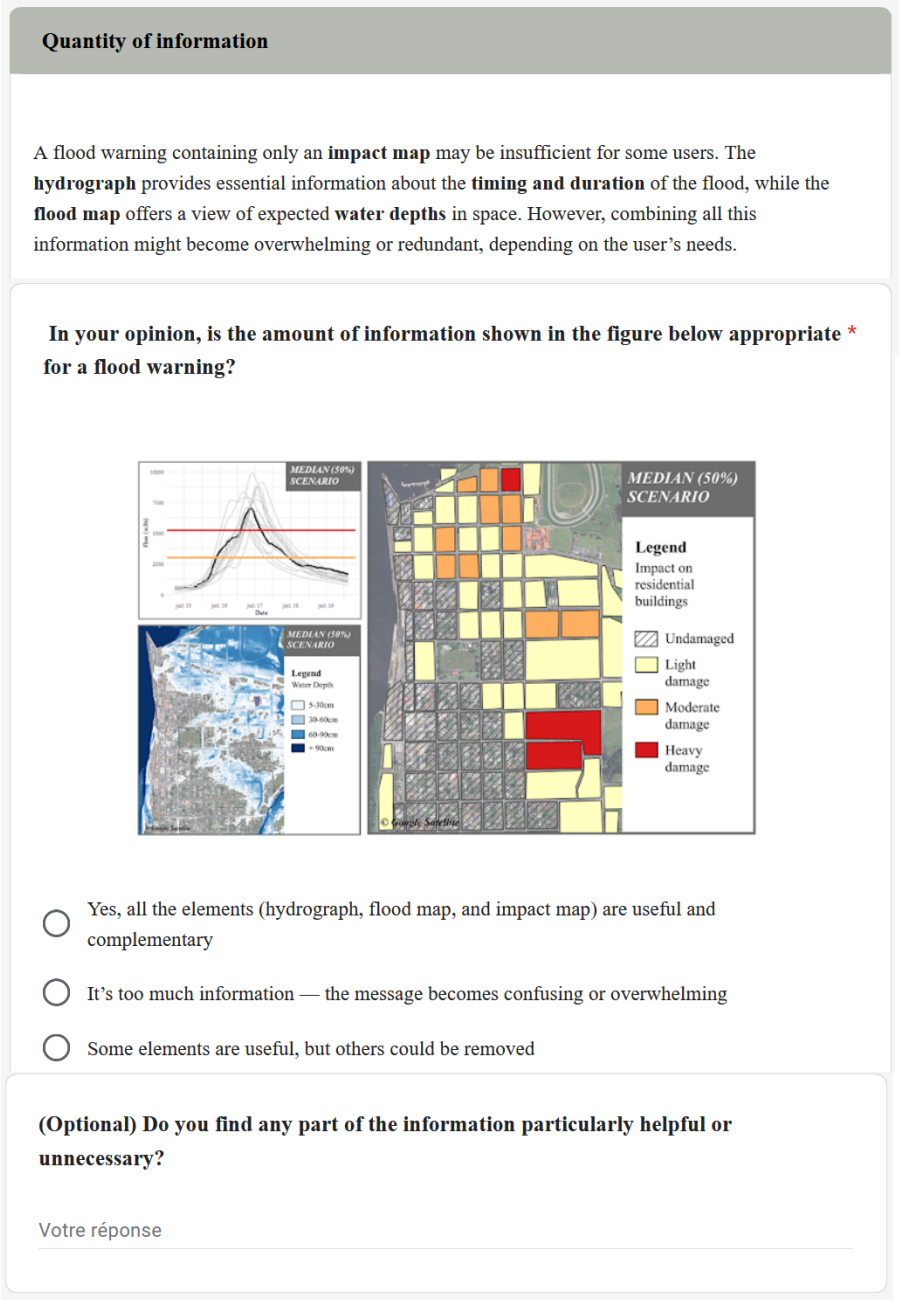


FIGURE F.25 – Sixth question of the survey and results obtained - Representation of ensemble-based impact forecasts

7.1.7 Quantity of information



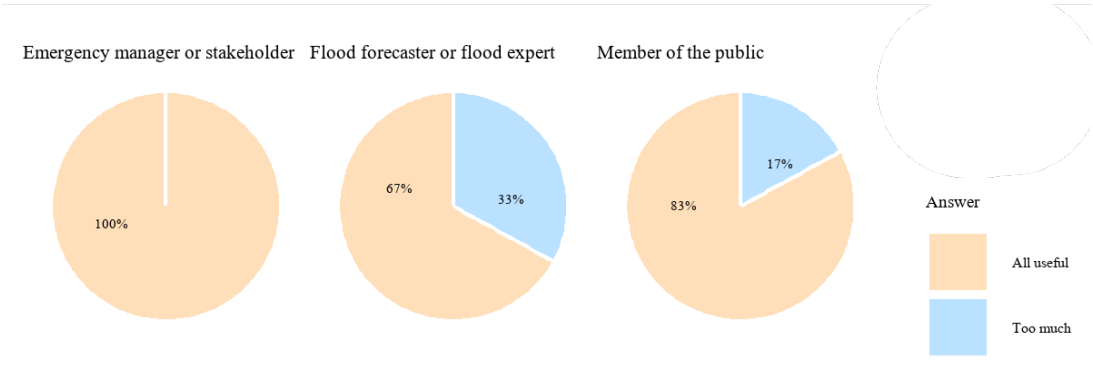


FIGURE F.26 – Seventh question of the survey and results obtained - Quantity of information

7.1.8 Types of impacts

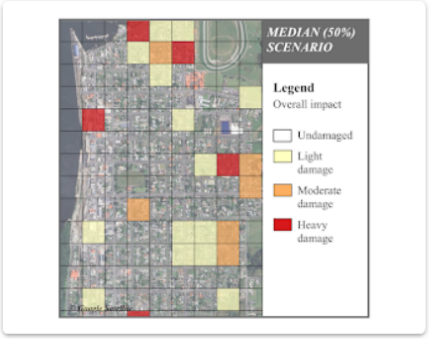
Types of impacts

In the previous examples, we focused exclusively on **residential buildings**. However, for emergency management and decision-making, knowing the expected impacts on **roads**, **community services**, **utilities**, or **commercial properties** can also be essential — depending on the user's needs.

Which format would you prefer for visualizing this information?



○ Separate maps for each impact category (e.g., one map for roads, one for services, etc.)



○ A single map combining all impacts, where categories are integrated and weighted based on their relative importance

(Optional) Can you explain why?

Votre réponse

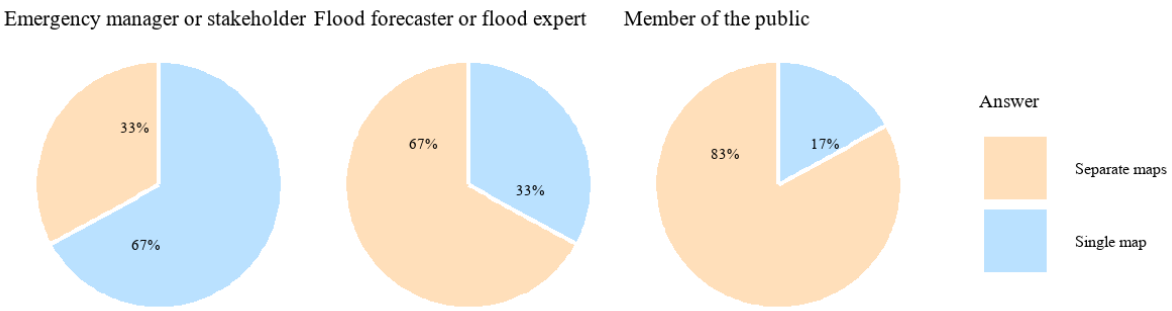


FIGURE F.27 – Eighth question of the survey and results obtained - types of impacts

7.2 Analysis of the results of the survey

Profile of respondents

During the webinar, the poll respondents were relatively well distributed across the three categories : members of the public (22%), emergency managers or stakeholders (43%), and flood forecasters or flood experts (35%). Among those who completed the follow-up survey (24 out of the 37 poll respondents), the majority identified as flood forecasters or flood experts (50%), while the remaining respondents were evenly split between the other two categories (25% each).

In the survey, participants were invited to specify their job titles. The "Flood forecaster or flood expert" category included roles such as River/Water Engineer, Catchment Manager, and Civil Engineer. Those identifying as "Emergency manager or stakeholder" included Project Managers, Risk Officers, and Analysts, with affiliations such as Auckland Emergency Management and the Climate Change Commission. The "Member of the public" category was more difficult to interpret, as the webinar targeted a relatively specialized audience. Nevertheless, six respondents selected this category, typically holding positions related to—though not directly involved in—flood response, such as Environmental Planner or Transportation Planner.

The different types of information in a flood warning, and the added value of the impact component

Questions 2, 3, and 7 explored the value of different components of flood forecasts—namely rainfall, flow, inundation, and impacts—to assess how respondents perceived the added value of impact-based information.

In response to the statement *"An impact map provides more understandable, valuable, and actionable information than a water depth map"* (Question 3), a majority of respondents agreed (13 out of 24). Interestingly, three respondents "totally disagreed"—all of whom identified as "Members of the public." This result is somewhat surprising, given the common assumption that impact maps are more accessible to non-experts. Conversely, disagreement was lowest among "Emergency managers or stakeholders" (1 out of 6), which aligns with expectations.

When asked to identify the most important type of information in a flood warning (Question 2), preferences varied across categories (Figure F.21). For flood forecasters and experts, rainfall was ranked highest, followed by flow, with inundation and impact information rated similarly. However, for both members of the public and emergency managers/stakeholders, inundation information was considered most important, followed by impact, then flow and rainfall.

Regarding whether presenting all three information types—hydrographs, flood maps, and impact maps—in a warning would be excessive (Question 7), most respondents agreed that the amount of information was appropriate. Nevertheless, 33% of flood forecasters or experts (4 out of 12) disagreed. Of those, two elaborated in optional comments : both believed that while flood maps and hydrographs are useful for experts or emergency managers, the general public would benefit more from a simplified message—preferably just the impact map, complemented by clear text describing timing and duration (rather than a hydrograph). One also suggested using the space to show minimum and maximum impact scenarios.

Interestingly, some respondents identifying as "Members of the public" expressed appreciation for flood maps, as they could easily relate them to their own location and assess personal risk, including danger to life based on water depth.

All emergency managers and stakeholders found the level of information appropriate. One noted that including multiple components helps communicate effectively with diverse audiences. Two respondents

highlighted the value of hydrographs over time, and another suggested adding historical context to the hydrograph for non-technical users. Finally, one respondent emphasized that, while building-focused impact maps are useful, they should be complemented with information about infrastructure impacts (e.g., on roads).

Challenges in communicating ensemble-based impact forecasts, managing multiple impact types and damage categories

This topic was explored in the final question of the survey, which asked respondents to choose between two approaches for presenting impact forecasts. The first option consisted of a single, integrated impact map combining various types of impacts (e.g., on buildings, roads, and services), with each type weighted according to its relative importance. This aggregated approach, adopted by Saint-Martin et al. [2016] in a Mediterranean case study, is widely used in the risk assessment community but remains less common in real-time forecasting contexts.

The second option proposed providing separate maps for each impact category. While this approach results in a larger volume of information, it is more reflective of the diversity of on-the-ground impacts.

Respondents' preferences varied by category. Members of the public and flood forecasters largely favoured the "separate maps" approach (83% and 67% respectively), whereas emergency managers and stakeholders tended to prefer the "single map" option. One emergency management respondent justified this choice by emphasizing the need to limit the number of documents consulted during an emergency. Another indicated that the integrated map was easier to interpret.

However, two other emergency managers/stakeholders preferred the "separate maps" approach, arguing that New Zealand's emergency response is multi-agency (e.g., local councils and NZTA for roads, utility companies for electricity and fibre), and that each agency requires information specific to their domain. This view was supported by several flood forecasters who also favoured disaggregated impact maps. Members of the public who preferred this approach commonly cited that it felt more tailored and allowed them to more clearly identify the consequences relevant to their situation.

Even when communicating impact forecasts focused on a single impact category (e.g. buildings), there are many ways of providing information. The easiest, and usually safest way is to provided binary information : a simple "affected / not affected" information for each building or area. Another option, which requires an additional level of precision, consists in providing different damage categories, e.g. "light"/"moderate"/"heavy damage". Question 24 asked the end-users whether they found this extra information useful, even at the cost of greater errors. The greatest agreement to usefulness was found with flood forecasters and flood experts (97%), then emrgency managers and stakeholders (67%), and member of the public were divided (50-50 %).

Finally, beyond the questions of which impact types to communicate and the level of detail to provide, representing impact-based forecasts becomes more complex when dealing with ensemble forecasts. Question 6 specifically addressed this issue by presenting respondents with three prototypes of ensemble-based impact forecasts, each based on a different communication approach :

- Prototype 1 showed the probability of an area being affected—based on a binary outcome. This is the most straightforward method for representing ensemble forecasts in a threshold-based framework : the probability of damage is calculated as the proportion of ensemble members forecasting damage for a given building or area.
- Prototype 2 was based on the Flood Risk Matrix used in the UK to communicate impact-based

forecasts [Cole et al., 2016], combining both the probability and the severity of damage.

- Prototype 3 illustrated uncertainty through min / median / max scenarios, without explicitly using probabilities. This approach was previously shown to be the most preferred for communicating flood maps [Jean et al., 2023].

This question was posed in both the live poll during the webinar and the longer survey. When combining responses from both sources across all respondent categories, a slight preference emerged for Prototype 3 (51%), confirming the earlier findings of Jean et al. [2023]. Prototype 2 followed closely behind (43%), while Prototype 1 was the least favored (7%).

Figures F.25a and F.25b show that preferences were fairly consistent across respondent categories. Flood forecasters and experts were equally split between Prototypes 2 and 3, whereas both emergency managers/stakeholders and members of the public showed a preference for Prototype 3. This again supports previous findings that probability-based representations are often misunderstood by non-technical audiences.

Among those who provided additional comments, respondents favouring Prototype 2 emphasized its clarity for the general public, the completeness of the information, and its usefulness in making critical decisions such as evacuation. Those who preferred Prototype 3 highlighted that it better reflects uncertainty, is easier to understand without a technical background, is familiar to some users (particularly flood forecasters), and offers more complete and visually intuitive information.

Trade-off between Accuracy and Familiarity

Although less directly linked to the previous questions, the trade-off between accuracy and familiarity in impact-based forecasts represents a key challenge in risk communication. Question 4 explored this issue by asking respondents whether they would prefer inundation and/or impact maps based on the closest known scenario (i.e., a map they may already be familiar with and have potentially practiced with), or a map based on the most accurate forecast available, even if it results in unfamiliar outputs.



French perspective : preference for known scenarios

This question resonates with the French experience from the ANR PICS project, where practitioners expressed a clear preference for a scenario-based approach. During group discussions with researchers, many highlighted the importance of using pre-established maps that could serve as references for predefined actions and planning. In contrast, initial feedback from New Zealand practitioners indicated some reluctance to rely on forecasts. However, as collaboration with researchers progressed, trust in the forecasting process grew, and the perceived value of accurate, even unfamiliar, forecasts increased accordingly.

Survey results aligned with this New Zealand perspective : respondents overall preferred to receive maps based on the most accurate forecast, even if they were unfamiliar. This preference was particularly strong among members of the public and emergency managers/stakeholders, suggesting a growing willingness to prioritise precision and up-to-date information over pre-existing familiarity.

8 Conclusions and perspectives

This internship enabled the application of an impact-based flood forecasting routine—similar to that developed during the PhD in France—to a New Zealand case study : the July 2021 flood event in Westport. By adapting an ensemble forecasting chain using NIWA's tools and data (including NZens rainfall forecasts, TopNet hydrological simulations, a precomputed flood library, and the RiskScape software), the study assessed both the potential and limitations of integrating real-time impact information into flood warnings.

Validation against placard data and insurance records helped to evaluate the contribution of downstream forecast components, investigate how uncertainty propagates through the modelling chain, and analyse the correlations between errors at different stages. In parallel, a user survey provided valuable insights into how impact-based products are interpreted and used by different audiences.

This case study offered a rare opportunity for high-resolution validation of impact forecasts and enabled the implementation of a full "cascading uncertainty" analysis across the entire forecasting chain—from rainfall to impacts—something that had not previously been attempted.

This conclusion highlights only the main points addressed in the report and is not intended to be exhaustive.

Several avenues of improvement and future research emerged from this work :

— Validation methods : Introduce additional metrics such as the Gerrity Skill Score (GSS) to assess multi-category impact forecasts.

— Variable thresholds : Explore the integration of building-specific floor height data to replace the fixed 50 cm water depth threshold currently used for damage detection, allowing for more accurate impact prediction.

— Transfer of user survey to France : Replicating the user survey with French stakeholders (including civil protection and the general public) could provide comparative insights into cross-cultural differences in how impact forecasts are perceived and used.

— Additional Westport case studies : Incorporate the February 2022 and August 2022 Westport flood events. Although impacts were limited, data on 6 placarded buildings in February could be used for partial validation. Proxy flood maps for both events could be derived by simulating inundation based on observed flows at Te Kuha.

— Cross-country analysis of cascading uncertainties : Apply the same correlation and divergence analysis used in this study to past French events (e.g., Argens 2010, Alpes-Maritimes 2015, Aude 2018). While PIAF-EPS forecasts are unavailable for these dates, a similar analysis could be performed on the November–December 2019 event, for which observed flood maps and impact data were collected.

These perspectives reflect both elements that were initially planned during the internship but could not be completed due to time or data constraints, and potential next steps that can reasonably be undertaken after the internship and before the end of the PhD. These actions aim to support the preparation of a potential publication based on this work.

COLLEGE DOCTORAL PAYS DE LA LOIRE
ED MATIERE, MOLECULES, MATERIAUX & GEOSCIENCES

Université Gustave Eiffel

Titre : Evaluation d'une chaîne de prévision immédiate des impacts des crues soudaines à l'échelle de l'arc méditerranéen français

Mots clés : crue soudaine, inondation, impacts, prévision, évaluation, système d'avertissement

Résumé : Les crues soudaines, qui touchent les petits cours d'eau, représentent une part importante des dommages liés aux inondations en France. Mieux les anticiper est essentiel pour en limiter les conséquences. Cette thèse construit et évalue une chaîne innovante de prévision immédiate des crues soudaines, intégrant une représentation des impacts attendus. La chaîne repose sur des outils existants de modélisation hydrologique et hydraulique, enrichis par des réflexions sur l'articulation des blocs de simulation. Elle a été appliquée sur environ 20 000 km de cours d'eau (bassins versants > 5 km²) des départements côtiers de l'arc méditerranéen français, sur une période continue de plus de dix ans (2010–2022). Dans un premier temps, la chaîne de simulation a été évaluée bloc par bloc, sur sa capacité à prévoir l'intensité des crues et à modéliser les zones inondées. Par la suite, la chaîne complète a été évaluée à l'aide de plusieurs jeux de données d'impacts : arrêtés CatNat, données assurantielles et interventions des services de secours. Enfin, la pertinence des avertissements basés sur l'intensité des crues, dans la logique du système opérationnel Vigicrues Flash actuel, a été comparée à celle d'avertissements basés sur les impacts. Les résultats ont montré que la prévision des impacts permet de réduire sigificativement le nombre de fausses alertes et d'améliorer la hiérarchisation des zones impactées.

Title: Evaluation of an impact-based nowcasting chain for flash floods across the French Mediterranean region

Keywords: flash flood, flood, impacts, forecasting, evaluation, early warning system

Abstract: Flash floods in small catchments account for a substantial share of flood-related damages in France, making accurate anticipation crucial to reducing their impacts. This PhD research develops and evaluates an innovative nowcasting chain for flash floods, with a particular focus on predicting expected impacts. The chain builds on existing hydrological and hydraulic modelling tools, complemented by methodological work on integrating the different simulation components. The study area covers ~20,000 km of river network (contributing areas > 5 km²) across the coastal departments of the French Mediterranean arc, and the chain was operated continuously over a hindcast period of more than ten years (2010–2022). First, the simulation chain was evaluated component by component, focusing on its ability to predict flood intensity and to model inundated areas. Then, the complete chain was assessed using several impact datasets: CatNat declarations, insurance data, and emergency services interventions. Finally, the effectiveness of hazard-based warnings, which focus solely on flood intensity, in line with the operational Vigicrues Flash system, was compared with that of impact-based warnings (IBWs). The results show that IBWs significantly reduce false alarms and improve the detection of the most severely affected areas.