THÈSE DE DOCTORAT DE L'UNIVERSITE MARIE ET LOUIS PASTEUR

École doctorale n°554

Environnement-Santé

Doctorat en Structure et évolution de la Terre

Par

Olivier Béjean-Maillard

SIMULATING SLOW-MOVING LANDSLIDES WITH AI: MULTI-SOURCE DATA, HYDRO-METEOROLOGICAL FORCINGS, AND INTERPRETABILITY

Thèse soutenue à Besançon, le 11 Mars 2026

Composition du Jury :

M. Thom Bogaard	Professeur, Univ. de Delft, Pays-Bas	Rapporteur
M. Filippo Catani	Professeur, Univ. de Padove, Italie	Rapporteur
Mme Anne Johannet	Professeure, IMT Mines Alès	Présidente / Examinatrice
Mme Séverine Bernardie	Ingénieure, BRGM Orléans	Examinatrice
Mme Héloïse Cadet	Ingénieure, SAGE INGENIERIE	Examinatrice
M. Clément Hibert	Physicien adjoint, Univ. de Strasbourg	Examinateur
Mme Catherine Bertrand	Professeure, Univ. Marie et Louis Pasteur	Directrice de thèse
M. Jean-Philippe Malet	Directeur de recherche CNRS, Univ. de Strasbourg	Co-directeur de thèse

THÈSE DE DOCTORAT DE L'UNIVERSITE MARIE ET LOUIS PASTEUR

École doctorale n°554

Environnement-Santé

Doctorat en Structure et évolution de la Terre

Par

Olivier Béjean-Maillard

SIMULATION DES GLISSEMENTS DE TERRAIN LENTS À L'AIDE DE L'IA : DONNÉES MULTI-SOURCES, FORÇAGES HYDRO-MÉTÉOROLOGIQUES, ET INTERPRÉTABILITÉ

Thèse soutenue à Besançon, le 11 Mars 2026

Composition du Jury :

M. Thom Bogaard	Professeur, Univ. de Delft, Pays-Bas	Rapporteur
M. Filippo Catani	Professeur, Univ. de Padove, Italie	Rapporteur
Mme Anne Johannet	Professeure, IMT Mines Alès	Présidente / Examinatrice
Mme Séverine Bernardie	Ingénieure, BRGM Orléans	Examinatrice
Mme Héloïse Cadet	Ingénieure, SAGE INGENIERIE	Examinatrice
M. Clément Hibert	Physicien adjoint, Univ. de Strasbourg	Examinateur
Mme Catherine Bertrand	Professeure, Univ. Marie et Louis Pasteur	Directrice de thèse
M. Jean-Philippe Malet	Directeur de recherche CNRS, Univ. de Strasbourg	Co-directeur de thèse

REMERCIEMENTS

Lors de la rédaction de mon mémoire de fin d'études, je me souviens avoir lu plusieurs manuscrits de thèse pour nourrir ma bibliographie, et avoir ressenti une émotion particulière en parcourant les paragraphes consacrés aux remerciements. À travers ces mots, je percevais l'aboutissement d'une longue aventure scientifique et humaine, riche d'apprentissages, d'anecdotes et de reconnaissances, et j'espérais, un jour, atteindre à mon tour cet accomplissement. C'est aujourd'hui à moi d'ajouter une pierre à l'édifice, et les quelques paragraphes qui suivent témoignent de ma gratitude tout en traduisant fidèlement les sentiments que j'éprouvais alors, quelques années plus tôt.

Ces premiers mots sont adressés à mes directeurs de thèse, Catherine Bertrand et Jean-Philippe Malet, qui m'ont embarqué avec eux dans cette grande aventure sur les pentes instables des glissements de terrain.

Merci Catherine d'avoir accepté de prendre ce jeune étudiant de L3 en stage de recherche, malgré les conditions sanitaires de 2020 rendant la tâche particulièrement complexe. Ces quelques semaines de terrain à Viella m'ont fait découvrir le monde des risques gravitaires et le rôle de l'hydrogéologie dans ces aléas, et, au final, n'ont été que les prémices de cette thèse ! Une aventure marquée par cette fameuse question autour d'un (ou plusieurs) café(s) : « Mais pourquoi Séchilienne bouge toujours ? », de glissades sur l'A39, ou de sushis partagés à 22 h à Pontarlier au retour des JAG.

Merci Jean-Philippe d'avoir co-dirigé cette thèse avec Catherine. Notre première rencontre date aussi de Viella, et tu as su m'introduire et me guider dans le monde des data sciences et du machine learning, complètement inconnus à mes débuts en thèse. Nous avons fait de nombreux terrains ensemble, qui sont autant de sources d'anecdotes mémorables. Des Alpes aux Pyrénées, nous nous sommes retrouvés embourbés sur les pistes forestières de la vallée d'Aspe, à court d'AdBlue vers la crête de Soleilboeuf (il ne faut jamais être à court d'AdBlue, c'est toujours bon à savoir), en pleine épopée pour rejoindre la galerie Catherine à Séchilienne, ou spectateurs par hasard d'un concert de saxophone à Chamrousse, et j'en passe ! À tous les deux, merci pour ces années de thèse passées à vos côtés.

Mes remerciements s'adressent également aux membres de mon jury. Merci à mes rapporteurs, Thom Bogaard et Filippo Catani, pour leur intérêt et leur évaluation attentive de mon manuscrit, ainsi qu'à mes examinateurs, Anne Johannet, Séverine Bernardie, Héloïse Cadet et Clément Hibert, pour nos échanges enrichissants et instructifs lors de la soutenance.

Merci aux membres de mon comité scientifique de thèse, Michel Jaboyedoff et Clément Dombry, pour tous nos échanges et vos conseils. Merci à tous mes collaborateurs qui ont participé à l'aboutissement de ce travail, et avec qui j'ai eu grand plaisir à échanger : Guillaume Cinkus, Naomi Mazzilli, Pierre Nevers, Laurent Dubois, Davide Bertolo, Patrick Thuegaz, Alessandro Corsini, Mélissa Tondo, José Moya, Marcel Hürlimann, Jérôme Latron, Olivier Maquaire, Mathieu Fressard et Mario Rega.

Merci à l'ensemble des membres du thème GEODE pour leur dynamisme, leur bienveillance et leur sympathie. J'ai pu vous côtoyer d'abord comme étudiant, puis comme collègue, et cela a toujours été un plaisir d'échanger avec vous. Merci à Sophie Denimal pour avoir été là dès les répétitions pour le concours de l'ED ; nous pourrions parler des heures autour d'un thé et de quelques biscuits. Merci à Henri Leclère pour ta bonne humeur que tu apportais à chaque fois que tu rentrais dans le bureau. Merci aux teams SNO-Karst et/ou EPOS : Marc Steinmann, Christophe Loup, Anne Boetch et Julie Albaric, pour votre bonne humeur durant ces workshops. Merci à Blandine Delaby d'avoir veillé à ce que l'on aille toujours bien, surtout dans les périodes de rush. Je tiens à remercier les équipes de Chrono-Environnement : Sébastien Landry, Yvon Thiriot, Valmir Jashari, Najat Ouhaik, Stéphanie Jubin, Nathalie Puillet, Estelle Franc, Estelle Camizuli, Charles-Henri Falconnet, Kevin Van Keulen et Nicolas Carry pour votre efficacité permanente.

REMERCIEMENTS

Merci à la team de l'EOST pour nos échanges et pour les moments de terrain que nous avons partagés ensemble : Floriane Provost, Julien Jean-Baptiste, Florian Bourcier et Yann Smith. Bastien, mon acolyte fan de MrJDay, j'espère que ce jambon d'Aoste qui n'est pas d'Aoste était bon ! Bon courage pour la fin de ta thèse, on pourra bientôt s'appeler docteurs mutuellement.

Merci aux stagiaires et étudiants que j'ai pu encadrer ou côtoyer pour cette thèse : Eloïse Granier (bon courage pour ta thèse !), Noémie Schuck (quel trio avec Josh pour aller à « Séssilienne »), Latte Lawson et Felix Quéré. Felix, merci pour ta sympathie, ton sérieux et ton investissement pour le modèle de La Saxe. Je te souhaite tout le meilleur pour ta carrière professionnelle.

Je me dois de dédier quelques mots à deux chaînes YouTube : StatQuest et Machine Learnia. Nous vivons aujourd'hui dans un monde où le partage des connaissances est largement facilité par les réseaux, et certains créateurs font preuve d'une pédagogie remarquable qu'il me semble important de saluer. Je ne peux qu'encourager tout débutant dans le domaine du machine learning et des statistiques à aller découvrir leur travail.

Merci à mes co-bureaux : Kilian, Alban, Antonin, Mario, Kevin, Tiphaine et Noémie, sans qui les journées de recherche auraient paru bien longues. La -227L est marquée à jamais par notre jeu de la carotte (encore bravo Kevin). Merci en particulier à Kilian pour toutes ces doses de rires, à toutes heures de la journée, jusque tard en soirée, à transiter entre blagues douteuses, bruits d'oiseaux et notre générique revisité de Monsieur Flap. Merci Mario pour ton optimisme constant, ta gentillesse et ton humour (douteux), auquel j'ai toujours pris un grand plaisir à contribuer. Nos nombreuses œuvres d'art me manqueront, mais j'espère que ce fameux « Hooon » demeurera un souvenir impérissable. Merci Kevin pour ta bonne humeur, ton soutien et tes conseils. Tu as été la force tranquille de ce bureau, une très bonne rencontre, et une grande source d'inspiration. Ton jumeau en carton restera la star de ce bureau.

Ces remerciements ne pourraient être complets sans parler de Josh, mon acolyte et MacGyver de tous les terrains. Trois années de péripéties vécues à deux, des centaines de kilomètres partagés, et une quantité innombrable d'anecdotes à raconter. Le traditionnel « Highway to Hell » pour rouler jusqu'à G710, la pizzeria de Vizille et ses fameux arancini, ou notre bon vieux VW, tout cela me manquera. Certains disent même que le glissement de Séchilienne s'attristera de ne plus nous voir arpenter ses sentiers.

Comment ne pas remercier les stagiaires, doctorants et post-doctorants du labo avec qui j'ai partagé quotidiennement ma vie durant trois ans : Anthony, Camille, Charles, Claire, Elisa, Fleur, Heikel (bonjour d'ailleurs), Jolan, Jonathan, Jordan, Justine, Léa, Manon (géol), Manon (racoon), Marie, Milva, Neven, Solofo et Tristan. Merci pour tous ces moments, ces pauses-café, ces parties de tarot, de yanniv (ou Yann-Yves), et ces potins partagés.

Colette, merci pour toutes ces discussions insolites et (presque) quotidiennes au RU. Rappelle-moi de ne plus jamais jouer au Traître à bord avec toi, sinon je me sentirais obligé de dire que la Maurienne surpasse de loin la Tarentaise. Laurie, ou devrais-je dire le « Ja Kid », merci d'avoir cru avec moi qu'une année zen était possible. Elle n'aura duré que quelques jours, mais l'illusion était là. Europapark doit encore nous entendre crier « Jaaaa »…

Bien sûr, je ne peux faire des remerciements convenables sans parler de ma dream team de la -219K, à commencer par Lise. Je ne pourrai pas modifier ce texte, donc pour une fois je vais être gentil, mais je t'invite au moins à le lire en québécois. Merci énormément pour toute ta bonne humeur, tes histoires improbables, ou pour ton humour décalé qui me sied tant. J'espère que tu feras à nouveau appel à moi, à mon sens de la narration (et mon charisme), pour te raconter l'histoire des attractions d'Europapark ou la biographie de Bernard Tapie, incanter Véronique Sanson dans ce beau chalet de Maurienne, ou rendre visite à Sam Gamgie dans le petit port de Polruan (par exemple). Nous avons, j'espère, encore beaucoup d'aventures à vivre autour d'ateliers aquarelle et de randonnées.

REMERCIEMENTS

Évidemment, merci sincèrement à Lorraine, ma bestie, l'âne soubirant, mon fidèle binôme de streaming, ou la collègue idéale pour notre future carrière d'ASVP. Tu as été la meilleure rencontre que j'ai pu faire durant cette thèse, un soutien essentiel jusqu'au dernier jour, et je me remercie encore de vous avoir proposé, avec Alexis, cette sortie aux cascades du Hérisson. Aujourd'hui, nous sommes tous les deux docteurs, mais nous pouvons nous satisfaire d'un autre diplôme, celui de Geoguessr pro. Sans voyager, les routes d'Afrique du Sud, les ruelles étroites de Montevideo, ou les plaines sèches de Mongolie n'ont plus aucun secret pour nous. Nos jeudis après-midi au Colombus, nos victoires au Toc, ou nos sorties par monts et par vaux me manqueront, mais ils ne marquent que la fin d'un chapitre, et le début d'une longue histoire à venir.

Merci à tous mes amis qui m'ont soutenu jusqu'à l'aboutissement de cette thèse. J'ai énormément de chance de tous vous avoir. En particulier, merci à Tristan, Snipe ou Sparcky, d'être devenu mon frère de cœur depuis le collège. Nous en avons vécu, des histoires complètement folles ensemble, et celle-ci n'en sera qu'une parmi tant d'autres. Comme un vieux sage l'a dit : « À chaque âge ses épreuves. » Ta présence encore aujourd'hui m'honore. Ben et Thomas, merci pour ce trinôme exceptionnel que nous formons depuis la L1. Ben, souviens-toi quand nous parlions de la thèse dès nos premiers terrains de géol, c'est chose faite. J'aurais bien commencé ma présentation par « Les hommes ont toujours eu besoin d'énergie », mais je suis sûr qu'au fond de toi, tu l'entendais. Thomas, nous avons fait un sacré bout de chemin ensemble, des photo-canis canis aux carrières de Caumont, en passant par le Trappiste, en compagnie d'Anne, Hugo, Théo et Nico, le hasard a toujours fait en sorte que nous nous retrouvions, et pour le mieux !

Claire, merci d'avoir été là tout du long. Je nous vois encore monter tes meubles avant de céder sans résistance pour un détour chez Oktay. Captain Brigsy frémit à l'idée de nous revoir. Lucie, tu as été une très belle rencontre durant cette thèse. Merci pour ton optimisme et ton soutien, ces barres chocolatées étaient vraiment bonnes pour le moral. Clément, il est loin le temps des courses sur Mario Kart DS, et pourtant notre amitié est toujours la même aujourd'hui. Encore bravo pour ton agrégation, et merci d'être présent à mes côtés depuis si longtemps. Molina, tu es quelqu'un d'exceptionnel. J'ai adoré passer tous ces moments autour d'une bière à tes côtés, à discuter des heures la nuit tombée. Le road trip en Auvergne, les Molinanniversaires, ou l'introduction aux Molipotes (big up au vieux Joe), merci pour tout. Bien sûr, merci aux chiengs, Anto et Hugo. Que dire à part que l'on ne peut être mieux entouré que par deux idiots comme vous. Merci pour toutes ces doses de rire que vous m'avez apportées. Ces soirées à jouer ensemble ont été salvatrices. En ce qui concerne la suite de nos périples, je ne vois qu'une seule chose à dire : « Aaaaah I see something ahead. »

Je souhaite tout particulièrement remercier ma famille, et surtout mes parents, Christine et Patrick. Si je suis arrivé jusqu'ici, c'est entièrement grâce à vous. Merci d'avoir alimenté ma curiosité dans tous les domaines, merci de m'avoir aidé à apprendre les identités remarquables un après-midi entier, merci de m'avoir orienté vers la géologie, merci d'avoir toujours été là pour m'écouter, m'aider à réviser, et d'avoir toujours su apporter un soutien sans faille. Je vous suis à jamais redevable. Merci à ma sœur, Audrey, pour tous ces délires et ces sorties. Le hasard a voulu que nous finissions nos études la même année ! Tu es une grande sœur exceptionnelle, ne change jamais. Merci à ma belle-famille, les Béjean, et quelle famille ! Je crois que je ne pouvais pas mieux tomber. Je suis très heureux que vous fassiez à présent partie intégrante de la mienne.

Enfin, merci à toi Perry, pour ton amour et ton soutien indéfectibles. Je te suis infiniment reconnaissant pour tout ce que tu m'as apporté depuis plus de huit ans. Nous en aurons vécu des péripéties durant cette thèse, des très heureuses comme des plus tristes, mais nous les avons toujours traversées à deux. Merci d'être comme tu es : rayonnante, captivante, créative, attentionnée, dégourdie, maladroite, même parfois brutus. Merci d'être devenue ma femme, c'est le plus beau cadeau que je pouvais avoir. Continuons à vivre toutes ces prochaines aventures qui nous attendent. Je t'aime ma patate. Ton docteur Rocks.

TABLE OF CONTENTS

Remerciements ... **1**
Table of contents ... **5**
List of figures ... **7**
List of Tables ... **12**
CHAPTER I GENERAL INTRODUCTION ... **13**
I.1. Context and objectives of the thesis ... **15**
I.2. Structure of the manuscript ... **19**
CHAPTER II SLOW-MOVING LANDSLIDES: FROM HYDRO-MECHANICAL PROCESSES TO AI-BASED DISPLACEMENT MODELLING ... **23**
II.1. Introduction to natural hazards ... **25**
II.2. Definition, classification, and controlling factors of slope movements ... **27**
II.3. Structural, geomorphological, and hydrological controls of slow-moving landslide dynamics ... **33**
II.4. Data collection methods for monitoring slow-moving landslides ... **36**
II.5. Simulating models for Slow-moving landslide displacement ... **39**
II.5.1. Physically-based models ... 39
II.5.2. Data-driven models ... 40
II.6. Chapter II – Conclusion ... **49**
CHAPTER III CONVENTIONAL MACHINE LEARNING FRAMEWORK FOR SIMULATING SLOW-MOVING LANDSLIDE DISPLACEMENTS ... **51**
III.1. Chapter III – Introduction ... **53**
III.2. Data preparation and exploratory analysis of field monitoring time series ... **54**
III.2.1. Time series collection and preparation ... 54
III.2.2. Exploratory statistical analysis ... 55
III.3. XGBoost modelling strategy steps ... **56**
III.3.1. Feature engineering ... 56
III.3.2. Train-test splitting ... 57
III.3.3. Hyperparameter tuning ... 57
III.3.4. Model performance evaluation ... 59
III.4. Model interpretability and explainability ... **60**
III.4.1. XGBoost feature importance ranking ... 61
III.4.2. SHAP post-hoc analysis ... 62
III.5. Chapter III – Conclusion ... **64**
CHAPTER IV RAINFALL-DISCHARGE-BASED MODELLING TO RECONSTRUCT LANDSLIDE GROUNDWATER LEVEL PROXIES ... **65**
IV.1. Chapter IV – Summary ... **67**
IV.2. Paper - Hydrogeological forecasting of deep-seated landslides dynamics: structure and sensitivity of tank models (Landslides, 2025) ... **70**
IV.2.1. Introduction ... 70
IV.2.2. Study site presentation and data collection ... 72
IV.2.3. Modelling methods ... 81
IV.2.4. Results ... 88
IV.2.5. Discussion ... 96

TABLE OF CONTENTS

IV.2.6. Conclusion 98

IV.3. Chapter IV - Conclusion 100

CHAPTER V FROM HYDRO-METEOROLOGICAL SERIES TO INTERPRETABLE VELOCITY SIMULATIONS: A XAI-BASED WORKFLOW 101

V.1. Chapter V – Summary 103

V.2. Paper – A Scalable XAI-based Forecasting of Landslide Velocities From Environmental Forcings (Nature Geosciences, submitted) 106

V.2.1. Introduction 106

V.2.2. Multi-site monitoring datasets 107

V.2.3. Explainable Artificial Intelligence framework 110

V.2.4. Model performance: simulated landslide velocities 118

V.2.5. Interpretability and explainability: environmental predictor controls 121

V.2.6. Conclusion: Implications for interpreting landslide dynamics with XAI framework 125

V.2.7. Extended data 127

V.3. Chapter V – conclusion 132

CHAPTER VI EXTENDING THE AI-BASED WORKFLOW TOWARDS PROBABILISTIC DISPLACEMENT SIMULATIONS 133

VI.1. Chapter VI – Summary 135

VI.2. Paper - Probabilistic Simulation of Monthly Landslide Velocity Under Hydro-meteorological Variability (in preparation) 137

VI.2.1. Introduction 137

VI.2.2. Methods 140

VI.2.3. Study sites and model calibration 145

VI.2.4. Results 154

VI.2.5. Discussion 160

VI.2.6. Conclusion 163

VI.2.7. Extended data 165

VI.3. Chapter VI – Conclusion 168

CHAPTER VII CONCLUSION AND OUTLOOKS 169

VII.1. General conclusion 171

VII.2. Discussion and outlooks 173

VII.2.1. Extending the methodology to other deep-seated landslides 173

VII.2.2. Using satellite products to address data gaps 174

VII.2.3. Towards spatio-temporal prediction of surface deformation velocities 175

VII.2.4. Advances in interpretability and explainability of AI models for time series 176

VII.2.5. Open-access contribution 178

Références bilibographiques 179

Appendix – Résumé long 197

Annexes 219

LIST OF FIGURES

Figure 1 | Main application domains of artificial intelligence in landslide science and hazard mitigation. 16

Figure 2 | Graphical abstract summarising the thesis chapters, their connections, and the associated research questions. 21

Figure 3 | Global landslide hazard worldmap, representing combined rainfall-triggered landslide hazard and earthquake-triggered landslide hazard. Hazards is defined with three categories, ranging from Low (yellow) to High (red) (from *Redshaw et al., 2020*)..... 26

Figure 4 | (a) Forces involved in the destabilisation of landslides (adapted from *Vallet, 2014*). **(b)** Stresses acting along a potential failure surface (adapted from *McColl, 2015*). The light-grey area denotes the unstable mass. 27

Figure 5 | Schematic representation of the most common types of landslides (adapted from *Highland & Johnson, 2004*). 29

Figure 6 | Hydro-mechanical coupling in a fractured medium. (i) and (ii) corresponds to direct couplings related to pore volume variations (short-term), while (iii) and (iv) corresponds to indirect couplings related to material properties changes (medium- and long-term) (adapted from *Rutqvist & Stephansson, 2003*)..... 31

Figure 7 | Displacement and groundwater levels measured in two boreholes on the Vallcebre landslide from November 1996 to December 2015. Red dashed lines illustrate the temporal relationship between groundwater level changes and acceleration phases recorded in borehole S2 (adapted from *Moya et al., 2017*)..... 32

Figure 8 | Typical morphological structures of deep-seated deformations (from *Agliardi et al., 2013*). 34

Figure 9 | Conceptual model of groundwater flows in unstable rock slopes, derived from the analysis of hillslopes in the Upper Tinée Valley (South-East French Alps) (from *Binet et al., 2007*)..... 35

Figure 10 | Synthetic representation of landslide monitoring data acquisition methods and displacement modelling approaches. Green and red boxes summarise the main positive and negative aspects of each methods..... 38

Figure 11 | Schemes for the structure description of the artificial neural networks (ANNs): **(a)** Description of neural network architecture. **(b)** Description of neural network active node (adapted from *Bekesiene et al., 2021*)..... 41

Figure 12 | Traditional versus Recurrent Neural Networks (adapted from *Kıraç et al., 2019*). 42

Figure 13 | Structures of a traditional RNN, a LSTM and a GRU modules (adapted from *Liu et al., 2020*). 43

Figure 14 | Structure of a decision tree. 45

Figure 15 | Schematic illustration of **(a)** Bagging and **(b)** Boosting workflows (adapted from *González et al., 2020).* 46

Figure 16 | Conceptual representation of XGBoost functioning in a regression task. □ represents the learning rate applied to each output of a regression tree. 48

Figure 17 | Workflow for time-series processing and conventional machine-learning model implementation. 53

Figure 18 | Conceptual representation of the method for extracting descriptive features from rainfall and groundwater level time series over multiple time windows..... 56

Figure 19 | Chapter IV – Schematic overview of the discharge-driven strategy used to reconstruct groundwater-state proxies. 68

Figure 20 | (a) Localisation of the Séchilienne and Mont de La Saxe landslides. **(b)** The Mont de La Saxe and **(c)** the Séchilienne landslides are represented by a yellow surface, along with meteorological stations (blue circles) and discharge stations (green squares) localisations..... 73

Figure 21 | (a) Shaded 1-meter DEM of the right flank of the Romanche Valley with localisation of the landslides, fractures, streams, springs and measurements stations. **(b)** Schematic cross-section of the hydrogeological functioning of the Séchilienne landslide (adapted from *Vallet et al., 2014*). **(c)** Daily rainfall and discharge time series acquired at the Mont Sec weather station and the G1 outlet. 75

Figure 22 | (a) Shaded 1-meter DEM of the left flank of the Ferret Valley with localisation of the landslides, fractures, streams, springs and measurements stations (adapted from Crosta et al., 2013). **(b)** Schematic cross-section of the Mont de La Saxe landslide (adapted from *Crosta et al., 2013*). **(c)** Daily rainfall and discharge time series acquired at the Mont de La Saxe weather station and the MS02 spring. 78

Figure 23 | Several cases of flow duration curves and their hydrodynamic interpretation (adapted from *Marsaud, 1996*). 79

LIST OF FIGURES

Figure 24 | Flow duration curve based for the **(a)** Séchilienne and **(b)** Mont de La Saxe landslides. Slopes are represented with alpha, slope breaks are shown with dotted lines and associated discharge values, and numbers indicate different behavioural phases. 80

Figure 25 | Tank structures for **(a)** the Séchilienne model and **(b)** the Mont de La Saxe model. Reservoirs are named according to their mechanisms: Infiltration (I), Matrix (M) and Fracture (F) reservoirs. The flux nomenclature is defined by the two reservoirs they connect, and the parameters $T1_{hy}$ and $T2_{hy}$ represent the two thresholds (T) of the hysteresis function. 84

Figure 26 | Results of the Séchilienne model simulation during the calibration and validation periods. **(a)** Observed (black) and simulated (red) discharges with the confidence interval of the simulation (orange area) and the best discharge simulation (dashed brown line). **(b)** Internal discharges representing the matrix (blue) and fracture (green) contributions to the total simulated discharge. **(c)** Water levels in reservoirs I (pink) and M (blue). 89

Figure 27 | Diagnostic efficiency polar plot for the Séchilienne model simulations during calibration (black) and validation (red) periods. 90

Figure 28 | Results of the Mont de La Saxe model simulation during the calibration and validation periods. **(a)** Observed (black) and simulated (red) discharges with the confidence interval of the simulation (orange area) and the best discharge simulation (dashed brown line). **(b)** Internal discharges representing the fracture (blue) and the hysteresis (green) contributions to the total simulated discharge. **(c)** Water levels in reservoirs I (pink) and F (blue). 91

Figure 29 | First-order and total-order Sobol sensitivity index values for **(a)** the Séchilienne and **(b)** the Mont de La Saxe models' parameters. 93

Figure 30 | Discharge simulations during the validation period with their performance values, based on the different calibration period lengths tested represented by P for **(a)** the Séchilienne and **(b)** the Mont de La Saxe models. The P0 periods represent the calibration periods used in this study to calibrate the models. The colours of the calibration periods correspond to the colours of the discharge simulations and the objective function values. 94

Figure 31 | Discharge simulations from the calibrated and validated Séchilienne model using historical meteorological data up to 1992. The discharge acquisition period is shown in yellow. The model simulated time series (internal discharges and water levels) follow the same colour code as in **Figure 26.** 95

Figure 32 | Comparison of the internal discharge time series representing the matrix contribution (blue) and fracture contribution (green) from the Séchilienne model with electrical conductivity (E.C.) measurements (purple). The observed discharge time series is shown in black dashed lines, and the pink areas indicate periods where the matrix contribution is the only input to the outlet. 97

Figure 33 | Inputs and target used in the XAI workflow. Rainfall and groundwater level describe hydro-meteorological forcing and hydrological state (predictors), while surface velocity is the target simulated by XGBoost. 103

Figure 34 | Maps of cumulative surface displacement measured on the Séchilienne landslide over four distinct periods between 1997 and 2022, superimposed on a 1-m-resolution shaded DEM. Periods are defined according to the progressive densification of the monitoring network, with newly installed stations shown as yellow triangles and pre-existing velocity plots as orange triangles. Cumulative displacements are displayed using a colour gradient. Black areas mask zones where interpolation is considered unreliable due to the absence of nearby plots. 104

Figure 35 | Overview of the three landslide sites: **(a)** Séchilienne (SH), **(b)** Viella (VI), and **(c)** Villerville (VV). Insets show the spatial distribution of displacement monitoring plots, tacheometric station (TS) and meteorological stations, including Météo-France (MFWS), Mont Sec (MSWS), Luitel (LWS) and Villerville (VWS). **(d–f)** Cumulative displacement at each monitoring location, colour-coded by monitoring plot, shown together with the corresponding effective rainfall and groundwater level time series. 108

Figure 36 | XAI modelling framework. The modelling framework is organised in six modules. M1 pre-processes cumulative displacement time series. M2 assesses time series stationarity, isolate long-term trend and transient components, and quantifies correlations with environmental forcing. M3 forecasts the trend component. M4 performs features extraction and XGBoost training and testing. M5 quantifies predictor importance using features-importance ranking, and attributes predictor effects using SHAP analysis. M6 evaluates the model predictive performance over independent test periods. 111

Figure 37 | Segmentation of the Séchilienne velocity record (plot A13) using Bai–Perron breakpoint detection. The Bai–Perron method is applied with three breakpoints to partition the time series into four quasi-linear periods consistent with changes in series evolution. Identified periods are shown using alternating red and blue shading. 116

Figure 38 | Simulated landslide velocities, interpretability and explainability of the predictors (feature importance, SHAP attribution). **(a-c)** Séchilienne (plot A13), **(d-f)** Viella (plot VIL-PL35), and **(g-i)** Villerville (plot VLRB). Observed and simulated daily velocities are shown in grey and black, respectively **(a, d, g)**, with predictive performance over test periods (periods 1-5) quantified using RMSE (triangle) and MAE (square) metrics. Feature importance (left) and SHAP-attribution (right) highlight the dominant hydrometeorological controls on model behaviour (b, e, h) and their influence on predicted velocities **(c, f, i)** across the training and testing periods, as indicated by the horizontal bars at the

top of each panel. For clarity, only the twenty most influential predictors are shown, capturing the majority of explained variance (see section V.2.5.2). In the SHAP plots, each point represents a single testing simulation, coloured by the corresponding predictor value. Predictors are colour-coded by source: effective rainfall (sky blue), net rainfall (steel blue), and groundwater level (magenta). At Séchilienne, two groundwater level series are shown, representing shallow and deep hydrological reservoirs (*Béjean-Maillard et al., 2025*). Simulated velocities correspond to the simulated total series (trend × transient); transient-component simulations are reported in the Extended Data (see ***section V.2.7.*** ; **Figure 45**). At Viella, the leading predictor exceeds 70% of the total importance score; accordingly, an axis break is used to retain a comparable scale across panels. 119

Figure 39 | Model sensitivity to increasing subsets of ranked predictors. **(a)** Séchilienne (plot A13), **(b)** Viella (plot VIL-PL35), and **(c)** Villerville (plot VLRB). Explained variance and error metrics (RMSE, MAE) are evaluated over the final test period (period 5). For Séchilienne, prediction errors are computed for the dimensionless transient-component forecast (i.e., the direct output of the XAI model), as the reconstructed total series (trend × transient) exhibits near-zero velocities, yielding low-contrast residuals. Séchilienne transient forecasts are reported in the Extended Data (see ***section V.2.7.*** ; **Figure 45**). For Viella, explained cumulative variance increases rapidly; the y-axis is therefore zoomed to the 90–100% range. 123

Figure 40 | Simulated landslide velocities, interpretability and explainability of the predictors (feature importance, SHAP attribution) for additional Viella monitoring plots (part 1). **(a–c)** Observed and simulated daily velocities are shown in grey and black, respectively, for **(a)** VIL-14, **(b)** MID-03, and **(c)** CAH-I8. Legends are provided in panel **(a)**. Feature importance (left) and SHAP attribution (right) highlight the dominant hydrometeorological controls and their influence on predicted velocities across the training and testing periods, as indicated by the horizontal bars at the top of each panel. Predictors are colour-coded as in **Figure 38**: effective rainfall (sky blue), net rainfall (steel blue), and groundwater level (magenta). Axis breaks (double bars) are used for feature importance and SHAP value axes to enable comparison across plots while accommodating large values. The January 2022 velocity peak at VIL-14 (~5 cm d^{-1}) is indicated by an arrow. 127

Figure 41 | Simulated landslide velocities, interpretability and explainability of the predictors for additional Viella monitoring plots (part 2) for **(a)** VIL-16, **(b)** VIL-11, and **(c)** BAV-01. Legends are shown in panel **(a)**. 128

Figure 42 | Simulated landslide velocities, interpretability and explainability of the predictors for additional Séchilienne monitoring plots for **(a)** A16, **(b)** C2, **(c)** P1103, and **(d)** P1300. Predictors and colour- 129

Figure 43 | Simulated landslide velocities, interpretability and explainability of the predictors for additional Villerville monitoring VLRH plot. Predictors and colour-coding follow **Figure 38.** 130

Figure 44 | Zonation map of cumulative surface displacement at Viella landslide. Analysed tacheometric targets and their associated time series are shown in colour. Additional monitoring plots are indicated in white; their cumulative displacement curves are overlaid using the colour of the corresponding displacement zone with partial transparency. 130

Figure 45 | Simulated transient-component velocities (dimensionless) for the Séchilienne monitoring plots. **(a)** plot A13, **(b)** plot A16, **(c)** plot C2, **(d)** plot P1103, and **(e)** plot P1300. Observed and simulated daily velocities are shown in grey and black, respectively, with predictive performance across the five test periods (periods 1–5) quantified using RMSE (triangles) and MAE (squares). 131

Figure 46 | Chapter VI - Schematic overview of the probabilistic simulation framework developed. 139

Figure 47 | Calibration and simulation workflow for **(a)** the rainfall, and **(b)** temperature time series generators (adapted from *Breinl et al., 2015, 2016*). 141

Figure 48 | **(a)** Location of the three studied deep-seated landslides, with details on the geomorphological features and monitoring network for **(b)** Villerville (adapted from *Lissak, 2009*), **(c)** Viella, and **(d)** Vallcebre. MFWS and VWS refer to Meteo-France Weather Station and Vallcebre Weather Station. 145

Figure 49 | Measured daily surface velocities, groundwater level, and effective rainfall on **(a)** Viella, **(b)** Villerville, and **(c)** Vallcebre landslides. 148

Figure 50 | Rainfall amount and number of wet days values distribution from measured and generated series for **(a)** Viella (blue), **(b)** Villerville (red), and **(c)** Vallcebre (green). 150

Figure 51 | Monthly **(a)** mean, **(b)** standard deviation, **(c)** maximum, and **(d)** minimum temperature value from measured series (solid line) and generated series (shaded area) for Viella (blue), Villerville (red), and Vallcebre (green). 151

Figure 52 | Measured (black) and simulated (blue) groundwater level series for **(a)** Viella, **(b)** Villerville, and **(c)** Vallcebre. 153

Figure 53 | Generated net rainfall, effective rainfall, temperature, and groundwater level series for **(a)** Viella, **(b)** Villerville, and **(c)** Vallcebre. For each site, the 15-month simulation period (corresponding to the XGBoost test set) is shown, with the observed series plotted as a black line. Total net rainfall and effective rainfall are summarised using box plots. 155

LIST OF FIGURES

Figure 54 | Monthly cumulative velocity distribution for **(a)** Viella, **(b)** Villerville and **(c)** Vallcebre landslides. Observed monthly cumulative velocities are shown as red dots, and those corresponding to the target period as black dots. Whiskers represent the 5th–95th percentile range, consistent with the Mean Interval Score (MIS). 157

Figure 55 | Monthly cumulative velocity distribution for **(a)** Viella, **(b)** Villerville and (c) Vallcebre landslides, obtained from optimised XGBoost model. 161

Figure 56 | Results of the XGBoost model applied to borehole VC_S2 at the Vallcebre landslide. **(a)** Simulated velocities (black) compared with observed velocities (grey) for each test block representing 20% of the data (Blocks 1 to 5). The total error for each test block is quantified using RMSE (triangles) and MAE (squares). **(b)** Ranking of the 20 most important features in explaining the variance of the model. Feature colours refer to their source time series (light blue for effective rainfall, purple for groundwater levels). **(c)** SHAP analysis of the effects of the 20 most important features on each velocity value in the test set, expressed as Shapley values. Feature values are shown using a blue-to-red colour scale, while their influence on the model output is represented along the x-axis. According to the feature importance ranking and SHAP analysis, short-term features related to effective rainfall and groundwater levels are the most influential in controlling velocities. In particular, the groundwater level on the current day exerts the strongest influence on the model predictions for new data. For further details on the method, refer to ***section V.2.3.2.*** 165

Figure 57 | Autocorrelation functions (ACF) of the observed series and the generated ensemble for lags up to 10 days at **(a)** Viella, **(b)** Villerville, and **(c)** Vallcebre. 165

Figure 58 | Simulated daily velocity series derived from 500 hydro-meteorological scenarios (grey), together with observed daily velocities for all available measurement years at **(a)** Viella, **(b)** Villerville and **(c)** Vallcebre. 166

Figure 59 | Simulated velocity time series on the training and testing datasets obtained with the optimised XGBoost model for **(a)** Viella, **(b)** Villerville, and **(c)** Vallcebre. 167

Figure 60 | Schematic overview of future research directions. 175

Figure A1 | Représentation synthétique des méthodes d'acquisition des données de surveillance des glissements de terrain et des approches de modélisation des déplacements. Les encadrés verts et rouges résument les principaux avantages et inconvénients de chaque méthode. 198

Figure A2 | Workflow de traitement des séries temporelles et de mise en œuvre d'un modèle de machine learning conventionnel. 199

Figure A3 | Courbes de débits classés établies pour les glissements de terrain de **(a)** Séchilienne et **(b)** Mont de La Saxe. Les pentes sont représentées par alpha, les ruptures de pente sont indiquées par des lignes pointillées accompagnées des valeurs de débit associées, et les nombres indiquent les différentes phases de fonctionnement. 202

Figure A4 | Structures en réservoirs des modèles de **(a)** Séchilienne et **(b)** Mont de La Saxe. Les réservoirs sont nommés en fonction des mécanismes qu'ils représentent : réservoir d'infiltration (I), réservoir matriciel (M) et réservoir de fracture (F). La nomenclature des flux est définie par les deux réservoirs qu'ils relient, et les paramètres T1hy et T2hy représentent les deux seuils (T) de la fonction d'hystérésis. 202

Figure A5 | Résultats de la simulation du modèle de Séchilienne durant les périodes de calibration et de validation. **(a)** Débits observés (noir) et simulés (rouge), avec l'intervalle de confiance de la simulation (zone orange) et la meilleure simulation de débit (ligne brune en pointillés). **(b)** Débits internes représentant les contributions de la matrice (bleu) et des fractures (vert) au débit total simulé. **(c)** Niveaux d'eau dans les réservoirs I (rose) et M (bleu). 203

Figure A6 | Résultats de la simulation du modèle de Mont de La Saxe durant les périodes de calibration et de validation. **(a)** Débits observés (noir) et simulés (rouge), avec l'intervalle de confiance de la simulation (zone orange) et la meilleure simulation de débit (ligne brune en pointillés). **(b)** Débits internes représentant les contributions des fractures (bleu) et de l'hystérésis (vert) au débit total simulé. **(c)** Niveaux d'eau dans les réservoirs I (rose) et F (bleu). 205

Figure A7 | Vue d'ensemble des trois sites de glissement de terrain : **(a)** Séchilienne (SH), **(b)** Viella (VI) et **(c)** Villerville (VV). Les encarts montrent la distribution spatiale des points de suivi des déplacements, de la station tachéométrique (TS) et des stations météorologiques, incluant Météo-France (MFWS), Mont Sec (MSWS), Luitel (LWS) et Villerville (VWS). **(d–f)** Déplacements cumulés à chaque point de suivi, codés en couleur selon le point de mesure, présentés conjointement avec les séries temporelles correspondantes de pluie efficace et de niveau d'eau souterraine. 207

Figure A8 | Cadre de modélisation XAI. Le cadre de modélisation est organisé en six modules. M1 prétraite les séries temporelles de déplacements cumulés. M2 évalue la stationnarité des séries temporelles, isole la tendance de long terme et les composantes transitoires, et quantifie les corrélations avec les forçages environnementaux. M3 prévoit la composante de tendance. M4 réalise l'extraction des variables descriptives ainsi que l'entraînement et le test du modèle XGBoost. M5 quantifie l'importance des prédicteurs à l'aide d'un classement d'importance des variables et attribue les effets des prédicteurs au moyen d'une analyse SHAP. M6 évalue les performances prédictives du modèle sur des périodes de test indépendantes. 208

LIST OF FIGURES

Figure A9 | Vitesses de glissement simulées, interprétabilité et explicabilité des prédicteurs (importance des variables, attribution SHAP). **(a-c)** Séchilienne (point A13), **(d-f)** Viella (point VIL-PL35) et **(g-i)** Villerville (point VLRB). Les vitesses journalières observées et simulées sont représentées respectivement en gris et en noir **(a, d, g)**, avec des performances prédictives sur les périodes de test (périodes 1 à 5) quantifiées à l'aide des métriques RMSE (triangle) et MAE (carré). L'importance des variables (à gauche) et l'attribution SHAP (à droite) mettent en évidence les contrôles hydrométéorologiques dominants du comportement du modèle (b, e, h) et leur influence sur les vitesses prédites (c, f, i) au cours des périodes d'entraînement et de test, comme l'indiquent les barres horizontales en haut de chaque panneau. Par souci de lisibilité, seuls les vingt prédicteurs les plus influents sont représentés, ceux-ci expliquant la majeure partie de la variance expliquée. Dans les graphiques SHAP, chaque point correspond à une simulation individuelle de test, colorée selon la valeur du prédicteur correspondant. Les prédicteurs sont codés par couleur selon leur source : pluie efficace (bleu ciel), pluie nette (bleu acier) et niveau d'eau souterraine (magenta). À Séchilienne, deux séries de niveau d'eau souterraine sont représentées, correspondant à des réservoirs hydrologiques superficiel et profond (Béjean-Maillard et al., 2025). Les vitesses simulées correspondent à la série totale simulée (tendance × transitoire). À Viella, le prédicteur principal dépasse 70 % du score total d'importance ; une rupture d'axe est donc utilisée afin de conserver une échelle comparable entre les graphiques. 210

Figure A10 | Chapitre VI – Vue d'ensemble schématique du cadre de simulation probabiliste développé. 213

Figure A11 | Workflow de calibration et de simulation des générateurs de séries temporelles de **(a)** pluie et **(b)** température (adapté de Breinl et al., 2015, 2016). 213

Figure A12 | Séries de niveaux d'eau souterraine mesurées (noir) et simulées (bleu) pour **(a)** Viella, **(b)** Villerville et **(c)** Vallcebre. 215

Figure A13 | Distribution des vitesses cumulées mensuelles pour les glissements de terrain de **(a)** Viella, **(b)** Villerville et **(c)** Vallcebre. Les vitesses cumulées mensuelles observées sont représentées par des points rouges, et celles correspondant à la période cible par des points noirs. Les moustaches représentent l'intervalle compris entre les 5e et 95e percentiles, en cohérence avec le Mean Interval Score (MIS). 217

LIST OF TABLES

Table 1 | Overview of studied slow-moving landslide features.....17
Table 2 | Risk level assessment by combining a hazard level and a vulnerability level (modified from *Maquaire, 2002*) 25
Table 3 | Common multi-source in situ monitoring systems used for slow-moving landslides (adapted from *Uhlemann et al., 2016*).....37
Table 4 | Main hyperparameters used in the XGBoost model (adapted from *Tarwidi et al., 2023*).....58
Table 5 | The Séchilienne and Mont de La Saxe landslides geographical and geological features 74
Table 6 | Summary of the Séchilienne landslide hydrometeorological dataset 76
Table 7 | Summary of the Mont de La Saxe hydrometeorological dataset.....79
Table 8 | Summary of the parameters for the Séchilienne and Mont de La Saxe models, including the defined ranges for simulation and the optimal values retained. The order of parameters presents the structural parameters (from I_0 to x_{hy}), the PET routine parameters (K_1 and K_2) and the Snow routine parameters (from T_S to CWH).....85
Table 9 | Warm-up, calibration, and validation periods for the Séchilienne and Mont de La Saxe model simulations .86
Table 10 | Summary of performance criteria and parameters values for the Séchilienne and Mont de La Saxe simulations during the calibration and validation periods.....88
Table 11 | Landslide geomorphology and observation network.....109
Table 12 | Hydro-meteorological predictors used in the XAI framework. Predictors are grouped by data source: net Rainfall (R), Effective Rainfall (ER), and GroundWater Level (GWL). W denotes sliding-window lengths, and L denotes time lags, in days. Feature names encode the data source, the extracted features, and the associated window and lag parameters.....110
Table 13 | Pre-processing parameter values used in Module M1. Values are reported for each monitoring plot. Parameters are grouped by processing step: Outlier Detection (OD), Adaptive Smoothing (AS), and interpolation. Séchilienne time series do not include AS_Wmax values, as the long-term trend affects the median-based standard-deviation classification used to assign short or long smoothing windows.....113
Table 14 | ADF and KPSS stationarity test results. Both tests are evaluated using p-values with a significance level of $\alpha = 0.05$. Rejection of the ADF null hypothesis (p-value $< \alpha$) indicates that the series is stationary. Rejection of the KPSS null hypothesis (p-value $< \alpha$) indicates that the series is non-stationary. Differences between ADF and KPSS outcomes reflect their complementary null hypotheses and can arise from near-threshold behaviour, weak or non-linear trends, or changes in variance.....115
Table 15 | ADF and KPSS stationarity test results across the A13 plot periods. Series segmentation is based on the Bai–Perron breakpoint detection method (**Figure 37**) 115
Table 16 | Geomorphological and datasets descriptions of the studied landslides 147
Table 17 | BIC values for Markov chain orders and RMSE of fitted cumulative distribution functions used for rainfall generator calibration. Reported values are the median over the 12 months.....149
Table 18 | Lambda, RMSE, and Lilliefors test results for the modulus transformation applied to the seasonal temperature distributions at the three sites. Seasons letters depict considered months. p-values are reported before and after the transformation, whereas RMSE refers to the minimum value achieved after transformation. The Lilliefors test evaluates the null hypothesis that the data are drawn from a normal distribution. 150
Table 19 | Total temperature probabilistic metrics for 90% prediction intervals 152
Table 20 | Optimal PASTAS model parameters and R^2 values for groundwater level simulations at Viella, Villerville, and Vallcebre.....152
Table 21 | Monthly cumulative velocity distribution ranges metrics for 90% prediction intervals. Results for the distributions obtained with the standard (S) and optimised (Opt) XGBoost models are shown for each site. N denotes the number of available measured values. 159
Table 22 | RMSE and MAE metric values between observed and simulated velocities over the test period, with and without XGBoost hyperparameter optimisation 160
Table 23 | Optuna-optimised XGBoost hyperparameters and corresponding optimal values obtained on the training–calibration sets 167

CHAPTER I

GENERAL INTRODUCTION

CHAPTER I GENERAL INTRODUCTION 13
I.1. Context and objectives of the thesis **15**
I.2. Structure of the manuscript **19**

I.1. CONTEXT AND OBJECTIVES OF THE THESIS

Landslides affect numerous regions worldwide and cause substantial damage to infrastructure, ecosystems and socio-economic activities each year. Among these gravity-driven processes, slow-moving, deep-seated landslides involve several million cubic metres of rock and exhibit slow, often continuous motion with marked seasonal acceleration/deceleration phases. Although catastrophic failures are relatively rare, events such as the Vajont landslide in 1963 illustrate the potentially devastating consequences of such instabilities.

In this contexte, data-driven approaches and, in particular, Artificial Intelligence (AI) models have recently emerged as powerful tools to analyse, interpret and anticipate landslide behaviour. Current AI applications in landslide science can be broadly structured into five main directions (**Figure 1**):

- **Regional detection,** in which AI models are trained on satellite imagery or LiDAR-derived products to identify signs of instability and support the completion of landslide inventories at large spatial scales *(Amatya et al., 2023; Stumpf et al., 2014)*.
- **Susceptibility mapping,** where static geo-environmental variables (e.g., slope gradient, lithology, land cover, distance to fractures) are used to assign a susceptibility index to each spatial unit *(Catani et al., 2013; Mondini et al., 2021; Reichenbach et al., 2018)*.
- **Runout and intensity estimation** (e.g., velocity, thickness, energy) during rapid failure, where surrogate models are trained on large ensembles of numerical simulations to approximate physically-based outputs at reduced computational cost *(Hibert et al., 2025)*.
- **Operational monitoring of unstable areas,** where AI-based classification is applied to satellite data or seismic signal to detect and track temporal changes in landslide activity *(Hibert et al., 2019, Li et al., 2025)*.
- **Forcing-displacement relationships understanding,** where AI models are trained on multi-source time series to reproduce observed displacements and identify the dominant controls through non-linear relationships *(Krkač et al., 2017)*.

This last direction remains comparatively less explored for three main reasons. First, surface dynamic in slow-moving landslides is primarily driven by hydro-meteorological forcings, including intense or prolonged rainfall, and the resulting groundwater level and pore-water pressure variations *(Bontemps et al., 2020; Crosta et al., 2013; Dille et al., 2019; Fiolleau et al., 2020)*. The influence of pore-water pressure on landslide displacement has been demonstrated in many case studies *(Binet et al., 2007; Cappa et al., 2004; Malet et al., 2005; Vallet et al., 2016)*, yet formalising robust forcing–response relationships remains challenging. The dynamic of slow-moving landslides are typically non-linear, heterogeneous in space and time, and sensitive to changing multi-scale hydrological conditions *(Desrues et al., 2019; Federico et al., 2012)*, which complicates process understanding.

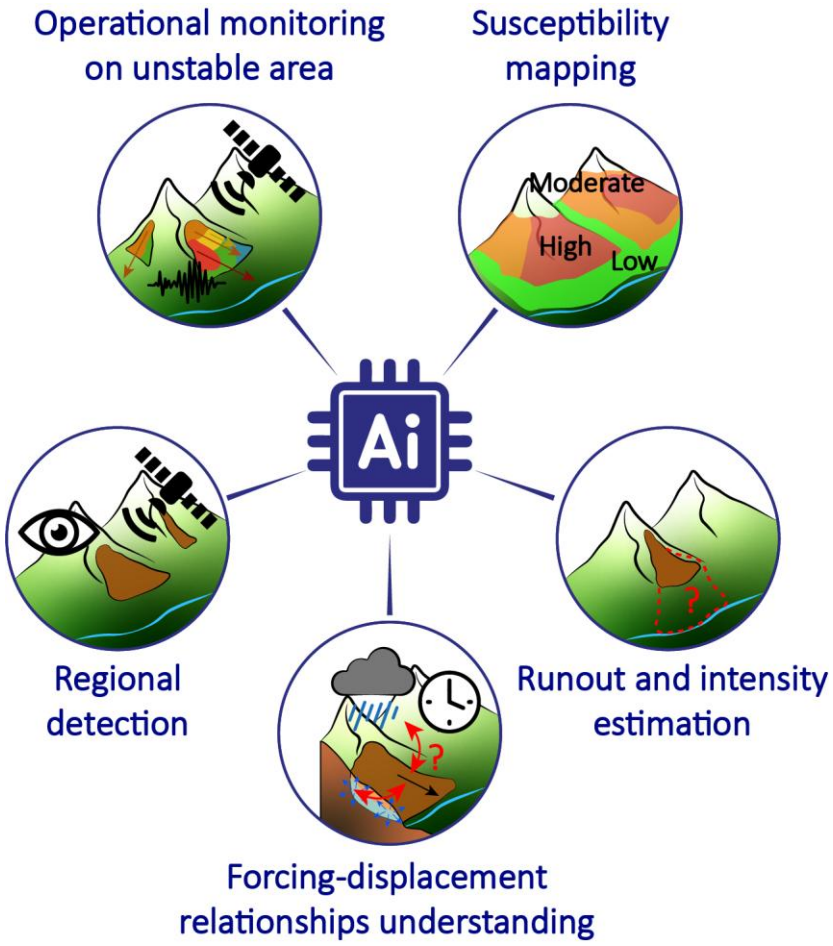


Figure 1 | Main application domains of artificial intelligence in landslide science and hazard mitigation.

Second, AI-based displacement forecasting requires long, consistent, multi-source datasets (i.e., both displacement and hydro-meteorological variables). Acquiring and maintaining such records over long periods is technically demanding, particularly in active landslide areas where instrumentation is exposed to damage, displacement-induced failure, and data gaps.

Third, many AI models still operate as black boxes, providing limited insight into which variables drive predictions and through which mechanisms. Questions of interpretability and explainability have therefore gained increasing attention, but methodological developments for time-series applications remain recent and often site-specific *(Gu et al., 2024; Z. Ma & Mei, 2025)*.

The objective of this thesis is to improve our understanding of hydro-meteorological controls on slow-moving landslide dynamics by developing and applying interpretable and explainable AI frameworks to simulate observed surface displacements across multiple sites. **The thesis addresses the following research questions:**

- **How can long, informative and representative multi-source time series be acquired – or reconstructed – to support robust AI model training in slow-moving landslides?**

- **How do hydro-meteorological controls shape the temporal dynamics of surface displacement, and over which characteristic time scales?**

- **How does slow-moving landslide displacement evolve under contrasted and plausible hydro-meteorological conditions?**

To address these questions, this work relies on multi-parameter datasets collected on several slow-moving landslides characterised by contrasting lithological settings, displacement dynamics, and meteorological contexts: the **Séchilienne landslide** (French Alps), the **Mont de La Saxe landslide** (western Italian Alps), the **Villerville landslide** (French Normandy coast), the **Viella landslide** (central French Pyrenees), and the **Vallcebre landslide** (central Spanish Pyrenees). Not all sites are examined in every chapter. Their mobilisation depends on the scientific question addressed, data availability, and the methodological objectives of each chapter (**Table 1**).

Table 1 | Overview of studied slow-moving landslide features.

	Séchilienne	**Mont de La Saxe**	**Villerville**	**Viella**	**Vallcebre**
Geographical coordinates (lat., lon. ; °C)	45.06, 5.80	45.82, 6.97	49.40, 0.12	42.87, 0.01	42.21, 1.83
Country	France	Italy	France	France	Spain
Area (km^2)	1.0	0.15	0.47	0.59	0.80
Altitudes (m a.s.l.)	450 to 1100	1400 to 1870	0 to 77	782 to 1344	966 to 1254
Lithology description	Micaschist bedrock locally overlain by morainic deposits	Micaschist and calcschist bedrock with overlain by colluvial and glacial deposits	Chalk, sand, and colluvial debris overlying marl substrate	Morainic and colluvial debris overlying calcareous-schist bedrock	Clayey siltstone and colluvial debris overlying limestone substrate
Meteorological data	2004/07/24 - 2024/12/31	2013/01/01 - 2018/12/01	2000/01/01 - 2025/10/26	2008/01/01 - 2025/10/26	1996/11/22 - 2015/12/31
Groundwater level data	-	-	2008/08/21 - 2022/03/11	2020/03/12 - 2024/12/17	1996/11/22 - 2015/09/13
Discharge data	2014/03/07 - 2024/12/31	2013/01/01 - 2018/12/01	-	-	-
Displacement data	2004/07/24 - 2019/03/31	-	2009/07/26 - 2025/03/05	2020/08/21 - 2024/12/17	1996/11/22 - 2015/12/09
Chapter(s) of analysis	IV, V	IV	V, VI	V, VI	VI

The availability of these multi-source datasets results from several collaborations, foremost with the French National Observatory **SNO-OMIV** (Observatoire Multidisciplinaire des Instabilités de Versants), which ensures the long-term monitoring and maintenance of instrumentation on key slow-moving landslides across metropolitan France, and places particular emphasis on providing open-access time series enabling detailed characterisation of landslide kinematics, geometry, hydrogeological functioning and meteorological forcing. These efforts are complemented by national academic and operational partners, including **OSU-THETA** (Observatoire des Sciences de l'Univers – Terre Homme Environnement Temps Astronomie; University Marie et Louis Pasteur), **EOST** (Ecole et Observatoire des Sciences de la Terre; University of Strasbourg), **IDESS** (Identités et Différenciations de l'Espace, de l'Environnement et des Sociétés; University Caen Normandie), **EMMAH** (Environnement Méditerranéen et Modélisation des Agro-Hydrosystèmes ; Avignon University), **CEREMA** (Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement), and **ONF-RTM** (Office National des Forêts - Restauration des Terrains de Montagne). In addition, this thesis was supported by two international collaborations: the **Department of Planning, Soil Protection and Water Resources of the Autonomous Region of Valle d'Aosta** (Italy) and the **Department of Civil and Environmental Engineering, Polytechnic University of Catalonia (UPC BarcelonaTech)** (Spain).

I.2. STRUCTURE OF THE MANUSCRIPT

To address the questions outlined above, the manuscript is organised into five chapters (**Figure 2**).

Chapter II introduces the conceptual and methodological background required for the subsequent developments. It reviews the current state of the art on monitoring and modelling slow-moving landslide and outlines the main AI models considered for time series forecasting.

Chapter III presents the key components of the modelling framework, including data preparation, time-series exploratory analysis, feature engineering, modelling strategy, and performance evaluation. Particular emphasis is placed on interpretability and explainability, and on how these concepts can move AI approaches beyond purely predictive models towards more transparent and physically-grounded frameworks.

Chapter IV addresses the first research question by focusing on hydrological data availability, which is critical for training reliable AI models. Access to groundwater and pore-water pressure information is essential to understand and reproduce displacement dynamics *(Bogaard & Greco, 2016; Matsuura et al., 2008; Prokešová et al., 2013)*, yet long-term piezometric records are rare. Deep instrumentation requires substantial resources and is frequently compromised by ongoing deformation. The chapter develops conceptual reservoir models that represent the functioning of fractured aquifers within the landslide mass. These semi-empirical, rainfall–discharge-based models can be calibrated without direct piezometric data and used to infer the temporal evolution of the internal hydrological state. They also enable the reconstruction of hydrological time series beyond the monitoring period, provided that external forcing data (rainfall, temperature) are available, thereby filling gaps and extending incomplete records. The methodology is applied to two outlets located on the Séchilienne and Mont de La Saxe landslides, both occurring in fractured crystalline bedrock. The KarstMod model, originally developed within SNO-KARST for karst systems, is adapted to fractured aquifers. Beyond providing an alternative source of groundwater proxies for subsequent AI modelling, the calibrated models improve understanding of the flow regimes governing aquifer dynamics and their relative contributions to observed discharge, in agreement with existing knowledge.

Chapter V addresses the second research question and aims to improve understanding of the relationships between hydro-meteorological forcing and surface velocity by introducing a scalable eXplainable Artificial Intelligence (XAI) framework based on long multi-parameter time series. Model transparency is pursued through both interpretability and explainability tools. Interpretability is ensured by using XGBoost, a decision-tree ensemble algorithm that provides native predictor-importance metrics during training. Explainability is addressed through post-hoc SHAP (Shapley) analysis, which

quantifies the contribution of each predictor to each simulated velocity value. Because AI performance on time series critically depends on data preparation, this chapter also introduces a complete workflow for pre-processing and analysing displacement and forcing series to ensure that inputs are representative and suitable for learning. Improving transparency alone is insufficient if models remain strongly site-specific. The chapter therefore also tackles the issue of genericity. Instead of relying on site-dependent predictors, a generic set of descriptive hydro-meteorological features is derived from forcing time series (net and effective rainfall, groundwater-level variations). This site-independent feature set is used to train XGBoost models that simulate observed surface velocities at multiple velocity monitoring plots across the Séchilienne, Viella, and Villerville landslides. The results show that hydro-meteorological predictors allow the models to reproduce velocity dynamics without explicitly encoding site-specific conditions. Moreover, the inferred forcing–velocity relationships reveal contrasted responses across and within landslides, consistent with differences in lithology and hydrological regime, and provide insight into dominant controls, characteristic time scales, and their influence on simulated velocities.

Chapter VI addresses the third research question by extending the deterministic framework towards probabilistic simulations of landslide velocity sensitivity under contrasting hydro-meteorological contexts, building on the XAI models developed in ***Chapter V***. Hydro-meteorological variability is introduced by coupling modified Richardson-type weather generators with a multi-transfer-function hydrological model to simulate groundwater responses under multiple plausible forcing scenarios. For each scenario, the pre-trained AI models produce daily velocity simulations, which are then aggregated into monthly probabilistic distributions over a complete hydrological cycle. The results highlight contrasted seasonal patterns and expected month-to-month variability at each site. The approach is evaluated on the Viella, Villerville, and Vallcebre landslides and demonstrates its potential to provide probabilistic, physically-informed inputs for landslide hazard mitigation.

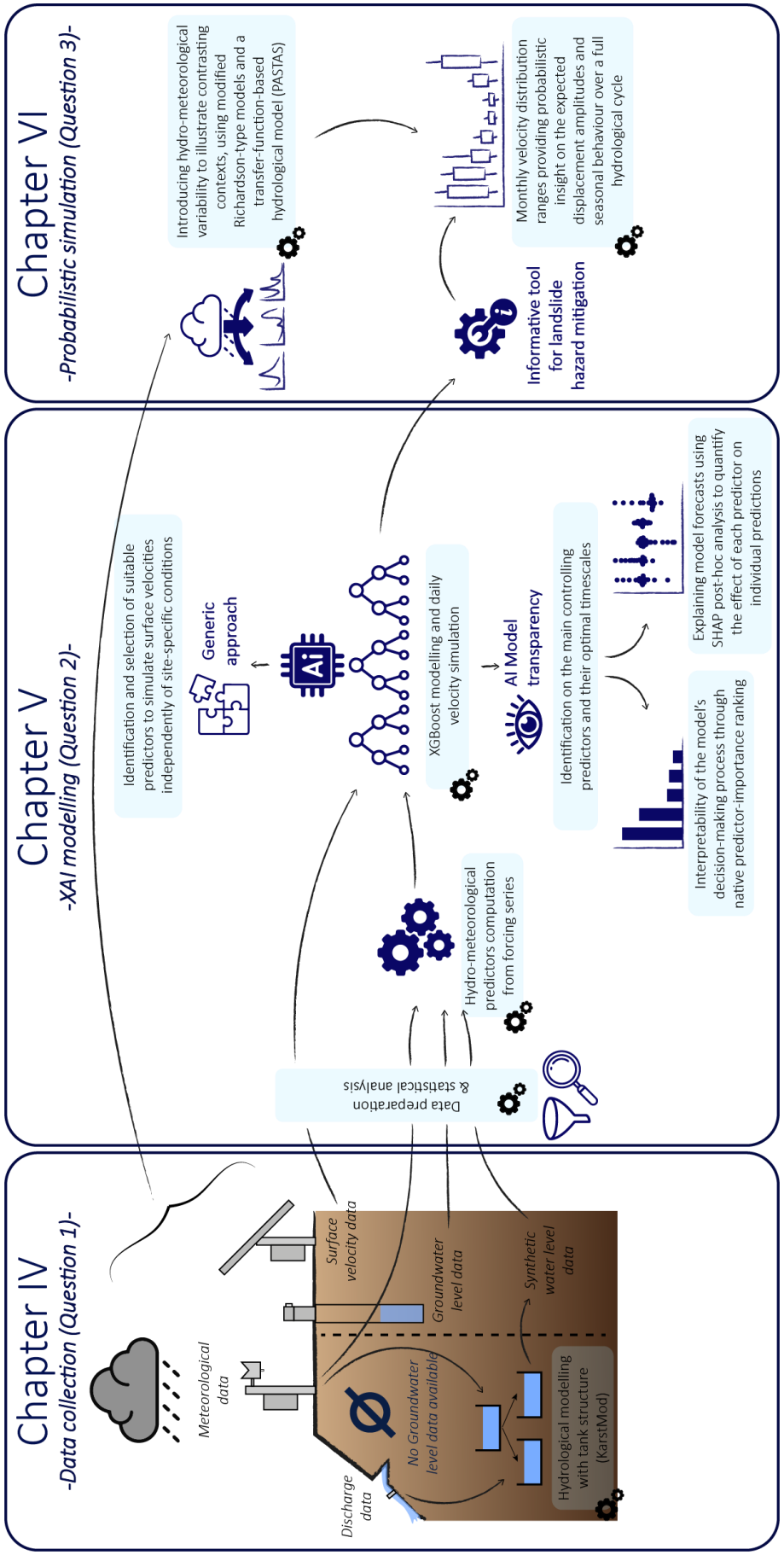


Figure 2 | Graphical abstract summarising the thesis chapters, their connections, and the associated research questions.

CHAPTER II

SLOW-MOVING LANDSLIDES: FROM HYDRO-MECHANICAL PROCESSES TO AI-BASED DISPLACEMENT MODELLING

CHAPTER II SLOW-MOVING LANDSLIDES: FROM HYDRO-MECHANICAL PROCESSES TO AI-BASED DISPLACEMENT MODELLING 23

II.1. Introduction to natural hazards 25

II.2. Definition, classification, and controlling factors of slope movements 27

II.3. Structural, geomorphological, and hydrological controls of slow-moving landslide dynamics 33

II.4. Data collection methods for monitoring slow-moving landslides 36

II.5. Simulating models for Slow-moving landslide displacement 39

II.5.1. Physically-based models 39

II.5.2. Data-driven models 40

II.6. Chapter II – Conclusion 49

II.1. INTRODUCTION TO NATURAL HAZARDS

Natural hazards are defined as threats or hazardous events of natural origin that may adversely affect human life and activities. The concept of risk combines the notions of hazard and vulnerability, which can be related through the following relationship:

$$Risk = Hazard \times Vulnerability$$

These two notions describe both the potential danger represented by a given hazard and the extent to which it may have a significant impact on human systems. According to *Maquaire (2002)*:

- ***Hazard*** refers to the "potential threat" of a natural phenomenon, characterised by a certain intensity, affecting a given area over a certain period of time.
- ***Vulnerability*** refers to the expected severity of the consequences of the hazard, encompassing all anthropogenic domains that may be affected (population, property, housing, economic activities, infrastructure, heritage, and the environment).

Expressed as a product, this relationship implies that the overall risk will increase if either hazard or vulnerability is high, and will become particularly critical if both components are elevated. Risk can therefore be assessed using a double-entry matrix that accounts for the various possible combinations (**Table 2**).

Table 2 | Risk level assessment by combining a hazard level and a vulnerability level (adapted from *Maquaire, 2002*).

		Vulnerability		
	Level	***Low***	***Medium***	***High***
Hazard	***Low***	Low	Low	Medium
	Medium	Low	Medium	High
	High	Medium	High	High

The evolution of risk and its perception is largely governed by vulnerability. On human time scales, hazard is often considered relatively stable, although this assumption can be seriously questioned in light of accelerated climate change regarding the scenarios proposed by IPCC reports *(Intergovernmental Panel on Climate Change (IPCC), 2023)*, and the clear increase in the frequency of extreme climatic events. In contrast, vulnerable areas are constantly evolving due to changes in land use and human exploitation. Vulnerability has progressively increased since the Industrial Era, as human

activities expanded into more hazardous environments such as coastal areas, mountain valleys, or regions subject to high seismicity.

Natural hazards include a wide range of phenomena of climatic (cyclones, tornadoes, extreme rainfall, drought), hydrological (tsunamis, floods) and geological origin (earthquakes, volcanic eruptions, slope movements).

Among these hazards, slope movements pose a major threat to populations and infrastructure and can cause severe damage *(Froude & Petley, 2018)*. They are defined as gravity-driven mass movements with variable kinematics. At the global scale, such phenomena occur predominantly in mountainous regions such as the Himalayas, the Andes, the Rockies, or the Alps, where relief is most pronounced (**Figure 3**). However, slope gradient alone does not fully control their spatial distribution. Other geomorphological factors, including lithology, soil mechanical properties and fracturing, can also contribute to landslide triggering. Consequently, other environments, including coastal or tropical areas, may also be exposed to these hazards.

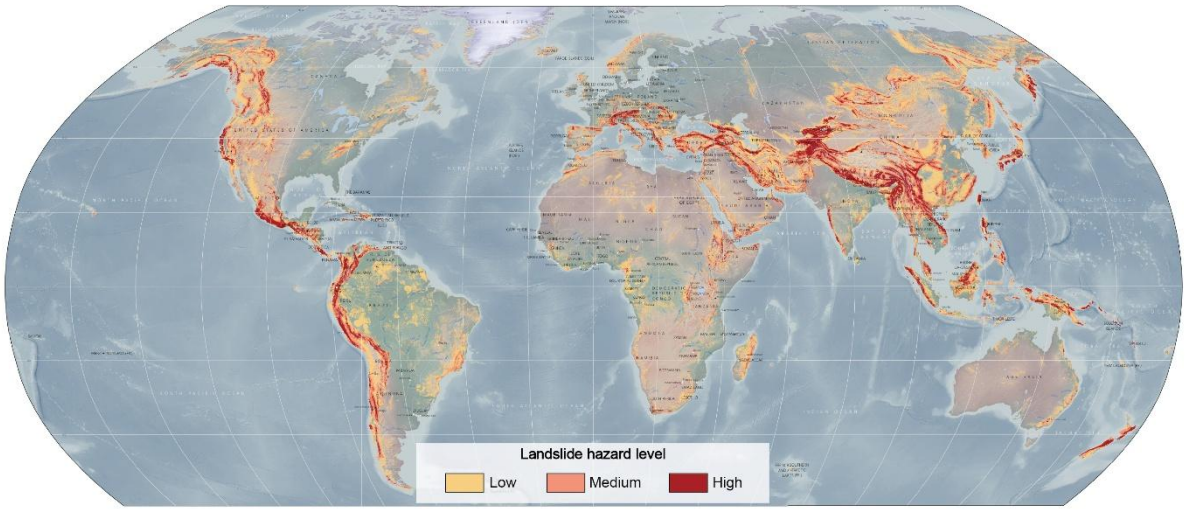


Figure 3 | Global landslide hazard worldmap, representing combined rainfall-triggered landslide hazard and earthquake-triggered landslide hazard. Hazards is defined with three categories, ranging from Low (yellow) to High (red) (from *Redshaw et al., 2020*).

II.2. DEFINITION, CLASSIFICATION, AND CONTROLLING FACTORS OF SLOPE MOVEMENTS

Slope movements are defined as the downslope and outward movement of slope-forming materials under the influence of gravity. They can involve volumes ranging from a few cubic metres to several million cubic metres, with displacement rates varying from a few $mm.yr^{-1}$ to, in the most extreme cases, several $m.day^{-1}$. They represent one of the main agents of topographic evolution, contributing both to relief degradation and to sediment transfer in fluvial systems and valley floors *(Agliardi et al., 2013; Clubb et al., 2023)*.

According to *Vallet (2014)*, a slope movement can be conceptualised as a body resting on an inclined plane and subjected to two main components: the normal component (P_n) and the tangential component (P_t) (**Figure 4a**). When referred to the contact area, these components define the normal stress (σ_n) and the shear stress (τ) (**Figure 4b**). Failure occurs when the ratio $\frac{\sigma_n}{\tau}$ exceeds a critical value, defined as the friction coefficient (μ). *Vallet (2014)* further defines the shear strength (R_s) as the resisting force to movement, controlled by the material cohesion (C), the friction coefficient, and the normal stress acting on the body (**Equation 1**):

$$R_s = C + \mu\sigma_n \quad (1)$$

A slope is therefore considered stable if the shear strength (R_s) compensates the applied shear stress (τ).

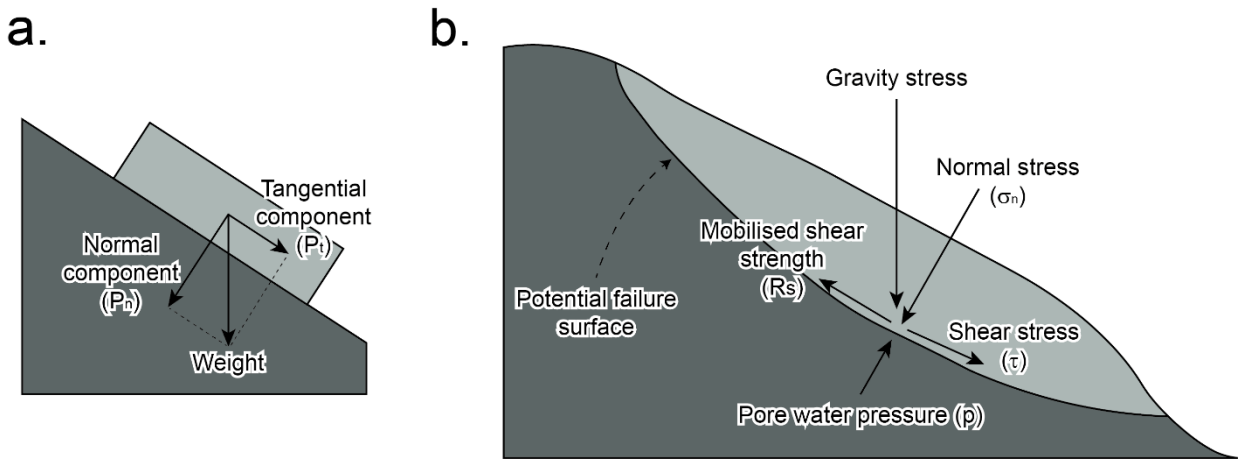


Figure 4 | (a) Forces involved in the destabilisation of landslides (adapted from *Vallet, 2014*). **(b)** Stresses acting along a potential failure surface (adapted from *McColl, 2015*). The light-grey area denotes the unstable mass.

Slope movements can be classified according to the nature of the displaced material (rock or soil), the movement mechanism, and the geometry of the rupture surface (rotational or planar). Based on these three descriptors, several classification systems have been proposed, including those of *Varnes (1954, 1978)*, *Hutchinson (1988)*, *Cruden & Varnes (1996)*, and the updated classification of Varnes proposed by *Hungr et al. (2014)*. The latter distinguishes six main types of movement – fall, topple, slide, spread, flow, and slope deformation – subdivided into 32 subtypes depending on the material involved and the rupture surface geometry. A schematic representation of the most common slope movements types is presented in **Figure 5**.

- **Falls** are characterised by the free fall of a rock or soil mass from a steep slope, producing a talus or debris cone downslope **(Figure 5a)**.
- **Topples** involve the forward rotation of a block from a cliff or steep slope about an axis located below its centre of gravity. They are often favoured by pre-existing discontinuities (fractures, bedding planes) and result in columnar or slab structures **(Figure 5b)**.
- **Slides** consist of a coherent mass of rock or soil moving along a rupture surface, without marked internal deformation. Rotational slides occur along concave-upward rupture surfaces, while translational slides move along planar surfaces **(Figures 5c-d)**.
- **Spreads** (or lateral spreads) develop on gentle slopes and involve the extension of surficial materials over a weaker underlying layer, leading to lateral extension, bulging, fracturing, and subsidence **(Figure 5e)**.
- **Flows** encompass movements in which the displaced material deforms internally in a continuous manner without a clearly defined rupture plane. Depending on the nature and cohesion of the material, they may take the form of rock avalanches, debris flows (non-cohesive mixtures with high water content), or earth flows (cohesive, lower-water-content masses) **(Figures 5f-g)**.
- **Slope deformations** correspond to large-scale gravitational deformations of steep slopes without a fully developed rupture surface, generally moving at very slow rates. In rock slopes, they are marked by geomorphological features such as sackungen, scarps, cracks, trenches, and bulges; in soil slopes, they are characterised by surface warping and creep-related deformation of surface elements **(Figures 5h-i)**.

In many cases, slope movements exhibit a combination of mechanisms, referred to as complex landslides. For example, slides or debris flows can destabilise steep headscarp areas, triggering secondary rockfalls, while rotational slides may evolve downslope into translational movement.

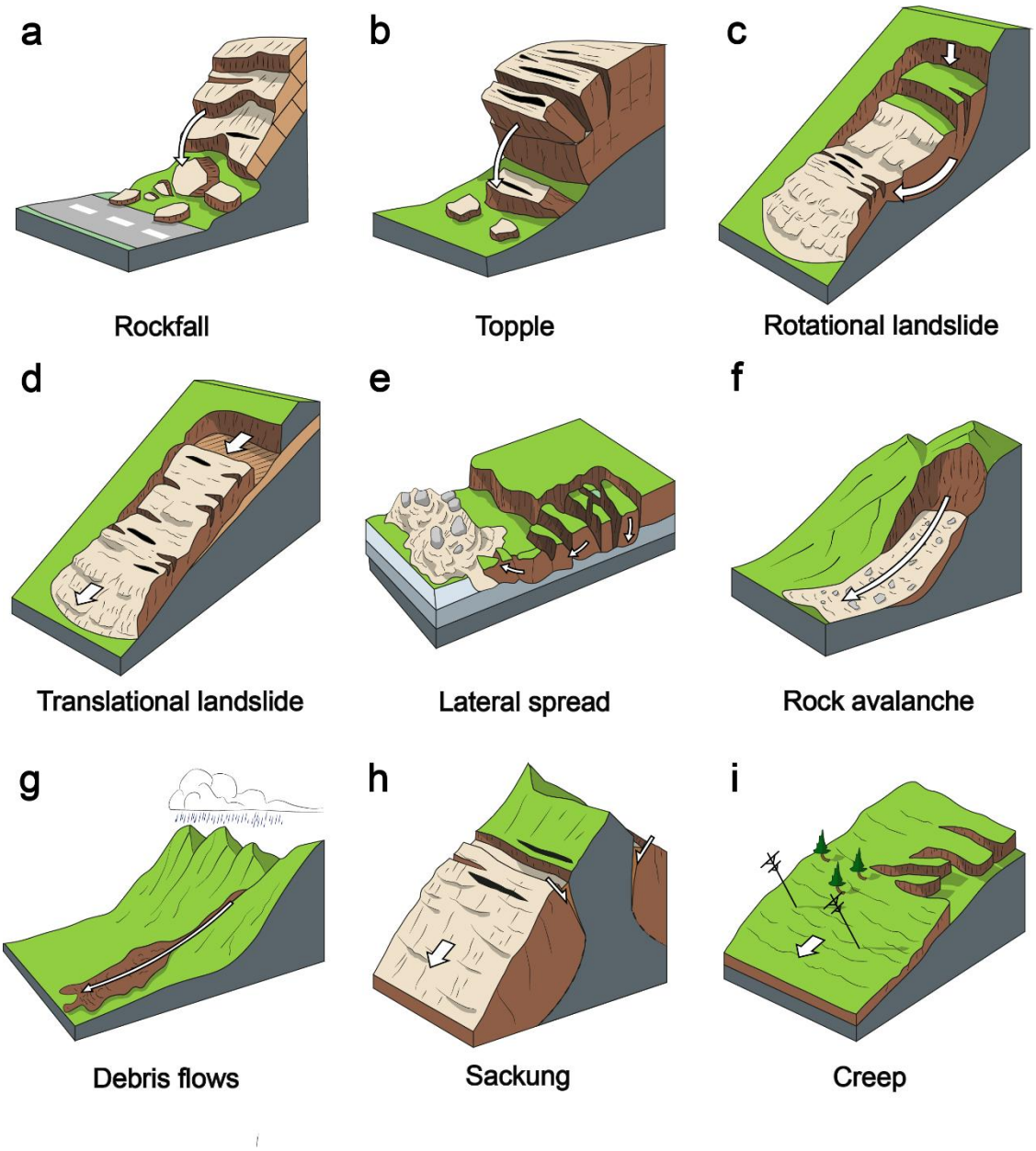


Figure 5 | Schematic representation of the most common types of landslides (adapted from *Highland & Johnson, 2004*).

The diversity of slope movements types reflects the wide range of factors controlling their initiation and triggering. These factors can be grouped into three categories: predisposing, preparatory, and triggering factors, each operating over different timescales:

- **Predisposing factors** define the intrinsic, long-term conditions that make a slope susceptible to instability. They include structural, geological, topographic, and climatic controls – such as

lithology, geological structures, slope gradient, aquifer presence, and glacial or tectonic heritage – and typically acting over of 10^4 – 10^6 years timescales

- **Preparatory factors** progressively modify the mechanical and hydrological conditions of a slope over intermediate timescales (10 – 10^4 years). They include processes such as weathering, surface erosion, frost wedging, and stress release following glacier retreat.
- **Triggering factors** induce rapid acceleration or sudden slope failure. They act over short timescales and include intense and/or prolonged rainfall, rapid snowmelt or permafrost thaw, groundwater level fluctuations, sudden loading, toe excavation, or seismic shaking. The most widely recognised dynamic triggers are earthquake shaking and pore-water pressure increases associated with rainfall infiltration.

Water infiltration from rainfall or snowmelt, through porous media or fracture networks, modifies pore-water pressure (p) within the unstable mass (**Figure 4b**). Pore-water pressure is critical because it reduces the effective normal stress (σ'_n) according to Terzaghi's principle *(Rutqvist & Stephansson, 2003; Terzaghi, 1951)* (**Equation 2**):

$$\sigma'_n = \sigma_n - p \tag{2}$$

An increase in pore-water pressure reduces the effective stress, and consequently, decreases both frictional resistance and the shear strength of the material. The influence of pore-water pressure depends on the depth of slope movement. For shallow landslides (<2 m), it is primarily controlled by soil moisture, with instability occurring when the soil approaches saturation *(Lu & Godt, 2008; Van Asch et al., 1999)*. For deep-seated slope movements (>5 m depth), pore-water pressure is mainly governed by fluctuations in the groundwater table *(Binet, Guglielmi, et al., 2007; Hibert et al., 2012; Van Asch et al., 1999)*. Such hydro-mechanical coupling is therefore considered as a key mechanism linking external forcing to slope instability *(Binet, 2006)* (**Figure 6**).

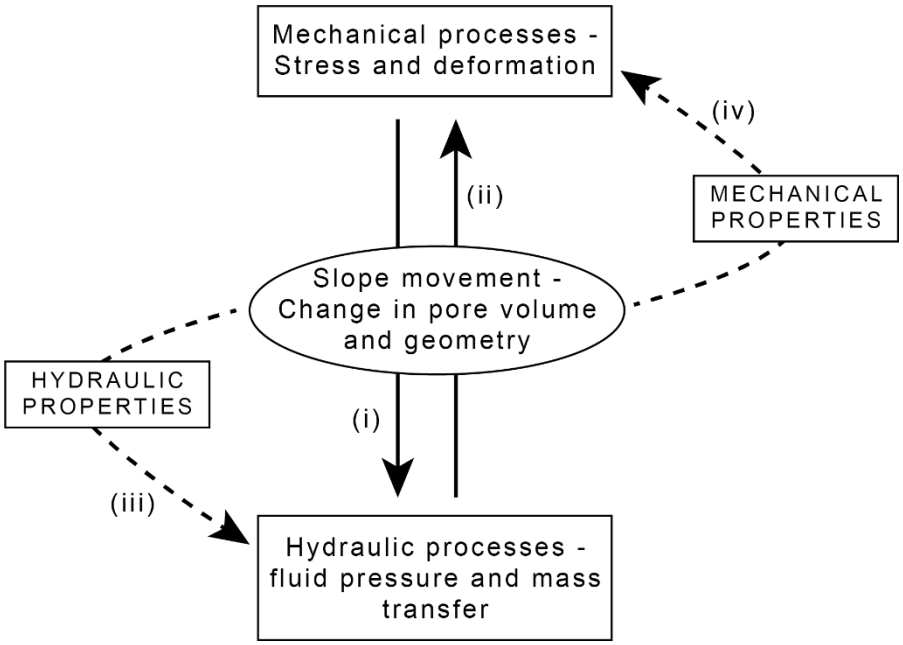


Figure 6 | Hydro-mechanical coupling in a fractured medium. (i) and (ii) corresponds to direct couplings related to pore volume variations (short-term), while (iii) and (iv) corresponds to indirect couplings related to material properties changes (medium- and long-term) (adapted from *Rutqvist & Stephansson, 2003*).

Their influences on instability can occur over the short-term, directly driven by pore volume changes. In such cases, slopes exhibit a fast hydro-mechanical response, with displacement closely correlated with rainfall, and, where the infiltration is rapid, with groundwater level fluctuations. The Vallcebre landslide provides a clear example of this behaviour, where acceleration phases recorded from 1996 to 2005 coincide with groundwater level changes (**Figure 7**).

Over the medium- and long-term, hydro-mechanical couplings represent indirect interactions through progressive changes in material properties. As a result, the rainfall–displacement relationship may evolve and become more complex. The response to rainfall is no longer immediate, and time-dependent deformation mechanisms such as long-term creep can emerge *(Brückl, 2001; Desrues et al., 2019; Federico et al., 2012; Vallet, 2014)*. These medium- and long-term hydro-mechanical couplings therefore act as preparatory factors *(Nevers, 2021; Vallet, 2014)*.

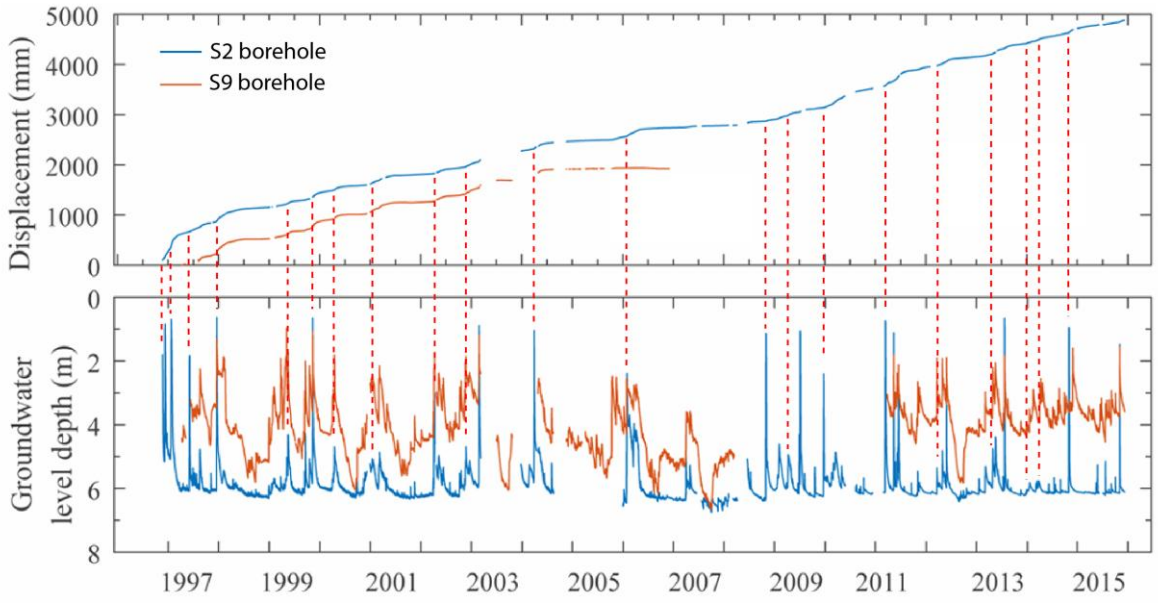


Figure 7 | Displacement and groundwater levels measured in two boreholes on the Vallcebre landslide from November 1996 to December 2015. Red dashed lines illustrate the temporal relationship between groundwater level changes and acceleration phases recorded in borehole S2 (adapted from *Moya et al., 2017*)

II.3. STRUCTURAL, GEOMORPHOLOGICAL, AND HYDROLOGICAL CONTROLS OF SLOW-MOVING LANDSLIDE DYNAMICS

Slow-moving, deep-seated landslides are mass movements characterised by a basal rupture surface located several metres to tens of metres below the ground surface, and by displacement rates ranging from a few centimetres to several metres per year *(Hungr et al., 2014)*. Their evolution is generally slow and progressive, typically following creep stages I and II, and rarely reaching complete failure *(Borrelli & Gullà, 2017; Desrues et al., 2019; El Bedoui et al., 2009)*. These persistent movements can nevertheless cause continuous damage to infrastructure *(Mansour et al., 2011)*. Under certain conditions, however, a sudden acceleration phase may occur, leading to the catastrophic collapse of the slope *(Hendron & Patton, 1987; Intrieri et al., 2019; Lacroix et al., 2018)*.

Their spatial distribution spans a wide range of geomorphological settings (coastal ranges, volcanic slopes, low and high mountain areas) but is most frequent in mountain chains, particularly where bedrock lithologies present low mechanical strength (e.g., schists, marls, clays, weathered or freactured crystalline rocks). In mountainous environments, initiation often occurs several thousand years after the Last Glacial Maximum, following glacial debuttressing *(Ballantyne, 2002; Hippolyte et al., 2012; Le Roux et al., 2009; Schwartz et al., 2017)*. Steep slopes resulting from glacial valley incision, together with tectonic inheritance (major discontinuities) and geological structures (schistosity, joint sets, bedding planes), form the main predisposing factors. Preparatory factors include progressive rock weathering due to meteorological processes, toe erosion, permafrost degradation, and land-use changes. Triggering factors are most often related to seismic activity, prolonged or intense precipitation, and snowmelt, which modify pore-water pressures and can accelerate movements *(Bontemps et al., 2020; Crosta et al., 2013; Dille et al., 2019; Fiolleau et al., 2020; Handwerger et al., 2013; Lacroix et al., 2020; Malet et al., 2002; Strozzi et al., 2010)*. Anthropogenic activities, such as excavation or the rapid drawdown of reservoir levels, can also contribute to accelerating slow-moving landslide motion *(Strauhal et al., 2016; Zangerl et al., 2010; H. Zhang et al., 2022)*.

The deformation of slow-moving landslides is complex, often involving composite gravitational processes such as flexural toppling, deep-seated sliding, and downslope block rotation *(Bovis, 1990)*. They also exhibit characteristic geomorphological features indicative of gravitational slope deformation, including counterscarps, double-crested ridges, uphill-facing scarps, trenches, and grabens *(Agliardi et al., 2009; Crosta, 1996)* (**Figure 8**). These morphological depressions can act as sediment traps and are often associated with lithological contrasts.

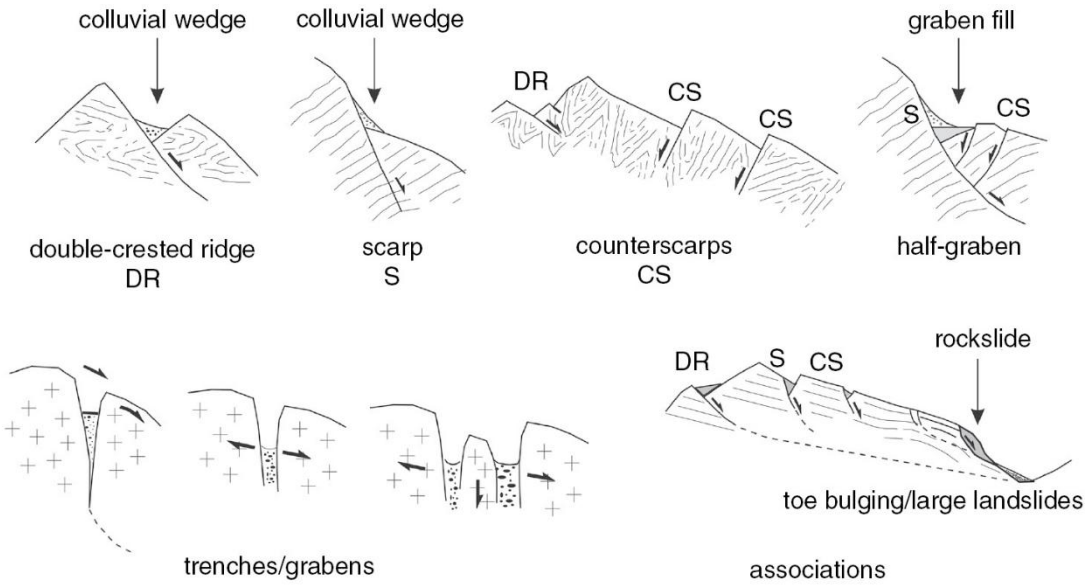


Figure 8 | Typical morphological structures of deep-seated deformations (from *Agliardi et al., 2013*).

The high degree of fracturing and anisotropy within deep-seated landslides results in complex hydrogeological systems. Infiltration from precipitation and snowmelt follows preferential flow paths controlled by discontinuities, which may extend to significant depths *(Bièvre et al., 2012; Krzeminska et al., 2012; van Asch et al., 1996)*. The intensity of these flows is governed by fracture aperture and connectivity. Such preferential pathways generate strong contrasts in hydraulic conductivity, leading to bimodal flow regimes combining fracture and matrix flows *(Cappa et al., 2004; Maréchal, 1998; Vallet et al., 2015; Van Asch et al., 1999)*. This promotes dual-aquifer configurations, with a perched aquifer within the unstable mass, consisting of discontinuous saturated zones, and a deeper aquifer located in the intact bedrock beneath the basal rupture surface *(Guglielmi et al., 2002; Meric et al., 2005; Pisani et al., 2010; Vengeon, 1998)* (**Figure 9**). Such flow heterogeneity induces strong spatial variability in pore-water pressure distribution and highly variable temporal responses to infiltration events.

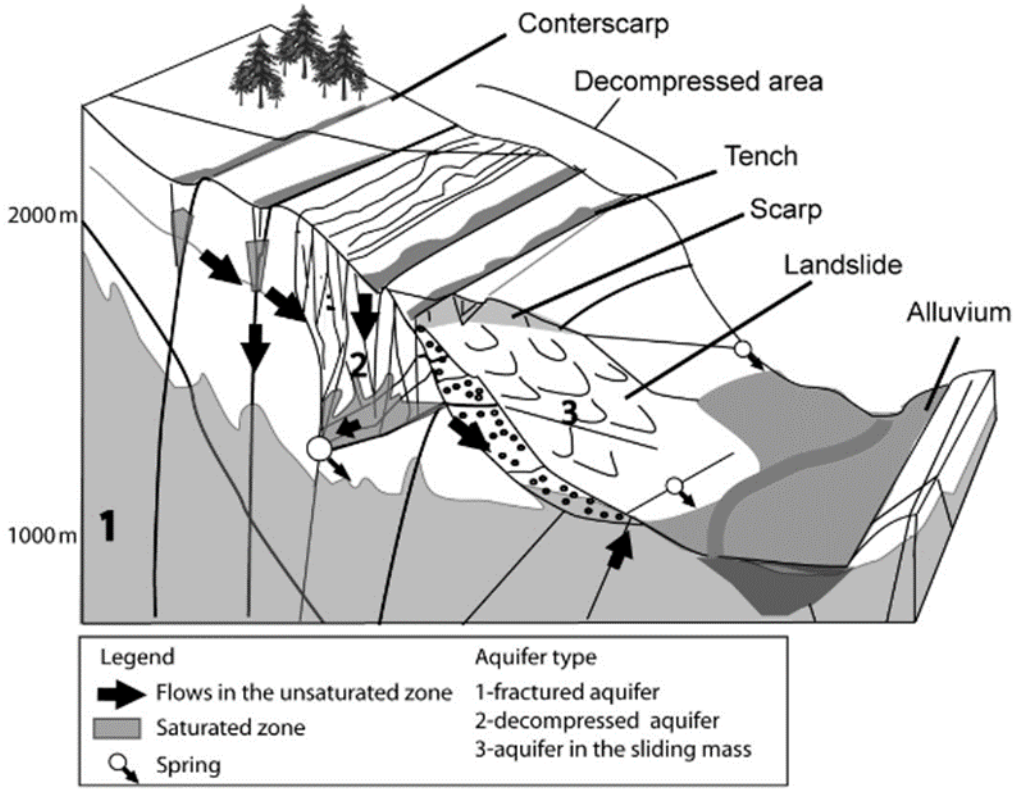


Figure 9 | Conceptual model of groundwater flows in unstable rock slopes, derived from the analysis of hillslopes in the Upper Tinée Valley (South-East French Alps) (from *Binet et al., 2007*).

The temporal dynamics of slow-moving landslides commonly exhibit a seasonal cycle, with acceleration phases correlated – sometimes with lags of several weeks *(Iverson & Major, 1987)* – with precipitation and associated increases in pore-water pressures. Landslide motion can therefore be described as a superposition of slow, continuous displacement, driven by long-term slope weakening, punctuated by transient acceleration episodes in response to hydrometeorological forcing *(Corominas et al., 2005; Vallet, 2014)*. However, displacement is spatially non-uniform, being modulated by the distribution of pore-water pressures, the local sensitivity to infiltration, and morphological features such as basal-surface geometry and the presence of major discontinuities *(Barla et al., 2010; Coe et al., 2009; Nereson & Finnegan, 2019; Schulz et al., 2017)*.

Progressive slope deformation and transient acceleration episodes also modify the fracture network though opening-closure dynamics *(Krzeminska et al., 2013)*, thereby altering flow paths and pore-water pressure distributions in both space and time *(Bonzanigo et al., 2007)*. As a result, the velocity–precipitation–pore-water pressure relationships are inherently non-linear, making displacement forecasting especially challenging.

II.4. DATA COLLECTION METHODS FOR MONITORING SLOW-MOVING LANDSLIDES

The development of displacement simulating models relies on three complementary sources of information: in situ monitoring, remote sensing, and numerical modelling (**Figure 10**).

In situ measurements provide direct observations of ground displacement and environmental conditions. Subsurface displacements can be measured using inclinometers, which capture the shear deformation and help localise the failure surface depth *(Uhlemann et al., 2016)*. Surface displacement is monitored using extensometers *(Gili et al., 2021)*, tacheometric target networks tracked by total stations *(Crosta et al., 2013)*, or Global Navigation Satellite System (GNSS) stations, which have become widespread owing to their low cost and ease of installation *(J. Wang et al., 2021)*. Environmental variables such as net rainfall, groundwater levels, and pore-water pressure are monitored using weather stations, rain gauges and piezometer networks *(Bednarczyk, 2018; Segoni et al., 2018; Wei et al., 2019)*. More occasionally, additional sensors are deployed to characterise snowmelt, water conductivity, or discharge at outlets *(Béjean-Maillard et al., 2025; Bertrand et al., 2015; Crosta et al., 2013)* (**Table 3**).

Remote sensing techniques provide high-resolution topographic and kinematic datasets. LiDAR and optical photogrammetry provide high-resolution Digital Elevation Models (DEMs) that support detailed characterisation of slope morphology (e.g., slope gradients, scarps, fracture networks, etc.) *(Handwerger et al., 2013; Travelletti et al., 2012)*. Image correlation from multi-temporal optical acquisitions and Synthetic Aperture Radar Interferometry (InSAR) enable sub-metric to centimetric detection of surface displacements *(Carlà et al., 2019; Casagli et al., 2023; Chang et al., 2025)*. Additionally, meteorological satellites provide gridded estimates of key surface variables such as surface temperature, snow cover and precipitation *(Latif et al., 2023)*.

Numerical models formalise the physical processes controlling slope dynamics. Geological models reconstruct the slope internal structure (e.g., failure planes, landslide thickness, discontinuities, geological bedding) using geotechnical and geophysical constraints *(Havenith et al., 2017; Vulliet & Hutter, 1988)*. Hydrological models simulate infiltration and pore-water pressure evolution, either spatially (e.g., modelling flow paths and water storage) *(Conte et al., 2014; Malet et al., 2005; Strauhal et al., 2016)*, or temporally, using conceptual or semi-empirical formulations to reproduce groundwater level fluctuations *(Krzeminska et al., 2013; G.-W. Lin et al., 2020; Weerasinghe et al., 2019)*. Finally, mechanical models simulate material response to hydraulic and gravitational loading through elastic, plastic, or viscous rheologies, often implemented with finite-element methods *(X. Li et al., 2023)*. These approaches generally require calibration using laboratory-derived mechanical parameters (e.g., direct shear or triaxial tests) *(Angeli et al., 1998; van Asch et al., 2007)*.

Collectively, these observations and models provide the fondation for landslide displacement modelling.

Table 3 | Common multi-source in situ monitoring systems used for slow-moving landslides (adapted from *Uhlemann et al., 2016*).

Measurement domain	Instrument	Field observable	Practical meaning (what it constrains / quantifies)
Subsurface deformation	Conventional borehole inclinometer	Casing inclination vs depth -> horizontal displacement profile (with depth)	Identifies the shear zone(s) depth and quantifies internal deformation magnitude and rate
	In-place inclinometer / SAA (Shape Acceleration Array)	Continuous 3D deformation along a borehole -> time series of displacement/strain with depth	Resolves temporal variability of internal deformation (accelerations/decelerations) and refines shear-zone activity
	AE waveguide (Accoustic Emission)	AE event rate (counts per unit time above threshold)	Proxy of shear deformation activity and relative displacement-rate changes
Surface deformation / kinematics	GNSS (RTK survey or permanent station)	3D position time series of benchmarks (X, Y, Z)	Quantifies surface displacement vectors and velocities, seasonaly cycles, and long-term trends
	Total station tacheometry (prism/target network)	Repeated 3D coordinates of multiple targets	Maps spatial displacement (changes) across the landslide body
	Tiltmeter	Local titl/rotation (angle vs time)	Sensitive indicator of local kinematic changes, particularly for rotational components or block tilting
	Extensometer / crackmeter	Relative displacement (distance change) or crack opening	Direct measure of local strain / fracture opening, useful for tracking episodic accelerations
Hydro-meteorological forcing	Automatic weather station + rain gauge	Rainfall (amount/intensity), air temperature (and optionally relative humidity, wind, solar radiation).	Provides external forcing boundary conditions driving water infiltration, recharge, and seasonal variability
Hydrogeological state	Piezometer (head / pore-water pressure sensor)	Groundwater level or pore-water pressure within a screened interval	Contrains hydro-mechanical coupling through pressure changes, thresholds, and lags relative to water infiltration
	Outlet discharge monitoring	Discharge at springs/outlets (often stage + rating curve)	Integrated signature of drainage dynamics, connectivity, and hydrological response times
	Electrical conductivity probe	Water electrical conductivity	Proxy for water mixing and mineralisation, informating flow-path changes and residence-time variations

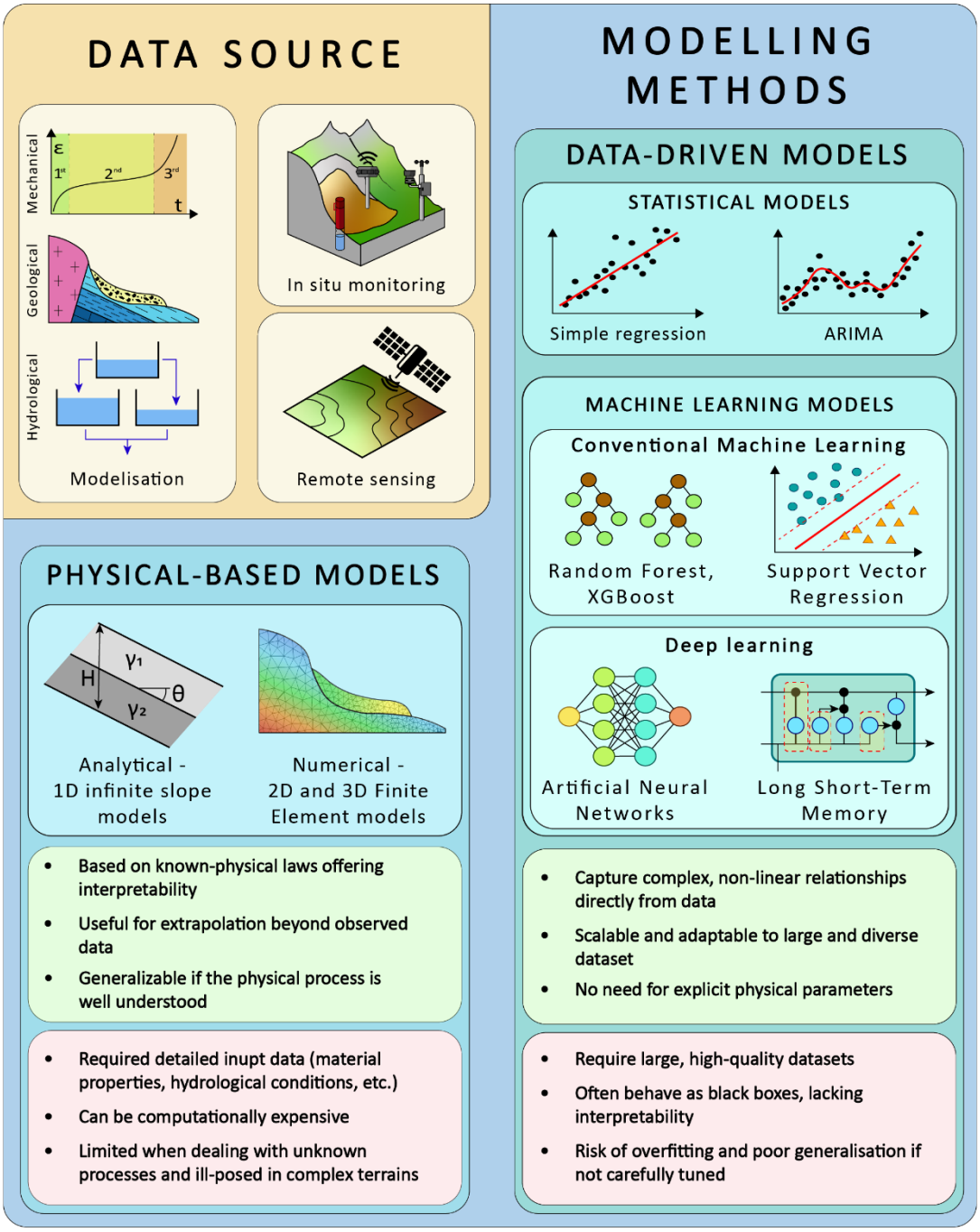


Figure 10 | Synthetic representation of landslide monitoring data acquisition methods and displacement modelling approaches. Green and red boxes summarise the main positive and negative aspects of each methods.

II.5. SIMULATING MODELS FOR SLOW-MOVING LANDSLIDE DISPLACEMENT

Landslide displacement dynamic can be modelled using two main approaches: physically-based models, which explicitly represent mechanical processes, and data-driven models, which rely on statistical or probabilistic relationships among observed variables *(Gong et al., 2024)* (**Figure 10**). The choice between these approaches depends largely on the modelling objectives, as well as the availability, quality, and resolution of the observational data. Although they are rooted in different methodologies, they are best viewed as complementary rather than mutually exclusive. In recent years, data-driven approaches have gained significant popularity in landslide forecasting, particularly because they can capture complex, non-linear system responses from multi-parameter observations *(Gong et al., 2024)*.

II.5.1. Physically-based models

Physically-based models aim to simulate landslide behaviour by explicitly representing geomorphological, geological, hydrological and mechanical properties.

Early approaches relied on simplified analytical formulations. Among these, the 1D-infinite slope model remains the most widely used to assess slope stability through the computation of a factor of safety *(Iverson, 2000; Lu & Godt, 2008)*. This model assumes several simplifying assumptions, such as a failure plane parallel to the ground surface, spatially uniform pore-water pressure and homogeneous material properties *(Conte et al., 2017; Herrera et al., 2009; Wienhöfer et al., 2011)*. Although effective for shallow landslides, these assumptions are rarely appropriate for slow-moving landslides, which typically exhibit strong heterogeneity in internal structure and geomaterial properties *(Gong et al., 2024)*.

With recent advances in computational methods, physically-based modelling has increasingly relied on numerical solutions, notably finite element methods (FEM) and finite difference (FDM) methods. These techniques discretise the unstable mass into a mesh of interconnected elements, enabling the simulation of complex mechanical interactions across space and time. Compared with analytical models, they can accommodate realistic 2D or 3D geometries and incorporate heterogeneous materials properties *(Conte et al., 2014; Havenith et al., 2017; Vassallo et al., 2015)* (**Figure 10**).

While particularly suited for reproducing the spatio-temporal evolution of displacement, such models require comprehensive and detailed datasets, including landslide geometry, geotechnical parameters, hydrological properties, and constitutive laws. Such parameters typically involve extensive field campaigns, numerical modelling and laboratory testing, which are costly and time-consuming. Furthermore, the computational cost associated with high-resolution numerical simulations are substantial. As a result, physically-based models are generally restricted to well-instrumented sites with extensive monitoring and well-characterised subsurface conditions.

II.5.2. Data-driven models

Data-driven models offer an alternative to physically-based approaches by bypassing the need to explicitly model the underlying mechanical processes. These models rely on identifying statistical or probabilistic relationships between multivariate time series and aim to predict landslide displacements (the target variable) based on influential external factors. Consequently, they require continuous, high-quality time series data recovering sufficiently long periods to capture system dynamics.

The target variable is typically derived from surface displacement measurements, whose availability has increase significantly with the development of GNSS systems and remote sensing techniques *(Benoit et al., 2015; Glueer et al., 2019; Zou et al., 2020)*. While these data only provide partial insight into the full deformation mechanism, they offer robust empirical basis for predictive modelling.

The most commonly used predictive variables include rainfall, reservoir level fluctuations, and groundwater level or pore-water pressure changes *(Fiolleau et al., 2020; Malet et al., 2002; Mondini et al., 2023; S. Xu & Niu, 2018)*. Other environmental drivers, such as snowmelt or regional climatic indices, are more rarely considered *(Crosta et al., 2013; Krkač et al., 2017)*.

Due to the internal heterogeneity and complex mechanical behaviour of slow-moving landslides, the external forcing and measured displacement relationship is often non-linear, multivariate, and subject to time lags (e.g., the slope's response to rainfall may be strongly conditioned by its antecedence saturation state). This complexity requires incorporating sequential data characteristics, memory effects, and variable interactions into the modelling framework.

Data-driven models can be broadly categorised into two main groups: statistical models and machine learning models. Statistical models, such as ARIMA, have historically been used for short-term forecasting of linear time series *(Aggarwal et al., 2020)*. However, their capacity to reproduce complex or non-linear dynamics remains limited, making them unsuitable for slow-moving landslide forecasting unless integrated with more advanced techniques such as deep learning models *(Shao et al., 2024)*.

Progress in computer science has facilitated the development of machine learning models able to autonomously learning complex, non-linear relationships from multivariate datasets without requiring explicit programming. Overall, these models are commonly distinguished between deep learning (DL) models, which rely on deep neural network architectures, and conventional machine learning (CML) models, including decision trees and support vector machines (**Figure 10**).

The following sections aim to detail the principles and contributions of these two approaches. In particular, deep learning models are first introduced to highlight the limitations inherent to their core operating mechanisms, thereby providing a basis to emphasise how conventional machine learning models can offer practical solutions to address these constraints.

II.5.2.1. Deep learning models

To capture the non-linearity and temporal dependence inherent in landslide displacement time series, Artificial Neural Networks (ANNs), originally proposed by *Mcculloch & Pitts (1943)*, have merged as a powerful and flexible modelling approach. Their architecture typically consists of an input layer, one or more hidden layers, and an output layer, where interconnected artificial neurones propagate information through non-linear activation functions (**Figure 11**).

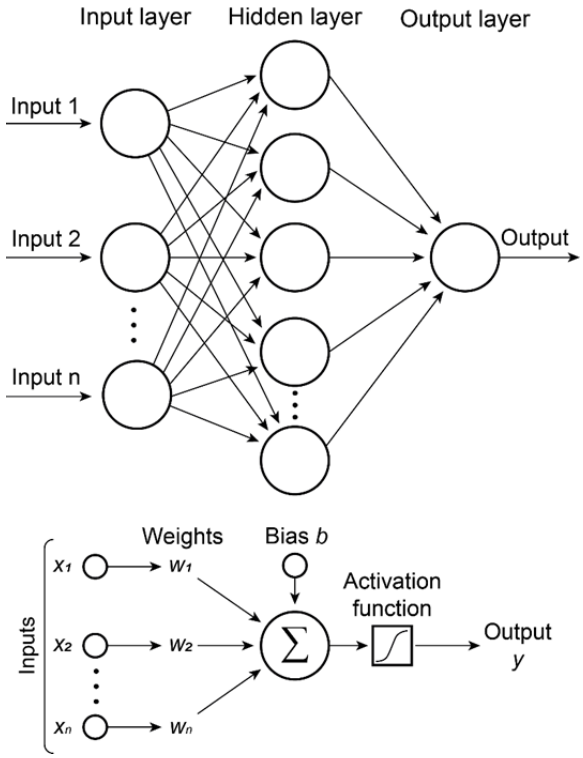


Figure 11 | Schemes for the structure description of the artificial neural networks (ANNs): **(a)** Description of neural network architecture. **(b)** Description of neural network active node (adapted from *Bekesiene et al., 2021*).

Several ANN variants have been developed to handle the sequential nature of time series data. Backpropagation Neural Networks (BPNNs) have been applied to landslide displacement forecasting *(H. Chen & Zeng, 2011; J. Du et al., 2013)*, but Recurrent Neural Networks (RNNs) are more explicitly designed to represent temporal dependence by introducing feedback connections that allow information to persist across time steps *(Z. Ma et al., 2021; J. Wang et al., 2021; Y. Zhang et al., 2021)* (**Figure 12**). However, conventional RNNs are prone to vanishing and exploding gradient during training, which limits their ability to learn long-term dependencies.

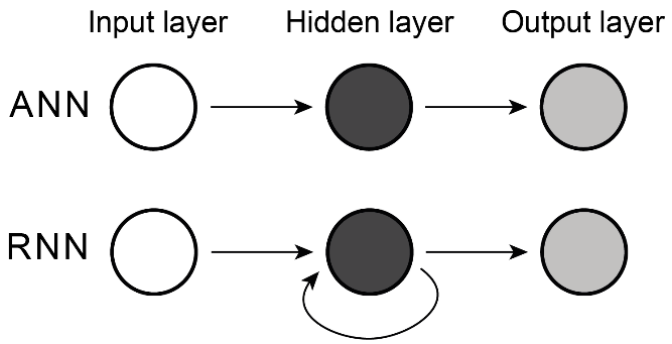


Figure 12 | Traditional versus Recurrent Neural Networks (adapted from *Kıraç et al., 2019*).

To address these limitations, more advanced and time-aware architectures have been introduced. Long Short-Term Memory (LSTM) network *(H. Jiang et al., 2020; Y. Wang et al., 2022; Xie et al., 2019; B. Yang et al., 2019)* and Gated Recurrent Units (GRU) *(Meng et al., 2024; Nava et al., 2023; B. Yang et al., 2022; W. Zhang et al., 2022)* extend the RNN framework by incorporating gated memory cells that regulate information flow through input, output, and forget mechanisms (**Figure 13**). This design allows models to selectively retain or discard information over multiple timescales, thereby capturing both short- and long-term dependencies. As a result, LSTM and GRU models generally outperform conventional RNNs and are now widely used in landslide displacement forecasting.

Despite their potential, deep learning models face several recurring challenges:

- They require large datasets to enable robust training, limit overfitting, and stabilise optimisation. This is often restrictive in landslide applications, where monitoring time series are frequently short, noisy, and discontinuous;
- Their black-box nature limits interpretability, as the links between inputs and predictions remain difficult to trace, which can hinder trust and operational uptake (e.g., in early warning systems).
- Their outputs arise from the optimisation of weight and biases and therefore lack an explicit physical meaning, unless additional contraints are introduced.

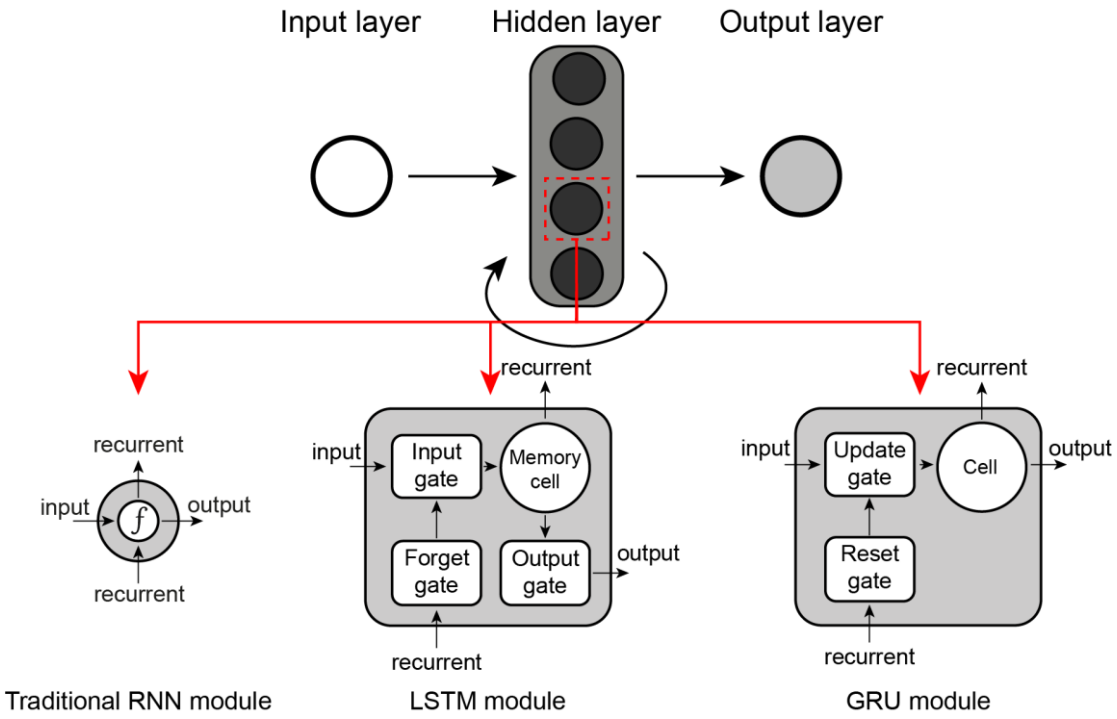


Figure 13 | Structures of a traditional RNN, a LSTM and a GRU modules (adapted from *Liu et al., 2020*).

To address these limitations, recent work has broadly followed two directions. The first is to introduce physical constraints directly into the learning process. Physics-Informed Neural Networks (PINNs), introduced by *Raissi et al. (2019)*, incorporate governing equations into the loss function, thereby encouraging model predictions to satisfy physical laws. *Moeineddin et al. (2023)*, for instance, used PINNs to model temperature variations along a basal shear surface, as a precursor to catastrophic failure. Similarly, *Liu et al. (2024)* developed the PIDA framework (Physics-Informed Data Assimilation), which dynamically updates model parameters through a hybrid combination of physical rules and data assimilation, and reported improved performance relative to conventional LSTM and GRU models.

The second direction focuses on improving transparency. Transformer-based architectures, initially developed for computer vision and natural language processing, have recently been adapted for landslide applications *(Dahal & Lombardo, 2023; Ge et al., 2024; Z. Ma & Mei, 2025; J. Sun et al., 2022; Wu et al., 2025)*. These models rely on self-attention mechanism, that can, in principle, highlight which past observations contributes most strongly to a prediction. For example, *Ge et al. (2024)* introduced LiteTransNet, a lightweight Transformer capable of forecasting displacements while visualising temporal dependencies through attention maps. *Ma & Mei (2025)* further incorporated spatio-temporal correlations among multiple monitoring points to capture both global displacement trends and local heterogeneities, and used a Temporal Fusion Transformer, to identify key drivers of slope deformation.

Despite these advances, deep learning approaches remain data-intensive and computationally demanding, and their generalisation to new sites is still often limited by model complexity, training costs, and the scarcity of long, homogeneous monitoring records.

II.5.2.2. Conventional machine learning models

Conventional Machine Learning (CML) approaches provide a valuable alternative to deep learning models for landslide displacement modelling. They include widely used algorithms such as Support Vector Machines (SVM), Decision Trees (DT), Random Forests (RF), and eXtreme Gradient Boosting (XGBoost). Their lower sensitivity to dataset size, faster training times, and, in many cases, greater interpretability make them well suited for field-based time series, which are often short, noisy, and incomplete *(Chauhan & Singh, 2018; J. Xu et al., 2022a)*. Some studies further suggested that, under certain configurations, CML models can match or outperform DL models in terms of predictive accuracy and robustness *(Y. Wang et al., 2022; W. Zhang et al., 2022)*.

Within the landslide literature, CML models have been most widely applied to susceptibility mapping. In this context, the problem is framed as a supervised classification, where each spatial unit (e.g., pixel or polygon) is assigned to a susceptibility class (e.g., low to very high), based on static predictors such as slope gradient, lithology, or distance to faults *(Catani et al., 2013; Kavzoglu & Teke, 2022; Lee et al., 2017; Sahin, 2020; Samodra et al., 2024)*.

By contrast, the use of CML in displacement forecasting remains comparatively limited. Unlike susceptibility mapping, time series prediction requires capturing dynamic dependencies between variables. This is challenging because most CML models do not include an intrinsic memory mechanisms, unlike recurrent architectures such as RNNs or LSTMs. As a result, CML forecasting typically relies on an explicit feature engineering step *(Chauhan & Singh, 2018)*, where raw forcing time series (e.g., rainfall, groundwater level) are derived into predictive descriptors through statistical aggregations (e.g., rolling means, cumulative sums), lagged variables, or event-based indicators. These calculations are needed to represent short-, medium-, and long-term behaviours, and to encode the non-linear, multi-temporal nature of landslide responses.

This phase can be time-consuming and methodologically demanding, which partly explains the appeal for deep learning models that can, in practice, learn these representations directly from raw inputs or with limited feature engineering.

Nevertheless, some studies have explored CML approaches for time series forecasting, often in comparative frameworks. For instance, *Y. Wang et al. (2022)* compared five CML models optimised with Particle Swarm Optimisation (PSO), with an LSTM model, and found that the PSO-LLSVM and PSO-SVM configurations yielded more stable and coherent results than LSTM. Similarly, *(J. Xu et al.,*

2022a) evaluated XGBoost and SVR for hourly displacement forecasting and reported better performance than RNNs when data availability was limited.

Among CML models, decision tree ensembles, particularly Random Forests and XGBoost, are frequently adopted. Their structure supports intuitive visualization of decision processes, while ensembling improves predictive skill and limits overfitting. In addition, these models provide native estimates of predictor importance, thereby enhancing interpretability and transparency, two crucial aspects in hazard assessment applications. For instance, *Krkač et al. (2017)* applied a Random Forest model to forecast daily displacements at the Kostanjek landslide (Serbia), using predictors derived from observed and simulated groundwater levels, and used variable importance rankings to inform interpretation of the hydro-mechanical response.

II.5.2.2.1. Decision trees and ensemble methods

Decision trees are non-parametric supervised learning algorithms applicable to both classification and regression tasks. Their operate by recursively partitioning the predictor space into increasingly homogeneous regions, using simple conditional rules of the form "if… then" *(Fürnkranz, 2011)*. Each internal node (child and split nodes) corresponds to a test on a predictor (e.g., cumulative rainfall over 10 days < 60 mm), each branch represents an outcome of that test, and each terminal leaf provides a prediction – typically the mean value in regression, or the majority class in classification (**Figure 14**).

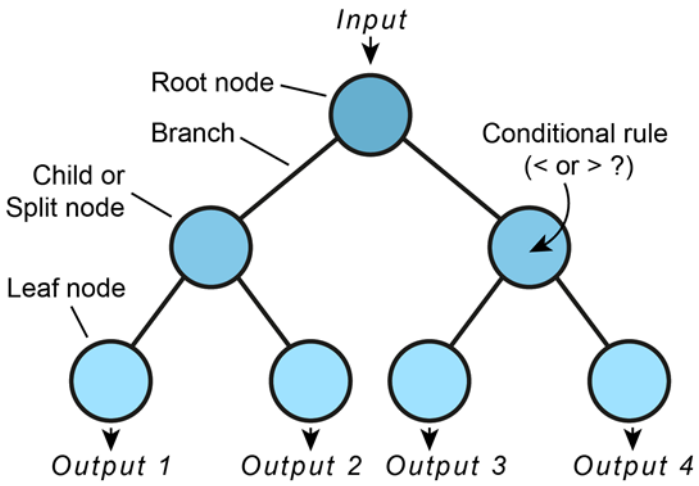


Figure 14 | Structure of a decision tree.

The construction of the tree seeks to maximise within-node homogeneity by minimising an impurity measure – such as Gini index, entropy, or variance, depending on the learning task. This process is generally performed recursively through a greedy strategy, selecting at each step the predictor that yields the greatest reduction in impurity.

Decision trees offer several advantages: their are intuitive, straightforward to interpret, and can accomodate both categorial and continuous predictors. However, their generalisation capacity is limited. A single tree is prone to overfitting, especially for noisy or sparse datasets, and often exhibits high variance, which can reduce predictive accuracy *(James et al., 2013)*.

Ensemble learning strategies were developed to address these limitations by combining multiple models to reduce bias and variance *(Sagi & Rokach, 2018)*. Two main groups are commonly distinguished: bagging, which trains models independentely in parallel, and boosting, which builds models sequentially in an adaptative manner (**Figure 15**).

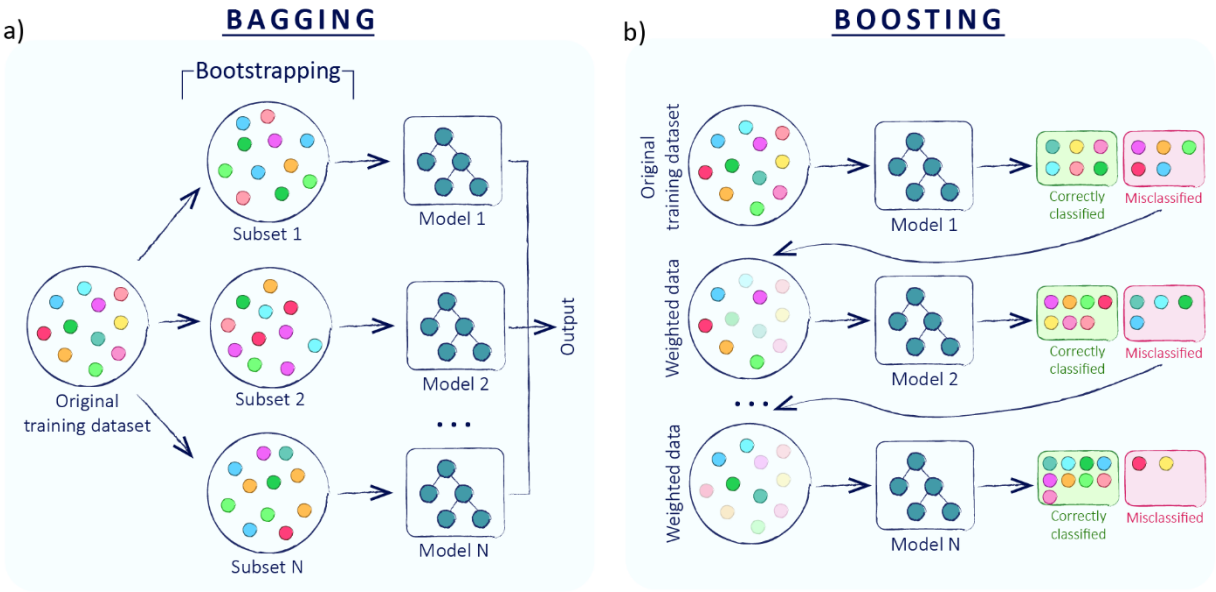


Figure 15 | Schematic illustration of **(a)** Bagging and **(b)** Boosting workflows (adapted from *González et al., 2020).*

Among bagging approaches, Random Forests *(Breiman, 2001)* are widely used. They aggregate predictions from multiple trees trained on bootstrap resamples of the training data. Additional diversity is introduced by considering a random subset of predictors at each split, which limits the dominance of highly predictive variables and reduces correlation among trees. Random Forests are robust, and generally insensitive to hyperparameter setting, but they have a structural limitation: trees are built independently, and therefore do not explicitly correct one another's errors. Boosting strategies address this gap by iteratively refining the ensemble through sequential error correction.

II.5.2.2.2. EXtreme Gradient Boosting (XGBoost)

Gradient boosting is an ensemble learning strategy that builds decision trees sequentially, with each new tree designed to improve the current model by correcting its remaining errors. Unlike bagging, boosting is explicitly error-driven. At each iteration, a new tree is fitted to the residuals, that is, the discrepancy between the observations and the predictions of the current ensemble (**Figure 16**). The procedure is formulated as an optimisation probleme and is typically solved through gradient descent, where the loss is progressively reduced by following the gradient of the objective with respect to the model predictions. For regression, the loss is often defined as the sum of squared residuals *(Friedman, 2001)*.

Among gradient boosting implementations, eXtreme Gradient Boosting (XGBoost) (**Figure 16**), introduced by *T. Chen & Guestrin (2016)*, has become a widely used reference for tabular data owing to its strong predictive performance, robustness, and efficient implementation *(Bentéjac et al., 2021; Gómez-Ríos et al., 2017; Sahin, 2020)*. XGBoost minimises an objective function composed of two terms:

- A loss function $L(y_i, F(x_i))$ measuring the discrepancy between the observed values y_i and the model prediction $F(x_i)$;
- A regularisation term Ω, penalising model complexity.

The objective function can be written as (**Equation 3**):

$$Obj = \sum_{i=1}^{N} L(y_i, F(x_i)) + \sum_{m=1}^{M} \Omega(h_m) \quad (3)$$

Where N is the number of observations, M the number of trees, and h_m the m-th tree in the sequence. The regularisation term is defined as (**Equation 4**):

$$\Omega(\mathrm{h}) = \gamma T + \frac{1}{2}\lambda\|w\|^2 \quad (4)$$

Where:

- γT represents the tree score, with T being the number of leaves in the tree, and γ the structural complexity penalty parameter;
- $\lambda\|w\|^2$ is the regularisation term, where λ is the L2 regularisation coefficient and $\|w\|^2$ is the L2 norm of the scores (or weights) assigned to each leaf.

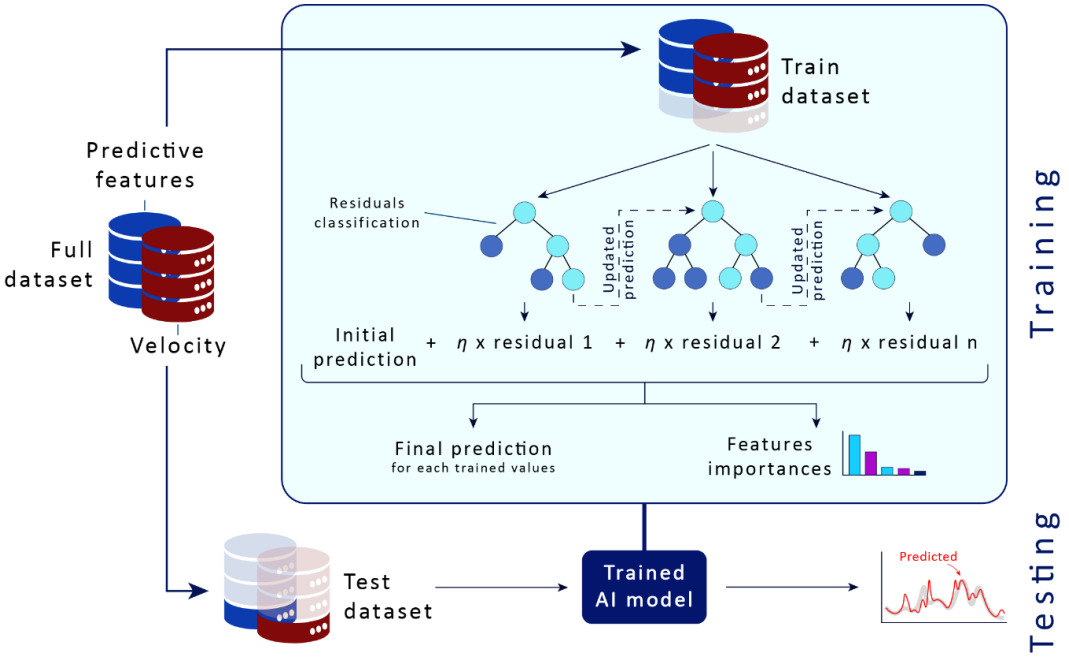


Figure 16 | Conceptual representation of XGBoost functioning in a regression task. η represents the learning rate applied to each output of a regression tree.

This term plays a central role in controlling model complexity: γ penalises excessively deep trees, favouring more compact structures, while the L2 penalty constrains the magnitude of the leaf scores, reducing the risk of overfitting to noise or outliers.

Beyond regularisation, XGBoost includes implementation-level optimisations that reduce training cost. These notably include efficient handling of feature values through column-wise storage and reuse of sorted structure, and an approximate greedy split algorithm, that discretises predictor distributions into bins, limiting the number of candidate thresholds evaluated at each split *(Bentéjac et al., 2021)*. XGBoost also supports tree pruning, whereby weak or non-contributory branches are removed to further improve generalisation.

II.6. CHAPTER II – CONCLUSION

This chapter has outlined the societal and scientific context that underpins the need for reliable landslide forecasting, with a focus on slow-moving, deep-seated slope movements. These landslides occur across a wide range of geomorphological settings and can generate substantial and persistent impacts on exposed populations and infrastructure. Forecasting objectives remain inherently multi-dimensional, spanning the questions of where instability may develop, when hazardous acceleration phases may occur, and how displacement evolves under the coupled influence of geological structure, geomorphological setting, and hydro-meteorological forcing.

By revisiting the main controlling factors of slope instability, this chapter emphasised the central role of hydro-mechanical coupling, whereby infiltration-driven pore-water pressure variations reduce effective stress and modulate shear strength. In deep-seated contexts, this response is strongly conditioned by heterogeneous internal structures, preferential flow paths, and complex groundwater configurations, which together promote non-linear, multivariate, and time-lagged forcing–response relationships. As a result, the kinematics of slow-moving landslides are best viewed as a superposition of long-term deformation and transient acceleration episodes, whose timing and magnitude depend on both external forcing and the evolving internal state of the slope.

The chapter further reviewed the main strategies available to characterise and model such systems, highlighting the complementary roles of in situ monitoring, remote sensing, and numerical approaches. Physically-based models provide process-based representations of hydrological and mechanical mechanisms and are particularly valuable when site geometry and parameters are well constrained. However, their applicability remains limited by the need for extensive site characterisation and the cost of high-resolution simulations. Data-driven approaches offer an alternative by learning empirical relationships from multivariate time series, but their performance and robustness rely on the availability of continuous, high-quality records that adequately capture the diversity of kinematic regimes and forcing conditions.

Finally, the chapter positioned recent AI-based developments within this modelling landscape. Deep learning architectures can reproduce complex, multi-temporal behaviours, yet they remain constrained by data intensity, computational demands, and limited transparency. In contrast, conventional machine-learning models – particularly tree-based ensemble methods such as Random Forests and XGBoost – often provide a favourable balance between predictive skill, robustness to field-data limitations, and interpretability. Within displacement forecasting applications, their effectiveness nevertheless depends on constructing predictors that explicitly encode temporal dependence (e.g., lags,

cumulative descriptors, or event-based indicators), thereby translating hydro-meteorological variability into informative representations of slope response.

Because these predictors are derived from the forcing time series, CML model performance is therefore strongly conditioned by the quality, continuity, and representativeness of the source data. This, in turn, implies combining these models with dedicated data preparation and analysis procedures to ensure that both the simulations and the inferred forcing–velocity relationships are robust. ***Chapter III*** consequently introduces the key methodological components of this framework, which provide the necessary basis for simulating slow-moving landslide displacement time series using CML approaches.

CHAPTER III

CONVENTIONAL MACHINE LEARNING FRAMEWORK FOR SIMULATING SLOW-MOVING LANDSLIDE DISPLACEMENTS

CHAPTER III CONVENTIONAL MACHINE LEARNING FRAMEWORK FOR SIMULATING SLOW-MOVING LANDSLIDE DISPLACEMENTS 51

III.1. Chapter III – Introduction 53

III.2. Data preparation and exploratory analysis of field monitoring time series 54

III.2.1. Time series collection and preparation 54

III.2.2. Exploratory statistical analysis 55

III.3. XGBoost modelling strategy steps 56

III.3.1. Feature engineering 56

III.3.2. Train-test splitting 57

III.3.3. Hyperparameter tuning 57

III.3.4. Model performance evaluation 59

III.4. Model interpretability and explainability 60

III.4.1. XGBoost feature importance ranking 61

III.4.2. SHAP post-hoc analysis 62

III.5. Chapter III – Conclusion 64

III.1. CHAPTER III – INTRODUCTION

Building on the background presented in ***Chapter II***, this chapter introduces the methodological framework required to apply robust conventional machine learning models to slow-moving landslide displacement time series. As discussed previously, tree-based ensemble models such as Random Forests and XGBoost provide a favourable compromise between predictive skill, robustness to field-data limitations, and interpretability, but their performance remains strongly conditioned by the quality, consistency, and representativity of the input observations and by the capacity of the predictors to encode temporal dependence. **Figure 17** synthesises the main processing steps adopted in this manuscript, from raw-data pre-processing and exploratory analysis to feature engineering, model calibration, and performance assessment, complemented by an interpretability-explainability layer when transparency is required. The following sections detail each of these steps and discuss practical considerations for time-series applications in landslide displacement modelling.

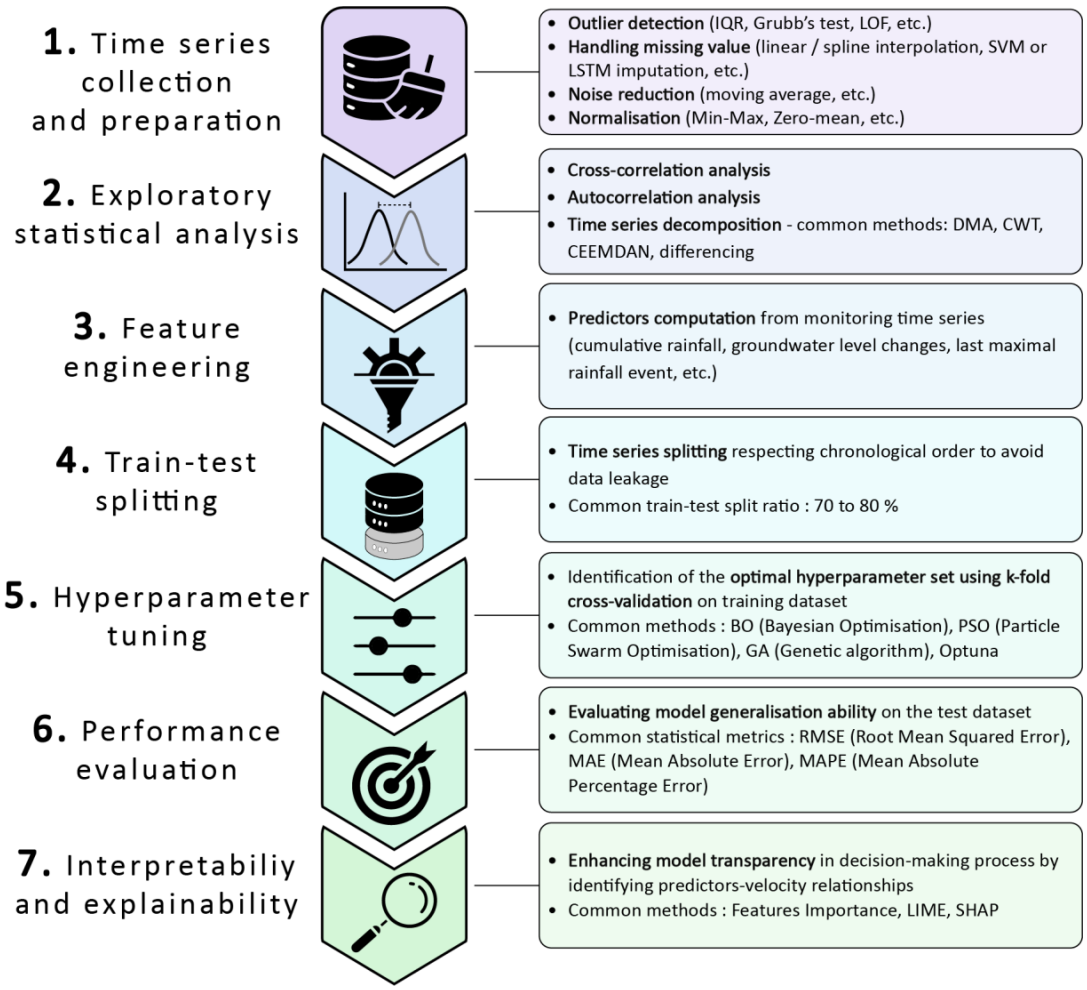


Figure 17 | Workflow for time-series processing and conventional machine-learning model implementation.

III.2. DATA PREPARATION AND EXPLORATORY ANALYSIS OF FIELD MONITORING TIME SERIES

III.2.1. Time series collection and preparation

Field monitoring time series acquired from in situ sensors or remote sensing are often affected by acquisition limitations, including temporal gaps, outliers, instrumental noise, measurement errors, and occasional sensor malfunctions. If not properly addressed, these issues can degrade model training and bias predictions. A preliminary cleaning phase is therefore essential to ensure that the time series are representative of both the displacement signal and the associated environmental factors.

In practice, three types of operations are commonly applied:

- **Outlier Detection and treatment –** Outliers can be identified using classical statistical criteria such as InterQuartile Range (IQR) filters or Z-score thresholds *(Chikodili et al., 2021; Mramba et al., 2024)*, as well as more robust approaches including Isolation Forest, Grubb's test, or Local Outlier Factor *(Blázquez-García et al., 2020; Z. Cheng et al., 2019)*.
- **Missing values handling –** Linear or spline interpolation can be used for short data gaps, whereas longer gaps in more complexe dependencies may require dedicated imputation approaches (SVM-based imputation or LSTM-based imputation, etc.) *(Pratama et al., 2016; Tawakuli et al., 2025)*.
- **Noise reduction –** Smoothing methods, such as moving averages or adaptative filters, are often applied to reduce non-informative variability while preserving the underlying signal dynamics *(Abi Nader et al., 2023; Krkač et al., 2017)*.

Normalisation or standardisation is also frequently used in machine learning pipelines to harmonise variable scales. However, this step is not required for tree-based models, which are invariant to monotonic rescaling of the inputs because they rely on recursive partioning rules rather than distance metrics or scale-sensitive gradients *(James et al., 2013)*.

III.2.2. Exploratory statistical analysis

Exploratory statistical analysis aims to characterise the main statistical properties of the variables by examining their distributions, identifying cross-correlations between variables, quantifying internal autocorrelations, and detecting potential seasonal and trend components. This step directly informs the subsequent feature engineering strategy.

Cumulative displacement time series often exhibit periodic variations linked to reservoir water levels fluctuations or seasonal groundwater level changes *(He et al., 2008; Vallet et al., 2016; van Asch et al., 2009)*, superimposed on long-term trend associated with slow, progressive internal slope deformation *(D. Bai et al., 2022; D. Huang et al., 2022; Niu et al., 2021)*. Such trends can amplify periodic signals, and introduce non-stationarity, manifested as time-varying statistical properties (e.g. mean, variance) *(Chatfield, 2013; Shumway & Stoffer, 2011)*. Stationarity is a prerequisite for many statistical models and generally improves the stability of predictive performance in machine learning applications *(Petrică et al., 2017)*.

To address this issue, cumulative displacement time series are often decomposed into trend and periodic components *(Montesino Pouzols & Lendasse, 2010)*. Proposed approaches include Double Moving Average (DMA) *(Xing et al., 2019)*, and additive or multiplicative decomposition *(B. Yang et al., 2019)*. More advanced methods aim to separate multiple time-frequency components to capture temporal irregularities, including CEEMDAN, VMD, or CWT-based approaches *(H. Li et al., 2020; Meng et al., 2024; Vallet et al., 2016; R. Wang et al., 2022)*.

However, these decomposition strategies require methodological choices regarding window size and filter parameters, that can introduce additional uncertainty and unnecessary complexity. As an alternative, *Nava et al. (2023)* proposed a first-order differencing, which computes changes between successive observations. This simpler and more reproducible approach is equivalent to working directly with velocity time series.

III.3. XGBOOST MODELLING STRATEGY STEPS

III.3.1. Feature engineering

Feature engineering consists of calculating explanatory predictors from the original forcing time series (e.g., rainfall, groundwater level) based on expert knowledge of the mechanisms controlling landslide motion. The aim is to represent the system state more explicitly and to help the model learn non-linear relationships and delayed responses. Common transformations include the computation of rolling statistics (e.g., mean, maximum, cumulative sums), first derivatives to capture temporal changes, and event-based descriptors targetting extremes or thresholds (e.g., intensity, duration, frequency) *(Krkač et al., 2017; R. Wang et al., 2022)*.

To account for system memory and temporal dependence, features are typically computed over multiple time scales, allowing short-, medium- and long-term effects to be represented. Additionally, introducing time lags in feature computation enables the modelling of latency between an external forcing (e.g., rainfall infiltration) and the velocity, as identified from cross-correlation analyses (**Figure 18**).

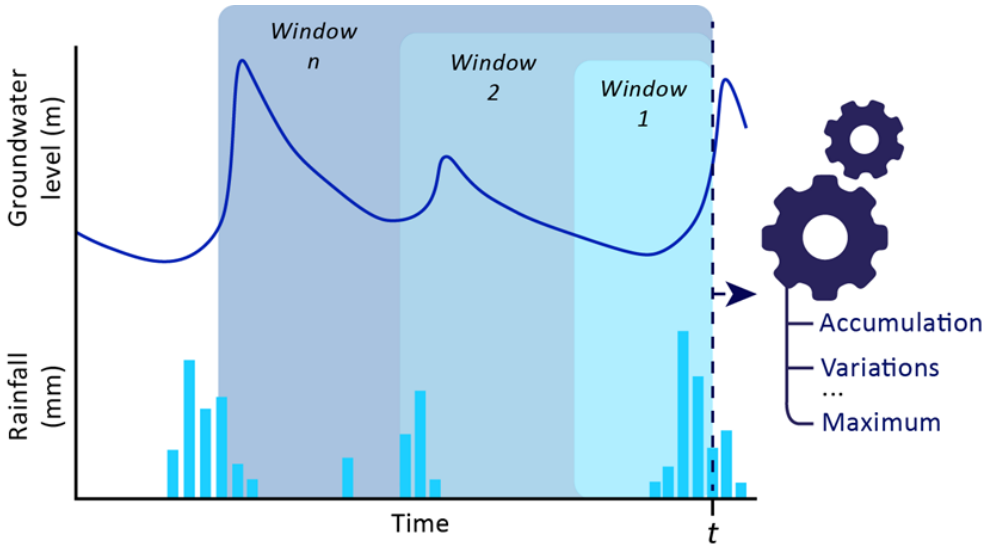


Figure 18 | Conceptual representation of the method for extracting descriptive features from rainfall and groundwater level time series over multiple time windows.

III.3.2. Train-test splitting

Splitting data into a training and a testing sets is a fundamental step in machine learning. For time series applications, the temporal structure must be strictly preserved to avoid data leakage that would bias model performance estimates *(Ebrahim, Fares, et al., 2024)*. Train-test splits should therefore not be performed randomly: observations must be remain in chronological so that the test set represents genuinely unseen conditions relative to the training period *(Lyu et al., 2021)*. This strategy provides realistic assessment of the model's capacity to generalise beyond the calibration interval.

The proportion allocated to training and testing depends on the length of the record and the presence of seasonal or cyclic behaviour. In practice, the training set should span at least one representative cycle of system variability – for instance, a full hydrological year when seasonality is pronounced. Conversely, an excessively long training set can increase the risk of overfitting and leave too little data for an informative evaluation. Although no single split ratio is universally accepted, many studies adopt a training proportion between 65% and 90%, with 70% to 80% being the most common range *(Ebrahim, Fares, et al., 2024; Ebrahim, Gomaa, et al., 2024)*.

Finally, hyperparameter tuning must be carried out using the training data only, prior to testing, to prevent leakage. This is typically achieved through k-fold cross-validation procedure (or time series-aware validation schemes) restricted to the training period. For models such as XGBoost, the relatively large hyperparameters space often motivates the use of dedicated optimisation strategies to identify robust configurations efficiently.

III.3.3. Hyperparameter tuning

Hyperparameter optimisation is a key step for improving the predictive performance and generalisation of XGBoost models. Early strategies rely on Grid Search (GS) or Random Search (RS), which evaluate parameter combinations either exhaustively or by random sampling. Although straightforward to implement, these approaches are computationally expensive and prone to convergence toward local minima of the loss function, without guaranteeing optimal performance *(Chowdhury et al., 2022; Q. Huang et al., 2012)*. To reduce computational cost and improve search efficiency, more advanced methods have been introduced. Bayesian Optimisation (BO), for instance, builds a probabilistic surrogate on the loss function – often using a Gaussian Process – to guide the most promising hyperparameters combinations selection *(Chang et al., 2025; L. Liu et al., 2024)*. Although BO can significantly reduce the number of trials required, its performance may degrade as the dimensionality of the search space increases.

Bio-inspired approaches such as Particle Swarm Optimisation (PSO) *(Han et al., 2021; Lian et al., 2013)* and Genetic Algorithm (GA) *(Cai et al., 2016)* have also been widely adopted. These methods have demonstrated a strong ability to efficiently explore the hyperparameter space and avoide local minima. Their success lies in their robustness and overall computational efficiency *(Gong et al., 2024; Han et al., 2021)*. However, their main limitations are the sensitivity to their own meta-paramaters and, in some case, substantial computational demand *(Vaiyapuri, 2025)*.

More recently, Optuna has emerged as a practical alternative for automated tuning *(Akiba et al., 2019)*. Optuna implements a BO-inspired strategy based on the Tree-structured Parzen Estimator (TPE) and includes pruning mechanisms that stop unpromising trials early. In existing applications, this strategy yield performance comparable to, or better than, PSO and GA at reduced computational cost *(N. Lin et al., 2023; Vaiyapuri, 2025)*. Nonetheless, Optuna-based optimisation for regression problems in landslide displacement forecasting remains largely unexplored *(M. Huang et al., 2024)*.

XGBoost poses a set of tunable hyperparameters controlling tree complexity (e.g., number of trees and maximum depth), stochasticity (e.g., subsampling of observations and predictors), and regularisation strength. Admissible ranges are bounded for some parameters, and default values are povided by the implementation. These hyperparameters, together with their default settings and functional meaning, are summarised in **Table 4**.

Table 4 | Main hyperparameters used in the XGBoost model (adapted from *Tarwidi et al., 2023*).

Hyperparameter	Range	Default value	Description
n_estimators	[1, ∞)	100	Maximum number of boosting rounds (trees)
max_depth	[1, ∞)	6	Maximum depth of each tree
learning_rate	[0, 1]	0.3	Step size applied at each boosting iteration
colsample_bytree	[0, 1]	1	Fraction of features sampled to build each tree
subsample	[0, 1]	1	Fraction of training instances sampled to build each tree
min_child_weight	[0, ∞]	1	Minimum sum of instance weights required in a child node
gamma	[0, ∞]	0	Minimum loss reduction required to make a further partition on a leaf node
alpha	[0, ∞]	0	L1 regularisation term on weights
lambda	[0, ∞]	1	L2 regularisation term on weights

Kavzoglu & Teke (2022), for instance, applied an optimised XGBoost model for shallow landslide spatial prediction, and reported a ~ 13% gain in predictive accuracy after tuning hyperparameters using multiple optimisation strategies, including BO, RS, and GA. Based on literature review, they focused on a subset of influential hyperparameters : *learning_rate, max_depth, min_child_weight, gamma, colsample_by_tree, n_estimators, and subsample*. More generally, both the

choice of hyperparameters to tune and the definition of their search ranges (i.e., the lower and upper bounds) are problem-specific and should be guided by prior studies or expert knowledge, as these choices directly condition the optimisation outcome.

III.3.4. Model performance evaluation

The performance of a regression model is typically assessed by quantifying the discrepancy (i.e., residual errors) between predicted and observed values. The most commonly reported metrics are the Root Mean Squared Error (RMSE) (**Equation 5**), the Mean Absolute Error (MAE) (**Equation 6**), and the Mean Absolute Percentage Error (MAPE) (**Equation 7**) *(Ebrahim, Fares, et al., 2024)*. These metrics are defined as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - Y_i)^2} \tag{5}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|X_i - Y_i| \tag{6}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{X_i - Y_i}{X_i}\right| \tag{7}$$

Where X_i denotes the observed value, Y_i the predicted value, and N the total number of observations.

Other measures, such as the Mean Squared Error (MSE) or the coefficient of determination (R^2), are also widely used in machine learning. However, they are less frequently reported in landslide displacement forecasting studies. *(Ebrahim, Fares, et al., 2024)*.

III.4. MODEL INTERPRETABILITY AND EXPLAINABILITY

Interpretability aims to make machine learning models more transparent by clarifying the logical that underpins their predictions. Complementarily, explainability focuses on providing human-understandable reasons for specific model outputs, by attributing a prediction to the most influential inputs. In regression time series applications – especially for complex models such as XGBoost or neural networks – interpretability helps mitigate the "black-box" nature of these approaches and support the development of eXplainable Artificial Intelligence (XAI) *(Rojat et al., 2021)*.

In practice, interpretability and explainability seeks to identify which explanatory variables most strongly influence the predictions to characterise the direction of their effects (positive or negative), and to provide explanations of both global scale (model-wide behaviour) and local level (individual predictions) *(Murdoch et al., 2019)*.

Historically, model transparency has been framed as being in opposition to predictive accuracy, under the assumption that a simple, readable model would be less accurate than a more complex one able to capture non-linear relationships *(Murdoch et al., 2019; Rojat et al., 2021)*. However, a wide range of methods has been developed to reduce this trade-off, including within deep-learning architectures such as Transformers, where self-attention mechanisms can be used to visualise the temporal dependencies informing the predictions *(Ge et al., 2024; Z. Ma & Mei, 2025)*.

For XGBoost, two broad families of interpretability-explainability approaches are commonly distinguished:

- **Native approach**, embedded in the algorithm's structure, which provides feature importance measures derived from the contribution of predictors to the decision-tree construction;
- **Post-hoc approaches**, model-agnostic techniques applied after training to analyse model behaviour without modifying its architecture (e.g., feature perturbation, additive attribution methods).

Among post-hoc methods, SHAP (Shapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are widely used. While they rely on different principles, SHAP has gained particular traction because of its solid theoretical grounding and its ability to provide consistent, multi-scale explanations *(Lundberg et al., 2019)*.

III.4.1. XGBoost feature importance ranking

According to *T. Chen & Guestrin (2016)*, feature importance in XGBoost is computed natively during training by quantifying how each predictor contributes to decision tree construction. Each time a predictor is selected to split a node, its contribution is recorded and then aggregated across all trees in the ensemble to derive a global importance score. Three indicators are typically reported:

- **Gain:** the average reduction in the loss function (loss) brought by splits using the predictor.
- **Cover:** the average number of training samples affected by splits using the predictor
- **Weight:** the total number of times the predictor is used for splitting across all trees.

Among these, Gain is generally the most informative because it directly reflects the model improvement associated with using a given predictor (i.e., loss reduction). In practice, the gain compares the score of a parent node to the combined scores of its two child nodes, as expressed in **Equation 8**:

$$Gain_t = \frac{1}{2}\left(\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_P)^2}{H_P + \lambda}\right) - \gamma \quad (8)$$

Where:

- G_i et H_i are the sum of the gradients and Hessians for node i, with $i \in \{P, L, R\}$ (Parent, Left and Right). In regression tasks, G corresponds to the sum of residuals, and H typically scales with their count.
- λ and γ are the L2 regularisation parameter and the structural complexity penalty, respectively.

The global importance of a predictor X_j is then calculated as the sum of the gains over all splits where this predictor is used across the M trees of the forest (**Equation 9**):

$$Imp(X_j) = \sum_{m=1}^{M} \sum_{t \epsilon T_m} \delta_{j_t j} \times Gain_t \quad (9)$$

Where:

- T_m is the set of splits in tree m
- $\delta_{j_t j} = 1$ if variable X_j is used at split t, and 0 otherwise.

However, this native approach presents two main limitations. First, tree construction can introduce selection bias: predictors chosen early in the splitting process may receive inflated importance, even when their contribution to generalisation is modest *(Takefuji, 2025)*. Second, these scores provide a global, averaged view of influence and do not inform on (i) the direction of effects, (ii) local

contributions to individual predictions, or (iii) interaction effects between predictors – predictors can strongly shape the model behaviour but remain hidden in standard importance rankings.

For these reasons, native feature importance measures are best complemented with post-hoc explainability approaches that explicitly quantify predictor contributions and effects both globally and at the level of individual predictions.

III.4.2. SHAP post-hoc analysis

The SHAP method, introduced by *Lundberg & Lee (2017)*, is a post-hoc, model-agnostic explainability approach rooted in cooperative game theory. For a given prediction, it quantifies the marginal contribution of each explanatory predictor and the direction of its influence (positive or negative), while accounting for potential interactions with other variables jointly support the prediction *(Xiao et al., 2024)*.

The core principle is to express a model prediction $f(x)$ as an additive decomposition of predictor contributions (**Equation 10**):

$$f(x) = \phi_0 + \sum_{j=1}^{M} \phi_j(x) \tag{10}$$

Where:

- ϕ_0 is the baseline value, typically the mean prediction over the training dataset;
- $\phi_j(x)$ is the Shapley value associated with variable X_j, representing its marginal contribution to the prediction.
- M is the total number of variables.

In its original formulation, SHAP estimates these contributions by averaging the marginal effect of each predictor across all possible subset of predictors *(Lundberg & Lee, 2017)*. However, as shown by *Lundberg et al. (2019)*, this computation becomes computationally expensive as M increases.

For decision tree models, an efficient implementation named TreeSHAP was proposed to compute exact Shapley values in polynomial time *(Lundberg et al., 2019)*. This approach exploits how an observation x traverses the nodes of each tree to assign contribution along the decision paths. The contribution ${\phi_j}^{(t)}(x)$ of variable X_j to the prediction of the tree t is defined as (**Equation 11**):

$${\phi_j}^{(t)}(x) = \sum_{i \in leaves} p_i(x) \times \left(\sum_{S \subseteq B_i\{j\}} \frac{|S|!\,(M - |S| - 1)!}{M!} \times \nabla_j^i(S) \right) \tag{11}$$

Where:

- $p_i(x)$ is the probability that observation x falls into leaf i ;
- B_i denotes the set of variables present in the decision path leading to leaf i ;
- $\nabla_j^i(S)$ is the marginal contribution of X_j within the context of subset S.

The final Shapley value of $\phi_j(x)$ is then obtained by averaging these contributions over all trees in the ensemble.

SHAP satisfies three desirable properties *(Lundberg et al., 2019)*:

- **Local accuracy** : the sum of contributions exactly matches the model prediction;
- **Consistency** : if a variable has a greater impact on the model output, it cannot be assigned a lower importance;
- **Nullity** : variables with no predictive power are assigned zero contribution.

In addition, SHAP provides a rich set of visualisation tools (e.g., beeswarm plot, dependence plots, force plots) that makes it particularly attractive for interpreting complex models. In the landslide studies, SHAP has mainly been applied to landslide susceptibility mapping. For instance, *Xiao et al. (2024)* used SHAP analysis to interpret the most influential variables in producing susceptibility maps in southern China. Similarly, *Al-Najjar et al. (2023)* combined SHAP with a Random Forest model to interpret the contribution of 37 spatial predictors to model decisions in Bhutan. *D. Sun et al. (2024)* proposed a framework combining boosting methods with SHAP for susceptibility mapping in the Wanzou District (China). A comparable approach based on XGBoost was adopted by *Kavzoglu & Teke (2022)* in northern Turkey.

By contrast, applications to dynamic predictors derived from time series remain limited. To date, *Collini et al. (2022)* was found as the only study that combines SHAP with XGBoost to characterise the role of time-varying predictors (e.g., cumulative rainfall, event duration) in landslide occurence.

Finally, it is important to note that SHAP values are conditional on the trained model structure and may therefore inherit model-specific biases, including those related to correlated predictors or to preferential selection of variables early in the trees *(Takefuji, 2025)*.

III.5. CHAPTER III – CONCLUSION

This chapter has introduced the methodological framework required to simulate slow-moving landslide displacement time series using conventional machine-learning models. The workflow summarised in **Figure 17** highlights that predictive performance is not solely driven by the choice of algorithm, but is strongly conditioned by the quality, consistency, and representativity of the input observations.

A first set of steps concerns the preparation and exploration of the raw time series. Cleaning procedures (outlier treatment, gap handling, and noise reduction) are required to mitigate acquisition artefacts, while exploratory analyses help characterise distributions, seasonalities, trends, autocorrelation structures, and potential time lags between forcing and response. These diagnostics directly inform the feature engineering stage, in which multi-scale and lagged descriptors are derived from hydro-meteorological variables to encode system memory and delayed responses in a form that can be exploited by tree-based models.

The chapter further detailed the core modelling steps required for robust evaluation, including chronological train-test splitting to prevent data leakage, hyperparameter optimisation strategies for XGBoost, and the use of complementary error metrics to quantify predictive skill. Finally, it introduced model transparency tools, combining native feature-importance measures with post-hoc SHAP analysis to attribute predictions to the most influential predictors and to examine both global tendencies and local contributions.

Building on this framework, the following chapters apply and extend the workflow to address the research questions of this thesis. In the context of **Figure 17**, ***Chapter IV*** primarily focuses on the upstream stages of data collection and preparation (Step 1), by developing approaches to acquire, reconstruct, and extend hydrological information required for AI training, including reservoir-based modelling to infer groundwater proxies from rainfall–discharge records. ***Chapter V*** then operationalises the core modelling workflow (Steps 1-6) together with the interpretability-explainability step, by applying XGBoost to long multi-parameter time series and by making the forcing velocity relationships more explicit through model transparency analyses. ***Chapter VI*** finally extends the resulting AI-based framework towards hazard mitigation perspectives by exploring landslide velocity sensitivity under contrasting hydro-meteorological scenarios and by providing probabilistic information relevant to decision support.

CHAPTER IV

RAINFALL-DISCHARGE-BASED MODELLING TO RECONSTRUCT LANDSLIDE GROUNDWATER LEVEL PROXIES

CHAPTER IV RAINFALL-DISCHARGE-BASED MODELLING TO RECONSTRUCT LANDSLIDE GROUNDWATER LEVEL PROXIES 65

IV.1. Chapter IV – Summary 67

IV.2. Paper - Hydrogeological forecasting of deep-seated landslides dynamics: structure and sensitivity of tank models (Landslides, 2025) 70

IV.2.1. Introduction 70

IV.2.2. Study site presentation and data collection 72

IV.2.3. Modelling methods 81

IV.2.4. Results 88

IV.2.5. Discussion 96

IV.2.6. Conclusion 98

IV.3. Chapter IV - Conclusion 100

IV.1. CHAPTER IV – SUMMARY

This chapter focuses on the upstream requirement of the methodological framework introduced in ***Chapter III***, that conditions any robust application of conventional machine-learning models to landslide displacement time series: the availability of long, informative, and representative multi-source datasets. In particular, because slow-moving landslide dynamics are strongly driven by hydro-meteorological forcing through groundwater-level and pore-water pressure variations, reliable predictors must capture not only rainfall inputs but also the internal hydrological state of the slope. ***Chapter IV*** therefore directly addresses the first research question of this thesis: **How can long, informative and representative multi-source time series be acquired – or reconstructed – to support robust AI model training in slow-moving landslides?**

Continuous long-term piezometric monitoring remains difficult to achieve in slow-moving, deep-seated landslides. Instrumentation requires demanding field deployments (drilling, maintenance, access constraints), and measurements are frequently disrupted by ongoing deformation processes (e.g., borehole shearing), resulting in incomplete, discontinuous, or short-duration groundwater-level records. These limitations are critical because CML model skill is strongly conditioned by the quality, continuity, and representativeness of the input time series.

To overcome this constraint, ***Chapter IV*** develops an alternative data-reconstruction strategy based on semi-empirical conceptual tank modelling. Rather than relying on direct piezometric observations, the approach uses the rainfall–discharge relationship to reproduce the temporal dynamics of subsurface water storage and release, thereby providing consistent proxies of the internal hydrological state that can complement – or replace – missing groundwater-level measurements. The methodology is implemented with KarstMod, an open-access conceptual modelling platform initially designed for karst systems and adapted here to fractured aquifers within landslide masses. The system is represented as interconnected tanks arranged in vertical and parallel configurations, enabling the simulation of contrasted flow components and storage-driven dynamics (**Figure 19**).

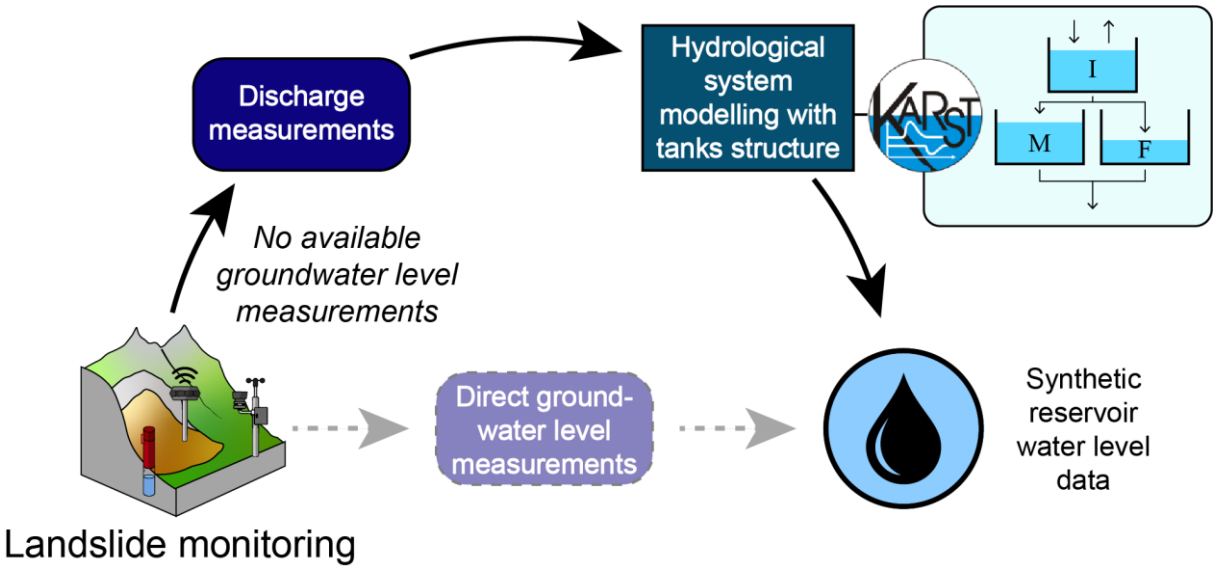


Figure 19 | Chapter IV – Schematic overview of the discharge-driven strategy used to reconstruct groundwater-state proxies.

The framework is applied to two slow-moving landslides located in the Western Alps: **Séchilienne** (French Alps) and **Mont de La Saxe** (Italian Alps), both developing within fractured crystalline contexts exhibiting low-permeability, karst-like hydrological behaviour. Based on flow duration curve analyses and leveraging KarstMod modularity, site-specific conceptual structures are defined, including:

- **The Séchilienne model:** composed of two tanks connected to the outlet by both a slow storage component and a direct flow path, reproducing a **bimodal** discharge regime that reflects the combined contribution of microfissured seepage and more conductive fracture networks.
- **The Mont de La Saxe model:** integrating a snowmelt routine and consisting of two tanks coupled with a feeding flow component, reproducing a **unimodal** regime mainly controlled by episodic inflows through preferential drainage paths associated with morphostructural deformation features.

Model calibration and validation are performed on multi-year daily discharge records using a composite objective function, coupling the Balance Error (BE) and the Kling-Gupta Efficiency (KGE) metrics.

The results show that the calibrated models reproduce observed discharge fluctuations under contrasted internal and external conditions. Importantly, the simulations provide access to reservoir water-level time series that can be used as synthetic yet physically consistent indicators of water-storage variability within the landslide system. Because the models are driven by meteorological inputs (rainfall

and temperature), they also offer the capacity to extend hydrological reconstructions beyond the monitoring period, thereby filling gaps and extending incomplete records. In the perspective of the thesis modelling framework introduced in **Figure 17** (see section ***III.1.***), ***Chapter IV*** **thus contributes to Step 1 (*time series collection and preparation*) by supplying reconstructed groundwater proxies that can be integrated into the subsequent feature-engineering and AI modelling stages.**

The following section (section IV.2.) consists on a manuscript published in *Landslides*:

O. Béjean-Maillard, C. Bertrand, J.-P. Malet, G. Cinkus, P. Nevers, J. Ducasse, D. Bertolo, and P. Thuegaz. *Hydrogeological forecasting of deep-seated landslides dynamics: structure and sensitivity of tank models*. Landslides 22, 2199–2217 (2025). https://doi.org/10.1007/s10346-025-02482-2

IV.2. PAPER - HYDROGEOLOGICAL FORECASTING OF DEEP-SEATED LANDSLIDES DYNAMICS: STRUCTURE AND SENSITIVITY OF TANK MODELS (LANDSLIDES, 2025)

IV.2.1. Introduction

Deep-seated landslides predominantly occur in high mountain ranges characterized by steep slopes *(Crosta et al., 2013; Palis et al., 2017; Vallet et al., 2016)*. In contrast to shallow landslides, these events are marked by extended periods of slow motion, interspersed with phases of acceleration. However, these acceleration phases typically stabilize without leading to catastrophic failures, though they often result in localized rockfalls in the most unstable areas *(Bonzanigo et al., 2007; Borrelli & Gullà, 2017; Zangerl et al., 2010)*.

Seasonal landslide acceleration phases are frequently initiated by various external factors, with hydro-meteorological events, such as intense rainfall and substantial snowmelt, serving as primary triggers *(Crosta et al., 2013)*. These events lead to fluctuations in pore fluid pressures and groundwater level, as well as a reduction in suction along the slip surface, which control landslide movements *(Bogaard & Greco, 2016; Matsuura et al., 2008; Preisig et al., 2016; Prokešová et al., 2013; Rosone et al., 2018)*. *H.-Y. Sun et al. (2019)* have highlighted the strong correlation between landslide displacement and hydrogeological parameters, demonstrating how changes in groundwater levels directly influence slope stability. *Preisig et al. (2016)* further showed that variations in pore fluid pressures within fractures contribute to the progressive degradation of slopes, facilitating regular reactivation events with minor pore pressure variations. Understanding these dynamics is crucial, particularly in deep-seated landslides, where acquiring accurate groundwater fluctuations data becomes a key factor in evaluating displacement rates. Approaches to modelling these processes can generally be divided into two categories: physical-based and empirical models. Physical-based models typically utilize seepage finite-element methods to simulate the hydrogeological behaviour of individual system units, relying on well-established flow laws such as Darcy's equations and Richards' equations, alongside models of unsaturated soil hydraulic characteristics like the Van Genuchten-Mualem model and the Brooks and Correy model. These models are grounded in robust physical principles and offer detailed representations of the processes occurring within the system. However, they require extensive input data (hydraulic conductivity, matrix suction, volumetric water content, fractures volume and aperture), as well as precise definition of the boundary conditions and adjustments of the 2D/3D meshes to the studied system *(Conte et al., 2014; Hung et al., 2018; Malet et al., 2005)*. The complexity and scale of deep-seated landslides often make it challenging to obtain the necessary input parameters, limiting the applicability of these models in such environments.

In contrast, empirical models rely on field observations and permanent instrumental data to establish statistical relationships between input variables (precipitation, snowmelt, evapotranspiration) and output variables (groundwater level, pore fluid pressure, discharge). This approach is often preferred in cases where the system's physical parameters are incomplete or difficult to obtain *(Hong & Wan, 2011a)* but requires long-period, high frequency, and accurate field monitoring data to calibrate the models. Among the empirical approaches, tank models are widely used for modelling groundwater levels or pore pressures in landslides *(Angeli et al., 1998; Buma, 2000; G.-W. Lin et al., 2020; Malet et al., 2005; Matsuura et al., 2008)*. These models conceptualize the system as a series of reservoirs that simulate groundwater infiltration, storage, and drainage processes. *G.-W. Lin et al. (2020)* employed a three-layered reservoir model to estimate the soil water index (SWI) for multiple rainfall-induced landslides events in Taiwan, identifying a critical SWI threshold for landslide triggering. *Weerasinghe et al. (2019)* proposed a multi-reservoir model to simulate the groundwater level variations observed in the Nawalpitiya landslide (Sri Lanka), demonstrating a positive correlation with landslide movements. More recently, *Zheng et al. (2024)* introduced a modified tank model with three superimposed trapezoidal-shaped reservoirs to account for non-linear water-level changes caused by intense precipitation events. However, most of these models use continuous groundwater level or pore fluid pressure measurements acquired from monitored boreholes, which may not always exist for the study sites.

An alternative approach involves utilizing discharge data from springs within the unstable zone and applying tank models based on rainfall-discharge relationship. This method is commonly employed in highly fractured media such as karstic systems *(Baudement et al., 2017; Cinkus et al., 2023; Fleury et al., 2007; Geyer et al., 2008; Jukić & Denić-Jukić, 2009)*, which offer several potential reservoir configurations and modelling software to simulate discharge at outlets. These discharge-based models provide insight into water level variations within the reservoirs that compose the model and help to understand fluctuations in water quantities within the studied system. *Baudement et al. (2017)* used a three-reservoir model to characterize the hydrological behaviour of the Dardennes spring (France), showing that simulated water levels in the conduit reservoir closely matched observed flood period measurements.

The objective of this work is to propose models capable of reproducing discharge series at the outlets of landslide fractured media using the KarstMod (a rainfall-discharge modelling software developed by SNO Karst) open-source library *(Mazzilli et al., 2019)*. The approach is implemented and tested at two deep-seated landslides: the Séchilienne landslide (French Alps) and the Mont de la Saxe landslide (Italian Alps). Modelling the discharge series allows for extracting qualitative information on water level fluctuations within these two systems and extend the simulation beyond the studied period to complement information on past periods before the setup of monitoring systems. First, we introduce the

hydrogeological context and data available for the Séchilienne and Mont de la Saxe models. Second, we present the architecture of the proposed tank models, the simulation strategies, and the performance criteria used to measure the simulation's accuracy. Third, we interpret the simulations' results and the internal series (reservoir water levels, internal discharge) obtained and discuss the sensitivity of the model parameters in the simulation. Finally, we discuss the use of these models to acquire hydrological information, first on the studied period and then beyond the available data period, as well as the relevance of the data to the landslides' hydrological understanding of water flows in fractured media.

IV.2.2. Study site presentation and data collection

IV.2.2.1. The Séchilienne landslide

IV.2.2.1.1. Geographical context and hydrological processes

The Séchilienne landslide is located on the right bank of the Romanche River, on the southwestern slope of the Mont Sec Massif, approximately 20km southeast of Grenoble (Isère, France) (**Figure 20**). The Romanche Valley exhibits geomorphological features shaped by recent glacial activity, with slopes reaching up to 45° in inclination *(Delunel et al., 2010; Le Roux et al., 2011; Schwartz et al., 2017)*. The Mont Sec massif, positioned at the southwestern edge of the external crystalline massif of Belledonne, is primarily composed of micaschist, with an overlay of sedimentary rocks dating from the Carboniferous to the Lias period, and Quaternary deposits from the Würm glaciation *(Vallet, 2014; Vengeon, 1998)*.

The landslide spans an area of approximately 1 km^2 and ranges in altitude from 450 m to 1100 m asl, encompassing a total volume of about 60 million m^3 (**Table 5**). The plateau above 1100 m asl, formed by glacial action, is primarily composed of Lias sedimentary deposits covered by Quaternary materials. The slope instability is marked by a well-developed fracture network consisting of two sets of nearly vertical fractures filled with detrital material, oriented N110-N120 and N70. Additionally, two major N20-oriented faults, the Séchilienne and Sabot faults, cross the Mont Sec Massif to the east of the instability *(Le Roux et al., 2011)* (**Figure 21a**). The significant fracturing and anisotropy of the massif result in two distinct groundwater flow dynamics within the unstable zone: a fast, responsive flow within the open fractures, and a slower, more inertial flow within the micro-fissured rock matrix of the micaschists *(Nevers, 2021; Vallet et al., 2015)*. This dual-permeability behaviour is typical in fractured

rock systems *(Cappa et al., 2004; Maréchal, 1998; Vallet et al., 2015)* and extends to depths of 150 to 200 meters, reaching the theoretical basal sliding surface of the landslide *(Le Roux et al., 2011).*

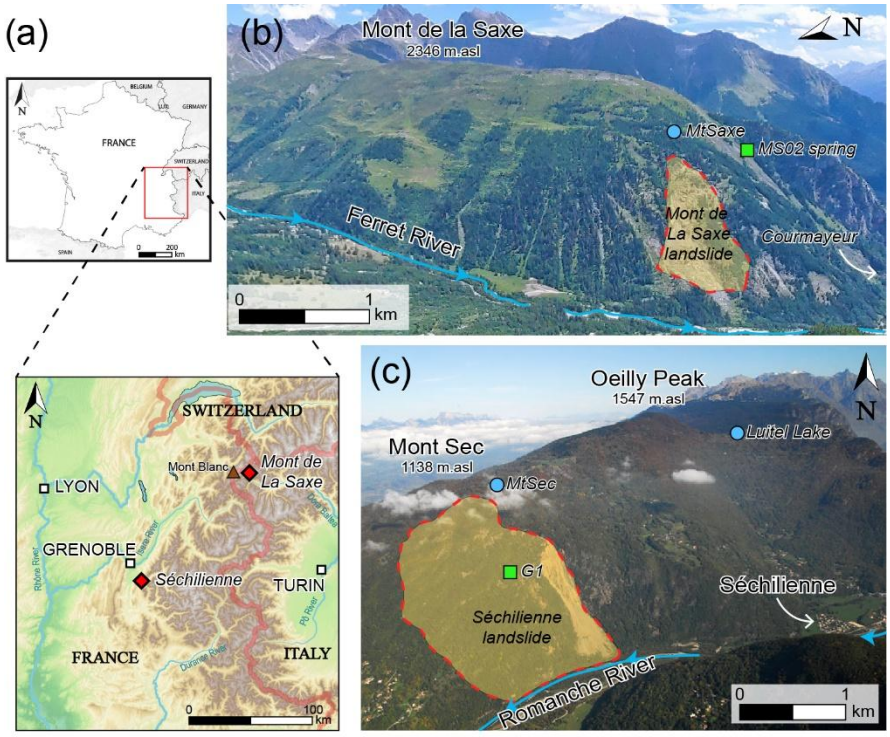


Figure 20 | (a) Localisation of the Séchilienne and Mont de La Saxe landslides. **(b)** The Mont de La Saxe and **(c)** the Séchilienne landslides are represented by a yellow surface, along with meteorological stations (blue circles) and discharge stations (green squares) localisations.

The hydraulic conductivity gradient between the highly fractured rocks of the unstable slope and the non-fractured rocks beneath creates a dual aquifer system, with a perched aquifer developing within the landslide body *(Guglielmi et al., 2002; Meric et al., 2005; Vengeon, 1998)* (**Figure 21b**). Due to the heterogeneous and anisotropic properties of the landslide media, the landslide perched aquifer is better described by numerous disconnected saturated zones rather than a single saturated entity. The extent and dynamics of the perched aquifer fluctuate with seasonal high- and low-flow periods *(Binet et al. 2007; Cappa et al., 2004; Guglielmi et al., 2002; Vallet et al., 2015)*. During high-flow periods, water inflow through vertical fractures maintains a basal continuous perched aquifer. When the water inflow is sufficient, the numerous disconnected saturated zones may temporarily connect, resulting in horizontal groundwater flows parallel to the slope down to the landslide toe. In contrast, during low-flow periods, the water inflow is insufficient to sustain continuous perched aquifer, resulting in disconnected saturated zones localized within the micro-fissured micaschists slowly draining towards the landslide's basal surface *(Vallet et al., 2015)* (**Figure 21b**).

Table 5 | The Séchilienne and Mont de La Saxe landslides geographical and geological features.

	Séchilienne landslide	**Mont de La Saxe landslide**
Geographical coordinates	45.06, 5.80	45.82, 6.97
Area (km^2)	1.0	0.15
Volume (M m^3)	60	8
Altitudes (m a.s.l.)	450 to 1100	1400 to 1870
Lithology	Micaschist bedrock with overlied morainic deposits	Micaschist and calcschist bedrock with overlied colluvial and glacial deposits

The landslide recharge area varies depending on the aquifer's saturation state. *Vallet et al. (2015)* demonstrated, using δO18 analyses, that the recharge area extends beyond the landslide's surface during high-flow periods, with water inputs derived from the sedimentary cover of the Mont Sec plateau, usually drained by the Sabot and Séchilienne faults. However, *Cappa et al. (2004)* and *Vallet et al. (2015)* conclude that the extensive fracturing within the unstable zone significantly reduces surface runoff, leading to a highly localized infiltrating area. This hypothesis is supported by *Nevers (2021)*, who estimated the recharge area of the G1 outlet within the monitoring gallery using the methodology developed by *Bonacci & Andrić (2015)*, determining that the recharge area ranges between 9 and 204 m^2.

IV.2.2.1.2. Field data collection and processing

The climate at Séchilienne is mountainous, with an average annual precipitation of 1238 mm. While there are no significant seasonal variations in precipitation throughout the year, notable increases in rainfall during May and from September to December are observed. Extreme rainfall events, with values ranging between 30 to 80 mm/day, primarily occur during summer. The Mont Sec Massif experiences moderate snowfall, with an average accumulation of 132 mm of water equivalent per winter. Potential evapotranspiration is calculated using the method proposed by *Oudin et al. (2005)*, yielding an annual average of 418 mm. Air temperatures exhibit a clear seasonal variation with an annual average of 6.80°C.

Meteorological data are collected from two weather stations. One station is situated at the Mont Sec summit, at 1120 m asl, while the other is located at the Luitel Lake, 4 km northeast of the landslide, at 1261 m asl. The Luitel station is used as a backup for missing data from the Mont Sec station. Both stations are equipped with snow gauges to estimate snowmelt in terms of water height (mm). The precipitation data used in the study encompass both rainfall and the water equivalent of snowfall, collected at hourly intervals and then summed up into daily data. The meteorological dataset covers 30 years, from 1992 to 2022 (**Table 6**).

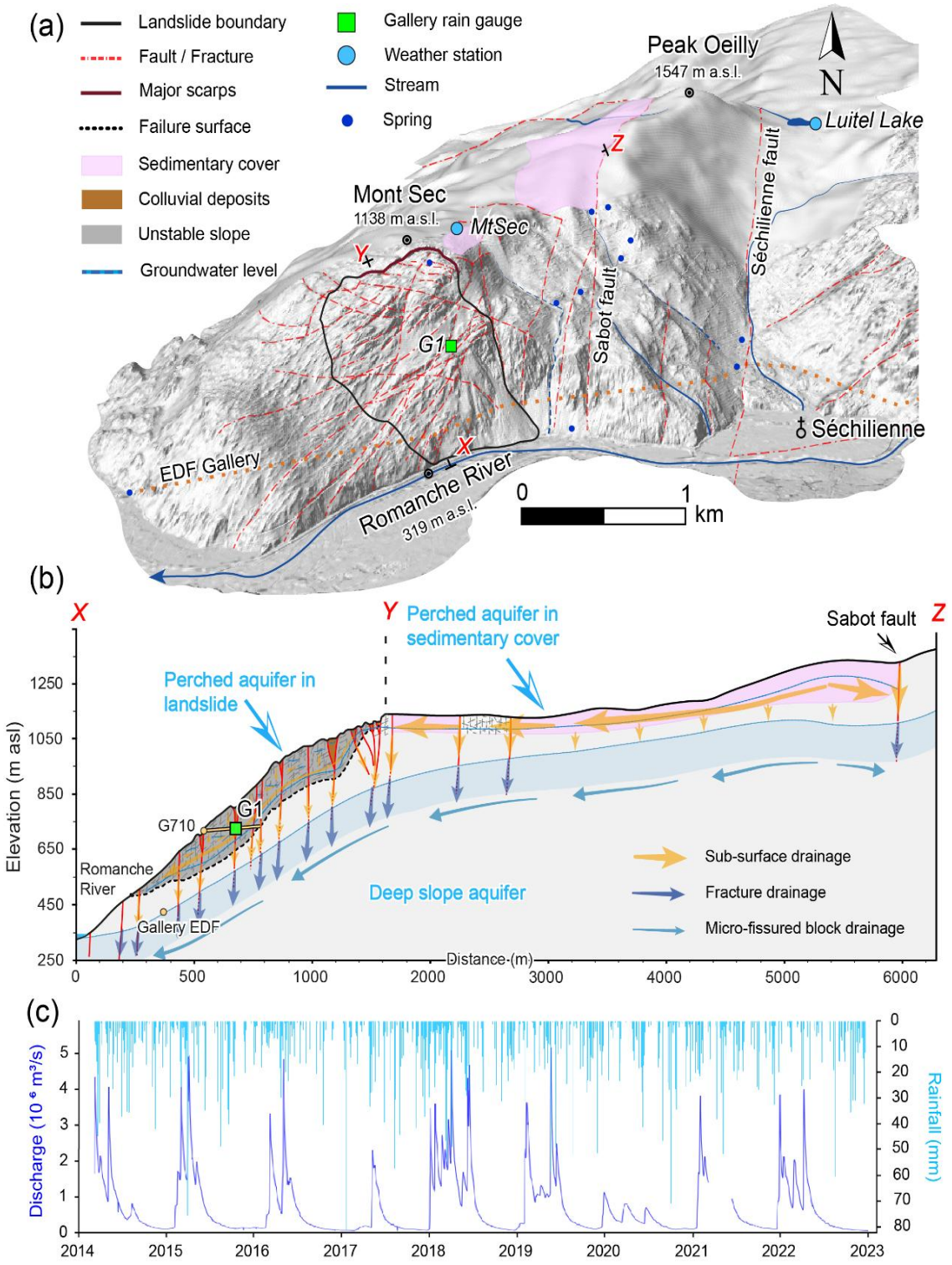


Figure 21 | (a) Shaded 1-meter DEM of the right flank of the Romanche Valley with localisation of the landslides, fractures, streams, springs and measurements stations. **(b)** Schematic cross-section of the hydrogeological functioning of the Séchilienne landslide (adapted from *Vallet et al., 2014*). **(c)** Daily rainfall and discharge time series acquired at the Mont Sec weather station and the G1 outlet.

Discharge data are collected in the G710 observation gallery, located at 710 m asl, along a fracture within the unstable slope, approximately 140 m from the gallery entrance (**Figure 21b**). Runoff waters from the fractured gallery roof are collected using a tarpaulin to maximize water collection, which is then channelled to the G1 measuring station. The measurements are acquired hourly using a water recovery system to capture the fractured and fissured rock flow. The system employs a Rain Gauge RM Young 52203 (10.13 mL/tip) paired with a CR200(X) data logger from Campbell Scientific, which records water volume in $mL.h^{-1}$. These raw measurements are subsequently converted into $m^3.s^{-1}$ and averaged daily. The discharge dataset spans eight years, from 2014 to 2022, with a 3-month interruption from March to June 2021 (**Table 6**). The recorded discharge rates are characterized by relatively low magnitudes, with occasional flood peak values. The variations in discharge follow a seasonal pattern, with high-flow periods observed between February and June, often showing one or multiple discharge peak events. During low-flow periods, a continuous base flow is maintained, sustained by the drainage of micro-fissured rock, which supports the G1 discharge (**Figure 21c**).

Table 6 | Summary of the Séchilienne landslide hydrometeorological dataset.

Time series	Studied period	Data collection frequency	Unit	Complete time series (%)	Geographical coordinates
Rainfall	2014/03/07 - 2024/12/31	Hourly	mm	100	Mont Sec station : 45.07, 5.81 Luitel station: 45.09, 5.85
Temperature		Hourly	°C	100	
Potential EvapoTranspiration (PET)		-	mm	100	-
Discharge		Hourly	$m^3.s^{-1}$	97	45.06, 5.80
Eelectrical conductivity		Half-hourly	$ms.cm^{-1}$	85	

IV.2.2.2. The Mont de La Saxe landslide

IV.2.2.2.1. Geographical context and hydrological process

The Mont de La Saxe landslide is situated in the northern part of the Aosta Valley, at the southwestern edge of the Mont de La Saxe Massif, on the left bank of the Ferret Valley, approximately 3 km north of Courmayeur (**Figure 20**). The Ferret Valley exhibits a landscape shaped by past glacial episodes, with steep slopes averaging an inclination of 37° *(Alberti, 2018)*. The valley is characterized by prominent NW-SE morphostructures, including faults, scarps, and counterscarps, inherited from the region's tectonic history, and features numerous deep-seated landslides, including the Mont de La Saxe landslide.

The Mont de La Saxe Massif is predominantly composed of micaschist and calcschist, part of the Ultra-Helvetic Basal Decollement Units, which are discordantly overlain by colluvium and glacial deposits (**Figures 22a-b**) *(Alberti, 2018; Leloup et al., 2005)*. The landslide covers an area of approximately 150,000 m², with a maximum depth of 420 m, extending from 1400 to 1870 m asl, and represents a volume of about 8 million m³ (**Table 5**) *(Alberti, 2018; Crosta et al., 2014)*. Two failure surfaces, referred as the rockslide failure surface and the deep-seated gravitational deformation failure surface, are identified within the slope. The rockslide failure surface is located at a depth of 30 to 90 meters and has been identified through geophysical investigation campaigns using seismic refraction tomography coupled with inclinometric measurements from several boreholes. The deep-seated gravitational deformation failure surface is not directly observed through field measurements, however, by modelling groundwater flow using a 2D transient groundwater model within the landslide, *Crosta et al. (2014)* supply the existence of a second deeper brecciated failure surface for assessing the observed oscillations of groundwater level.

Numerous surface springs are located at the base and along the southern edge of the landslide (**Figure 22a**). *Crosta et al. (2014)* identified a landslide recharge area of approximately 0.54 km², including surface water flows along the morphostructures during intense rainfall and snowmelt periods. However, *Alberti (2018)* proposes a larger recharge area of 4 km², accounting for the discharge rates observed at the springs within the massif at the landslide site. Tracer campaigns carried out by the Valle d'Aosta region's geological monitoring service corroborate these findings, indicating a more distant origin of meltwater, specifically from the Arminaz and Malatra torrents, located 4 and 6 km northwest of the Ferret Valley and aligned with the regional morphostructures.

IV.2.2.2.2. Field data collection and processing

The climate of Mont de La Saxe is typical of mountainous regions, with an average annual precipitation of 1470 mm. A pronounced seasonality is observed with peak rainfall occurring in May, October and November. The summer months are characterized by intense rainfall events, with daily precipitation exceeding 150 mm. Winters bring substantial snowfall, with an average water equivalent of approximately 810 mm and snow depths reaching up to 2.5 m. Snow cover persists for about six months, from November to April, resulting in considerable snowmelt during the spring. The average potential evapotranspiration is 370 mm per year, while temperatures follow a seasonal pattern, with an annual average of 6.18°C *(Alberti, 2018)*.

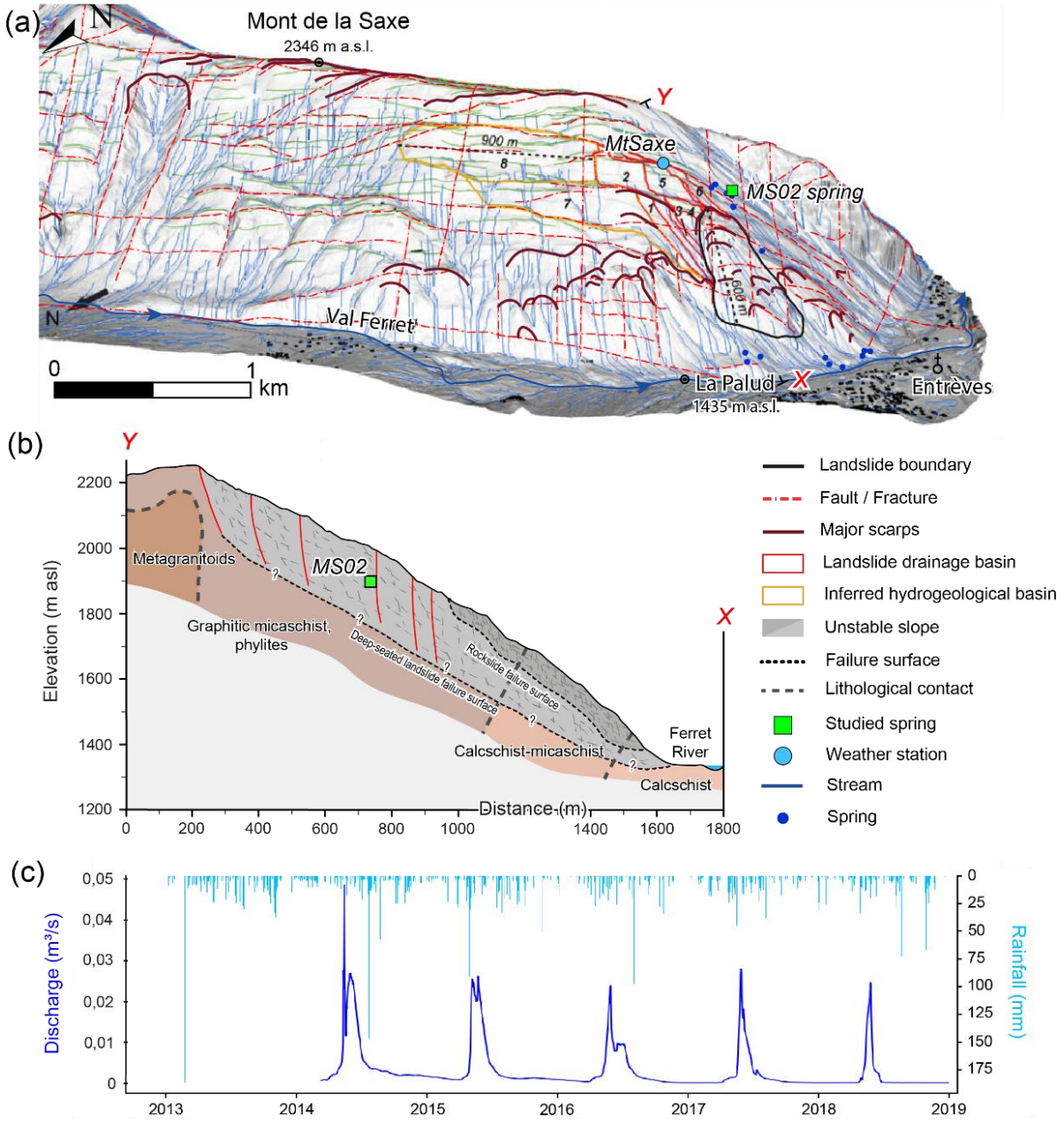


Figure 22 | (a) Shaded 1-meter DEM of the left flank of the Ferret Valley with localisation of the landslides, fractures, streams, springs and measurements stations (adapted from Crosta et al., 2013). **(b)** Schematic cross-section of the Mont de La Saxe landslide (adapted from *Crosta et al., 2013*). **(c)** Daily rainfall and discharge time series acquired at the Mont de La Saxe weather station and the MS02 spring.

Daily rainfall and air temperature data were collected over 10 years, from 2013 to 2023, at the Mont de La Saxe weather station, located at 2076 m asl (**Table 7**). Discharge data for the same period were obtained from the MS02 spring, situated on the southern boundary of the landslide at an altitude of 1837 m asl. These measurements were recorded hourly as water height and converted into daily average discharge ($m^3.s^{-1}$) (**Table 7).** The discharge data reveal notable seasonal variations, with high-flow periods extending from late March to early August, primarily driven by snowmelt, during which

discharge peaks range from 0.025 to 0.045 $m^3.s^{-1}$. After these peak periods, discharge rates decline sharply to base flow levels, often resulting in the drying up of the spring for the remainder of the year (**Figure 22c**). Due to numerous gaps in rainfall and discharge data from 2019 to 2023, the analysis period extends until 2018/12/01.

Table 7 | Summary of the Mont de La Saxe hydrometeorological dataset.

Time series	Studied period	Data collection frequency	Unit	Complete time series (%)	Geographical coordinates
Rainfall	2013/01/01 - 2018/12/01	Hourly	mm	100	45.82, 6.97
Temperature		Hourly	°C	100	
Discharge		Hourly	$m^3.s^{-1}$	79	45.81, 6.97

IV.2.2.3. Outlet's hydrogeological behaviour

The flow duration curve method, proposed by *Mangin (1975)*, is a technique aimed at identifying specific hydrogeological events (e.g., leakage, overflow, storage, water input). The method is based on calculating ranked flow curves where discharges are categorized into classes to highlight changes in the hydrographs. Anomalies are identified and interpreted using the classification proposed by *Marsaud (1996)* based on the slope breaks' angle and position (**Figure 23**).

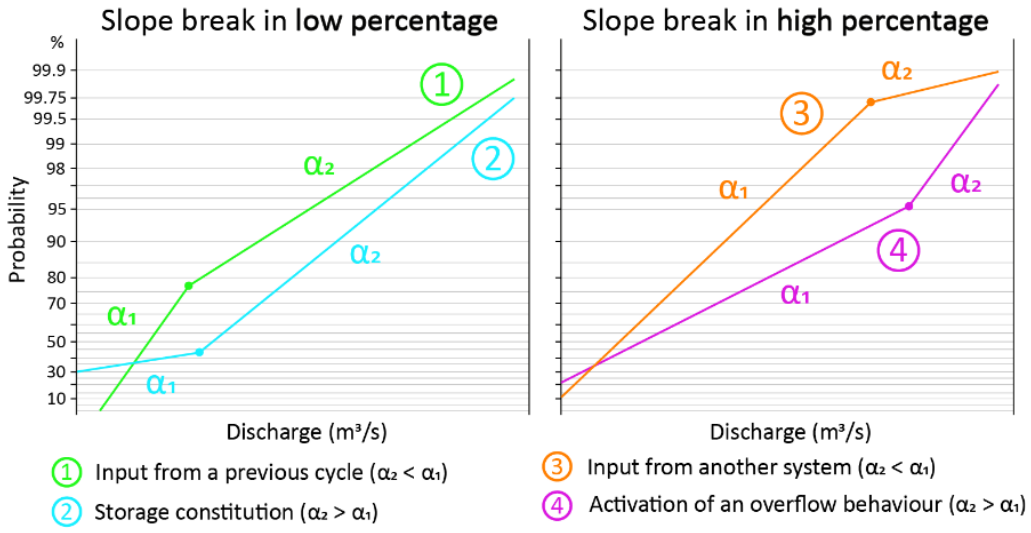


Figure 23 | Several cases of flow duration curves and their hydrodynamic interpretation (adapted from *Marsaud, 1996*).

The flow duration curve applied to the observed discharges of the Séchilienne landslide indicates a three-phase behaviour (**Figure 24a**):

- An initial phase, where the outlet is fed by a previous hydrological cycle ($\alpha_2 < \alpha_1$ at low percentages);
- A storage constitution phase ($\alpha_3 > \alpha_4$ in low percentages);
- An overflow phase ($\alpha_5 > \alpha_4$ in high percentages).

This three-phase behaviour depict a storage capacity within the system with long-term water release illustrated by the observed baseflow during low-flow periods. Furthermore, it highlights an overflow dynamic when the system reaches saturation, resulting in significant water contributions to the outlet illustrated with discharge peaks.

In contrast, the flow duration curve from the observed discharges of the Mont de La Saxe landslide exhibits a two-phase behaviour (**Figure 24b**):

- A storage constitution phase ($\alpha_2 > \alpha_1$ in low percentages);
- An overflow phase ($\alpha_3 > \alpha_2$ in high percentages).

It displays a lack of sufficient storage to sustain a long-term baseflow, as evidenced by observed spring drying periods. Further, the spring's discharge is driven by a gradual groundwater accumulation or, at least, the sudden arrival of a large amount of water from spring snowmelt, which is partly drained during the overflow phases.

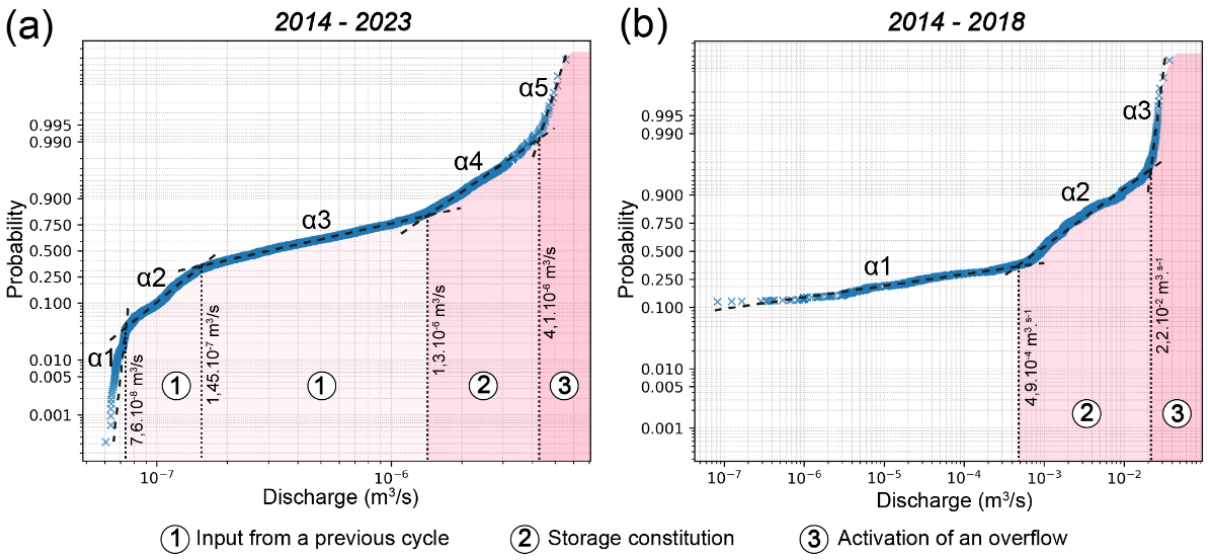


Figure 24 | Flow duration curve based for the **(a)** Séchilienne and **(b)** Mont de La Saxe landslides. Slopes are represented with alpha, slope breaks are shown with dotted lines and associated discharge values, and numbers indicate different behavioural phases.

IV.2.3. Modelling methods

IV.2.3.1. KarstMod structure and implementation

KarstMod is an open-source modelling library developed within the framework of the French National Karst Observatory (SNO) initiative *(Mazzilli et al., 2019)*, designed to provide a modular interface for simulating rainfall-discharge interactions in karst systems. The software has been tested and validated across various karst sites, such as the Fontaine de Vaucluse (France) and the Source du Lez (France), which exhibit diverse geological characteristics and long-term field data records. This allows for consideration of the heterogeneity and temporal variability in their responses to precipitation events. KarstMod represents karst systems through multiple parallel flow paths, which may include reservoirs, accounting for the dual porosity commonly found in karst media (slow matrix flow and rapid fracture flow) and the water exchanges between different hydrological compartments. This approach fully adapts to less developed fractured media, such as those observed in landslide systems.

The software offers a comprehensive range of features, including various options for connecting tanks, flexibility in selecting study periods, multiple performance criteria, and integrated modules for calculating potential evapotranspiration (PET routine) and snowmelt (Snow routine). The KarstMod user guide *(Mazzilli et al., 2023)* provides detailed descriptions of these functions.

The software is structured on two levels, supporting up to four tanks. The infiltration tank (I) simulates the infiltration zone or soil layers, while a maximum of three lower tanks (M, F, and L) represent the potential storage of the system's slow and fast flow components. The different tanks and flow paths can be activated in the user interface. Each tank is connected through one or more flow paths, which can be linear, threshold-based, or exhibit hysteretic behaviour. Changes in the water levels within the tanks are governed by a balance equation based on the inflows and outflows of water. These flows are quantified by the system's recession coefficients k, expressed in $mm.day^{-1}$. The flows are denoted as Q, with nomenclature representing flow from one reservoir to another (e.g., Q_{AB} for the flow from reservoir A to reservoir B). The outlet flow, Q_O, is the sum of the contributing flows to the system's outlet, scaled by the catchment area RA, which is estimated by the software.

The discharge simulation process divides the time series into three periods: warm-up, calibration, and validation. These periods are essential for initializing the model, adjusting parameters, and testing the calibrated parameter set over the validation period. The simulation performance is evaluated using an objective function (W_{OBJ}), which can either be a single performance criterion, such as Nash-Sutcliffe Efficiency (NSE), Kling-Gupta Efficiency (KGE), or Non-Parametric Kling-Gupta Efficiency (KGENP), or a weighted combination of multiple criteria. The model retains simulations that exceed a user-defined threshold for the objective function. KarstMod also provides analytical tools for

interpreting model results and assessing parameter performance, including Diagnostic Efficiency plots, performance radar plots, KGE decomposition radar plots, and Sobol sensitivity indices.

The Diagnostic Efficiency method (DE), proposed by *Schwemmle et al. (2021)*, offers a hydrogeological interpretation of model errors by assessing the combination of three types of errors: constant error (B_{REL}), dynamic error (B_{SLOPE}), and timing error (r) (**Equation 12**). The constant and dynamic error are derived from the flow duration curve of the simulated and observed series, representing, respectively, the average deviation and the integrated relative bias between simulated and observed discharge over the entire flow duration curve domain. Timing error evaluates the temporal agreement between simulated and observed discharge time series. Constant error informs on the extent to which simulated flows are consistently underestimated or overestimated. A negative value indicates a persistent underestimation, and inversely. Dynamic error determines whether the model underestimates low flows and overestimates high flows and vice versa. A negative value indicates a tendency to underestimate low flows and overestimate high flows, whereas a positive value indicates the opposite behaviour. Timing error measures the synchronization of variations between the simulated and observed discharge time series. A value close to 1 signifies near-perfect alignment of the fluctuations of both series. The final output of DE ranges from 0 to +∞, where a value close to 0 indicates a minimal error. *Schwemmle et al. (2021)* depict the contributions of these errors using a diagnostic polar plot where axis represent constant and dynamic error. The contribution of low flows (0th-50th percentile) and high flows (50th-100th percentile), defined as ε_{LF} and ε_{HF}, are displayed by pies. Timing error is indicated with a colour code on the pies.

$$DE = \sqrt{\overline{B_{REL}}^2 + |B_{SLOPE}|^2 + (r-1)^2} \quad (12)$$

The Sobol sensitivity indices *(Saltelli, 2002)* decompose the variance of the model's output into contributions from individual input variables (first-order index) and their interactions (total effect index), allowing for the distinction between the exclusive and total influence of each parameter. By default, KarstMod calculates sensitivity indices based on a number of simulations of $Z = 1000 \times (n + 2)$, where *n* represents the number of model parameters to be calibrated *(Mazzilli et al., 2019)*.

IV.2.3.2. Models calibration and validation

IV.2.3.2.1. Models structure development

Assuming that the two fractured environments behave similarly to karst systems, an initial model structure was proposed for both sites. This configuration consists of an infiltration tank (I) and two flow paths arranged in parallel with lower tanks (M and F), representing the dual-permeability behaviour, with one flow path simulating slow matrix flow with a storage capacity (reservoir M) and another capturing rapid fracture flow with a storage capacity (reservoir F). Unlike typical landslide tank models, which often require multiple superimposed tanks to represent different lithologies, the karst system models usually focus on a single lithology, such as limestone. In the case of the Séchilienne and Mont de La Saxe landslides, where the lithology is homogeneous (micaschist), using multiple stacked reservoirs to represent different layers is unnecessary. Instead, parallel reservoirs are employed to capture the heterogeneity of the flow behaviour within micaschist, with separate tanks representing the slow matrix flow and the rapid fracture flow. Additionally, considering the effect of snow accumulation in the mountainous terrains of Séchilienne and Mont de La Saxe, a snow routine is integrated to redistribute precipitation inputs.

Preliminary simulations were conducted using the initial model structure, and modifications were made iteratively based on the outlet's operating phases identified with the flow duration curves analysis to achieve models that accurately simulate discharge time series. The models were then simplified and optimized to reduce the number of parameters while maintaining performance and interpretability. The Séchilienne model (**Figure 25a**) consists of an infiltration reservoir (I) connected to:

- a reservoir M that simulates slow matrix flow and storage processes before feeding the outlet;
- a direct flow path to the outlet, without storage capacity, characterized by a dual-threshold hysteresis function to represent fracture flows. This flow is activated once the second threshold ($T2_{hy}$) is reached and remains active as long as the water level in the reservoir I does not drop below the first threshold ($T1_{hy}$). The dual-threshold function represents the outlet overflow dynamics.

Although a snow routine was initially included, it did not significantly redistribute precipitation or affect the simulation, so it was excluded from the final model. The Séchilienne model has seven parameters, all related to the model's structure (**Table 8**).

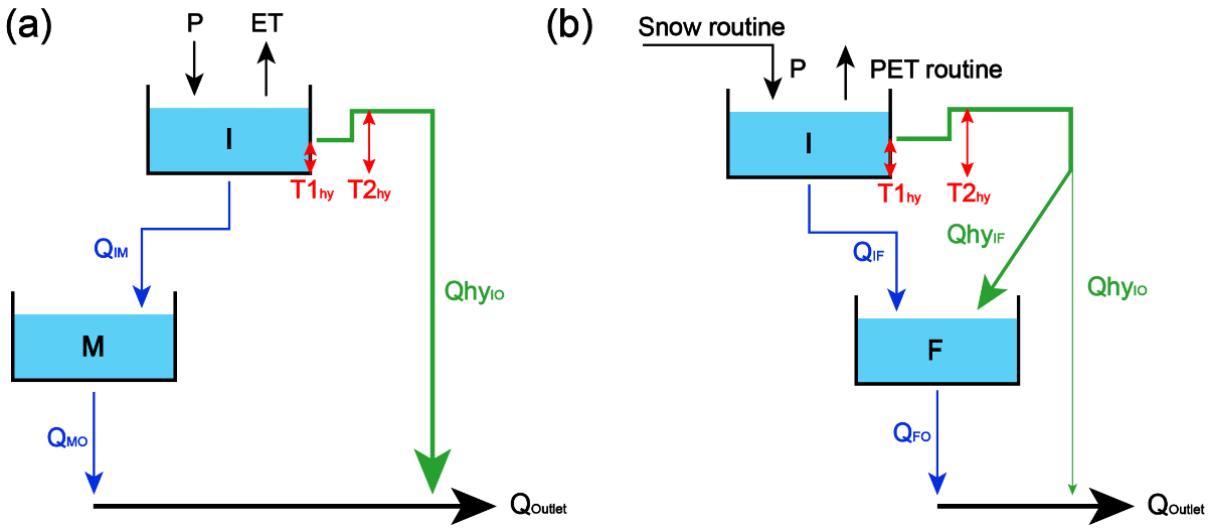


Figure 25 | Tank structures for **(a)** the Séchilienne model and **(b)** the Mont de La Saxe model. Reservoirs are named according to their mechanisms: Infiltration (I), Matrix (M) and Fracture (F) reservoirs. The flux nomenclature is defined by the two reservoirs they connect, and the parameters $T1_{hy}$ and $T2_{hy}$ represent the two thresholds (T) of the hysteresis function.

The Mont de La Saxe model displays an unimodal flow controlled by the slope morphostructure without sustaining a baseflow. Thus, a matrix flow path with a reservoir M is unnecessary. However, to slightly buffer the direct effect of input water on the outlet, a reservoir (Reservoir F) is integrated into the flow path. The unimodal flow is then described (**Figure 25b**):

- through a direct flow from reservoir I to reservoir F;
- through a dual-threshold flow ($T1_{hy}$ and $T2_{hy}$) that also captures the hysteretic behaviour of rapid flows. This flow primarily feeds reservoir F, with a smaller proportion directly contributing to the outlet. The distribution factor x_{hy} governs the flow partitioning between reservoir F and the outlet.

Due to the substantial snow accumulation and snowmelt dynamics on the Mont de La Saxe massif, the snow and PET routine modules were retained in the model to account for precipitation redistribution properly. The Mont de La Saxe model includes 16 parameters, with 10 related to the model structure, 4 associated with the snow routine, and 2 linked to the PET routine (**Table 8**).

Table 8 | Summary of the parameters for the Séchilienne and Mont de La Saxe models, including the defined ranges for simulation and the optimal values retained. The order of parameters presents the structural parameters (from I_0 to x_{hy}), the PET routine parameters (K_1 and K_2) and the Snow routine parameters (from T_S to CWH).

Parameter	Description	Séchilienne model		Mont de La Saxe model		Unit
		Range	Best value	Range	Best value	
I_0	Initial water level in tank I	220 to 300	284.95	-	-	mm
$T1_{hy}$	Desactivation threshold of the hysteresis function	0 to 50	6.357	200 to 250	241.457	mm
$T2_{hy}$	Activation threshold of the hysteresis function	200 to 270	259.027	270 to 300	283.387	mm
k_{IF}	Recession coefficient from the tank I to the tank F	-	-	$4.0\ 10^{-4}$ to $1.0\ 10^{-3}$	$8.822\ 10^{-4}$	$mm.day^{-1}$
k_{FO}	Recession coefficient from the tank F to the outlet	-	-	$1.0\ 10^{-2}$ to 1.0	$7.1\ 10^{-2}$	$mm.day^{-1}$
k_{IM}	Recession coefficient from the tank I to the tank M	$1.0\ 10^{-4}$ to $3.0\ 10^{-2}$	$7.0\ 10^{-3}$	-	-	$mm.day^{-1}$
k_{MO}	Recession coefficient from the tank M to the outlet	$1.0\ 10^{-3}$ to $1.0\ 10^{-1}$	$1.5\ 10^{-2}$	-	-	$mm.day^{-1}$
k_{hy}	Recession coefficient of the hysteresis flow	$2.0\ 10^{-2}$ to $9.0\ 10^{-2}$	$4.5\ 10^{-2}$	$3.0\ 10^{-2}$ to 1.0	$6.3\ 10^{-2}$	$mm.day^{-1}$
x_{hy}	Distribution coefficient of the hysteresis flow	-	-	$5.0\ 10^{-1}$ to 1.0	$8.78\ 10^{-1}$	-
RA	Recharge area	$2.3\ 10^{-5}$ to $3.0\ 10^{-5}$	$2.639\ 10^{-5}$	$1.3\ 10^{-1}$ to $2.0\ 10^{-1}$	$1.67\ 10^{-1}$	km^2
K_1	PET constant with Oudin's formula	-	-	80 to 110	90.633	°C
K_2	PET constant with Oudin's formula	-	-	1 to 10	5.475	°C
Ts	Snowmelt temperature threshold	-	-	2.0 to 4.6	3.26	°C
MF	Melt factor	-	-	3 to 7	4.38	$mm.°C^{-1}$
CFR	Refreezing factor	-	-	1.5 to 3.0	2.20	-
CWH	Water holding capacity of snow	-	-	6 to 10	9.58	-

IV.2.3.2.2. Simulation strategies and performance measurements

The definition of the model's warm-up, calibration, and validation periods depends on the data availability at each site. The warm-up period does not need to be particularly long, as its primary function is to initialize the model state before calibration. For effective model calibration, the calibration period should cover several hydrological cycles to capture the full range of rainfall–discharge interactions and

to understand the relationships involved. The validation period covers at least two hydrological cycles to ensure that the model can accurately reproduce both the timing and the magnitude of discharge events. The starts of each period are set at low-flow periods to align with the beginning of the hydrological cycle.

For the Séchilienne model simulation, the warm-up period spans six months from the beginning of the discharge data series. As this period coincides with a high-flow period, an initial water level for tank I was defined using the parameter I_0 to account for prior saturation. The calibration period extends over five years, from 2014/02/09 to 2019/01/09, and the validation period spans three years, from 2019/02/09 to 2022/12/31 (**Table 9**), excluding the three-month data gap discussed in the **"***Field data collection and processing"* section (see ***section IV.2.2.1.2****.)*.

For the Mont de La Saxe simulations, the warm-up period spans from 2013/01/01 to 2014/04/13, preceding the discharge measurements start. This strategy was chosen because rainfall data has been available since 2013/01/01, and the available discharge series is relatively short, less than five years. As such, the discharge data is used exclusively for model calibration and validation. The calibration period covers three years, from 2014/04/14 to 2017/02/05, and the validation period spans approximately two years, from 2017/02/06 to 2018/12/01, extending until the end of the discharge series (**Table 9**).

Table 9 | Warm-up, calibration, and validation periods for the Séchilienne and Mont de La Saxe model simulations.

	Séchilienne model	**Mont de La Saxe model**
Data collection	2014/03/07 to 2022/12/31	2013/01/01 to 2015/12/01
Warm-up period	2014/03/07 to 2014/09/01	2013/01/01 to 2014/04/13
Calibration period	2014/09/02 to 2020/03/08	2014/04/14 to 2017/02/05
Validation period	2020/03/09 to 2022/12/31	2017/02/06 to 2018/12/01

The objective function used for both study sites is a composite function that combines the Kling-Gupta Efficiency (KGE) and the Balance Error (BE), weighted by a factor of w = 0.8 (**Equation 13**).

Based on the model's performance, the threshold for this function is set to 0.75 for Séchilienne and 0.8 for Mont de La Saxe. These threshold values are deemed sufficiently high to ensure the selection of the most reliable simulations.

$$W_{OBJ} = wKGE + (1 - w)BE \quad (13)$$

The Kling-Gupta Efficiency (KGE), introduced by *Gupta et al. (2009)* (**Equation 14**), has become a widely used performance metric due to its ability to provide a more balanced evaluation of model performance compared to the Nash-Sutcliffe Efficiency *(Cinkus et al., 2023; Knoben et al., 2019; Willmott et al., 2012)*. The KGE integrates three performance components: the Pearson correlation coefficient r_P, the ratio of variances α, and the ratio of the means β. The KGE values range from $-\infty$ to 1, where a value close to 1 indicates that the simulation closely matches the observed data.

$$KGE = 1 - \sqrt{\left(r_p - 1\right)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \tag{14}$$

The Balance Error (BE) metric evaluates the model's accuracy in estimating total discharge over the simulation period (**Equation 15**). It ranges between 0 and 1, where a value of 1 indicates no error in the balance between observed and simulated discharge. Q_{OBS} and Q_{SIM} represent observed and simulated discharge values, respectively.

$$BE = 1 - \left|\frac{\sum Q_{OBS} - \sum Q_{SIM}}{\sum Q_{OBS}}\right| \tag{15}$$

The simulated discharge time series obtained (Q_{oPROB}) aims to represent the most probable discharge value at each time step by integrating the results of each retained simulation. This approach uses the weighted sum of simulated discharge values $Q_{O(i)}$ with the objective function values W_{OBJ}, calculated over the entire Y simulations exceeding the objective function threshold (**Equation 16**).

$$Q_{O_{PROB}} = \frac{\sum_{i=1}^{Y} W_{OBJ(i)} Q_{O(i)}(t)}{\sum_{i=1}^{Y} W_{OBJ(i)}} \tag{16}$$

The initial parameter ranges are broad to identify values that yield simulations exceeding predefined objective function thresholds. After analysing the parameter value distributions across all simulations, these ranges are refined to focus on the most relevant intervals, optimizing the model's computational cost (**Table 8**).

IV.2.4. Results

IV.2.4.1. Séchilienne and Mont de La Saxe hydrogeological simulations

Both models were run to achieve 10,000 simulations with a performance value exceeding the thresholds defined in the "*Simulation strategies and performance measurements*" section (see ***section IV.2.3.2.2.***) within a maximum runtime of six hours. The selected simulation, shown in red, represents the average simulation based on the weighted mean of 10,000 retained simulations, following **Equation 16**. The simulation with the best performance is also depicted as a brown dashed line. Finally, the 90% confidence interval, calculated from the 10,000 simulations, is displayed in orange to highlight the variability of the results (**Figures 26a, 28a**). **Table 10** summarizes the performance criteria and their parameters values obtained for simulations in calibration and validation phases.

The Séchilienne model simulation shows satisfactory performance during the calibration phase and moderate performance during the validation phase, with results of 0.77 and 0.62, respectively. The simulated discharge time series closely matches the observed discharge values. However, while the model accurately reproduces the timing of discharge peaks, it struggles to capture their amplitude. Moreover, the simulations tend to exhibit noise during low-flow periods. The observed noise is attributed to the hysteresis flow mechanism, which regulates the peaks' timing, amplitude, and duration. However, this structure remains essential for the model's overall functionality. The best simulation obtained, represented by the dashed line, is slightly less affected by the noise during low-flow periods (**Figure 26a**).

Table 10 | Summary of performance criteria and parameters values for the Séchilienne and Mont de La Saxe simulations during the calibration and validation periods.

	Séchilienne model		**Mont de La Saxe model**	
	Calibration	**Validation**	**Calibration**	**Validation**
α	0.88	0.95	0.91	0.87
β	1.02	1.13	1.00	1.31
r_P	0.75	0.58	0.89	0.84
WOBJ	0.77	0.62	0.86	0.65
BE	0.98	0.87	0.99	0.69
KGE	0.72	0.56	0.86	0.63
DE	0.38	0.49	-	-

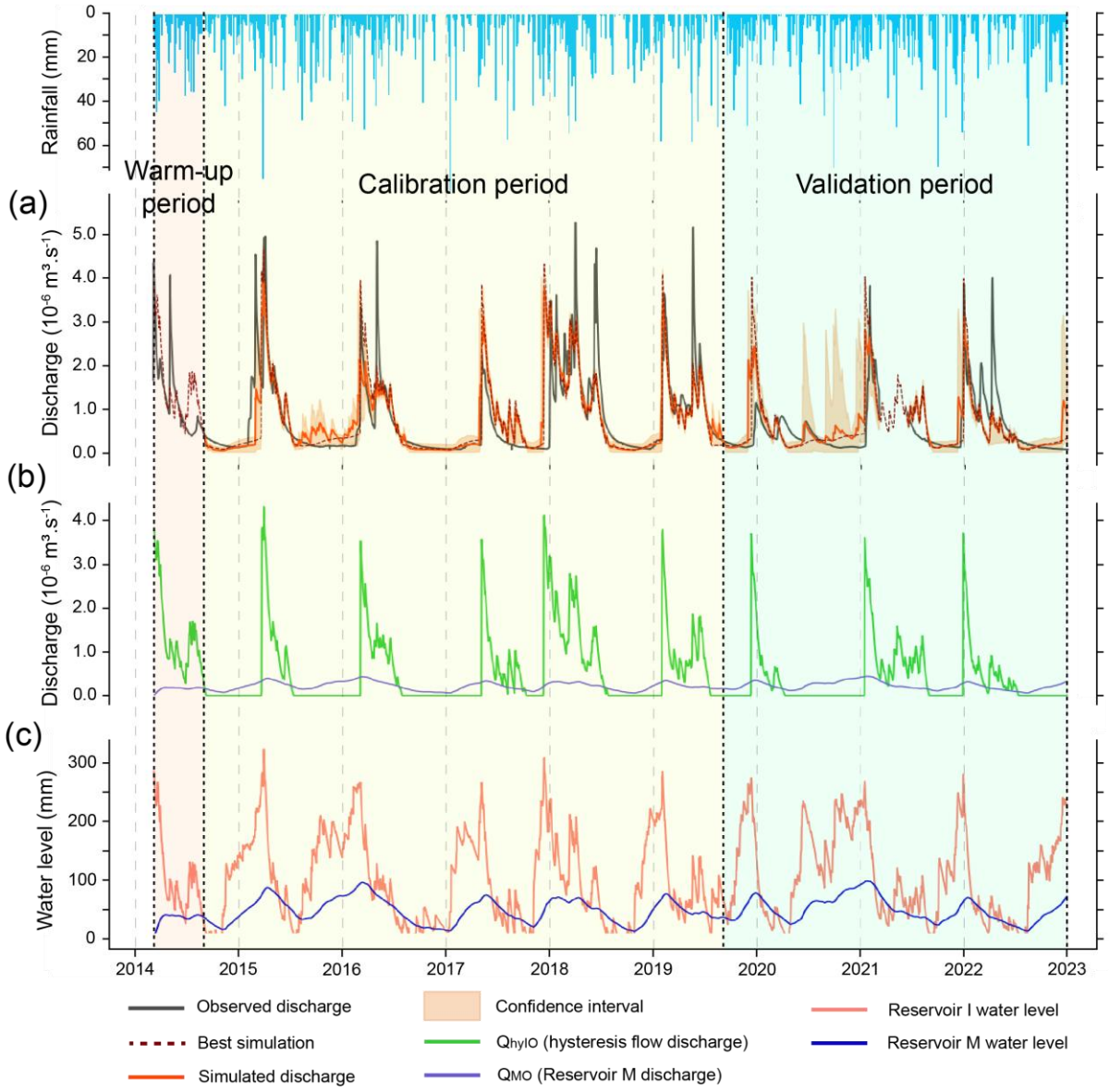


Figure 26 | Results of the Séchilienne model simulation during the calibration and validation periods. **(a)** Observed (black) and simulated (red) discharges with the confidence interval of the simulation (orange area) and the best discharge simulation (dashed brown line). **(b)** Internal discharges representing the matrix (blue) and fracture (green) contributions to the total simulated discharge. **(c)** Water levels in reservoirs I (pink) and M (blue).

The BE remains high, while the KGE significantly decreases during the validation period, dropping from 0.72 to 0.56 (**Table 6**). For each parameter, calibration and validation periods are denoted by the "cal" and "val" indexes. The simulated time series reveals a slight underestimation of variability (α_{CAL} = 0.882 and α_{VAL} = 0.952) and a tendency to overestimate flow volumes (β_{CAL} = 1.019 and β_{VAL} = 1.126). During calibration, the model shows a correct representation of shape and timing variation (r_{pCAL} = 0.75) but only moderate performance during validation (r_{pVAL} = 0.59). The Diagnostic Efficiency (DE) is acceptable during the calibration period, represented by the black dot, with a value of 0.38 and a very

low timing error. The error is primarily associated with low-flow periods when the simulations tend to be consistently underestimated ($B_{SLOPE} < 0$). In contrast, the simulation for the validation period presents a moderate DE value (0.49) with contributions from all three DE error components. The model error can be attributed to the constant error, indicating that simulated flows are generally overestimated compared to observations ($B_{REL} > 0$), with close contributions from high flows and low flows. The dynamic error is slightly positive, showing a tendency to overestimate low flows ($B_{SLOPE} > 0$), while the timing error is middling, highlighting the model's difficulty in accurately simulating flow fluctuations, particularly during low-flow periods (**Figure 27**). This issue is especially pronounced in hydrological years with lower discharges, such as 2020 (**Figure 26a**).

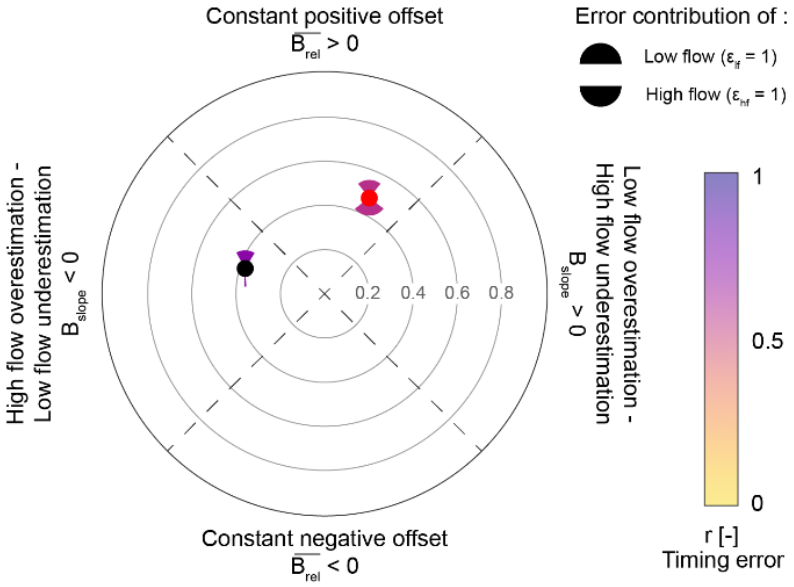


Figure 27 | Diagnostic efficiency polar plot for the Séchilienne model simulations during calibration (black) and validation (red) periods.

The Mont de La Saxe model simulation presents highly satisfactory performance during the calibration phase and acceptable performance during the validation phase, with results of 0.89 and 0.65, respectively. The redistribution of precipitation using the snow routine module introduces a substantial and punctual water supply into the system, allowing the model to simulate the timing and amplitude of discharge peaks accurately. However, the snowmelt modelling also increases the model's sensitivity during the first drying period of 2014 and leads to a gradual decrease in flood peak amplitude over time (**Figure 28a**).

The BE shows a significant decrease from 0.99 to 0.69, suggesting that the model struggles to accurately replicate observed discharges during the validation period. The KGE parameter values corroborate this observation (Table 6). The simulations exhibit slightly underestimated variability (α_{CAL}

= 0.91; α_{VAL} = 0.87) but accurately represent shape and timing variations (r_{pCAL} = 0.89; r_{pVAL} = 0.84) for calibration and validation phases. However, the relative bias is significantly overestimated during the validation period compared to the calibration period (β_{CAL} = 1.002; β_{VAL} = 1.306). This overestimation can be attributed to the model's difficulty in accurately replicating the spring's drying periods and the faster recession of recent discharge peaks, resulting in a more prolonged recession of simulated discharge than observed discharge. The diagnostic efficiency measurement for the Mont de La Saxe model simulation was not applied, as this method requires non-zero discharge values.

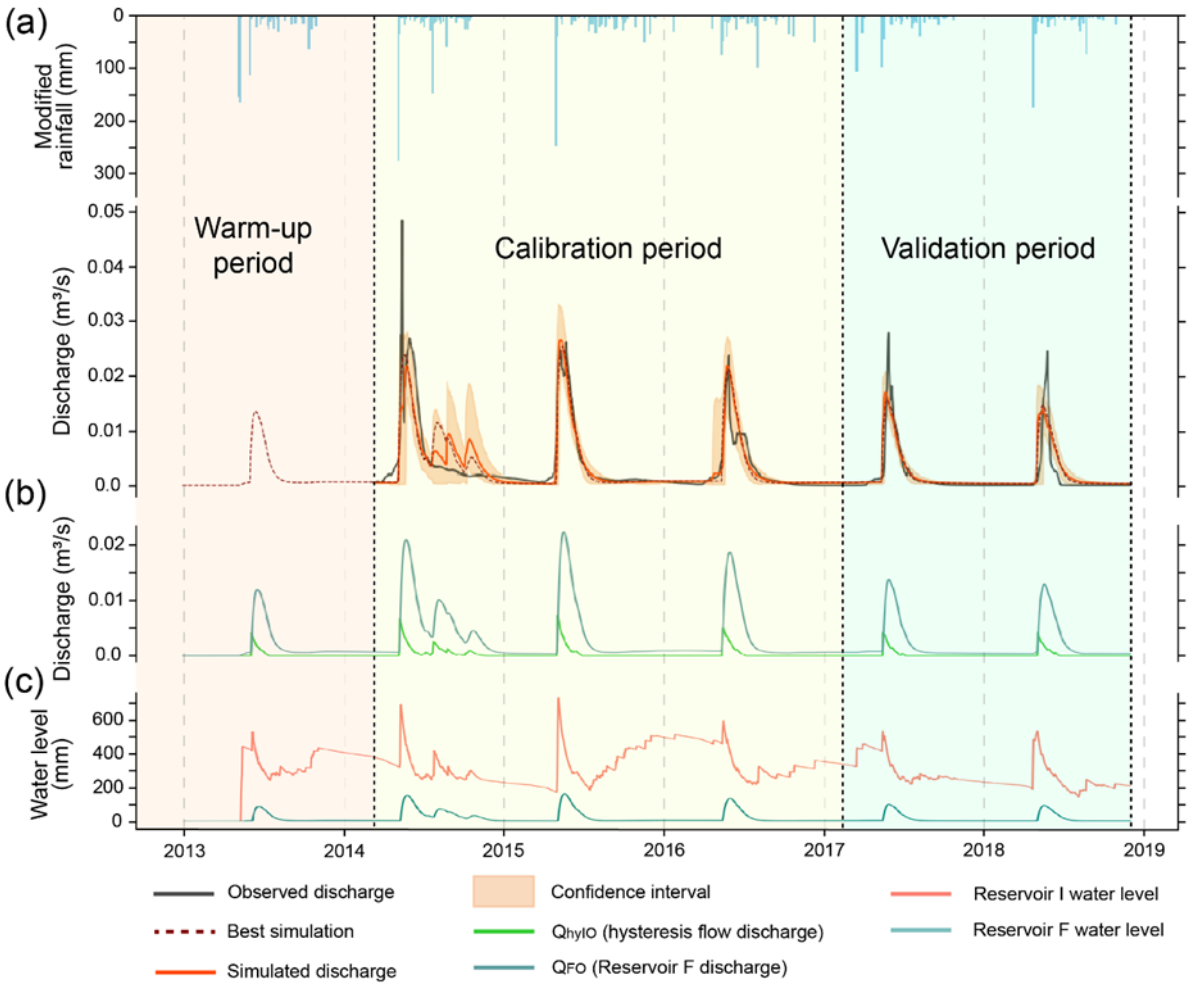


Figure 28 | Results of the Mont de La Saxe model simulation during the calibration and validation periods. **(a)** Observed (black) and simulated (red) discharges with the confidence interval of the simulation (orange area) and the best discharge simulation (dashed brown line). **(b)** Internal discharges representing the fracture (blue) and the hysteresis (green) contributions to the total simulated discharge. **(c)** Water levels in reservoirs I (pink) and F (blue).

IV.2.4.2. Impact of parameter sensitivity and calibration length on model performances

The sensitivity of the simulations to parameter variations is quantified using first-order and total-effect Sobol sensitivity indices. These indices capture the direct contribution of a parameter and its total contribution, including interactions with other parameters, to the simulation's variance. Sensitivity analysis was performed using 9,000 simulations for the Séchilienne model and 18,000 for the Mont de La Saxe model, with the number of simulations proportional to the number of parameters in each model. For this analysis, parameter ranges were expanded compared to the refined ranges used in high-performance simulations (**Table 8**) to allow larger variability and ensure robust random sampling during the sensitivity evaluation. The most sensitive parameters for both models were those related to the flow recession coefficients connecting the reservoirs and the recharge area (**Figure 29**).

In the Séchilienne model, the recession coefficients for matrix flow and hysteretic fracture flow, k_{hy} and k_{IM}, exhibit moderate to low sensitivity in first-order indices (0.22 and 0.03, respectively), but become more sensitive when interactions with other parameters are considered (0.65 and 0.42, respectively). The recharge area RA consistently remains a sensitive parameter for the simulation (0.19 and 0.40 for first-order and total-effect indices, respectively), while the structural parameters $T1_{hy}$ and $T2_{hy}$ exhibit low or nearly negligible first-order sensitivity indices (**Figure 29a**). However, based on expert knowledge and tests conducted without incorporating the hysteretic function, these parameters remain crucial for accurately simulating the non-linear discharge variations at the outlet in response to system saturation. The dominance of k_{hy} and k_{IM} highlights the importance of a bimodal flow process in accurately reproducing discharge dynamics over the study period.

In the Mont de La Saxe model, the recharge area *RA* and the recession coefficients k_{FO} and k_{hy} exhibit significant first-order sensitivity (0.25, 0.18, and 0.11, respectively), which nearly doubles when interactions with other parameters are considered (0.44, 0.33, and 0.40). The first-order sensitivity index of the recession coefficient for the flow connecting Reservoir I to Reservoir F (k_{IF}) is almost negligible (0.007) but becomes significant in the total-effect index (0.31). This indicates that k_{IF} does not directly affect discharge when other variables are constant. However, it is significant in interaction with other variables as it regulates the infiltration rate from Reservoir I to Reservoir F, which can significantly influence the timing of surpassing the $T2_{hy}$ threshold. Structural parameters x_{hy}, $T1_{hy}$, and $T2_{hy}$ are low-sensitive, with slightly higher values for the total-effect index. However, as in the Séchilienne model, structural tests conducted without the dual threshold produce noisier discharge simulations that are less aligned with the observed series, indicating that $T1_{hy}$ and $T2_{hy}$ remain essential parameters for ensuring accurate discharge simulations. Lastly, the parameters related to the Snow routine and PET routine modules do not significantly affect result variance. The four Snow routine parameters represent

'physical' processes for simulating snow dynamics and are restricted within defined value ranges. The PET routine parameters, K_1 and K_2, are solely involved in evapotranspiration calculation and could be omitted if external evapotranspiration is estimated using Oudin's method. Lastly, k_{IF}, representing direct flow between reservoirs I and F, is the only parameter with negligible influence. This implies that the model operates primarily through the hysteresis-fed reservoir F, functioning predominantly unimodally with a main drainage component (**Figure 29b**).

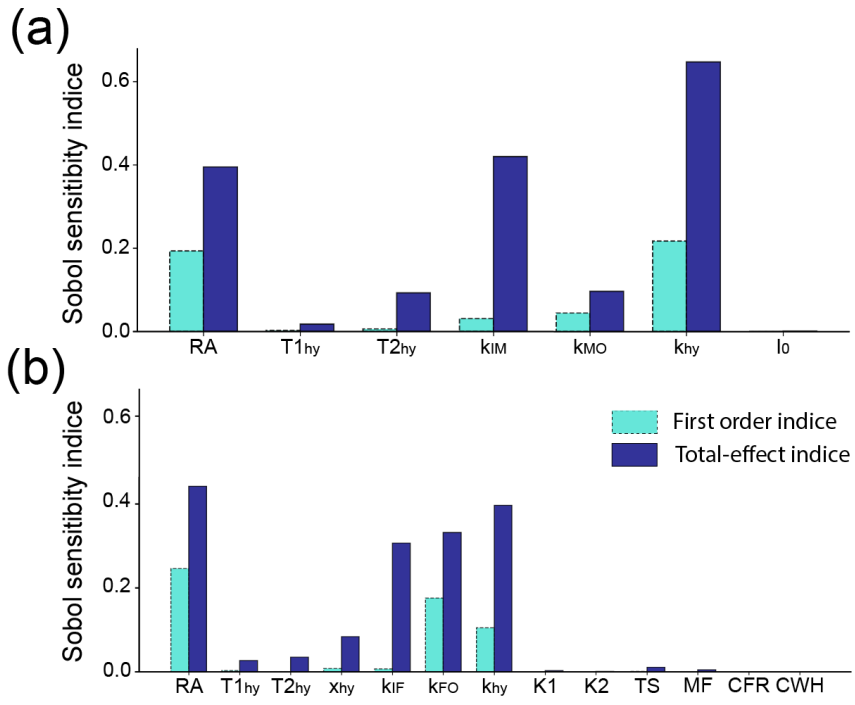


Figure 29 | First-order and total-order Sobol sensitivity index values for **(a)** the Séchilienne and **(b)** the Mont de La Saxe models' parameters.

Finally, the effect of calibration period length on model performance during validation is evaluated by progressively reducing the calibration period to observe model performance when fewer hydrological cycles are provided for training. This reduction is performed incrementally by one hydrological year. For periods less than one year, it was done over 6- and 3-month intervals preceding the validation period.

The different training periods for both models are labelled PX (e.g., P0, P1, etc.), where P0 represents the calibration periods used in this study. **Figure 30** presents discharge forecasts for the validation period and their objective function values. Reducing the calibration period systematically results in a slight decline in model performance and tends to stabilize as long as at least one full hydrological year is included in the calibration period. In the Séchilienne model, model performance decreases when the calibration period is reduced to less than one hydrological year, with simulations

becoming less precise during recession and matrix recharge periods (**Figure 30a**). This observation appears to be similar for the Mont de La Saxe model (**Figure 30b**). However, no simulations reaching the objective function threshold were retained when less than one hydrological year was provided (periods P3 and P4).

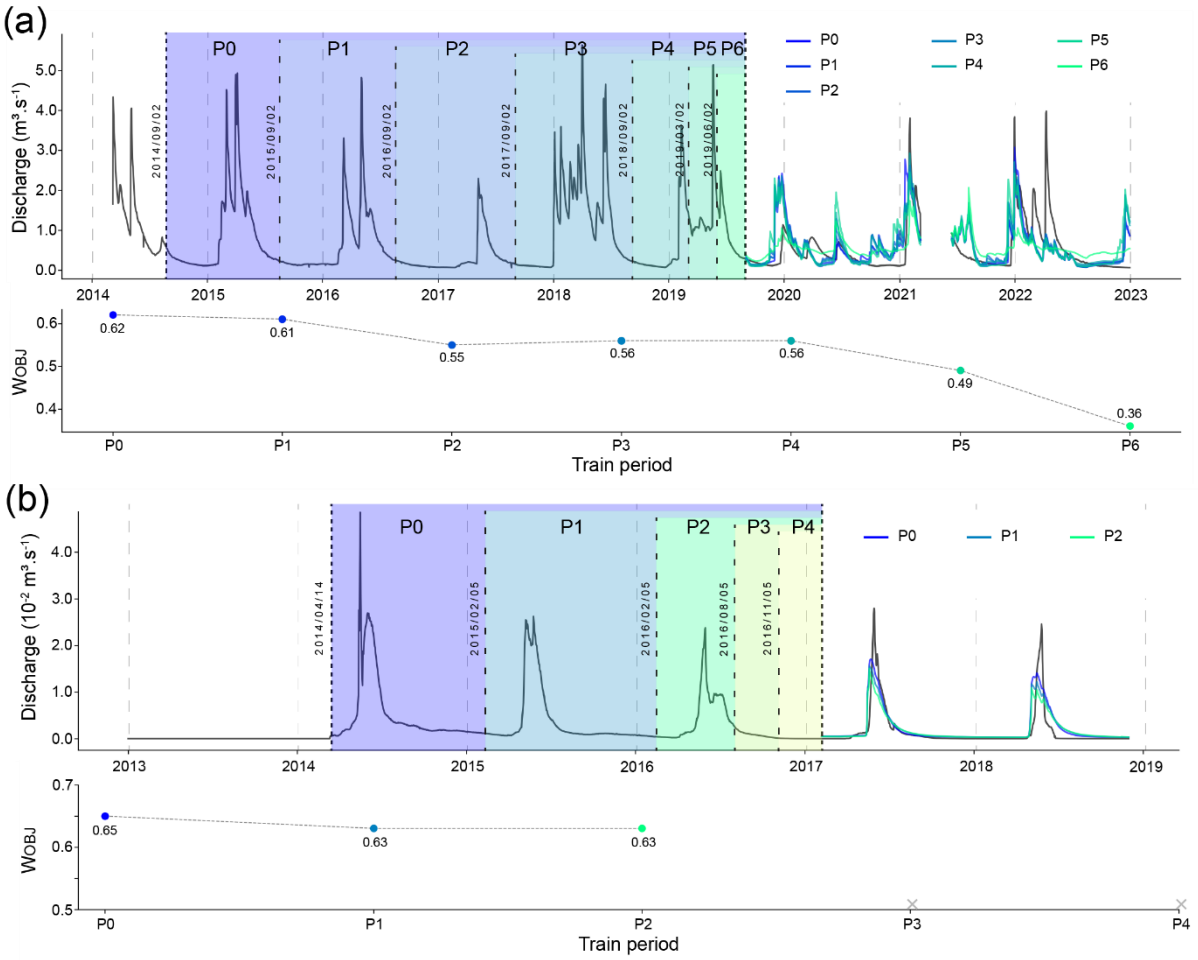


Figure 30 | Discharge simulations during the validation period with their performance values, based on the different calibration period lengths tested represented by P for **(a)** the Séchilienne and **(b)** the Mont de La Saxe models. The P0 periods represent the calibration periods used in this study to calibrate the models. The colours of the calibration periods correspond to the colours of the discharge simulations and the objective function values.

IV.2.4.3. Retrospective simulations using historical data

One of the notable advantages of the KarstMod platform lies in its possibility to extend discharge time series beyond the studied period, even in the absence of direct discharge measurements, as long as the input data are consistent and continuous and the model has been appropriately trained and validated. This functionality is particularly useful for predicting discharge variations based on meteorological forecasts and reconstructing past discharges before installing measurement stations. The

latter application is especially valuable, as it generates an extended hydrological time series, enabling the identification of correlations between hydrological behaviour and landslide dynamics.

The approach was implemented for the Séchilienne model using the meteorological records from the Séchilienne weather stations, which provide continuous records since October 1992. To extend the discharge simulations to 1992, we set the model parameters to their optimized values, as determined during the calibration phase (**Table 8**).

The projections successfully replicated the hydrological cycles, including low-flow and high-flow periods, based on the model's calibrated behaviour and in response to the incoming precipitation data. Notably, some years exhibited prolonged matrix recharge and periods of high system reactivity, particularly between 2002 and 2003, offering valuable insights into the historical internal state of the Séchilienne system. Furthermore, the simulated discharge is accompanied by internal discharge and reservoir water levels time series, providing additional insights into the matrix-fracture flow distribution and variations in water storage within the system (**Figure 31**).

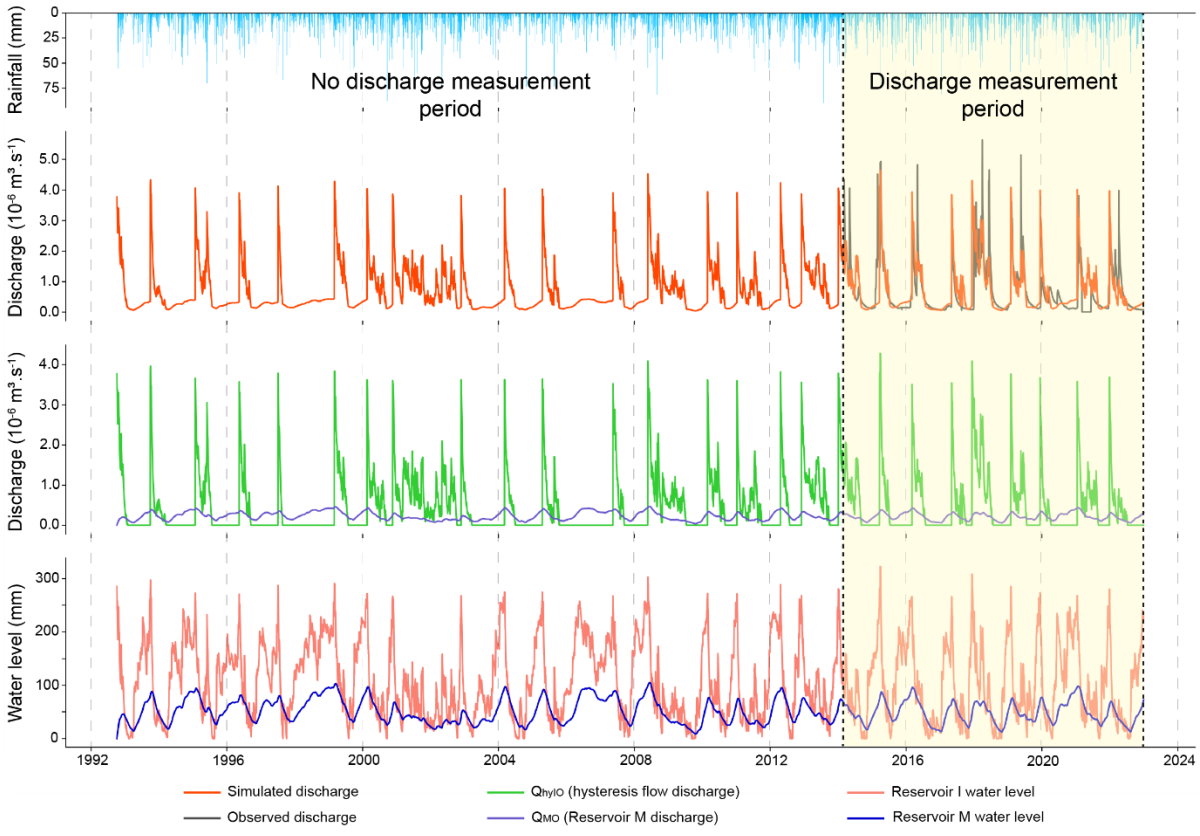


Figure 31 | Discharge simulations from the calibrated and validated Séchilienne model using historical meteorological data up to 1992. The discharge acquisition period is shown in yellow. The model simulated time series (internal discharges and water levels) follow the same colour code as in **Figure 26**.

IV.2.5. Discussion

The tank models developed to simulate discharge in fractured media using the KarstMod platform demonstrated good performance, highlighting two distinct flow dynamics within a homogeneous lithology. The Séchilienne model represents a bimodal flow, where the rock matrix stores water and releases it slowly while the fractures respond quickly to precipitation events. In contrast, the Mont de La Saxe model simulates an unimodal flow, where water mainly drains through a network of fractures. A key feature shared by both models is the fast flow component with hysteretic behaviour, meaning the flow only activates when the system is saturated, contributing to the observed discharge. This characteristic allows the models to replicate the system's reactivity during high-flow periods while reducing the impact of heavy rainfall during low-flow periods by promoting water storage, similar to the approach used by *Zheng et al. (2024)* with trapezoidal reservoirs to capture non-linear changes due to intense precipitation.

The use of parallel reservoirs effectively represents the different flow patterns within the same lithology and helps provide a better understanding of the internal processes, particularly how each component contributes to the observed discharge at the outlet. The models are also well-suited to the specific climatic conditions of the study sites, as shown by the Mont de La Saxe model, where the inclusion of the snow routine module for precipitation redistribution allowed for accurate simulation of snowmelt and the timing of discharge peaks.

Analysis of the training period showed that a large dataset is not required to accurately simulate flood and recession periods as long as at least one full hydrological year is available for model calibration. However, having as much data as possible is still preferable to better capture the system's response to precipitation.

Despite their strengths, the models exhibit limitations due to their structure and the functionality of KarstMod in managing water distribution across different reservoirs and flows. The Séchilienne model struggles to consistently reproduce the amplitude of discharge peaks during high-flow periods, except for the first peak. These results indicate that the inflow volume might be insufficient due to a limited recharge area or the contribution of external waters during saturation periods. The recharge area values derived from the simulation are tightly constrained, offering limited flexibility for model adjustment. The optimal value is consistent with previous observations *(Cappa et al., 2004; Nevers, 2021; Vallet et al., 2014)*, which estimated a recharge area between 9 and 204 m². The missing volume could originate from the Mont Sec sedimentary plateau's perched aquifer, as *Vallet et al. (2014)* suggested, who demonstrated that this aquifer connects to the landslide system only during saturation periods. However, KarstMod cannot dynamically adjust the recharge area based on system saturation,

nor can it incorporate an external flow or another infiltration tank triggered by specific water level thresholds in reservoir I, highlighting a potential area for future model enhancements.

While the Mont de La Saxe model accurately simulates discharge using the snow routine and precipitation redistribution, it also has some limitations. The continuous increase in water levels in reservoir I does not realistically reflect water fluctuations in the system, limiting its interpretative value. The model sets a high threshold for activating hysteretic flow, allowing it to store water from small precipitation events and only activate the flow during significant snowmelt. This dynamic may be related to the spring's altitude, which is often dry and only becomes active during snowmelt. Matrix water is likely present, contrary to the chosen model structure, but continuous measurements from downstream springs would be needed to apply another tank model and confirm this.

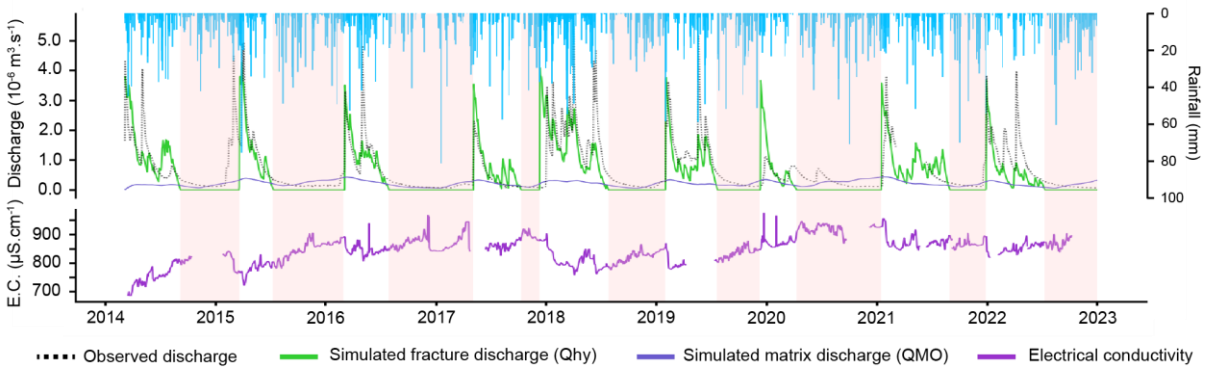


Figure 32 | Comparison of the internal discharge time series representing the matrix contribution (blue) and fracture contribution (green) from the Séchilienne model with electrical conductivity (E.C.) measurements (purple). The observed discharge time series is shown in black dashed lines, and the pink areas indicate periods where the matrix contribution is the only input to the outlet.

The projection of discharge over historical periods provides theoretical insights into the hydrological state of the system at different times, especially in terms of water level fluctuations within the reservoirs. Although these levels are theoretical, they give useful information on the system's reactivity to precipitation and its behaviour during both high- and low-flow periods. The internal discharge contributions from the matrix and fracture flow to the outlet are compared with electrical conductivity measurements taken at the same outlet to validate the relevance of the model's internal time series (**Figure 32**). *Nevers (2021)* demonstrated that the outlet's electrical conductivity fluctuations are linked to the water residence time within the rock matrix and its subsequent enrichment in elements like sulphate. Matrix water exhibits higher electrical conductivity due to its longer residence time than faster-flowing water through fractures. The comparison revealed that most periods with increasing conductivity signal coincide when the discharge is primarily sustained by matrix flow or when fracture flow contributions drop rapidly (**Figure 32**). These observations suggest that the internal discharge simulations are consistent with field observations. However, the lack of on-site piezometric

measurement stations prevents a full validation of the model concerning the fluctuations of the reservoir water levels.

These data could later be further integrated into machine learning models, as proposed by *Krkač et al. (2017)*, to explore the relationship between hydrology and landslide dynamics, demonstrating that theoretical water level data from indirect methods can be effectively used. However, this approach lies beyond the scope of this article and will be considered in future research.

IV.2.6. Conclusion

Deep-seated landslides are primarily driven by hydrometeorological events such as intense rainfall or significant snowmelt, which cause fluctuations in pore pressure and groundwater levels that can lead to acceleration phases. However, direct hydrological monitoring in these environments is often challenging. An effective alternative involves using conceptual tank models to simulate hydrological processes based on the rainfall–runoff relationship. This study applied such models to two fractured rock environments, the Séchilienne and Mont de La Saxe deep-seated landslides.

The models could accurately simulate discharge time series through calibration and validation over several years. The performance of the models, assessed through an objective function combining the Kling-Gupta Efficiency (KGE) and Balance Error (BE), demonstrated satisfactory results with calibration values of 0.77 and 0.89, and acceptable validation results of 0.62 and 0.65. These outcomes highlight two distinct hydrological behaviours in fractured environments. The Séchilienne model exhibits bimodal behaviour, with a slow matrix contribution maintaining baseflow and a rapid, more reactive flow through fractures. On the other hand, the Mont de La Saxe model is controlled predominantly by fracture drainage, which results in periodic drying of the outlet. Integrating the snow routine in the Mont de La Saxe model allowed for precise replication of snowmelt dynamics, accurately capturing the timing and amplitude of discharge peaks. Furthermore, by incorporating historical meteorological records, the models can reconstruct past discharges, providing valuable insights such as groundwater levels fluctuations and the contribution of different flow modes within deep-seated landslides before the installation of monitoring systems. Using rainfall–discharge-based tank models presents a practical approach to simulating hydrological processes in fractured rock environments. The acquisition of longer groundwater level time series helps to supplement landslide monitoring databases and addresses the gap in hydrological monitoring data. Additionally, the growing use of machine learning algorithms, leveraging a data-driven approach to develop predictive displacement models, particularly at a daily frequency, underscores the need for comprehensive data associated with triggering factors such as rainfall and groundwater level fluctuations. The more extensive the data series, the better these models can identify complex and non-linear relationships between movements and external factors, thereby improving their predictive capabilities. Reconstructing hydrological series using tank

models thus addresses the data availability issue for applying these models. Future research will aim to integrate these reconstructed data into machine learning-based models to identify relationships between groundwater level fluctuations and acceleration-deceleration phases and explore their potential in predicting landslide movements.

IV.3. CHAPTER IV - CONCLUSION

This chapter addressed the first research question of the thesis and the first step of the methodological workflow introduced in ***Chapter III***, namely the need to collect, consolidate, and prepare long, informative, and representative multi-source time series before any robust application of conventional machine learning models to landslide displacement simulation. Because slow-moving landslide kinematics are strongly modulated by hydro-meteorological forcing through subsurface water storage and groundwater level, the availability of hydrological state information is a limiting factor for building predictors that correctly encode system memory and delayed responses.

Chapter IV therefore proposed an alternative strategy by reconstructing groundwater state proxies from discharge observations, using a semi-empirical conceptual reservoir representation driven by meteorological inputs. In the perspective of the overall thesis workflow, these reconstructed proxies enabling the completion and extension of key predictor series that would otherwise be missing for displacement simulations.

Beyond the site-specific results, the chapter also illustrates the broader value of conceptual reservoir modelling as a form of scale transfer in heterogeneous fractured media. Although the available observations are generally non-spatialised (time series at meteorological stations and outlets), the reservoir representation integrates the effects of internal heterogeneity into a parsimonious set of effective compartments and flow paths. In this sense, the calibrated tank structures provide a global and operational description of the hydrological functioning of the landslide mass, capable of reproducing contrasted behaviours (e.g., bimodal matrix–fracture contributions versus unimodal preferential drainage) while remaining compatible with the limited information typically accessible in deep-seated systems.

In addition, conceptual reservoir models calibrated on discharge offer a practical tool to (i) fill gaps and lengthen incomplete time series, (ii) reconstruct retrospective hydrological states prior to monitoring installation, and (iii) support process-oriented interpretation of storage–release dynamics and system reactivity under contrasted hydro-meteorological conditions, including under future climate scenarios. These outcomes are relevant both for improving the representativeness of datasets intended for CML modelling, and for broader hydrogeological analyses where the objective is to characterise storage variations and hydrological regimes.

CHAPTER V

FROM HYDRO-METEOROLOGICAL SERIES TO INTERPRETABLE VELOCITY SIMULATIONS: A XAI-BASED WORKFLOW

CHAPTER V FROM HYDRO-METEOROLOGICAL SERIES TO INTERPRETABLE VELOCITY SIMULATIONS: A XAI-BASED WORKFLOW 101

V.1. Chapter V – Summary 103

V.2. Paper – A Scalable XAI-based Forecasting of Landslide Velocities From Environmental Forcings (Nature Geosciences, submitted) 106

V.2.1. Introduction 106

V.2.2. Multi-site monitoring datasets 107

V.2.3. Explainable Artificial Intelligence framework 110

V.2.4. Model performance: simulated landslide velocities 118

V.2.5. Interpretability and explainability: environmental predictor controls 121

V.2.6. Conclusion: Implications for interpreting landslide dynamics with XAI framework 125

V.2.7. Extended data 127

V.3. Chapter V – conclusion 132

V.1. CHAPTER V – SUMMARY

This chapter focuses on the core modelling and interpretation stages of the methodological framework introduced in ***Chapter III***, which aim to simulate slow-moving landslide displacement time series while making the underlying forcing–response relationships explicit. Building on long-term multi-source monitoring datasets, and potentially complemented by the groundwater-state proxies introduced in ***Chapter IV***, ***Chapter V*** therefore directly adresses the second research question of this thesis: **How do hydro-meteorological controls shape the temporal dynamics of surface displacement, and over which characteristic time scales?**

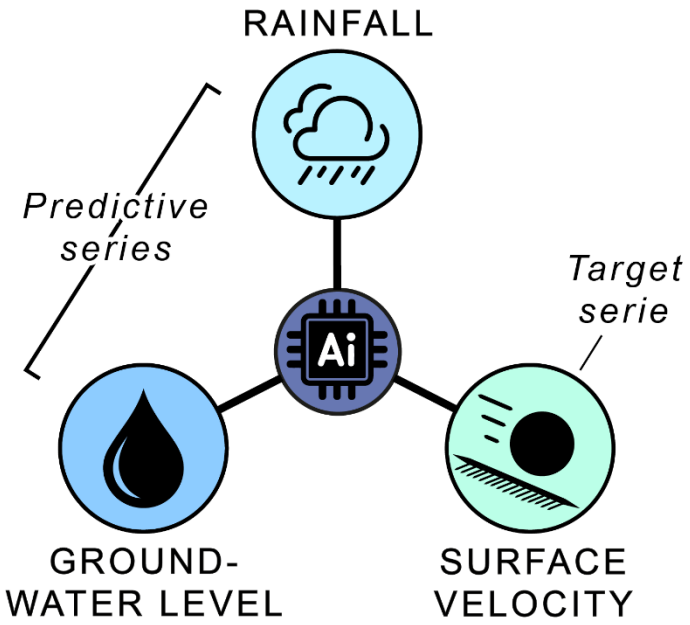


Figure 33 | Inputs and target used in the XAI workflow. Rainfall and groundwater level describe hydro-meteorological forcing and hydrological state (predictors), while surface velocity is the target simulated by XGBoost.

Forecasting the daily surface velocity dynamics of slow-moving landslides remains challenging because forcing–velocity relationships are multivariate, strongly non-linear, and expressed across multiple time scales, with pronounced memory effects and time lags. As outlined in ***Chapter II***, data-driven approaches have increasingly been adopted to reproduce these complex behaviours. However, two limitations persist: (i) many approaches remain data-intensive and difficult to interpret, and (ii) modelling efforts are often restricted to single site demonstrations, providing limited constraints on robustness and genericity under contrasted geological and hydrological settings.

This chapter addresses these limitations by introducing a scalable eXplainable Artificial Intelligence (XAI) framework designed to (i) simulate daily velocity dynamics from hydro-

meteorological time series and (ii) make the inferred forcing–response relationships transparent and comparable across sites. In continuity with ***Chapter III***, the framework combines time series pre-processing, exploratory statistical analysis, feature engineering, a structured train-test splitting strategy, and performance evaluation, complemented by a dedicated interpretability/explainability layer.

The framework is built around XGBoost and a generic set of 248 hydrometeorological predictors physically-grounded and site-independent, derived from net rainfall, effective rainfall, and groundwater level series (**Figure 33**). Predictors are computed over multiple time windows (1-90 days) to encode short-, intermediate-, and long-term responses, as well as delayed kinematic effects informed by exploratory correlation analyses. Model training and testing rely on a blocked time-series validation scheme, in which the record is partitioned into consecutive periods that are sequentially withheld for testing, thereby preserving temporal structure.

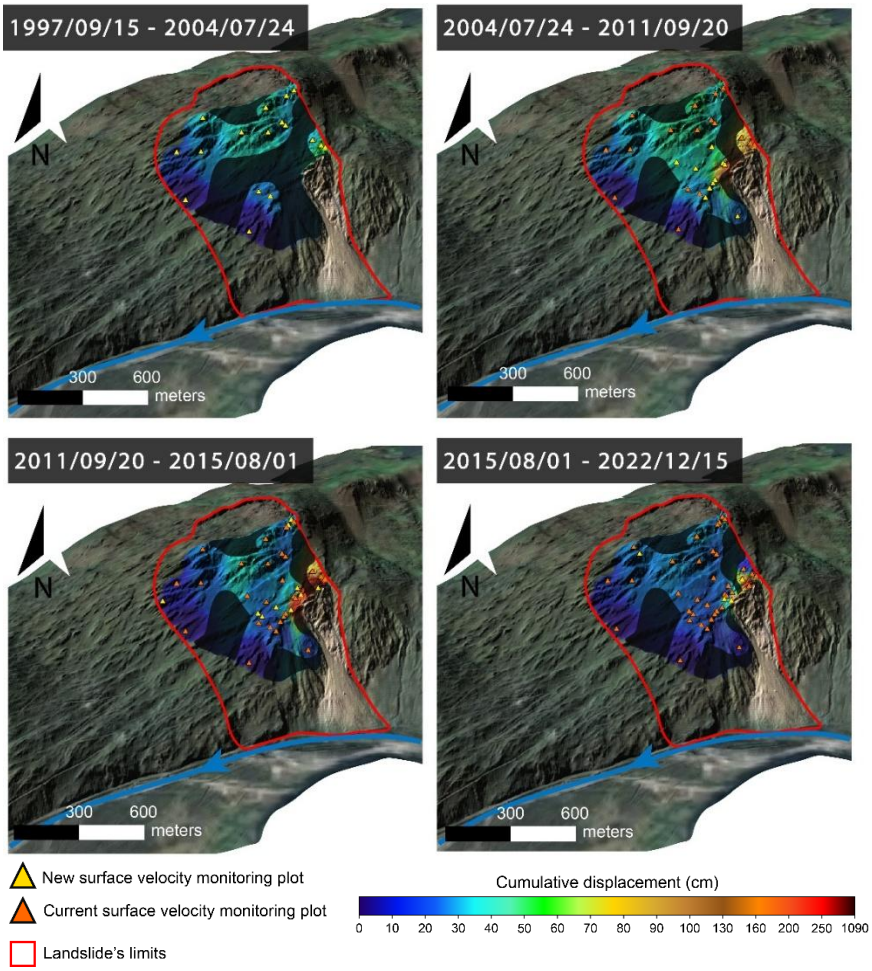


Figure 34 | Maps of cumulative surface displacement measured on the Séchilienne landslide over four distinct periods between 1997 and 2022, superimposed on a 1-m-resolution shaded DEM. Periods are defined according to the progressive densification of the monitoring network, with newly installed stations shown as yellow triangles and pre-existing velocity plots as orange triangles. Cumulative displacements are displayed using a colour gradient. Black areas mask zones where interpolation is considered unreliable due to the absence of nearby plots.

The approach is applied to three contrasting slow-moving landslides (Séchilienne, Viella, and Villerville) and to multiple monitoring plots per site, acknowledging that forcing–response relationships can vary spatially within a single unstable mass. At Séchilienne, for instance, the surface displacement monitoring network enables the production of time-evolving maps of observed cumulative displacement, which highlight both the spatial variability of displacement and its heterogeneous evolution through time (**Figure 34**).

Across sites and plots, the models reproduce the alternation between low-activity phases and acceleration/deceleration episodes with satisfactory error levels, although some rare extreme peaks remain difficult to capture. Model transparency is achieved by combining native XGBoost feature-importance ranking with post-hoc SHAP attribution, revealing contrasted dominant controls and characteristic time scales between sites, and highlighting within-site variability consistent with local heterogeneity and hydrological connectivity.

The results show that a single, physics-informed and site-independent predictor set can support transferable velocity simulations while providing interpretable constraints on hydro-meteorological controls and their temporal signatures.

In the perspective of the thesis modelling framework introduced in **Figure 17** (see ***section III.1.***), ***Chapter V*** **thus contributes to steps 1-6 by implementing the conventional machine-learning modelling chain – including data preparation, feature engineering, velocity modelling, and performance evaluation – while further addressing the interpretability/explainability stage to make the inferred forcing–velocity relationships explicit and comparable across sites.**

The following section (section V.2.) consists on a manuscript submitted to Nature Geoscience special collection "Landslide trigger and hazards":

O. Béjean-Maillard, C. Bertrand, J.-P. Malet, L. Dubois, C. Batailles, L. Lespine, O. Maquaire, M. Fressard, and J. Ducasse. *Scalable XAI-Based Forecasting of Landslide Velocities from Environmental Forcings* (submitted).

V.2. PAPER – A SCALABLE XAI-BASED FORECASTING OF LANDSLIDE VELOCITIES FROM ENVIRONMENTAL FORCINGS (NATURE GEOSCIENCES, SUBMITTED)

V.2.1. Introduction

Forecasting the evolution of slow-moving landslides remains a fundamental challenge in geoscience. These deep-seated slope instabilities pose a persistent threat to populations and infrastructure worldwide *(Froude & Petley, 2018; Mansour et al., 2011)*, yet their kinematics typically evolve over years to decades, hindering reliable anticipation of acceleration phases and failure *(Cascini et al., 2014; H. Du & Song, 2022; Federico et al., 2012)*. Slow-moving landslides occur across a broad range of geomorphic settings *(Kirschbaum et al., 2015)*, from ridge crests to valley flanks, and commonly involve low-strength materials such as clay-rich horizons or highly weathered bedrock *(Lacroix et al., 2020)*.

Forecasting changes in landslide surface velocities is difficult *(Carlà et al., 2019; Intrieri & Gigli, 2016)* as landslide motion is modulated by hydro-meteorological forcing (rainfall, snowmelt, and groundwater fluctuations) acting across multiple timescales *(Bonzanigo et al., 2007; Handwerger et al., 2013; Iverson & Major, 1987; Malet et al., 2005; Schulz et al., 2018)*. The complexity in landslide mechanical behaviour thus limits the predictive capability and transferability of physically-based models. By leveraging long-term multi-parameter landslide monitoring and the forcing response relationship *(Bernardie et al., 2015; Corominas et al., 2005)*, eXplainable Artificial Intelligence (XAI) offers a scalable pathway toward transferable velocity forecasting across diverse landslide settings.

Although AI models have achieved some predictive skill for simulating landslide velocity across many sites, two limitations remain. First, most approaches rely on deep learning architectures, ranging from Artificial and Recurrent Neural Networks (ANNs, RNNs ; *Yao et al., 2015*) to more recent time-aware models such as Gated Recurrent Units (GRUs ; *W. Ma et al., 2022; W. Zhang et al., 2022*) and Long Short-Term Memory (LSTMs ; *Nava et al., 2023; L. Wang et al., 2023*) networks that successfully reproduce landslide non-linear dynamics *(van Natijne et al., 2022; Zou et al., 2020)* but do not constrain landslide physics and have limited interpretability and transferability. Recent efforts to improve interpretability through Transformer-based architectures *(Z. Ma & Mei, 2025)*, represent important progress, yet they remain computationally demanding and challenging to generalize across sites.

Second, few AI applications address landslides governed by hydro-meteorological forcing *(Neaupane & Achet, 2004; Xie et al., 2019)* where velocity is controlled by the combined effects of rainfall,

snowfall, shallow and deep groundwater flows. This imbalance restricts the broader applicability of AI-based forecasting approaches to the majority of slow-moving landslides worldwide.

By contrast, conventional machine learning approaches, such as Random Forests *(Krkač et al., 2017)* and Gradient Boosting *(Collini et al., 2022; J. Xu et al., 2022a)* demonstrated strong predictive skill in landslide velocity simulation while offering improved transparency. However, existing applications of conventional machine learning are largely site-specific, relying on tailored predictor sets and local calibration that limit transferability across landslides with contrasting geological and hydrological conditions. Addressing this limitation requires interpretable modelling frameworks capable of operating across multiple landslide sites including data-scarce settings.

Here we introduce a scalable modelling framework using eXtreme Gradient Boosting (XGBoost) and based on a large set of physically motivated hydro-meteorological predictors. The approach simulates daily landslide velocities and provides interpretable and explainable constraints on the relative influence of the environmental drivers. We evaluate and discuss the modelling framework on three slow-moving landslides (Séchilienne, Viella, Villerville; France) spanning contrasting lithologies, deformation mechanisms and kinematics to demonstrate the scalability of the approach (**Figure 35 ; Table 11**). This combination of temporal depth, multi-sensor information, and cross-site variability provides a robust basis for training and evaluating transferable, physically consistent XAI-based framework.

V.2.2. Multi-site monitoring datasets

The Séchilienne landslide (western French Alps) affects a micaschist bedrock locally overlain by morainic deposits and includes a distinct, higher-acitivity unit dominated by toppling *(Vallet et al., 2016)* (**Figure 35a**). The Viella landslide (central French Pyrénées) affects morainic and colluvial deposits overlying calcareous schist bedrock, and develops multiple sliding surfaces and high spatial variability in kinematics *(Loke et al., 2022)* (**Figure 35b**). The Villerville landslide (French Normandy coast) affects chalk and sand layers with colluvial deposits overlying marl susbtrate, and is characterized by multiple rotational slip surfaces *(Lissak et al., 2014)* (**Figure 35c**).

Each site is instrumented with a multi-sensor monitoring network combining surface-displacement measurements (tacheometric targets, GNSS, or extensometers), groundwater level observations, and local meteorological records (precipitation, temperature). Data are acquired at daily resolution, yielding long, multivariate time series that capture several amplitudes of landslide responses to hydro-meteorological forcing.

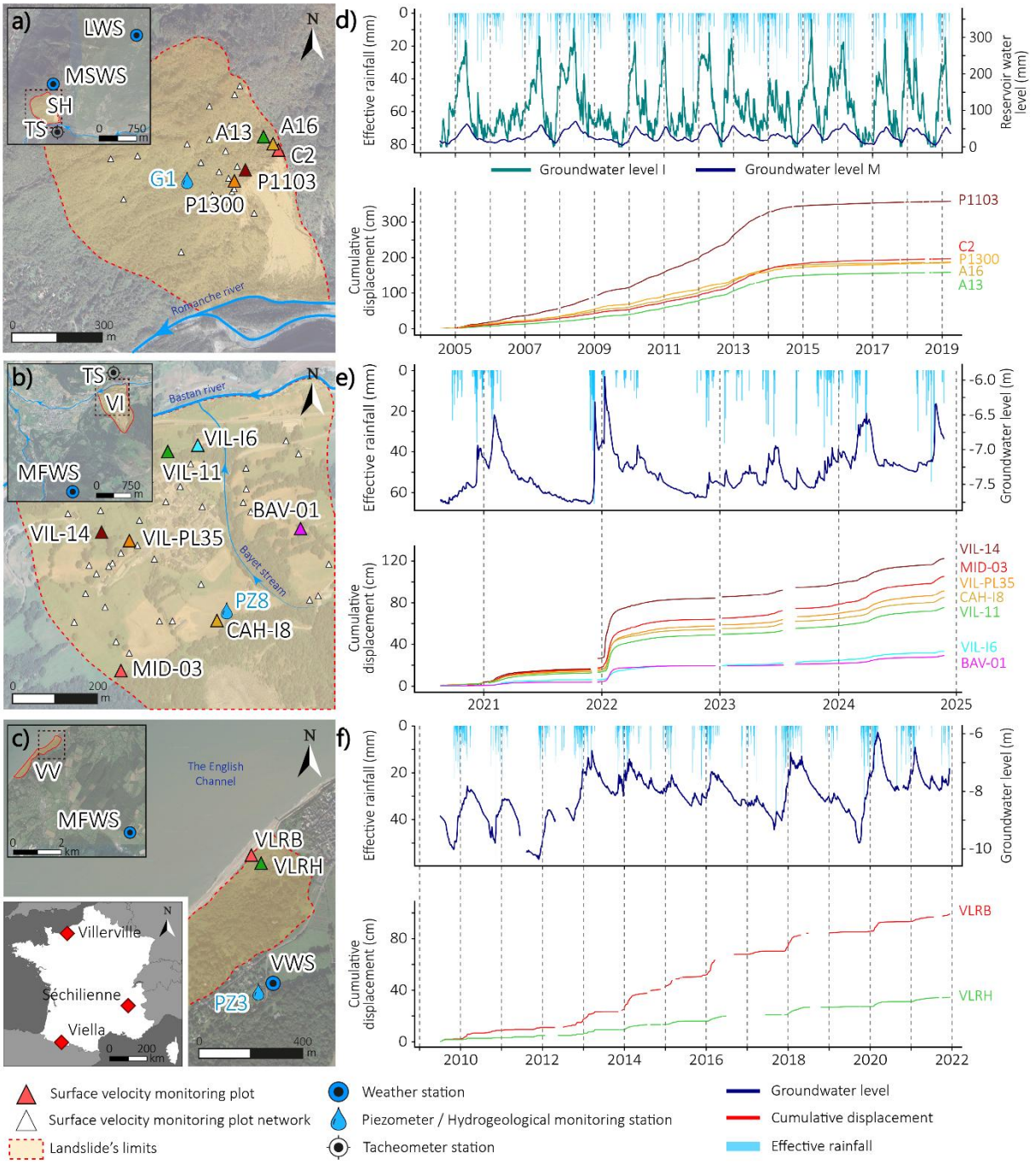


Figure 35 | Overview of the three landslide sites: **(a)** Séchilienne (SH), **(b)** Viella (VI), and **(c)** Villerville (VV). Insets show the spatial distribution of displacement monitoring plots, tacheometric station (TS) and meteorological stations, including Météo-France (MFWS), Mont Sec (MSWS), Luitel (LWS) and Villerville (VWS). **(d–f)** Cumulative displacement at each monitoring location, colour-coded by monitoring plot, shown together with the corresponding effective rainfall and groundwater level time series.

Monitoring plots are designed to ensure spatial representativeness and kinematic behaviours and provide diverse observational constraints. A subset of monitoring plots, defined using spatial zoning at Viella and dominant kinematic behaviour at Séchilienne and Villerville, is used to simulate landslide velocities. Five plots are analysed at Séchilienne, seven at Viella, and two at Villerville, with velocity records spanning from 4 to 15 years (**Table 11; Figure 35**). Meteorological observations are drawn from stations located close to each landslide. Nearby stations (Luitel for Séchilienne, Saint-Gatien-des-Bois for Villerville; **Figure 35a, c**) are used to ensure temporal completeness. Effective rainfall is estimated as precipitation minus potential evapotranspiration, with evapotranspiration computed using the Oudin formulation *(Oudin et al., 2005)*, well suited to data-scarce environments (see ***section V.2.3.2.1**.)*

Table 11 | Landslide geomorphology and observation network

		Séchilienne	**Viella**	**Villerville**
Geomorphological description	Geographical coordinates	45.06, 5.79	42.87, 0.01	49.40, 0.12
	Altitudes (m asl)	450 to 1100	782 to 1344	0 to 77
	Mean slope (°)	38	27	10
	Area (km^2)	1.00	0.58	0.47
	Lithology description	Micaschist bedrock with overlied morainic deposits	Morainic and colluvial debris overlying calcareous-schist bedrock	Chalk, sand, and colluvial debris overlying marl substrate
Monitoring network	Displacement stations	28	44	2
	Selected displacement stations	5	7	2
	Weather stations	45.07, 5.81 (MSWS) 45.09, 5.85 (LWS)	42.86, -0.02	49.39, 0.12 (VWS) 49.36, 0.16 (MFWS)
	Groundwater level stations	45.06, 5.80	42.87, 0.02	49.39, 0.12
Displacement serie	Studied period	2004/07/24 – 2019/03/31	2020/08/21 – 2024/12/17	2009/07/06 – 2021/12/11
	Frequency	Daily	Daily	Daily
	Missing data (%)	2 to 9	5 to 12	16 to 18

At each landslide, groundwater level observation capture site-specific hydrogeology within the landslide body. At Séchilienne, where direct observations are not available or incomplete (due to borehole shearing), groundwater levels are reconstructed using rainfall discharge models that conceptualise hillslope hydrology. Subsurface water discharge is monitored within an underground gallery, from which two synthetic groundwater proxies are derived *(Béjean-Maillard et al., 2025)*: a shallow, rapidly responding reservoir (GWLI) and a deeper, more inertial matrix-type reservoir (GWLM) (**Figure 35d**).

At Viella and Villerville, cumulative displacements display a step-like evolution, characterised by alternating phases of acceleration/deceleration (**Figures 35e, f**). Acceleration episodes commonly coincide with intense or sustained rainfall, often accompanied by progressive or rapid groundwater level variations. At Séchilienne, this step-like behaviour is less pronounced, as displacement is dominated by a long-term deformation trend that intensified until 2014 before decelerating, consistent with mechanically-driven processes *(Vallet et al., 2016)* (**Figure 35d**). At each site, landslide responses is strongly non-linear with displacement amplitudes only weakly related to the magnitude of the environmental forcing.

V.2.3. Explainable Artificial Intelligence framework

We developed a scalable XAI framework structured in six modules (M1-M6) (**Figure 36**). The framework decomposes landslide velocities into trend and transient components, simulates these components separately, and quantifies their responses to hydro-meteorological forcing using an XGBoost approach (1.000 decision trees). Landslide velocities are simulated using a combination of physics-motivated hydro-meteorological predictors that capture both the internal hydrogeological state of the slope and rapid infiltrated rainfall perturbations. In total, 248 predictors are constructed and grouped into ten categories (**Table 12**), enabling consistent application across landslide sites. The XAI framework simulates daily velocities over independent testing periods, quantifies predictor importance over the training period, and attributes predictor effects for each individual test prediction.

Table 12 | Hydro-meteorological predictors used in the XAI framework. Predictors are grouped by data source: net Rainfall (R), Effective Rainfall (ER), and GroundWater Level (GWL). W denotes sliding-window lengths, and L denotes time lags, in days. Feature names encode the data source, the extracted features, and the associated window and lag parameters.

Data source	Feature label	Description	Time windows and shift values
(E)R – (**E**ffective) **R**ainfall series	(E)R_*W1*	Cumulative (E)R over the previous W1 days	*W1* : [1, 2, 5, 10, 20, 30, 60, 90] *W2* : [3, 5, 10, 20, 30, 40, 60] *W3* : [30, 60, 90]
	(E)R_max*W1*	Maximum daily (E)R over the previous W1 days	
	(E)R_anomaly*W1*	Ratio of the mean (E)R over the previous W1 days to the long-term mean (E)R across rainy days	
	(E)R_sat*W2*/*W3*	Ratio of cumulative (E)R over the previous *W2* days to cumulative (E)R over the preceeding (*W3*-*W2*) days	
	(E)R_event_strenght	Ratio of the cumulative rainfall of the most recent i-day rainfall event to the mean cumulative rainfall of all i-day events.	
GWL – **G**round**W**ater **L**evel series	GWL_l*L*	*L*-day lagged GWL value	*L* : [0, 1, 5, 8, 20, 30] *W4* : [5, 10, 20, 30, 60, 90]
	GWL_l*L*_min*W4*	*L*-day lagged minimum GWL over the previous *W4* days	
	GWL_l*L*_max*W4*	*L*-day lagged maximum GWL over the previous *W4* days	
	GWL_l*L*_mean*W4*	*L*-day lagged mean GWL over the previous *W4* days	
	GWL_l*L*_diff*W4*	*L*-day lagged change in GWL over the previous *W4* days	
	GWL_l*S*_diff_Gmean	Difference between the *L*-day lagged GWL and the long-term mean GWL.	

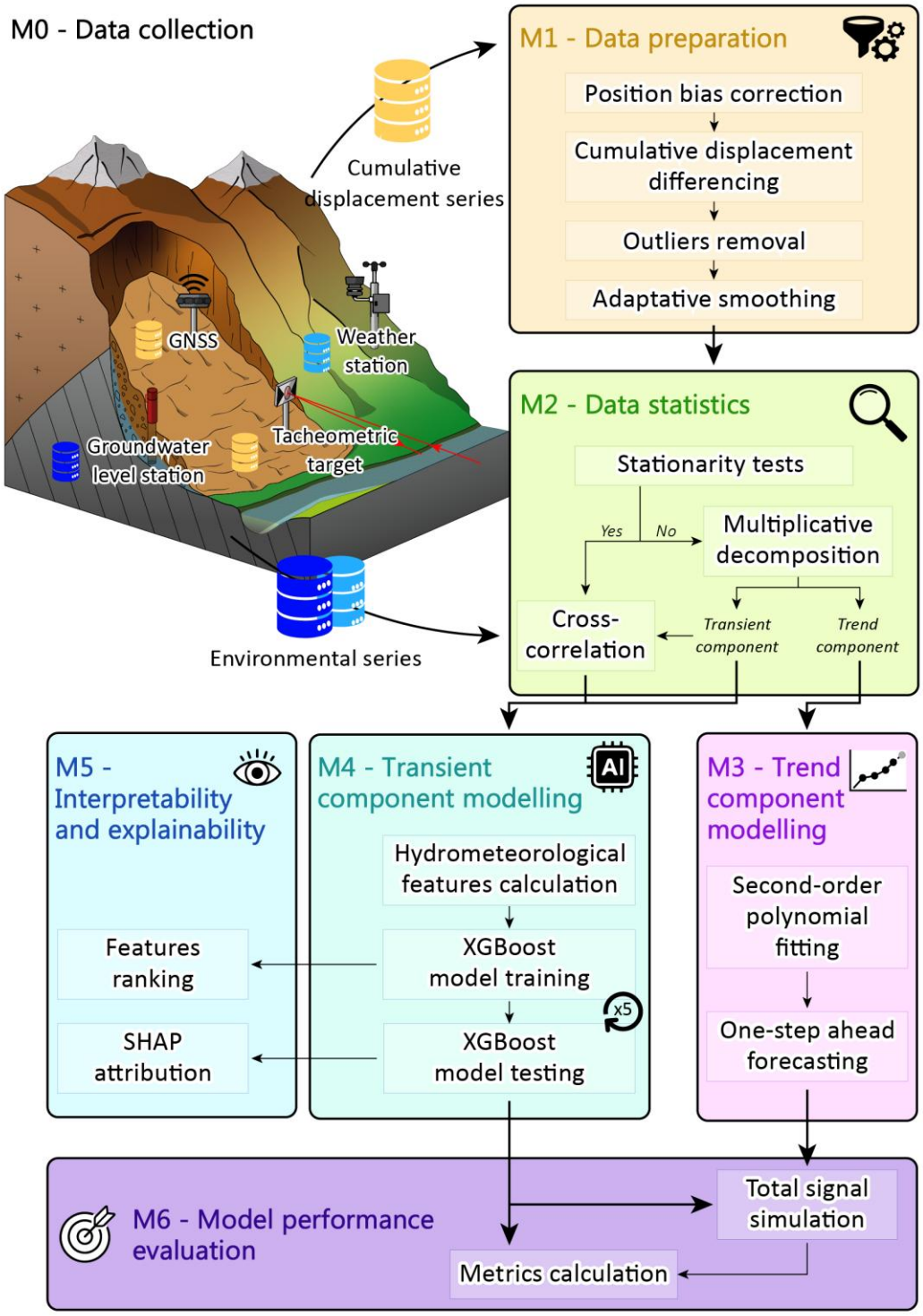


Figure 36 | XAI modelling framework. The modelling framework is organised in six modules. M1 pre-processes cumulative displacement time series. M2 assesses time series stationarity, isolate long-term trend and transient components, and quantifies correlations with environmental forcing. M3 forecasts the trend component. M4 performs features extraction and XGBoost training and testing. M5 quantifies predictor importance using features-importance ranking, and attributes predictor effects using SHAP analysis. M6 evaluates the model predictive performance over independent test periods.

V.2.3.1. Data collection

Surface velocity monitoring plots were firstly selected based on their observed sensitivity to hydro-meteorological forcing, identified by recurrent acceleration phases alternating with periods of low motion *(Lissak et al., 2014; Vallet et al., 2016*). Applying this criterion excluded much of the Séchilienne network (plots in white in **Figure 35**), where velocity is predominantly very-slow and continuous and shows limited direct response to external forcing, with the exception of the active southeastern sector. Secondly, where the monitoring plots density was sufficiently high (e.g., Viella; **Figure 35b**), we delineate displacement units using inverse-distance-weighted interpolation of cumulative displacements. A representative station was selected for each unit, prioritizing the location with the largest cumulative displacement and the most complete temporal coverage. At Viella, this approach reveals seven distinct kinematic units, with the highest cumulative displacements located in the southern and western sectors (see ***section V.2.7.*** ; **Figure 44**).

Meteorological forcing was characterized through Potential Evapo-Transpiration (PET) estimated using the Oudin formulation *(Oudin et al., 2005)* (*PE-Oudin library v0.3*), which derives extra-terrestrial radiation from mean daily air temperature, latitude and day of year. This parsimonious approach is well suited to data-limited settings and avoids reliance on poorly constrained atmospheric variables. Effective rainfall was computed using a bucket-type water balance scheme parameterized by site-specific surface lithology. In the absence of local soil measurements, available water capacity (AWC) was assigned from published reference values *(Le Bas, 2018)*, assuming a uniform rooting depth of 1 m. Maximum AWC was set to, respectively 100 mm, 50 mm, and 70 mm for the Séchilienne, Viella, and Villerville sites.

V.2.3.2. XAI modelling framework

The XAI modelling framework comprises six modules (M1-M6), whose functionalities are described below.

V.2.3.2.1. Module M1: Data preparation

Daily velocity time series (GNSS, tacheometric targets, extensometers) are affected by instrumental errors during low velocity periods. Such errors can obscure the displacement signal and markedly impair the performance and interpretability of the XAI model. To mitigate these effects, *Module M1* implements a sequential pre-processing workflow that combines bias correction, outlier detection, and adaptative smoothing. First, systematic position biases are corrected by fitting a linear regression to the full position record at each station. The inferred trend defines the main displacement axis, onto which all observations are orthogonally projected. Cumulative displacement is then calculated by accumulating daily displacements computed between consecutive projected positions using great-

circle distances (Haversine formula; *haversine library v2.9.0*). Because cumulative displacements are physically constrained to be non-decreasing, spurious decreases are corrected using a forward-filling procedure that enforces monotonicity (i.e. each invalid value is replaced with the last increasing observation). Daily velocity time series are subsequently derived by differencing the corrected cumulative displacement.

Table 13 | Pre-processing parameter values used in Module M1. Values are reported for each monitoring plot. Parameters are grouped by processing step: Outlier Detection (OD), Adaptive Smoothing (AS), and interpolation. Séchilienne time series do not include AS_Wmax values, as the long-term trend affects the median-based standard-deviation classification used to assign short or long smoothing windows.

		Outlier Detection (OD)		Adaptative Smoothing (AS)		Missing values interpolation	
Landslide sites	**Velocity plots**	**OD_W**	**OD_z**	**AS_Wmax**	**AS_Wmin**	**Interp_lim**	**Wsmooth**
Séchilienne	A13	10	1.5	-	7	10	3
	A16	10	1.5	-	7	10	3
	C2	10	1.5	-	7	10	3
	P1300	10	1.5	-	7	10	5
	P1103	10	1.5	-	7	10	5
Viella	VIL-14	10	1.5	60	3	5	3
	MID-03	10	1.5	60	5	5	3
	VIL-PL35	10	1.5	60	3	5	3
	CAH-I8	10	1.5	60	3	5	3
	VIL-11	10	1.5	60	3	5	3
	VIL-16	10	1.5	60	3	5	3
	BAV-01	10	1.5	60	3	5	3
Villerville	VLRB	15	1.0	60	7	5	5
	VLRV	15	1.0	60	7	5	5

Second, outliers in the velocity series are identified and removed using a sliding-window z-score filter based on mean (μ) and standard-deviation (σ) computed over a window size (*OD_W*), with a z-factor (*OD_z*) adapted to data quality (**Equation 17 ; Table 13**). Smaller *OD_z* values and larger *OD_W* windows are used when residual noise remains high after position correction, requiring stricter filtering (e.g., Villerville). Remaining noise is reduced using an adaptive smoothing scheme designed to preserve velocity variations associated with acceleration–deceleration phases. Velocity values are first classified as low-activity or active periods based on the median of the velocity standard-deviation distribution. Smoothing window lengths are then adjusted accordingly, with longer windows during stable periods (*AS_Wmax*) and shorter windows during active periods (*AS_Wmin*). *AS_Wmin* typically ranges from 3–

7 days and is set according to series quality (**Table 13**). At Séchilienne, velocities are smoothed using the short-window parameter only, as the long-term trend component obscure the step-like evolution that affects standard-deviation distribution. Finally, missing data gaps arising from acquisition issues or outlier removal are reconstructed by linear interpolation. A final short-window smoothing pass (3–5 days) is applied to reduce residual noise.

$$Thresholds_W = \mu_W \pm z \times \sigma_W \quad (17)$$

V.2.3.2.2. Module M2: Data statistics

Module M2 integrates statistical tools applied to the velocity and predictor time series. It aggregates three components: velocity series stationarity assessment and decomposition, and velocity-environmental series cross-correlation analysis.

Stationarity is a critical prerequisite for robust XGBoost-based model training, as non-stationary behavior can induce scale inconsistencies between the training and testing datasets possibly degrading predictive skill. The module integrates the Augmented Dickey-Fuller *(Dickey & Fuller, 1979)* (ADF; *statsmodels adfuller library, v0.15.0*) test and the Kwiatkowski-Phillips-Schmidt-Shin *(Kwiatkowski et al., 1992)* (KPSS; *statsmodels kpss library, v0.15.0*) test to leverage their complementary null hypothesis. Test trend parameters are set either as constant (Viella and Villerville) or deterministic (Séchilienne), considering velocity fluctuations around the trend component. Stationarity is verified for the Viella and Villerville velocity series. However, because tests assume linear trends, the Séchilienne series do not satisfy both criteria when applied to the full record (**Table 14**). To adress this, the module integrates the Bai-Perron breakpoint detection method *(J. Bai & Perron, 2003)* (*ruptures library, v1.1.10*) to identify periods of non-linear evolution, and segment the time series into four periods of quasi-linear behaviour (plot A13) (**Figure 37**). Stationarity tests are applied to each segment independently. ADF and KPSS tests yield consistent results for the main active phase (2009–2015). Minor discrepancies persist in the earliest and latest segments, likely because KPSS is more sensitive to small variance shifts or residual trends (**Table 15**).

Non-stationary series (Séchilienne) are further handled using statistical multiplicative decomposition. Velocity time series (y_t) are expressed as the product of a trend component (T_t), a seasonal component (S_t), and an irregular component (I_t), assuming an annual (365 days) periodicity (**Equation 18**). Stationarity of the transient component (i.e., the product of seasonal and irregular components) is subsequently verified using the ADF and KPSS tests.

$$y_t = T_t \times S_t \times I_t \quad (18)$$

Table 14 | ADF and KPSS stationarity test results. Both tests are evaluated using p-values with a significance level of $\alpha = 0.05$. Rejection of the ADF null hypothesis (p-value < α) indicates that the series is stationary. Rejection of the KPSS null hypothesis (p-value < α) indicates that the series is non-stationary. Differences between ADF and KPSS outcomes reflect their complementary null hypotheses and can arise from near-threshold behaviour, weak or non-linear trends, or changes in variance.

		ADF test		KPSS test	
Landslide site	**Velocity plots**	**p-value**	**Stationarity**	**p-value**	**Stationarity**
Séchilienne	A13	$2.30\ 10^{-2}$	Yes	$1.0\ 10^{-2}$	No
	A16	$9.00\ 10^{-2}$	Yes	$1.0\ 10^{-2}$	No
	C2	$1.00\ 10^{-2}$	Yes	$1.0\ 10^{-2}$	No
	P1300	$1.00\ 10^{-3}$	Yes	$1.0\ 10^{-2}$	No
	P1103	$3.30\ 10^{-2}$	Yes	$1.0\ 10^{-2}$	No
Viella	VIL-14	$1.51\ 10^{-6}$	Yes	$1.0\ 10^{-1}$	Yes
	MID-03	$4.21\ 10^{-5}$	Yes	$1.0\ 10^{-1}$	Yes
	VIL-PL35	$3.00\ 10^{-6}$	Yes	$1.0\ 10^{-1}$	Yes
	CAH-I8	$9.60\ 10^{-7}$	Yes	$1.0\ 10^{-1}$	Yes
	VIL-11	$3.24\ 10^{-6}$	Yes	$1.0\ 10^{-1}$	Yes
	VIL-16	$3.42\ 10^{-5}$	Yes	$1.0\ 10^{-1}$	Yes
	BAV-01	$1.15\ 10^{-5}$	Yes	$1.0\ 10^{-1}$	Yes
Villerville	VLRB	$7.36\ 10^{-10}$	Yes	$1.0\ 10^{-1}$	Yes
	VLRV	$3.05\ 10^{-9}$	Yes	$1.0\ 10^{-1}$	Yes

Table 15 | ADF and KPSS stationarity test results across the A13 plot periods. Series segmentation is based on the Bai–Perron breakpoint detection method (**Figure 37**).

			ADF test		KPSS test	
Velocity plot	**Segment**	**Period**	**p-value**	**Stationarity**	**p-value**	**Stationarity**
A13	1	2004/07/24 – 2009/12/25	$6.52\ 10^{-6}$	Yes	$1.00\ 10^{-2}$	No
	2	2009/12/26 – 2014/03/09	$7.00\ 10^{-2}$	No	$3.00\ 10^{-2}$	No
	3	2014/03/10 – 2015/04/13	$3.30\ 10^{-1}$	No	$1.00\ 10^{-2}$	No
	4	2014/04/14 – 2019/03/31	$1.00\ 10^{-2}$	Yes	$1.00\ 10^{-2}$	No

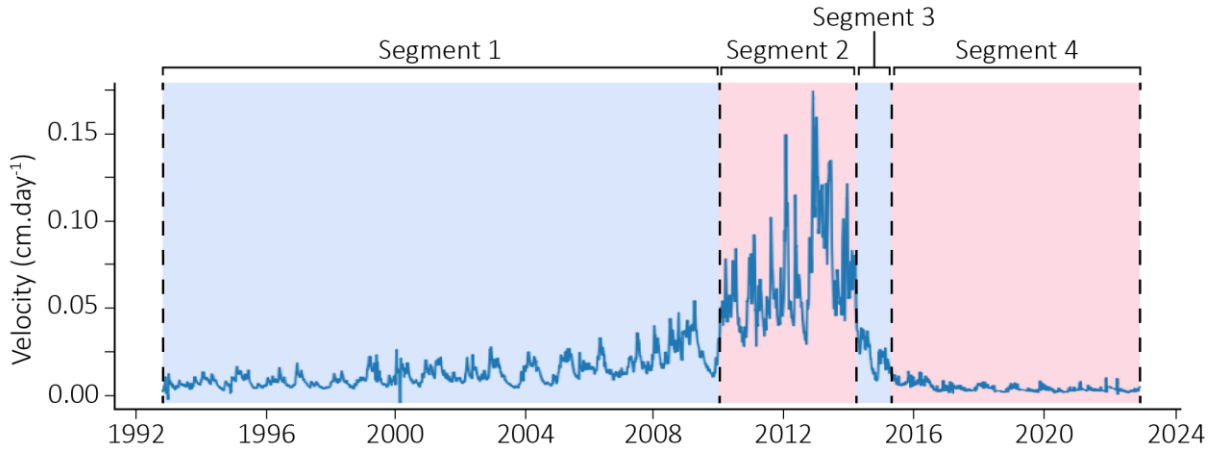


Figure 37 | Segmentation of the Séchilienne velocity record (plot A13) using Bai–Perron breakpoint detection. The Bai–Perron method is applied with three breakpoints to partition the time series into four quasi-linear periods consistent with changes in series evolution. Identified periods are shown using alternating red and blue shading.

Before calculating the XGBoost training features, cross-correlation techniques are applied to the stationary velocity and groundwater level time series. The analysis is used to identify the temporal lag that maximizes coupling between hydrogeological forcing and the landslide kinematic response. According to the landslide, optimal lags vary between 1 and 30 days, and multiple lag values within this range are used to calculate the site-specific training features.

V.2.3.2.3. Module M3: Trend component modelling

Module M3 integrates tools to model the long-term landslide velocity trend which captures the progressive and irreversible deformation of the landslides induced by changes in their mechanical states. Because this trend is largely decoupled from short-term hydro-meteorological forcing, *Module M3* uses a one-step-ahead forecasting method *(Chang et al., 2025)* to simulate this deformation with a 2nd-order polynomial fitted within a 30-day sliding window.

V.2.3.2.4. Module M4: Transient component modelling

To enrich the XAI-model input space beyond the raw hydro-meteorological forcing time series, *Module M4* calculates a generic set of physics-informed features and associates them with the daily landslide velocity observations. The features are designed to represent three complementary aspects of landslide water-related forcing: (i) the hydrological state of the system (e.g., saturation versus dryness), (ii) hydrological memory effects (e.g., cumulative rainfall, trends and extrema in groundwater levels), and (iii) short-term hydrological transient processes (e.g., intense rainfall events, rapid groundwater rises/lowering). Based on three forcing series (i.e., net rainfall, effective rainfall and groundwater level), *Module M4* generates 248 predictors grouped into ten categories (**Table 12**). Features are computed

over multiple time windows (spanning 1 to 90 days) enabling the XAI-model to resolve short-, intermediate-, and long-term landslide hydromechanical responses. To account for delayed kinematic responses, groundwater-level features are computed with an additional time lag defined from the cross-correlation analysis implemented in *Module M2*.

XGBoost is a supervised ensemble-learning framework that iteratively constructs decision trees to minimize prediction error *(T. Chen & Guestrin, 2016)*. XGBoost is used to simulate daily velocities from the feature sets. Model training and testing is performed using a blocked time-series cross-validation strategy designed to sample contrasting meteorological and hydrological periods. The landslide time series is partitioned into five contiguous periods (labelled 1-5), each comprising 20% of the data. Periods are sequentially withheld for testing, while the remaining data is used for training. Although this scheme allows limited temporal overlap between training and testing periods for the early periods (1 to 4), the duration of each test window constrains information leakage to the period boundaries, thereby minimizing its impact on velocity simulations.

V.2.3.2.5. Module M5: Interpretability and explainability

Module M5 is designed to assess model interpretability, both in terms of features contribution and decision rules. Variable importance is assessed using the Gini impurity-based metric *(Breiman, 2001; Louppe et al., 2013)*, which measures the mean reduction in node variance attributable to each predictor across the ensemble of decision trees. While this metric captures how frequently and effectively variables are ranked in the tree splitting, it does not evaluate probability quality and integrates performance over all possible thresholds, many of which are irrelevant in practice.

Therefore, *Module M5*, integrates SHAP (SHapley Additive exPlanations ; *Lundberg & Lee, 2017*) metric which quantifies the marginal contribution of each feature by averaging its effect over all possible feature orderings, yielding consistent, locally accurate, and directionally interpretable explanations. *Module M5* computes feature attribution with the TreeSHAP algorithm (*shap library* ; *Lundberg et al., 2019*). The median absolute SHAP values, and their directional effects on the predicted velocities are visualised using beeswarm plots.

V.2.3.2.6. Module M6: Model performance evaluation

Module M6 is designed to evaluate model performance over the test periods using the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) metrics. Metrics are computed from the XAI simulation outputs. When the original series is decomposed into trend and transient components, the multiplicative decomposition is inverted to simulate the total velocity series, and metrics are computed for both the simulated transient component and total series.

V.2.4. Model performance: simulated landslide velocities

The XAI framework accurately reproduces landslide velocity time series across sites and testing periods, capturing both landslide long-term displacement trends and the timing of transient acceleration/deceleration episodes. Simulated velocities (**Figures 38a, d, g**) closely follow observed dynamics at representative plots from A13 (Séchilienne), VIL-PL35 (Viella), and VLRB (Villerville). Consistent performance is observed across all the monitoring network (see ***section V.2.7.*** ; **Figures 40-43**). Residual errors are small relative to the amplitude of observed velocity variations. Root-Mean-Square (RMSE) and Mean Absolute (MAE) Errors range from 0.005-0.015 $cm.d^{-1}$ at Séchilienne, 0.01-0.035 $cm.d^{-1}$ at Viella, and 0.02-0.06 $cm.d^{-1}$ at Villerville, indicating robust predictive skill across contrasting kinematic regimes. Model misfits are primarily associated with peak velocity amplitudes, which are intermittently under- or overestimated.

These discrepancies are most pronounced at Villerville (**Figure 38g**), where the maximum simulated velocities (~0.23 $cm.d^{-1}$), are lower than the observed maxima (0.30-0.45 cm. d^{-1}). Other mismatches occur during short acceleration episodes (early 2015, early 2019, and early 2021) with velocities either overestimated or not captured. Three mechanisms likely contribute to these discrepancies. First, velocity at Villerville exhibits rapid, high-amplitude fluctuations, for which small differences in predictor combinations can translate into disproportionate changes in simulated velocity amplitudes, leading to local misestimation of velocity peaks. Second, progressive structural evolution within the landslide material (fracture opening and closure, changes in water pathways, changes in geomechanical properties) can modify pore-water-pressure distributions and hydrological sensitivity through time, inducing behavioural shifts not captured by the trained XAI model. Third, the coupling between hydro(meteorological forcing and velocity may involve event-scale dynamics that are insufficiently resolved by the current predictor set, limiting the model's ability to reproduce isolated or transient acceleration episodes within the ensemble of decision trees.

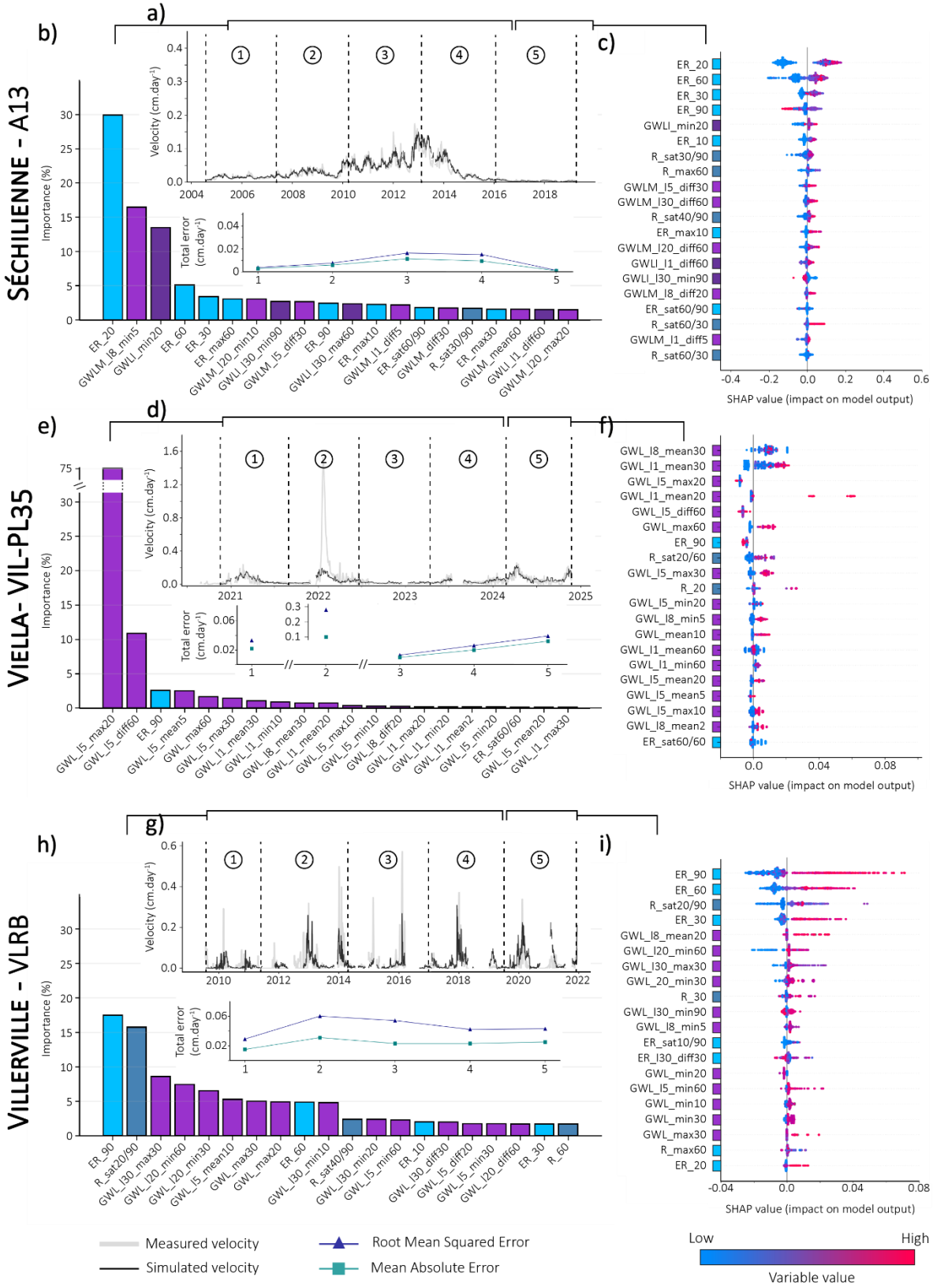


Figure 38 | Simulated landslide velocities, interpretability and explainability of the predictors (feature importance, SHAP attribution). **(a-c)** Séchilienne (plot A13), **(d-f)** Viella (plot VIL-PL35), and **(g-i)** Villerville (plot VLRB). Observed and simulated daily velocities are shown in grey and black,

respectively **(a, d, g)**, with predictive performance over test periods (periods 1-5) quantified using RMSE (triangle) and MAE (square) metrics. Feature importance (left) and SHAP-attribution (right) highlight the dominant hydrometeorological controls on model behaviour **(b, e, h)** and their influence on predicted velocities **(c, f, i)** across the training and testing periods, as indicated by the horizontal bars at the top of each panel. For clarity, only the twenty most influential predictors are shown, capturing the majority of explained variance (see ***section V.2.5.2***). In the SHAP plots, each point represents a single testing simulation, coloured by the corresponding predictor value. Predictors are colour-coded by source: effective rainfall (sky blue), net rainfall (steel blue), and groundwater level (magenta). At Séchilienne, two groundwater level series are shown, representing shallow and deep hydrological reservoirs (*Béjean-Maillard et al., 2025*). Simulated velocities correspond to the simulated total series (trend × transient); transient-component simulations are reported in the Extended Data (see ***section V.2.7.*** ; **Figure 45**). At Viella, the leading predictor exceeds 70% of the total importance score; accordingly, an axis break is used to retain a comparable scale across panels.

Another discrepancy is observed for the early 2022 period at Viella (**Figure 38d**) marked by an extreme acceleration that is not fully reproduced by the XAI model. Observed velocities peak at ~1.4 $cm.d^{-1}$, far exceeding those of other acceleration phases (~0.3 $cm.d^{-1}$) represented in the training data. While the onset of this event is captured, its amplitude is underestimated, reflecting that predictions are constrained by the range of behaviours observed in the time series and thus sampled during model training. The early 2022 event is nevertheless retained in the training dataset to expose the model to extreme dynamics and to support discrimination between the drivers of exceptional accelerations and those governing more moderate displacement episodes.

V.2.5. Interpretability and explainability: environmental predictor controls

V.2.5.1. Controls of hydro-meteorological predictors on landslide dynamics

Figure 38 illustrates the control of environmental predictors to the XAI model and their influence on predicted velocities for the three landslide sites. Feature importance and SHAP-based attribution reveal how environmental forcings structure the learned relationships (**Figures 38b, e, h**) and modulate velocity simulations (**Figures 38c, f, i**). Importance and attribution are derived from models trained on periods 1–4 and evaluated on an independent final interval (period 5), preserving temporal continuity and preventing information leakage that could bias interpretation. For clarity, only the twenty most influential predictors are shown, capturing the dominant share of explained variance while maintaining interpretability (see ***section V.2.5.2.***).

Each landslide displays a distinct hierarchy of controlling predictors, reflecting contrasts in lithology, hydrological response and landslide deformation style. At Séchilienne (plot A13), velocity is primarily governed by medium-term effective rainfall, with cumulative forcing over 20-90 days exerting the strongest influence on both model structure and predicted acceleration magnitudes (**Figures 38b, c**). Predictors operating over 20-60 days dominate, indicating peak sensitivity to hydro-meteorological forcing integrated over several weeks. SHAP attribution reveals a positive relationship between medium-term effective cumulative rainfall and velocity, whereas longer integration windows contribute negatively, reflecting post-rainfall event deceleration phases. Secondary predictors, notably R_sat60/90, capture persistently wet antecedent conditions that precondition acceleration, highlighting a state-dependent control (**Figure 38c**). Together, these patterns point a mid-term hydro-meteorological control that integrates both antecedent hydrological conditions and recent rainfall forcing. These patterns are consistent with deformation mechanisms governed by a fracture-modulated hydrological response within the weathered micaschist bedrock.

At Viella (plot VIL-PL35), model attribution indicates that groundwater evolution dominates landslide velocity dynamic (**Figures 38e, f**). Feature importance ranking and SHAP analysis consistently identify groundwaterlevel descriptors as the primary controls, reflecting the strong hydromechanical coupling at this site. Interpretation is nevertheless complicated by the uneven temporal distribution of active phases and the presence of an extreme acceleration event in the training period. Variance reduction is largely driven by a short-lag groundwater feature (i.e., GWL_lag5_max20) that isolates the January 2022 acceleration from the otherwise active periods. Lower-ranked predictors differentiate minor acceleration episodes (~0.2 $cm.d^{-1}$) from predominantly very low and constant velocity periods (<0.1 $cm.d^{-1}$), explaining their limited contribution to overall variance. SHAP attribution (**Figure 38f**) reveals strongly non-linear groundwater effects. Mean groundwater levels

computed over 20-30 days contribute negligibly across most of their range, but shift to positive contributions once a critical level is exceeded, consistent with a threshold-type response. Predictors computed with 5-8 days lags exert weaker influence and primarily refine the temporal context of recent groundwater fluctuations rather than sustained trends. Several influential predictors exhibit net negative SHAP contributions (e.g., ER_90). These do not indicate stabilisation but instead reflect post-event deceleration, whereby elevated groundwater levels or rainfall are consistently associated with reduced velocities following major acceleration episodes. Together, these patterns highlight a landslide behaviour dominated by short-term groundwater fluctuations and threshold-driven responses, punctuated by extreme events that locally constrain model interpretability.

At Villerville (plot VLRB), rainfall and groundwater exert complementary controls on landslide velocity (**Figures 38h, i**). Model features ranking (**Figure 38h**) identifies long-term effective rainfall and soil saturation metrics as the dominant contributors, with secondary influence from slope hydrogeology descriptors. Most influential predictors integrate forcing over 60-90 days with groundwater lags of 20-30 days, indicating a delayed and cumulative hydrological control on landslide velocity. SHAP analysis reveals a strong positive response to medium- to long-term cumulative rainfall (60-90 days) (**Figure 38i**), consistent with progressive pore-water pressures build-up in a low-permeability (clay-rich) landslide material *(Lissak et al., 2014)*. The R_sat20/90 variable, capturing the transition from dry to wet conditions in the sub-surface, emerges as a key modulator: it dampens system sensitivity to rainfall events following prolonged dry periods while exerting limited influence during sustained wet or dry regimes, highlighting the role of antecedent saturation as a prerequisite for triggering a velocity response. Other groundwater levels predictors (e.g., GWL_lag8_mean20, GWL_lag0_max30) refine the temporal structure of the response, contributing episodically during extreme hydro-meteorological events. Together, these patterns point to an inertial, state-dependent hydro-mechanical behaviour governing landslide motion at Villerville.

V.2.5.2. Model sensitivity to environmentals predictors

We assessed model sensitivity to predictor dimensionality by evaluating predictive errors over the final test period (period 5) using progressively larger subsets of ranked feature, from the top 5 to the full features set (**Figure 39**).

At Séchilienne (plot A13) and Viella (plot VIL-PL35) (**Figures 39a, b**), predictive skill improves monotonically as additional predictors are included, with MAE and RMSE decreasing and stabilising once the dominant fraction of variance is captured. Performance (corresponding to nearly 99% of the explained model variance) stabilizes at ~200 predictors for A13 and ~100 for VIL-PL35, beyond which further features yield negligible gains. For these two landslides, reducing the number of predictors lowers computational cost and slightly enhances model interpretability.

Villerville (plot VLRB) exhibits a contrasting behaviour (**Figure 39c**). Although prediction errors stabilize beyond ~30 predictors, the lowest residuals are obtained using only the top 5-10 features. This sensitivity reflects the landslide abrupt acceleration/deceleration dynamics and its strong inertia to hydrogeological processes. For this landslide, reduced feature dimensionality yields the most stable and physically consistent predictions.

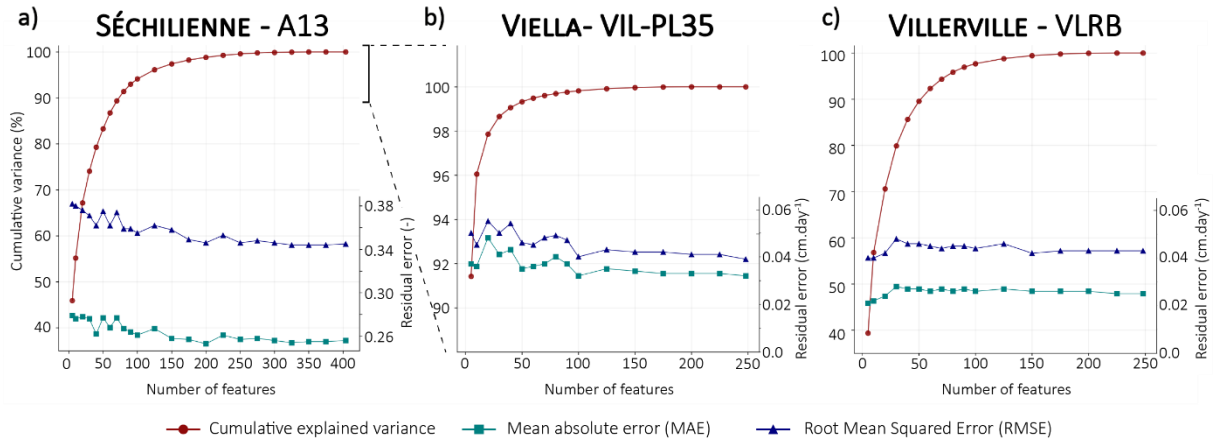


Figure 39 | Model sensitivity to increasing subsets of ranked predictors. **(a)** Séchilienne (plot A13), **(b)** Viella (plot VIL-PL35), and **(c)** Villerville (plot VLRB). Explained variance and error metrics (RMSE, MAE) are evaluated over the final test period (period 5). For Séchilienne, prediction errors are computed for the dimensionless transient-component forecast (i.e., the direct output of the XAI model), as the reconstructed total series (trend × transient) exhibits near-zero velocities, yielding low-contrast residuals. Séchilienne transient forecasts are reported in the Extended Data (see ***section V.2.7.*** ; **Figure 45**). For Viella, explained cumulative variance increases rapidly; the y-axis is therefore zoomed to the 90–100% range.

V.2.5.3. Model transferability to other monitoring plots

Application of the XAI framework to additional monitoring plots (see ***section V.2.7.***) reveals plot-specific feature rankings and SHAP attribution patterns. These differences arise from contrasting time series characteristics (amplitudes, timing and duration of acceleration/deceleration episodes) that shape model learning (i.e., the prioritization of predictors during the tree growth and splitting), differences in data quality and signal-to-noise ratio that modulates feature-velocity relationships, and intrinsic local slope heterogeneity (lithology, fracturing, pore-water pressure dynamics). Together, these factors underscore the site-specific expression of hydro-meteorological forcing on landslide motion and highlight the need for modelling approaches that accommodate both common hydro-meteorological controls and local variability.

At Séchilienne, effective rainfall consistently emerges as the dominant control on displacement, although the characteristic response time varies across the slope (see ***section V.2.7.*** ; **Figure 40**). Monitoring plots (P1300, P1103) located west of the main active zone (**Figure 35a**) exhibit strongest sensitivity to longer cumulative rainfall windows (60-90 days) indicating pronounced hydrological memory and delayed pore-water pressure transmission within the fractured micaschist bedrock. In contrast, monitoring plots A13, A16 and C2, installed across open fractures respond preferentially to shorter cumulative rainfall (20-30 days), reflecting more direct hydraulic coupling to the shallow fracture networks.

Groundwater level predictors display clear spatially dependent influences. At Viella, time-lagged groundwater responses are most informative, with stations (i.e., VIL-16, BAV-01; **Figure 35b**) located farther from the representative borehole exhibiting longer effective delays (see ***section V.2.7.*** ; **Figures 41-42**). This pattern is consistent with their peripheral positions within the landslide body and reflects spatial variability in subsurface hydrogeological connectivity.

At Villerville, spatial variability in velocity amplitudes combined with short-term acceleration/deceleration phases, modulate predictor ranking. At VLRH monitoring plot (**Figure 35c**), groundwater level emerges as the dominant control, indicating a direct forcing–response relationships (see ***section V.2.7. ;*** **Figure 43**). Across stations, however, long-term hydrology influences remain consistently important, highlighting the role of hydrological memory in governing landslide dynamics.

V.2.6. Conclusion: Implications for interpreting landslide dynamics with XAI framework

Our results demonstrate the robustness and modularity of the proposed scalable XAI framework, which reproduces daily landslide velocities across three contrasting (lithology, deformation mechanisms, velocity amplitude) landslides. Using a single, generic set of physically grounded hydro-meteorological predictors (without site-specific calibration) the model captures key deformation dynamics and autonomously identifies the dominant hydro-meteorological controls and their characteristic timescales.

The identified predictors per landslide align with contrasting physical processes, including delayed pore-water pressure build-up driven by slow matrix infiltration in impermeable slope material (Villerville, **Figures 38h, i**) and rapid responses to rainfall in more permeable (Viella, **Figures 38e-f**) or fractured (Séchilienne, **Figures 38b, c**) slope materials. Together, these results show that XAI frameworks can recover site-specific landslide behaviour while preserving physical interpretability across diverse settings.

However, the XAI model does not systematically reproduce the full amplitude, with occasional underestimation and rarer overestimation of velocities. These discrepancies reflect limitations in training data representativeness (uneven sampling of acceleration/deceleration phases; Viella, **Figure 38d**), abrupt short-term regime shifts that reduce the separability of velocity values from environmental predictors (Villerville, **Figure 38g**), and constraints in the learned predictor velocity relationships, either because the generic feature set does not fully encode site-specific processes or because rare high-impact predictors receive limited importance during training (Villerville). Additional limitations may arise from progressive structural evolution within the landslide material, which can bias attribution and complicate long-term learning.

Despite these limitations, our results demonstrate that XAI models is transferable using a scalable set of physically grounded predictors reproducing across a range of landslide velocity regimes and forcing–response relationships. The framework allows analysing multi-sensor monitoring time series and simulating velocity (timing and amplitude of acceleration/deceleration episodes, long-term trend and short-term responses) across long datasets (spanning 4-15 years), and identify first-order hydro-meteorological controls through features ranking and attribution-based interpretation. To our knowledge, this constitutes one of the first applications of a XAI workflow across multiple distinct landslide case studies using the same model structure and non-site-specific predictors.

V.2.6.1. Acknowledgments

Authors acknowledge support from the French Landslide Observatory OMIV / Observatoire Multidisciplinaire des Instabilités de Versants (ano-omiv.cnrs.fr) in supporting long-term field instrumentation and data dissemination on the three studied landslides. O. Béjean-Mailard and C. Bertrand are members of OSU THETA, and J.-P. Malet is member of OSU EOST, which both contributed to support on-site data collection.

V.2.6.2. Data availability

The multi-parameter landslide dataset on which this study grounds is available on the EasyData FAIR repository under DOIs: Séchilienne, https://doi.org/10.57932/d75b152e-b617-4032-a48d-ed4d3e9410be; Viella, https://doi.org/10.57932/1ac1fcc4-fcbd-4bcf-a6e9-aeb2f749b1dd; and Villerville, https://doi.org/10.57932/f238a562-8741-4aee-9eaf-1443faa22cad.

V.2.6.3. Code availability

The code used to compute the hydro-meteorological predictor set and to implement the XGBoost modelling workflow (including feature importance ranking and SHAP analysis) is available from the certified IN2P3 GitLab repository theia/theia-landslide-forecaster (https://gitlab.in2p3.fr/theia/theia-landslide-forecaster).

V.2.6.4. Declarations

V.2.6.4.1. Funding

The French National Landslides Observatory (SNO OMIV) and the Observatory of Earth Sciences (OSU-THETA) from the National Centre for Scientific Research (CNRS) funded the installation of the monitoring system and the data used in this study.

V.2.6.4.2. Conflict of Interest

The authors declare no competing interests.

V.2.7. Extended data

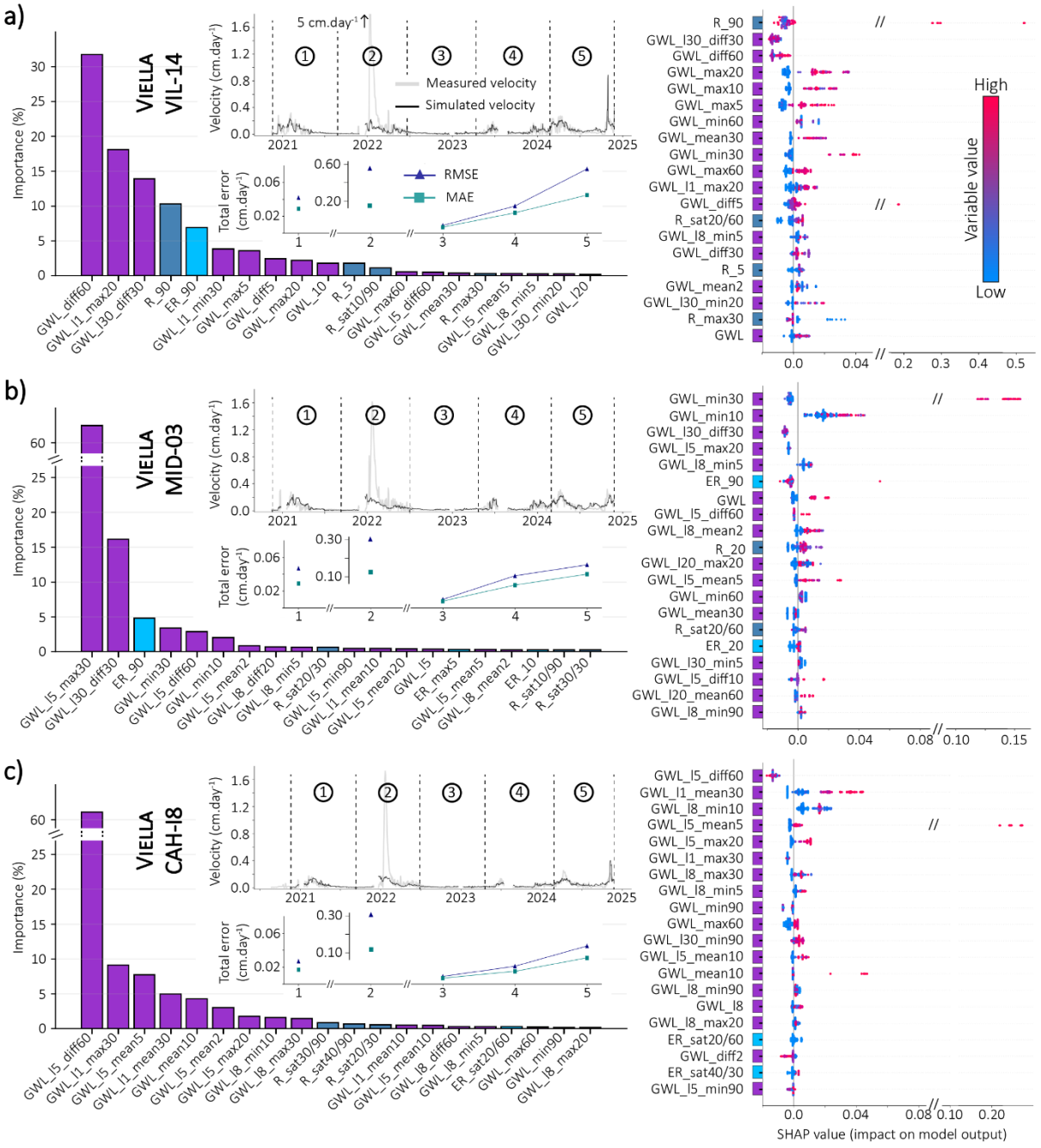


Figure 40 | Simulated landslide velocities, interpretability and explainability of the predictors (feature importance, SHAP attribution) for additional Viella monitoring plots (part 1). **(a–c)** Observed and simulated daily velocities are shown in grey and black, respectively, for **(a)** VIL-14, **(b)** MID-03, and **(c)** CAH-I8. Legends are provided in panel **(a)**. Feature importance (left) and SHAP attribution (right) highlight the dominant hydrometeorological controls and their influence on predicted velocities across the training and testing periods, as indicated by the horizontal bars at the top of each panel. Predictors are colour-coded as in **Figure 38**: effective rainfall (sky blue), net rainfall (steel blue), and groundwater level (magenta). Axis breaks (double bars) are used for feature importance and SHAP value axes to enable comparison across plots while accommodating large values. The January 2022 velocity peak at VIL-14 (~5 cm d^{-1}) is indicated by an arrow.

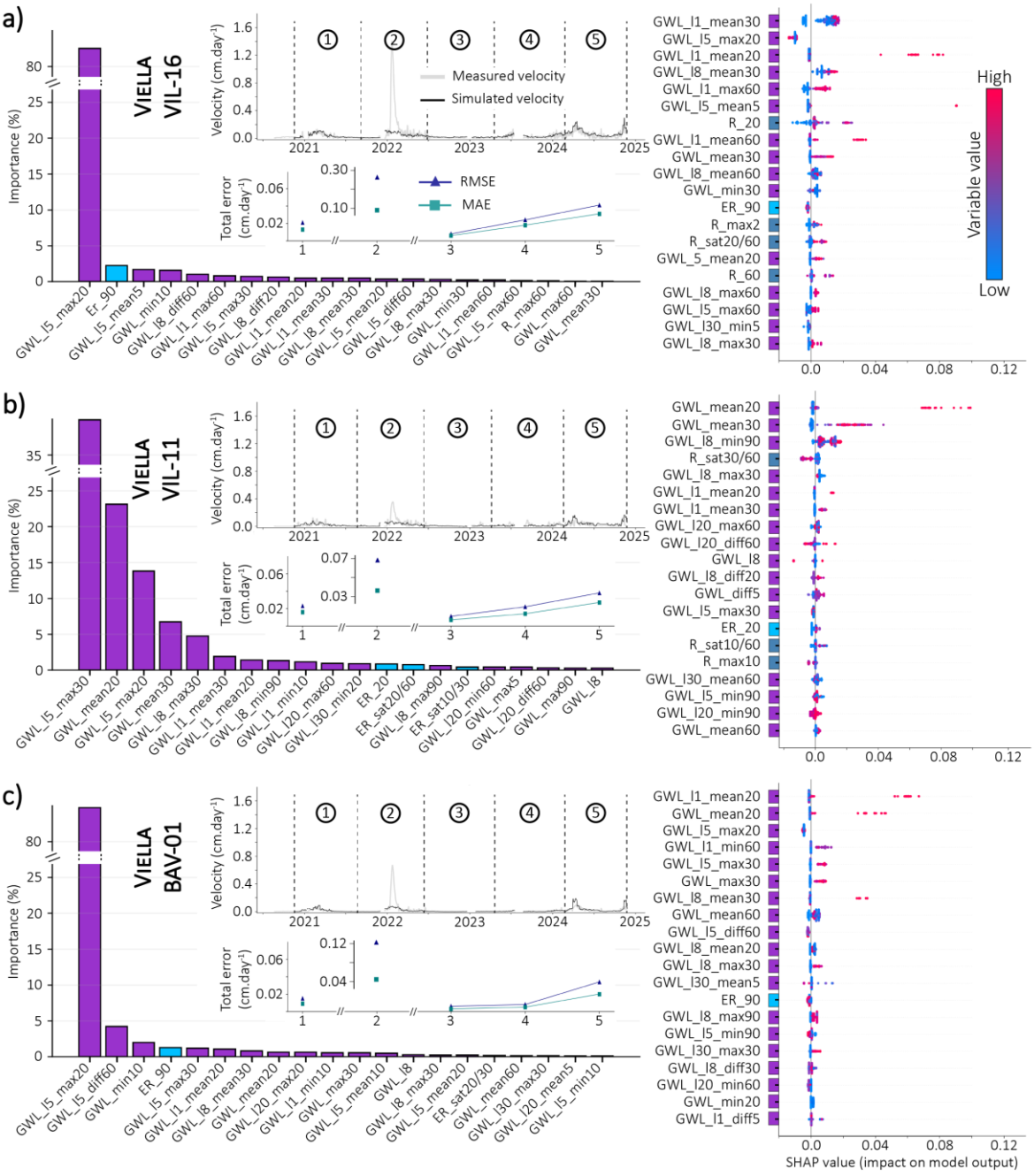


Figure 41 | Simulated landslide velocities, interpretability and explainability of the predictors for additional Viella monitoring plots (part 2) for **(a)** VIL-16, **(b)** VIL-11, and **(c)** BAV-01. Legends are shown in panel **(a)**.

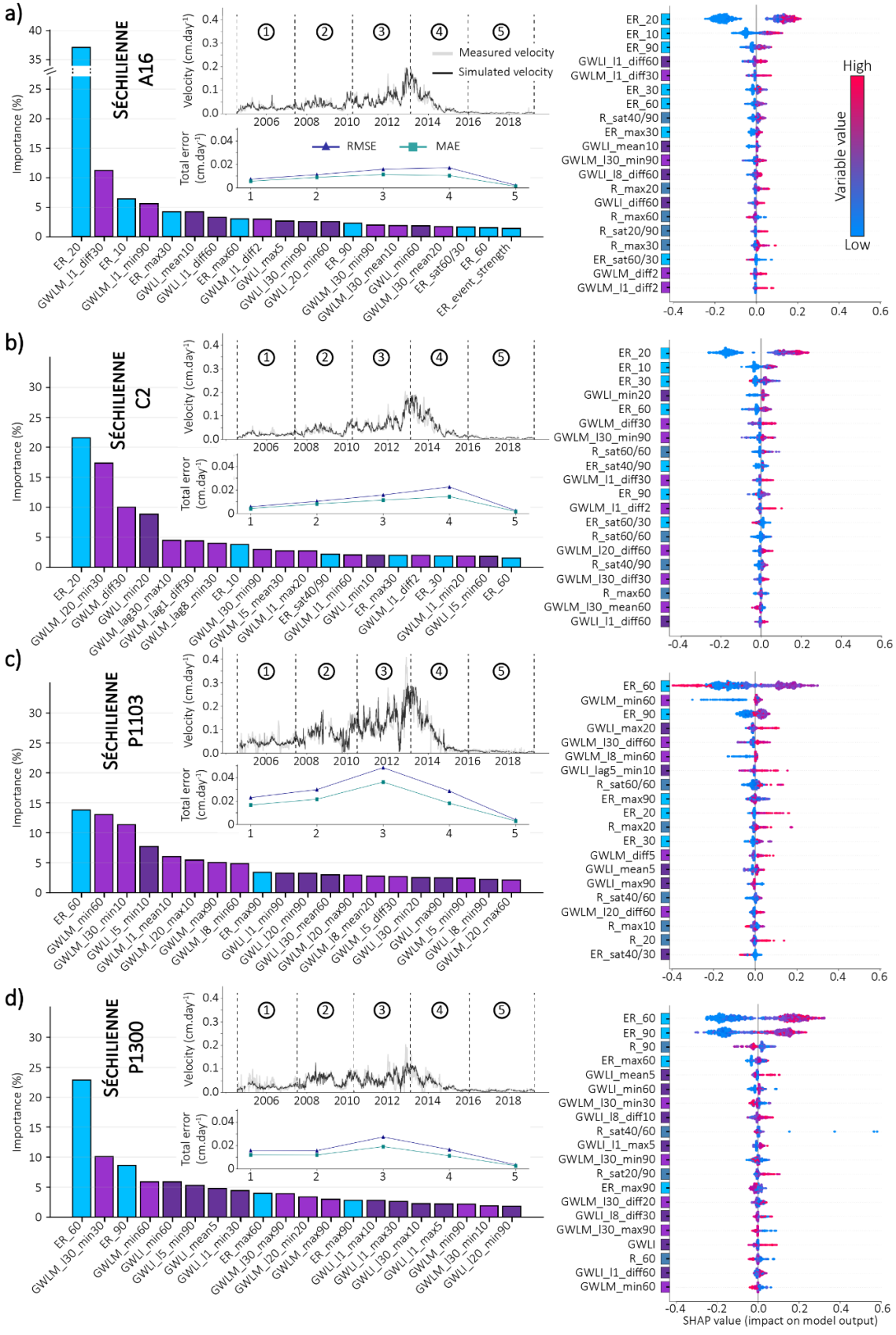


Figure 42 | Simulated landslide velocities, interpretability and explainability of the predictors for additional Séchilienne monitoring plots for **(a)** A16, **(b)** C2, **(c)** P1103, and **(d)** P1300. Predictors and colour-

coding follow **Figure 38**. At Séchilienne, two groundwater-level series are considered, representing shallow (GWLI) and deep hydrological reservoirs (GWLM).

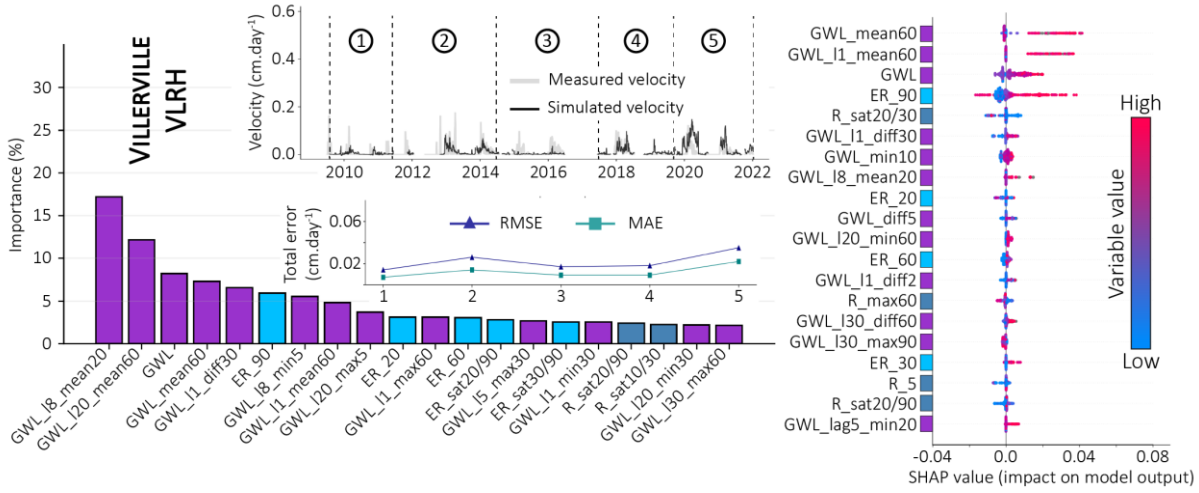


Figure 43 | Simulated landslide velocities, interpretability and explainability of the predictors for additional Villerville monitoring VLRH plot. Predictors and colour-coding follow **Figure 38.**

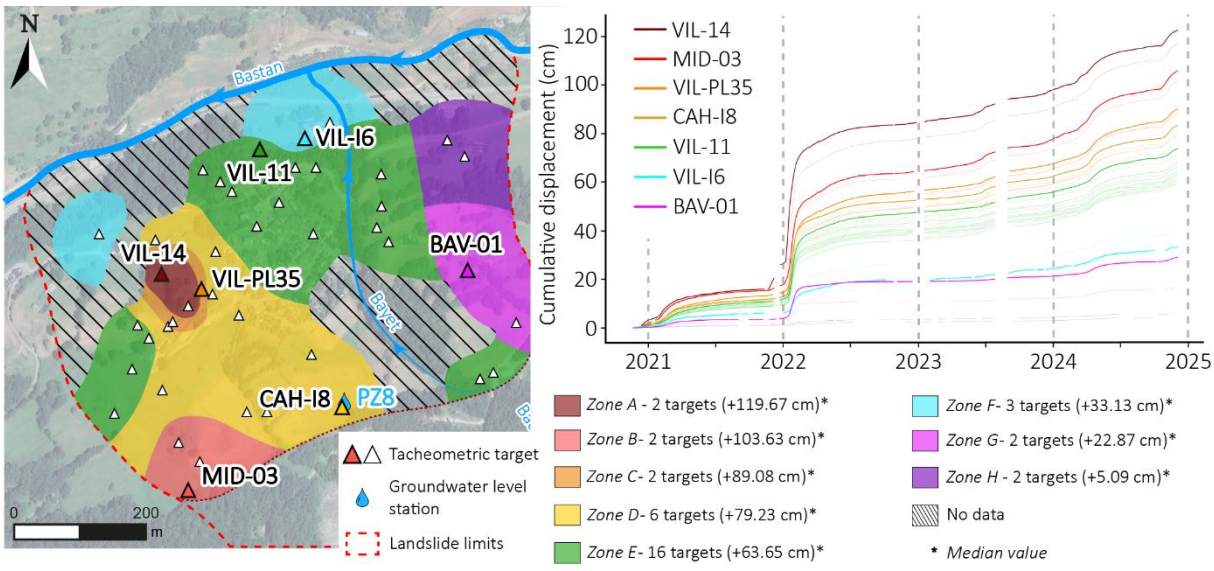


Figure 44 | Zonation map of cumulative surface displacement at Viella landslide. Analysed tacheometric targets and their associated time series are shown in colour. Additional monitoring plots are indicated in white; their cumulative displacement curves are overlaid using the colour of the corresponding displacement zone with partial transparency.

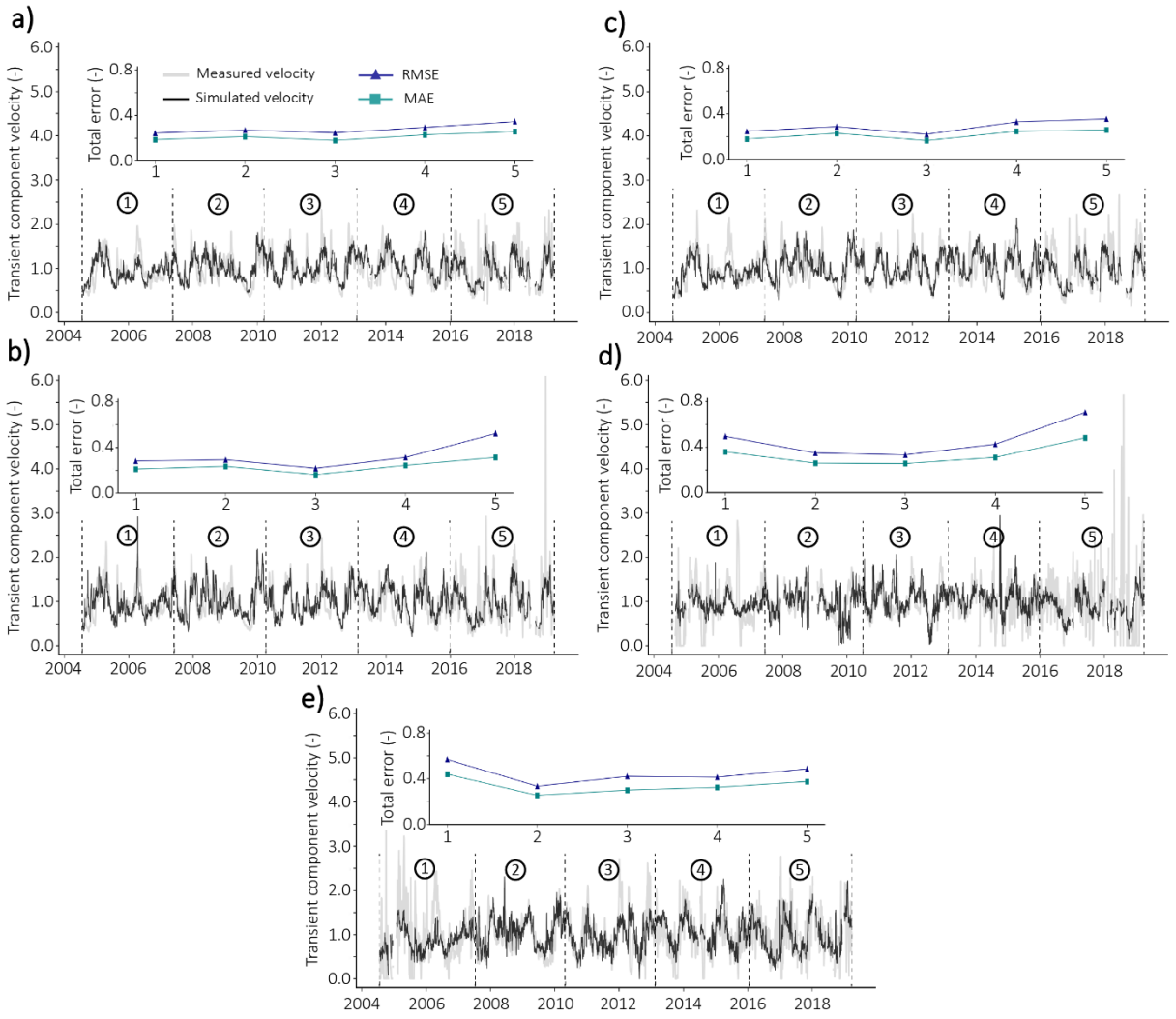


Figure 45 | Simulated transient-component velocities (dimensionless) for the Séchilienne monitoring plots. **(a)** plot A13, **(b)** plot A16, **(c)** plot C2, **(d)** plot P1103, and **(e)** plot P1300. Observed and simulated daily velocities are shown in grey and black, respectively, with predictive performance across the five test periods (periods 1–5) quantified using RMSE (triangles) and MAE (squares).

V.3. CHAPTER V – CONCLUSION

This chapter addressed the second research question of the thesis and the core modelling and interpretation stages of the methodological framework introduced in ***Chapter III***. Its objective was to reproduce daily surface velocity dynamics from hydro-meteorological forcing while making the underlying forcing–response relationships explicit in terms of dominant controls, non-linear behaviours, and characteristic response time scales.

Chapter V therefore developed and applied a transferable conventional machine learning workflow based on XGBoost and a generic set of physically-grounded predictors derived from net rainfall, effective rainfall, and groundwater level variations. By relying on long-term multi-source time series – including groundwater information obtained either from in situ monitoring or from reconstructed proxies – the framework supports the simulation of contrasted velocity regimes without requiring site-dependent predictors or explicit encoding of local geological structures. Importantly, the multi-plot application also acknowledges that forcing–response relationships may vary within a single unstable mass, and provides a consistent basis to document this internal variability.

Beyond the ability to reproduce observed dynamics, a central outcome of this chapter is that model transparency can be systematically integrated into the workflow. Combining native XGBoost features importance ranking with post-hoc SHAP explainability enables the extracted forcing–velocity relationships to be examined as physically consistent constraints, rather than as purely empirical correlations.

These results open several perspectives within and beyond the thesis workflow. First, the identified dominant predictors provide operational guidance to orient local monitoring strategies, by prioritising the forcing components and hydrological-state observations that most strongly condition velocity dynamics. Second, the inferred response signatures offer a valuable basis to inform process-based modelling efforts, by helping refine which mechanisms, parameters, and characteristic time scales should be represented when constructing physically based formulations. Third, the generic and physically grounded feature representation developed here provides a structured entry point for evolving towards more complex AI architectures (e.g., deep learning), while preserving the interpretability gains achieved. Finally, by making forcing–response relationships explicit, the framework establishes the methodological foundation required to extend deterministic simulations towards scenario-based and probabilistic applications, which is the focus of ***Chapter VI*** and its hazard mitigation perspective.

CHAPTER VI

EXTENDING THE AI-BASED WORKFLOW TOWARDS PROBABILISTIC DISPLACEMENT SIMULATIONS

CHAPTER VI EXTENDING THE AI-BASED WORKFLOW TOWARDS PROBABILISTIC DISPLACEMENT SIMULATIONS 133

VI.1. Chapter VI – Summary 135

VI.2. Paper - Probabilistic Simulation of Monthly Landslide Velocity Under Hydro-meteorological Variability (in preparation) 137

VI.2.1. Introduction 137

VI.2.2. Methods 140

VI.2.3. Study sites and model calibration 145

VI.2.4. Results 154

VI.2.5. Discussion 160

VI.2.6. Conclusion 163

VI.2.7. Extended data 165

VI.3. Chapter VI – Conclusion 168

VI.1. CHAPTER VI – SUMMARY

Chapter V demonstrated that XGBoost models, combined with a generic and physically-grounded hydro-meteorological feature set, can reproduce daily surface velocity dynamics across slow-moving landslides occurring in contrasted settings, while making the inferred forcing–response relationships explicit through interpretability and explainability analyses. Yet, this deterministic modelling strategy remains conditioned by a single historical condition of hydro-meteorological forcing. Even when model skill is satisfactory, it does not inform on the range of surface displacement that may be expected over a given period under alternative, yet plausible, hydro-meteorological conditions.

Building on the workflow introduced in ***Chapter III*** and implemented in ***Chapter V***, ***Chapter VI*** therefore directly addresses the third research question of this thesis: **How does slow-moving landslide displacement evolve under contrasting and plausible hydro-meteorological conditions?**

This chapter extends the deterministic framework towards probabilistic simulations by explicitly sampling hydro-meteorological variability and propagating it through the same feature engineering and velocity simulation stages. In line with the data requirements discussed in ***Chapter II***, the approach relies on long-term multi-source time series, including, here, solely monitored groundwater level records that provide a direct description of the slope hydrological state.

This study proposes a probabilistic framework to estimate monthly distributions of surface velocity by coupling stochastic hydro-meteorological scenario generation, groundwater level modelling, and machine learning-based velocity forecasting. The workflow consists of three successive steps. First, a modified Richardson-type weather generator is calibrated on historical rainfall and temperature series and used to generate an ensemble of 500 daily meteorological scenarios over a 15-month period (i.e., a full hydrological cycle, preceded by an initialisation window). Second, groundwater level variations are simulated for each scenario from the corresponding effective rainfall using a transfer-function hydrological model (implemented in PASTAS) with two response functions. Third, the generated hydro-meteorological and groundwater level series are derived into the same generic descriptive predictors as in ***Chapter V*** and used as inputs to pre-trained XGBoost models to simulate daily surface velocities. Daily simulations are finally aggregated to the monthly time step to derive probability distributions of monthly cumulative velocities.

The method is applied to three slow-moving landslides characterised by contrasted forcing–response regimes: Viella and Villerville (already analysed in ***Chapter V***) and Vallcebre (Spanish Pyrenees). The generated meteorological series preserve the main statistical properties of the observed records, while the simulated groundwater level series reproduce the dominant seasonal and event-driven behaviours observed at each site, providing coherent hydrological-state evolutions under a broad range of plausible forcing conditions.

The resulting monthly cumulative velocity distributions reveal site-specific seasonal signatures and quantify the expected monthly variability. When compared with historical monthly cumulative velocities, the simulated distributions generally encompass the observed values and reproduce the main seasonal contrasts, while also highlighting residual over- and underestimation during specific periods. These mismatches primarily reflect the combined sensitivity of the framework to (i) the accuracy of groundwater level simulations, and (ii) the representativeness of the velocity training datasets used to build the XGBoost models.

Overall, ***Chapter VI*** provides a first basis to translate deterministic AI-based velocity simulations into probabilistic monthly displacement distribution ranges under plausible hydro-meteorological variability, thereby offering a pathway towards scenario-based products that are directly relevant for hazard mitigation perspectives.

The following section (section VI.2.) is based on a manuscript in preparation for submission to Engineering Geology:

O. Béjean-Maillard, C. Bertrand, J.-P. Malet, J. Moya, and O. Maquaire. *Probabilistic Simulation of Monthly Landslide Velocity Under Hydro-meteorological Variability* (in preparation).

VI.2. PAPER - PROBABILISTIC SIMULATION OF MONTHLY LANDSLIDE VELOCITY UNDER HYDRO-METEOROLOGICAL VARIABILITY (IN PREPARATION)

VI.2.1. Introduction

Groundwater recharge driven by rainfall is a key factor for controlling the destabilisation of landslides, as it reduces the effective strength of slope materials through pore-water pressure variations *(Van Asch et al., 1999; Binet, Jomard, et al., 2007; Vallet et al., 2016; Prokešová et al., 2013; Lacroix et al., 2020)*. In slow-moving landslides undergoing long-term creep, the hydro-mechanical processes that lead to surface acceleration are highly complex and non-linear, and are strongly influenced by time-dependent factors that reflect the memory of the system *(Federico et al., 2012; Osawa et al., 2018; Pfeiffer et al., 2021; Ren et al., 2023)*.

Data-driven models, notably machine learning models, have emerged as powerful tools to reproduce the complex and non-linear relationships between hydro-meteorological forcing and landslide kinematics *(Nava et al., 2023; van Natijne et al., 2020)*. Numerous algorithms, such as Random Forests (RF) *(Krkač et al., 2017)* and Extreme Gradient Boosting (XGBoost) *(J. Xu et al., 2022b)*, as well as recurrent neural networks (e.g., LSTM, GRU) *(H. Jiang et al., 2020; W. Ma et al., 2022; Nava et al., 2023; L. Wang et al., 2023)*, have been successfully applied to model landslide velocities in a range of geological contexts and deformation mechanisms.

Most of these models are trained and tested on historical time series, using features derived from observed hydro-meteorological series. However, these time series represent only one specific hydro-meteorological context and therefore do not capture the full probability range of expected landslide surface velocity for a given period under other plausible hydro-meteorological conditions.

In hydrology, explicitly accounting for meteorological variability is a common approach when assessing streamflow, drought risk, or water availability *(Keller et al., 2015; Y.-H. Li et al., 2018; W. Liu et al., 2015; Ullrich et al., 2021)*. Stochastic weather generators, such as the Richardson type *(Katz, 1977; Richardson, 1981)*, are widely used to produce ensembles of daily rainfall and temperature series consistent with observed climatology. These models typically simulate wet/dry occurrence using a two-state Markov chain and sample precipitation amounts from parametric distributions, before generating other variables (e.g., temperature, solar radiation) conditional on rainfall occurrence. When parameterised on a monthly basis, they reproduce seasonal variations in occurrence, intensity and persistence, and provide synthetic time series that preserve the main statistical properties of the observed record.

To ensure consistency between meteorological forcing and groundwater dynamics, the groundwater-level response to the generated meteorological series must be reproduced while preserving the main characteristics of the observed fluctuations. Groundwater level modelling from rainfall has been extensively investigated, with approaches broadly divided into physically based models and data-driven models. Physically based models provide high accuracy groundwater dynamics simulations *(Hong & Wan, 2011b; Hu & Jiao, 2015; Kishor et al., 2025)*, but they often require explicit representation and calibration of numerous physical parameters, which may limit their transferability and operational use. Data-driven models, such as deep learning models, circumventing the need for detailed parameterisation while achieving high predictive skill *(H.-Y. Chen et al., 2023; Müller et al., 2021; Pham et al., 2022; Sahu et al., 2020; Vu et al., 2021)*. However, most of these approaches rely on sequences including past groundwater levels *(J. Sun et al., 2022; Vu et al., 2021)*, whereas using only rainfall-based inputs typically result in poorer skill *(J. Sun et al., 2022)*. Other studies have explored machine learning techniques, such as ensemble decision trees. For instance, *Krkač et al. (2017)* applied features extracted from rainfall series to reproduce daily groundwater variations with Random Forests, but obtained simulations remains smoother than the observed fluctuations.

More recently, lumped transfer-function models, notbaly those implemented in the PASTAS open-source framework *(Collenteur et al., 2019)*, have been proposed as an efficient way to simulate groundwater level variations by convolving different stresses (e.g., effective rainfall, pumping) with predefined response functions. These models have a limited number of parameters, rely only on stress inputs, and provide accurate reproduction of groundwater level variations.

To our knowledge, no previous study has explicitly accounted for hydro-meteorological variability when transposing this type of groundwater-based approach to the simulation of landslide velocities. Yet, considering such variability could provide information on the probability ranges of landslide surface velocity over predefined periods, which would be directly useful for hazard management services.

This paper proposes a framework to estimate monthly probabilities of surface velocity for slow-moving landslides by combining stochastic hydro-meteorological scenario generation, groundwater level modelling and machine-learning-based velocity simulation. The approach consists of three main steps. First, a modified Richardson-type weather generator is calibrated on historical rainfall and temperature data and is used to generate an ensemble of 500 daily meteorological scenarios over a 15-month period. Second, a lumped hydrological model with two transfer functions is employed to simulate groundwater level variations from the effective rainfall derived from each scenario. Third, the resulting hydro-meteorological time series are transformed into a set of generic descriptive features used as predictors to simulate daily surface velocities based on pre-trained XGBoost models (**Figure 46**). The

simulated velocities are then aggregated at a monthly scale to define the monthly cumulative displacement distribution ranges over a full year. The proposed method is applied to three slow-moving landslides occurring in distinct lithological settings: Villerville (French Normandy coast), Viella (central French Pyrennees), and Vallcebre (central Spanish Pyrennees), for which several years of daily measurements are available.

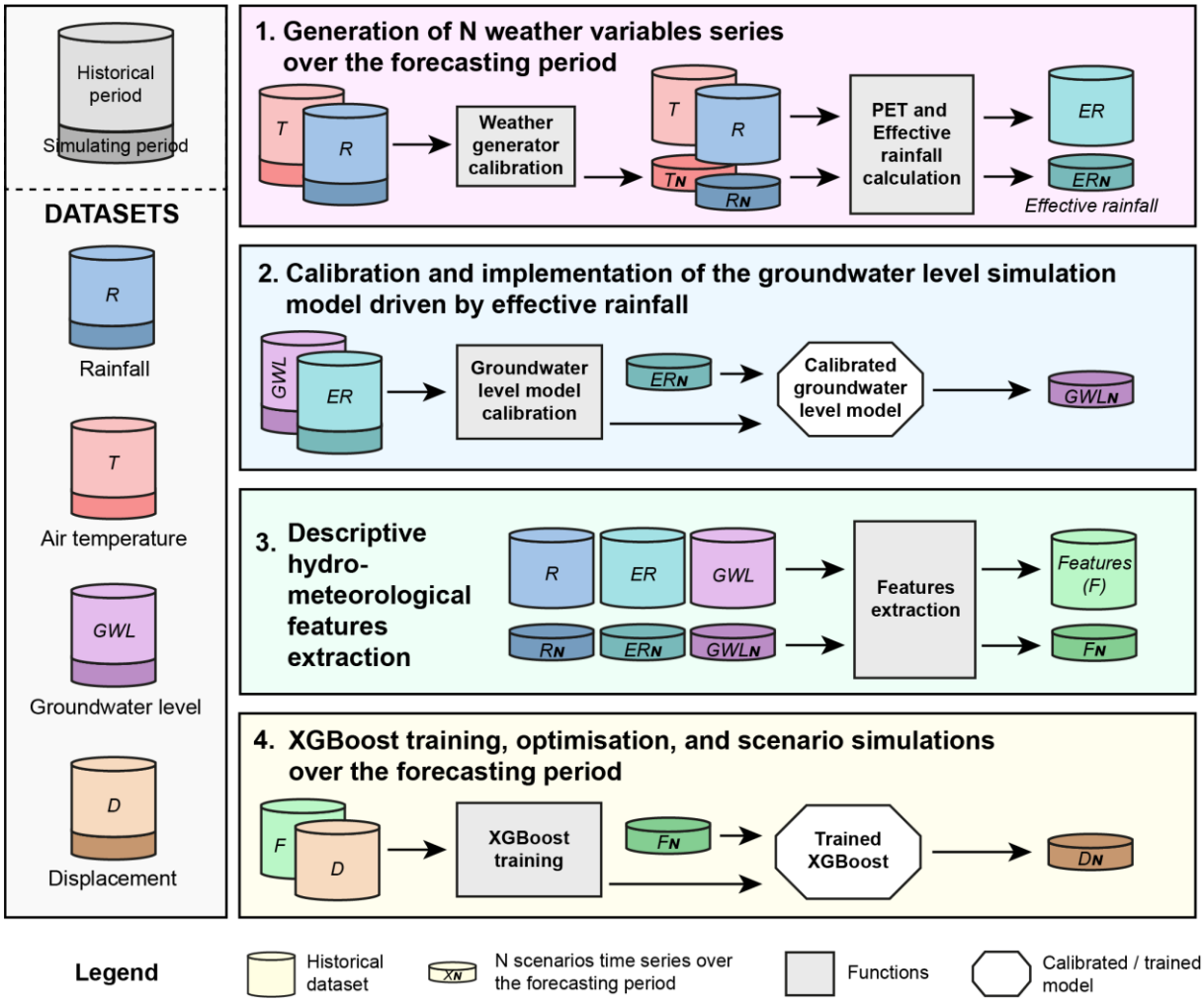


Figure 46 | Chapter VI - Schematic overview of the probabilistic simulation framework developed.

VI.2.2. Methods

VI.2.2.1. Meteorological variables generation

Richardson-type weather generators are well documented in the hydrological literature and are only briefly summarised here *(W. Liu et al., 2015; Ullrich et al., 2021)*. Their principle is to generate synthetic daily weather time series from a model calibrated on several years of historical daily observations. Following Richardson's original formulation, meteorological variables such as air temperature or solar radiation are assumed to be influenced by the occurrence of wet days. Consequently, the model first simulates daily rainfall occurrence, and then simulates the other meteorological variables conditional on the wet–dry sequence.

VI.2.2.1.1. Rainfall generation

In the original Richardson model *(Richardson, 1981)*, daily rainfall is simulated in two steps. First, rainfall occurrence is generated using a two-state Markov chain, which defines a binary wet–dry sequence. Second, rainfall amounts on wet days are sampled from a parametric probability density function, typically an exponential distribution in the original formulation. However, this formulation presents two main limitations. First, using a two-state Markov chain can restrict the coherence between simulated rainfall amounts. Second, while the exponential distribution may accurately represent the empirical cumulative distribution, it may be less suitable for reproducing the distribution tails.

To address these limitations, *Breinl (2016)* proposed a modified version of the Richardson model based on a six-state Markov chain (one dry state and five wet states corresponding to different magnitude orders of daily rainfall), and compared several candidate distributions. They showed that increasing the number of wet states helps preserve the coherence of simulated rainfall amounts across a range of intensities, and that other distributions, such as gamma and hybrid distributions, are more appropriate than the exponential distribution to reproduce the empirical probability density function (PDF).

For rainfall amounts, we test exponential and gamma probability density functions on a monthly basis. Their performances are compared by computing the RMSE between each theoretical PDF and the empirical PDF. Hybrid distributions, which combine two PDFs (one for the bulk and one for the distribution tail) (e.g., Gamma-GPD), are not considered here, as the choice of the probability threshold for using PDFs remains an open issue *(Langousis et al., 2016; Mínguez, 2025; Naveau et al., 2016)*, and the gamma distribution was found to provide a sufficiently good fit of the distribution tails *(Martinez-Villalobos & Neelin, 2019; Wójcik et al., 2014)*.

Following *Breinl (2016)*, rainfall amounts are generated in several steps once the wet–dry sequence has been simulated. First, for each wet state, rainfall amounts are sampled from the historical observations belonging to that state. Then, parameters of the rainfall amount distribution are sampled on the fitted PDF. Finally, the sampled amounts are adjusted based on the parametric distribution to generate distinct values from the historical measurements (**Figure 47a**). 300 series are generated over the full historical period to evaluate the generator's performance by comparing statistical properties of observed and simulated rainfall. Then, 500 series are generated over a 15-month target period, capturing seasonal behaviour across a full year while accounting for initial condition variability through the additional three months.

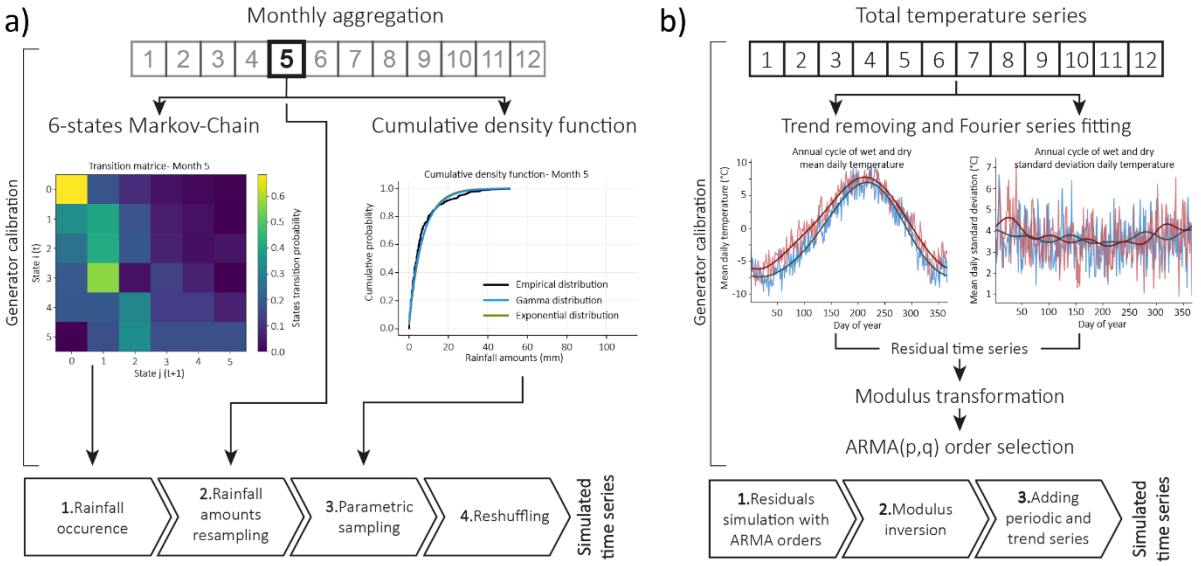


Figure 47 | Calibration and simulation workflow for **(a)** the rainfall, and **(b)** temperature time series generators (adapted from *Breinl et al., 2015, 2016*).

VI.2.2.1.2. Temperature generation

Temperature is simulated by superimposing a Fourier-series representation of the seasonal signal on a stochastic model for the residual component, thereby preserving both periodicity and temporal dependence. Following *Breinl et al. (2015)*, the residuals are first subjected to a modulus transformation to obtain an approximately normal distribution, and an ARMA model is then fitted to the transformed residuals, rather than a simple AR(1), which tends to underestimate autocorrelation persistence.

The temperature generator is built in four main steps. First, the long-term trend is removed from the historical temperature series. Second, periodic components associated with the annual cycle of the mean and standard deviation are modelled using third- and fifth-order Fourier series, respectively, and subtracted to obtain a residual time series. Unlike *Breinl et al. (2015)*, we distinguish between temperatures values on wet and dry days, resulting in separate Fourier fits for the mean and standard deviation under wet and dry conditions. Third, the residuals distribution is adjusted using the modulus transformation proposed by *John & Draper (1980)* which reduces skewness and brings the residuals closer to normality *(Box & Cox, 1964; John & Draper, 1980)*, thereby aligning more closely with the Gaussian assumptions of ARMA models. Transformation adequacy are assessed using Q-Q plots and the Lilliefors test at the 5% significance level *(Lilliefors, 1967)*. Fourth, an ARMA(p, q) model is fitted to the transformed residuals, with (p, q) selected for each case by minimising the Akaike Information Criterion (AIC) *(Akaike, 1974)*. Further details are provided in *Breinl et al. (2015)*. Complete temperature series are then simulated by generating residuals with the fitted ARMA model, applying the inverse modulus transformation to return to the original residual scale, and finally adding back the seasonal and trend components (**Figure 47b**).

300 synthetic temperature series are first generated over the full historical period to compare their statistical properties (i.e., monthly mean, standard deviation, maximum and minimum, and autocorrelation decay) with those of the observed data. Subsequently, 500 temperature series are generated over the 15-month simulation period.

VI.2.2.1.3. Effective rainfall generation

Effective rainfall has been shown to exert stronger influence than raw precipitation on landslide velocity dynamics *(Cappa et al., 2004; Prokešová et al., 2013; Vallet et al., 2016)*. To derive effective rainfall series for the 500 generated meteorological series, we first compute daily potential evapotranspiration using Oudin's method *(Oudin et al., 2005)*, which requires air temperature and site latitude. Effective rainfall is then calculated using a bucket-type approach, with the maximum soil available water capacity ($SAWC_{max}$) set according to the lithological context of each landslide.

VI.2.2.2. Groundwater level series generation

The PASTAS package simulates groundwater level variations as the superposition of multiple stresses (e.g. rainfall, evapotranspiration, pumping) *(Collenteur et al., 2019)*. In this study, effective rainfall is considered as the sole stress driving groundwater level variations. The groundwater response to effective rainfall is represented using two transfer functions: (i) an exponential response function (θ_{exp}) to capture fast responses (**Equation 19**), and (ii) an inverse gamma response function (θ_{γ}) to capture slow and delayed responses (**Equation 20**). For both response functions, *A* is a scaling factor that controls the amplitude of the head change induced by a unit stress, whereas *a* is a characteristic time scale, controlling the response rate.

$$\theta_{exp}(t) = \frac{A}{a} e^{-t/a}, \qquad t \geq 0 \tag{19}$$

$$\theta_{\gamma}(t) = A \frac{t^{n-1}}{a^{n}\Gamma(n)} e^{-t/a}, \qquad t \geq 0 \tag{20}$$

The inverse-gamma response function includes an additional shape parameter (*n*), which govern the curvature of the groundwater response. Each transfer function is defined by a set of parameters optimised by minimising an integrated objective function. In total, six parameters are calibrated over the full time series, with parameter bounds set to standard calibration ranges during optimisation.

To ensure appropriate initial conditions for groundwater level conditions, the hydrological model is warmed up over a one-year period. The calibrated model is then used to simulate groundwater level variations driven by the 500 effective rainfall series over the 15-month simulation period.

VI.2.2.3. Hydrometeorological features set and XGBoost application

The forcing time series (i.e., net rainfall, effective rainfall, and groundwater levels) are transformed into a set of descriptive hydro-meteorological features computed over multiple time windows (see ***section V.2.3. ;*** **Table 12**). These features are designed to enrich the training dataset and enabling the XGBoost model to capture non-linear and lagged forcing-velocity relationships.

The performance of XGBoost models based on this generic feature set has been demonstrated in ***section V.2.4.***. Consistent with that workflow, models are trained on the first 80% of the available time series. Although the 15-month simulation period may not match the remaining 20% originally used for testing, this does not affect model behaviour. The generated forcing series simply provide additional inputs that are independent from the training data.

In a second step, XGBoost hyperparameter are optimised to assess potential performance gains. A validation set corresponding to the last 20% of the training period is defined, and optimisation is performed using the Optuna package *(Akiba et al., 2019)*. Optuna relies on a Bayesian optimisation strategy with a Tree-structured Parzen Estimator (TPE) to efficiently explore the hyperparameter space. The optimisation is run over 100 trials to identify the best-performing hyperparameter set.

VI.2.2.4. Probabilistic forecast evaluation

To estimate the accuracy of probabilistic forecast, we used the Prediction Interval Coverage Probability (PICP) probabilistic metric *(Dogulu et al., 2015)*. PICP is the probability that the observed values (y_t) lie within the estimated prediction limits (*PL*) computed for a significance confidence level 1 - α (e.g., 90%) (**Equation 21**). Ideally, the PICP value should be equal to, or close to, the specified confidence level.

$$PICP = \frac{1}{n}\sum_{t=1}^{n} C \qquad \text{Where } C = \begin{cases} 1, & PL_t^{lower} \leq y_t \leq PL_t^{upper} \\ 0, & otherwise \end{cases} \tag{21}$$

Although PICP is a useful metric to evaluate how many observations fall within the estimated prediction intervals, it can be sensitive to the limited number of observations, leading to abrupt changes in coverage values. In addition, it does not account for the width of the intervals. As a result, very wide, poorly informative intervals will systematically include most or all observed values.

To address these limitations, we combine PICP with the Mean Interval Score (MIS) *(Gneiting & Raftery, 2007; Winkler, 1972)*. Interval scores simultaneously penalise wide prediction intervals and lack of coverage. In this setup, lower interval scores correspond to better prediction intervals. More specifically, *Gneiting & Raftery (2007)* define the interval score for a (1-α) % prediction interval [*l*,*u*] and observed value *y* to be (**Equation 22**):

$$IS_\alpha(l,u;y) = (u-l) + \frac{2}{\alpha}(l-y)\mathbb{1}_{\{y<l\}} + \frac{2}{\alpha}(y-u)\mathbb{1}_{\{y>u\}} \tag{22}$$

where $\mathbb{1}$ is the standard indicator function.

VI.2.3. Study sites and model calibration

Three deep-seated landslides – Viella, Villerville and Vallcebre – are studied (**Figure 48a**). Each site provides several years of daily measurements (**Table 16**), enabling analysis of surface velocity responses to variations from external forcing across contrasting lithological settings.

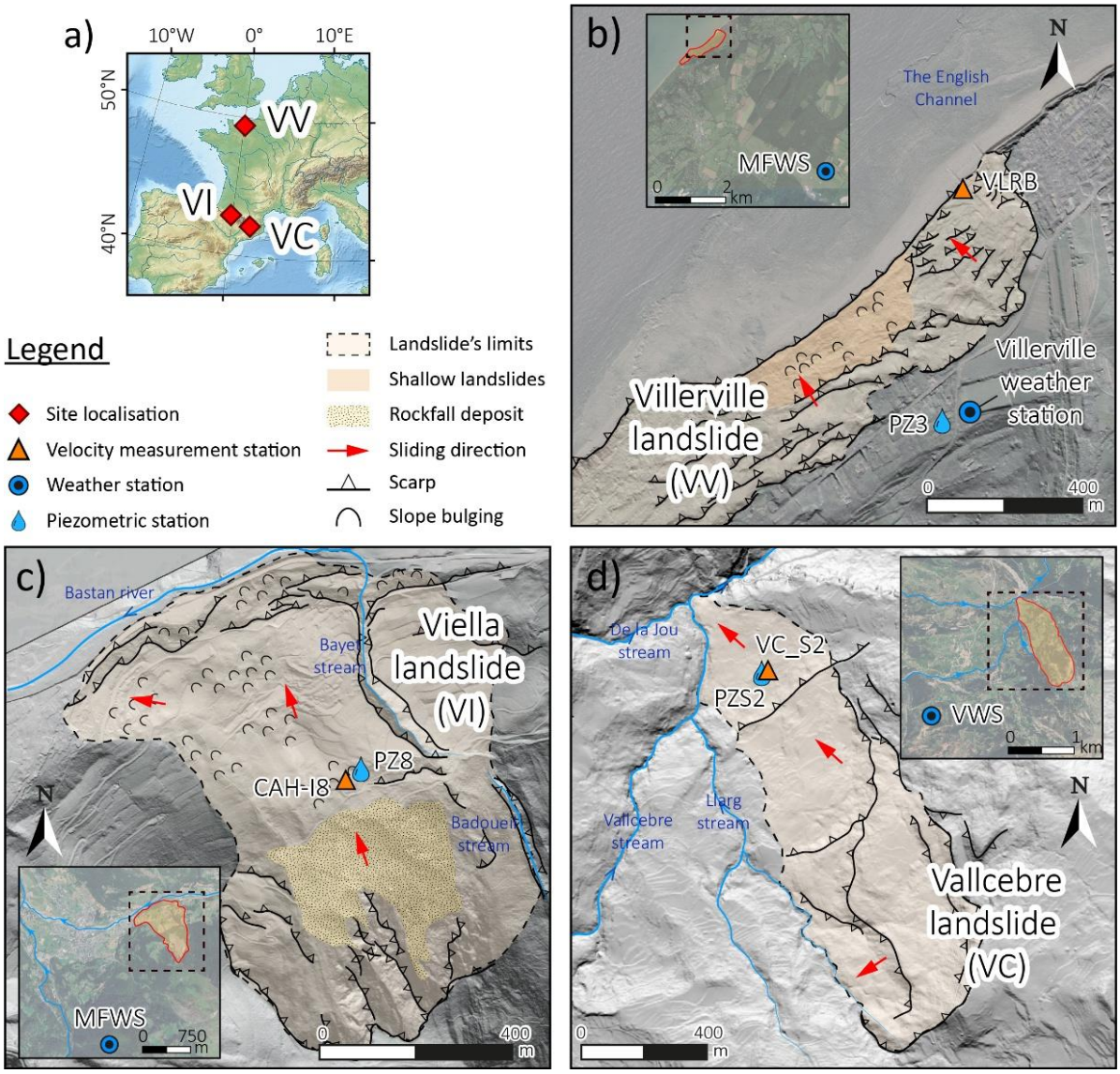


Figure 48 | **(a)** Location of the three studied deep-seated landslides, with details on the geomorphological features and monitoring network for **(b)** Villerville (adapted from *Lissak, 2009*), **(c)** Viella, and **(d)** Vallcebre. MFWS and VWS refer to Meteo-France Weather Station and Vallcebre Weather Station.

VI.2.3.1. Villerville landslide

The Villerville landslide is located on the Normandy coast, approximately 10 km south of Le Havre, and affects a coastal area of about 0.47 km², with basal rupture depths ranging from 10 to 40 m (**Figure 48b**). The unstable mass involves a stratified lithological succession composed of chalk blocks, sandy layers and colluvial debris overlying a marl substrate. The landslide can be subdivided into two main sectors, with predominantly translational deformation in the upper part and rotational deformation in the lower part. The site is influenced by an oceanic climate, with a mean annual precipitation of about 870 mm *(Lissak et al., 2009)*. Surface velocity is monitored at station VLRB, located in the north-eastern sector near the landslide toe. Groundwater level data are measured at piezometer PZ3, situated upslope of the unstable area. Meteorological data are recorded by a weather station near PZ3 and complemented by a secondary station located about 4.5 km to the southeast to reduce data gaps (**Figure 48b**). Depending on variables, the data availability ranges from 12 to 25 years (**Table 16**).

The velocity time series exhibits periodic acceleration phases, typically ranging from 0.2 to 0.5 cm.day^{-1}. These accelerations mostly occur between December and March and coincide with groundwater level rises at PZ3. Groundwater level fluctuations, ranging from −10.0 to −6.0 m depth, are primarily controlled by effective rainfall inputs. During the first four years of monitoring, groundwater levels showed lower mean values and a reduced reactivity to rainfall, which was reflected in smaller velocity amplitudes (0.1–0.3 cm.day^{-1}). In contrast, the subsequent period was characterised by higher mean groundwater levels (around −8.0 m depth) (**Figure 48b**). Given the low-permeability materials composing the landslide, the forcing–velocity relationship is mainly explained by long-term hydro-meteorological features over the 60 days preceding the current velocity value (see ***section V.2.5.1.***).

VI.2.3.2. Viella landslide

The Viella landslide is located in the French Pyrenees and affects an unstable area of approximately 0.59 km² along the left bank of the Bastan River, directly impacting the village of Viella (**Figure 49a**). The instability was reactivated following a rockfall in 2018 that occurred upslope. The landslide mass involves a heterogeneous lithology composed of colluvial and alluvial debris, varved clays and morainic deposits from the last glaciation, overlying a calcareous–schist substrate. The internal structure is complex, with several compartments exhibiting distinct kinematics and bounded by rupture surfaces. This structural heterogeneity strongly influences the hydrogeological system. Groundwater level variations depend on internal discontinuities within the landslide body, local infiltration from surface streams, and deeper groundwater flow originating from the rockfall aquifer and upslope snowmelt. The site is subject to a mountain climate, with a mean annual precipitation of about 1203

mm, calculated over 14 years (2008-2024), and is regularly influenced by snowfall and snowmelt in the upper catchment.

Surface velocity monitoring is performed at station CAH-I8, selected from the local monitoring network for its proximity to piezometer PZ8, which records groundwater level fluctuations. Meteorological data are provided by a weather station situated about 1.5 km to the southwest of the unstable area (**Figure 48c**). Data availability ranges from 4 to 17 years depending on the variable considered, and is summarised in **Table 16.**

The velocity time series reveals annual acceleration phases, with amplitudes reaching up to 0.25 $cm.day^{-1}$. Surface velocity dynamic is correlated with groundwater level variations, and the most pronounced accelerations occur during abrupt groundwater-level rises. A notable event occurred in early 2022, when the landslide experienced an exceptional acceleration phase with velocities up to 1.7 $cm.day^{-1}$, following intense precipitation and two successive groundwater level rise periods of about 2 m and 1 m, respectively, before rapidly decreasing (**Figure 49a**). The forcing–velocity relationship is primarily explained by medium-term groundwater-related variables computed over 20 to 60 days, for which groundwater level plays a key role in reproducing the landslide acceleration phases (see ***section V.2.5.1.***).

Table 16 | Geomorphological and datasets descriptions of the studied landslides

		Viella (VI)	**Villerville (VV)**	**Vallcebre (VC)**
Geomorphological description	Geographical coordinates	42.87, 0.01	49.40, 0.12	42.21, 1.83
	Altitudes (m a.s.l.)	732 to 1344	0 to 77	966 to 1264
	Mean slope (°)	27	10	9
	Area (km^2)	0.59	0.47	0.80
	Lithology description	Morainic and colluvial debris overlying calcareous-schist bedrock	Chalk, sand, and colluvial debris overlying marl substrate	Calyey siltstone and colluvial debris overlying limestone substrate
Dataset description	Meteorological record period	2008/01/01 – 2025/10/26	2000/01/01 – 2025/10/26	1996/11/22 – 2015/12/31
	Groundwater level record period	2020/03/12 – 2024/11/25	2008/08/21 – 2022/03/11	1996/11/22 – 2015/09/13
	Surface velocity record period	2020/08/21 – 2024/12/02	2009/07/26 – 2025/03/05	1996/11/22 – 2015/12/09
	Target simulating period	2023/09/01 – 2024/11/30	2020/09/01 – 2021/11/30	2012/08/01 – 2013/10/31

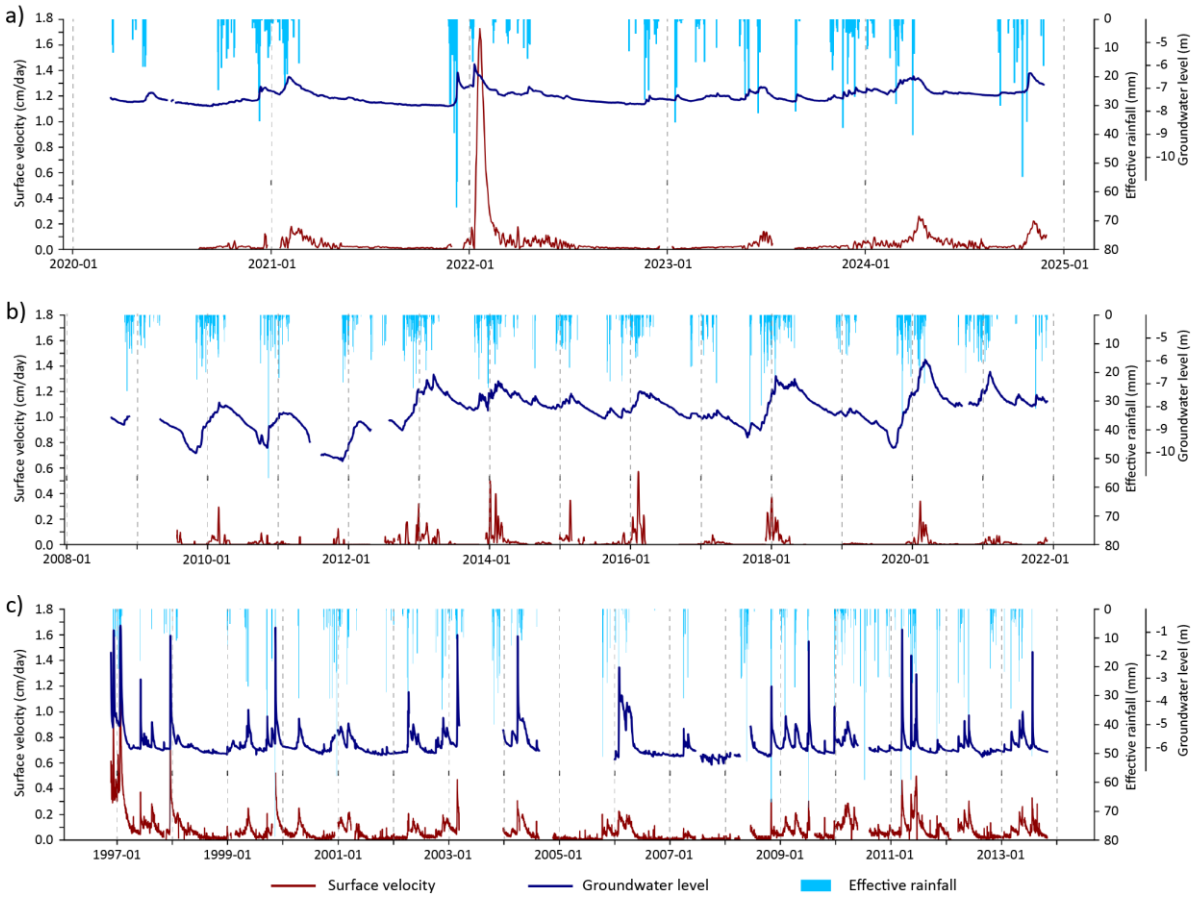


Figure 49 | Measured daily surface velocities, groundwater level, and effective rainfall on **(a)** Viella, **(b)** Villerville, and **(c)** Vallcebre landslides.

VI.2.3.3. Vallcebre landslide

The Vallcebre landslide is located in the Spanish Pyrenees, about 140 km north of Barcelona, and affects an area of approximately 0.8 km² *(Gili et al., 2021)*. The landslide is structured into three units, with the southern units being the most active (**Figure 48d**). The unstable mass involves colluvial debris, clayey siltstones and fissured shales overlying a limestone substrate. The basal rupture surface is broadly planar and located within a thin fissured shale layer at 15 to 25 m depths *(Corominas et al., 2005)*. The site is subject to a relatively dry mountain climate, with a mean annual precipitation of about 852 mm, calculated over 19 years (1996-2015).

Surface velocity and groundwater levels are monitored in borehole VC_S2, where velocity is measured using an extensometer. Meteorological data are provided by a weather station located about 2 km to the southwest of the unstable area (**Figure 48d**). Data are available over a 19-year period, from early 1996 to late 2015 (**Table 16**).

The velocity time series exhibits several seasonal acceleration phases, generally occurring either in spring or between October and November, with amplitudes ranging from 0.2 to 0.4 cm.day^{-1}. These accelerations are strongly correlated with groundwater level fluctuations and effective rainfall events (**Figure 49c**). *Corominas et al. (2005)* attributed this behaviour to fractures within the landslide body, which act as preferential drainage pathways through ground fissures and bypass matrix infiltration in the clayey material. As this landslide was not included in ***Chapter V***, a similar application was carried out here to train the XGBoost model and identify the main controlling hydro-meteorological features. Daily groundwater level variation emerged as the primary predictor of surface velocity. The XGBoost results and feature controls are presented in the Extended Data section (see ***section VI.2.7.*** ; **Figure 56**).

VI.2.3.4. Models calibration

The weather generator was first calibrated by identifying the Markov-chain order and the parametric probability density function that best fitted the empirical distribution of daily rainfall. For the three sites and for each month, a first-order Markov chain yielded the lowest BIC values, and the gamma distribution function provided a better fit than the exponential distribution (**Table 17**). Using the monthly transition probability matrices and the fitted gamma distributions, 500 rainfall time series were then generated over the full observation period. For each site, their statistical distributions closely match those of the observed series (**Figure 50**).

Table 17 | BIC values for Markov chain orders and RMSE of fitted cumulative distribution functions used for rainfall generator calibration. Reported values are the median over the 12 months.

		Viella	Villerville	Vallcebre
Mean Markov-Chain BIC value	Order = 1	1382.3	2162.6	1288.8
	Order = 2	2019.2	2857.3	1848.1
	Order = 3	3656.7	5195.1	3125.0
Mean Gamma RMSE		0.037	0.049	0.062
Mean Exponential RMSE		0.074	0.098	0.145

Temperature generator calibration is carried out by selecting the ARMA(p, q) orders that minimised the AIC when fitted to the residual temperature series. Prior to ARMA fitting, residuals are transformed using the modulus transformation to obtain a distribution closer to normality. This transformation involves estimating two λ parameters to adjust the lower and upper tails of the distribution. A Monte Carlo procedure with 10,000 runs was applied, exploring λ values ranging between −2 and 2. Because tail shape may vary seasonally, the procedure was applied separately to four seasonal residual distributions. As confirmed by the Lilliefors test, the modulus transformation yields

residuals that are approximately normally distributed. The resulting seasonal λ values for each site are summarised in **Table 18.** Based on AIC, the selected ARMA model orders are (3, 1) for Viella and Villerville, and (3, 2) for Vallcebre.

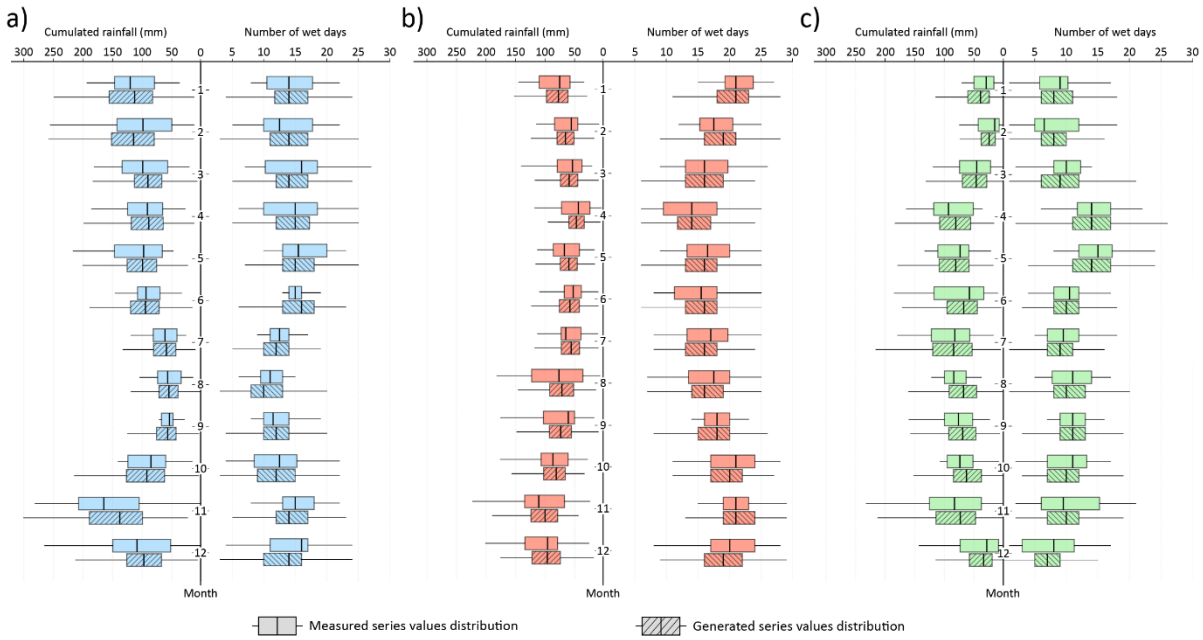


Figure 50 | Rainfall amount and number of wet days values distribution from measured and generated series for **(a)** Viella (blue), **(b)** Villerville (red), and **(c)** Vallcebre (green).

Table 18 | Lambda, RMSE, and Lilliefors test results for the modulus transformation applied to the seasonal temperature distributions at the three sites. Seasons letters depict considered months. P-values are reported before and after the transformation, whereas RMSE refers to the minimum value achieved after transformation. The Lilliefors test evaluates the null hypothesis that the data are drawn from a normal distribution.

		Modulus transformation			**Lilliefors test**		
	Season	**λ for x ≥ 0**	**λ for x ≤ 0**	**RMSE**	**p-value pre-transformation**	**p-value post-transformation**	**Rejection of the Null hypothesis**
Viella	Winter (DJF)	1.1371	0.7922	0.0235	0.0001	0.9358	No
	Spring (MAM)	0.9399	1.0850	0.0200	0.0721	0.8016	No
	Summer (JJA)	0.9800	1.4934	0.0283	0.0001	0.2422	No
	Autmun (SON)	1.1572	1.2996	0.0180	0.0695	0.8359	No
Villerville	Winter (DJF)	1.3744	0.9379	0.0154	0.0001	0.2958	No
	Spring (MAM)	0.8037	1.1778	0.0382	0.0001	0.2094	No
	Summer (JJA)	0.4777	1.3637	0.0160	0.0001	0.7816	No
	Autmun (SON)	0.9004	1.0592	0.0231	0.0413	0.4981	No
Vallcebre	Winter (DJF)	0.9255	0.8120	0.0142	0.1466	0.9290	No
	Spring (MAM)	0.9153	0.9294	0.0235	0.3269	0.5970	No
	Summer (JJA)	1.0717	0.9766	0.0228	0.2450	0.7185	No
	Autmun (SON)	1.0919	1.0784	0.0322	0.1860	0.0748	No

By applying the inverse transformation and reconstruction steps, 300 temperature series were generated over the full observation period. The statistical properties of the generated series are close to those of the observations (**Figure 51**). Monthly means and standard deviations yield low MIS values, and the corresponding prediction intervals contain 75.0 % to 83.3 % of the measured values (**Table 19**). Uncertainties are mainly associated with the intervals for September to December at the three sites (**Figure 51**). The reproduction of extreme values is less accurate. PICP ranges from 75.0 % to 91.7%, whereas higher MIS values indicate wider prediction intervals, with notable discrepancies during the summer months at Vallcebre and the winter months at Viella. These uncertainties may be related to the wet-dry days distinction, but they remain within acceptable ranges. The autocorrelation function decay over the first 10 lags are satisfactorily reproduced and presented in the Extended Data section (see ***section VI.2.7*** ; **Figure 57**).

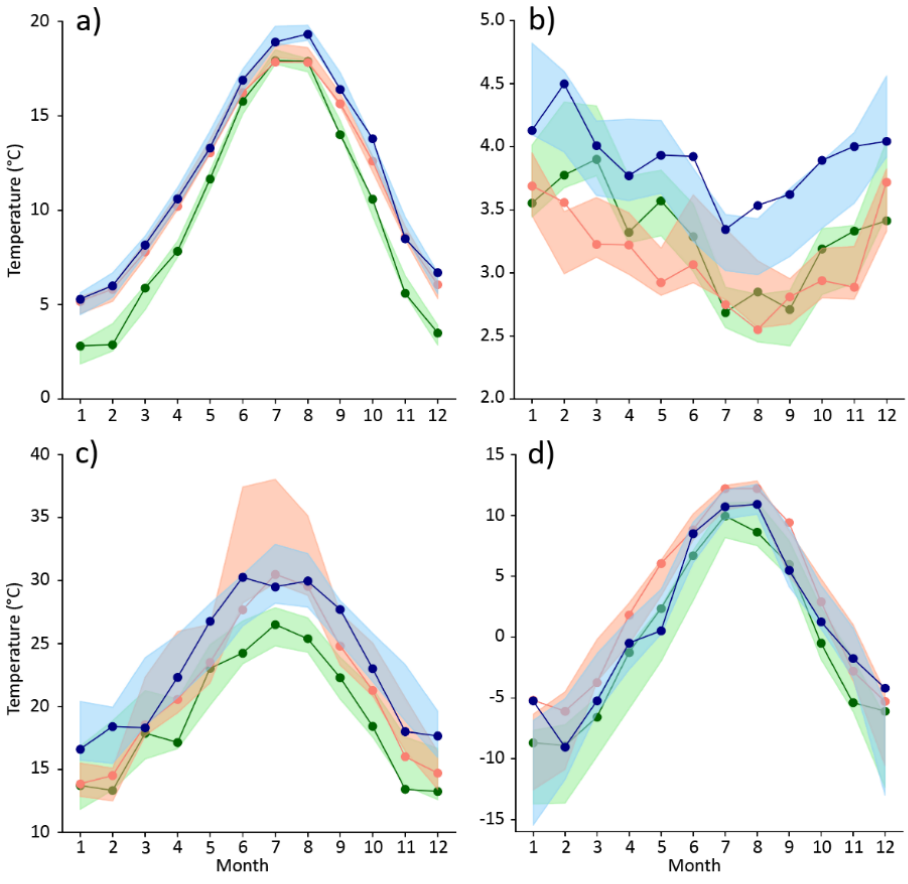


Figure 51 | Monthly **(a)** mean, **(b)** standard deviation, **(c)** maximum, and **(d)** minimum temperature value from measured series (solid line) and generated series (shaded area) for Viella (blue), Villerville (red), and Vallcebre (green).

Table 19 | Total temperature probabilistic metrics for 90% prediction intervals.

	Viella		Villerville		Vallcebre	
	PICP (%)	**MIS (°C)**	**PICP (%)**	**MIS (°C)**	**PICP (%)**	**MIS (°C)**
Mean	75.0	1.59	83.3	1.37	83.3	1.32
Standard deviation	75.0	0.91	75.0	0.80	83.3	0.63
Maximum	83.3	6.11	83.3	6.68	83.3	4.65
Minimum	83.3	9.28	75.0	6.97	91.7	5.29

Groundwater level models were calibrated over the full groundwater time series, with an additional one-year warm-up period used to initialise model states from effective rainfall. The PASTAS package estimates the parameter sets that best describe the exponential and inverse gamma response functions by minimising its native objective function. Calibrated parameter values for both response functions are summarised in **Table 20**. Simulated groundwater levels closely match observations at the three sites, with R^2 values of 0.84, 0.76, and 0.53 for Viella, Villerville, and Vallcebre, respectively. Some site-specific behaviours are less well reproduced, notably the first years of the Villerville record and the short-duration extreme events at Vallcebre (**Figure 52**).

Table 20 | Optimal PASTAS model parameters and R^2 values for groundwater level simulations at Viella, Villerville, and Vallcebre.

	Viella	Villerville	Vallcebre
A_exp	0.15	1.95	0.24
a_exp	46.18	304.87	11.49
A_gamma	1.04	0.89	0.30
a_gamma	370.81	1059.26	57.91
n_gamma	1.21	1.33	1.29
d	-8.72	-10.86	-6.30
R^2	0.84	0.76	0.53

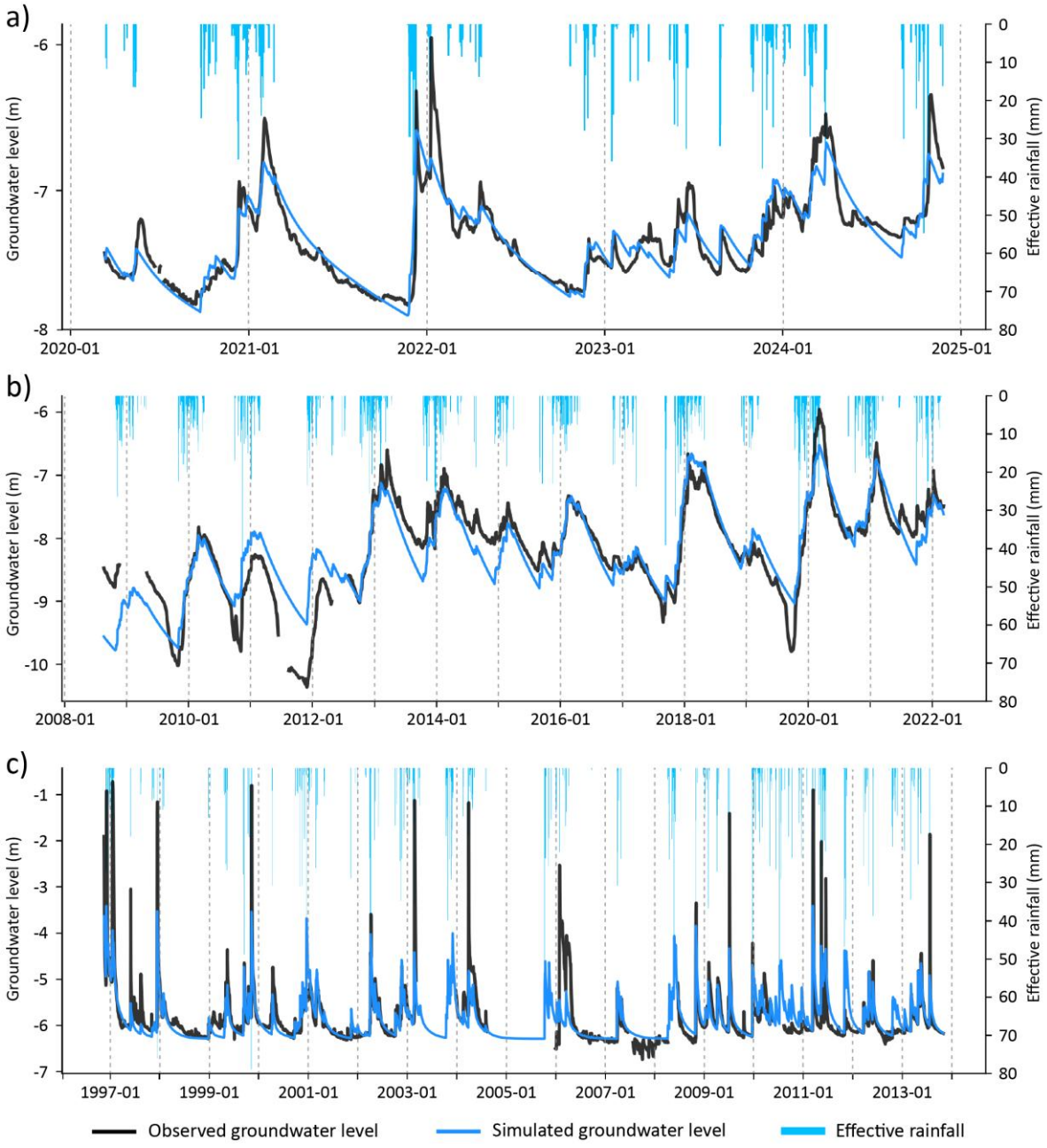


Figure 52 | Measured (black) and simulated (blue) groundwater level series for **(a)** Viella, **(b)** Villerville, and **(c)** Vallcebre.

VI.2.4. Results

VI.2.4.1. Generated forcing time series

Figure 53 shows, for each site, the generated and measured forcing time series over the 15-month simulation period.

The generated cumulative rainfall series display long, progressive increases without a marked seasonal pattern. Comparison with the measured cumulative rainfall indicates that the target simulation years are relatively dry at Viella, wet at Villerville, and close to the median conditions at Vallcebre.

Viella and Villerville effective rainfall series display marked seasonality, where low infiltration broadly prevails from April to early October, and higher infiltration occurs from autumn to spring (**Figures 53a-b**). For instance, at Viella, the 2024-series illustrates an earlier onset of recharge following a overall drier preceding year, resulting in a near-median cumulative effective rainfall. At Villerville, the 2021 series displays broadly typical behaviour, but with an earlier response in October. At Vallcebre, seasonal recharge variations are less pronounced. Both the generated series and the observed 2013-series show a step-like response closely tied to rainfall events, although this relationship is less systematic during summer months (**Figure 53c**).

The generated groundwater level series for Viella and Villerville reproduce the seasonal recharge patterns. From late autumn to spring, groundwater levels rise, with amplitudes depending on recharge intensity, whereas summer is characterised by a gradual recession punctuated by occasional low-amplitude rises. Relative to the observed groundwater records, the envelope of generated series spans the full range of observed variations and includes both lower and higher saturation states (**Figures 53a-b**).

At Vallcebre, the generated groundwater series exhibit sharp, high-amplitude events throughout the year, superimposed on a 6 m depths base level. This behaviour is consistent with historical groundwater dynamics, which are directly driven by rainfall events through preferential flow paths in the fissure network. Despite the strong event-to-event variability, reactivity is generally higher from October to January and from March to June. Events also occur during summer, but their amplitudes are typically lower than during the rest of the year (**Figure 53c**).

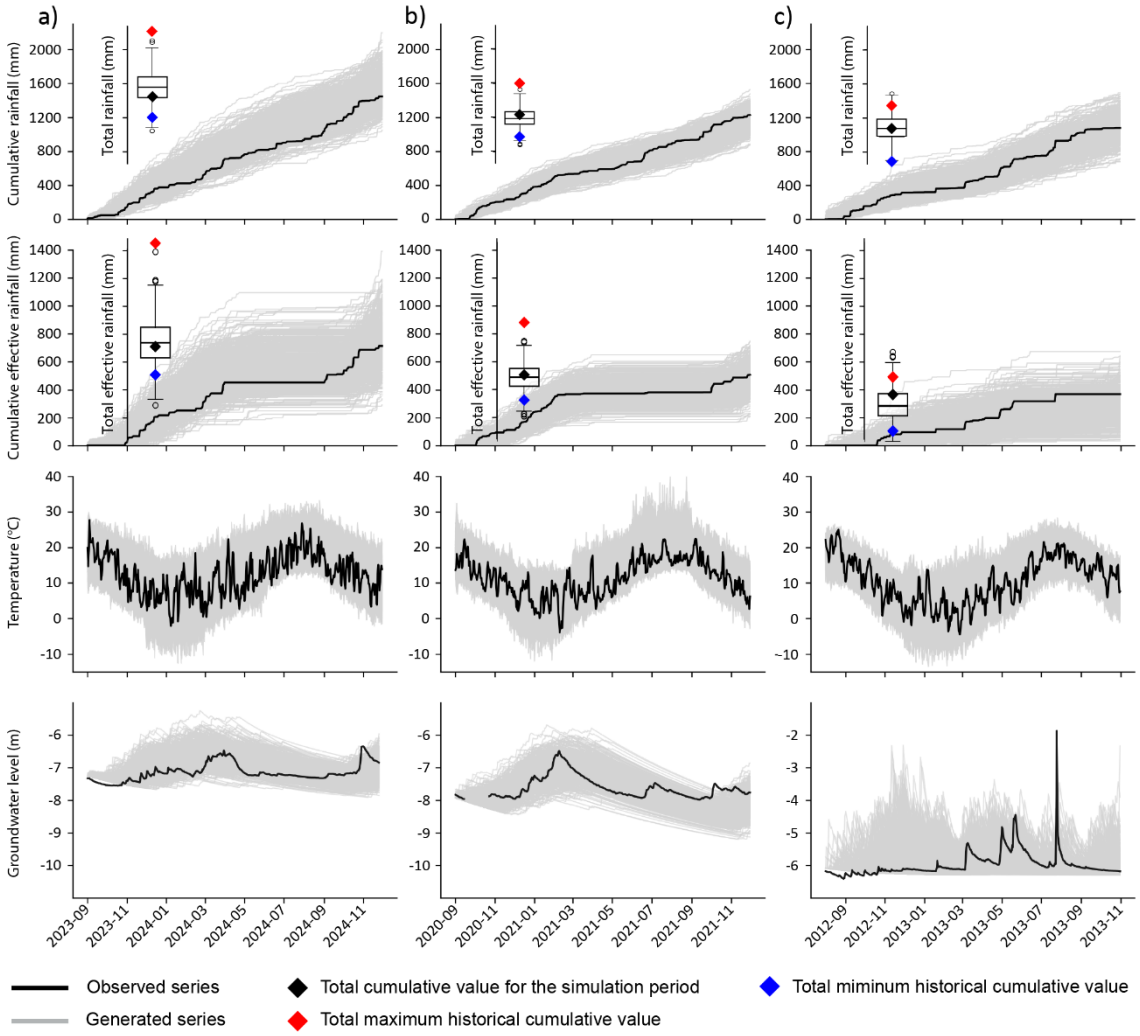


Figure 53 | Generated net rainfall, effective rainfall, temperature, and groundwater level series for **(a)** Viella, **(b)** Villerville, and **(c)** Vallcebre. For each site, the 15-month simulation period (corresponding to the XGBoost test set) is shown, with the observed series plotted as a black line. Total net rainfall and effective rainfall are summarised using box plots.

VI.2.4.2. Monthly cumulative velocity distribution

A total of 500 daily velocity series are simulated from hydro-meteorological features derived from the generated forcing series. Daily velocities are then aggregated to a monthly time step, and the resulting distributions are displayed as boxplots. Observed monthly cumulative velocities from the historical record are superimposed as individual points (**Figure 54**).

VI.2.4.2.1. Viella monthly cumulative velocity distribution

At Viella (**Figure 54a**), monthly cumulative velocities exhibit pronounced seasonality. Median values are highest from January to May and lowest from September to December, while the remaining months represent transition periods with gradual shifts in the distributions. Variability is particularly high in winter and spring. Monthly cumulative velocity typically ranges from 2.5 to 9 cm.month^{-1} and can reach 11-19 cm.month^{-1} when considering 90% of the distribution. During the rest of the year, total monthly cumulative velocity generally does not exceed 5 cm.month^{-1} , with narrow distributions clustered near zero, which indicates sustained periods of low velocities. However, the distributions derived from the generated series tend to systematically overestimate monthly cumulative velocities.

According to MIS results, monthly cumulative velocities from December to June are associated with high MIS values ranging from 10.3 to 131.5 cm.month^{-1} . For the other months, MIS values are lower, between 3.0 and 7.6 cm.month^{-1}, with October showing the best agreement. However, these performance metrics must be interpreted with caution because only 2 to 4 observed monthly values are available to evaluate each distribution (**Table 21**).

These discrepancies mainly reflect limitations of the XGBoost simulation and the representativeness of the training dataset. The available velocity record spans only four years and include four major acceleration phases. Two occur in the testing period, and one corresponds to an exceptional high-velocity event, leaving only a single “typical” acceleration phase in the training set. As a result, simulated daily velocity series often display step-like behaviour, with maxima driven either by this single representative event or by extreme events (see ***section VI.2.7*** ; **Figure 58**). Transition phases are poorly reproduced and frequently appear as abrupt shifts rather than gradual, physically realistic changes. Overall, this lack of temporal coherence leads to overestimated monthly cumulative velocities and inflated variability during seasons when extreme events may occur.

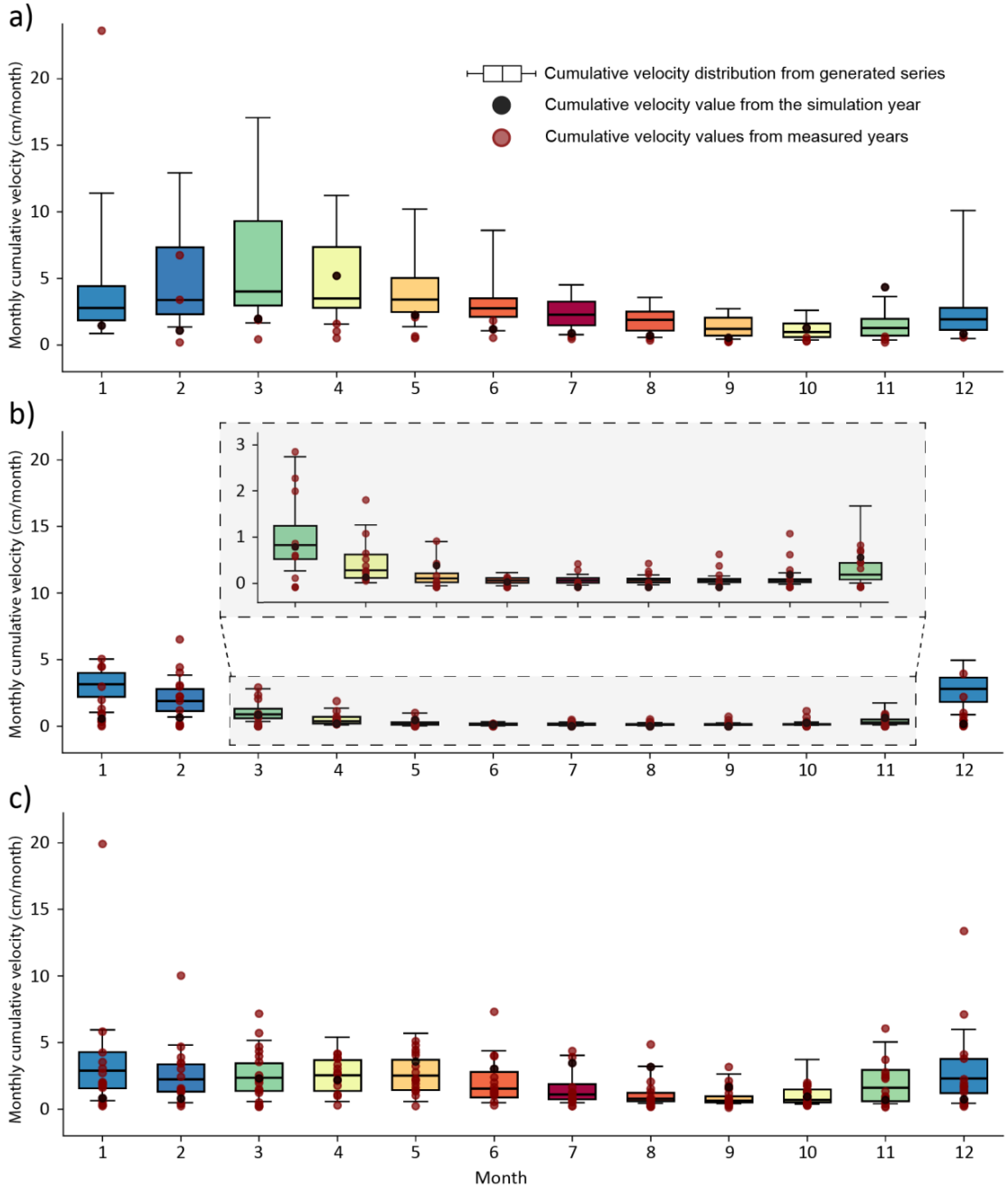


Figure 54 | Monthly cumulative velocity distribution for **(a)** Viella, **(b)** Villerville and **(c)** Vallcebre landslides. Observed monthly cumulative velocities are shown as red dots, and those corresponding to the target period as black dots. Whiskers represent the 5th–95th percentile range, consistent with the Mean Interval Score (MIS).

VI.2.4.2.2. Villerville monthly cumulative velocity distribution

Villerville monthly cumulative velocities also display marked seasonal pattern. Winter months show the highest median values (about 2-3 cm.month^{-1}) and the widest distributions, whereas from June to October distributions are very narrow and near zero. March and April show a progressive decrease in both median values and dispersion, while November marks the transition towards the active winter season (**Figure 54b**).

PICP values for the generated distributions range from 8.3% (September) to 91.7% (April), with the lowest coverage observed from July to September. Regarding MIS, winter months (December–February) exhibit the highest values (10.5-13.0 cm.month^{-1} ; **Table 21**). Although the MIS value partly reflects the scale of the underlying velocities, most of the uncertainty arises from a lack of coverage of the observed values rather than from overly wide prediction intervals. As shown in **Figure 54b**, the generated winter distributions tend to overestimate monthly cumulative velocities, although their upper bounds remain comparable to the observed maxima. In contrast, the generated distributions for July to October strongly underestimate the observed monthly cumulative velocities, with 90% of simulated values remaining close to zero, whereas some measured monthly cumulative velocities reach 0.5–1 cm.month^{-1} .

The underestimation in late summer and early autumn likely reflects combined limitations of the hydrological and XGBoost models. Simulated groundwater recession periods tend to decline faster than observed and to show reduced sensitivity to recharge events during these low-flow conditions. In addition, the model must compromise between the first years of the record, characterised by lower mean groundwater levels and weaker rainfall reactivity, and the more responsive behaviour observed later (**Figure 52b**), which can affect the optimal parameter set and the reproducibility of groundwater dynamics. Because velocity sensitivity at Villerville is mainly controlled by long-term groundwater variations (see ***section V.2.5.1.***), stronger simulated recessions, lower summer groundwater levels, and more episodic recharge responses translate into reduced simulated velocities and, therefore, underestimated monthly distributions during these months.

The overestimation of winter distributions is primarily related to XGBoost simulation limitations. Winter values is characterised by sharp, short-term velocity variations that are difficult to reproduce. Consequently, simulated velocities tend to follow smoother and more persistent trajectories rather than abrupt peaks, inflating monthly cumulative velocities. Notably, the lowest observed winter values correspond to the first monitoring years, when the landslide exhibited lower reactivity associated with smaller groundwater fluctuations and lower base levels compared with the later, more reactive period (**Figure 52b**).

Table 21 | Monthly cumulative velocity distribution ranges metrics for 90% prediction intervals. Results for the distributions obtained with the standard (S) and optimised (Opt) XGBoost models are shown for each site. N denotes the number of available measured values.

	Viella					Villerville					Vallcebre				
		PICP (%)		MIS (cm.month^{-1})			PICP (%)		MIS (cm.month^{-1})			PICP (%)		MIS (cm.month^{-1})	
Month	N	S	Opt	S	Opt	N	S	Opt	S	Opt	N	S	Opt	S	Opt
1	2	50.0	100.0	131.2	25.1	14	35.7	35.7	10.5	8.4	18	72.2	72.2	21.9	22.0
2	4	50.0	50.0	16.6	31.9	14	42.9	35.7	11.7	14.8	16	75.0	75.0	11.4	11.4
3	4	75.0	75.0	21.6	35.0	12	58.3	41.7	4.6	6.1	17	64.7	64.7	8.9	9.1
4	4	50.0	50.0	18.2	31.2	12	91.7	33.3	2.1	4.0	14	92.9	92.9	5.3	5.6
5	4	50.0	50.0	16.6	32.2	10	60.0	10.0	1.1	4.8	16	93.6	93.6	5.5	5.7
6	4	75.9	75.0	10.3	23.4	11	45.5	0.0	0.6	5.0	14	85.8	85.8	8.4	8.3
7	3	33.3	33.3	6.4	9.2	11	18.2	9.1	1.2	4.9	15	80.0	80.0	4.7	4.7
8	3	33.3	33.3	4.7	6.0	11	18.2	9.1	1.3	4.6	15	53.3	46.7	6.3	6.3
9	4	25.0	25.0	4.6	6.6	12	8.3	0.0	2.0	5.2	17	70.6	70.6	3.5	3.6
10	4	50.0	50.0	3.0	4.4	13	23.1	7.7	3.0	5.5	16	81.3	81.3	3.7	3.5
11	4	50.0	50.0	7.6	9.0	12	58.3	58.3	2.3	3.7	15	73.3	73.3	6.6	6.6
12	2	100.0	100.0	10.4	22.7	12	25.0	33.3	13.0	7.9	16	68.8	68.8	16.8	16.6

VI.2.4.2.3. Vallcebre monthly cumulative velocity distribution

Vallcebre monthly cumulative velocities distributions show a weaker seasonal contrast, consistent with the generated forcing series dynamics (**Figure 54c**). Median values are around 2 cm.month^{-1} , with cumulative velocities ranging from zero to about 7.0 cm.month^{-1} between December and May. Summer months are characterised by a gradual decrease in the cumulative velocity, with lower values around 1 cm.month^{-1} in September. Velocity increase during October and November, reflecting sensitivity to the recharge events and associated groundwater level responses.

According to PICP and MIS values, most observed monthly cumulative velocities fall within the generated distributions. MIS values exceed 10 cm.month^{-1} for December to February, which reflects larger discrepancies for these months. Although some measured values lie below the 90% prediction intervals, this mainly results from the strong contrast in velocities observed between 1996 and 1998, which are much higher than those represented by the generated distributions (**Figure 49c**). Conversely, the generated distributions tend to underestimate monthly cumulative velocities during the lowest-activity period, from August to September, with PICP values of 53.3% and 70.6%, respectively (**Table 21**). As for Villerville, simulated groundwater levels at Vallcebre show stronger recessions than observed and appear less responsive to late-summer recharge events (**Figure 22c**). Landslide's surface velocity is primarily driven by short-term groundwater level variations (see ***section VI.2.7*** ; **Figure 56**).

Given these biases, the resulting distributions are likely to underestimate the observed cumulative velocities during summer months, when short-duration, high-amplitude velocity events can occur.

VI.2.5. Discussion

VI.2.5.1. Effects of XGBoost hyperparameters optimisation

Limitations in the monthly cumulative velocity distributions are primarily driven by (i) the learning capacity of the XGBoost model and (ii) the accuracy of the simulated groundwater levels. To assess whether these distributions could be better constrained, we tested the effect of hyperparameter optimisation during model training (see ***section VI.2.7*** ; **Table 23**).

For Villerville and Vallcebre, the optimised models show a slight overall improvement over the test period, with marginally lower RMSE and MAE values (**Table 7**). However, at Villerville, although some spurious reactivity is reduced, simulated baseline velocities remain systematically above zero (see ***section VI.2.7*** ; **Figure 58**). This behaviour reflects a trade-off during optimisation process between local accuracy and global error minimisation. Errors at low-velocity contribute less to the objective than errors at high velocities, which can lead the optimiser to favour a small positive bias in the baseline level.

For Viella, hyperparameter optimisation increases the overall error over the test period (**Table 22**). This loss of performance is linked to the limited data availability and the small number of representative acceleration events. Although the optimised model better captures the April 2024 acceleration amplitude, it substantially overestimates the November 2024 event, likely because the exceptional January 2022 acceleration has a disproportionate influence on the learned response (see ***section VI.2.7*** ; **Figure 58**).

Table 22 | RMSE and MAE metric values between observed and simulated velocities over the test period, with and without XGBoost hyperparameter optimisation

	Viella		Villerville		Vallcebre	
	S	**Opt**	**S**	**Opt**	**S**	**Opt**
RMSE (cm.month^{1})	0.038	0.081	0.045	0.038	0.036	0.034
MAE (cm.month^{1})	0.029	0.037	0.027	0.023	0.023	0.020

Figure 55 shows the monthly cumulative velocity distributions obtained with the optimised models. Seasonal patterns remain very similar to those of the standard models at all three sites. At Viella, the increased sensitivity of the optimised model to extreme velocities is reflected in the distributions. From January to June, monthly cumulative velocity distributions become wider and extreme values are more frequent, indicating an over-reactive response to external forcing, whereas summer and autumn months remain essentially unchanged. With the exception of January, MIS values increase and penalise these wider generated distributions (**Table 21**).

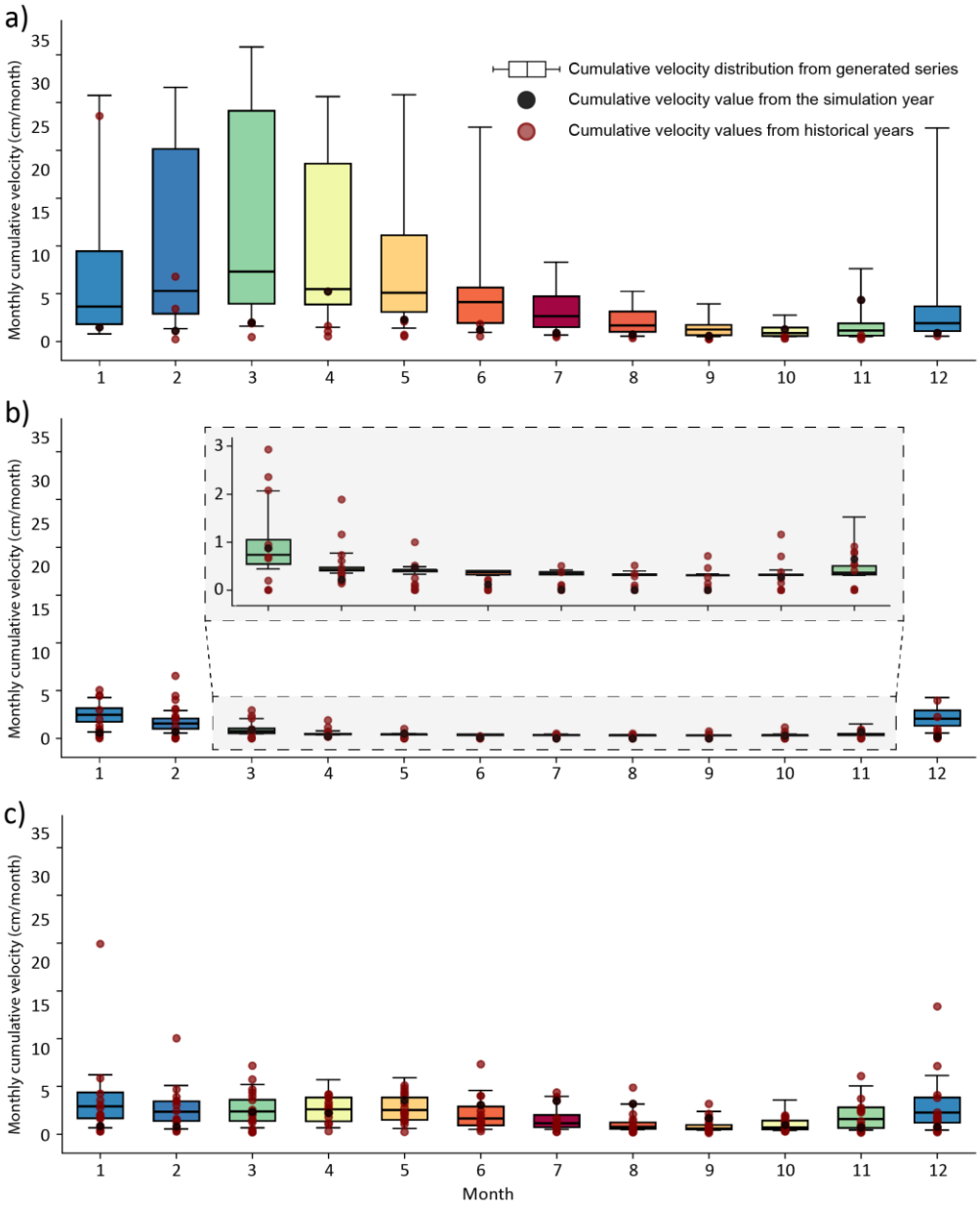


Figure 55 | Monthly cumulative velocity distribution for **(a)** Viella, **(b)** Villerville and (c) Vallcebre landslides, obtained from optimised XGBoost model.

At Villerville, December and January distributions are slightly reduced, with monthly velocities generally ranging from 1 to 3 cm.month^{-1}, which is consistent with the MIS values. In contrast, summer distributions are strongly affected by the systematically positive baseline velocities. As a consequence, most observed values lie below the simulated ranges (**Figure 55b**), which results in higher MIS and lower PICP values compared with the standard XGBoost simulations. For Vallcebre, hyperparameter optimisation does not yield any meaningful improvement in the monthly distributions, and PICP and MIS values remain similar (**Figure 55c ; Table 21**).

Overall, hyperparameter optimisation does not substantially refine the monthly cumulative velocity distributions. Limited representativeness of the training data (as at Viella) and the occurrence of sharp, short-term instability events (as at Villerville) both reduce the simulation robustness. At this stage, given the current data availability and representativeness, and the absence of clear gains even where longer series are available (Vallcebre case), the standard XGBoost configuration remains the most appropriate option. Future work could explore cross-validation rather than a single validation split to reduce the risk of hyperparameters tuning over a non-representative period.

VI.2.5.2. Groundwater modelling effects

Groundwater level fluctuations, over both short- and long-term time windows, play a key role in controlling velocity dynamics (see ***section V.2.5.1.***). Most limitations identified in this study, especially the underestimation of monthly cumulative velocity distributions, stem from a combination of factors: simulated groundwater levels that decrease too rapidly, reduced sensitivity to recharge episodes during recession, and, the resulting bias in the descriptive features derived from these biased series and used as inputs to the XGBoost models.

Transfer-function models are well suited to generic application across contrasting lithological contexts because they involve relatively few parameters and can be calibrated automatically on historical time series. However, their calibration relies on minimising a global error metric, which can lead to local mismatches, such as underestimated peak amplitudes or overly pronounced recession periods.

To improve groundwater simulation accuracy, alternative groundwater models, particularly semi-empirical lumped reservoir approaches, could be considered *(Angeli et al., 1998; G.-W. Lin et al., 2020; Weerasinghe et al., 2019)*. These models may provide a more realistic representation of groundwater dynamics and help refine the ranges of the monthly cumulative velocity distributions. However, they typically require more extensive calibration and site-specific parameterisation, which reduces their ease of use and limits their applicability in a fully generic framework across multiple landslide sites.

VI.2.6. Conclusion

Hydro-meteorological forcing are the primary control on surface velocities in slow-moving landslides. Although many studies use such forcing to predict observed velocity time series, the full probabilistic range of expected velocity over a defined period – at least at the monthly scale and under different plausible hydro-meteorological conditions – has not yet been explicitly addressed. To fill this gap, this paper proposes (i) a complete workflow for generating daily hydrometeorological series by coupling a modified Richardson-type weather generator with a transfer-function-based groundwater model, and (ii) an approach to estimate monthly cumulative surface velocity probability distributions over a full year, based on daily velocity simulations produced by a pre-trained XGBoost model. The framework is applied to three deep-seated landslides in contrasting lithological settings and therefore characterised by distinct hydro-meteorological – velocity relationships.

The generated meteorological time series reproduce the main properties and variability of the observed records with good accuracy. The calibrated groundwater level models capture the overall historical groundwater dynamics, although they tend to under-reproduce peak amplitudes and to simulate recession periods that are too rapid. Combining the generated meteorological and groundwater series yields 500 plausible scenarios of hydro-meteorological conditions over a full year for each site.

The resulting monthly cumulative velocity distributions offer valuable insight into both velocity periodicity and expected ranges. Depending on the landslide, seasonality is more or less pronounced, consistent with the observed behaviour and its dependence on hydrometeorological forcing. Compared with observed monthly cumulative velocities, the simulated distributions generally encompass most observations, although some over- and underestimation persist particularly in winter and late summer. These discrepancies mainly reflect limitations in groundwater level modelling and the limited length and representativeness of the available velocity records to train the XGBoost model. Semi-empirical groundwater models with more physically based calibration could help refine the distribution ranges, but at the expense of reduced generic applicability across sites.

Overall, providing access to probabilistic monthly distributions of cumulative surface velocities under a broad ensemble of plausible hydro-meteorological conditions offers a potentially valuable tool for operational landslide monitoring and early warning systems.

VI.2.6.1. Acknowledgments

The authors are grateful to Jérôme Latron (Institute of Environmental Assessment and Water Research, IDAEA-CSIC, Barcelona) and Marcel Hürlimann (Department of Civil and Environmental Engineering, UPC BarcelonaTECH, Barcelona, Spain) for their collaboration and for sharing information and data on the Vallcebre landslide. We also thank Joshua Ducasse for acquiring field data on the Villerville and Viella landslides. Finally, we acknowledge the OMIV Observatory for providing monitoring data for the Villerville and Viella landslides.

VI.2.6.2. Data availability

The hydro-meteorological and velocity data for the Viella and Villerville landslides are available on the OMIV website (http://bdoh.unice.fr/OMIV-METEOROLOGY/) and can be cited using the following DOI: doi:10.25666/DATAOSU-2016-06-27. Additional meteorological observations were retrieved from the Météo-France open data portal (https://donneespubliques.meteofrance.fr/), specifically from the Luz-Saint-Sauveur and Saint-Gatien-des-Bois stations. The hydro-meteorological datasets for the Vallcebre landslide are available and can be cited using the following DOIs : https://doi.org/10.5281/zenodo.17207473 and https://doi.org/10.5281/zenodo.17207888.

VI.2.6.3. Declarations

VI.2.6.3.1. Funding

The French National Landslides Observatory (SNO OMIV) and the Observatory of Earth Sciences (OSU-THETA) from the National Centre for Scientific Research (CNRS) funded the installation of the monitoring system of the Viella and Villerville landslides. At the Vallcebre landslides, the installation and long-term operation of the monitoring equipment were funded mainly through European Commission and Spanish national research projects (e.g., TESLEC, NEWTECH, SAFELAND; CICYT, MODEVALL) with additional support from Catalan and Interreg POCTEFA/ERDF programmes (PYRMOVE).

VI.2.6.3.2. Conflict of Interest

The authors declare no competing interests.

VI.2.7. Extended data

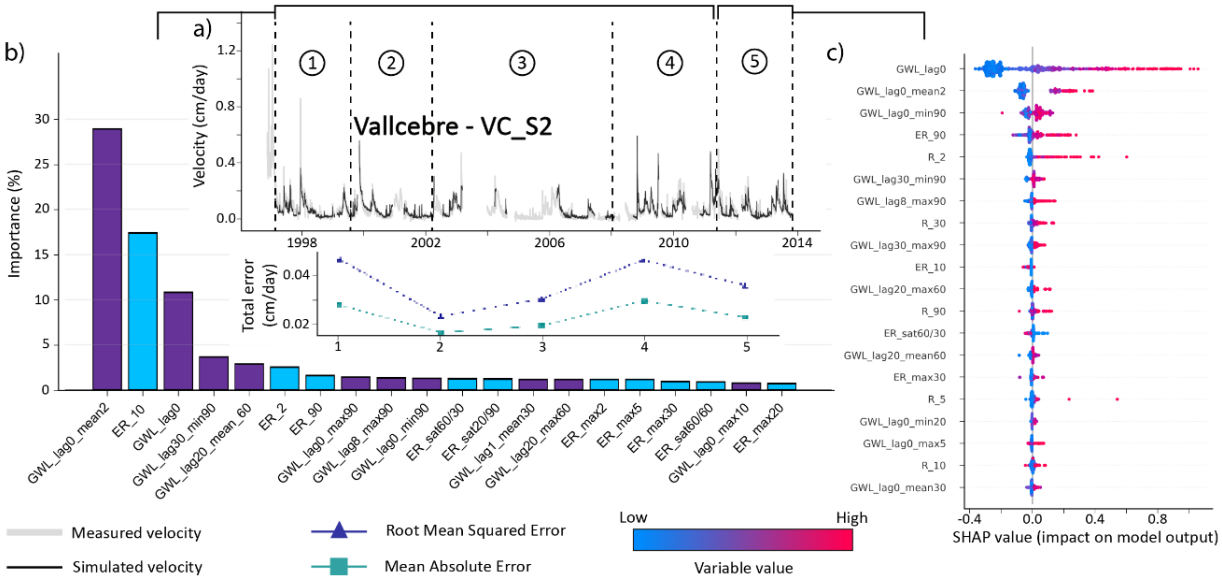


Figure 56 | Results of the XGBoost model applied to borehole VC_S2 at the Vallcebre landslide. **(a)** Simulated velocities (black) compared with observed velocities (grey) for each test block representing 20% of the data (Blocks 1 to 5). The total error for each test block is quantified using RMSE (triangles) and MAE (squares). **(b)** Ranking of the 20 most important features in explaining the variance of the model. Feature colours refer to their source time series (light blue for effective rainfall, purple for groundwater levels). **(c)** SHAP analysis of the effects of the 20 most important features on each velocity value in the test set, expressed as Shapley values. Feature values are shown using a blue-to-red colour scale, while their influence on the model output is represented along the x-axis. According to the feature importance ranking and SHAP analysis, short-term features related to effective rainfall and groundwater levels are the most influential in controlling velocities. In particular, the groundwater level on the current day exerts the strongest influence on the model predictions for new data. For further details on the method, refer to ***section V.2.3.2.***

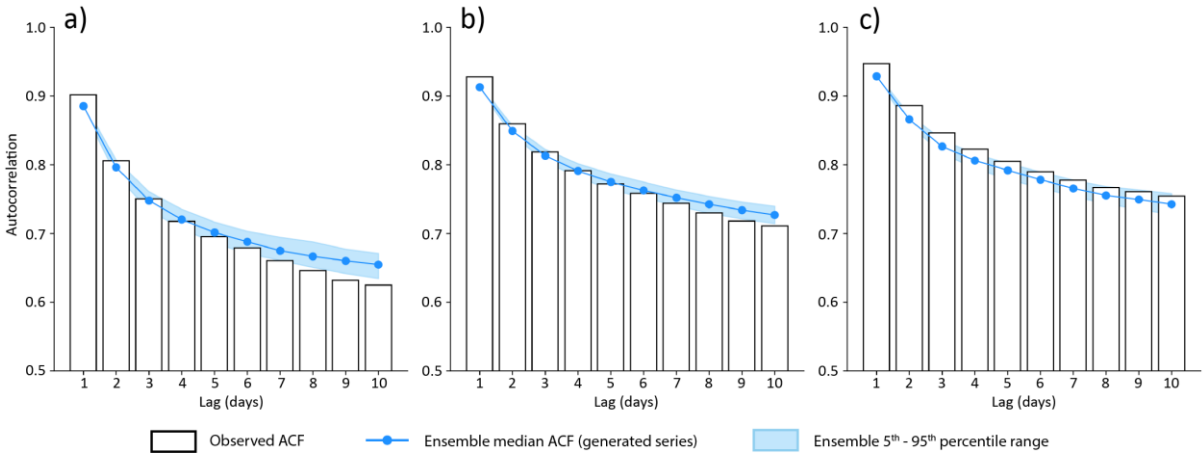


Figure 57 | Autocorrelation functions (ACF) of the observed series and the generated ensemble for lags up to 10 days at **(a)** Viella, **(b)** Villerville, and **(c)** Vallcebre.

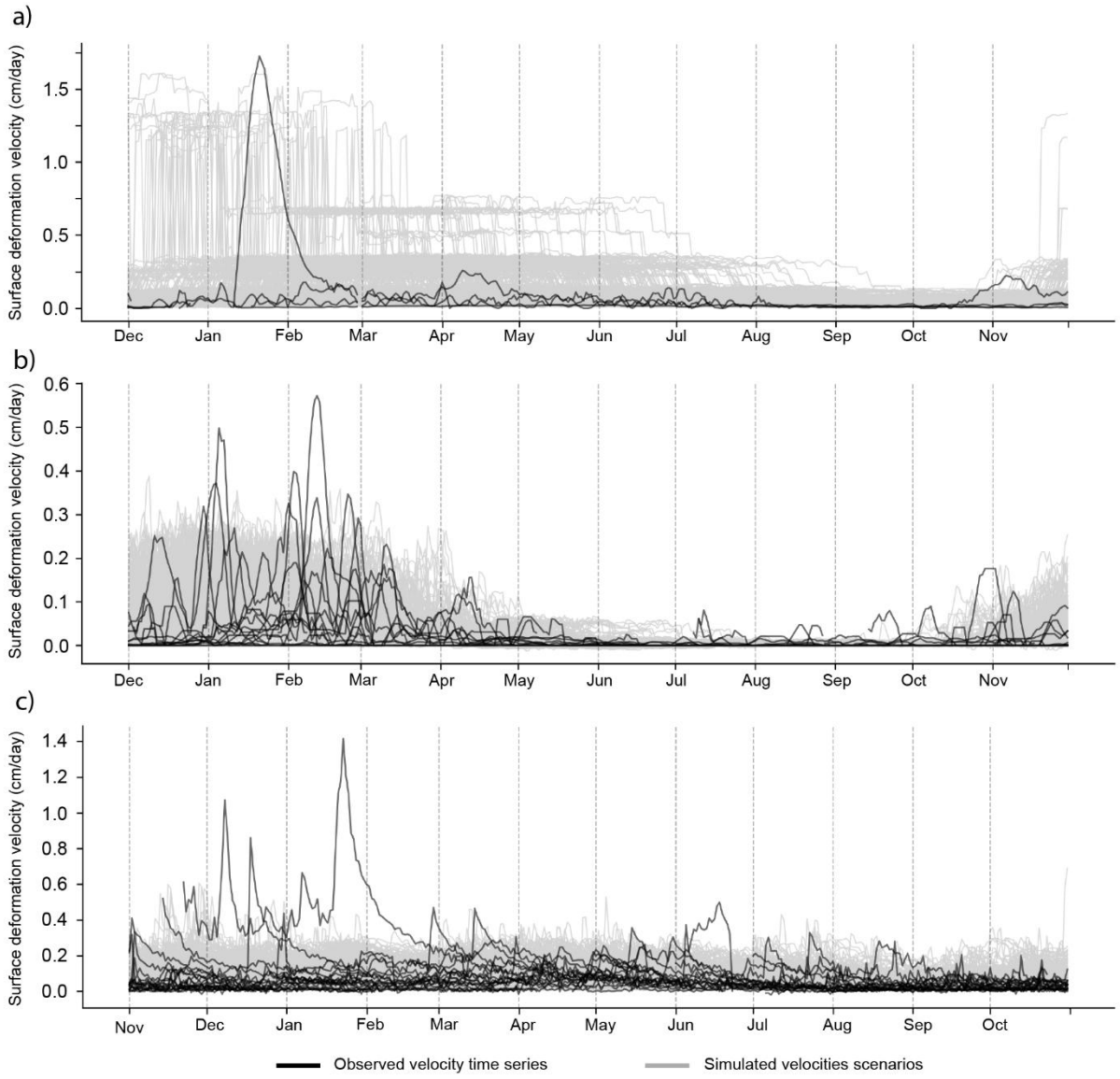


Figure 58 | Simulated daily velocity series derived from 500 hydro-meteorological scenarios (grey), together with observed daily velocities for all available measurement years at **(a)** Viella, **(b)** Villerville and **(c)** Vallcebre.

Table 23 | Optuna-optimised XGBoost hyperparameters and corresponding optimal values obtained on the training–calibration sets.

			Optimised value		
Parameter	**Standard value**	**Range**	**Viella**	**Villerville**	**Vallcebre**
N_estimators	100	50 to 2000	929	324	1506
Max_depth	6	2 to 10	10	5	7
Learning_rate	$3.0\ 10^{-1}$	$1.0\ 10^{-3}$ to $2.0\ 10^{-1}$	$1.219\ 10^{-2}$	$4.631\ 10^{-2}$	$4.811\ 10^{-2}$
Subsample	1.0	$5.0\ 10^{-1}$ to 1.0	$9.284\ 10^{-1}$	$5.195\ 10^{-1}$	$8.159\ 10^{-1}$
Colsample_bytree	1.0	$5.0\ 10^{-1}$ to 1.0	$9.098\ 10^{-1}$	$8.673\ 10^{-1}$	$5.562\ 10^{-1}$
Min_child_weight	1	1 to 10	6	5	5
Reg_alpha	0	$1.0\ 10^{-8}$ to $1.0\ 10^{1}$	$1.0958\ 10^{-3}$	$8.283\ 10^{-5}$	$3.311\ 10^{-4}$
Reg_lambda	1	$1.0\ 10^{-4}$ to $1.0\ 10^{1}$	$1.205\ 10^{-1}$	$5.858\ 10^{-1}$	$4.815\ 10^{-3}$
gamma	0	0 to 5	$2.086\ 10^{-2}$	$1.082\ 10^{-1}$	$1.149\ 10^{-1}$

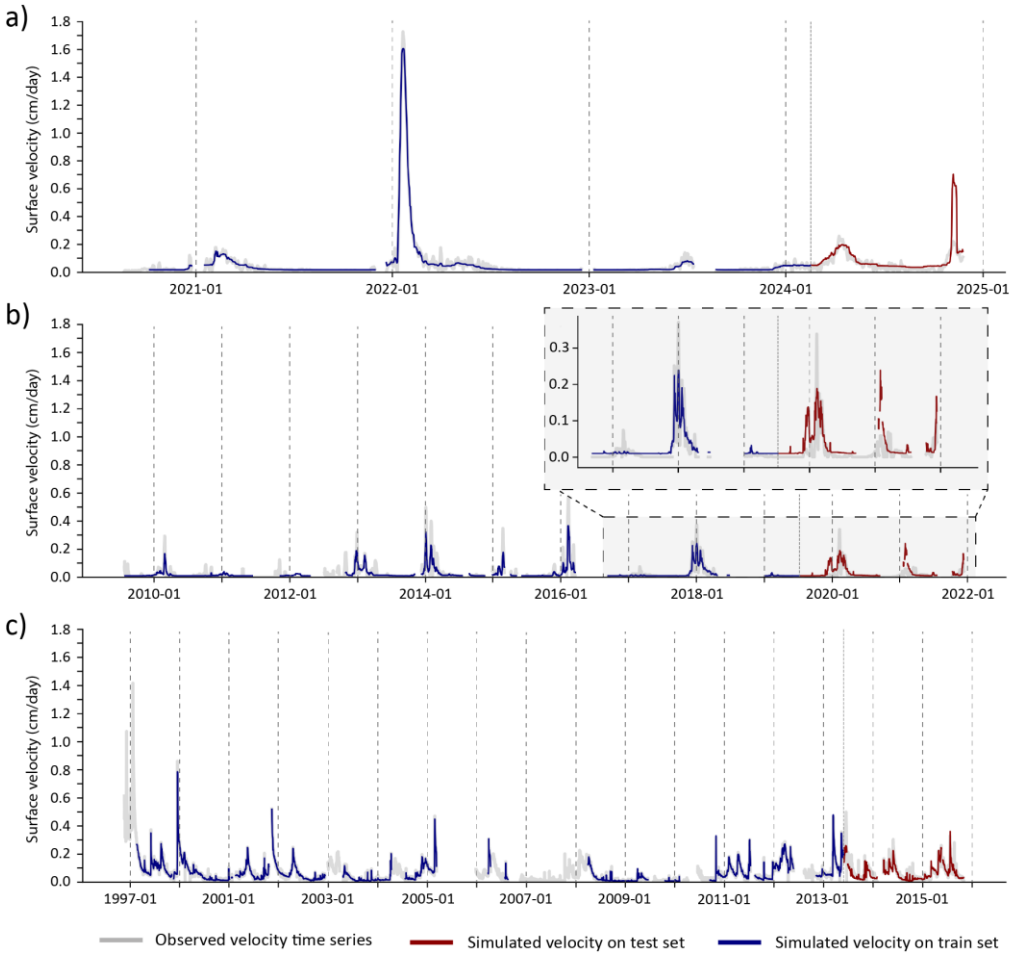


Figure 59 | Simulated velocity time series on the training and testing datasets obtained with the optimised XGBoost model for **(a)** Viella, **(b)** Villerville, and **(c)** Vallcebre.

VI.3. CHAPTER VI – CONCLUSION

This chapter addressed the third research question of the thesis and the probabilistic extension of the methogological framework introduced in ***Chapter III***. Its objective was to quantify how slow-moving landslide displacement evolves under contrasted and plausible hydro-meteorological conditions by propagating forcing variability through the velocity simulation chain. Rather than evaluating model performance under a single historical realisation of forcing, the chapter aims to derive probabilistic displacement ranges that capture the expected variability in monthly velocity.

Chapter VI therefore complemented the deterministic workflow by introducing a scenario-based layer that samples daily rainfall and air temperature variability, simulates coherent groundwater level responses, and applies the same generic hydro-meteorological predictors and pre-trained XGBoost models applied in ***Chapter V***, to generate ensembles of daily velocity time series. In this chapter, groundwater information is essentialy provided by monitored groundwater level records, which offer a direct description of the slope hydrological state and provide a robust basis to evaluate the probabilistic framework under well-constrained conditions.

The results show that daily scenario ensembles can be translated into monthly cumulative velocitiy probability distributions that reproduce contrasted seasonal signatures and provide an explicit quantification of expected month-to-month variability. The comparison with historical monthly cumulative velocities indicates that the distributions generally encompass observed values while also revealing residual misfits during specific periods. These mismatches emphasise a key outcome of the chapter: probabilistic displacement ranges remain conditioned by the consistency of the simulated groundwater level trajectories and by the representativeness of the training datasets used to learn forcing–response relationships.

In the perspective of the thesis framework, ***Chapter VI*** extends the contributions of ***Chapter V*** towards hazard-oriented applications by shifting from deterministic simulations to probabilistic displacement ranges under plausible forcing variability. By coupling ensemble hydro-meteorological scenarios with the transparent forcing–response constraints obtained through conventional machine learning models, the approach provides a structured basis to interpret scenario sensitivity and seasonal contrasts in a process-oriented manner. More broadly, this framework offers practical perspectives for operational monitoring, including the production of probabilistic displacement products for decision support, and for prospective applications in which contrasted forcing scenarios are explored to anticipate potential evolutions in slow-moving landslide activity.

CHAPTER VII

CONCLUSION AND OUTLOOKS

CHAPTER VII CONCLUSION AND OUTLOOKS 169
VII.1. General conclusion 171
VII.2. Discussion and outlooks 173
VII.2.1. Extending the methodology to other deep-seated landslides 173
VII.2.2. Using satellite products to address data gaps 174
VII.2.3. Towards spatio-temporal prediction of surface deformation velocities 175
VII.2.4. Advances in interpretability and explainability of AI models for time series 176
VII.2.5. Open-access contribution 178

VII.1. GENERAL CONCLUSION

This thesis set out to develop and assess modelling strategies to better characterise, understand, and simulate the surface dynamics of deep-seated, slow-moving landslides and their complex response to hydro-meteorological forcing across contrasted sites and geological settings. Building on the state of the art reviewed in ***Chapter II***, the manuscript adopted the structured methodological framework introduced in ***Chapter III***, in which robust displacement simulation relies on (i) long, informative and representative time series, (ii) predictors that explicitly encode temporal dependence and system memory, and (iii) modelling approaches whose results can be examined beyond purely predictive skill.

The work was structured around three research questions that follow the successive needs of this workflow.

The first research question focused on the upstream requirement that conditions any robust application of conventional machine-learning models: **How can long, informative and representative multi-source time series be acquired – or reconstructed – to support robust AI model training in slow-moving landslides?** ***Chapter IV*** addressed this issue by targeting the limiting availability of groundwater level information, which is critical to capture subsurface storage and delayed responses. By developing semi-empirical, rainfall–discharge-based conceptual tank models calibrated on outlet discharge, the chapter showed that the temporal evolution of the internal hydrological state can be inferred even when long, continuous piezometric records are sparse or unavailable. Beyond gap-filling, these reservoir-based reconstructions provide a parsimonious, site-consistent representation of storage–release dynamics in heterogeneous fractured media, thereby offering transferable groundwater state proxies that can complement monitoring datasets and support subsequent feature engineering and modelling stages.

The second research question then addressed the core modelling and interpretation stages of the thesis framework: **How do hydro-meteorological controls shape the temporal dynamics of surface displacement, and over which characteristic time scales?** ***Chapter V*** answered this question by developing and applying a transferable conventional machine-learning workflow based on XGBoost and a generic set of physically-grounded predictors derived from net rainfall, effective rainfall and groundwater level variations. The multi-site and multi-plot applications demonstrated that daily surface-velocity dynamics can be reproduced across contrasted settings without relying on site-dependent predictors or explicit encoding of local geological structures, while still acknowledging that forcing–response signatures may vary within a single unstable mass. Importantly, the chapter established that transparency can be systematically integrated into the workflow: combining native XGBoost feature importance ranking with post-hoc SHAP analysis makes the inferred forcing–response relationships

explicit in terms of dominant controls, non-linear behaviours, and response time scales, thereby providing physically consistent constraints that can be compared across sites.

Finally, the third research question extended these deterministic and interpretable simulations towards hazard-oriented applications: **How does slow-moving landslide displacement evolve under contrasted and plausible hydro-meteorological conditions?** ***Chapter VI*** addressed this question by introducing a scenario-based probabilistic layer that explicitly samples forcing variability and propagates it through the velocity simulation chain. By coupling stochastic generation of daily rainfall and air temperature series, coherent groundwater level simulations driven by scenario-based effective rainfall, and the same generic predictors and pre-trained XGBoost models as in ***Chapter V***, the chapter translated daily ensembles into monthly cumulative velocities distribution ranges. These distributions quantify expected month-to-month variability, preserve site-specific seasonal signatures, and provide a direct basis for hazard mitigation by expressing potential displacement as probability ranges rather than single values.

Taken together, the thesis demonstrates that combining **(i) conceptual hydrological modelling to reconstruct hydrological-state information, (ii) a transferable and transparent XGBoost-based workflow to simulate daily velocity dynamics and extract forcing–response constraints, and (iii) probabilistic scenario propagation to derive displacement ranges under plausible forcing variability, provides a coherent framework to move from monitoring time series to hazard-relevant products.** Beyond their respective applications, the three methodological components are complementary: groundwater state reconstructions strengthen the data foundations required for robust modelling, interpretable predictors and explainability analyses make learned relationships examinable and comparable, and probabilistic simulations translate these constraints into scenario-based displacement ranges aligned with hazard-mitigation perspectives. This integrated approach therefore offers a scalable pathway to support both scientific understanding of hydro-meteorologically driven landslide dynamics and the development of operationally meaningful products for monitoring and prospective assessments.

VII.2. DISCUSSION AND OUTLOOKS

The developed approaches in this thesis rely on a series of specific methodological choices designed to address the questions raised about deep-seated landslide behaviour and its hydro-meteorological controls. At the same time, the rapid advances in artificial intelligence and in hydrogeological modelling continuously open new avenues to refine, extend and challenge the proposed frameworks. The results obtained here also reveal a number of limitations – related to data availability, model structure, and transferability – that delimit the current scope of application of these methods. Taken together, these elements provide a basis for a broader discussion on the robustness and generality of the proposed approaches, and point towards several promising directions for future developments (**Figure 60**).

VII.2.1. Extending the methodology to other deep-seated landslides

The methodology developed in this thesis relies on long, multi-parametric time series acquired from dense in situ monitoring networks, which allow for detailed characterisation of both hydrological forcing and landslide response. Such monitoring configurations remain exceptional and are currently limited to a small number of well-studied deep-seated landslides. Extending the proposed framework to a broader set of sites – ultimately with the aim of building a generic forecasting tool – will therefore have to confront with shorter, incomplete, or discontinuous records of deformation and forcing.

In this context, several recent studies highlight transfer learning as a promising strategy to mitigate data scarcity. *Long et al. (2022)*, for instance, introduced a multi-feature fusion transfer learning (MFTL) approach in which a model is first trained on a well-instrumented reservoir-induced landslide and then transferred to other landslides exhibiting similar deformation characteristics, yielding a clear performance gain compared with models trained independently on each site. Similarly, *H. Zheng et al. (2022)* and *J. Ma et al. (2025)* employed Bi-GRU or CNN–BiLSTM architectures pre-trained on one landslide and subsequently fine-tuned on a second site, maintaining satisfactory predictive accuracy despite short or discontinuous time series. More recently, *C. Cheng et al. (2025)* extended this principle to several hundred rainfall-induced landslides by combining static factors (lithology, geometry, climate) and dynamic variables (rainfall, displacement) within a Temporal Fusion Transformer and applying transfer learning between groups of landslides sharing similar characteristics.

These approaches, however, remain largely confined to relatively homogeneous regional contexts (e.g. reservoir-induced landslides in the Three Gorges area), and the question of transferability across contrasting lithologies, climatic regimes and monitoring configurations is still only weakly explored. A possible way forward would be to develop forecasting models by groups of landslides

grouped according to lithology and deformation mechanism. Within each family, models could be trained on several well-instrumented sites and then adapted by transfer learning to sparsely monitored slopes, while explicitly accounting for differences in climate and hydrology.

VII.2.2. Using satellite products to address data gaps

The approaches developed in this thesis primarily rely on in situ measurements from monitoring networks installed on unstable slopes. However, such dense and long-term monitoring systems remain rare and are typically restricted to a limited number of well-documented landslides. In many mountainous and hard-to-access regions (e.g., the Andes, the Himalaya), where human activities are often concentrated and landslide hazard is high, ground-based monitoring is limited or even absent. Yet, in these poorly monitored areas, it remains crucial to be able to assess hazard levels.

To address these limitations, the use of satellite-based products represents a particularly promising avenue. So far, most AI applications in landslide science have used satellite data mainly for susceptibility mapping *(Peiro et al., 2024)* or for the detection and delineation of unstable areas *(Boehm et al., 2022)*. In contrast, their exploitation as time series of predictive forcings and deformation observables for landslide-specific forecasting models remains largely unexplored, despite the strong potential offered by satellite-derived hydro-meteorological products at regional to global scales.

The accuracy of satellite forcing products is sensitive to orography, climate regime and season *(Stampoulis & Anagnostou, 2012)*, but their complementarity is a major asset. Combining satellite-based precipitation estimates (e.g., SAFRAN, CHIRPS, PERSIANN, GPM), snow products (e.g., MODIS, Sentinel-3) and soil-moisture datasets (e.g., SMAP, ASCAT) can provide daily time series at spatial resolutions on the order of 0.05°–0.25°, suitable for regional-scale landslide studies. In parallel, the increasing availability of multi-sensor satellite imagery (SAR and optical), together with mature processing chains *(Provost et al., 2022; Stumpf et al., 2017)*, and metre- to decametre-resolution digital elevation models (e.g., Pleiades, SPOT DEM) offers powerful means to compensate both for missing deformation data and for incomplete geomorphological information.

In combination, these satellite products could be integrated into the type of workflow proposed in this thesis by (i) providing spatially distributed hydro-meteorological forcings where gauge networks are sparse, and (ii) supplying deformation proxies and high-resolution morphological descriptors to constrain or calibrate AI-based models. Such developments would extend the applicability of the present methodology beyond intensively monitored reference sites, towards data-poor regions where landslide hazard is high but ground observations are limited or absent.

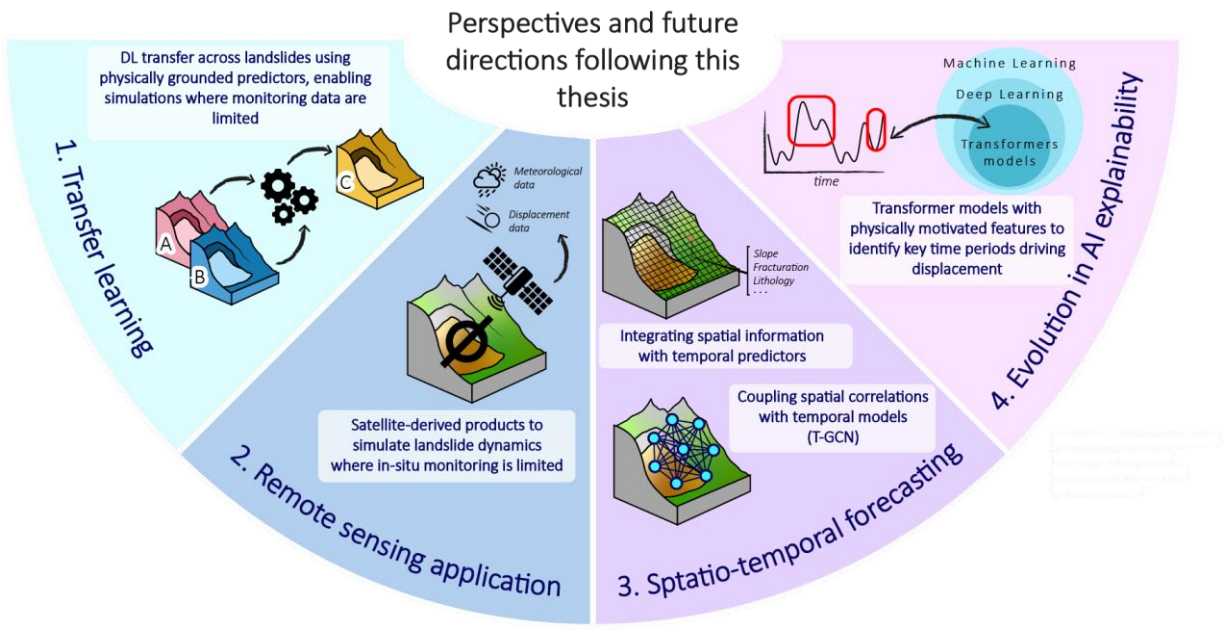


Figure 60 | Schematic overview of future research directions.

VII.2.3. Towards spatio-temporal prediction of surface deformation velocities

The velocity-forecasting methodology developed in this thesis is applied at the scale of individual monitoring stations. Even when several stations are instrumented on the same slope, the models remain essentially point-based: each simulation describes the temporal evolution of deformation at a given location, without explicitly accounting for spatial structure or interactions between stations.

Yet, forecasting the behaviour of deep-seated landslides is inherently a spatio-temporal problem. The spatial distribution of deformation, although heterogeneous, is controlled by local geomorphological and geological conditions, the internal architecture of the landslide mass, and the connectivity of fractured and permeable zones. Ignoring these spatial dependencies limits the ability of station-based models to provide a consistent slope-scale characterisation of landslide dynamics.

In this context, the explicit integration of spatial information has recently motivated the use of new modelling strategies, in particular Graph Convolutional Networks (GCN) *(J. Li et al., 2025; J. Wang et al., 2025; C. Yang et al., 2024)*. In GCN-based frameworks, the monitoring network is represented as a graph, where stations are nodes connected through edges defined by an adjacency matrix. For each node, the model aggregates information from neighbouring nodes via graph convolutions to learn latent representations that explicitly encode the topological structure of the slope.

For example, *Y. Jiang et al. (2022)* proposed a Temporal GCN (T-GCN) that combines a GCN module, used to capture spatial correlations in forcing and deformation, with a GRU module, used to represent temporal dependencies, thereby providing holistic predictions over the full station network. *J. Li et al. (2025)* further developed a "geology-aware" GCN approach, in which the graph structure and edge weights account not only for spatial proximity but also for geological similarity, thereby better capturing the physical dependencies that influence landslide dynamics. Across these studies, models that integrate spatial correlations systematically outperform classical LSTM or GRU architectures that treat stations independently and only consider temporal forcing–response relationships.

In light of these advances, incorporating spatial characteristics into the deformation-forecasting process appears as a logical and necessary evolution if the objective is to move towards truly slope-scale predictive models. A first step, compatible with the XGBoost-based framework developed in this thesis, would be to enrich the current set of hydro-meteorological time-series predictors with static spatial descriptors derived from geomorphological analyses (e.g., lithology, slope, curvature, degree of fracturing), similar to variables commonly used in susceptibility mapping. This would preserve the interpretability of tree-based models while introducing an explicit spatial context at each station.

A second step would consist in fully accounting for multi-station spatial correlations by adopting emerging GCN-type architectures. In this perspective, hybrid approaches could be explored, in which the physically motivated hydro-meteorological features developed here are embedded within graph-based deep learning models, or where GCN-derived latent spatial representations are coupled with XGBoost-type modules. Such developments would extend the present station-wise framework towards spatio-temporal forecasts at the scale of unstable slopes.

VII.2.4. Advances in interpretability and explainability of AI models for time series

A central objective of this thesis was to design an AI framework that is not only accurate, but also as transparent as possible: interpretable in terms of its decision process and explainable in terms of how hydro-meteorological forcing drives observed and simulated deformation. By combining tree-based models with hydro-meteorological features, it was possible to identify (i) which characteristics had the greatest influence during training and to what extent they contributed positively or negatively to the predictions, and (ii) where the models reached their limits, as illustrated for instance by the Viella landslide through feature importance and SHAP analyses.

This explicit search for transparency motivated the choice of decision tree ensembles such as XGBoost, rather than purely deep learning architectures, and the systematic use of post-hoc explainability tools to explore forcing–response relationships.

The rapid development of AI in all fields has been accompanied by an equally intense effort to improve model transparency and mitigate the "black-box" character of complex architectures. As emphasised by *Gu et al. (2024)*, eXplainable AI (XAI) for time series is especially challenging compared with text or image applications, due to the sequential structure of the data and the strong temporal dependence between observations and features. Developing explanation methods explicitly tailored to time series therefore remains an open and active research area.

In this context, *Gu et al. (2024)* proposed a taxonomy of XAI methods adapted to time series and assessed their relative effectiveness. The SHAP-based analyses used in this thesis belong to one of the most widely adopted families of model-agnostic methods, notably because they can be applied to a broad range of models and offer intuitive and visual representations. Other explanation strategies are gaining momentum for time series models, including distribution-aware approaches, decomposition-based methods and attention-based scores. Decomposition-based explanations and attention mechanisms have in particular emerged as promising tools to enhance the interpretability of deep learning architectures.

Transformer-type models, which underlie many modern generative AIs, rely on attention mechanisms rather than recurrent neural networks and are naturally suited to sequence data. Their attention scores can in principle serve as indicators of importance, by highlighting which time steps or features the model focuses on when issuing a prediction. At the same time, the ongoing "attention is (not) explanation" debate underscores that attention weights do not necessarily constitute faithful or sufficient explanations, and that additional validation of explanation quality is required.

In the field of landslide deformation forecasting, transparency and explainability in deep learning models have only recently started to be explored. *J. Wang et al. (2025)*, for example, combined a Temporal GCN with an attention module to quantify the relative importance of different forcing variables in the simulated deformation. *Z. Ma & Mei, (2025)* used transformer-based architectures to identify internal and external factors controlling deformation, while simultaneously accounting for spatial correlations along unstable slopes.

Overall, XAI models for landslide deformation evolution remains a very active research topic. Future work could build on the physically motivated hydro-meteorological features developed in this thesis, while integrating more advanced XAI methods specifically adapted to sequential and spatio-temporal models. As deep learning frameworks become more transparent, transformer-type models may become increasingly attractive for landslide forecasting – provided that the predictive features remain physically meaningful.

VII.2.5. Open-access contribution

Beyond the methodological developments presented in this thesis, a key enabling contribution lies in strengthening open-access practices for both data and models. Landslide hazard mitigation ultimately requires robust and reliable forecasting tools, yet – as highlighted throughout this manuscript – such robustness is conditional on the availability of long, informative, and representative multi-source time series. In this context, the landslide research landscape benefits from a particularly favourable complementarity between academic and operational stakeholders, which makes it possible to sustain high-quality spatio-temporal monitoring networks and to translate them into broadly reusable datasets.

National observatory initiatives such as SNO-OMIV contribute directly to this objective by ensuring long-term instrumentation maintenance on key sites and by prioritising the dissemination of open-access time series that document landslide kinematics, hydrogeological functioning and meteorological forcings. In line with these efforts, the work developed in this thesis reinforces both accessibility and transparency by making available the constructed datasets, the processing pipelines used to consolidate and analyse monitoring series, and the modelling codes required to reproduce the generic predictor set and the AI-based simulations, thereby supporting reproducibility, cross-site comparisons, and future extensions towards operational and probabilistic applications.

REFERENCES BILIBOGRAPHIQUES

A

Abi Nader, A., Albaric, J., Steinmann, M., Hibert, C., Malet, J.-P., Sue, C., Fores, B., Marchand, A., Gros, M., Celle, H., Pohl, B., Stefani, V., & Boetsch, A. (2023). Machine learning prediction of groundwater heights from passive seismic wavefield. *Geophysical Journal International*, *234*(3), 1807-1818. https://doi.org/10.1093/gji/ggad160

Aggarwal, A., Alshehri, M., Kumar, M., Alfarraj, O., Sharma, P., & Pardasani, K. R. (2020). Landslide data analysis using various time-series forecasting models. *Computers & Electrical Engineering*, *88*, 106858. https://doi.org/10.1016/j.compeleceng.2020.106858

Agliardi, F., Crosta, G. B., Frattini, P., & Malusà, M. G. (2013). Giant non-catastrophic landslides and the long-term exhumation of the European Alps. *Earth and Planetary Science Letters*, *365*, 263-274. https://doi.org/10.1016/j.epsl.2013.01.030

Agliardi, F., Crosta, G. B., Zanchi, A., & Ravazzi, C. (2009). Onset and timing of deep-seated gravitational slope deformations in the eastern Alps, Italy. *Geomorphology*, *103*(1), 113-129. https://doi.org/10.1016/j.geomorph.2007.09.015

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716-723. https://doi.org/10.1109/TAC.1974.1100705

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2623–2631). ACM. https://doi.org/10.1145/3292500.3330701

Alberti, S. (2018). *Geological analysis and numerical modelling of La Saxe landslide (Courmayeur) to improve understanding of geomorphological and geotechnical mechanisms and of the potential landslide evolution* (Doctoral thesis, Università degli Studi di Milano-Bicocca).

Alhubail, A., Fahs, M., Lehmann, F., & Hoteit, H. (2024). Modeling fluid flow in heterogeneous porous media with physics-informed neural networks : Weighting strategies for the mixed pressure head-velocity formulation. *Advances in Water Resources*, *193*, 104797. https://doi.org/10.1016/j.advwatres.2024.104797

Al-Najjar, H. A. H., Pradhan, B., Beydoun, G., Sarkar, R., Park, H.-J., & Alamri, A. (2023). A novel method using explainable artificial intelligence (XAI)-based Shapley Additive Explanations for spatial landslide prediction using Time-Series SAR dataset. *Gondwana Research*, *123*, 107-124. https://doi.org/10.1016/j.gr.2022.08.004

Amatya, P., Scheip, C., Déprez, A., Malet, J.-P., Slaughter, S. L., Handwerger, A. L., Emberson, R., Kirschbaum, D., Jean-Baptiste, J., Huang, M.-H., Clark, M. K., Zekkos, D., Huang, J.-R., Pacini, F., & Boissier, E. (2023). Learnings from rapid response efforts to remotely detect landslides triggered by the August 2021 Nippes earthquake and Tropical Storm Grace in Haiti. Natural Hazards, 118(3), 2337-2375. https://doi.org/10.1007/s11069-023-06096-6

Angeli, M. G., Buma, J., Gasparetto, P., & Pasuto, A. (1998). A combined hillslope hydrology/stability model for low-gradient clay slopes in the Italian Dolomites. *Engineering Geology*, *49*(1), 1-13. https://doi.org/10.1016/S0013-7952(97)00033-1

B

Bai, D., Lu, G., Zhu, Z., Zhu, X., Tao, C., Fang, J., & Li, Y. (2022). Prediction Interval Estimation of Landslide Displacement Using Bootstrap, Variational Mode Decomposition, and Long and Short-Term Time-Series Network. *Remote Sensing*, *14*(22), 5808. https://doi.org/10.3390/rs14225808

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1-22. https://doi.org/10.1002/jae.659

REFERENCES BILBIOGRAPHIQUES

Ballantyne, C. K. (2002). Paraglacial geomorphology. *Quaternary Science Reviews, 21(18–19)*, 1935–2017. https://doi.org/10.1016/S0277-3791(02)00005-7

Barla, G., Antolini, F., Barla, M., Mensi, E., & Piovano, G. (2010). Monitoring of the Beauregard landslide (Aosta Valley, Italy) using advanced and conventional techniques. *Engineering Geology*, *116*(3-4), 218-235. https://doi.org/10.1016/j.enggeo.2010.09.004

Baudement, C., Arfib, B., Mazzilli, N., Jouves, J., Lamarque, T., & Guglielmi, Y. (2017). Groundwater management of a highly dynamic karst by assessing baseflow and quickflow with a rainfall-discharge model (Dardennes springs, SE France). *Bulletin de La Société Géologique de France*, *188*(6), 40. https://doi.org/10.1051/bsgf/2017203

Bednarczyk, Z. (2018). Identification of flysch landslide triggers using conventional and 'nearly real-time' monitoring methods – An example from the Carpathian Mountains, Poland. *Engineering Geology*, *244*, 41-56. https://doi.org/10.1016/j.enggeo.2018.07.012

Béjean-Maillard, O., Bertrand, C., Malet, J.-P., Cinkus, G., Nevers, P., Ducasse, J., Bertolo, D., & Thuegaz, P. (2025). Hydrogeological forecasting of deep-seated landslides dynamics: Structure and sensitivity of tank models. *Landslides, 22(7)*, 2199–2217. https://doi.org/10.1007/s10346-025-02482-2

Bekesiene, S., Smaliukiene, R., & Vaicaitiene, R. (2021). Using Artificial Neural Networks in Predicting the Level of Stress among Military Conscripts. *Mathematics*, *9*(6), 626. https://doi.org/10.3390/math9060626

Benoit, L., Briole, P., Martin, O., Thom, C., Malet, J.-P., & Ulrich, P. (2015). Monitoring landslide displacements with the Geocube wireless network of low-cost GPS. *Engineering Geology*, *195*, 111-121. https://doi.org/10.1016/j.enggeo.2015.05.020

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review, 54(3)*, 1937–1967. https://doi.org/10.1007/s10462-020-09896-5

Bernardie, S., Desramaut, N., Malet, J.-P., Gourlay, M., & Grandjean, G. (2015). Prediction of changes in landslide rates induced by rainfall. *Landslides*, 12(3), 481-494. https://doi.org/10.1007/s10346-014-0495-8

Bertrand, C., Vallet, A., & Mudry, J. (2015). *Hydrochemical approach of mechanical degradation of the Séchilienne unstable slope*. In G. Lollino, D. Giordan, G. B. Crosta, J. Corominas, R. Azzam, J. Wasowski, & N. Sciarra (Eds.), Engineering geology for society and territory (Vol. 2, pp. 2137–2141). Springer. https://doi.org/10.1007/978-3-319-09057-3_383

Bièvre, G., Jongmans, D., Winiarski, T., & Zumbo, V. (2012). Application of geophysical measurements for assessing the role of fissures in water infiltration within a clay landslide (Trièves area, French Alps). *Hydrological Processes*, *26*(14), 2128-2142. https://doi.org/10.1002/hyp.7986

Binet, S. (2006). *L'hydrochimie, marqueur de l'évolution à long terme des versants montagneux fracturés vers de grands mouvements de terrain : Application à plusieurs échelles sur la haute vallée de la Tinée (Mercantour, France) et sur le versant de Rosone (Gran Paradiso, Italie)* (Thèse de doctorat en cotutelle, Université de Franche-Comté & Politecnico di Torino)

Binet, S., Guglielmi, Y., Bertrand, C., & Mudry, J. (2007). Unstable rock slope hydrogeology : Insights from the large-scale study of western Argentera-Mercantour hillslopes (South-East France). *Bulletin de La Société Géologique de France*, *178*(2), 159-168. https://doi.org/10.2113/gssgfbull.178.2.159

Binet, S., Jomard, H., Lebourg, T., Guglielmi, Y., Tric, E., Bertrand, C., & Mudry, J. (2007). Experimental analysis of groundwater flow through a landslide slip surface using natural and artificial water chemical tracers. *Hydrological Processes*, 21(25), 3463-3472. https://doi.org/10.1002/hyp.6579

Blázquez-García, A., Conde, A., Mori, U., & Lozano, J. A. (2020). *A review on outlier/anomaly detection in time series data* (arXiv:2002.04236). arXiv. https://doi.org/10.48550/arXiv.2002.04236

Boehm, V., Leong, W. J., Mahesh, R. B., Prapas, I., Nemni, E., Kalaitzis, F., Ganju, S., & Ramos-Pollan, R. (2022). *Deep Learning for Rapid Landslide Detection using Synthetic Aperture Radar (SAR) Datacubes* (arXiv:2211.02869). arXiv. https://doi.org/10.48550/arXiv.2211.02869

Bogaard, T. A., & Greco, R. (2016). Landslide hydrology : From hydrology to pore pressure. *WIREs Water*, *3*(3), 439-459. https://doi.org/10.1002/wat2.1126

Bonacci, O., & Andrić, I. (2015). Karst spring catchment : An example from Dinaric karst. *Environmental Earth Sciences*, *74*(7), 6211-6223. https://doi.org/10.1007/s12665-015-4644-8

Bontemps, N., Lacroix, P., Larose, E., Jara, J., & Taipe, E. (2020). Rain and small earthquakes maintain a slow-moving landslide in a persistent critical state. *Nature Communications*, *11*(1), 780. https://doi.org/10.1038/s41467-020-14445-3

Bonzanigo, L., Eberhardt, E., & Loew, S. (2007). Long-term investigation of a deep-seated creeping landslide in crystalline rock. Part I. Geological and hydromechanical factors controlling the Campo Vallemaggia landslide. *Canadian Geotechnical Journal*, *44*(10), 1157-1180. https://doi.org/10.1139/T07-043

Borrelli, L., & Gullà, G. (2017). Tectonic constraints on a deep-seated rock slide in weathered crystalline rocks. *Geomorphology*, *290*, 288-316. https://doi.org/10.1016/j.geomorph.2017.04.025

Bovis, M. J. (1990). Rock-slope deformation at Affliction Creek, southern Coast Mountains, British Columbia. *Canadian Journal of Earth Sciences*, *27*(2), 243-254. https://doi.org/10.1139/e90-024

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological), 26(2),* 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Breinl, K. (2016). Driving a lumped hydrological model with precipitation output from weather generators of different complexity. *Hydrological Sciences Journal*, *61*(8), 1395-1414. https://doi.org/10.1080/02626667.2015.1036755

Breinl, K., Turkington, T., & Stowasser, M. (2015). Simulating daily precipitation and temperature : A weather generation framework for assessing hydrometeorological hazards: Simulating daily precipitation and temperature. *Meteorological Applications*, *22*(3), 334-347. https://doi.org/10.1002/met.1459

Brückl, E. P. (2001). Cause-effect models of large landslides. *Natural Hazards*, *23*, 291–314. https://doi.org/10.1023/A:1011160810423

Buma, J. (2000). Finding the most suitable slope stability model for the assessment of the impact of climate change on a -landslide in southeast France. *Earth Surface Processes and Landforms*, *25*(6), 565-582. https://doi.org/10.1002/1096-9837

C

Cai, Z., Xu, W., Meng, Y., Shi, C., & Wang, R. (2016). Prediction of landslide displacement based on GA-LSSVM with multiple factors. *Bulletin of Engineering Geology and the Environment*, *75*(2), 637-646. https://doi.org/10.1007/s10064-015-0804-z

Cappa, F., Guglielmi, Y., Soukatchoff, V. M., Mudry, J., Bertrand, C., & Charmoille, A. (2004). Hydromechanical modeling of a large moving rock slope inferred from slope levelling coupled to spring long-term hydrochemical monitoring : Example of the La Clapière landslide (Southern Alps, France). *Journal of Hydrology*, *291*(1-2), 67-90. https://doi.org/10.1016/j.jhydrol.2003.12.013

Carlà, T., Intrieri, E., Raspini, F., Bardi, F., Farina, P., Ferretti, A., Colombo, D., Novali, F., & Casagli, N. (2019). Perspectives on the prediction of catastrophic slope failures from satellite InSAR. *Scientific Reports*, *9*(1), 14137. https://doi.org/10.1038/s41598-019-50792-y

Casagli, N., Intrieri, E., Tofani, V., Gigli, G., & Raspini, F. (2023). Landslide detection, monitoring and prediction with remote-sensing techniques. *Nature Reviews Earth & Environment*, *4*(1), 51-64. https://doi.org/10.1038/s43017-022-00373-x

Cascini, L., Calvello, M., & Grimaldi, G. M. (2014). Displacement trends of slow-moving landslides : Classification and forecasting. *Journal of Mountain Science*, 11(3), 592-606. https://doi.org/10.1007/s11629-013-2961-5

Catani, F., Lagomarsino, D., Segoni, S., & Tofani, V. (2013). Landslide susceptibility estimation by random forests technique : Sensitivity and scaling issues. *Natural Hazards and Earth System Sciences*, *13*(11), 2815-2831. https://doi.org/10.5194/nhess-13-2815-2013

Chang, F., Dong, S., Yin, H., Ye, X., Wu, Z., Zhang, W., & Zhu, H. (2025). 3D displacement time series prediction of a north-facing reservoir landslide powered by InSAR and machine learning. *Journal of Rock Mechanics and Geotechnical Engineering*, S1674775525000204. https://doi.org/10.1016/j.jrmge.2024.10.033

REFERENCES BILBIOGRAPHIQUES

Chatfield, C. (2003). *The analysis of time series: An introduction* (6th ed.). Chapman and Hall/CRC. https://doi.org/10.4324/9780203491683

Chauhan, N. K., & Singh, K. (2018). A Review on Conventional Machine Learning vs Deep Learning. *2018 International Conference on Computing, Power and Communication Technologies (GUCON)*, 347-352. https://doi.org/10.1109/GUCON.2018.8675097

Chen, H., & Zeng, Z. (2011). Deformation prediction of landslide based on genetic-simulated annealing algorithm and BP neural network. *The Fourth International Workshop on Advanced Computational Intelligence*, 675-679. https://doi.org/10.1109/iwaci.2011.6160092

Chen, H.-Y., Vojinovic, Z., Lo, W., & Lee, J.-W. (2023). Groundwater Level Prediction with Deep Learning Methods. *Water*, *15*(17), 3118. https://doi.org/10.3390/w15173118

Chen, T., & Guestrin, C. (2016). XGBoost : A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794. https://doi.org/10.1145/2939672.2939785

Cheng, C., Zhao, W., Wu, L., Chang, X., Scheuer, B., Zhang, J., Huang, R., & Tian, Y. (2025). Short-Term Displacement Prediction of Rainfall-Induced Landslides Through the Integration of Static and Dynamic Factors : A Case Study of China. *Water*, *17*(19), 2882. https://doi.org/10.3390/w17192882

Cheng, Z., Zou, C., & Dong, J. (2019). Outlier detection using isolation forest and local outlier factor. *Proceedings of the Conference on Research in Adaptive and Convergent Systems*, 161-168. https://doi.org/10.1145/3338840.3355641

Chikodili, N. B., Abdulmalik, M. D., Abisoye, O. A., & Bashir, S. A. (2021). Outlier Detection in Multivariate Time Series Data Using a Fusion of K-Medoid, Standardized Euclidean Distance and Z-Score. In S. Misra & B. Muhammad-Bello (Éds.), *Information and Communication Technology and Applications* (Vol. 1350, p. 259-271). Springer International Publishing. https://doi.org/10.1007/978-3-030-69143-1_21

Chowdhury, A. A., Das, A., Hoque, K. K. S., & Karmaker, D. (2022). A Comparative Study of Hyperparameter Optimization Techniques for Deep Learning. In M. S. Uddin, P. K. Jamwal, & J. C. Bansal (Éds.), *Proceedings of International Joint Conference on Advances in Computational Intelligence* (p. 509-521). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-0332-8_38

Cinkus, G., Wunsch, A., Mazzilli, N., Liesch, T., Chen, Z., Ravbar, N., Doummar, J., Fernández-Ortega, J., Barberá, J. A., Andreo, B., Goldscheider, N., & Jourde, H. (2023). Comparison of artificial neural networks and reservoir models for simulating karst spring discharge on five test sites in the Alpine and Mediterranean regions. *Hydrology and Earth System Sciences*, *27*(10), 1961-1985. https://doi.org/10.5194/hess-27-1961-2023

Clubb, F. J., Mudd, S. M., Schildgen, T. F., van der Beek, P. A., Devrani, R., & Sinclair, H. D. (2023). Himalayan valley-floor widths controlled by tectonically driven exhumation. *Nature Geoscience*, *16*(8), 739-746. https://doi.org/10.1038/s41561-023-01238-8

Coe, J. A., McKenna, J. P., Godt, J. W., & Baum, R. L. (2009). Basal-topographic control of stationary ponds on a continuously moving landslide. *Earth Surface Processes and Landforms*, *34*(2), 264-279. https://doi.org/10.1002/esp.1721

Collenteur, R. A., Bakker, M., Caljé, R., Klop, S. A., & Schaars, F. (2019). Pastas : Open Source Software for the Analysis of Groundwater Time Series. *Groundwater*, *57*(6), 877-885. https://doi.org/10.1111/gwat.12925

Collini, E., Palesi, L. A. I., Nesi, P., Pantaleo, G., Nocentini, N., & Rosi, A. (2022). Predicting and Understanding Landslide Events With Explainable AI. *IEEE Access*, *10*, 31175-31189. https://doi.org/10.1109/ACCESS.2022.3158328

Conte, E., Donato, A., & Troncone, A. (2014). A finite element approach for the analysis of active slow-moving landslides. *Landslides*, *11*(4), 723-731. https://doi.org/10.1007/s10346-013-0446-9

Conte, E., Donato, A., & Troncone, A. (2017). A simplified method for predicting rainfall-induced mobility of active landslides. *Landslides*, *14*(1), 35-45. https://doi.org/10.1007/s10346-016-0692-8

Corominas, J., Moya, J., Ledesma, A., Lloret, A., & Gili, J. A. (2005). Prediction of ground displacements and velocities from groundwater level changes at the Vallcebre landslide (Eastern Pyrenees, Spain). *Landslides*, *2*(2), 83-96. https://doi.org/10.1007/s10346-005-0049-1

Crosta, G. B. (1996). Landslide, spreading, deep seated gravitational deformation: Analysis, examples, problems and proposals. *Geografia Fisica e Dinamica Quaternaria*, 19(2), 297–313.

Crosta, G. B., di Prisco, C., Frattini, P., Frigerio, G., Castellanza, R., & Agliardi, F. (2014). Chasing a complete understanding of the triggering mechanisms of a large rapidly evolving rockslide. *Landslides*, *11*(5), 747-764. https://doi.org/10.1007/s10346-013-0433-1

Crosta, G. B., Frattini, P., & Agliardi, F. (2013). Deep seated gravitational slope deformations in the European Alps. *Tectonophysics*, *605*, 13-33. https://doi.org/10.1016/j.tecto.2013.04.028

Cruden, D. M., & Varnes, D. J. (1996). Landslide types and processes. In A. K. Turner & R. L. Schuster (Eds.), *Landslides: Investigation and mitigation* (Transportation Research Board Special Report No. 247, pp. 36–75). National Academies Press.

D-F

Dahal, A., & Lombardo, L. (2023). Explainable artificial intelligence in geoscience : A glimpse into the future of landslide susceptibility modeling. *Computers & Geosciences*, *176*, 105364. https://doi.org/10.1016/j.cageo.2023.105364

Delunel, R., Hantz, D., Braucher, R., Bourlès, D. L., Schoeneich, P., & Deparis, J. (2010). Surface exposure dating and geophysical prospecting of the Holocene Lauvitel rock slide (French Alps). *Landslides*, *7*(4), 393-400. https://doi.org/10.1007/s10346-010-0221-0

Desrues, M., Lacroix, P., & Brenguier, O. (2019). Satellite Pre-Failure Detection and In Situ Monitoring of the Landslide of the Tunnel du Chambon, French Alps. *Geosciences*, *9*(7), 313. https://doi.org/10.3390/geosciences9070313

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*(366a), 427–431. https://doi.org/10.2307/2286348

Dille, A., Kervyn, F., Mugaruka Bibentyo, T., Delvaux, D., Ganza, G. B., Ilombe Mawe, G., Kalikone Buzera, C., Safari Nakito, E., Moeyersons, J., Monsieurs, E., Nzolang, C., Smets, B., Kervyn, M., & Dewitte, O. (2019). Causes and triggers of deep-seated hillslope instability in the tropics – Insights from a 60-year record of Ikoma landslide (DR Congo). *Geomorphology*, *345*, 106835. https://doi.org/10.1016/j.geomorph.2019.106835

Dogulu, N., López López, P., Solomatine, D. P., Weerts, A. H., & Shrestha, D. L. (2015). Estimation of predictive hydrologic uncertainty using the quantile regression and UNEEC methods and their comparison on contrasting catchments. *Hydrology and Earth System Sciences*, *19*(7), 3181-3201. https://doi.org/10.5194/hess-19-3181-2015

Du, H., & Song, D. (2022). Investigation of failure prediction of open-pit coal mine landslides containing complex geological structures using the inverse velocity method. *Natural Hazards*, *111*(3), 2819-2854. https://doi.org/10.1007/s11069-021-05159-w

Du, J., Yin, K., & Lacasse, S. (2013). Displacement prediction in colluvial landslides, Three Gorges Reservoir, China. *Landslides*, *10*(2), 203-218. https://doi.org/10.1007/s10346-012-0326-8

Ebrahim, K. M. P., Fares, A., Faris, N., & Zayed, T. (2024). Exploring time series models for landslide prediction : A literature review. Geoenvironmental Disasters, 11(1), 25. https://doi.org/10.1186/s40677-024-00288-3

Ebrahim, K. M. P., Gomaa, S. M. M. H., Zayed, T., & Alfalah, G. (2024). Rainfall-induced landslide prediction models, part ii : Deterministic physical and phenomenologically models. *Bulletin of Engineering Geology and the Environment, 83(3)*, 85. https://doi.org/10.1007/s10064-024-03563-7

El Bedoui, S., Guglielmi, Y., Lebourg, T., & Pérez, J.-L. (2009). Deep-seated failure propagation in a fractured rock slope over 10,000 years : The La Clapière slope, the south-eastern French Alps. *Geomorphology, 105(3-4)*, 232-238. https://doi.org/10.1016/j.geomorph.2008.09.025

Federico, A., Popescu, M., Elia, G., Fidelibus, C., Internò, G., & Murianni, A. (2012). Prediction of time to slope failure : A general framework. *Environmental Earth Sciences, 66(1)*, 245-256. https://doi.org/10.1007/s12665-011-1231-5

Fiolleau, S., Jongmans, D., Bièvre, G., Chambon, G., Baillet, L., & Vial, B. (2020). Seismic characterization of a clay-block rupture in Harmalière landslide, French Western Alps. *Geophysical Journal International, 221(3)*, 1777-1788. https://doi.org/10.1093/gji/ggaa050

Fleury, P., Plagnes, V., & Bakalowicz, M. (2007). Modelling of the functioning of karst aquifers with a reservoir model : Application to Fontaine de Vaucluse (South of France). *Journal of Hydrology, 345(1-2)*, 38-49. https://doi.org/10.1016/j.jhydrol.2007.07.014

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences, 18(8)*, 2161-2181. https://doi.org/10.5194/nhess-18-2161-2018

Fürnkranz, J. (2011). Decision tree. In C. Sammut & G. I. Webb (Eds.*), Encyclopedia of Machine Learning (pp. 263–267)*. Springer. https://doi.org/10.1007/978-0-387-30164-8_204

G

Ge, Q., Li, J., Wang, X., Deng, Y., Zhang, K., & Sun, H. (2024). LiteTransNet : An interpretable approach for landslide displacement prediction using transformer model with attention mechanism. *Engineering Geology*, *331*, 107446. https://doi.org/10.1016/j.enggeo.2024.107446

Geyer, T., Birk, S., Liedl, R., & Sauter, M. (2008). Quantification of temporal distribution of recharge in karst systems from spring hydrographs. *Journal of Hydrology*, *348*(3-4), 452-463. https://doi.org/10.1016/j.jhydrol.2007.10.015

Gili, J. A., Moya, J., Corominas, J., Crosetto, M., & Monserrat, O. (2021). Past, Present and Future Monitoring at the Vallcebre Landslide (Eastern Pyrenees, Spain). *Applied Sciences*, *11*(2), 571. https://doi.org/10.3390/app11020571

Glueer, F., Loew, S., Manconi, A., & Aaron, J. (2019). From Toppling to Sliding : Progressive Evolution of the Moosfluh Landslide, Switzerland. *Journal of Geophysical Research: Earth Surface*, *124*(12), 2899-2919. https://doi.org/10.1029/2019JF005019

Gneiting, T., & Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. Journal of the American Statistical Association, 102(477), 359-378. https://doi.org/10.1198/016214506000001437

Gómez-Ríos, A., Luengo, J., & Herrera, F. (2017). A study on the noise label influence in boosting algorithms: AdaBoost, GBM and XGBoost. In F. J. Martínez de Pisón, R. Urraca, H. Quintián, & E. Corchado (Eds.), *Hybrid Artificial Intelligent Systems* (Lecture Notes in Computer Science, Vol. 10334, pp. 268–280). Springer. https://doi.org/10.1007/978-3-319-59650-1_23

González, S., García, S., Del Ser, J., Rokach, L., & Herrera, F. (2020). A practical tutorial on bagging and boosting based ensembles for machine learning : Algorithms, software tools, performance study, practical perspectives and opportunities. *Information Fusion, 64*, 205-237. https://doi.org/10.1016/j.inffus.2020.07.007

Gong, W., Zhang, S., Juang, C. H., Tang, H., & Pudasaini, S. P. (2024). Displacement prediction of landslides at slope-scale : Review of physics-based and data-driven approaches. *Earth-Science Reviews*, *258*, 104948. https://doi.org/10.1016/j.earscirev.2024.104948

Gu, X., Yang, L., & Sun, L. (2024). *Explainable artificial intelligence for time series: A systematic survey* (Preprint). ResearchGate.

Guglielmi, Y., Vengeon, J., Bertrand, C., Mudry, J., Follacci, J., & Giraud, A. (2002). Hydrogeochemistry : An investigation tool to evaluate infiltration into large moving rock masses (case study of La Clapière and Séchilienne alpine landslides). *Bulletin of Engineering Geology and the Environment*, *61*(4), 311-324. https://doi.org/10.1007/s10064-001-0144-z

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria : Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1-2), 80-91. https://doi.org/10.1016/j.jhydrol.2009.08.003

H

Han, H., Shi, B., & Zhang, L. (2021). Prediction of landslide sharp increase displacement by SVM with considering hysteresis of groundwater change. *Engineering Geology, 280*, 105876. https://doi.org/10.1016/j.enggeo.2020.105876

Handwerger, A. L., Roering, J. J., & Schmidt, D. A. (2013). Controls on the seasonal deformation of slow-moving landslides. *Earth and Planetary Science Letters, 377-378*, 239-247. https://doi.org/10.1016/j.epsl.2013.06.047

Havenith, H.-B., Mreyen, A.-S., Torgoev, A., & Micu, M. (2017). Numerical models of unstable slopes in seismic areas—Based on 3D geomodels. In M. Mikoš, N. Casagli, Y. Yin, & K. Sassa (Eds.), *Advancing culture of living with landslides* (pp. 47–57). Springer. https://doi.org/10.1007/978-3-319-53485-5_6

Hendron, A. J., & Patton, F. D. (1987). The Vaiont slide—A geotechnical analysis based on new geologic observations of the failure surface. *Engineering Geology, 24(1-4)*, 475-491. https://doi.org/10.1016/0013-7952(87)90080-9

Herrera, G., Fernández-Merodo, J. A., Mulas, J., Pastor, M., Luzi, G., & Monserrat, O. (2009). A landslide forecasting model using ground based SAR data : The Portalet case study. *Engineering Geology, 105(3-4)*, 220-230. https://doi.org/10.1016/j.enggeo.2009.02.009

Hibert, C., Grandjean, G., Bitri, A., Travelletti, J., & Malet, J.-P. (2012). Characterizing landslides through geophysical data fusion : Example of the La Valette landslide (France). *Engineering Geology, 128*, 23-29. https://doi.org/10.1016/j.enggeo.2011.05.001

Hibert, C., Michéa, D., Provost, F., Malet, J.-P., & Geertsema, M. (2019). Exploration of continuous seismic recordings with a machine learning approach to document 20 yr of landslide activity in Alaska. *Geophysical Journal International, 219*(2), 1138-1147. https://doi.org/10.1093/gji/ggz354

Highland, L., & Johnson, M. (2004). *Landslide types and processes* (Fact Sheet 2004–3072). U.S. Geological Survey. https://doi.org/10.3133/fs20043072

Hippolyte, J.-C., Bourlès, D., Léanni, L., Braucher, R., Chauvet, F., & Lebatard, A. E. (2012). 10Be ages reveal >12 ka of gravitational movement in a major sackung of the Western Alps (France). *Geomorphology, 171-172*, 139-153. https://doi.org/10.1016/j.geomorph.2012.05.013

Hong, Y.-M., & Wan, S. (2011). Forecasting groundwater level fluctuations for rainfall-induced landslide. *Natural Hazards, 57(2)*, 167-184. https://doi.org/10.1007/s11069-010-9603-9

Hu, L., & Jiao, J. J. (2015). Calibration of a large-scale groundwater flow model using GRACE data : A case study in the Qaidam Basin, China. *Hydrogeology Journal, 23(7)*, 1305-1317. https://doi.org/10.1007/s10040-015-1278-6

Huang, D., He, J., Song, Y., Guo, Z., Huang, X., & Guo, Y. (2022). Displacement Prediction of the Muyubao Landslide Based on a GPS Time-Series Analysis and Temporal Convolutional Network Model. *Remote Sensing, 14(11)*, 2656. https://doi.org/10.3390/rs14112656

Huang, M., Yang, H., & Yang, F. (2024). Application of the Optuna-NeuralProphet model for predicting step-like landslide displacement. *AIP Advances, 14(12)*, 125125. https://doi.org/10.1063/5.0233619

Huang, Q., Mao, J., & Liu, Y. (2012). An improved grid search algorithm of SVR parameters optimization. *2012 IEEE 14th International Conference on Communication Technology*, 1022-1026. https://doi.org/10.1109/ICCT.2012.6511415

Huang, Y., & Zhao, L. (2018). Review on landslide susceptibility mapping using support vector machines. *CATENA, 165*, 520-529. https://doi.org/10.1016/j.catena.2018.03.003

Hungr, O., Leroueil, S., & Picarelli, L. (2014). The Varnes classification of landslide types, an update. *Landslides, 11(2)*, 167-194. https://doi.org/10.1007/s10346-013-0436-y

Hutchinson, J. N. (1988). General report: Morphological and geotechnical parameters of landslides in relation to geology and hydrogeology. In C. Bonnard (Ed.), *Proceedings of the 5th International Symposium on Landslides* (Lausanne, 10–15 July 1988) (Vol. 1, pp. 3–35). A.A. Balkema.

I-J

Intergovernmental Panel on Climate Change. (2023). Climate change 2021: The physical science basis: Working Group I contribution to the *Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. https://doi.org/10.1017/9781009157896

Intrieri, E., Carlà, T., & Gigli, G. (2019). Forecasting the time of failure of landslides at slope-scale : A literature review. *Earth-Science Reviews, 193*, 333-349. https://doi.org/10.1016/j.earscirev.2019.03.019

Intrieri, E., & Gigli, G. (2016). Landslide forecasting and factors influencing predictability. *Natural Hazards and Earth System Sciences, 16*(12), 2501-2510. https://doi.org/10.5194/nhess-16-2501-2016

Iverson, R. M. (2000). Landslide triggering by rain infiltration. *Water Resources Research, 36(7)*, 1897-1910. https://doi.org/10.1029/2000WR900090

Iverson, R. M., & Major, J. J. (1987). Rainfall, ground-water flow, and seasonal movement at Minor Creek landslide, northwestern California : Physical interpretation of empirical relations. *Geological Society of America Bulletin, 99(4)*, 579-584.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. Springer. https://doi.org/10.1007/978-1-4614-7138-7

Jiang, H., Li, Y., Zhou, C., Hong, H., Glade, T., & Yin, K. (2020). Landslide Displacement Prediction Combining LSTM and SVR Algorithms : A Case Study of Shengjibao Landslide from the Three Gorges Reservoir Area. *Applied Sciences, 10(21)*, 7830. https://doi.org/10.3390/app10217830

Jiang, Y., Luo, H., Xu, Q., Lu, Z., Liao, L., Li, H., & Hao, L. (2022). A Graph Convolutional Incorporating GRU Network for Landslide Displacement Forecasting Based on Spatiotemporal Analysis of GNSS Observations. *Remote Sensing, 14(4)*, 1016. https://doi.org/10.3390/rs14041016

John, J. A., & Draper, N. R. (1980). An alternative family of transformations. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 29(2)*, 190–197. https://doi.org/10.2307/2986305

Jukić, D., & Denić-Jukić, V. (2009). Groundwater balance estimation in karst by using a conceptual rainfall–runoff model. *Journal of Hydrology, 373(3-4)*, 302-315. https://doi.org/10.1016/j.jhydrol.2009.04.035

K

Katz, R. W. (1977). Precipitation as a Chain-Dependent Process. *Journal of Applied Meteorology, 16(7)*, 671-676.

Kavzoglu, T., & Teke, A. (2022). Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (XGBoost). *Bulletin of Engineering Geology and the Environment, 81(5)*. https://doi.org/10.1007/s10064-022-02708-w

Keller, D. E., Fischer, A. M., Frei, C., Liniger, M. A., Appenzeller, C., & Knutti, R. (2015). Implementation and validation of a Wilks-type multi-site daily precipitation generator over a typical Alpine river catchment. *Hydrology and Earth System Sciences, 19(5)*, 2163-2177. https://doi.org/10.5194/hess-19-2163-2015

Keqiang, H., Xiangran, L., Xueqing, Y., & Dong, G. (2008). The landslides in the Three Gorges Reservoir Region, China and the effects of water storage and rain on their stability. *Environmental Geology, 55*(1), 55-63. https://doi.org/10.1007/s00254-007-0964-7

Kıraç, M. F., Aktemur, B., Sözer, H., & Gebizli, C. Ş. (2019). Automatically learning usage behavior and generating event sequences for black-box testing of reactive systems. *Software Quality Journal, 27*(2), 861-883. https://doi.org/10.1007/s11219-018-9439-1

Kirschbaum, D., Stanley, T., & Zhou, Y. (2015). Spatial and temporal analysis of a global landslide catalog. *Geomorphology, 249*, 4-15. https://doi.org/10.1016/j.geomorph.2015.03.016

Kishor, K., Aggarwal, A., Srivastava, P. K., Sharma, Y. K., Lee, J., & Ghobadi, F. (2025). A Systematic Literature Review of MODFLOW Combined with Artificial Neural Networks (ANNs) for Groundwater Flow Modelling. *Water, 17(16)*, 2375. https://doi.org/10.3390/w17162375

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note : Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences, 23(10)*, 4323-4331. https://doi.org/10.5194/hess-23-4323-2019

Krkač, M., Špoljarić, D., Bernat, S., & Arbanas, S. M. (2017). Method for prediction of landslide movements based on random forests. *Landslides, 14(3)*, 947-960. https://doi.org/10.1007/s10346-016-0761-z

Krzeminska, D. M., Bogaard, T. A., Malet, J.-P., & van Beek, L. P. H. (2013). A model of hydrological and mechanical feedbacks of preferential fissure flow in a slow-moving landslide. *Hydrology and Earth System Sciences, 17(3)*, 947-959. https://doi.org/10.5194/hess-17-947-2013

Krzeminska, D. M., Bogaard, T. A., van Asch, Th. W. J., & van Beek, L. P. H. (2012). A conceptual model of the hydrological influence of fissures on landslide activity. *Hydrology and Earth System Sciences, 16(6)*, 1561-1576. https://doi.org/10.5194/hess-16-1561-2012

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics, 54(1-3)*, 159-178. https://doi.org/10.1016/0304-4076(92)90104-Y

L

Lacroix, P., Bièvre, G., Pathier, E., Kniess, U., & Jongmans, D. (2018). Use of Sentinel-2 images for the detection of precursory motions before landslide failures. *Remote Sensing of Environment*, *215*, 507-516. https://doi.org/10.1016/j.rse.2018.03.042

Lacroix, P., Handwerger, A. L., & Bièvre, G. (2020). Life and death of slow-moving landslides. *Nature Reviews Earth & Environment*, *1*(8), 404-419. https://doi.org/10.1038/s43017-020-0072-8

Langousis, A., Mamalakis, A., Puliga, M., & Deidda, R. (2016). Threshold detection for the generalized Pareto distribution : Review of representative methods and application to the NOAA NCDC daily rainfall database. *Water Resources Research*, *52*(4), 2659-2681. https://doi.org/10.1002/2015WR018502

Latif, S. D., Hazrin, N. A. B., Koo, C. H., Ng, J. L., Chaplot, B., Huang, Y. F., El-Shafie, A., & Ahmed, A. N. (2023). Assessing rainfall prediction models: Exploring the advantages of machine learning and remote sensing approaches. *Alexandria Engineering Journal, 82*, 16–25. https://doi.org/10.1016/j.aej.2023.09.060

Le Bas, C. (2018). Carte de la Réserve Utile en eau issue de la *Base de Données Géographique des Sols de France* (V2) [Data set]. Recherche Data Gouv. https://doi.org/10.15454/JPB9RB

Le Roux, O., Jongmans, D., Kasperski, J., Schwartz, S., Potherat, P., Lebrouc, V., Lagabrielle, R., & Meric, O. (2011). Deep geophysical investigation of the large Séchilienne landslide (Western Alps, France) and calibration with geological data. *Engineering Geology*, *120*(1-4), 18-31. https://doi.org/10.1016/j.enggeo.2011.03.004

Le Roux, O., Schwartz, S., Gamond, J. F., Jongmans, D., Bourles, D., Braucher, R., Mahaney, W., Carcaillet, J., & Leanni, L. (2009). CRE dating on the head scarp of a major landslide (Séchilienne, French Alps), age constraints on Holocene kinematics. *Earth and Planetary Science Letters*, *280*(1-4), 236-245. https://doi.org/10.1016/j.epsl.2009.01.034

Lee, S., Hong, S.-M., & Jung, H.-S. (2017). A Support Vector Machine for Landslide Susceptibility Mapping in Gangwon Province, Korea. *Sustainability*, *9*(1), 48. https://doi.org/10.3390/su9010048

Leloup, P. H., Arnaud, N., Sobel, E. R., & Lacassin, R. (2005). Alpine thermal and structural evolution of the highest external crystalline massif: The Mont Blanc. *Tectonics, 24(4)*, TC4002. https://doi.org/10.1029/2004TC001676

Li, B., Liu, K., Wang, M., He, Q., Jiang, Z., Zhu, W., & Qiao, N. (2022). Global Dynamic Rainfall-Induced Landslide Susceptibility Mapping Using Machine Learning. *Remote Sensing*, *14*(22), 5795. https://doi.org/10.3390/rs14225795

Li, J., Qin, J., Kang, K., Liang, M., Liu, K., & Ding, X. (2025). Enhanced Spatiotemporal Landslide Displacement Prediction Using Dynamic Graph-Optimized GNSS Monitoring. *Sensors*, *25*(15), 4754. https://doi.org/10.3390/s25154754

Li, X., Handwerger, A. L., & Buscarnera, G. (2023). Viscoplastic modelling of rainfall-driven slow-moving landslides : Application to California Coast Ranges. *Landslides*, *20*(6), 1101-1113. https://doi.org/10.1007/s10346-023-02039-1

Li, Y.-H., Tung, C.-P., Li, Y.-H., & Chen, P.-Y. (2018). Assessing climate change adaptations for community-scale water resources using a low-frequency weather generator. *Paddy and Water Environment*, *16*(1), 55-69. https://doi.org/10.1007/s10333-017-0613-z

Li, Z., Chen, J., Cao, C., Zhang, W., Li, Y., Hu, J., & Li, Z. (2025). Detecting subtle nonlinear changes in slow-moving landslides with a hybrid InSAR change point detection framework : Methodological assessment and validation. International *Journal of Applied Earth Observation and Geoinformation*, *144*, 104949. https://doi.org/10.1016/j.jag.2025.104949

Lian, C., Zeng, Z., Yao, W., & Tang, H. (2013). Displacement prediction of landslide based on PSOGSA-ELM with mixed kernel. *2013 Sixth International Conference on Advanced Computational Intelligence (ICACI)*, 52-57. https://doi.org/10.1109/icaci.2013.6748473

Lilliefors, H. W. (1967). On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association, 62(318)*, 399–402. https://doi.org/10.1080/01621459.1967.10482916

Lin, G.-W., Kuo, H.-L., Chen, C.-W., Wei, L.-W., & Zhang, J.-M. (2020). Using a Tank Model to Determine Hydro-Meteorological Thresholds for Large-Scale Landslides in Taiwan. *Water*, *12*(1), 253. https://doi.org/10.3390/w12010253

Lin, N., Zhang, D., Feng, S., Ding, K., Tan, L., Wang, B., Chen, T., Li, W., Dai, X., Pan, J., & Tang, F. (2023). Rapid Landslide Extraction from High-Resolution Remote Sensing Images Using SHAP-OPT-XGBoost. *Remote Sensing*, *15*(15), 3901. https://doi.org/10.3390/rs15153901

Lissak, C. (2012). *Les glissements de terrain des versants côtiers du pays d'Auge (Calvados) : Morphologie, fonctionnement, et gestion du risque*. (Thèse de doctorat). Université de Caen Basse-Normandie.

Lissak, C., Maquaire, O., & Malet, J.-P. (2009). *Role of hydrological process in landslide occurrence: Villerville–Cricqueboeuf landslides (Normandy coast, France)*. In J.-P. Malet, A. Remaître, & T. A. Bogaard (Eds.), *Proceedings of the International Conference on Landslide Processes: From geomorphologic mapping to dynamic modelling* (pp. 175–180). CERG Éditions. https://hal.science/hal-00533741

Liu, L., Yin, H., Xiao, T., Yang, B., & Lacasse, S. (2024). Ensemble learning for landslide displacement prediction : A perspective of Bayesian optimization and comparison of different time series analysis methods. *Stochastic Environmental Research and Risk Assessment*, *38*(8), 3031-3058. https://doi.org/10.1007/s00477-024-02730-2

Liu, W., Zhang, A., Wang, L., Fu, G., Chen, D., Liu, C., & Cai, T. (2015). Projecting streamflow in the Tangwang River basin (China) using a rainfall generator and two hydrological models. *Climate Research*, *62*(2), 79-97. https://doi.org/10.3354/cr01261

Liu, Y., Long, J., Li, C., & Zhan, W. (2024). Physics-informed data assimilation model for displacement prediction of hydrodynamic pressure-driven landslide. *Computers and Geotechnics*, *167*, 106085. https://doi.org/10.1016/j.compgeo.2024.106085

Liu, Z., Guo, D., Lacasse, S., Li, J., Yang, B., & Choi, J. (2020). Algorithms for intelligent prediction of landslide displacements. *Journal of Zhejiang University-SCIENCE A*, 21(6), 412-429. https://doi.org/10.1631/jzus.A2000005

Logar, J., Turk, G., Marsden, P., & Ambrožič, T. (2017). Prediction of rainfall induced landslide movements by artificial neural networks. *Natural Hazards and Earth System Sciences Discussions* [Preprint, withdrawn]. https://doi.org/10.5194/nhess-2017-253

Long, J., Li, C., Liu, Y., Feng, P., & Zuo, Q. (2022). A multi-feature fusion transfer learning method for displacement prediction of rainfall reservoir-induced landslide with step-like deformation characteristics. *Engineering Geology*, *297*, 106494. https://doi.org/10.1016/j.enggeo.2021.106494

Louppe, G., Wehenkel, L., Sutera, A., & Geurts, P. (2013). Understanding variable importances in forests of randomized trees. In *Advances in Neural Information Processing Systems* (Vol. 26, pp. 431–439).

Lu, N., & Godt, J. W. (2008). Infinite slope stability under steady unsaturated seepage conditions. *Water Resources Research*, *44*(11), W11404. https://doi.org/10.1029/2008WR006976

Lundberg, S. M., Erion, G. G., & Lee, S.-I. (2019). *Consistent Individualized Feature Attribution for Tree Ensembles* (arXiv:1802.03888). arXiv. https://doi.org/10.48550/arXiv.1802.03888

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in *Neural Information Processing Systems* (Vol. 30, pp. 4765–4774).

Lyu, Y., Li, H., Sayagh, M., Jiang, Z. M. (Jack), & Hassan, A. E. (2021). An Empirical Study of the Impact of Data Splitting Decisions on the Performance of AIOps Solutions. *ACM Transactions on Software Engineering and Methodology*, *30*(4), 1-38. https://doi.org/10.1145/3447876

M

Ma, J., Yao, L., Deng, L., Yang, Q., Chen, Y., & Ouyang, C. (2025). Spatio-Temporal Deformation Prediction of Large Landslides in the Three Gorges Reservoir Area Based on Time-Series Graph Convolutional Network Model. *Applied Sciences*, *15*(8), 4491. https://doi.org/10.3390/app15084491

Ma, W., Dong, J., Wei, Z., Peng, L., Wu, Q., Chen, C., Wu, Y., & Xie, F. (2022). Landslide Displacement Prediction With Gated Recurrent Unit and Spatial-Temporal Correlation. *Frontiers in Earth Science*, *10*, 950723. https://doi.org/10.3389/feart.2022.950723

Ma, Z., & Mei, G. (2025). Forecasting landslide deformation by integrating domain knowledge into interpretable deep learning considering spatiotemporal correlations. *Journal of Rock Mechanics and Geotechnical Engineering*, *17*(2), 960-982. https://doi.org/10.1016/j.jrmge.2024.02.034

Ma, Z., Mei, G., & Piccialli, F. (2021). Machine learning for landslides prevention : A survey. *Neural Computing and Applications*, *33*(17), 10881-10907. https://doi.org/10.1007/s00521-020-05529-8

Malet, J.-P., Maquaire, O., & Calais, E. (2002). The use of Global Positioning System techniques for the continuous monitoring of landslides : Application to the Super-Sauze earthflow (Alpes-de-Haute-Provence, France). *Geomorphology*, *43*(1-2), 33-54. https://doi.org/10.1016/S0169-555X(01)00098-8

Malet, J.-P., van Asch, Th. W. J., van Beek, R., & Maquaire, O. (2005). Forecasting the behaviour of complex landslides with a spatially distributed hydrological model. *Natural Hazards and Earth System Sciences*, *5*(1), 71-85. https://doi.org/10.5194/nhess-5-71-2005

Mangin, A. (1975). *Contribution à l'étude hydrodynamique des aquifères karstiques* (Thèse de doctorat). Université de Dijon, Dijon, France

Mansour, M. F., Morgenstern, N. R., & Martin, C. D. (2011). Expected damage from displacement of slow-moving slides. *Landslides*, *8*(1), 117-131. https://doi.org/10.1007/s10346-010-0227-7

Maquaire, O. (2002). *Aléas géomorphologiques (mouvements de terrain) : Processus, fonctionnement, cartographie* (Mémoire d'habilitation à diriger des recherches). Université Louis Pasteur, Strasbourg, France.

Maréchal, J.-C. (1998). *Les circulations d'eau dans les massifs cristallins alpins et leurs relations avec les ouvrages souterrains* (Thèse de doctorat, no 1769). École polytechnique fédérale de Lausanne (EPFL), Lausanne, Suisse.

Marsaud, B. (1996). *Structure et fonctionnement de la zone noyée des karsts à partir des résultats expérimentaux* (Thèse de doctorat). Université Paris XI, Orsay, France.

Martinez-Villalobos, C., & Neelin, J. D. (2019). Why Do Precipitation Intensities Tend to Follow Gamma Distributions? *Journal of the Atmospheric Sciences*, *76*(11), 3611-3631. https://doi.org/10.1175/JAS-D-18-0343.1

Matsuura, S., Asano, S., & Okamoto, T. (2008). Relationship between rain and/or meltwater, pore-water pressure and displacement of a reactivated landslide. *Engineering Geology*, *101*(1-2), 49-59. https://doi.org/10.1016/j.enggeo.2008.03.007

Mazzilli, N., Guinot, V., Jourde, H., Lecoq, N., Labat, D., Arfib, B., Baudement, C., Danquigny, C., Dal Soglio, L., & Bertin, D. (2019). KarstMod : A modelling platform for rainfall - discharge analysis and modelling dedicated to karst systems. *Environmental Modelling & Software*, *122*, 103927. https://doi.org/10.1016/j.envsoft.2017.03.015

Mazzilli, N., Sivelle, V., Cinkus, G., Jourde, H., & Bertin, D. (2023). *KarstMod user guide – version 3.0*. HAL. https://hal.science/hal-01832693v2

McColl, S. T. (2015). Landslide causes and triggers. In J. F. Shroder & T. Davies (Eds.), *Landslide hazards, risks, and disasters* (pp. 17–42). Academic Press. https://doi.org/10.1016/B978-0-12-396452-6.00002-1

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, *5*, 115–133.

Meng, S., Shi, Z., Peng, M., Li, G., Zheng, H., Liu, L., & Zhang, L. (2024). Landslide displacement prediction with step-like curve based on convolutional neural network coupled with bi-directional gated recurrent unit optimized by attention mechanism. *Engineering Applications of Artificial Intelligence*, *133*, 108078. https://doi.org/10.1016/j.engappai.2024.108078

Meric, O., Garambois, S., Jongmans, D., Wathelet, M., Chatelain, J. L., & Vengeon, J. M. (2005). Application of geophysical methods for the investigation of the large gravitational mass movement of Séchilienne, France. *Canadian Geotechnical Journal*, *42*(4), 1105-1115. https://doi.org/10.1139/t05-034

Mínguez, R. (2025). Automatic Threshold Selection for Generalized Pareto and Pareto–Poisson Distributions in Rainfall Analysis : A Case Study Using the NOAA NCDC Daily Rainfall Database. *Atmosphere*, *16*(1), 61. https://doi.org/10.3390/atmos16010061

Moeineddin, A., Seguí, C., Dueber, S., & Fuentes, R. (2023). Physics-informed neural networks applied to catastrophic creeping landslides. *Landslides*, *20*(9), 1853-1863. https://doi.org/10.1007/s10346-023-02072-0

Mondini, A. C., Guzzetti, F., Chang, K.-T., Monserrat, O., Martha, T. R., & Manconi, A. (2021). Landslide failures detection and mapping using Synthetic Aperture Radar : Past, present and future. *Earth-Science Reviews*, *216*, 103574. https://doi.org/10.1016/j.earscirev.2021.103574

Mondini, A. C., Guzzetti, F., & Melillo, M. (2023). Deep learning forecast of rainfall-induced shallow landslides. *Nature Communications*, *14*(1), 2466. https://doi.org/10.1038/s41467-023-38135-y

Montesino Pouzols, F., & Lendasse, A. (2010). Effect of different detrending approaches on computational intelligence models of time series. *The 2010 International Joint Conference on Neural Networks (IJCNN)*, 1-8. https://doi.org/10.1109/IJCNN.2010.5596314

Moya, J., Corominas, J., Gili, J. A., Ledesma, A., & Lloret, A. (2017). *Twenty years of continuous monitoring of the Vallcebre landslide (Eastern Pyrenees): Lessons learned*. In E. Alonso & N. Pinyol (Eds.), *Proceedings of the JTC1 Workshop on Advances in Landslide Understanding* (Barcelona, Spain, May 24–26, 2017) (pp. 101–104). CIMNE. (ISBN 978-84-946909-1-4)

Mramba, L. K., Liu, X., Lynch, K. F., Yang, J., Aronsson, C. A., Hummel, S., Norris, J. M., Virtanen, S. M., Hakola, L., Uusitalo, U. M., & Krischer, J. P. (2024). Detecting potential outliers in longitudinal data with time-dependent covariates. *European Journal of Clinical Nutrition*, *78*(4), 344-350. https://doi.org/10.1038/s41430-023-01393-6

Müller, J., Park, J., Sahu, R., Varadharajan, C., Arora, B., Faybishenko, B., & Agarwal, D. (2021). Surrogate optimization of deep neural networks for groundwater predictions. *Journal of Global Optimization*, *81*(1), 203-231. https://doi.org/10.1007/s10898-020-00912-0

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Interpretable machine learning : Definitions, methods, and applications. *Proceedings of the National Academy of Sciences*, *116*(44), 22071-22080. https://doi.org/10.1073/pnas.1900654116

N-O

Nava, L., Carraro, E., Reyes-Carmona, C., Puliero, S., Bhuyan, K., Rosi, A., Monserrat, O., Floris, M., Meena, S. R., Galve, J. P., & Catani, F. (2023). Landslide displacement forecasting using deep learning and monitoring data across selected sites. *Landslides*, *20*(10), 2111–2129. https://doi.org/10.1007/s10346-023-02104-9

Naveau, P., Huser, R., Ribereau, P., & Hannart, A. (2016). Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resources Research*, *52*(4), 2753-2769. https://doi.org/10.1002/2015WR018552

Neaupane, K. M., & Achet, S. H. (2004). Use of backpropagation neural network for landslide monitoring : A case study in the higher Himalaya. *Engineering Geology*, *74*(3-4), 213-226. https://doi.org/10.1016/j.enggeo.2004.03.010

Nereson, A. L., & Finnegan, N. J. (2019). Drivers of earthflow motion revealed by an 80 yr record of displacement from Oak Ridge earthflow, Diablo Range, California, USA. *GSA Bulletin*, *131*(3-4), 389-402. https://doi.org/10.1130/B32020.1

Nevers, P. (2021). *Fonctionnement hydrogéologique et processus de transfert dans le massif fissuré instable de Séchilienne (Alpes Françaises). Apports du couplage de traceurs naturels et isotopiques.* (Thèse de doctorat). Université Bourgogne Franche-Comté, Besançon, France.

Niu, X., Ma, J., Wang, Y., Zhang, J., Chen, H., & Tang, H. (2021). A Novel Decomposition-Ensemble Learning Model Based on Ensemble Empirical Mode Decomposition and Recurrent Neural Network for Landslide Displacement Prediction. *Applied Sciences, 11*(10), 4684. https://doi.org/10.3390/app11104684

Osawa, H., Matsushi, Y., Matsuura, S., Okamoto, T., Shibasaki, T., & Hirashima, H. (2018). Seasonal transition of hydrological processes in a slow-moving landslide in a snowy region. *Hydrological Processes*, *32*(17), 2695-2707. https://doi.org/10.1002/hyp.13212

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model? *Journal of Hydrology*, *303*(1-4), 290-306. https://doi.org/10.1016/j.jhydrol.2004.08.026

P

Palis, E., Lebourg, T., Tric, E., Malet, J.-P., & Vidal, M. (2017). Long-term monitoring of a large deep-seated landslide (La Clapiere, South-East French Alps) : Initial study. *Landslides, 14*(1), 155-170. https://doi.org/10.1007/s10346-016-0705-7

Peiro, Y., Volpe, E., Ciabatta, L., & Cattoni, E. (2024). High Resolution Precipitation and Soil Moisture Data Integration for Landslide Susceptibility Mapping. Geosciences, 14(12), 330. https://doi.org/10.3390/geosciences14120330

Petrică, A.-C., Stancu, S., & Ghițulescu, V. (2017). Stationarity – The central concept in time series analysis. *International Journal of Emerging Research in Management and Technology*, *6*(1), 6–16. https://doi.org/10.23956/IJERMT/V6N1/107

Pfeiffer, J., Zieher, T., Schmieder, J., Rutzinger, M., & Strasser, U. (2021). Spatio-temporal assessment of the hydrological drivers of an active deep-seated gravitational slope deformation : The Vögelsberg landslide in Tyrol (Austria). *Earth Surface Processes and Landforms, 46*(10), 1865-1881. https://doi.org/10.1002/esp.5129

Pham, Q. B., Kumar, M., Di Nunno, F., Elbeltagi, A., Granata, F., Islam, A. R. Md. T., Talukdar, S., Nguyen, X. C., Ahmed, A. N., & Anh, D. T. (2022). Groundwater level prediction using machine learning algorithms in a drought-prone area. *Neural Computing and Applications, 34*(13), 10751-10773. https://doi.org/10.1007/s00521-022-07009-7

Pisani, G., Castelli, M., & Scavia, C. (2010). Hydrogeological model and hydraulic behaviour of a large landslide in the Italian Western Alps. *Natural Hazards and Earth System Sciences, 10*(11), 2391-2406. https://doi.org/10.5194/nhess-10-2391-2010

Pratama, I., Permanasari, A. E., Ardiyanto, I., & Indrayani, R. (2016). A review of missing values handling methods on time-series data. In *2016 International Conference on Information Technology Systems and Innovation (ICITSI)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICITSI.2016.7858189

Preisig, G., Eberhardt, E., Smithyman, M., Preh, A., & Bonzanigo, L. (2016). Hydromechanical Rock Mass Fatigue in Deep-Seated Landslides Accompanying Seasonal Variations in Pore Pressures. *Rock Mechanics and Rock Engineering, 49*(6), 2333-2351. https://doi.org/10.1007/s00603-016-0912-5

Prokešová, R., Medveďová, A., Tábořík, P., & Snopková, Z. (2013). Towards hydrological triggering mechanisms of large deep-seated landslides. *Landslides, 10*(3), 239-254. https://doi.org/10.1007/s10346-012-0330-z

Provost, F., Michéa, D., Malet, J.-P., Boissier, E., Pointal, E., Stumpf, A., Pacini, F., Doin, M.-P., Lacroix, P., Proy, C., & Bally, P. (2022). Terrain deformation measurements from optical satellite imagery : The MPIC-OPT processing services for geohazards monitoring. *Remote Sensing of Environment, 274*, 112949. https://doi.org/10.1016/j.rse.2022.112949

R

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks : A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

Redshaw, P., Bottomley, J., & Free, M. (2020). *The global landslide hazard map: Final project report*. Ove Arup & Partners International Ltd. (Report prepared for the Global Facility for Disaster Reduction and Recovery [GFDRR] and the World Bank).

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. Earth-Science Reviews, 180, 60-91. https://doi.org/10.1016/j.earscirev.2018.03.001

Ren, S., Zhang, Y., Li, J., Zhou, Z., Liu, X., & Tao, C. (2023). Deformation Behavior and Reactivation Mechanism of the Dandu Ancient Landslide Triggered by Seasonal Rainfall : A Case Study from the East Tibetan Plateau, China. Remote Sensing, 15(23), 5538. https://doi.org/10.3390/rs15235538

Richardson, C. W. (1981). Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resources Research, 17(1), 182-190. https://doi.org/10.1029/WR017i001p00182

Rojat, T., Puget, R., Filliat, D., Del Ser, J., Gelin, R., & Díaz-Rodríguez, N. (2021). Explainable artificial intelligence (XAI) on time series data: A survey (arXiv:2104.00950). arXiv. https://arxiv.org/abs/2104.00950

Rosone, M., Ziccarelli, M., Ferrari, A., & Farulla, C. A. (2018). On the reactivation of a large landslide induced by rainfall in highly fissured clays. Engineering Geology, 235, 20-38. https://doi.org/10.1016/j.enggeo.2018.01.016

Rutqvist, J., & Stephansson, O. (2003). The role of hydromechanical coupling in fractured rock engineering. Hydrogeology Journal, 11(1), 7-40. https://doi.org/10.1007/s10040-002-0241-5

S

Sagi, O., & Rokach, L. (2018). Ensemble learning : A survey. *WIREs Data Mining and Knowledge Discovery, 8*(4), e1249. https://doi.org/10.1002/widm.1249

Sahin, E. K. (2020). Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. *SN Applied Sciences, 2*(7). https://doi.org/10.1007/s42452-020-3060-1

Sahu, R. K., Müller, J., Park, J., Varadharajan, C., Arora, B., Faybishenko, B., & Agarwal, D. (2020). Impact of Input Feature Selection on Groundwater Level Prediction From a Multi-Layer Perceptron Neural Network. *Frontiers in Water, 2*, 573034. https://doi.org/10.3389/frwa.2020.573034

Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications, 145*(2), 280-297. https://doi.org/10.1016/S0010-4655(02)00280-1

Samodra, G., Ngadisih, & Nugroho, F. S. (2024). Benchmarking data handling strategies for landslide susceptibility modeling using random forest workflows. *Artificial Intelligence in Geosciences, 5*, 100093. https://doi.org/10.1016/j.aiig.2024.100093

Schulz, W. H., Coe, J. A., Ricci, P. P., Smoczyk, G. M., Shurtleff, B. L., & Panosky, J. (2017). Landslide kinematics and their potential controls from hourly to decadal timescales : Insights from integrating ground-based InSAR measurements with structural maps and long-term monitoring data. *Geomorphology, 285*, 121-136. https://doi.org/10.1016/j.geomorph.2017.02.011

Schulz, W. H., Smith, J. B., Wang, G., Jiang, Y., & Roering, J. J. (2018). Clayey Landslide Initiation and Acceleration Strongly Modulated by Soil Swelling. *Geophysical Research Letters, 45*(4), 1888-1896. https://doi.org/10.1002/2017GL076807

Schwartz, S., Zerathe, S., Jongmans, D., Baillet, L., Carcaillet, J., Audin, L., Dumont, T., Bourlès, D., Braucher, R., & Lebrouc, V. (2017). Cosmic ray exposure dating on the large landslide of Séchilienne (Western Alps) : A

synthesis to constrain slope evolution. *Geomorphology, 278*, 329-344. https://doi.org/10.1016/j.geomorph.2016.11.014

Schwemmle, R., Demand, D., & Weiler, M. (2021). Technical note : Diagnostic efficiency – specific evaluation of model performance. *Hydrology and Earth System Sciences, 25*(4), 2187-2198. https://doi.org/10.5194/hess-25-2187-2021

Segoni, S., Piciullo, L., & Gariano, S. L. (2018). A review of the recent literature on rainfall thresholds for landslide occurrence. *Landslides, 15*(8), 1483-1501. https://doi.org/10.1007/s10346-018-0966-4

Shao, P., Wang, H., Long, G., Liao, J., Gan, F., Xu, B., Hu, K., & Teng, Y. (2024). A novel multi-step ahead prediction method for landslide displacement based on autoregressive integrated moving average and intelligent algorithm. *Engineering Applications of Artificial Intelligence, 137*, 109107. https://doi.org/10.1016/j.engappai.2024.109107

Shumway, R. H., & Stoffer, D. S. (2011). *Time series analysis and its applications: With R examples* (3rd ed.). Springer. https://doi.org/10.1007/978-1-4419-7865-3

Stampoulis, D., & Anagnostou, E. N. (2012). Evaluation of Global Satellite Rainfall Products over Continental Europe. *Journal of Hydrometeorology, 13*(2), 588-603. https://doi.org/10.1175/JHM-D-11-086.1

Strauhal, T., Loew, S., Holzmann, M., & Zangerl, C. (2016). Detailed hydrogeological analysis of a deep-seated rockslide at the Gepatsch reservoir (Klasgarten, Austria). *Hydrogeology Journal, 24*(2), 349-371. https://doi.org/10.1007/s10040-015-1341-3

Strozzi, T., Delaloye, R., Kääb, A., Ambrosi, C., Perruchoud, E., & Wegmüller, U. (2010). Combined observations of rock mass movements using satellite SAR interferometry, differential GPS, airborne digital photogrammetry, and airborne photography interpretation. *Journal of Geophysical Research: Earth Surface, 115*(F1), F01014. https://doi.org/10.1029/2009JF001311

Stumpf, A., Malet, J.-P., & Delacourt, C. (2017). Correlation of satellite image time-series for the detection and monitoring of slow-moving landslides. *Remote Sensing of Environment, 189*, 40-55. https://doi.org/10.1016/j.rse.2016.11.007

Sun, D., Wang, J., Wen, H., Ding, Y., & Mi, C. (2024). Landslide susceptibility mapping (LSM) based on different boosting and hyperparameter optimization algorithms : A case of Wanzhou District, China. *Journal of Rock Mechanics and Geotechnical Engineering, 16*(8), 3221-3232. https://doi.org/10.1016/j.jrmge.2023.09.037

Sun, H.-Y., Pan, P., Lü, Q., Wei, Z.-L., Xie, W., & Zhan, W. (2019). A case study of a rainfall-induced landslide involving weak interlayer and its treatment using the siphon drainage method. *Bulletin of Engineering Geology and the Environment, 78*(6), 4063-4074. https://doi.org/10.1007/s10064-018-1365-8

Sun, J., Hu, L., Li, D., Sun, K., & Yang, Z. (2022). Data-driven models for accurate groundwater level prediction and their practical significance in groundwater management. *Journal of Hydrology, 608*, 127630. https://doi.org/10.1016/j.jhydrol.2022.127630

T-U

Takefuji, Y. (2025). Beyond XGBoost and SHAP : Unveiling true feature importance. *Journal of Hazardous Materials, 488*, 137382. https://doi.org/10.1016/j.jhazmat.2025.137382

Tarwidi, D., Pudjaprasetya, S. R., Adytia, D., & Apri, M. (2023). An optimized XGBoost-based machine learning method for predicting wave run-up on a sloping beach. MethodsX, 10, 102119. https://doi.org/10.1016/j.mex.2023.102119

Tawakuli, A., Havers, B., Gulisano, V., Kaiser, D., & Engel, T. (2025). Survey: Time-series data preprocessing : A survey and an empirical analysis. *Journal of Engineering Research, 13*(2), 674-711. https://doi.org/10.1016/j.jer.2024.02.018

Terzaghi, K. (1951). *Mechanism of landslides* (Harvard Soil Mechanics Series No. 36). Harvard University, Department of Engineering.

Travelletti, J., Sailhac, P., Malet, J.-P., Grandjean, G., & Ponton, J. (2012). Hydrological response of weathered clay-shale slopes : Water infiltration monitoring with time-lapse electrical resistivity tomography. *Hydrological Processes, 26*(14), 2106-2119. https://doi.org/10.1002/hyp.7983

Uhlemann, S., Smith, A., Chambers, J., Dixon, N., Dijkstra, T., Haslam, E., Meldrum, P., Merritt, A., Gunn, D., & Mackay, J. (2016). Assessment of ground-based monitoring techniques applied to landslide investigations. *Geomorphology, 253*, 438-451. https://doi.org/10.1016/j.geomorph.2015.10.027

Ullrich, S. L., Hegnauer, M., Nguyen, D. V., Merz, B., Kwadijk, J., & Vorogushyn, S. (2021). Comparative evaluation of two types of stochastic weather generators for synthetic precipitation in the Rhine basin. *Journal of Hydrology, 601*, 126544. https://doi.org/10.1016/j.jhydrol.2021.126544

V

Vaiyapuri, T. (2025). An Optuna-Based Metaheuristic Optimization Framework for Biomedical Image Analysis. *Engineering, Technology & Applied Science Research, 15*(4), 24382-24389. https://doi.org/10.48084/etasr.11234

Vallet, A. (2014). *Mouvements de fluides et processus de déstabilisation des versants alpins : Apport de l'étude de l'instabilité de Séchilienne.* (Thèse de doctorat). Université de Franche-Comté, Besançon, France.

Vallet, A., Bertrand, C., Fabbri, O., & Mudry, J. (2015). A new method to compute groundwater recharge for the study of rainfall-triggered deep-seated landslides: Application to the Séchilienne unstable slope (Western Alps). *Hydrology and Earth System Sciences, 19*, 427–449. https://doi.org/10.5194/hess-19-427-2015

Vallet, A., Bertrand, C., Mudry, J., Bogaard, T., Fabbri, O., Baudement, C., & Régent, B. (2015). Contribution of time-related environmental tracing combined with tracer tests for characterization of a groundwater conceptual model : A case study at the Séchilienne landslide, western Alps (France). *Hydrogeology Journal, 23*(8), 1761-1779. https://doi.org/10.1007/s10040-015-1298-2

Vallet, A., Charlier, J. B., Fabbri, O., Bertrand, C., Carry, N., & Mudry, J. (2016). Functioning and precipitation-displacement modelling of rainfall-induced deep-seated landslides subject to creep deformation. *Landslides, 13*(4), 653-670. https://doi.org/10.1007/s10346-015-0592-3

Van Asch, Th. W. J., Buma, J., & Van Beek, L. P. H. (1999). A view on some hydrological triggering systems in landslides. *Geomorphology, 30*(1-2), 25-32. https://doi.org/10.1016/S0169-555X(99)00042-2

van Asch, Th. W. J., Hendriks, M. R., Hessel, R., & Rappange, F. E. (1996). Hydrological triggering conditions of landslides in varved clays in the French Alps. *Engineering Geology, 42*(4), 239-251. https://doi.org/10.1016/0013-7952(95)00082-8

van Asch, Th. W. J., Malet, J.-P., & Bogaard, T. A. (2009). The effect of groundwater fluctuations on the velocity pattern of slow-moving landslides. *Natural Hazards and Earth System Sciences, 9*(3), 739-749. https://doi.org/10.5194/nhess-9-739-2009

van Asch, Th. W. J., Van Beek, L. P. H., & Bogaard, T. A. (2007). Problems in predicting the mobility of slow-moving landslides. *Engineering Geology, 91*(1), 46-55. https://doi.org/10.1016/j.enggeo.2006.12.012

van Natijne, A. L., Bogaard, T. A., Zieher, T., Pfeiffer, J., & Lindenbergh, R. C. (2023). Machine-learning-based nowcasting of the Vögelsberg deep-seated landslide: Why predicting slow deformation is not so easy. *Natural Hazards and Earth System Sciences, 23*, 3723–3745. https://doi.org/10.5194/nhess-23-3723-2023

van Natijne, A. L., Lindenbergh, R. C., & Bogaard, T. A. (2020). Machine Learning : New Potential for Local and Regional Deep-Seated Landslide Nowcasting. *Sensors, 20*(5), 1425. https://doi.org/10.3390/s20051425

Varnes DJ (1954). Landslide types and processes. In: Eckel EB (Ed.), *Landslides and engineering practice, special report 28*. Highway research board. National Academy of Sciences, Washington, DC, pp. 20–47

Varnes DJ (1978). Slope movement types and processes. In: Schuster RL, Krizek RJ (Eds), *Landslides, analysis and control, special report 176*: Transportation research board, National Academy of Sciences, Washington, DC., pp. 11–33

Vassallo, R., Grimaldi, G. M., & Di Maio, C. (2015). Pore water pressures induced by historical rain series in a clayey landslide : 3D modeling. *Landslides, 12*(4), 731-744. https://doi.org/10.1007/s10346-014-0508-7

Vengeon, J.-M. (1998). *Déformation et rupture des versants en terrain métamorphique anisotrope. Apport de l' étude des Ruines de Séchilienne* (Thèse de doctorat). Université Joseph Fourier (Grenoble I), Grenoble, France.

Vu, M. T., Jardani, A., Massei, N., & Fournier, M. (2021). Reconstruction of missing groundwater level data by using Long Short-Term Memory (LSTM) deep neural network. *Journal of Hydrology, 597*, 125776. https://doi.org/10.1016/j.jhydrol.2020.125776

Vulliet, L., & Hutter, K. (1988). Viscous-type sliding laws for landslides. *Canadian Geotechnical Journal, 25*(3), 467–477. https://doi.org/10.1139/t88-052

W

Wang, J., Nie, G., Gao, S., Wu, S., Li, H., & Ren, X. (2021). Landslide Deformation Prediction Based on a GNSS Time Series Analysis and Recurrent Neural Network Model. *Remote Sensing, 13*(6), 1055. https://doi.org/10.3390/rs13061055

Wang, J., Zhu, H., Zhang, W., Tan, D., & Pasuto, A. (2025). Enhancing Landslide Displacement Prediction Using a Spatio-Temporal Deep Learning Model With Interpretable Features. *Journal of Geophysical Research: Machine Learning and Computation, 2*(3), e2025JH000592. https://doi.org/10.1029/2025JH000592

Wang, L., Wu, C., Yang, Z., & Wang, L. (2023). Deep learning methods for time-dependent reliability analysis of reservoir slopes in spatially variable soils. *Computers and Geotechnics, 159*, 105413. https://doi.org/10.1016/j.compgeo.2023.105413

Wang, R., Zhang, K., Qi, J., Xu, W., Long, Y., & Huang, H. (2022). A Prediction Model of Hydrodynamic Landslide Evolution Process Based on Deep Learning Supported by Monitoring Big Data. *Frontiers in Earth Science, 10*, 829221. https://doi.org/10.3389/feart.2022.829221

Wang, Y., Tang, H., Huang, J., Wen, T., Ma, J., & Zhang, J. (2022). A comparative study of different machine learning methods for reservoir landslide displacement prediction. *Engineering Geology, 298*, 106544. https://doi.org/10.1016/j.enggeo.2022.106544

Weerasinghe, A. R. P., Bandara, K. N., & Gunathilake, J. (2019). Application of tank model and quantitative assessment to predict rainfall induced displacement on landslide – Case study on Nawalapitiya landslide, Sri Lanka. *Ceylon Journal of Science, 48*(4), 311-318. https://doi.org/10.4038/cjs.v48i4.7670

Wei, Z., Lü, Q., Sun, H., & Shang, Y. (2019). Estimating the rainfall threshold of a deep-seated landslide by integrating models for predicting the groundwater level and stability analysis of the slope. Engineering Geology, 253, 14-26. https://doi.org/10.1016/j.enggeo.2019.02.026

Wienhöfer, J., Lindenmaier, F., & Zehe, E. (2011). Challenges in Understanding the Hydrologic Controls on the Mobility of Slow-Moving Landslides. *Vadose Zone Journal, 10*(2), 496-511. https://doi.org/10.2136/vzj2009.0182

Willmott, C. J., Robeson, S. M., & Matsuura, K. (2012). A refined index of model performance. *International Journal of Climatology, 32*(13), 2088-2094. https://doi.org/10.1002/joc.2419

Winkler, R. L. (1972). A Decision-Theoretic Approach to Interval Estimation. Journal of the American Statistical Association, 67(337), 187-191. https://doi.org/10.1080/01621459.1972.10481224

Wójcik, R., Pilarski, M., & Miętus, M. (2014). Statistical downscaling of probability density function of daily precipitation on the Polish coast. *Meteorology Hydrology and Water Management, 2*(1), 27-36. https://doi.org/10.26491/mhwm/21590

Wu, D., Zhou, B., & Zimin, M. (2025). Prediction of landslide displacement based on the CA-stacked transformer model. *Alexandria Engineering Journal, 124*, 389-403. https://doi.org/10.1016/j.aej.2025.03.140

X-Z

Xiao, X., Zou, Y., Huang, J., Luo, X., Yang, L., Li, M., Yang, P., Ji, X., & Li, Y. (2024). An interpretable model for landslide susceptibility assessment based on Optuna hyperparameter optimization and Random Forest. Geomatics, Natural Hazards and Risk, 15(1), 2347421. https://doi.org/10.1080/19475705.2024.2347421

Xie, P., Zhou, A., & Chai, B. (2019). The Application of Long Short-Term Memory (LSTM) Method on Displacement Prediction of Multifactor-Induced Landslides. IEEE Access, 7, 54305-54311. https://doi.org/10.1109/ACCESS.2019.2912419

Xing, Y., Yue, J., Chen, C., Cong, K., Zhu, S., & Bian, Y. (2019). Dynamic Displacement Forecasting of Dashuitian Landslide in China Using Variational Mode Decomposition and Stack Long Short-Term Memory Network. Applied Sciences, 9(15), 2951. https://doi.org/10.3390/app9152951

Xu, J., Jiang, Y., & Yang, C. (2022). Landslide Displacement Prediction during the Sliding Process Using XGBoost, SVR and RNNs. Applied Sciences, 12(12), 6056. https://doi.org/10.3390/app12126056

Xu, S., & Niu, R. (2018). Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China. Computers & Geosciences, 111, 87-96. https://doi.org/10.1016/j.cageo.2017.10.013

Yang, B., Xiao, T., Wang, L., & Huang, W. (2022). Using Complementary Ensemble Empirical Mode Decomposition and Gated Recurrent Unit to Predict Landslide Displacements in Dam Reservoir. Sensors, 22(4), 1320. https://doi.org/10.3390/s22041320

Yang, B., Yin, K., Lacasse, S., & Liu, Z. (2019). Time series analysis and long short-term memory neural network to predict landslide displacement. Landslides, 16(4), 677-694. https://doi.org/10.1007/s10346-018-01127-x

Yang, C., Yin, Y., Zhang, J., Ding, P., & Liu, J. (2024). A graph deep learning method for landslide displacement prediction based on global navigation satellite system positioning. Geoscience Frontiers, 15(1), 101690. https://doi.org/10.1016/j.gsf.2023.101690

Yao, W., Zeng, Z., Lian, C., & Tang, H. (2015). Training enhanced reservoir computing predictor for landslide displacement. Engineering Geology, 188, 101-109. https://doi.org/10.1016/j.enggeo.2014.11.008

Zangerl, C., Eberhardt, E., & Perzlmaier, S. (2010). Kinematic behaviour and velocity characteristics of a complex deep-seated crystalline rockslide system in relation to its interaction with a dam reservoir. Engineering Geology, 112(1-4), 53-67. https://doi.org/10.1016/j.enggeo.2010.01.001

Zhang, H., Song, Y., Xu, S., He, Y., Li, Z., Yu, X., Liang, Y., Wu, W., & Wang, Y. (2022). Combining a class-weighted algorithm and machine learning models in landslide susceptibility mapping : A case study of Wanzhou section of the Three Gorges Reservoir, China. Computers & Geosciences, 158, 104966. https://doi.org/10.1016/j.cageo.2021.104966

Zhang, W., Li, H., Tang, L., Gu, X., Wang, L., & Wang, L. (2022). Displacement prediction of Jiuxianping landslide using gated recurrent unit (GRU) networks. Acta Geotechnica, 17(4), 1367-1382. https://doi.org/10.1007/s11440-022-01495-8

Zhang, Y., Chen, X., Liao, R., Wan, J., He, Z., Zhao, Z., Zhang, Y., & Su, Z. (2021). Research on displacement prediction of step-type landslide under the influence of various environmental factors based on intelligent WCA-ELM in the Three Gorges Reservoir area. Natural Hazards, 107(2), 1709-1729. https://doi.org/10.1007/s11069-021-04655-3

Zheng, H., Duan, M., Sun, X., Han, G., & Jin, Q. (2022). Landslide Displacement Prediction Based on Transfer Learning and Bi-GRU. Journal of Sensors, 2022, 1-12. https://doi.org/10.1155/2022/4659853

Zheng, Y., Huang, D., Fan, X., & Shi, L. (2024). Groundwater Level Prediction for Landslides Using an Improved TANK Model Based on Big Data. Water, 16(16), 2286. https://doi.org/10.3390/w16162286

Zou, Z., Yang, Y., Fan, Z., Tang, H., Zou, M., Hu, X., Xiong, C., & Ma, J. (2020). Suitability of data preprocessing methods for landslide displacement forecasting. Stochastic Environmental Research and Risk Assessment, 34(8), 1105-1119. https://doi.org/10.1007/s00477-020-01824-

APPENDIX – RESUME LONG

A.I. Contexte et objectifs de la thèse

Les glissements de terrain sont des mouvements gravitaires constituant un aléa majeur pour les populations, les infrastructures, et les activités socio-économiques à l'échelle mondiale *(Froude et Petley, 2018)*. Leur occurrence dépends de la combinaison de facteurs de prédisposition (ex., topographie, fracturation, lithologie), de facteurs préparatoires (ex., relâchement des contraintes après le retrait glaciaire) favorisant la susceptibilité à l'instabilité du versant, ainsi que de facteurs de déclenchement (ex., fortes pluies, fonte de neige, activité sismique) agissant sur le court terme et entrainant l'accélération du mouvement *(Vallet, 2014)*.

En particulier, l'infiltration des eaux pluviales et nivales agissent directement sur les variations du niveau piezométrique au sein des instabilités, et par conséquent sur les fluctuations des pressions interstitielles *(Rutqvist and Stpehansson 2003 ; Therzaghi, 1951)*, favorisant l'instabilité et représentant ainsi un mécanisme de contrôle clé dans la dynamique des mouvements des glissements de terrain *(Binet, 2006)*.

La prévision des déplacements et leur relation avec les forçages hydro-météorologiques demeurent toutefois complexe et non-linéaire en raison de l'anisotropie et de l'hétérogénéité spatio-temporelle des propriétés hydro-mécaniques des versants instables, notamment pour les glissements de terrain lents et profonds *(Brièvre et al., 2012 ; van Asch et al., 1996, Schulz et al., 2017)*. Ces derniers se caractérisent par des mouvements lents et progressifs ($cm.an^{-1}$ à $m.an^{-1}$) ponctuées par des phases d'accélérations périodiques *(Corominas et al., 2005)* contrôlées par l'évolution spatio-temporelle des pressions interstitielles dans le versant à plusieurs échelles de temps.

La prévision des déplacements implique l'acquisition de données multi-sources de surveillance (déplacement, niveau piezométrique, événements météorologiques) au travers de plusieurs méthodes (in-situ, satellitetaire, modélisation) afin d'élaborer des modèles efficaces et précis. Les approches de modélisation mobilisées pour reproduire la dynamique des glissements de terrain se répartissent classiquement en deux familles : les modèles physiques et les modèles basés sur les données, employées selon les niveaux de description et les disponibilités des données (**Figure A1**)

Les modèles physiques visent à reproduire le comportement du glissement en représentant explicitement les propriétés géomorphologiques, géologiques, hydrologiques, et mécaniques, au travers de modèle analytiques (ex., modèle de pente infinie) ou numériques (ex., FEM, FDM) *(Herrera et al., 2009 ; Conte et al., 2014 ; Havenith et al., 2017)*. Toutefois, leur mise en œuvre nécessite une caractérisation détaillée du milieu et des jeux de données souvent difficiles à acquérir sur le terrain. De

plus, leur coût de calibration et de calcul limite leur applicabilité seulement aux sites les mieux contraints.

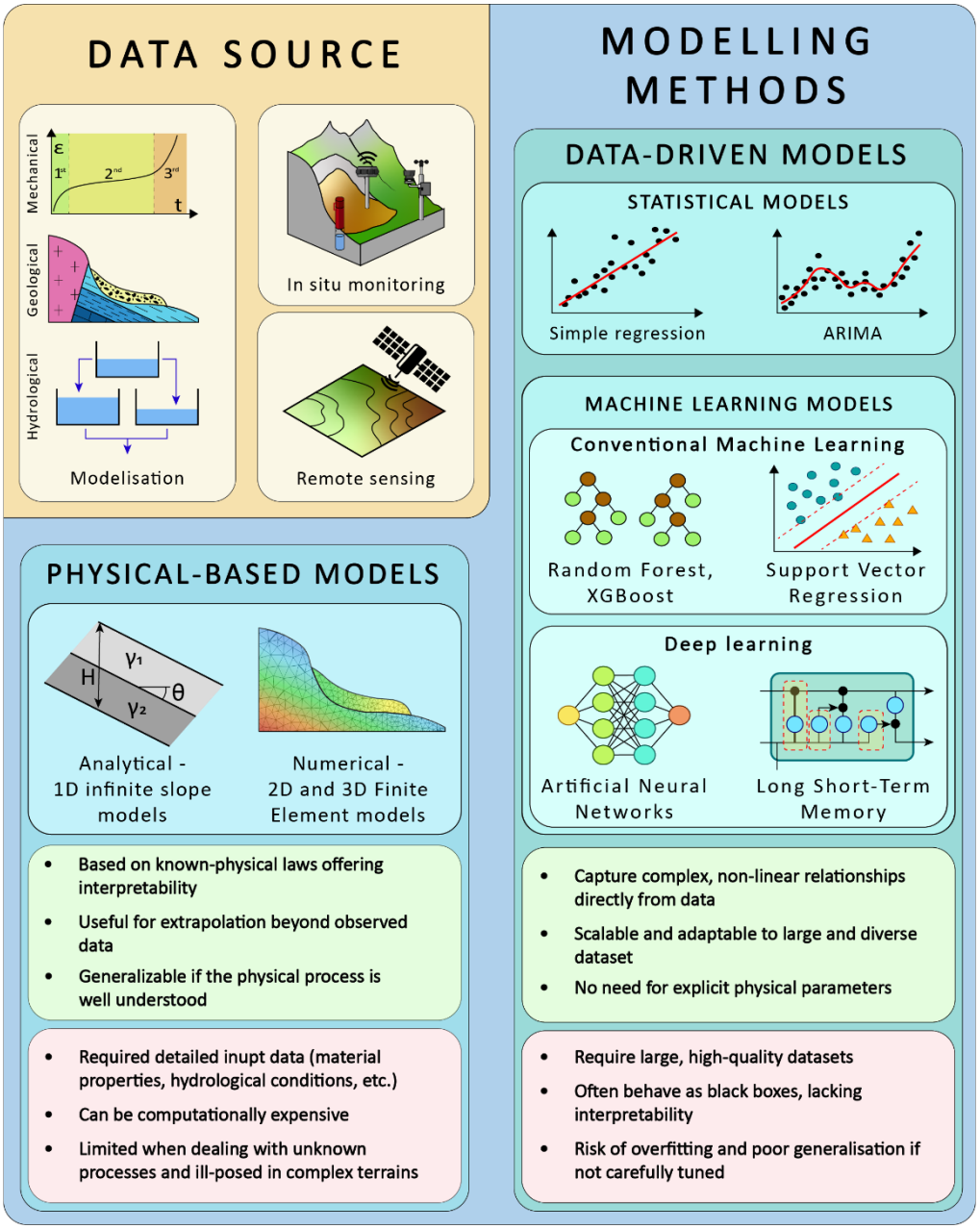


Figure A61 | Représentation synthétique des méthodes d'acquisition des données de surveillance des glissements de terrain et des approches de modélisation des déplacements. Les encadrés verts et rouges résument les principaux avantages et inconvénients de chaque méthode.

A l'inverse, les approches data-driven contournent la nécessité de représenter explicitement l'ensemble des mécanismes internes en apprenant directement les relations entre variables explicatives et variables cibles à partir de séries temporelles multivariées. Ces approches sont particuliérement attractives lorsque les relations forçages-réponse sont complexes, non-linéaires et difficiles à formaliser dans un cadre physique unique. Leur robustesse dépend toutefois fortement de la qualité, de la continuité

et de la représentativité des séries utilisées pour l'apprentissage, ainsi que de la capacité des variables construites à encoder la mémoire du systèmen les effets de cumul, les délais de réponse et les changements d'état hydrologique. Les modèles d'intelligence artificielle (IA), et en particulier les modèles de deep-learning (DL ; ex., LSTM, GRU) ont montré leur capacité à reproduire des comportements non-linéraires et multi-temporels *(Xie et al. 2019 ; Meng et al., 2024 ; Nava et al. 2023)*. Elles restent néanmoins fortement dépendants du volume de données disponibles et encore limitées en terme d'interprétabilité et de transférabilité.

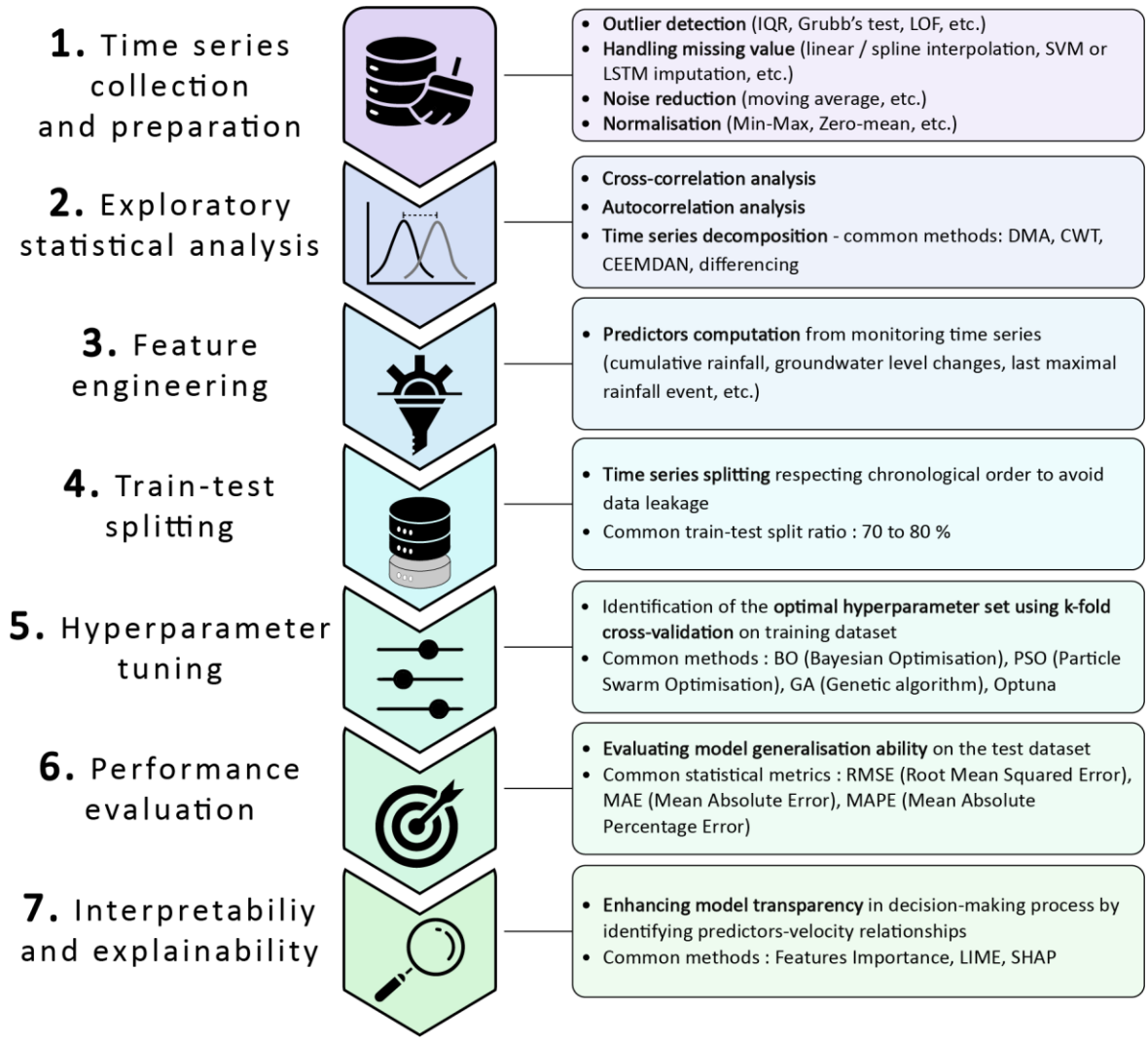


Figure A62 | Workflow de traitement des séries temporelles et de mise en œuvre d'un modèle de machine learning conventionnel.

Comme alternative, les approches de machine learning conventionnels (CML), et notamment les modèles d'ensembles basés sur les arbres de décision (Random Forest ; RF, et eXtreme Gradient Boosting ; XGBoost) présentent un compromis méthodologique incluant une bonne robustesse vis-à-vis des jeux de données imparfaits, une capacité à reproduire des relations non linéaires complexes, et un niveau d'interprétabilité supérieur à celui des architectures profondes *(Krkač et al., 2017)*. Leur

efficacité dépend toutefois de la construction de prédicteurs capables de traduire explicitement la dépendance temporelle, par exemple sous la forme de variables retardées, de cumuls glissants, d'indicateurs d'événements ou de descripteurs d'état hydrologique. En ce sens, la performance du modèle ne repose pas uniquement sur l'algorithme choisi, mais également sur la qualité du prétraitement, sur la représentativité des chroniques d'apprentissage, et sur la cohérence physique des variables dérivées à partir des forçages observés.

C'est dans cette perspective que s'inscrit cette thèse qui repose sur trois fondements méthodologiques complémentaires, suivant le workflow proposé en **Figure A2**. Le premier est que toute simulation robuste des déplacements nécessite des séries temporelles longues, informatives et représentatives, en particulier pour les variables décrivant l'état hydrologique du système. Le second est que les prédicteurs doivent encoder explicitement la dépendance temporelle et la mémoire du versant, afin de restituer les réponses multi-échelles et retardées qui caractérisent les glissements lents. Le troisième est que les approches de modélisation retenues doivent pouvoir être examinées au-delà de leur seule performance prédictive, de manière à expliciter les relations forçage-réponse en termes de contrôles dominants, de non-linéarités et d'échelles de temps caractéristiques.

Dans ce cadre, la thèse s'organise autour de trois questions principales :

- Comment acquérir, compléter ou reconstruire des séries temporelles multi-sources longues, informatives et représentatives, capables de soutenir un apprentissage robuste dans des contextes où les données hydrogéologiques sont souvent incomplètes ?
- Comment les forçages hydro-météorologiques contrôlent-ils la dynamique temporelle des déplacements de surface, et selon quelles échelles de temps ?
- Comment les déplacements peuvent-ils évoluer sous des conditions hydro-météorologiques contrastées mais plausibles, dans une perspective de simulation probabiliste orientée vers l'aléa ?

L'ensemble vise ainsi à développer un cadre méthodologique générique combinant reconstruction de l'état hydrologique, simulation explicable des vitesses et propagation scénarisée de la variabilité hydro-météorologique.

Pour répondre à ces questions, plusieurs glissements de terrain lents instrumentés sont étudiés dans des contextes contrastés : Séchilienne (Alpes françaises), Le Mont de La Saxe (Alpes italiennes), Villerville (côte normande), Viella (Pyrénées françaises), et Vallcebre (Pyrénées espagnoles). Ces sites diffèrent par leur lithologie, leur géométrie, leur régime hydrologique, la nature de leurs déplacements et les données disponibles, et permettent de tester la robustesse, la transférabilité et l'interprétabilité des approches proposées.

A.II. Chapitre IV - Reconstruction de proxys hydrogéologiques de glissements de terrain lents par modélisation conceptuelle pluie-débit

A.II.1. Introduction

Ce chapitre répond à la première question de la thèse, qui porte sur l'acquisition et la reconstruction de séries temporelles multi-sources longues et représentatives pour l'apprentissage automatique appliqué aux glissements lents. La disponibilité de chroniques hydrogéologiques continues constitue donc une condition essentielle pour construire des prédicteurs capables de représenter correctement la mémoire du système et ses réponses retardées. Or, dans les glissements lents et profonds, le suivi piézométrique de long terme reste difficile à mettre en œuvre en raison des contraintes de forage, de maintenance, d'accessibilité et de la dégradation fréquente des instruments sous l'effet de la déformation.

Pour répondre à cette limitation, le chapitre développe une stratégie de reconstruction indirecte de l'état hydrologique fondée sur des modèles conceptuels de type réservoir, calibrés sur la relation pluie-débit. L'objectif est de reconstituer des proxys cohérents de l'état hydrologique interne du glissement à partir de chroniques de débit observées à l'exutoire. L'approche est mise en œuvre avec la plateforme KarstMod appliquée à des aquifères fracturés au sein de glissements profonds. Elle est appliquée à deux glissements lents des Alpes occidentales, Séchilienne et Mont de La Saxe, afin de caractériser leur fonctionnement hydrologique, d'évaluer la robustesse des structures conceptuelles proposées, et d'explorer leur capacité à prolonger les séries hydrologiques au-delà des périodes instrumentées.

A.II.2. Matériels et méthodes

L'étude repose sur deux sites développés dans des contextes cristallins fracturés à faible perméabilité : le glissement de Séchilienne et le glissement du Mont de La Saxe. Les deux sites disposent de chroniques de pluie, de température et de débit acquises à fréquence journalière sur plusieurs années (9 et 6 ans respectivement).

La première étape méthodologique consiste à analyser le comportement hydrogéologique des exutoires à partir des courbes de débits classés (**Figure A3**). Cette analyse met en évidence des signatures de fonctionnement contrastées entre les deux glissements. À Séchilienne, la courbe traduit un comportement en trois phases, cohérent avec un système disposant d'une capacité de stockage et d'une composante de débit de base entretenue sur les périodes de basses eaux. Au Mont de La Saxe, la courbe révèle un comportement en deux phases, marqué par l'absence d'un stockage suffisant pour soutenir

durablement l'écoulement et par un contrôle plus direct des débits par les apports rapides, notamment lors de la fonte nivale. Ces diagnostics guident la définition de structures réservoirs à chaque site.

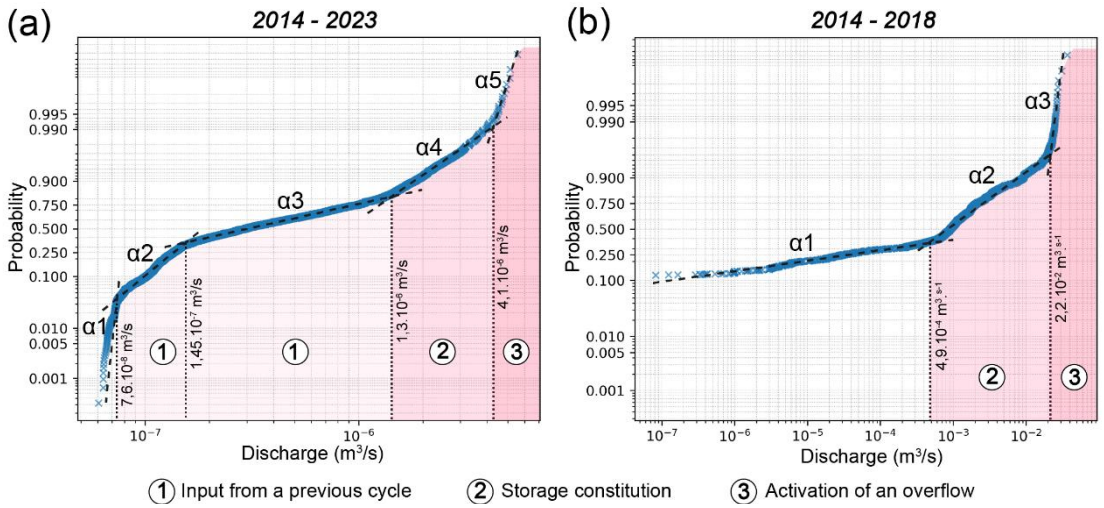


Figure A63 | Courbes de débits classés établies pour les glissements de terrain de **(a)** Séchilienne et **(b)** Mont de La Saxe. Les pentes sont représentées par alpha, les ruptures de pente sont indiquées par des lignes pointillées accompagnées des valeurs de débit associées, et les nombres indiquent les différentes phases de fonctionnement.

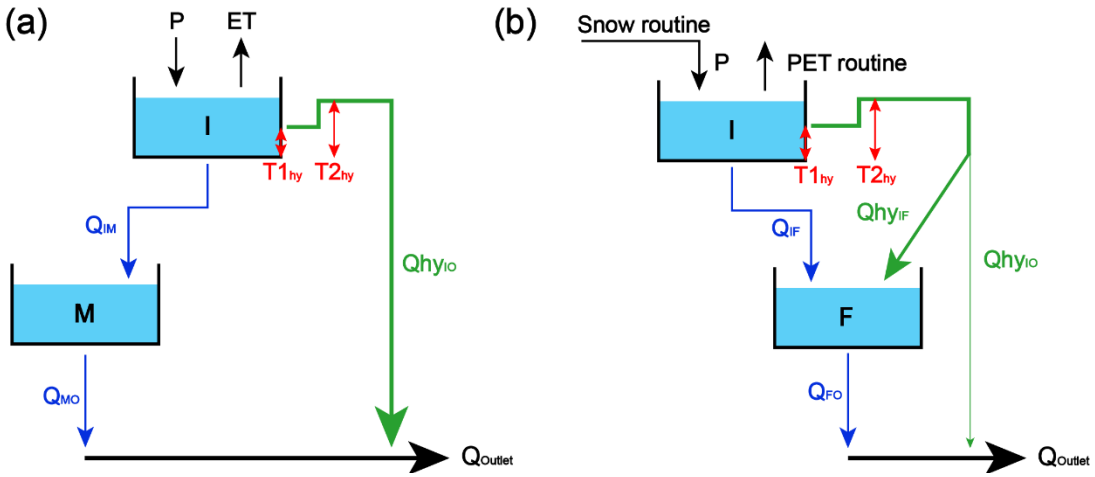


Figure A64 | Structures en réservoirs des modèles de **(a)** Séchilienne et **(b)** Mont de La Saxe. Les réservoirs sont nommés en fonction des mécanismes qu'ils représentent : réservoir d'infiltration (I), réservoir matriciel (M) et réservoir de fracture (F). La nomenclature des flux est définie par les deux réservoirs qu'ils relient, et les paramètres T1hy et T2hy représentent les deux seuils (T) de la fonction d'hystérésis.

La modélisation est ensuite réalisée avec KarstMod, en représentant chaque glissement par un nombre limité de compartiments effectifs reliés par des flux linéaires ou hysterétiques. À Séchilienne, la structure retenue associe un réservoir d'infiltration à un réservoir matriciel lent et à un écoulement direct vers l'exutoire contrôlé par une fonction hysterétique à double seuil, afin de reproduire la coexistence d'une contribution lente liée à la matrice fissurée et d'une réponse rapide liée aux fractures (**Figure A4a**). Au Mont de La Saxe, la structure retenue comprend un réservoir d'infiltration, un

réservoir de drainage rapide et un flux hysterétique, complétés par une routine neige et une routine d'évapotranspiration, nécessaires pour représenter la dynamique de fonte des neiges via une redistribution des valeurs de pluie (**Figure A4b**).

Les simulations sont organisées en périodes de warm-up, calibration et validation, et évaluées au moyen d'une fonction objectif composite combinant le Kling-Gupta Efficiency (KGE) et le Balance Error (BE). Le chapitre complète cette analyse par une étude de sensibilité de type Sobol, un test de la longueur minimale de calibration, et une extension rétrospective des simulations à partir des seules données météorologiques disponibles.

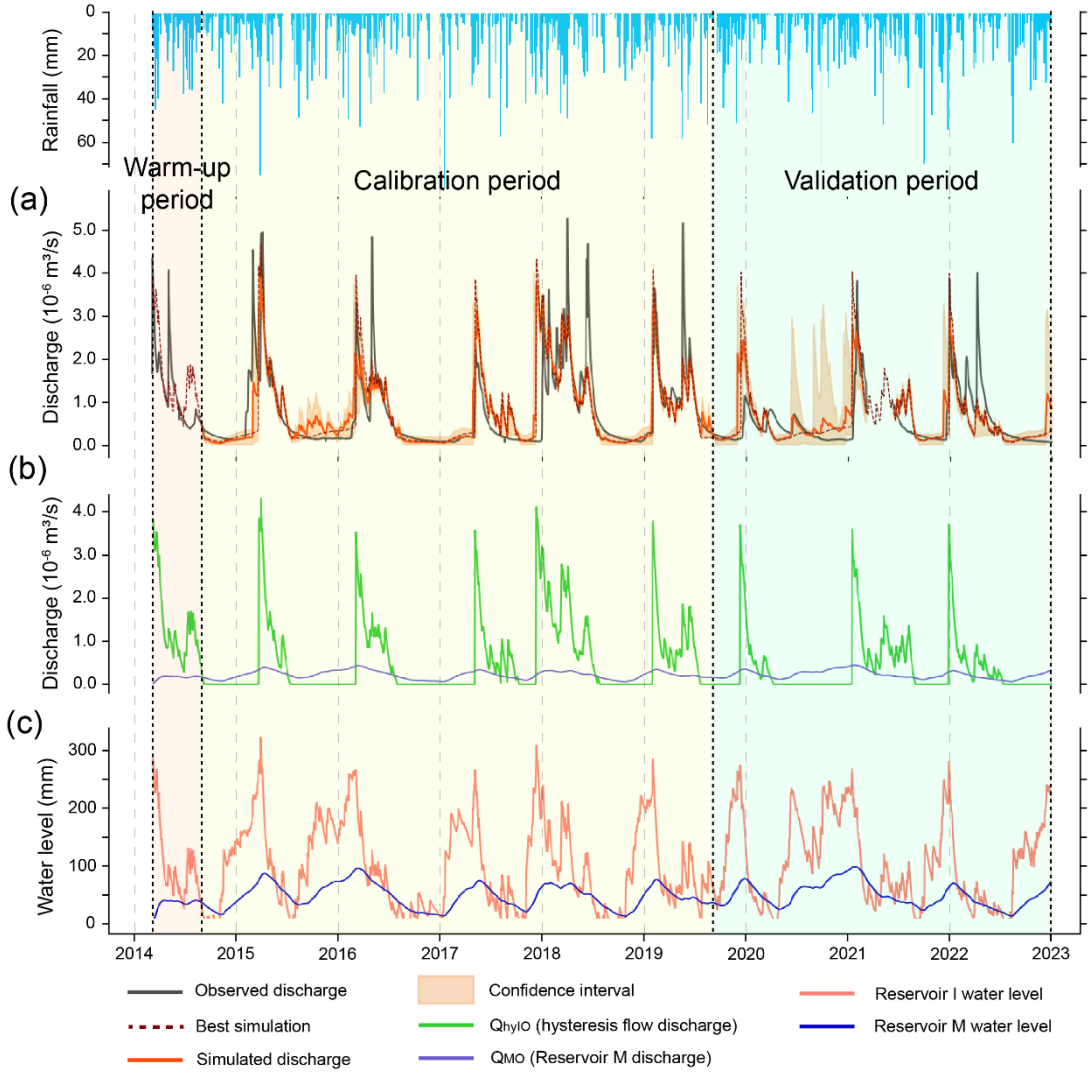


Figure A65 | Résultats de la simulation du modèle de Séchilienne durant les périodes de calibration et de validation. **(a)** Débits observés (noir) et simulés (rouge), avec l'intervalle de confiance de la simulation (zone orange) et la meilleure simulation de débit (ligne brune en pointillés). **(b)** Débits internes représentant les contributions de la matrice (bleu) et des fractures (vert) au débit total simulé. **(c)** Niveaux d'eau dans les réservoirs I (rose) et M (bleu).

A.II.3. Résultats et conclusion

Les résultats montrent que les deux modèles reproduisent de manière satisfaisante les chroniques de débit observées, malgré les contrastes de fonctionnement hydrologique entre les sites. À Séchilienne, la structure calibrée restitue un comportement bimodal, dans lequel une contribution lente de la matrice entretient les débits de base, tandis qu'une contribution plus rapide liée aux fractures contrôle les épisodes de réactivité élevée. Le modèle reproduit correctement la temporalité des pics de débit, mais certaines amplitudes restent plus difficile à reproduire, notamment associées à des apports provenant de systèmes externes connectées ponctuellement en périodes de hautes-eaux à l'aquifère perché du glissement (**Figure A5**). Au Mont de La Saxe, la structure retenue reproduit un fonctionnement plus unimodal, dominé par l'effet de la fonte de neige, avec une amélioration nette de la restitution des pics grâce à l'intégration de la routine de fonte (**Figure A6**). Dans les deux cas, les performances apparaissent satisfaisantes en calibration (0.77 et 0.89) et acceptables en validation (0.62 et 0.65), ce qui confirme la capacité des structures conceptuelles à représenter la dynamique de stockage et de restitution de l'eau dans ces milieux fracturés.

L'analyse de sensibilité montre que les paramètres les plus influents sont principalement les coefficients de récession des flux entre compartiments et la surface de recharge, ce qui confirme le rôle déterminant des vitesses de transfert et du volume effectivement mobilisé dans la réponse de l'exutoire. Les tests sur la longueur de calibration indiquent en outre qu'au moins un cycle hydrologique complet est nécessaire pour stabiliser les performances, et que la réduction de la période d'apprentissage dégrade progressivement la qualité des simulations, en particulier sur les phases de récession et de recharge lente.

Enfin, un apport majeur du chapitre réside dans la possibilité d'extraire, à partir des modèles calibrés, des séries internes non observées directement, en particulier des fluctuations de niveau d'eau dans les réservoirs conceptuels et la contribution relative des différents modes d'écoulement. À Séchilienne, cette approche permet de prolonger la reconstruction hydrologique jusqu'en 1992 à partir des seules chroniques météorologiques, et de documenter rétrospectivement les phases de recharge, de stockage et de réactivité du système avant la mise en place du suivi de débit. Ces sorties ne constituent pas des mesures directes, mais des proxys hydrologiques synthétiques et physiquement cohérents, pouvant d'être intégrés dans les modèles d'IA de simulation des vitesses.

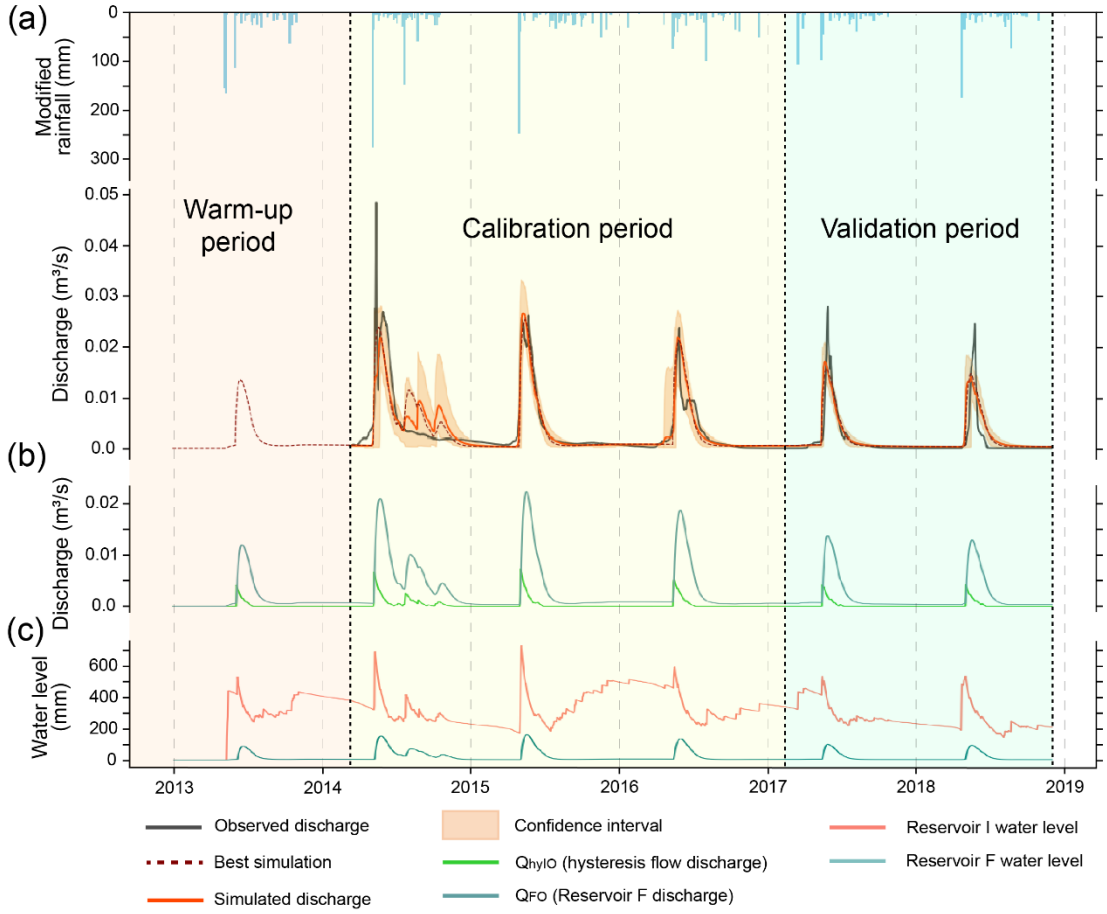


Figure A66 | Résultats de la simulation du modèle de Mont de La Saxe durant les périodes de calibration et de validation. **(a)** Débits observés (noir) et simulés (rouge), avec l'intervalle de confiance de la simulation (zone orange) et la meilleure simulation de débit (ligne brune en pointillés). **(b)** Débits internes représentant les contributions des fractures (bleu) et de l'hystérésis (vert) au débit total simulé. **(c)** Niveaux d'eau dans les réservoirs I (rose) et F (bleu).

Au-delà des résultats propres aux deux sites étudiés, le chapitre montre que la modélisation conceptuelle pluie-débit constitue une stratégie opérationnelle pour pallier le manque de données hydrogéologiques dans les glissements lents et profonds. En condensant l'hétérogénéité interne du massif dans un nombre restreint de compartiments effectifs, les modèles calibrés fournissent une description globale mais robuste du fonctionnement hydrologique du versant. Ils permettent ainsi de compléter, prolonger et interpréter des séries temporelles indispensables aux étapes ultérieures de la thèse. Dans le cadre général du manuscrit, le ***chapitre IV*** alimente donc directement l'étape de préparation des données en fournissant des proxys de l'état hydrologique interne, nécessaires à la construction de prédicteurs pertinents pour les approches de machine learning développées dans les chapitres suivants.

A.III. Chapitre V - Des séries hydro-météorologiques à la simulation interprétable des vitesses : un workflow fondé sur l'XAI

A.III.1. Introduction

Ce chapitre répond à la deuxième question de la thèse et constitue le cœur de la chaîne de modélisation. Son objectif est de reproduire la dynamique journalière des vitesses de surface de glissements lents à partir de chroniques hydro-météorologiques, tout en rendant explicites les relations forçage–réponse apprises par le modèle. Il s'inscrit dans un contexte où la prévision des vitesses reste difficile, car les relations entre forçages environnementaux et déplacements sont multivariées, fortement non linéaires, exprimées sur plusieurs échelles de temps, et marquées par des effets de mémoire et de décalage temporel. En outre, les approches existantes demeurent souvent soit difficiles à interpréter, soit limitées à des démonstrations mono-site, ce qui réduit leur portée méthodologique.

Pour répondre à ces limites, le chapitre développe un workflow d'intelligence artificielle explicable, transférable entre sites, fondé sur XGBoost et sur un ensemble générique de prédicteurs physiquement motivés. L'approche est appliquée à trois glissements lents contrastés – Séchilienne, Viella et Villerville – et à plusieurs points de mesure sur chacun d'eux (**Figure A7**), afin d'évaluer simultanément la robustesse prédictive du modèle, sa capacité à restituer des régimes cinématiques différents, et la cohérence physique des contrôles hydro-météorologiques identifiés. L'objectif est donc, en plus de simuler correctement les vitesses, de caractériser pour chaque site et chaque secteur les variables dominantes, leurs non-linéarités et leurs échelles de temps d'action.

A.III.2. Matériels et méthodes

Le workflow repose sur des séries longues multi-sources de déplacement, de pluie, de pluie efficace et de niveau d'eau souterraine (**Figure A7**). Les données de déplacement sont issues de réseaux de suivi de surface, tandis que les informations hydrologiques proviennent soit d'observations in situ, soit de proxys reconstruits, notamment à Séchilienne. Le cadre méthodologique est organisé en plusieurs modules successifs (**Figure A8**). Le premier est dédié au prétraitement des chroniques de déplacement, avec correction des biais, détection des valeurs aberrantes, lissage adaptatif et interpolation des lacunes. Le second module porte sur l'analyse statistique des séries, en particulier leur stationnarité, et, lorsque nécessaire, sur la séparation entre composante de tendance et composante transitoire. À Séchilienne, où la déformation de long terme domine, cette étape conduit à une décomposition multiplicative et à une segmentation en sous-périodes quasi linéaires.

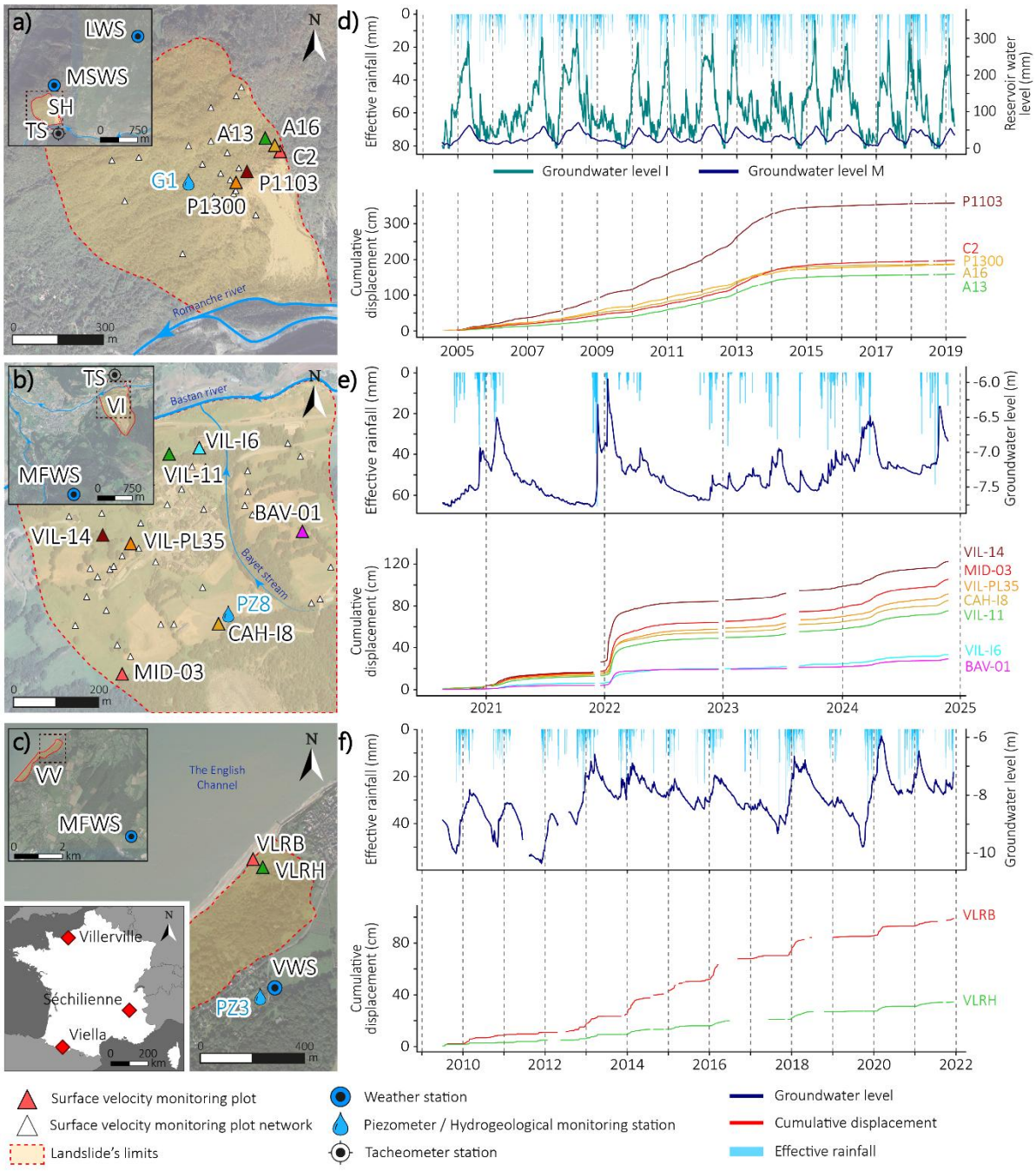


Figure A67 | Vue d'ensemble des trois sites de glissement de terrain : **(a)** Séchilienne (SH), **(b)** Viella (VI) et **(c)** Villerville (VV). Les encarts montrent la distribution spatiale des points de suivi des déplacements, de la station tachéométrique (TS) et des stations météorologiques, incluant Météo-France (MFWS), Mont Sec (MSWS), Luitel (LWS) et Villerville (VWS). **(d–f)** Déplacements cumulés à chaque point de suivi, codés en couleur selon le point de mesure, présentés conjointement avec les séries temporelles correspondantes de pluie efficace et de niveau d'eau souterraine.

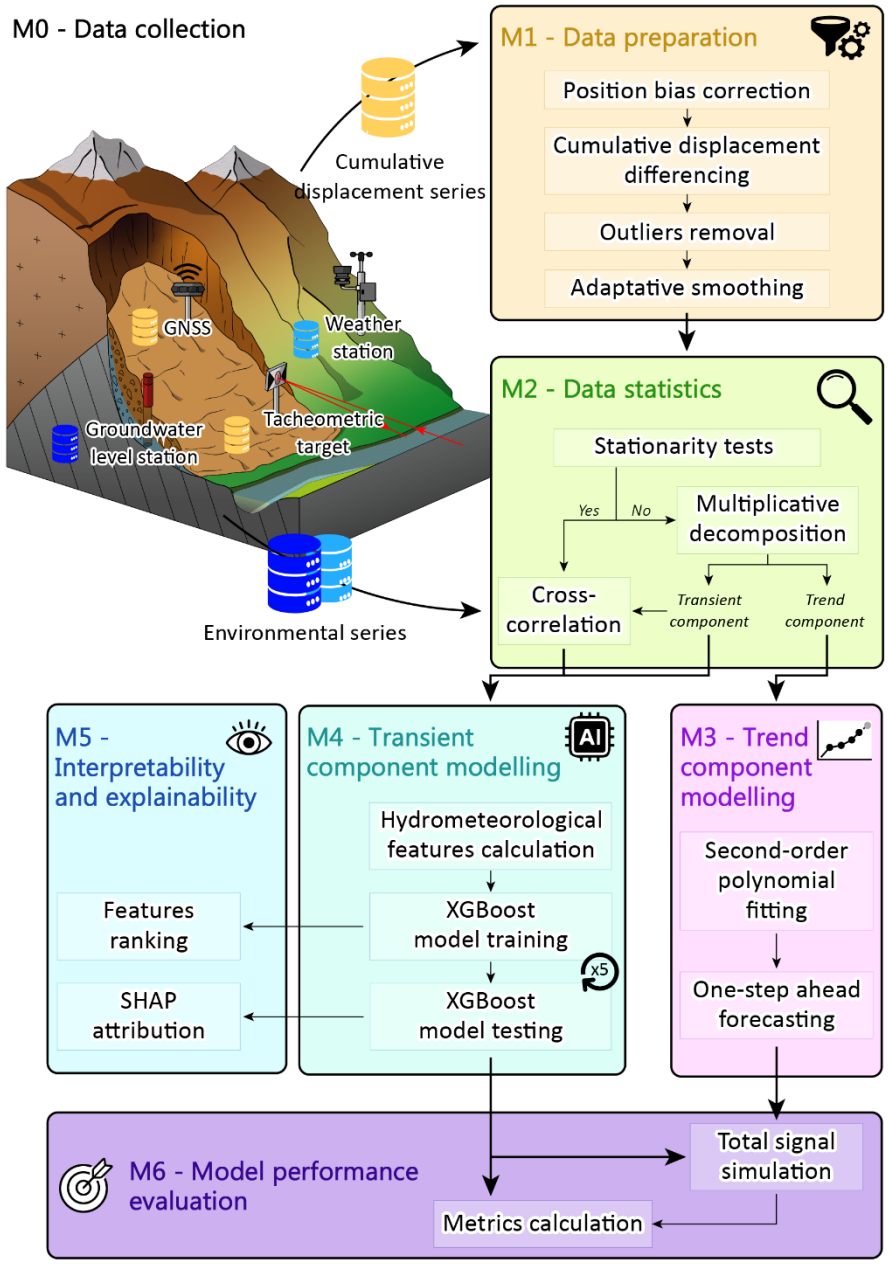


Figure A68 | Cadre de modélisation XAI. Le cadre de modélisation est organisé en six modules. M1 prétraite les séries temporelles de déplacements cumulés. M2 évalue la stationnarité des séries temporelles, isole la tendance de long terme et les composantes transitoires, et quantifie les corrélations avec les forçages environnementaux. M3 prévoit la composante de tendance. M4 réalise l'extraction des variables descriptives ainsi que l'entraînement et le test du modèle XGBoost. M5 quantifie l'importance des prédicteurs à l'aide d'un classement d'importance des variables et attribue les effets des prédicteurs au moyen d'une analyse SHAP. M6 évalue les performances prédictives du modèle sur des périodes de test indépendantes.

Le troisième module concerne la modélisation de la tendance de long terme, tandis que le quatrième repose sur l'extraction d'un ensemble générique de 248 prédicteurs construits à partir des pluies nettes, des pluies efficaces et des niveaux d'eau. Ces variables décrivent à la fois l'état hydrologique du système, les effets de mémoire, les réponses retardées et les perturbations transitoires, au moyen de

cumuls, maxima, moyennes, gradients, anomalies, rapports de saturation et variables décalées dans le temps. Les prédicteurs sont calculés sur plusieurs fenêtres glissantes comprises entre 1 et 90 jours, de manière à représenter les réponses à court, moyen et long termes. L'apprentissage est ensuite réalisé avec XGBoost selon une validation croisée bloquée sur cinq périodes temporelles contiguës, ce qui permet d'évaluer le modèle sur des intervalles indépendants tout en respectant la structure chronologique des séries. Enfin, l'interprétation du modèle combine les importances natives d'XGBoost et l'analyse SHAP, afin de quantifier à la fois l'importance globale des variables et leur contribution directionnelle aux vitesses simulées.

A.III.3. Résultats et conclusions

Les résultats montrent que le workflow reproduit correctement les vitesses journalières sur l'ensemble des sites étudiés, en restituant à la fois la dynamique générale de long terme, le timing et l'amplitude des épisodes transitoires d'accélération et de décélération.

Sur les points représentatifs A13 à Séchilienne, VIL-PL35 à Viella et VLRB à Villerville, les vitesses simulées suivent de manière satisfaisante les variations observées, avec des erreurs résiduelles globalement faibles au regard de l'amplitude des séries (**Figures A9a,d,g**). Les performances restent cohérentes à l'échelle de l'ensemble du réseau de points analysés, ce qui confirme la robustesse du cadre malgré la diversité des contextes cinématiques et hydrologiques. Les principaux écarts concernent surtout l'amplitude de certains pics, qui peuvent être sous-estimés ou, plus rarement, surestimés. Ces limites sont particulièrement visibles à Villerville (**Figure A9g**), où certaines accélérations brèves et marquées ne sont que partiellement reproduites, et à Viella lors de l'épisode extrême du début 2022, dont le déclenchement est capté mais dont l'amplitude reste sous-estimée (**Figure A9d**). Ces écarts sont relatifs à la représentativité inégale des phases d'accélération dans les données d'apprentissage, à des changements rapides de régime, et à la difficulté pour un ensemble générique de prédicteurs de capturer certains processus plus spécifiques ou plus rares.

Au-delà de la performance prédictive, le résultat central du chapitre réside dans l'interprétation des relations forçage-réponse. L'analyse conjointe des importances de variables et des valeurs SHAP met en évidence des hiérarchies de contrôle distinctes selon les sites. À Séchilienne (**Figures A9b-c**), la vitesse est principalement contrôlée par des variables de pluie efficace intégrées sur des durées intermédiaires, typiquement de l'ordre de 20 à 90 jours, ce qui traduit une sensibilité à une réponse hydrologique de moyen terme en milieu fracturé. Des variables décrivant des conditions durablement humides, telles que les rapports de saturation antécédente, jouent également un rôle important, ce qui confirme un contrôle dépendant de l'état hydrologique préalable. À Viella (**Figures A9e-f**), ce sont les descripteurs de niveau d'eau souterraine qui dominent, ce qui reflète un couplage hydromécanique fort et une dépendance marquée à l'évolution piézométrique.

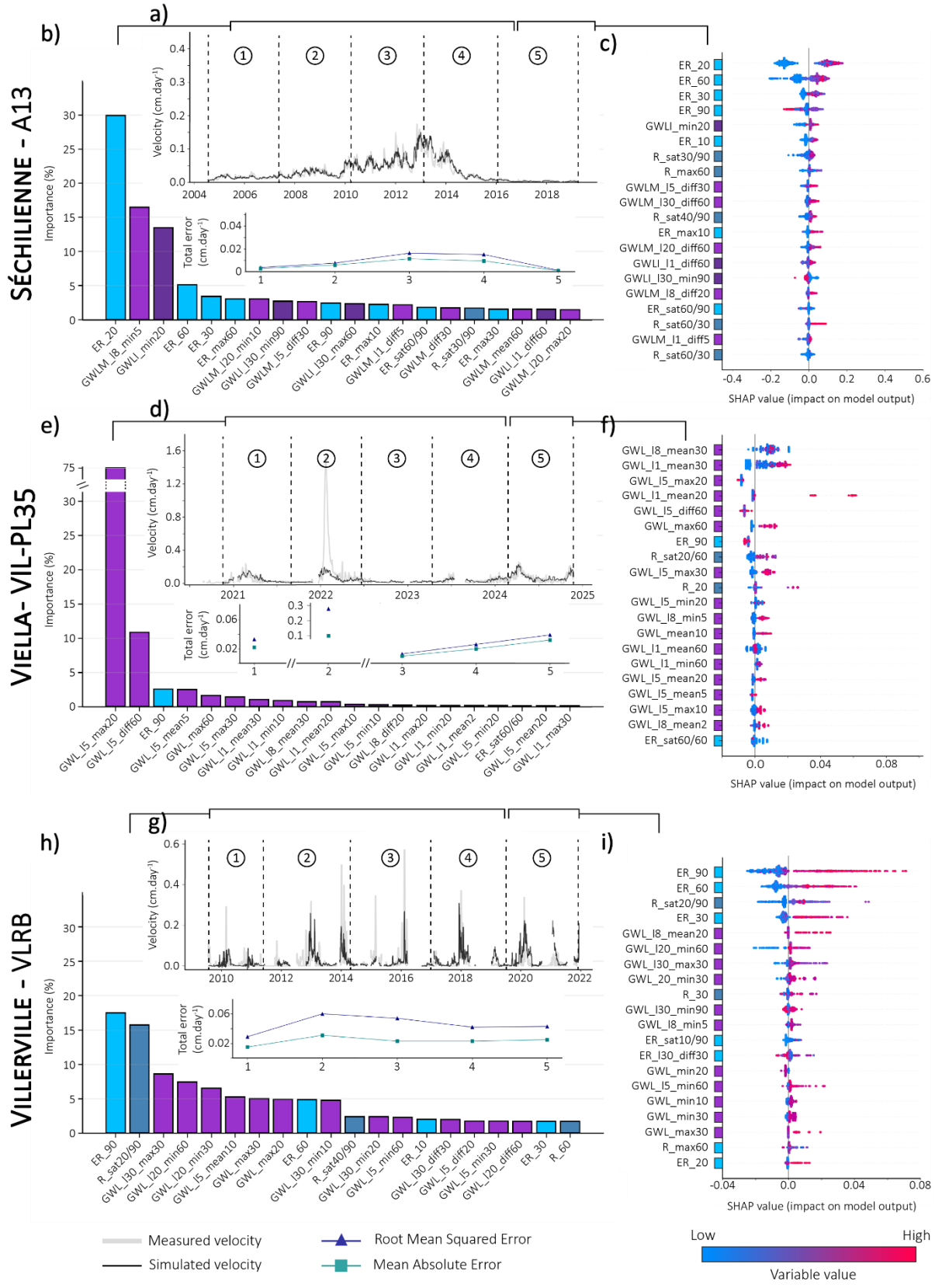


Figure A69 | Vitesses de glissement simulées, interprétabilité et explicabilité des prédicteurs (importance des variables, attribution SHAP). **(a-c)** Séchilienne (point A13), **(d-f)** Viella (point VIL-PL35) et **(g-i)** Villerville (point VLRB). Les vitesses journalières observées et simulées sont représentées respectivement en gris et en noir **(a, d, g)**, avec des performances prédictives sur les périodes de test

(périodes 1 à 5) quantifiées à l'aide des métriques RMSE (triangle) et MAE (carré). L'importance des variables (à gauche) et l'attribution SHAP (à droite) mettent en évidence les contrôles hydrométéorologiques dominants du comportement du modèle **(b, e, h)** et leur influence sur les vitesses prédites (c, f, i) au cours des périodes d'entraînement et de test, comme l'indiquent les barres horizontales en haut de chaque panneau. Par souci de lisibilité, seuls les vingt prédicteurs les plus influents sont représentés, ceux-ci expliquant la majeure partie de la variance expliquée. Dans les graphiques SHAP, chaque point correspond à une simulation individuelle de test, colorée selon la valeur du prédicteur correspondant. Les prédicteurs sont codés par couleur selon leur source : pluie efficace (bleu ciel), pluie nette (bleu acier) et niveau d'eau souterraine (magenta). À Séchilienne, deux séries de niveau d'eau souterraine sont représentées, correspondant à des réservoirs hydrologiques superficiel et profond *(Béjean-Maillard et al., 2025)*. Les vitesses simulées correspondent à la série totale simulée (tendance × transitoire). À Viella, le prédicteur principal dépasse 70 % du score total d'importance ; une rupture d'axe est donc utilisée afin de conserver une échelle comparable entre les graphiques.

À Villerville (**Figures A9h-i**), l'analyse fait ressortir l'importance de la mémoire hydrologique de plus long terme, ainsi qu'un contrôle significatif de la piézométrie sur certains points, cohérent avec une réponse plus inertielle dans un matériau moins perméable. La hierarchie des prédicteurs varient également au sein d'un même glissement, ce qui documente explicitement la variabilité interne des relations forçage-réponse.

En synthèse, ce chapitre montre qu'un workflow fondé sur XGBoost et sur un ensemble générique de prédicteurs physiquement fondés permet de reproduire des régimes cinématiques contrastés sans recourir à des variables spécifiques à chaque site ni à une représentation explicite de la structure géologique locale. Il montre également que la transparence du modèle peut être intégrée de façon systématique à la chaîne de traitement, de sorte que les relations extraites ne soient pas considérées comme de simples corrélations empiriques, mais comme des contraintes physiquement interprétables sur les contrôles hydro-météorologiques des déplacements. Dans le cadre de la thèse, le ***chapitre V*** établit ainsi la base méthodologique nécessaire pour passer d'une simulation déterministe explicable des vitesses à des applications scénarisées et probabilistes, développées au ***chapitre VI*****.**

A.IV. Chapitre VI - Extension du workflow fondé sur l'IA vers des simulations probabilistes des déplacements

A.IV.1. Introduction

Ce chapitre répond à la troisième question de la thèse et prolonge le cadre déterministe développé au ***chapitre V*** vers une approche probabiliste de la dynamique des glissements lents. Alors que le workflow XGBoost précédemment développé permet de reproduire les vitesses journalières à partir d'une contexte hydro-météorologiques historique et unique, il ne renseigne pas sur l'ensemble des déplacements susceptibles d'être observés sur une période donnée sous d'autres conditions hydro-météorologiques, pourtant plausibles. L'objectif du ***chapitre VI*** est donc de quantifier comment les déplacements de glissements lents peuvent évoluer sous des contextes météorologiques contrastés, en proposant un workflow permettant de propager explicitement la variabilité des forçages à travers la chaîne complète de simulation des vitesses.

La stratégie proposée repose sur l'idée que, dans les glissements lents contrôlés par les forçages hydro-météorologiques, une approche probabiliste peut fournir non plus une série unique de vitesse simulée, mais un intervalle de vitesses mensuelles attendues sur un cycle hydrologique complet de un an. Pour cela, le workflow intègre successivement la génération stochastique de chroniques météorologiques journalières, la simulation des fluctuations piézométriques cohérentes avec ces chroniques, puis l'application des modèles XGBoost pré-entraînés sur ces nouvelles réalisations. L'approche est appliquée à trois glissements présentant des régimes forçage-réponse contrastés : Viella, Villerville et Vallcebre. Elle vise ainsi à produire des distributions probabilistes mensuelles de vitesses cumulées, directement interprétables dans une perspective de suivi opérationnel et d'aide à la gestion de l'aléa.

A.IV.2. Matériels et méthodes

Le workflow dévelopé repose sur trois étapes successives (**Figure A10**). La première consiste à générer un ensemble de scénarios météorologiques journaliers plausibles à partir des séries historiques observées. Pour cela, un générateur météorologique stochastique de type Richardson modifié, calibré sur les chroniques de pluie et de température de chaque site, est utilisé. Contrairement à la formulation initiale fondée sur une chaîne de Markov à deux états, la version retenue repose sur une chaîne à six états, plus adaptée à la préservation de la cohérence entre occurrences pluvieuses et intensités (**Figure A11a**).

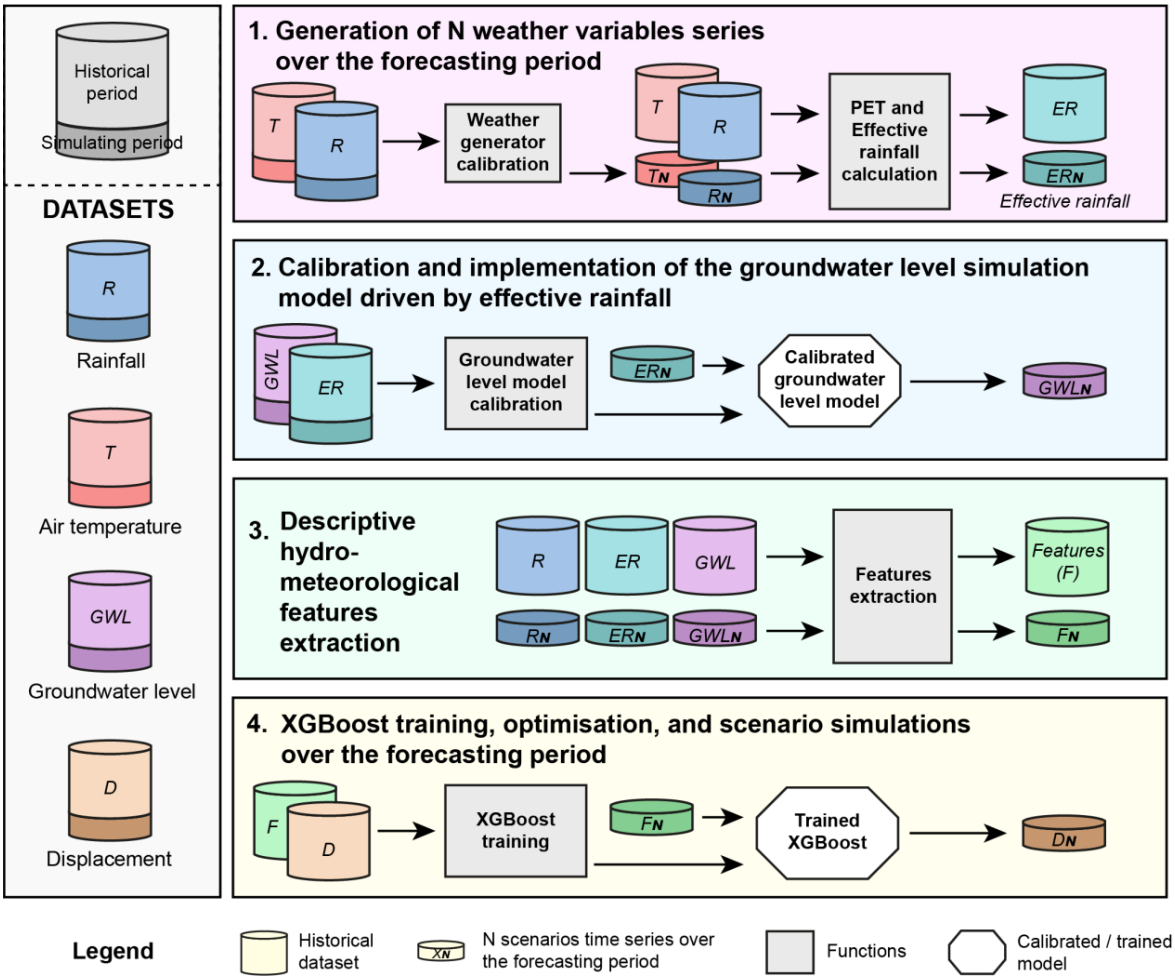


Figure A70 | Chapitre VI – Vue d'ensemble schématique du cadre de simulation probabiliste développé.

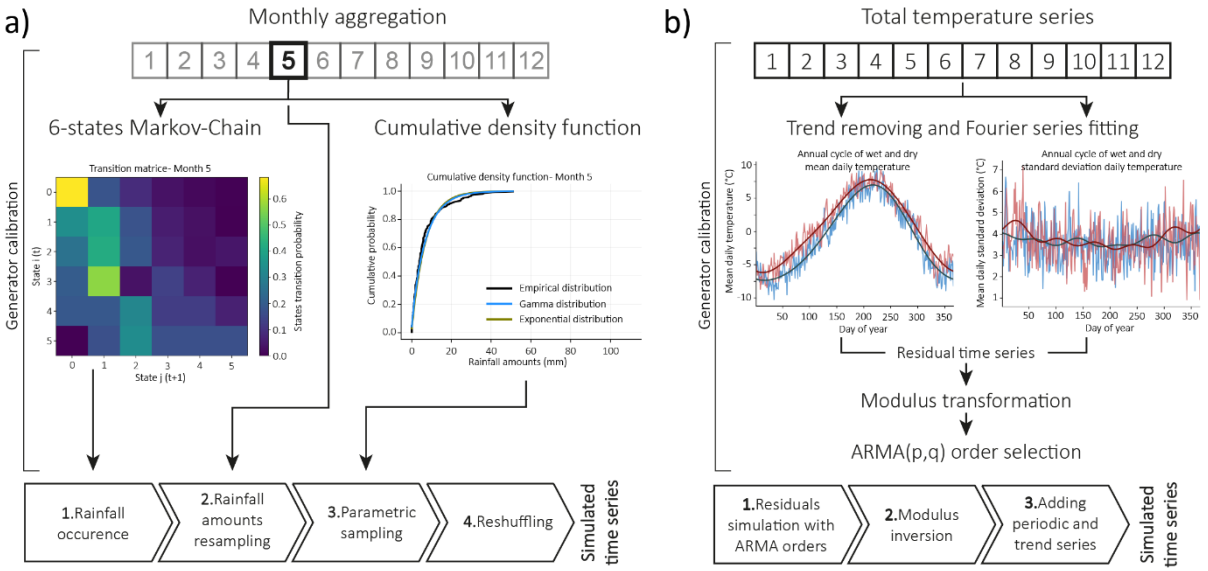


Figure A71 | Workflow de calibration et de simulation des générateurs de séries temporelles de **(a)** pluie et **(b)** température (adapté de Breinl et al., 2015, 2016).

Les occurrences pluie/non-pluie sont simulées, puis les quantités de pluie sur les jours humides sont échantillonnées à partir de distributions paramétriques testées à l'échelle mensuelle, tandis que les températures journalières sont générées conditionnellement à la séquence humide/sèche (**Figure A11b**). Cette étape permet de produire, pour chaque site, un ensemble de 500 scénarios journaliers de pluie et de température sur 15 mois, soit un cycle hydrologique complet précédé d'une période d'initialisation de 3 mois pour générer de la variabilité initiale dans le calcul des prédicteurs.

La deuxième étape vise à simuler les fluctuations du niveau d'eau souterraine associées à chacun de ces scénarios météorologiques. Le chapitre adopte pour cela un modèle hydrologique global à fonctions de transfert implémenté dans la librairie python open-access PASTAS. Les niveaux d'eau sont simulés à partir de la pluie efficace, au moyen de deux fonctions de réponse complémentaires représentant des dynamiques rapides, avec une fonction exponentielle (**Equation A.1**), et lentes, avec une fonction gamme inverse (**Equation A.2**).

$$\theta_{exp}(t) = \frac{A}{a} e^{-t/a}, \qquad t \geq 0 \tag{A.1}$$

$$\theta_{\gamma}(t) = A \frac{t^{n-1}}{a^n \Gamma(n)} e^{-t/a}, \qquad t \geq 0 \tag{A.2}$$

Ce choix méthodologique permet de conserver un modèle relativement parcimonieux, calibrable à partir des seules chroniques de forçage et de niveau d'eau, tout en reproduisant les principales signatures saisonnières et événementielles observées dans les séries piézométriques. La troisième étape consiste à transformer les séries générées de pluie, de pluie efficace, de température et de niveau d'eau en prédicteurs descriptifs de même nature que ceux développés dans le chapitre V, puis à injecter ces prédicteurs dans des modèles XGBoost pré-entraînés pour simuler les vitesses journalières. Les vitesses obtenues sont ensuite agrégées à l'échelle mensuelle afin de produire, pour chaque mois et pour chaque site, une distribution probabiliste des vitesses cumulées. Les distributions obtenus sont évaluées selon leur intégration des valeurs historiques mesurées à l'aide des métriques PICP et MIS.

A.IV.3. Résultats et conclusions

Les résultats montrent d'abord que les séries météorologiques générées reproduisent correctement les principales propriétés statistiques des chroniques observées. Les distributions de pluie, le nombre de jours humides et la variabilité saisonnière de la température sont globalement bien préservés, ce qui indique que les scénarios générés constituent une base cohérente pour propager la variabilité hydro-météorologique à travers la chaîne de simulation. De la même manière, les niveaux

d'eau simulés à partir des fonctions de transfert restituent les grandes tendances saisonnières et les principales fluctuations événementielles. Toutefois, cette reproduction est impactée par par une sous-éstimation des amplitudes de certains pics, ainsi que par des phases de récession simulées trop rapides, ce qui affecte ensuite la qualité des prédicteurs dérivés utilisés pour simuler les vitesses (**Figure A12**).

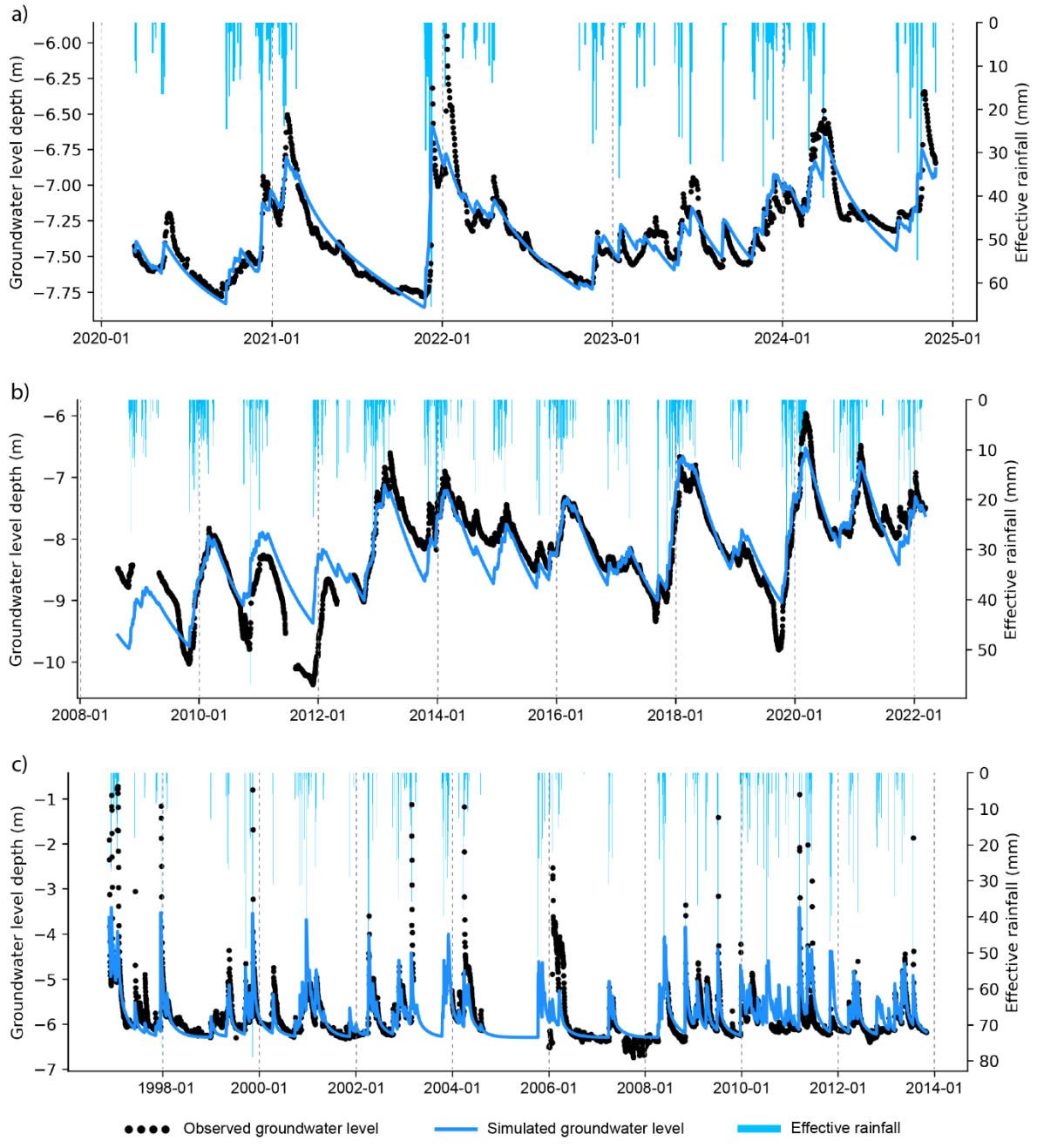


Figure A72 | Séries de niveaux d'eau souterraine mesurées (noir) et simulées (bleu) pour **(a)** Viella, **(b)** Villerville et **(c)** Vallcebre.

À partir de ces séries générées, les modèles XGBoost produisent des ensembles de vitesses journalières qui, une fois agrégés à l'échelle mensuelle, définissent des distributions de vitesses cumulées propres à chaque site. Ces distributions mettent en évidence des signatures saisonnières contrastées, cohérentes avec les régimes de fonctionnement identifiés dans les chapitres précédents. Selon les sites, la saisonnalité est plus ou moins marquée, et les gammes de vitesses mensuelles varient fortement d'un mois à l'autre. Comparées aux vitesses mensuelles historiques via les métriques PICP et MIS, les distributions simulées englobent globalement la majorité des observations, ce qui montre que le cadre probabiliste capte correctement les grands contrastes saisonniers et l'ordre de grandeur de la variabilité attendue (**Figure A13**).

Quelques écarts d'intégration des valeurs historiques subsistent et sont attribuées à deux origines. La première concerne la qualité des chroniques piézométriques simulées, dont certaines caractéristiques locales – notamment l'amplitude des pics et les récessions – restent imparfaitement reproduites par les modèles à fonctions de transfert simple. La seconde concerne la représentativité des séries de vitesses ayant servi à l'apprentissage des modèles XGBoost. Lorsque certaines combinaisons de forçage et de réponse sont peu représentées dans les données d'entraînement, la robustesse des distributions mensuelles produites s'en trouve réduite. Le chapitre montre également que l'optimisation des hyperparamètres XGBoost n'apporte pas ici d'amélioration substantielle des distributions mensuelles, ce qui confirme que les principales limites du cadre se situent davantage dans la représentation du signal hydrologique et dans la représentativité des données d'apprentissage que dans le réglage de l'algorithme de prévision lui-même.

En synthèse, ce chapitre fournit une première base méthodologique pour traduire des simulations déterministes de vitesses en plages probabilistes mensuelles sous variabilité hydro-météorologique plausible. Il montre qu'en combinant un générateur météorologique stochastique, un modèle piézométrique global et des modèles XGBoost pré-entraînés, il est possible de produire des distributions mensuelles de vitesses cumulées cohérentes avec les principales signatures saisonnières observées. Dans le cadre général de la thèse, le ***chapitre VI*** étend ainsi les contributions du ***chapitre V*** vers des applications orientées aléa, en proposant une approche structurée pour explorer la sensibilité des glissements lents à des scénarios contrastés de forçage. Il ouvre également des perspectives concrètes pour le suivi opérationnel, notamment via la production de gammes probabilistes de déplacement susceptibles d'alimenter des outils d'aide à la décision.

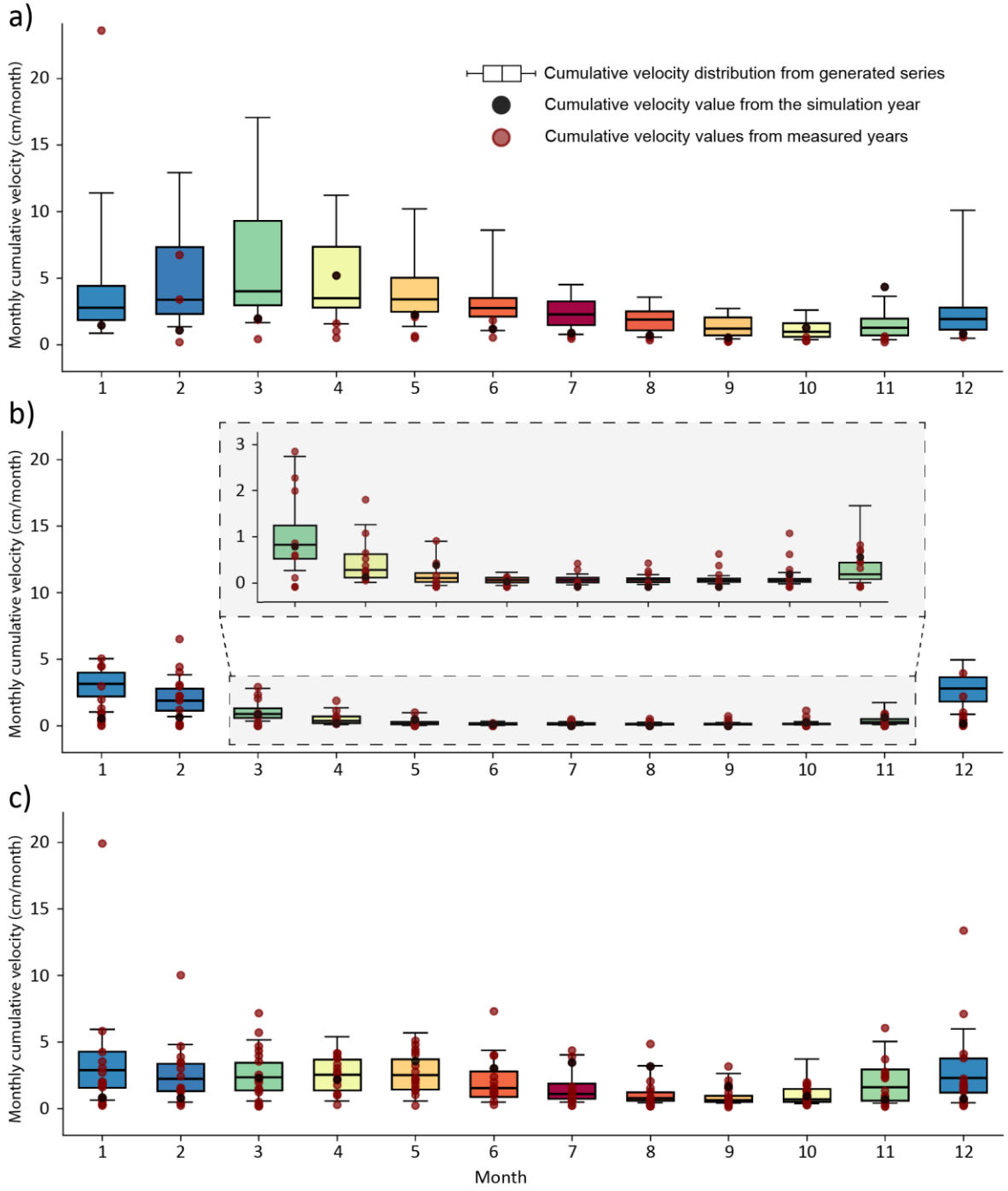


Figure A73 | Distribution des vitesses cumulées mensuelles pour les glissements de terrain de **(a)** Viella, **(b)** Villerville et **(c)** Vallcebre. Les vitesses cumulées mensuelles observées sont représentées par des points rouges, et celles correspondant à la période cible par des points noirs. Les moustaches représentent l'intervalle compris entre les 5e et 95e percentiles, en cohérence avec le Mean Interval Score (MIS).

A.V. Conclusions et perspectives

Cette thèse a proposé un cadre méthodologique visant à mieux caractériser, comprendre et simuler la dynamique de glissements de terrain lents et profonds en réponse aux forçages hydro-météorologiques. Trois contributions principales en ressortent. La première concerne la reconstruction de proxys cohérents de l'état hydrologique interne à partir de modèles conceptuels pluie-débit, afin de pallier l'absence ou l'incomplétude des chroniques piézométriques. La deuxième repose sur le développement d'un workflow générique fondé sur XGBoost et sur des prédicteurs hydro-météorologiques physiquement motivés, capable de reproduire les vitesses journalières sur plusieurs glissements contrastés tout en explicitant les relations forçage-réponse. La troisième réside dans l'extension de ce cadre vers des simulations probabilistes, par propagation de scénarios hydro-météorologiques plausibles jusqu'aux distributions mensuelles de vitesse. L'ensemble montre qu'une articulation entre modélisation hydrologique conceptuelle, apprentissage automatique explicable et simulation probabiliste permet de relier des séries de suivi multi-sources à des produits pertinents pour l'analyse de l'aléa.

Plusieurs perspectives prolongent directement ce travail. Une première consiste à étendre la méthodologie à un plus grand nombre de glissements lents, y compris dans des contextes moins bien instrumentés, notamment à l'aide de stratégies de transfert entre sites ou entre familles de glissements. Une deuxième concerne l'intégration de produits satellitaires pour compléter les données de forçage, de déformation et de géométrie, et ainsi élargir le champ d'application du cadre proposé à des zones où les observations in situ restent limitées. Une troisième perspective porte sur l'évolution du cadre actuel vers une prévision explicitement spatio-temporelle des vitesses, afin de mieux représenter l'organisation interne des déformations à l'échelle du versant. Une quatrième concerne les progrès méthodologiques en interprétabilité et explicabilité des modèles d'IA appliqués aux séries temporelles, en particulier dans le cas de modèles plus complexes (ex : modèle Transformers). Enfin, fondée sur l'application de modèles basés sur les données, la poursuite d'une démarche d'open access sur les données multi-sources, les chaînes de traitement et les outils développés apparaît comme un levier important pour renforcer la reproductibilité, les comparaisons inter-sites et les développements futurs de ces approches.

ANNEXES

SCIENCE OUTREACH

Une classe, un chercheur *25 May 2023*	Presentation of a mediation workshop on my PhD topic to school groups at a school in Besançon (25)
Festival Horizone *8 Juillet 2023*	Presentation of a mediation workshop for the general public on my PhD topic in Colmier-le-Bas (52).
Nuit européenne des chercheur.e.s *29 Septembre 2023* *27 Septembre 2024*	Presentation of my research and discussion with the general public through artworks exhibited at the museums of Besançon (25), as part of the "One artwork, one researcher" programme.

TEACHING

UFR Sciences & Techniques Université Marie et Louis Pasteur *2023-2024*	General Geology tutorial (1st-year Life and Earth Sciences): 42 hours

SUPERVISED INTERNSHIPS

Late Lawson	One-month internship, Master's degree in Mathematics and Statistics (UFR ST) *10 February - 10 March 2023*
Felix Queré	5-week internship, 2nd year of engineering school (ENSG Nancy) *5 February - 7 March 2024*
Alexandre Pédémonte	One-month internship, Bachelor's degree in Geology (UFR ST) *02 June - 07 July 2025*

LIST OF MANUSCRIPTS

Manuscrit 1 *Published*	**Béjean-Maillard, O.**, Bertrand, C., Malet, J.-P., Cinkus, G., Nevers, P., Ducasse, J., Bertolo, D., & Thuegaz, P. (2025). Hydrogeological forecasting of deep-seated landslides dynamics: Structure and sensitivity of tank models. Landslides, 22(7), 2199–2217. https://doi.org/10.1007/s10346-025-02482-2
Manuscrit 2 *Submitted*	**Béjean-Maillard O.**, Bertrand C., Malet J.-P., Dubois L., Batailles C., Lespines L., Ducasse J. A Scalable XAI-based Forecasting of Landslide Velocities From Environmental Forcings
Manuscrit 3 *In preparation*	**Béjean-Maillard O.**, Bertrand C., Malet J.-P., Moya J. Probabilistic Simulation of Monthly Landslide Velocity Under Hydro-meteorological Variability

LIST OF COMMUNICATIONS

6th World Landslide Forum *Poster* *13-17 November 2023*	**Béjean-Maillard O.**, Bertrand C., Malet J.-P. Forecasting landslide motion with machine learning models: the use case of Séchilienne landslide (French Alps) to identify the relevant predicting variables
European Geosciences Union *Oral communication* *15-19 April 2024*	**Béjean-Maillard, O.**, Bertrand, C., Malet, J.-P., Cinkus, G., Nevers, P., Ducasse, J., Bertolo, D., & Thuegaz, P. Forecasting landslide motion with EXplainable Machine Learning models: the use case of Séchilienne landslide (French Alps) to identify the relevant predicting variables
SNO Karst Workshop (FR) *Oral communication* *21-23 May 2024*	**Béjean-Maillard, O.**, Bertrand, C., Malet, J.-P., Cinkus, G., Nevers, P., Ducasse, J., Bertolo, D., & Thuegaz, P. Application des modèles en reservoirs dans la connaissance de transferts de masses d'eau en milieux fissurés instables : Exemple des glissements de terrain de Séchilienne et du Mont de La Saxe
14th International Symposium of Landslide *Oral communication* *08-12 July 2024*	**Béjean-Maillard, O.**, Bertrand, C., Malet, J.-P., Cinkus, G., Nevers, P., Ducasse, J., Bertolo, D., & Thuegaz, P. Forecasting landslide motion machine learning models: the use case of Séchilienne landslide (French Alps) to identify relevant predicting variables
Data Terra Workshop (FR) *Oral communication* *26-28 May 2025*	**Béjean-Maillard, O.**, Bertrand, C., Malet, J.-P. Comprendre et prévoir la dynamique des glissements de terrain en réponse aux forçages hydrométéorologiques à l'aide de l'IA
Journée des Aléas Gravitaires (FR) *Poster & Oral communication* *22-24 September 2025*	**Béjean-Maillard, O.**, Bertrand, C., Malet, J.-P. A generic and explainable AI-based workflow to reproduce landslide dynamics and provide insights into the contribution of hydrometeorological predictive variables
Rencontres Epos-France (FR) *Oral* *18-20 November 2025*	**Béjean-Maillard, O.**, Bertrand, C., Malet, J.-P. Intelligence Artificielle et Glissements de Terrain

Titre : Simulation des glissements de terrain lents à l'aide de l'IA : données multi-sources, forçages hydro-météorologiques, et interprétabilité

Mots clés : Glissements de terrain lents, intelligence artificielle explicable, simulation de séries temporelles, forçages hydro-météorologiques, données multi-sources, modèles hydrologiques

Résumé : Les glissements de terrain lents représentent un danger significatif pour les populations et les infrastructures à l'échelle mondiale. Pourtant, la prévision de leur vitesse reste un défi majeur en raison de l'hétérogénéité spatio-temporelle de leurs propriétés internes. Leur dynamique est néanmoins largement reconnue comme étant fortement contrôlée par les forçages hydro-météorologiques à court et à long terme. Dans ce contexte, les approches basées sur les données, en particulier l'Intelligence Artificielle (IA), offrent une alternative efficace aux modèles physiques pour capturer des relations complexes entre variables hydrométéorologiques et vitesse de surface.

Cette thèse vise à mieux caractériser et anticiper la dynamique des glissements de terrain lents en combinant de longues séries temporelles multi-sources avec des modèles d'IA interprétables. L'objectif est de proposer un cadre plus générique pour analyser les liens entre forçages hydro-météorologiques et vitesse de surface, tout en identifiant les mécanismes dominants à différentes échelles de temps.

Ce travail s'articule autour de trois axes complémentaires. Le premier axe répond au manque de mesures hydrogéologiques in situ en mobilisant une modélisation hydrologique de type réservoirs (KarstMod), calibrée sur des glissements profonds alpins, afin de reproduire les débits aux exutoires et de fournir des séries synthétiques représentatives du stockage interne.

Le second axe consiste à développer une méthodologie évolutive d'IA explicable (XAI), du prétraitement des données à la reproduction des déplacements, fondée sur XGBoost et l'interprétation SHAP, pour simuler les vitesses et quantifier l'influence des variables hydro-météorologiques à partir de chroniques multivariées acquises sur plusieurs glissements instrumentés.

Enfin, le troisième axe évalue l'effet de conditions hydro-météorologiques contrastées en couplant des générateurs de scénarios (pluie et piézométrie) aux modèles entraînés, afin de caractériser la variabilité interannuelle et l'évolution des probabilités des vitesses mensuelles au cours d'un cycle hydrologique.

L'ensemble de ces travaux améliore la transparence des relations entre forçages hydro-météorologiques et déplacement, en proposant un cadre méthodologique générique et des outils opérationnels pour l'analyse des glissements de terrain lents. À terme, ces approches pourront évoluer au rythme des progrès en acquisition de données, en performance et en explicabilité des modèles d'IA, afin d'intégrer davantage la dimension spatio-temporelle et de renforcer leur applicabilité dans des régions faiblement instrumentées, voire dépourvues de surveillance.

Title : Simulating slow-moving landslides with AI: Multi-source data, hydro-meteorological forcings, and interpretability

Keywords : Slow-moving landslides, explainable artificial intelligence, time series simulation, hydro-meteorological forcings, multi-source data, hydrological models

Abstract : Slow-moving landslides represent a significant hazard for populations and infrastructure worldwide. Yet, forecasting their velocity remains a major challenge because of the strong spatial and temporal heterogeneity of their internal properties. Their dynamics are nevertheless widely recognised as being strongly controlled by short- and long-term hydro-meteorological forcing. In this context, data-driven approaches, and in particular Artificial Intelligence (AI), provide an efficient alternative to physics-based models for capturing complex relationships between hydro-meteorological variables and landslide surface velocities.

This thesis aims to better characterise and anticipate the dynamics of slow-moving landslides by combining long, multi-source time series with interpretable AI models. The objective is to propose a more generic framework for analysing the links between hydrometeorological forcing and surface velocity, while identifying the dominant mechanisms at different time scales.

The work is structured around three complementary axes. The first axis addresses the lack of in situ hydrogeological measurements by using conceptual tank-type hydrological modelling (KarstMod), calibrated on deep Alpine landslides, to reproduce outlet discharge and to provide synthetic time series representative of internal storage.

The second axis consists in developing a scalable eXplainable AI (XAI) framework, from data preprocessing to velocity simulation, based on XGBoost and SHAP interpretation. This framework is used to simulate velocity and to quantify the influence of hydro-meteorological variables from multivariate records acquired on several instrumented landslides.

Finally, the third axis evaluates the effect of contrasting hydro-meteorological conditions by coupling rainfall and piezometric scenario generators with the trained models, in order to characterise the interannual variability and the evolution of the probabilities of monthly velocities over a hydrological year.

These contributions improve the transparency of hydro-meteorological controls on displacement by providing a scalable methodological framework and operational tools for analysing slow-moving landslides. Ultimately, these approaches may evolve in line with advances in data acquisition, performance and explainability of AI models, in order to further integrate the spatio-temporal dimension and enhance their applicability in regions with limited, or absent, monitoring.

UNIVERSITÉ MARIE & LOUIS PASTEUR
CHRONO ENVIRONNEMENT
ÉCOLE DOCTORALE Environnements - Santé Bourgogne | Franche-Comté