UNIVERSITE DE NICE-SOPHIA ANTIPOLIS - UFR SCIENCES

Ecole Doctorales des Sciences Fondamentales et Appliquées

THESE

pour obtenir le titre de

Docteur en Sciences

de l'Université de Nice-Sophia Antipolis

Spécialité : Sciences de l'Univers

présentée par

TRAN Dinh Trong

ANALYSE RAPIDE ET ROBUSTE DES SOLUTIONS GPS POUR LA TECTONIQUE

Thèse dirigée par **Jean-Mathieu NOCQUET**

Soutenue le 06 juin 2013 devant le jury composé de:

M. Pierre EXERTIER	Président du Jury
Mme. Andrea WALPERSDORF	Rapporteur
M. Frédéric MASSON	Rapporteur
M. Eric CALAIS	Examinateur
M. Jean-Mathieu NOCQUET	Directeur de thèse

Remerciements

Je tiens tout d'abord à remercier profondément mon directeur de thèse, Monsieur Jean-Mathieu NOCQUET, pour tous ses aides précieuses et son accueil dès mon arrivée, pour m'avoir conseillé, encouragé et soutenu tout au long de la thèse avec patience et disponibilité. Sous sa direction, j'ai appris de lui non seulement des connaissances scientifiques mais aussi l'expérience de recherche et la méthode de recherche. Grâce à sa gentillesse, et à sa compréhension pour ma connaissance limitée de la langue française, j'ai surmonté toutes les difficultés dans mon travail et dans ma vie pour réaliser cette thèse. Je le remercie aussi pour ses corrections et ses suggestions qui ont amélioré ce manuscrit.

Je remercie Phillippe Charvis pour son implication lors de ma première année de thèse.

Je tiens à remercier Monsieur Frédéric MASSON et Madame Andrea WALPERSDORF d'avoir accepté d'être les rapporteurs de la thèse. Je remercie aussi les membres de jury pour leur temps et l'intérêt qu'ils ont accordés à mon travail.

J'adresse mes remerciements à Monsieur Emmanuel TRIC, directeur du laboratoire Géoazur, pour m'avoir accueilli au sein du laboratoire, ainsi qu'à l'ensemble de l'équipe administrative du laboratoire, de l'école doctorale SFA pour leur aide précieuse.

Mes sincères remerciements s'adressent à Monsieur Pierre EXERTIER, responsable de l'équipe Astrogéo, pour m'avoir accueilli au sein de l'équipe et pour la solution financière en fin de ma thèse, et à l'ensemble de l'équipe Astrogéo pour leur amitié.

Je remercie les membres du club GPS, Emilie KLEIN pour plusieurs tests sur mon code dans le cadre de son stage, Lucie ROLLAND pour ses suggestions et ses corrections françaises de ma rédaction, Mathilde VERGNOLLE pour ses conseils sur l'algorithme de notre approche, LIN Kuan Chuan et Juan Carlos VILLEGAS LANZA pour leurs nombreuses discussions, suggestions et aides dans la vie ainsi que dans le travail.

Je vous remercie aussi mes amis Clara CASTELLANOS LOPEZ, Guanghui HU pour votre bonne humeur et votre gentillesse, les musiciens Alain DUJARDIN et Clément PERRIN pour les musiques que nous nous avons partagé, surtout les musiques du band Salerderien de Clément PERRIN, grace à cela, j'ai eu des moments agréables pour recharger l'énergie pour le travail.

Je tiens à remercier l'équipe de football géoazuréen : Quentin BLETERY, Clément PERRIN, Alain DUJARDIN, LUONG Ngoc Dung, Nestor CERPA GILVONIO, Thomas Dylan MIKESELL, Aurélien BIGOT, Sadrac SAINT-FLEUR,... pour les matchs très amicaux et inoubliables.

Je tiens à remercier Maurin VIDAL et Martin VALLEE pour m'avoir cherché à l'aéroport

de Nice pour ma première arrivée en France ainsi que ma première rencontre avec les géoazuréens que j'étais très impressionné par leur gentillesse.

Et je remercie les personnes du laboratoire que je ne peux pas toutes nommer ici, pour leur amitié et leurs aides diverses.

Je ne peux pas oublier mes amis vietnamiens en France : Dung, Huyen, Huu-Phuong, Thinh, Dan, Thanh-Phuong, Huy, Khoa, Nam, The, Nhan,... pour cette belle fraternité et le partage de moments heureux comme ceux de tristesse tout au long du séjour en France.

Toutes mes pensées vont à mes parents, ma sœur, mon frère qui sont toujours à mes côtés et m'ont soutenu pendant mes études et particulièrement dans les moments difficiles de ma vie. Un grand merci à mon épouse - mon grand amour - et notre petite fille qui m'ont encouragé à démarrer ce travail de thèse. Elles ont toujours été là pour me motiver et me réconforter du début jusqu'à la fin.

Cette thèse a été financée par le gouvernement vietnamien dans le cadre du projet 322. Je tiens sincèrement à remercier le gouvernement vietnamien, le Ministère de l'Éducation et de la Formation vietnamienne et tous les membres du projet 322. Je remercie également la Faculté de Géodésie et Cartographie - Département des Points et Chaussées - Ecole nationale supérieur de Génie civil de Hanoi, Vietnam d'avoir créé les conditions favorables à ma thèse.

Résumé

Analyse rapide et robuste des solutions GPS pour la tectonique

Le Global Positioning System (GPS) a de nombreuses applications scientifiques. En géophysique de précision, il est utilisé pour déterminer les mouvements des plaques tectoniques, quantifier la déformation aux frontières des plaques ou dans le domaine intraplaque, ou bien détecter les signaux transitoires associés au cycle sismique. Aujourd'hui, la surface du globe est recouverte de milliers de stations GPS permanentes permettant de générer les séries temporelles de coordonnées de stations GPS et de suivre en continu les mouvements de l'écorce terrestre.

Mon travail de thèse se situe dans le contexte du développement de grand réseaux GPS et GNSS (Global Navigation Satellite System) permanents. Par exemple, le réseau GEONET (GNSS Earth Observation Network System) du Japon comprend plus de 1000 stations, le réseau du Plate Boundary Observatory comprend lui aussi environ 1200 stations dans l'ouest des Etats Unis. A une échelle plus locales, les réseaux permanents en Europe possèdent environ 300 stations, et celui de Taiwan, plus de 400 stations.

Des nombreuses difficultés se posent en pratique pour réaliser des séries temporelles précises et analyser les solutions de ces grands réseaux GPS permanents. Une première difficulté réside dans l'expression des solutions journalières dans un repère précis et stable dans le temps. Le grand nombre de points et la longueur des séries temporelles maintenant disponibles rendent les calculs lourds en temps. Des erreurs dans les solutions journalières ou dans la solution de référence peuvent biaiser l'estimation des paramètres de la transformation et dégrader la précision des séries temporelles obtenues. A l'étape de l'analyse des séries temporelles, on rencontre fréquemment plusieurs problèmes causés soit par des causes artificielles ou des mouvements géophysiques parfois complexes. La détection de ces problèmes et leur résolution dans les séries temporelles GPS de plus d'une décennie d'observation par une approche manuelle n'est plus possible et des algorithmes d'analyse automatique doivent être développés.

L'objet de ma thèse est de déterminer des approches, des méthodes et des algorithmes robustes permettant (1) la réalisation rapide et précises de séries temporelles de position (2) l'identification rapide des problèmes présents dans les séries temporelles GPS (3) la résolution automatique des problèmes les plus courants (4) la manipulation facile des séries temporelles pour extraire les paramètres utiles aux analyses géophysiques.

Dans ce travail, je présente tout d'abord une approche basée sur la norme L1 pour estimer les paramètres de transformations des solutions libres vers une solution de référence. Ensuite, je présenterai différents algorithmes de recherches automatiques d'erreurs et de

détection, estimation, corrections des sauts. Enfin, je montrerai comment ces algorithmes peuvent être utilisés dans un modèle général des séries temporelles pour obtenir une analyse automatique et par exemple, extraire les paramètres des déformations co- et post-sismiques.

Les essais méthodologiques ont été en premier lieu testés sur le réseau national GPS permanent français RENAG et une solution combinée des réseaux GPS de Taiwan. Ces deux applications permettent d'évaluer la capacité des méthodes développées à obtenir des vitesses précises et modéliser des mouvements complexes liées au cycle sismique.

Mots clefs : GPS, série temporelle GPS, méthodologie géodésique.

Abtract

Rapid and robust analysis of GPS solutions for tectonics

The development of the Global Positioning System (GPS) allows numerous applications in Science. In high precision geophysics, GPS is used to determine the motion of tectonic plates, to quantify the plate-boundary and intra-plate deformations, and detect the time variable deformation associated with the seismic cycle. Today, the Earth's surface is covered with thousands of permanent GPS stations allowing to generate the position time series of GPS stations and continuously monitor the earth's crust displacement.

My thesis takes place in the context of the development of large permanent GPS networks. For example, the Japan GEONET (GNSS Earth Observation Network System) network includes over 1000 stations, the Plate Boundary Observatory (PBO) network that includes about 1200 stations in the western United States. At a smaller scale, the GPS permanent network in Europe comprises about 300 stations, and in Taiwan more than 400 stations.

The analysis of permanent GPS networks faces many difficulties mainly that falls into two categories: at the step of transformation of the free daily solutions into the reference solution (reference frame definition), the large number of points and length of the time series makes the calculation time consuming. It is furthermore subject to errors in either free or the reference solution, possibly biasing the transformation parameters estimates and ultimately decreasing the accuracy of final time series. At the step of GPS time series analysis, either artificial or geophysical signals can cause non-linear motion, offsets causing bias in the estimated velocities. The detection of these problems and their resolution in the GPS time series with decades of observations with a manual approach is difficult, if not impossible.

The subject of my thesis is to define rapid and robust approaches, methods and algorithms for (1) the rapid determination of time series, (2) the identification of problems in the derived time series (3) the automatic resolution of most common problems, (4) the easy manipulation of time series enabling the user extract the parameters useful for the geophysical analysis, for instance the co- and post-seismic displacements.

The methodologies developed are tested using the solutions for the RENAG network (French National GPS Network) and the GPS Taiwan network. The first analysis aims at assessing the capability to derive high precision velocity, while the Taiwan network defines a good test case to extract and modelize signals associated with the earthquake cycle.

Key words: GPS, GPS time series, geodesy methodology.

TABLE DES MATIÈRES

Introduction 1

1 Transformation de Helmert entre deux repères de référence terrestre 5

1.1 Introduction . . . 5
1.2 Le repère de référence terrestre ITRF . . . 6
1.3 La solution de référence IGS . . . 8
1.4 Transformation de Helmert entre repères géodésiques . . . 10
1.4.1 Transformation de Helmert . . . 10
1.4.2 Transformation de Helmert en coordonnées locales . . . 11
1.5 Estimation optimale des paramètres de transformation de Helmert . . . 13
1.5.1 Difficultés pratiques . . . 13
1.5.2 Estimation par moindres carrés . . . 14
1.5.3 Estimation robuste de Dikin . . . 15
1.5.3.1 Moindres valeurs absolues et programmation linéaire . . . 17
1.5.3.2 La méthode affine de Dikin . . . 19
1.6 Validation de l'estimateur robuste Dikin pour les transformations de Helmert 21
1.6.1 Objet des tests . . . 21
1.6.2 Fabrication du jeu synthétique . . . 21
1.6.3 Résultats . . . 22
1.6.3.1 Etude de l'influence d'une erreur accidentelle . . . 22
1.6.3.2 Etude de l'effet de réseau local . . . 25
1.6.3.3 Etude de l'effet d'asymétrie du réseau . . . 27
1.6.3.4 Etude de l'influence d'une erreur/plusieurs erreurs accidentelles sur les séries temporelles . . . 29
1.6.3.5 Etude des performances de calcul des différents estimateurs 37
1.6.3.6 Validation de la stratégie finale . . . 39
1.6.4 Conclusions . . . 40
1.7 Implémentation: le programme pyacs_make_time_series.py . . . 41
1.7.1 Le framework PYACS . . . 41
1.7.2 Le langage Python et ses bibliothèques scientifiques . . . 41

1.7.3 Programmation objet . . . 42
1.7.4 Structure de calcul du programme pyacs_make_time_series.py . . 43
1.7.5 Limitations de l'estimation robuste et rapide . . . 44
1.7.6 Exemples pratiques . . . 45
1.8 Conclusions . . . 47

2 Analyses des séries temporelles de position GPS 49
2.1 Introduction . . . 49
2.2 Les séries temporelles de positions GPS . . . 51
2.2.1 Les séries temporelles de position GPS . . . 51
2.2.2 Inventaire des cas à traiter lors de l'analyse des séries temporelles . 52
2.3 Analyse des séries temporelles GPS . . . 59
2.3.1 Modèle général des séries temporelles GPS . . . 59
2.3.2 Détection automatique des outliers et des offsets dans les séries temporelles . . . 62
2.3.2.1 Proposition . . . 62
2.3.2.2 Première identification par rms . . . 62
2.3.2.3 Détection des outliers . . . 63
2.3.2.4 Détection des offsets . . . 66
2.3.2.5 Vérification des offsets détectés . . . 67
2.3.3 Estimation du temps de relaxation post-sismique . . . 68
2.4 Conclusions . . . 72

3 Implémentation : la Classe Gts et le programme pygps_time_series.py 75
3.1 Philosophie . . . 75
3.2 Implémentation sous Python . . . 76
3.2.1 La Classe Gts . . . 76
3.2.2 Les méthodes primitives . . . 77
3.2.3 Les méthodes de second niveau . . . 77
3.2.4 La méthode de visualisation plot . . . 78
3.2.5 Le "wrapper" pygps_time_serie.py . . . 79
3.3 Test de validité . . . 81
3.4 Conclusions . . . 85

4 Etudes de cas : les solutions des réseaux RENAG et Taiwan 87
4.1 La solution RENAG . . . 87
4.1.1 Présentation du réseau RENAG . . . 87
4.1.2 Présentation de la solution étudiée . . . 88

4.1.3 Expression des solutions libres issues de Gamit dans l'ITRF2008 . . 89
4.1.4 Typologie des problèmes rencontrés dans les séries temporelles . . . 91
4.1.5 Résultat de l'analyse automatique avec les paramètres par défaut . 93
4.1.6 Ajustement des paramètres de l'analyse automatique 96
4.1.7 Résultat final 96
4.1.8 Analyse rapide de la qualité du champ de vitesses obtenus 98
4.1.9 Conclusions 100
4.2 La solution Réseau GPS du Taiwan 100
4.2.1 Présentation 100
4.2.2 Expression des solutions libres dans l'ITRF2008 101
4.2.3 Typologie des problèmes rencontrés dans les séries temporelles . . . 102
4.2.4 Résultat de l'analyse automatique avec les paramètres par défaut . 104
4.2.5 Analyse semi-automatique des déformations post-sismiques 105
4.2.6 Résultat final 109
4.2.7 Conclusions 112

Conclusion générale **113**

Bibliographie **117**

Annexes **122**
A Résultats du réseau RENAG 122
B Résultats du réseau GPS du Taiwan 127
C Méthode de la Classe Gts "Geodetic time series" 132

Liste des figures **137**

Liste des tableaux **141**

INTRODUCTION

Le GPS (Global Positioning System) a été développé dans les années 1970 par le Ministère de la Défense des Etats Unis, son objectif initial étant réservé à l'usage militaire [Langley, 2011].

Depuis son ouverture au public au milieu des années 1980, le GPS a permis d'améliorer considérablement la navigation et le positionnement à la surface de la Terre. Outre son usage toujours militaire, le GPS trouve aujourd'hui des applications dans de nombreux domaines comme l'automobile, la navigation, la géodésie. La communauté scientifique l'utilise à de nombreuses fins notamment en météorologie, en climatologie et également en géophysique de précision pour déterminer les mouvements des plaques tectoniques, quantifier la déformation aux frontières des plaques ou bien détecter des déplacements dans les zones de déformation intraplaque.

Les mesures des déformations de la croûte terrestre sont depuis longtemps une préoccupation des scientifiques qui, jusqu'à l'avènement du GPS, avaient recours à des mesures classiques comme celles obtenues par distancemètre. L'arrivée du GPS offre la possibilité d'effectuer des mesures en continu grâce aux stations permanentes. Aujourd'hui, la surface du globe est recouverte de milliers des stations GPS qui captent en permanence les signaux satellites, permettant de générer finalement les séries temporelles de coordonnées de stations GPS et de suivre en continu les mouvements de l'écorce terrestre.

Ainsi, le GPS permet aujourd'hui de calculer des coordonnées de points à la surface de la Terre avec une précision de l'ordre de quelques millimètre à l'échelle globale. Cependant, déterminer les vitesses avec une précision de quelques dixièmes de millimètre par an reste difficile, en particulier en raison des caractéristiques temporelles du bruit des séries temporelles GPS.

Généralement, obtenir des séries temporelles de positions précises à partir des observations de géodésie spatiale nécessite d'exprimer les coordonnées dans un repère précis et stable dans le temps. Celui-ci est généralement l'ITRF (International Terrestrial Reference Frame, Altamimi 2006, Altamimi et al. 2011), obtenu par combinaison de techniques différentes de géodésie spatiale. Classiquement, dans un premier temps les observations de géodésie spatiale sont réduites pour obtenir des jeux de coordonnées où seules les po-

sitions relatives sont bien déterminées. Dans un deuxième temps, la solution réduite est transformée pour exprimer les coordonnées finales dans le repère de référence choisi.

Mon travail de thèse se situe dans le contexte du développement de grand réseaux GPS et GNSS (Global Navigation Satellite System) permanents. Par exemple, le réseau GEONET (GNSS Earth Observation Network System) du Japon géré par le GSI (http://www.gsi.go.jp/ENGLISH/page_e30030.html) comprend plus de 1000 stations, le réseau PBO (Plate Boundary Observatory) géré par l'UNAVCO (http://pbo.unavco.org/) environ 1200 stations dans l'ouest des Etats Unis. A une échelle plus petite, les réseaux permanents en Europe comportent environ 200 stations, et ceux de Taiwan, étudiés pendant ma thèse, comportent plus de 400 stations. Avec une telle quantité d'information, il est pratiquement impossible pour un analyste d'étudier les séries temporelles individuellement, une à une.

L'objet de ma thèse est donc de déterminer des approches, des méthodes et des algorithmes robustes permettant (1) la réalisation rapide de séries temporelles, (2) la résolution automatique des problèmes les plus courants, (3) la manipulation facile des séries temporelles pour extraire les paramètres utiles aux analyses géophysiques, (4) l'identification rapide de problèmes.

Ainsi, le chapitre 1 de ce mémoire est tout d'abord consacré à l'estimation robuste et rapide des paramètres de la transformation de Helmert pour réaliser les séries temporelles. Nous avons développé une approche en coordonnées locales d'estimation des 7 paramètres. Nous avons aussi étudié l'estimation robuste de Dikin et évalué ses performances en robuste et temps de calcul. Nous proposons que l'estimation des 7 paramètres de transformation est l'association de l'estimation robuste (pour détecter les outliers) suivie d'une estimation par moindres carrés sur les données validées. Grâce à ces études, nous avons développé un programme pyacs_make_time_series.py, qui permet d'estimer rapidement et de manière robuste une transformation de Helmert et de réaliser les séries temporelles de positions GPS.

Le chapitre 2 aborde les problématiques des séries temporelles de positions GPS, en particulièrement la détection des outliers, des offsets et des mouvements non-linéaires de déformations post-sismique. Nous avons développé une méthode basée sur les variations de rms (root mean square) calculés par rapport à la position moyenne locale et la différenciation, qui permet d'identifier automatiquement les outliers et les offsets dans les séries temporelles. Nous présentons aussi l'algorithme de l'estimation du temps de relaxation post-sismique pour un modèle exponentiel.

Le chapitre 3 est l'implémentation des méthodes présentées dans le chapitre 2 avec un programme capable de détecter automatiquement les problèmes communs (outliers, offsets) dans les séries temporelles de positions GPS, de les résoudre, d'estimer les signaux

de post-séismes et de fournir l'ensemble des paramètres décrivant la série temporelle.

Le chapitre 4 présente l'analyse de résultats réels pour deux réseaux GPS permanents. Dans le premier cas, il s'agit d'obtenir une solution rapide et de haute qualité des vitesses pour l'étude de la faible déformation tectonique du réseau RENAG dans les Alpes occidentales (http://www.renag.fr). Le second est l'essai d'un traitement de différents réseaux du Taiwan, réalisé par Lin Kuan Chuan dans le cadre de sa thèse. Pour ce réseau, l'étude des discontinuités et des déformations non-linéaires est particulièrement difficile. Ces réseaux constituent donc deux jeux de données permettant d'évaluer les performances et les limitations des outils que j'ai développés.

CHAPITRE 1

TRANSFORMATION DE HELMERT ENTRE DEUX REPÈRES DE RÉFÉRENCE TERRESTRE

1.1 Introduction

Le développement des réseaux GPS permanents permet de suivre l'évolution quotidienne de la position d'un point, exprimée dans un référentiel géocentrique global. Les calculs ultra-précis se font en deux étapes. Dans une première étape, les mesures de pseudo-distances et de phases entre chaque satellite et chaque station au sol sont réduites pour obtenir une solution dite " libre ". Dans cette solution libre, les positions relatives des points sont déterminées avec une précision de quelques millimètres, tandis que la position dans un référentiel absolu n'est encore connue qu'à quelques centimètres près. Dans une deuxième étape, on transforme ces solutions libres pour exprimer les coordonnées dans un repère géocentrique précis et stable dans le temps. Celui-ci est généralement l'ITRF (International Terrestrial Reference Frame, Altamimi 2006, Altamimi et al. 2011). La transformation vers l'ITRF est une similitude (translation d'origine, rotation des axes, facteur d'échelle), appelée dans sa forme linéarisée transformation de Helmert. Si l'estimation des paramètres de la transformation de Helmert est un problème algébrique simple, de nombreuses difficultés se posent en pratique :

- La première difficulté est que l'ensemble des points disponibles pour estimer les paramètres de la transformation n'est pas stable dans le temps. Des stations nouvelles apparaissent et d'autres sont arrêtées ou manquent certains jours. L'apparition ou la disparition d'une station peut modifier significativement la position exprimée après transformation.

- La seconde est la présence d'offsets dans les positions, à la fois dans la solution de référence et dans la solution libre. Ces offsets peuvent être induits par des causes physiques comme le déplacement lié à un séisme ou bien être des artefacts liés à un changement d'équipement ou des problèmes de traitements. Ces offsets se traduisent par une erreur systématique, qui non corrigée, peut induire un biais des positions transformées.

- Enfin, il existe des erreurs accidentelles dans les solutions libres, par exemple dues à un malfonctionnement d'un récepteur GPS à une date donnée.

L'objectif de ce chapitre est de définir une stratégie optimale pour l'estimation des paramètres de la transformation de Helmert, afin d'obtenir la meilleure qualité possible des séries temporelles.

1.2 Le repère de référence terrestre ITRF

L'ITRF (International Terrestrial Reference Frame) est la réalisation de l'ITRS (International Terrestrial Reference System). Un système de référence terrestre est un repère, au sens mathématique du terme, proche de la Terre et tournant avec elle. L'ITRS est un système de référence conventionnel défini par les conditions suivantes [Boucher and Altamimi, 2001] :

- l'origine du repère est le barycentre des masses de la Terre, incluant ses enveloppes océaniques et atmosphériques.

- l'unité de longueur est le mètre (SI). Cette échelle est cohérente avec le temps TGC (Geocentric Coordinate Time) pour un repère géocentrique local.

- l'orientation a été fournie par le Bureau International de l'Heure (BIH) à l'époque 1984.0.

- l'évolution temporelle de l'orientation du repère est assurée par une condition de non-rotation globale des mouvements tectoniques sur l'ensemble de la surface terrestre.

l'ITRF est la réalisation concrète de l'ITRS donnant accès au repère de manière pratique. L'ITRF consiste en un jeu de coordonnées positions et vitesses et de la matrice variance-covariance associée, pour un ensemble d'environ 500 points de contrôle à la surface de la Terre. L'ITRF est obtenu par une combinaison de solutions positions, et de vitesses issues des techniques GPS, SLR (Satellite Laser Ranging), LLR (Lunar Laser Ranging), DORIS (Doppler Orbitography Radiopositioning Integrated by Satellite) et VLBI (Very Long Baseline Interferometry). Les sites où plusieurs techniques sont co-localisées et où un rattachement de qualité existe, permettent d'obtenir une solution unique et cohérente à partir des solutions individuelles des différentes techniques.

Les solutions ITRF sont publiées tous les 2 ans environ par l'IERS (International Earth Rotation Service). Le nombre accolé à chaque réalisation indique la dernière année de données incluse dans le calcul. La version courante est l'ITRF2008 (http://itrf.ensg.ign.fr/ITRF_solutions/2008/ITRF2008.php), réalisée par quatre techniques géodésiques spatiales : VLBI, SLR, GPS, DORIS couvrant respectivement 29, 26, 12.5 et 16 années

d'observation. L'implémentation du système de référence dans l'ITRF2008 a été réalisée par l'application des conditions suivantes [Altamimi et al., 2011] :

- L'origine est définie de telle façon que les paramètres de translation à l'époque 2005.0 et les vitesse de translation sont aussi zéros par rapport aux séries temporelles SLR.

- L'échelle est définie de telle façon que le facteur d'échelle à l'époque 2005.0 et la vitesse d'échelle sont nuls par rapport à l'échelle moyenne et la vitesse d'échelle moyenne des séries temporelles SLR et VLBI.

- L'orientation est définie de telle façon que les paramètres de rotation à l'époque 2005.0 et les vitesses de rotation entre ITRF2008 et ITRF2005 sont nuls. Ces deux conditions sont appliquées sur l'ensemble de 179 stations de référence qui sont localisées à 131 sites (voir la figure 1.1).

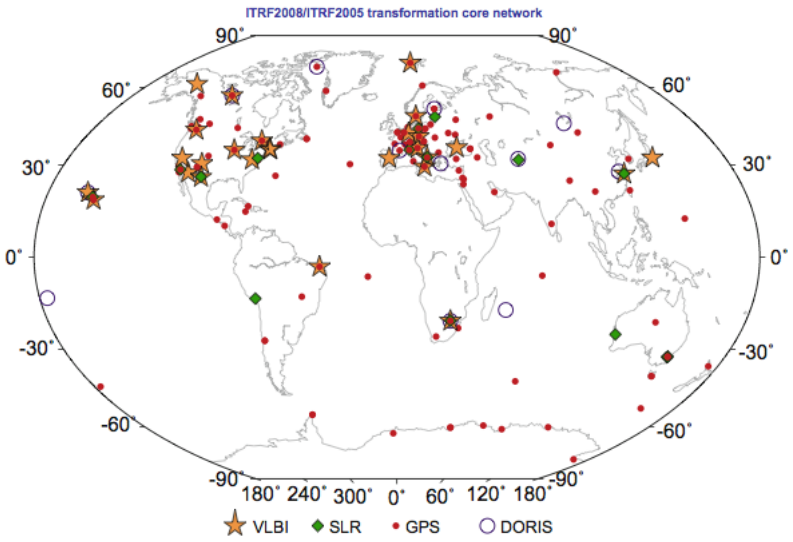


Figure 1.1: Réseau utilisé pour l'estimation des paramètres de transformation entre ITRF2008 et ITRF2005 (http://itrf.ensg.ign.fr)

L'ITRF2008 est réalisé par deux étapes : (1) l'empilement des séries temporelles individuelles pour estimer une solution à long terme pour une technique, comprenant les positions de station à une époque de référence, (2) la combinaison des solutions à long terme des quatre techniques avec les liens locaux dans les sites en co-location.

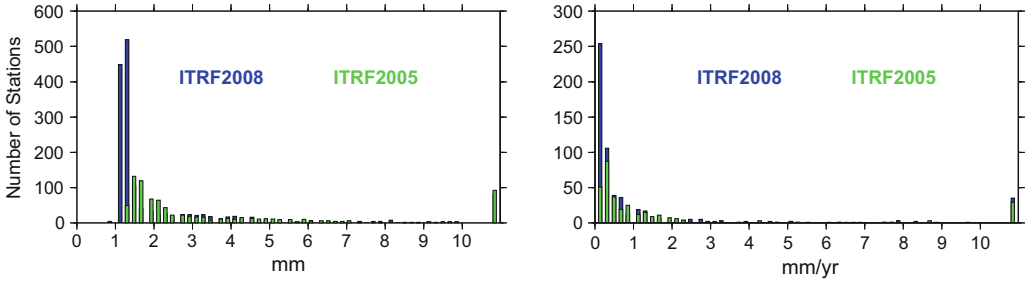


Figure 1.2: Histogrammes de l'erreur sphériques des positions et des vitesses de l'ITRF [Altamimi et al., 2011]

La méthodologie de calcul et la présentation des résultats de l'ITRF2008 sont détaillées dans Altamimi et al. [2011].

1.3 La solution de référence IGS

L'IGS (International GNSS Service) est un service de l'Association Internationale de Géodésie (IAG) depuis 1994. Il fournit des produits comme les orbites précises des satellites GPS et GLONASS, les paramètres de rotation de la Terre, les coordonnées et les vitesses des stations du réseau global IGS, les informations d'ionosphère et de troposphère,... (http://igs.org/overview/viewindex.html). Les produits de l'IGS sont utilisés essentiellement en géodésie et géophysique de haute précision.

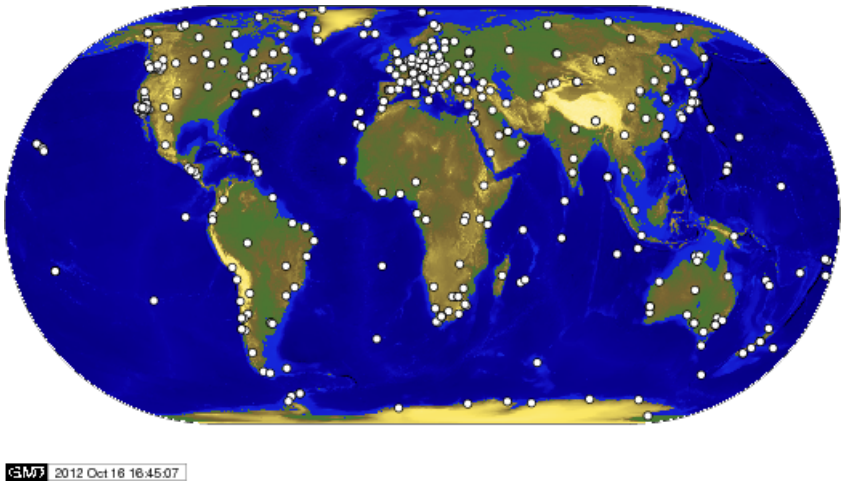


FIGURE 1.3: Réseau GPS permanent de l'IGS (http://igscb.jpl.nasa.gov/network/netindex.html)

Le réseau global de l'IGS contient plus de 350 stations GPS continues bi-fréquences (voir la figure 1.3) dont les données sont transférées directement aux centres de données. Ces centres convertissent les informations reçues en format RINEX (Receiver INdependent

EXchange format) et les centres d'analyses les récupèrent ensuite pour calculer les positions des stations, les paramètres d'horloge des satellites ou encore les orbites. Ces données sont ensuite transmises aux coordinateurs d'analyse qui combinent toutes les solutions des centres d'analyses pour former la solution globale de l'IGS.

Le repère de référence IGS actuel est l'IGS08 publié en Avril 2011, qui contient 232 stations stables, performant du réseau global de l'IGS dans l'ITRF2008 [Rebischung et al., 2012]. Il est dérivé d'une extraction de 232 stations GNSS stables de l'ITRF2008 (les stations de l'IGS08 ont été sélectionnées suivant les critères du tableau 1.1). Avec l'IGS08, un ensemble actualisé des calibrations absolues d'antenne (igs08.atx) a été adopté par l'IGS. Par rapport à l'igs05.atx, outre les corrections d'antenne satellite, la plupart des calibrations d'antenne de récepteur ont été modifiées. Alors que le retraitement des stations IGS a été basé sur les calibrations d'antenne igs05.atx, IGS08 a été rendu compatible avec les dernières calibrations d'antenne igs08.atx.

Criterion	Threshold
Data span	>5 years
Maximum time span between two discontinuities	>3 years
Number of discontinuities	<5
Number of velocity discontinuities	=0
3D formal error of latest velocity estimate	<0.3 mm/year
Residual time series	3D RMS <10 mm; visual inspection

All criteria refer to the ITRF2008 results

TABLE 1.1: Critères utilisés pour la sélection des stations IGS08 à partir des station IGS de l'ITRF2008 [Rebischung et al., 2012]

Les coordonnées, les vitesses de stations IGS08 ainsi que d'autres informations sont mises à jour et publiées sur le lien `ftp://igs-rf.ign.fr/pub/IGS08`.

L'avantage d'utiliser la solution IGS est que (1) elle est issue des stations de haute qualité de l'ITRF dont la cohérence comme solution de référence est plus élevée que celle de l'ITRF [Ray et al., 2004], (2) elle est généralement plus à jour que la réalisation courante de l'ITRF . Par ailleurs, (3) un point important pour la qualité de la solution est la prise en compte des discontinuités dans les sites servant à la mise en référence de la solution. La liste des discontinuités de la solution IGS est mise à jour sur `ftp://igs-rf.ign.fr/pub/discontinuities/soln.snx`. Nous choisissons donc la solution cumulée IGS la plus actuelle comme solution de référence pour l'expression des solutions libres dans l'ITRF.

1.4 Transformation de Helmert entre repères géodésiques

1.4.1 Transformation de Helmert

La transformation de Helmert est une similitude permettant de passer d'un repère à un autre.

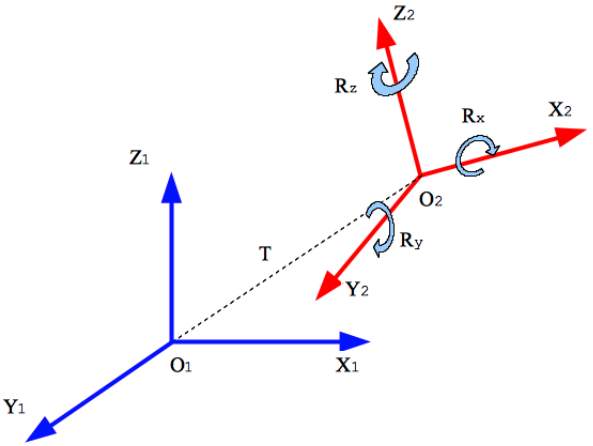


Figure 1.4: Schéma décrivant la transformation de Helmert. T : la translation d'origine. R_X, R_Y, R_Z : les rotations suivant les 3 axes X,Y,Z. La transformation de Helmert comprend en plus l'application d'un facteur d'échelle.

Pour deux repères de référence terrestre (RRT) $\mathbb{R}_1$ et $\mathbb{R}_2$, un point i de coordonnées $X_i^1 = (x_i^1, y_i^1, z_i^1)^T$ exprimées dans $\mathbb{R}_1$ et de coordonnées $X_i^2 = (x_i^2, y_i^2, z_i^2)^T$ exprimées dans $\mathbb{R}_2$, la transformation entre $\mathbb{R}_1$ et $\mathbb{R}_2$ s'écrit

$$X_i^2 = T + \lambda \Re X_i^1 \tag{1.1}$$

Où : $T = \begin{bmatrix} T_x & T_y & T_z \end{bmatrix}^T$ est le vecteur de translation entre les origines des deux RRT,

λ est le facteur d'échelle entre les deux RRT

$\Re$ est la matrice de rotation, et $\Re = R_X R_Y R_Z$

avec : $R_X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cosR_x & -sinR_x \\ 0 & sinR_x & cosR_x \end{bmatrix}$, $R_Y = \begin{bmatrix} cosR_y & 0 & -sinR_y \\ 0 & 1 & 0 \\ sinR_y & 0 & cosR_y \end{bmatrix}$,

$$R_Z = \begin{bmatrix} cosR_z & sinR_z & 0 \\ -sinR_z & cosR_z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

L'équation 1.1 n'est pas linéaire car la matrice $\Re$ inclut les fonctions circulaires. Pour les repères réalisés par des techniques de géodésie spatiale, les translations sont de l'ordre de

100 m, les différences d'échelle et d'orientation entre deux repères sont de l'ordre de 10^{-5} [McCarthy and Petit, 2003], de sorte que $sinR_i \approx R_i$, $cosR_i \approx 1$. On peut donc linéariser les coefficients de la matrice $\Re$ en faisant un développement limité d'ordre 1 :

$$\Re = I + R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & R_z & -R_y \\ -R_z & 0 & R_x \\ R_y & -R_x & 0 \end{bmatrix} \tag{1.2}$$

De plus en écrivant $\lambda = 1 + D$, l'équation 1.1 s'écrit sous la forme linéaire :

$$X_i^2 = X_i^1 + T + DX_i^1 + RX_i^1 \tag{1.3}$$

ou sous la forme du système linéaire :

$$E_i\theta = L_i \tag{1.4}$$

Où, L_i est le vecteur de la différence de coordonnées :

$$L_i = X_i^2 - X_i^1 \tag{1.5}$$

E_i est la matrice des dérivées partielles :

$$E_i = \begin{bmatrix} 1 & 0 & 0 & x_i^1 & 0 & z_i^1 & -y_i^1 \\ 0 & 1 & 0 & y_i^1 & -z_i^1 & 0 & x_i^1 \\ 0 & 0 & 1 & z_i^1 & y_i^1 & -x_i^1 & 0 \end{bmatrix} \tag{1.6}$$

et θ est le vecteur de 7 paramètres de la transformation de Helmert :

$$\theta = \begin{bmatrix} T_x & T_y & T_z & D & R_x & R_y & R_z \end{bmatrix}^T \tag{1.7}$$

1.4.2 Transformation de Helmert en coordonnées locales

Les coordonnées géocentriques cartésiennes (ou globales) $X_{xyz} = (x, y, z)^T$ d'un point P sont transformées en coordonnées topocentriques (ou locales) $X_{enu} = (e, n, u)^T$ (Est, Nord, Vertical) de ce point, par deux étapes :

- Changement d'origine du repère géocentrique O au point P en utilisant le vecteur de translation

$$T = (x_0, y_0, z_0)^T \tag{1.8}$$

- Rotation des axes du repère géocentrique vers le repère local en utilisant la matrice

de rotation

$$R_{xyz\to enu} = \begin{bmatrix} -sin\lambda & cos\lambda & 0 \\ -sin\varphi cos\lambda & -sin\varphi sin\lambda & cos\varphi \\ cos\varphi cos\lambda & cos\varphi sin\lambda & sin\varphi \end{bmatrix} \tag{1.9}$$

où φ et λ sont les latitudes et longitudes géographiques approchées.

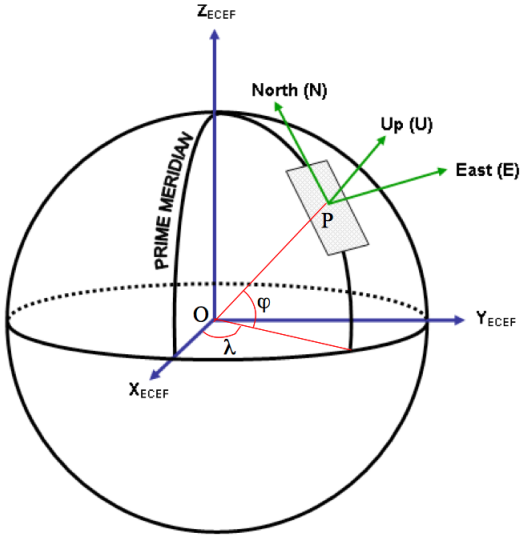


FIGURE 1.5: Coordonnées géocentriques cartésiennes et coordonnées locales (http://en.academic.ru/dic.nsf/enwiki/498677)

La passage des coordonnées globales aux coordonnées locales s'écrit

$$X_{enu} = R_{xyz\to enu} X_{xyz} + T \tag{1.10}$$

et la loi de propagation de variance donne

$$C_{enu} = R_{xyz\to enu} C_{xyz} R^T_{xyz\to enu} \tag{1.11}$$

où C_{enu} est la matrice variance-covariance en coordonnées locales, C_{xyz} est celle en coordonnées globales.

Multiplions l'équations 1.4 avec la matrice de rotation

$$[R_{xyz\to enu} E_{xyz}]\theta = R_{xyz\to enu} L_{xyz} \tag{1.12}$$

Notons les indices 1 et 2 correspondant aux repères géodésiques $\mathbb{R}_1$ et $\mathbb{R}_2$, nous pouvons écrire

$$R_{xyz\to enu} L_{xyz} = (R_{xyz\to enu} X^1_{xyz} + T) - (R_{xyz\to enu} X^2_{xyz} + T) = X^2_{enu} - X^1_{enu}$$

Puis

$$L_{enu} = R_{xyz \to enu} L_{xyz} \tag{1.13}$$

En notant $E_{enu=}R_{xyz \to enu}E_{xyz}$, l'équation de Helmert en coordonnées locales du point P donne

$$E_{enu}\theta = L_{enu} \tag{1.14}$$

Nous proposons d'utiliser la formulation en coordonnées locales 1.14 dans la suite de l'étude. Cette formulation offre plusieurs avantages. Les solutions GPS présentent d'avantage d'erreur sur la composante verticale, en raison soit des défauts de modélisation du retard troposphérique, soit des problèmes de calibration d'antenne et de changement d'équipement. La formulation en coordonnées locales permet d'isoler cette composante et de réaliser des tests statistiques spécifiques. Par ailleurs, elle permet aussi de choisir de dépondérer facilement cette composante lors du calcul de la transformation de Helmert.

1.5 Estimation optimale des paramètres de transformation de Helmert

1.5.1 Difficultés pratiques

Dans la pratique, on constate certaines difficultés lors de l'estimation des paramètres de transformation.

La première est que l'ensemble des points disponibles pour estimer les paramètres changent au cours du temps. Des stations nouvelles apparaissent et d'autres sont arrêtées ou manquent certains jours. L'apparition ou la disparition d'une station peut modifier significativement la position exprimée après transformation, particulièrement si elle est située en extrémité de réseau. Il faut donc définir des stratégies qui diminuent cet effet et assurent la stabilité des solutions.

La seconde est la présence d'offsets dans les positions, à la fois dans la solution de référence et dans la solution libre. Ces offsets peuvent être induits par des causes physiques comme le déplacement lié à un séisme ou à un changement d'équipement. Ces offsets se traduisent par une erreur systématique, qui non corrigée, peut induire un biais des paramètres de transformation et par conséquent des positions transformées. La stratégie à mettre en place est d'identifier ces problèmes pour retirer les points ou les composantes problématiques de l'estimation des paramètres de transformation.

Enfin, il existe des erreurs (outliers) dans les solutions libres, par exemple dues à un malfonctionnement d'un récepteur GPS à une date donnée. L'identification des erreurs est parfois délicate.

1.5.2 Estimation par moindres carrés

L'estimation par moindres carrés est une méthode classique largement utilisée dans les sciences de l'univers à vocation métrologique (astronomie, géodésie,...). Son algorithme a été développé indépendamment par Legendre et Gauss au 18ème siècle.

La méthode des moindres carrés est utilisée pour résoudre un ensemble d'équations linéaires présentant plus d'équations que de variables inconnues. Cette méthode utilise la norme L2

$$\sum_{i=1}^{n} e_i^2 \rightarrow min \tag{1.15}$$

où e_i est l'erreur d'observation ou le résidu d'observation.

La théorie et les applications de la méthode des moindres carrés sont présentées par plusieurs auteurs comme Sillard [2001], Press et al. [1992]. Dans cette étude, nous présentons cette méthode pour l'estimation des paramètres de transformation de Helmert.

Pour chaque point i dont les coordonnées sont exprimées dans les deux repères, on obtient 3 équations 1.14 suivantes

$$E_i\theta = L_i,\ P_i$$

Où P_i est la matrice poids associée aux observations

$$P_i = (C_i^1 + C_i^2)^{-1} \tag{1.16}$$

C_i^1 et C_i^2 sont les matrices variance-covariance de position correspondant aux repères $\mathbb{R}_1$ et $\mathbb{R}_2$ du point i.

La méthode des moindres carrés est réalisée pour le système d'équation 1.14 si le nombre d'équations est supérieur à celui d'inconnus. On doit donc avoir au moins 3 points communs pouvoir estimer les 7 paramètres de la transformation de Helmert. Supposons que l'on a n points communs entre les repères $\mathbb{R}_1$ et $\mathbb{R}_2$ ($3n > 7$), l'estimateur par moindres carrés est

$$\hat{\theta} = (E^T P E)^{-1} E^T P L \tag{1.17}$$

Où : $L = [L_1, L_2, ..., L_n]^T$

$E = [E_1, E_2, ..., E_n]^T$

La matrice de pondération P, si on considère que les point communs sont indépen-

dants, est

$$P = \begin{bmatrix} P_1 & 0 & ... & 0 \\ 0 & P_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & P_n \end{bmatrix} \tag{1.18}$$

de sorte que le vecteur des résidus V est

$$V = E\hat{\theta} - L \tag{1.19}$$

On vérifie la qualité d'une estimation par moindres carrés en regardant la variance des paramètres estimés, la cohérence entre les résidus de l'estimation et la variance a priori des observations.

$$\hat{\sigma}_0^2 = \frac{V^T P V}{3n - 7} \tag{1.20}$$

où n est le nombre de points communs.

Les variances des inconnues s'écrivent

$$\sigma_{\hat{\theta}}^2 = \hat{\sigma}_0^2 (E^T P E)^{-1} \tag{1.21}$$

Les variances des résidus s'écrivent aussi

$$\sigma_V^2 = \hat{\sigma}_0^2 (P^{-1} - E(E^T P E)^{-1} E^T) \tag{1.22}$$

La méthode des moindres carrés est une estimation qui suppose que les erreurs d'observation suivent la loi Normale, c'est-à-dire qu'il n'y a que des erreurs aléatoires sur les observations.

L'influence des composantes de coordonnées différentes L_i du point commun i sur l'estimation par moindres carrés des 7 paramètres Helmert, suit Yang [1999]

$$IF(L_i, \hat{\theta}) = (E^T P E)^{-1} E_i^T P_i V_i \tag{1.23}$$

La fonction 1.23 montre que toute erreur sur L_i aura des effets sur les 7 paramètres estimés, dont l'amplitude dépend de la configuration des problèmes (matrice E) et des poids donnés à cette observation (matrice P et P_i).

1.5.3 Estimation robuste de Dikin

Pour l'estimation des paramètres de transformation de Helmert, la méthode des moindres carrés est optimale si les observations ne sont influencées que par des erreurs aléatoires

suivant une distribution normale. En pratique, on rencontre de nombreuses erreurs grossières dans les séries temporelles d'observation de position GPS. La méthode des moindres carrés n'est donc pas encore optimale dans ce cas là. Il faut donc identifier et rejeter les erreurs grossières avant l'estimation par moindres carrés.

On considère qu'une erreur grossière (outlier) est ici définie comme une erreur significativement plus grande que son incertitude associée.

Plusieurs méthodes classiques de détection des erreurs grossières reposent sur des tests statistiques sur les résidus ajustés. Cependant ces méthodes ne sont efficaces que si un petit nombre d'erreurs grossières existe dans les observations [Koch, 1999]. Dans un cas faiblement redondant, les techniques d'estimation robuste sont utiles pour la détection des erreurs de données brutes avant une estimation finale effectuée par la méthode des moindres carrés [Khodabandeh and Amiri-Simkooei, 2011].

On trouve de nombreux articles dans la littérature géodésique concernant l'estimation robuste et les techniques de détection des erreurs grossières qui ont été publiés par exemple Gui and Zhang [1998], Koch [1999], Yang [1999], Khodabandeh and Amiri-Simkooei [2011].

Dans l'estimation robuste, la norme L1 est la plus utilisée. Elle s'écrit

$$\sum_{i=1}^{m} |e_i| \rightarrow min \tag{1.24}$$

Où $|e_i|$ représente les valeurs absolues du vecteur des résidus.

L'estimateur de la norme L1 est moins sensible aux erreurs grossières que l'estimateur des moindres carrés. Dans le cas d'estimation d'observations répétées d'une quantité inconnue, par exemple la série temporelle de position GPS, les moindres carrés donne la moyenne des observations, tandis que la minimisation de la norme L1 est proche de la médiane des observations [Press et al., 1992]. Il est évident que l'effet des erreurs grossières sur la moyenne des observations est plus important que sur la médiane.

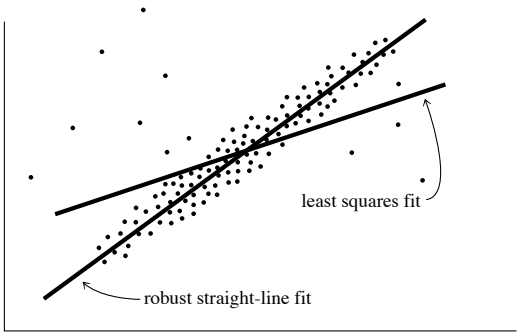


Figure 1.6: Exemple illustrant le comportement d'un estimateur robuste et de l'estimateur des moindres carrés. En présence d'outlier, pour le cas de la régression linéaire [Press et al., 1992]

La figure 1.6 montre l'estimation de la vitesse à partir d'une série temporelle par deux estimateurs : la méthode des moindres carrés et l'estimation robuste dans le cas de la présence d'outliers. L'estimation par moindres carrés a une sensibilité non désirée aux outliers.

Dans la suite, nous présentons et étudions la méthode de Dikin pour l'estimation robuste.

1.5.3.1 Moindres valeurs absolues et programmation linéaire

La programmation linéaire est une branche des mathématiques qui a pour but de résoudre des problèmes d'optimisation linéaire. Elle s'applique à une classe des problèmes où on est amené à minimiser (ou maximiser) une application linéaire, appelée fonction d'objectif (par exemple la minimisation 1.24), sur un ensemble d'équations, dites contraintes (par exemple l'équation 1.25).

Soit le modèle linéaire des équations de n observations :

$$\tilde{y} = \tilde{A}\beta + \tilde{e} \tag{1.25}$$

Où: $\tilde{y}$ est la vecteur de n observations, $\tilde{A}$ est la matrice modèle, $\tilde{e}$ est le vecteur d'erreurs, β est le vecteur de m inconnues

On divise l'observation $\tilde{y}_i$ par son écart-type $\sigma_{\tilde{y}_i}$ pour obtenir les observations dont l'écart-type est égal à 1. On obtient donc le modèle équivalent [Koch, 1999]:

$$y = A\beta + e \tag{1.26}$$

Où:

$$y = P^{\frac{1}{2}}\tilde{y},\ A = P^{\frac{1}{2}}\tilde{A},\ e = P^{\frac{1}{2}}\tilde{e} \tag{1.27}$$

P est la matrice de pondération

$$P = \begin{bmatrix} \sigma_{\widetilde{y_1}}^{-2} & 0 & ... & 0 \\ 0 & \sigma_{\widetilde{y_2}}^{-2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \sigma_{\widetilde{y_n}}^{-2} \end{bmatrix} \tag{1.28}$$

On définit la solution β tell que $\|A\beta - y\|$ soit minimale.

On assimile les e_i à la différence de deux réels positifs u_i, v_i soient $e_i = u_i - v_i$. Sous la forme de vecteur, on a

$$e = u - v \tag{1.29}$$

La condition minimale 1.24 devient

$$\sum_{i=1}^{m} |e_i| = \sum_{i=1}^{m} |u_i - v_i| = \sum_{i=1}^{m} (u_i + v_i) \to min \tag{1.30}$$

On constate que e_i est minimal si u_i est minimal avec $v_i = 0$, ou v_i est minimal avec $u_i = 0$.

On peut aussi analyser le vecteur d'inconnues β par deux autres vecteurs non-négatifs suivants

$$\beta = \lambda - \gamma \tag{1.31}$$

La programmation linéaire inclut [Khodabandeh and Amiri-Simkooei, 2011]:

- La fonction d'objet

$$z = f^T x = \left[0_m^T \vdots 0_m^T \vdots c_m^T \vdots c_m^T\right] \begin{bmatrix} \lambda \\ \gamma \\ u \\ v \end{bmatrix} \to min \tag{1.32}$$

- La fonction de contrante

$$B^T x = \left[A \vdots -A \vdots I_m \vdots -I_m\right] \begin{bmatrix} \lambda \\ \gamma \\ u \\ v \end{bmatrix} = y \tag{1.33}$$

Où: 0_m est le vecteur nul de taille m

c_m est le vecteur unitaire de taille m

I_m est la matrice identité de taille m.m

$$x = \begin{bmatrix} \lambda & \gamma & u & v \end{bmatrix}^T \geq 0$$

Les équations 1.32, 1.33 sont généralement résolues par des approches différentielles, Khodabandeh and Amiri-Simkooei [2011] montre que la méthode affine de Dikin est plus rapide en calcul et plus facile d'application que la méthode complexe.

1.5.3.2 La méthode affine de Dikin

La méthode de Dikin est une des méthodes les plus simples [Dantzig and Thapa, 1997]. Dans cette méthode, un point initial intérieur réalisable est donné ou trouvé et la méthode l'améliore dans un processus itératif. On voit qu'à partir des équations 1.32 et 1.33, le processus itératif s'écrit

$$B^T x^{(t+1)} = B^T(x^{(t)} - \alpha f) = y \tag{1.34}$$

Où α est la grandeur positive scalaire.

Dans la méthode de Dikin, les matrices x, B, f sont remplacées par les quantités transférées:

$$\breve{x} = D^{-1}x,\ \breve{B} = DB,\ \breve{f} = Df \tag{1.35}$$

Où D est la matrice diagonale du vecteur x

$$D = diag(x) \tag{1.36}$$

La précessus itératif du vecteur $x^{(t+1)}$ est donc

$$\breve{x}^{(t+1)} = \breve{x}^{(t)} + \alpha d \tag{1.37}$$

Où d est le vecteur de direction et $P_B^{\perp}$ la matrice de projection quand le vecteur f est projeté orthogonalement sur l'espace $N(B^T)$:

$$d = -P_B^{\perp}\breve{f} \tag{1.38}$$

$$P_B^{\perp} = I - \breve{B}(\breve{B}^T\breve{B})^{-1}\breve{B}^T \tag{1.39}$$

En pratique, la grandeur scalaire α est choisie telle que [Dantzig and Thapa, 1997]

$$\alpha = 0.9/\theta \tag{1.40}$$

Où θ est le plus petit élément du vecteur d avec le signe négatif, soit

$$\theta = -min(d) \tag{1.41}$$

La processus d'itération 1.37 est répétée jusqu'à ce que $x^{(t+1)}$ et $x^{(t)}$ soient similaires. En pratique, on choisit le seuil de la différence entre deux itérations, par exemple 10^{-3} ou 10^{-5}.

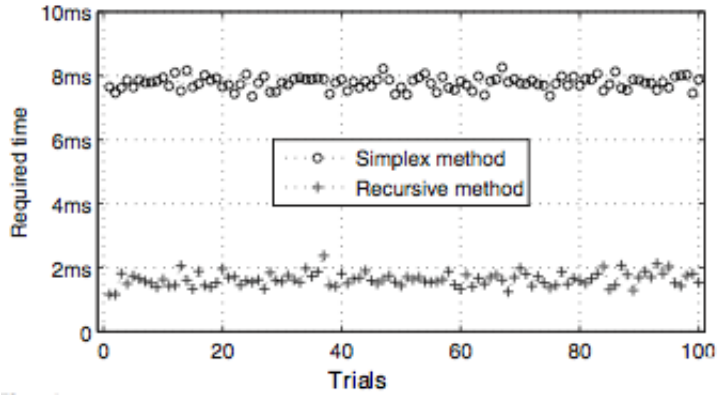


Figure 1.7: Temps de calcul pour une estimation robuste par la méthode simplex et la méthode Dikin [Khodabandeh and Amiri-Simkooei, 2011]

Cette méthode est rapide et plus rapide que la méthode standard du simplex. Des tests simples montrent qu'elle converge environ quatre fois plus rapidement que le simplex, mais seulement pour un problème linéaire (voir la figure 1.7, la méthode Dikin se nomme la méthode récursive). Nous testons ensuite la validité de cette approche dans le paragraphe suivant.

1.6 Validation de l'estimateur robuste Dikin pour les transformations de Helmert

1.6.1 Objet des tests

On cherche à évaluer la performance de différents estimateurs pour une transformation de Helmert en présence d'erreurs. Pour cela, on produit un jeu synthétique parfait, puis on introduit des erreurs et on étudie le comportement du système.

Tous les tests sont réalisés par le programme pyacs_make_time_series.py (développé dans la cadre de ma thèse). On s'intéressera à trois aspects : l'exactitude des paramètres estimés, la comparaison de l'estimation de Dikin par rapport à un moindres carrés, et les paramètres de temps du calcul.

1.6.2 Fabrication du jeu synthétique

Pour fabriquer le jeu de référence, on se donne un ensemble de points i de coordonnées $X_i(x_i, y_i, z_i)$ qui définissent les coordonnées du jeu de référence.

On fabrique ensuite le jeu libre en deux étapes. Premièrement, on perturbe le jeu de référence $X_i(x_i, y_i, z_i)$ avec des erreurs suivant la loi Normale et on vérifie ensuite le niveau

de précision avec lequel on retrouve les paramètres de transformation, qui doivent valoir tous 0. Deuxièmement, on introduit ensuite une ou des erreurs grossières de différentes magnitudes et on regarde le comportement du système.

Chaque simulation est caractérisée par le nombre de points n , la distribution spatiale des points à la surface de la Terre et la valeur de l'erreur aléatoire. Pour la plupart des simulations, on prend une erreur sphérique sur trois composantes égale entre les points.

La matrice variance-covariance est négligée.

Trois programmes permettent de réaliser les simulations :

1. GEN_SYNTHETIC_DATA(n, long_min,long_max,lat_min,lat_max) : Génère un ensemble de points au format SINEX (sans la matrice variance-covariance).

2. PERTURB_DATA (SINEX , sigma) : Rajoute des erreurs gaussiennes.

3. ADD_ERROR (SINEX, nom_du point, erreur) : Rajoute une erreur/plusieurs erreurs grossières sur un ou des points.

1.6.3 Résultats

Le test est réalisé pour étudier le comportement du système sur 3 critères : la précision ; l'influence d'erreur sur les paramètres de transformations estimées, l'influence sur les séries temporelles ; et les performances du calcul.

1.6.3.1 Etude de l'influence d'une erreur accidentelle

Pour ce test initial, on simule 10 sites distribués aléatoirement à la surface de la Terre (figure 1.8) définissant le fichier de référence (*ori.snx*).

On introduit ensuite une perturbation de ce jeu par des erreurs suivant la loi Normale avec $\sigma = 2mm$, puis on introduit une erreur allant de 1 mm à 30 m sur la composante X du site X000 (voir la figure 1.8) pour fabriquer le jeu libre erroné de 30 fichiers sinex.

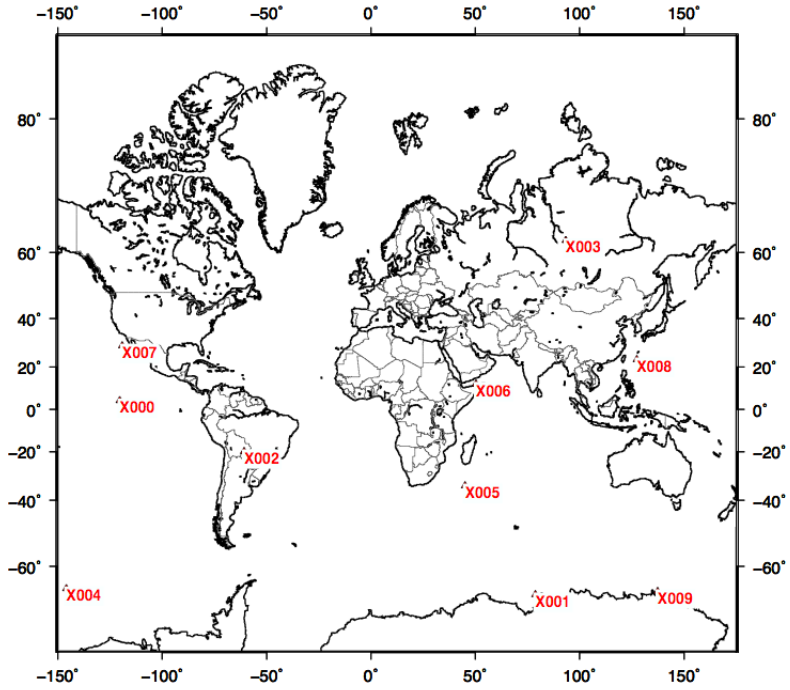


Figure 1.8: Carte des sites du réseau global utilisé pour le test 1.6.3.1

On estime successivement 30 transformations de Helmert du jeu libre erroné par l'estimateur robust de Dikin et l'estimateur par moindres carrés.

On caractérise l'erreur de paramètres de transformation de Helmert suivant la formule (nommée erreur Helmert dans la suite)

$$e_{Helmert} = \sqrt{\|T\|^2 + (DR_T)^2 + (206265^{-1}\,\|\omega\|\,R_T))^2} \tag{1.42}$$

Où: $\|T\| = \sqrt{T_X^2 + T_Y^2 + T_Z^2}$, $\|\omega\| = \sqrt{\omega_X^2 + \omega_Y^2 + \omega_Z^2}$

T_X, T_Y, T_Z, D , ω_X, ω_Y, ω_Z sont les 7 paramètres de la transformation de Helmert

et $R_T = 6371km$ est le rayon de la Terre.

L'indice l'erreur Helmert donne l'erreur maximum sur la Terre.

Le tableau 1.2 résume les résultats obtenus et la figure 1.9 montre l'évolution de l'erreur Helmert en fonction de l'amplitude de l'erreur introduite.

Erreur intr-	Estimation par moindres carrés								Estimation robuste de Dikin							
-oduite (m)	Tx(cm)	Ty(cm)	Tz(cm)	D(ppb)	Rx(mas)	Ry(mas)	Rz(mas)	e_H(m)	Tx(cm)	Ty(cm)	Tz(cm)	D(ppb)	Rx(mas)	Ry(mas)	Rz(mas)	e_H(m)
0,001	0,03	0,01	0,01	-0,07	0,01	0,00	0,00	0,001	0,01	0,01	-0,01	-0,03	0,02	0,01	0,01	0,001
0,005	-0,01	0,01	0,01	-0,04	0,01	0,00	-0,02	0,001	-0,04	0,02	-0,01	-0,04	0,01	0,01	-0,01	0,001
0,010	-0,07	0,01	0,02	0,00	0,01	-0,01	-0,04	0,001	-0,06	-0,01	-0,01	0,01	0,02	0,00	-0,02	0,001
0,015	-0,13	0,00	0,02	0,03	0,01	-0,01	-0,07	0,002	-0,07	-0,02	0,00	0,03	0,02	0,00	-0,02	0,001
0,020	-0,18	0,00	0,03	0,07	0,01	-0,02	-0,09	0,003	-0,07	-0,02	-0,01	0,04	0,02	0,00	-0,02	0,001
0,040	-0,41	-0,02	0,05	0,22	0,01	-0,05	-0,20	0,007	-0,07	-0,02	-0,01	0,04	0,02	0,00	-0,02	0,001
0,060	-0,63	-0,04	0,06	0,37	0,01	-0,07	-0,30	0,011	-0,07	-0,03	-0,01	0,04	0,02	0,00	-0,02	0,001
0,100	-1,08	-0,08	0,10	0,67	0,02	-0,13	-0,50	0,018	-0,07	-0,02	-0,02	0,04	0,02	0,00	-0,02	0,001
0,150	-1,65	-0,13	0,15	1,04	0,03	-0,19	-0,76	0,027	-0,06	-0,02	-0,02	0,04	0,02	0,00	-0,02	0,001
0,200	-2,21	-0,18	0,19	1,41	0,03	-0,26	-1,01	0,036	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,001
0,400	-4,46	-0,37	0,37	2,90	0,05	-0,52	-2,03	0,073	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,001
0,600	-6,71	-0,57	0,55	4,40	0,08	-0,78	-3,05	0,111	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,001
0,800	-8,96	-0,76	0,73	5,89	0,10	-1,05	-4,08	0,148	-0,08	-0,01	-0,02	0,02	0,02	0,00	-0,02	0,001
1,000	-11,22	-0,96	0,91	7,38	0,12	-1,31	-5,10	0,185	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,001
2,000	-22,48	-1,93	1,81	14,84	0,23	-2,63	-10,20	0,371	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,001
5,000	-56,26	-4,84	4,51	37,21	0,57	-6,58	-25,51	0,928	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,001
10,000	-112,56	-9,69	9,01	74,51	1,12	-13,16	-51,02	1,857	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,001
15,000	-168,86	-14,54	13,51	111,80	1,68	-19,74	-76,54	2,786	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,001
20,000	-225,16	-19,39	18,01	149,09	2,24	-26,33	-102,06	3,715	-0,07	-0,02	-0,02	0,02	0,02	0,00	-0,02	0,001
30,000	-337,77	-29,09	27,01	223,68	3,35	-39,49	-153,09	5,572	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,001

Table 1.2: Evolution de la valeur des 7 paramètres de la transformation de Helmert en fonction de l'amplitude de l'erreur introduite. e_H est définie par 1.42.

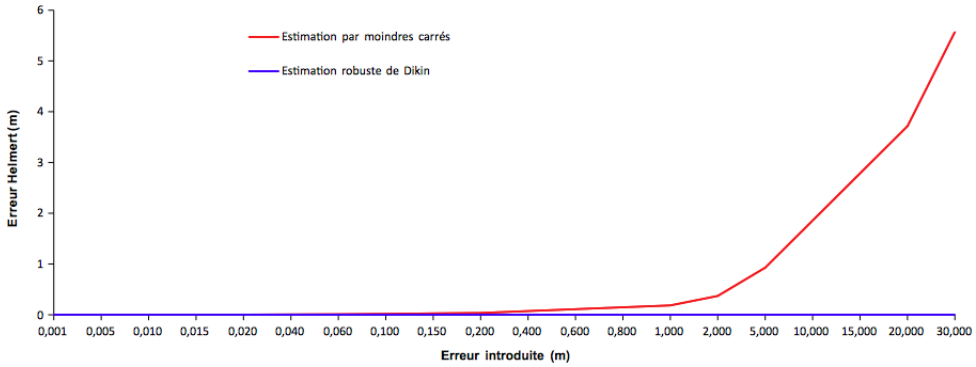


Figure 1.9: Diagramme de l'évolution des erreurs de la transformation de Helmert en fonction de l'amplitude de l'erreur introduite pour les estimateurs de Dikin et des moindres carrés

Résultats

La figure 1.9 montre que les valeurs des paramètres résolus par l'estimateur robuste de Dikin restent proche de 0 dans tous les cas. Par contre, les valeurs des paramètres résolus par la méthode des moindres carrés sont toujours différentes de 0 et elles grossissent avec l'augmentation de l'amplitude d'erreur introduite, particulièrement à partir de l'erreur à 100 mm.

On conclut que l'estimation robuste de Dikin reste stable à l'ajout d'une erreur, en particulier, son exactitude est très peu dépendante de l'amplitude de l'erreur introduite.

1.6.3.2 Etude de l'effet de réseau local

Dans le test précédent, la distribution globale des sites permet une bonne estimation indépendante de chaque paramètre de la transformation de Helmert. Dans ce test, on étudie l'effet de réseau local, c'est-à-dire, le comportement du système pour une petite partie de la sphère, ce qui correspond au cas pratique des études à but tectonique.

Dans ce test, pour étudier l'effet de réseau local, on fabrique le jeu des données comme le test 1.6.3.1, mais on choisit ici le cas extrême de 10 sites distribués sur région de surface 10x10 km, ils créent donc un réseau local (figure 1.10).

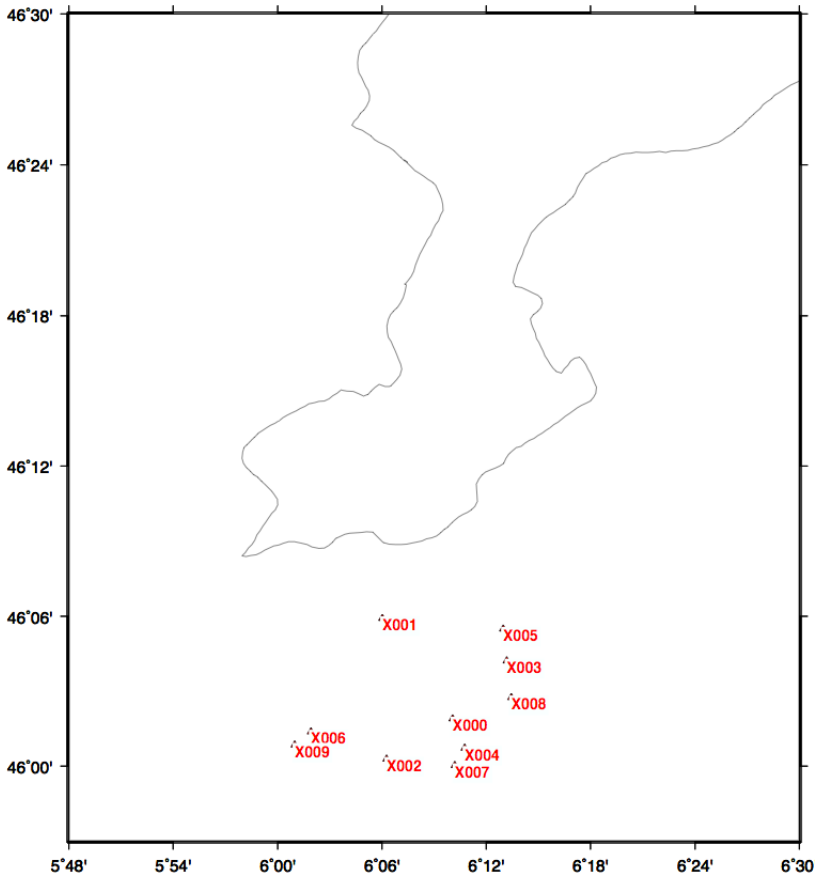


Figure 1.10: Carte des sites du réseau local utilisé pour le test 1.6.3.2

On utilise la formule 1.42 pour évaluer l'erreur maximum du réseau régional e'_H. R_T est remplacé par

$$R' = \sqrt{X_m^2 + Y_m^2 + Z_m^2} \tag{1.43}$$

Où (X_m, Y_m, Z_m) sont les coordonnées moyennes des sites.

Le tableau 1.3 résume les résultats obtenus et la figure 1.11 montre l'évolution de l'erreur Helmert en fonction de l'amplitude de l'erreur introduite.

Erreur intr-	Estimation par moindres carrés								Estimation robuste de Dikin							
-oduite (m)	Tx(cm)	Ty(cm)	Tz(cm)	D(ppb)	Rx(mas)	Ry(mas)	Rz(mas)	e'_H(m)	Tx(cm)	Ty(cm)	Tz(cm)	D(ppb)	Rx(mas)	Ry(mas)	Rz(mas)	e'_H(m)
0,001	82,68	26,79	-74,71	-9,05	-15,73	-38,55	-28,68	1,754	0,79	-49,97	48,00	-49,22	-40,78	7,56	-17,83	1,59
0,005	73,19	10,76	-60,82	-12,54	-17,96	-33,05	-23,42	1,508	-4,03	-63,41	74,76	-72,58	-41,51	15,01	-11,78	1,779
0,010	61,33	-9,28	-43,47	-16,90	-20,75	-26,16	-16,84	1,253	0,60	-62,72	78,05	-81,41	-46,30	14,14	-16,89	1,943
0,015	49,47	-29,31	-26,12	-21,27	-23,54	-19,28	-10,26	1,096	-4,62	-63,56	75,93	-73,24	-39,94	15,57	-10,10	1,746
0,020	37,62	-49,35	-8,76	-25,63	-26,33	-12,39	-3,68	1,079	-4,09	-63,43	74,70	-72,43	-41,55	15,00	-11,81	1,780
0,040	-9,82	-129,49	60,65	-43,09	-37,49	15,14	22,63	2,086	-0,85	-61,79	65,20	-65,46	-26,63	13,68	2,74	1,357
0,060	-57,25	-209,63	130,06	-60,54	-48,66	42,68	48,94	3,588	-31,20	-66,28	46,60	-10,95	-44,71	14,66	-15,11	1,757
0,100	-152,12	-369,91	268,88	-95,46	-70,98	97,75	101,56	6,777	-31,24	-66,29	46,54	-10,84	-44,75	14,66	-15,15	1,758
0,150	-270,71	-570,27	442,41	-139,10	-98,88	166,59	167,34	10,828	-31,18	-66,26	46,60	-10,98	-44,76	14,65	-15,16	1,758
0,200	-389,29	-770,62	615,94	-182,74	-126,79	235,43	233,12	14,897	-31,18	-66,29	46,61	-10,98	-44,73	14,66	-15,12	1,757
0,400	-863,63	-1572,04	1310,05	-357,30	-238,40	510,79	496,23	31,211	-31,54	-66,30	46,29	-10,23	-44,60	14,68	-15,00	1,753
0,600	-1337,97	-2373,46	2004,17	-531,85	-350,02	786,15	759,34	47,538	-31,17	-66,29	46,62	-11,01	-44,69	14,66	-15,08	1,756
0,800	-1812,31	-3174,87	2698,28	-706,41	-461,63	1061,51	1022,45	63,868	-31,24	-66,28	46,54	-10,84	-44,76	14,65	-15,17	1,758
1,000	-2286,65	-3976,29	3392,39	-880,97	-573,25	1336,87	1285,56	80,199	-31,21	-66,29	46,59	-10,93	-44,70	14,66	-15,10	1,756
2,000	-4658,36	-7983,38	6862,96	-1753,77	-1131,32	2713,67	2601,11	161,861	-32,28	-66,20	45,66	-8,73	-44,72	14,70	-15,19	1,757
5,000	-11773,47	-20004,64	17274,68	-4372,16	-2805,55	6844,07	6547,77	406,857	-31,17	-66,28	46,64	-11,02	-44,73	14,66	-15,13	1,757
10,000	-23631,99	-40040,08	34627,53	-8736,14	-5595,93	13728,07	13125,54	815,186	-31,27	-66,29	46,56	-10,83	-44,73	14,67	-15,13	1,757
15,000	-35490,52	-60075,52	51980,38	-13100,12	-8386,31	20612,06	19703,31	1223,516	-31,14	-66,27	46,66	-11,07	-44,67	14,67	-15,06	1,755
20,000	-47349,04	-80110,96	69333,23	-17464,10	-11176,69	27496,06	26281,08	1631,845	-31,15	-66,28	46,67	-11,09	-44,72	14,66	-15,11	1,757
30,000	-71066,08	-120181,84	104038,93	-26192,07	-16757,46	41264,06	39436,62	2448,505	-27,00	-66,29	52,56	-22,24	-46,75	14,82	-16,98	1,837

Table 1.3: Evolution de la valeur des 7 paramètres de la transformation de Helmert en fonction de l'amplitude de l'erreur introduite. e_H est définie par 1.42 et 1.43.

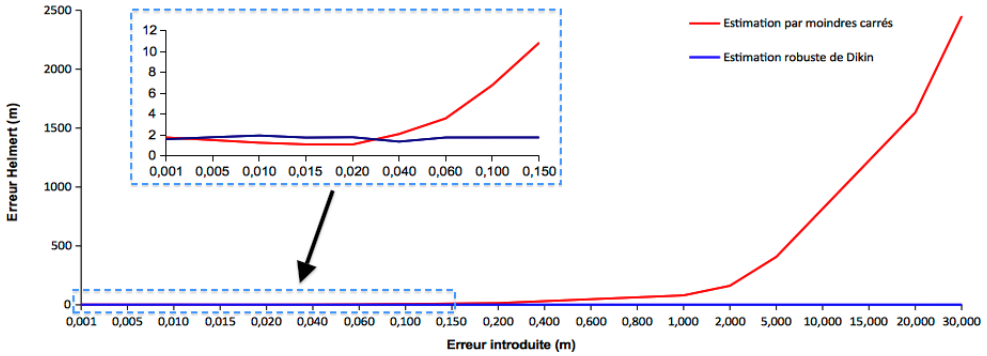


Figure 1.11: Diagramme de l'évolution des erreurs de la transformation en fonction de l'amplitude de l'erreur introduite pour les estimateurs de Dikin et des moindres carrés

Résultats

Le tableau 1.3 montre que les valeurs des paramètres résolus par l'estimateur robuste de Dikin et par la méthode des moindres carrés sont toujours différentes que 0. Pour la méthode des moindres carrés, elles grossissent avec l'augmentation de l'amplitude d'erreur introduite. Plus particulièrement, elles augmentent beaucoup plus rapidement à partir de 60 mm (voir la figure exagérée dans 1.11). Cependant, pour l'estimation robuste de Dikin, elles restent environ de 175 cm. Malgré que le résultat de la méthode Dikin est bien meilleur, on sait que, dans un réseau local, l'erreur influence très fortement à la transformation de Helmert estimée.

Les paramètres estimés du réseau local restent cependant très bons dans proches du barycentre du réseau, par contre ils sont très mauvais de l'autre côté de la Terre. L'erreur maximum du point de ces réseaux est toujours plus grande que celle du réseau global.

C'est l'effet de réseau local. La méthode Dikin réduit cet effet et fournit des résultats significativement meilleurs que la méthode des moindres carrés.

1.6.3.3 Etude de l'effet d'asymétrie du réseau

On continue à étudier l'effet de réseau local en regardant maintenant les effets d'asymétrie. Comme pour le test 1.6.3.1, on fabrique un jeu des données de 10 sites, mais ces sites forment un réseau asymétrique avec le site X009 est distant des autres d'environ 1500 km (figure 1.12). Dans ce test, on introduit une erreur allant de 1 mm à 30 m sur la composante Y du site X009, pour fabriquer le jeu libre.

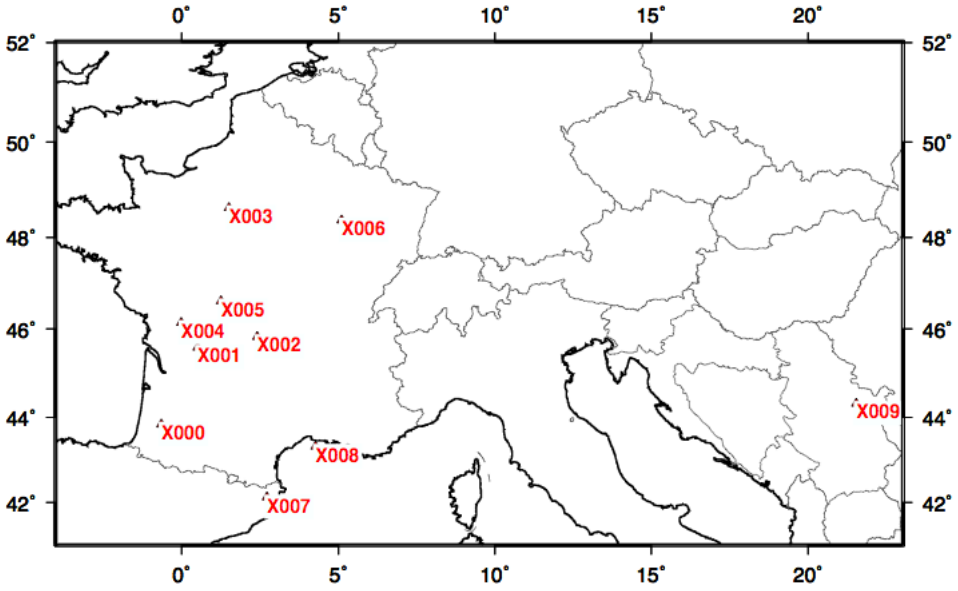


Figure 1.12: Carte des sites du réseau asymétrique utilisé pour le test 1.6.3.3

On continue à utiliser la formule 1.42 en évaluant de l'erreur maximum du réseau e'_H. R_T est calculé suivant la formule 1.43, à partir de sites en excluant le site distant X009.

Le résultats sont présentés dans le tableau 1.4 et la figure 1.13.

Erreur intr-	Estimation par moindres carrés								Estimation robuste de Dikin							
-oduite (m)	Tx(cm)	Ty(cm)	Tz(cm)	D(ppb)	Rx(mas)	Ry(mas)	Rz(mas)	e'_H(m)	Tx(cm)	Ty(cm)	Tz(cm)	D(ppb)	Rx(mas)	Ry(mas)	Rz(mas)	e'_H(m)
0,001	-0,08	-0,94	-0,33	0,62	-0,19	-0,05	0,26	0,016	-0,02	-0,93	-0,37	0,58	-0,17	-0,08	0,27	0,01
0,005	0,79	-1,14	0,57	-1,34	-0,22	-0,05	0,33	0,024	1,04	-1,14	0,74	-1,83	-0,19	-0,06	0,36	0,024
0,010	1,88	-1,38	1,69	-3,79	-0,27	-0,04	0,40	0,052	2,18	-1,38	2,42	-4,98	-0,24	0,06	0,43	0,050
0,015	2,97	-1,62	2,82	-6,24	-0,31	-0,04	0,48	0,083	2,93	-1,78	2,81	-6,19	-0,26	-0,03	0,60	0,063
0,020	4,06	-1,87	3,94	-8,69	-0,35	-0,03	0,56	0,114	2,94	-2,35	2,93	-6,35	-0,30	-0,01	0,80	0,068
0,040	8,42	-2,84	8,43	-18,48	-0,53	-0,01	0,88	0,239	3,54	-4,48	3,04	-7,05	-0,33	-0,12	1,74	0,096
0,060	12,77	-3,81	12,93	-28,28	-0,70	0,00	1,20	0,364	4,64	-6,49	3,69	-8,90	-0,38	-0,22	2,60	0,132
0,100	21,48	-5,75	21,92	-47,87	-1,05	0,04	1,83	0,614	7,48	-10,39	5,69	-14,11	-0,50	-0,42	4,28	0,213
0,150	32,37	-8,18	33,15	-72,35	-1,49	0,09	2,63	0,928	8,36	-15,25	8,15	-17,77	-1,11	-0,05	5,92	0,290
0,200	43,26	-10,61	44,39	-96,84	-1,92	0,13	3,42	1,241	7,64	-20,40	8,96	-17,83	-2,31	0,30	7,08	0,348
0,400	86,82	-20,33	89,33	-194,79	-3,66	0,32	6,60	2,495	7,12	-41,05	12,01	-20,32	-6,87	1,12	11,99	0,623
0,600	130,38	-30,05	134,27	-292,74	-5,41	0,50	9,78	3,748	6,82	-62,01	13,78	-21,51	-11,80	1,59	16,66	0,909
0,800	173,94	-39,77	179,21	-390,68	-7,15	0,69	12,95	5,002	6,69	-82,92	15,97	-23,37	-16,65	2,11	21,38	1,201
1,000	217,50	-49,49	224,16	-488,63	-8,89	0,87	16,13	6,256	6,43	-103,86	17,78	-24,66	-21,57	2,58	26,09	1,495
2,000	435,29	-98,09	448,86	-978,37	-17,61	1,80	32,02	12,524	5,31	-208,55	27,12	-31,62	-46,01	4,95	49,77	2,976
5,000	1088,66	-243,87	1122,99	-2447,59	-43,75	4,57	79,67	31,328	9,67	-521,06	63,67	-70,59	-116,04	12,33	123,23	7,427
10,000	2177,62	-486,85	2246,54	-4896,29	-87,31	9,19	159,09	62,669	16,51	-1041,99	123,93	-134,33	-232,79	24,58	245,67	14,846
15,000	3266,58	-729,82	3370,08	-7345,00	-130,88	13,81	238,52	94,009	23,19	-1562,95	184,05	-197,74	-349,57	36,84	368,11	22,265
20,000	4355,54	-972,80	4493,63	-9793,70	-174,45	18,43	317,94	125,350	29,96	-2083,89	244,25	-261,34	-466,33	49,09	490,55	29,684
30,000	6533,46	-1458,75	6740,72	-14691,10	-261,58	27,66	476,79	188,031	43,51	-3125,78	364,66	-388,56	-699,85	73,60	735,43	44,521

Table 1.4: Evolution de la valeur des 7 paramètres de la transformation de Helmert en fonction de l'amplitude de l'erreur introduite. e_H est définie par 1.42 et 1.43.

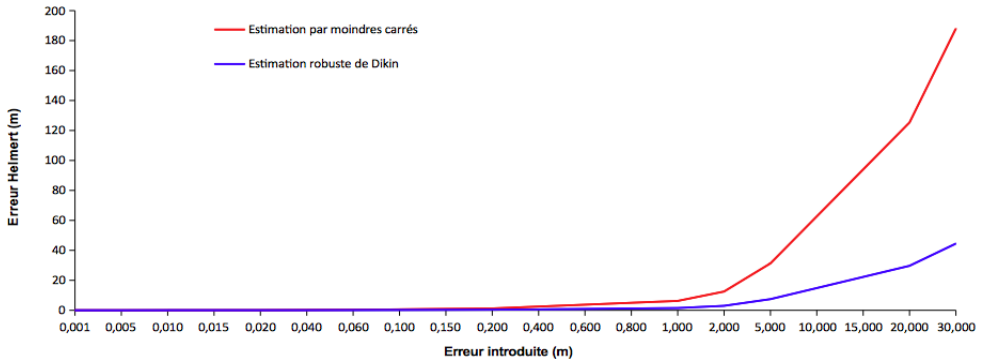


Figure 1.13: Diagramme de l'évolution des erreurs de la transformation en fonction de l'amplitude de l'erreur introduite pour les estimateurs de Dikin et des moindres carrés

Résultats

Le tableau 1.4 et la figure 1.13 montrent que les 7 paramètres de transformation de Helmert estimés par l'estimation robuste de Dikin et par la méthode des moindres carrés sont toujours différents de 0. Bien que l'erreur Helmert de transformation estimée par l'estimateur de Dikin soit plus base que celle par la méthode des moindres carrés, les erreurs Helmerts des deux estimateurs augmentent avec l'augmentation de l'amplitude d'erreur introduite. L'effet d'asymétrie de réseau réduit la précision de transformation estimée dans les 2 estimateurs, particulièrement en présence d'erreurs.

1.6.3.4 Etude de l'influence d'une erreur/plusieurs erreurs accidentelles sur les séries temporelles

Dans les tests précédents, on étude l'influence de l'erreur sur la transformation de Helmert en utilisant deux estimateurs. Dans ce test, on continue à étudier l'influence de l'erreur sur les séries temporelles de position GPS des sites après estimation de la transformation. On cherche à évaluer la possibilité de deux estimateurs pour la détection des erreurs et l'isolement des erreurs dans les séries temporelles.

On utilise le jeu de référence du premier test 1.6.3.1, avec 10 sites distribués aléatoirement à la surface de la Terre (figure 1.8).

Pour le jeu libre, on perturbe ce fichier par des erreurs suivants la loi Normal avec $\sigma = 2mm$ pour 10 jours consécutifs. On obtient un jeu libre qui est la série temporelle de 10 jours successifs (le jeu libre original).

Pour étudier l'influence d'une erreur accidentelle sur les séries temporelles, on introduit successivement une erreur à 1 cm, 10 cm et 1 m sur la composante X du site X000 au 3ème jour du jeu libre. On obtient donc trois jeux libres entachées d'erreur.

Pour étudier l'influence de plusieurs erreurs, on introduit alternativement et simultanément cinq erreurs de 1 cm, cinq erreurs de 10 cm et cinq erreurs de 1m sur les composantes X du site X000, X001, X003, X006, X009 au 3ème jour du jeu libre. Dans ce cas-là, on obtient aussi trois jeux libres erronés.

On regarde ensuite les séries temporelles obtenues pour le site erroné X000 et un site non-erroné du réseau (X004).

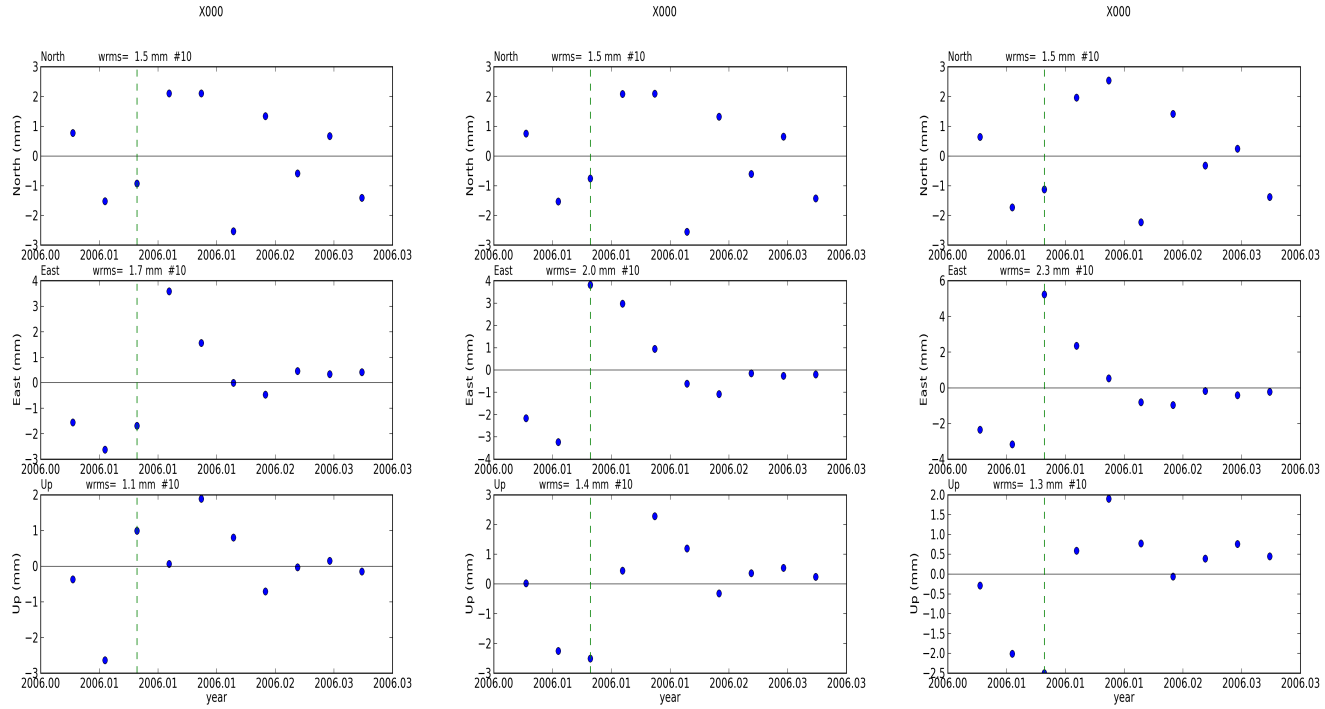
1.6.3.4.1 Résultats pour une erreur On compare les séries temporelles du site X000 (erroné) et du site X004 (non-erroné) pour trois résultats :

- sans l'erreur introduite : les transformations estimées par moindres carrés ;

- avec l'erreur introduite : les transformations estimées par l'estimateur robuste de Dikin ;

- et avec l'erreur introduite : les transformations estimées par moindres carrés.

Les séries temporelles sont présentées sur les figures 1.14, 1.15 et 1.16, la ligne verte verticale corresponds à la date de l'erreur introduite.



Sans erreur introduite

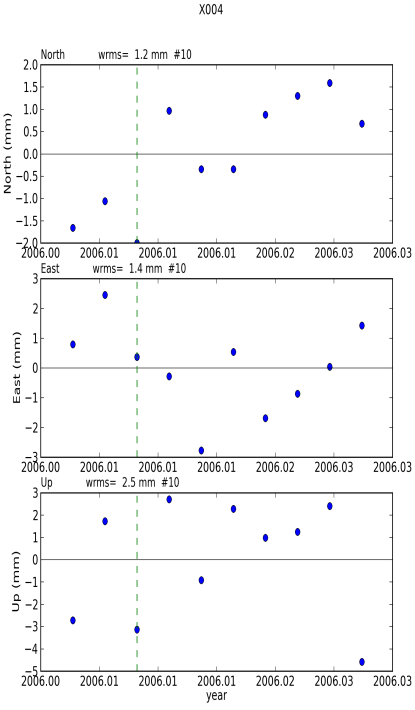
Estimation par moindres carrés

Erreur de 1cm introduite

Estimation par moindres carrés

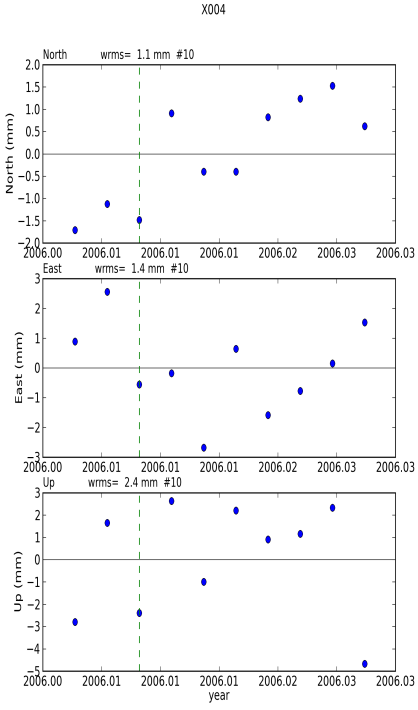
Erreur de 1cm introduite

Estimation robuste de Dikin



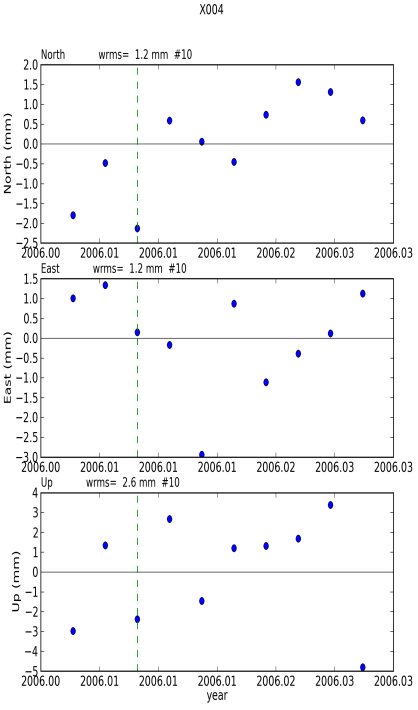
Sans erreur introduite

Estimation par moindres carrés



Erreur de 1cm introduite

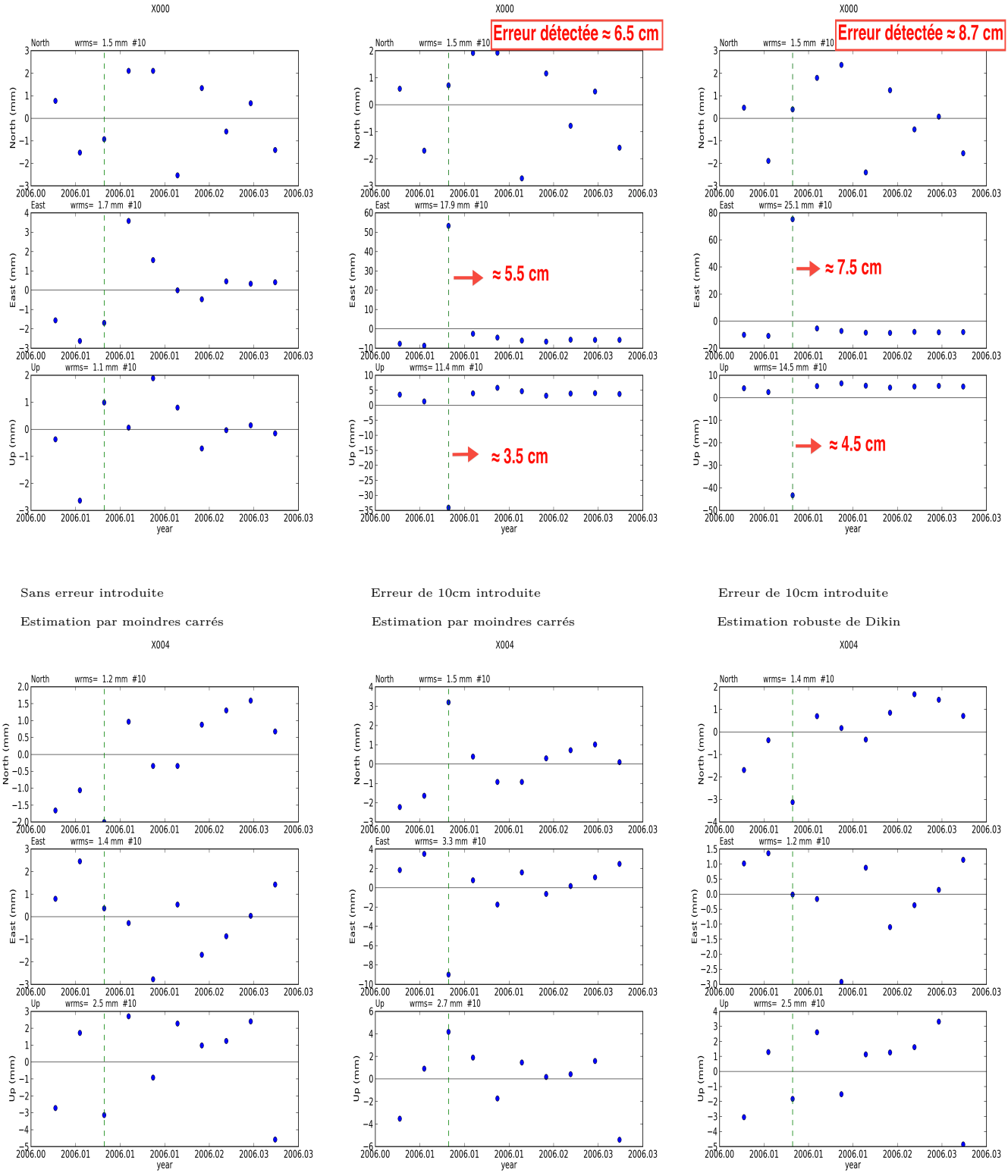
Estimation par moindres carrés



Erreur de 1cm introduite

Estimation robuste de Dikin

Figure 1.14: Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 1cm est introduite sur la composante X du site X000 au 3ème jour (ligne verticale)



Sans erreur introduite

Estimation par moindres carrés

Erreur de 10cm introduite

Estimation par moindres carrés

Erreur de 10cm introduite

Estimation robuste de Dikin

Sans erreur introduite

Estimation par moindres carrés

Erreur de 10cm introduite

Estimation par moindres carrés

Erreur de 10cm introduite

Estimation robuste de Dikin

Figure 1.15: Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 10 cm est introduite sur la composante X du site X000 au 3ème jour (ligne verte verticale). Cette erreur est détectée, calculée et montrée pour le site X000 (carré rouge).

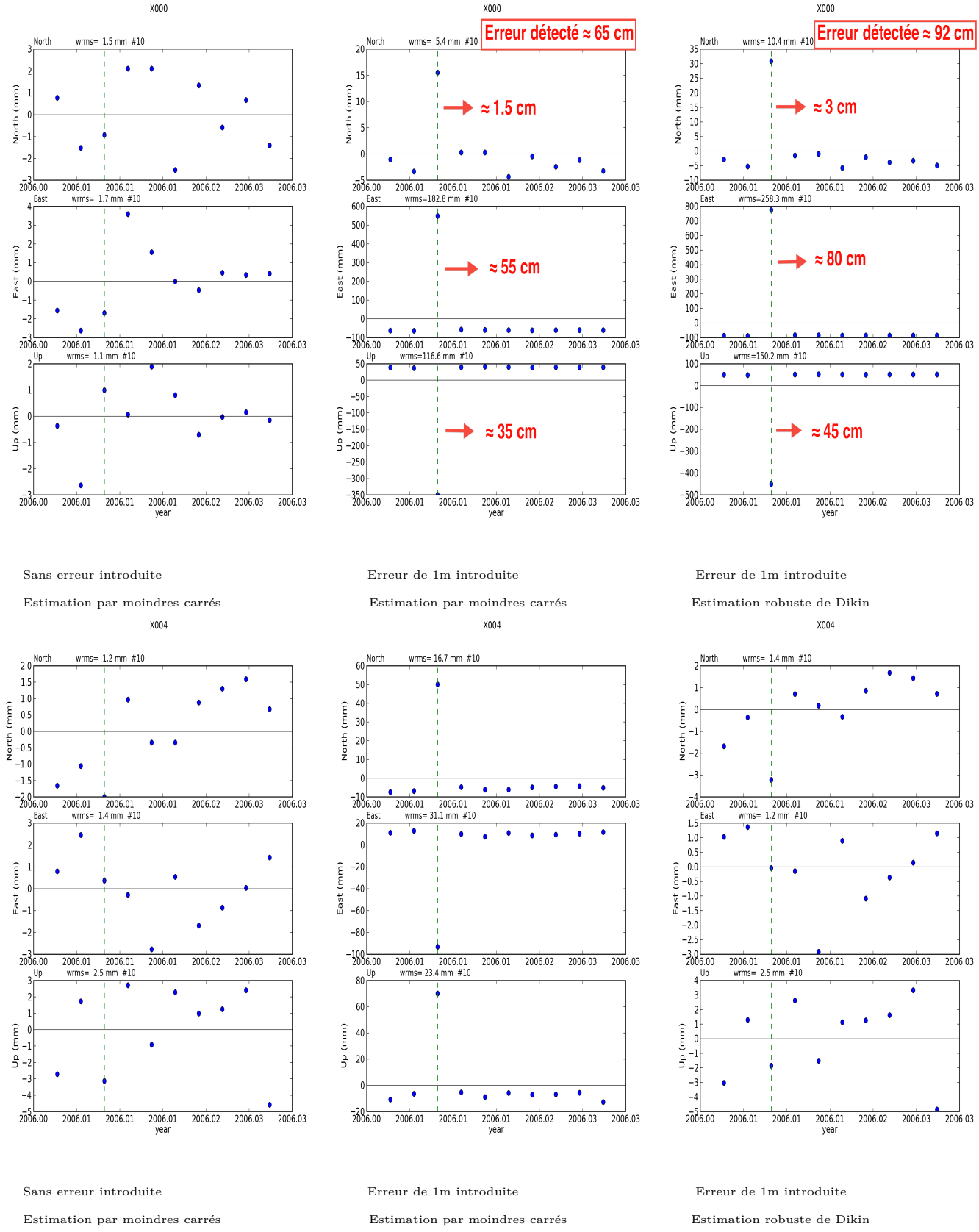


Figure 1.16: Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 1 m est introduite sur la composante X du site X000 au 3ème jour (ligne verte verticale). Cette erreur est détectée, calculée approximativement et montrée pour le site X000 (carré rouge).

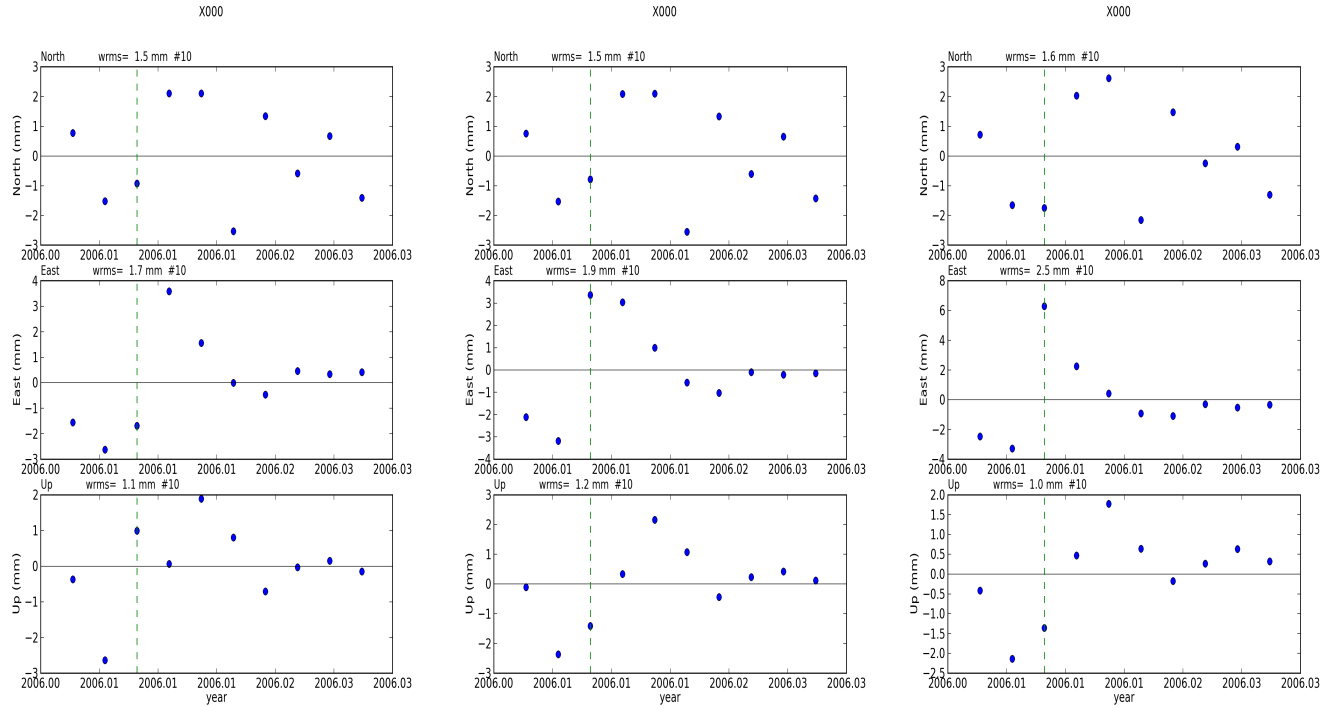
Résultats

Pour une erreur introduite de 1 cm (figure 1.14), on voit que les séries temporelles ne varient pas. Cette erreur sur le site X000 n'impacte presque pas les séries temporelles des autres sites. Cela montre que les deux méthodes sont également insensibles aux petites erreurs.

Pour des erreurs plus grandes (10 cm, figure 1.15 et 1m, figure 1.16), l'estimation robuste de Dikin détecte bien l'erreur sur le site X000 (en carré rouge) dont la valeur approche la valeur de l'erreur introduite. En outre, la série temporelle du site X004 (non-erroné) ne varie presque pas. Cela prouve que l'estimation de Dikin détecte l'erreur et ne la distribue pas aux autres observations. Au contraire, pour l'estimation par moindres carrés, cette erreur n'est pas isolée et impacte les autres sites.

1.6.3.4.2 Résultats pour multiples erreurs Maintenant, on présente les résultats obtenus lorsque le nombre d'erreur augmente par rapport au nombre d'observations total. On introduit une erreur sur la composante X de 5 sites (5/30 observations $\approx 17\%$) X000, X001, X003, X006, X009 au 3ème jour.

On regarde les série temporelles des sites X000 (site erroné) et X004 (site non-erroné) sur les figures 1.17, 1.18 et 1.19.



Sans erreur introduite

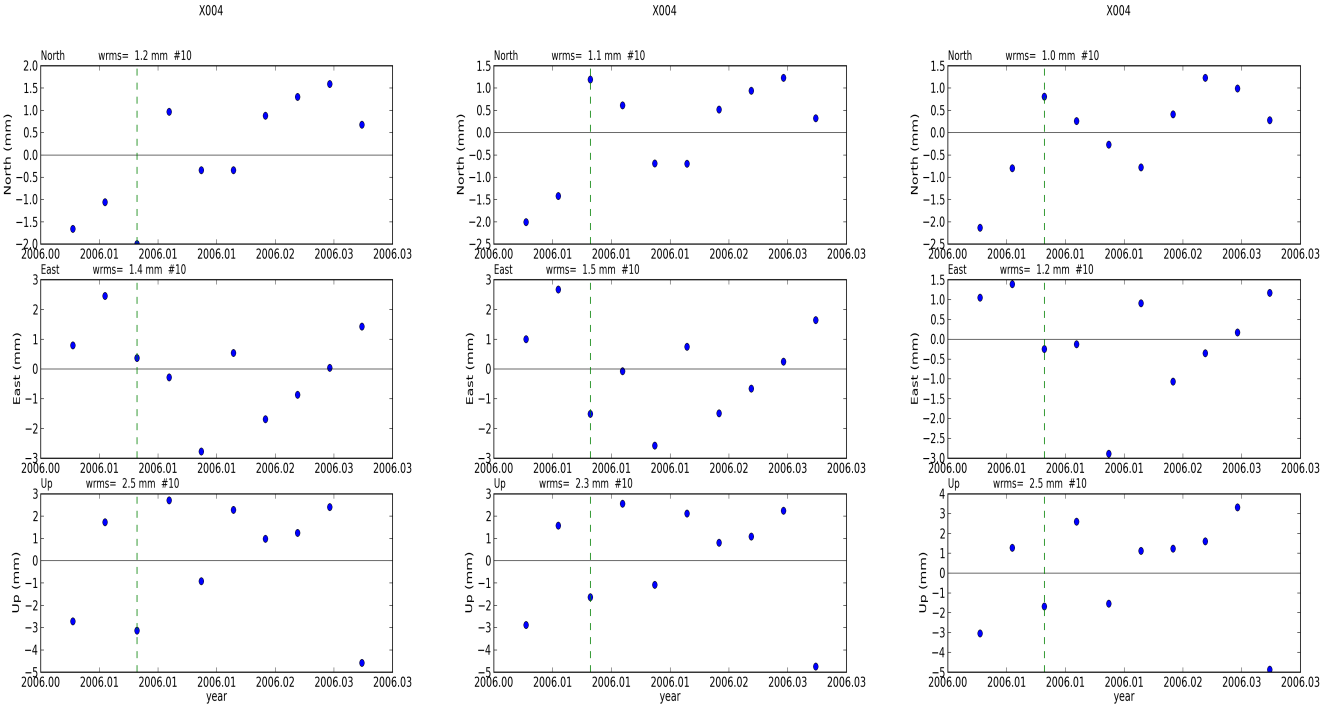
Estimation par moindres carrés

Erreurs de 1cm introduite

Estimation par moindres carrés

Erreurs de 1cm introduite

Estimation robuste de Dikin



Sans erreur introduite

Estimation par moindres carrés

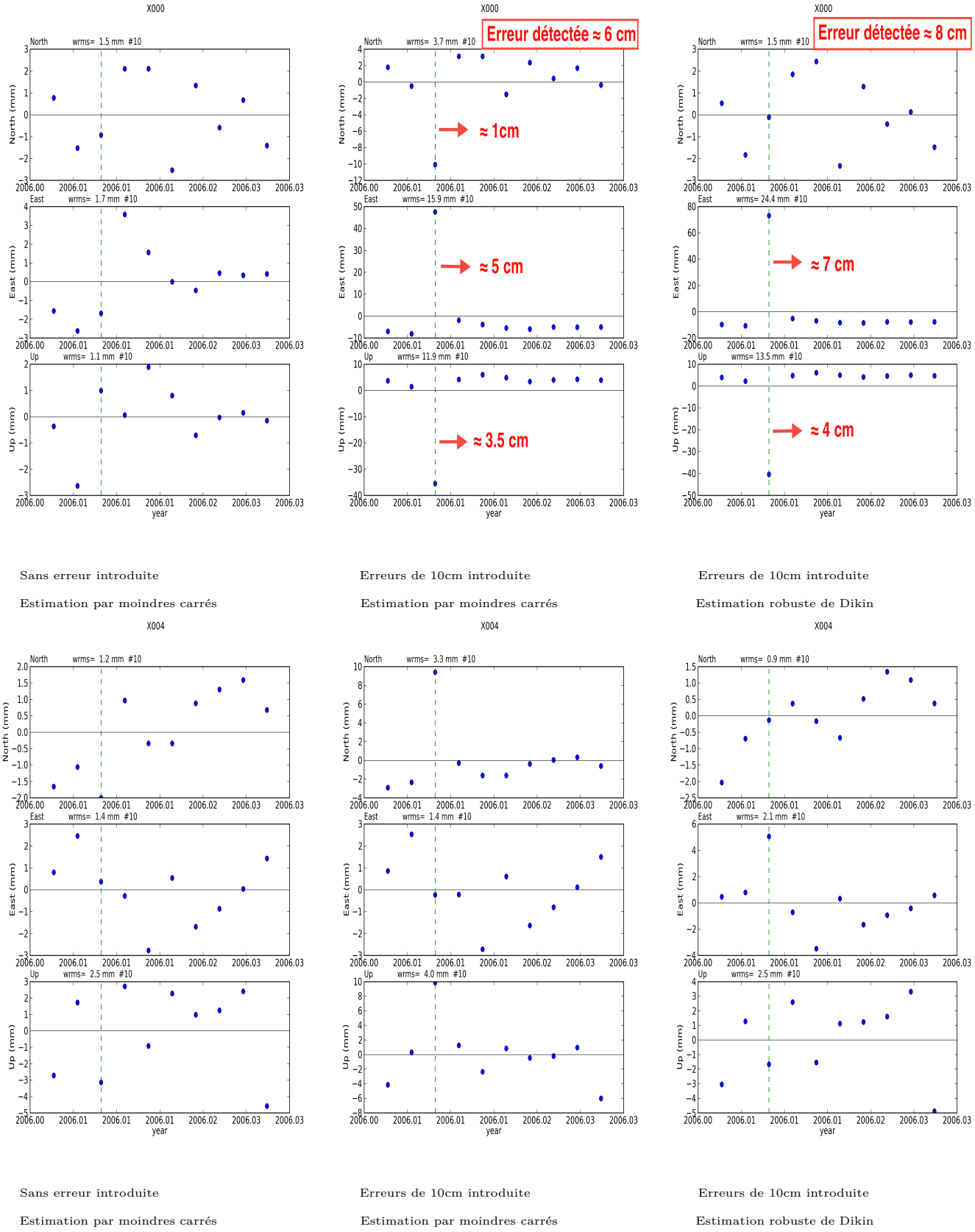
Erreurs de 1cm introduite

Estimation par moindres carrés

Erreurs de 1cm introduite

Estimation robuste de Dikin

Figure 1.17: Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 1 cm est introduite sur la composante X des 5 sites X000, X001, X003, X006 et X009 au 3ème jour (ligne verte verticale).



Sans erreur introduite

Estimation par moindres carrés

Erreurs de 10cm introduite

Estimation par moindres carrés

Erreurs de 10cm introduite

Estimation robuste de Dikin

Sans erreur introduite

Estimation par moindres carrés

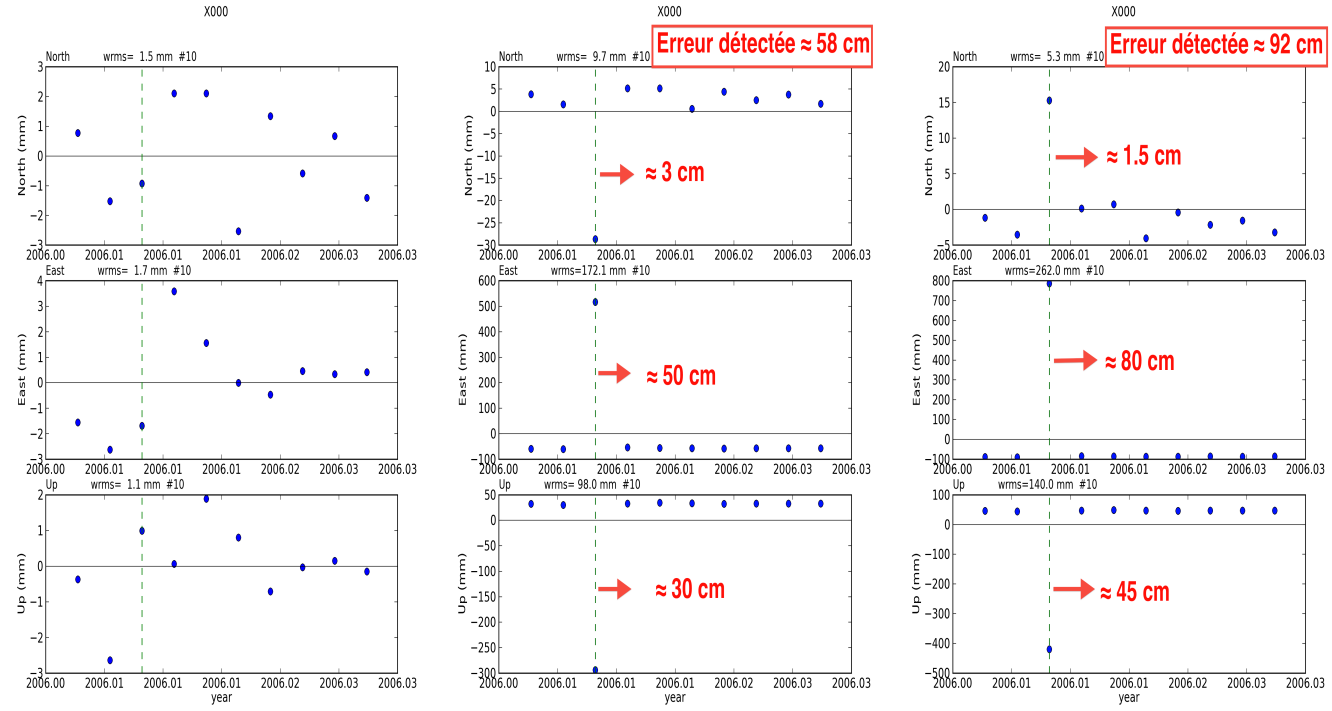
Erreurs de 10cm introduite

Estimation par moindres carrés

Erreurs de 10cm introduite

Estimation robuste de Dikin

Figure 1.18: Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 10 cm est introduite sur la composante X des 5 sites X000, X001, X003, X006 et X009 au 3ème jour (ligne verte verticale). Cette erreur est détectée, calculée approximativement et montrée pour le site X000 (carré rouge).



Sans erreur introduite

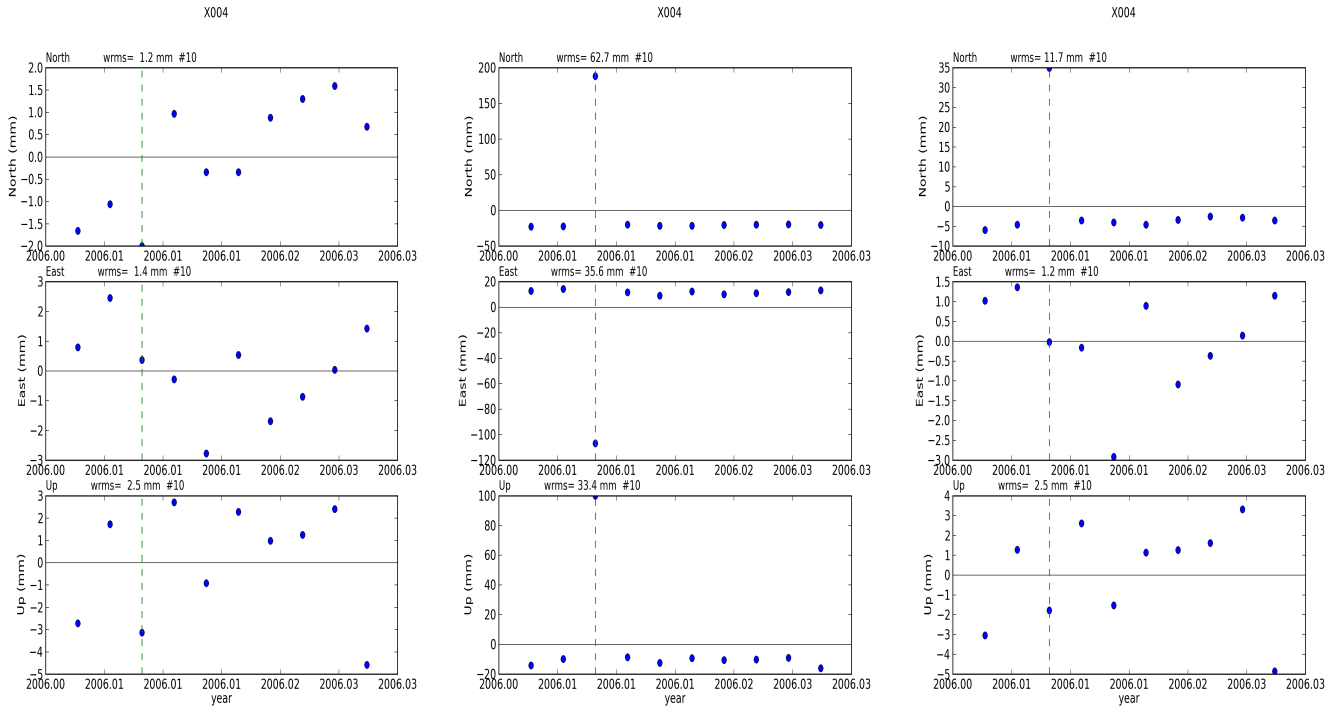
Estimation par moindres carrés

Erreurs de 1m introduite

Estimation par moindres carrés

Erreurs de 1m introduite

Estimation robuste de Dikin



Sans erreur introduite

Estimation par moindres carrés

Erreurs de 1m introduite

Estimation par moindres carrés

Erreurs de 1m introduite

Estimation robuste de Dikin

Figure 1.19: Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 1 m est introduite sur la composante X des 5 sites X000, X001, X003, X006 et X009 au 3ème jour (ligne verte verticale). Cette erreur est détectée, calculée approximativement et montrée pour le site X000 (carré rouge).

Résultats

Sur les figures 1.17, 1.18 et 1.19, on voit que les erreurs introduite perturbent les séries temporelles pour les deux estimations, à partir d'une valeur de 10cm. La série temporelle obtenue par l'estimation de Dikin est meilleure que celle obtenue par la méthode des moindres carrés. En particulier, elle permet encore d'identifier correctement l'origine des erreurs même pour des perturbations de 10cm et 1m.

En conclusion, bien que la méthode Dikin ne rejette pas complètement l'influence de ces 5 erreurs (correspondant à 17% d'observations), elle la diminue appréciablement. Il faudra donc faire attention pour les séries temporelles comportant des erreurs dont les valeurs sont grandes et nombreuses. Toutefois, on peut anticiper que l'approche Dikin sera suffisamment robuste pour résoudre la plupart des cas pratiques que l'on rencontre, où il est rare que plus de 6% des sites communs en estimation de la transformation soient grossièrement faux.

1.6.3.5 Etude des performances de calcul des différents estimateurs

L'estimation robuste de Dikin repose sur un processus itératif plus gourmand en temps de calcul que la méthode des moindres carrés. On cherche donc à évaluer la faisabilité pratique de cette approche.

Dans ce test, on étudie les performances de temps de calcul, dans 2 cas : le nombre de paramètres estimés et le nombre d'observations. Dans les 2 cas, on compare le temps du calcul de l'estimation par moindres carrés et celui par l'estimation robuste de Dikin. L'estimation robuste est appliquée avec un critère de convergence d'itération à 10^{-5} et à 10^{-3} pour évaluer sa sensibilité au seuil de convergence.

1.6.3.5.1 Nombre de transformations estimées Le jeu de données est fabriqué similairement au test 1.6.3.1, mais le nombre de sites communs est de 20.

On estime les paramètres de la transformation de Helmert quand le nombre de fichiers du jeu libre, c'est-à-dire le nombre des transformations, augmente peu à peu jusqu'à 10000 fichiers. On regarde la diagramme du temps de calcul (figure 1.20) est présenté sur un Macintosh avec un processeur cadencé à 2.66 Ghz.

Sur la figure 1.20, la ligne en rouge est le temps de calcul pour l'estimation par moindres carrés, en bleu est le temps de calcul pour l'estimation robuste de Dikin avec un critère de convergence d'itération à 10^{-3} et en vert est celui à 10^{-5}. Le temps du calcul par la méthode des moindres carrés est toujours beaucoup plus faible que celui par l'estimation de Dikin, particulièrement quand le nombre de transformations est grand. En comparaison

entre la même estimation de Dikin, on voit que celle avec un critère de convergence d'itération à 10^{-3} calcule plus rapidement (environ 1,5 fois) que celle à 10^{-5}.

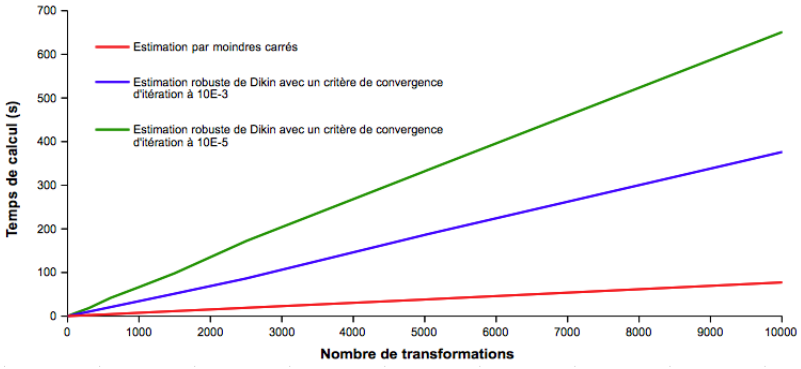


Figure 1.20: Evolution du temps de calcul des estimateurs suivant l'augmentation du nombre de transformation de Helmert (nombre de jours)

1.6.3.5.2 Nombre d'observations Dans ce test, le nombre de sites communs entre le jeu de référence et le jeu libre augmente pas à pas jusqu'à 600 (18000 observations). L'évolution du temps de calcul est présenté dans la figure 1.21.

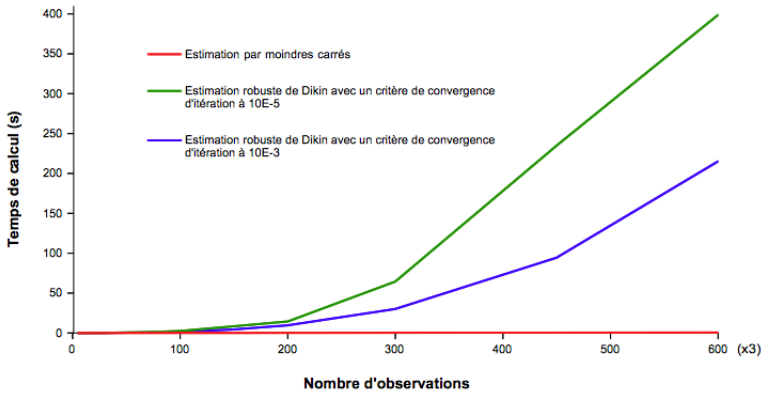


Figure 1.21: Evolution du temps de calcul de l'estimateur robuste de Dikin en fonction du nombre d'observations

La vitesse du calcul par 3 estimations (par moindres carrés, de Dikin avec un critère de convergence d'itération à 10^{-3} et de Dikin avec un critère de convergence d'itération à 10^{-5}) n'est pas très différente quand le nombre d'observations est inférieur à 300 (100 sites communs). Par contre, le temps de calcul de l'estimation de Dikin augmente très rapidement suivant l'augmentation du nombre d'observations quand le nombre est supérieur à 300. Par ailleurs, l'estimation de Dikin avec un critère de convergence d'itération à 10^{-3} se calcule plus rapidement (1 à 2 fois) que celle de 10^{-5}.

1.6.3.6 Validation de la stratégie finale

Notre stratégie préférée consiste à utiliser successivement l'approche par estimation de Dikin pour détecter les outliers et les retirer de l'estimation, puis de réaliser l'estimation finale par moindres carrés sur les données validées.

Dans ce test, on étudie la précision de transformation avec notre stratégie. On utilise le jeu des données du premier test 1.6.3.1.

Tout d'abord l'estimateur robuste de Dikin est appliqué pour estimer les 7 paramètres de Helmert. On détecte ensuite des erreurs grossières (outliers) dans l'ensemble de sites communs suivant les 3 composantes

$$\begin{aligned} |v_i^e| &\geq 3 * median(|v^e|) \\ |v_i^n| &\geq 3 * median(|v^n|) \\ |v_i^u| &\geq 3 * median(|v^u|) \end{aligned} \tag{1.44}$$

Ainsi, on tire profit de la formulation de la transformation de Helmert en coordonnées locales : celle-ci nous permet de choisir la composante horizontale si, seule, la composante verticale est erronée, sans avoir à rejeter les deux composantes. Si la composante verticale est un outlier, on la retire seulement et retient les composantes Est, Nord pour l'estimation finale. Par ailleurs, si la composante Est ou Nord est outlier, on rejette les trois composantes.

Finalement, on estime les 7 paramètres de la transformation de Helmert par moindres carrés avec les données validées.

Le tableau 1.5 et la figure 1.22 représentent les résultats et leur comparaison avec les résultats de l'estimation robuste de Dikin du premier test 1.6.3.1.

Erreur intr-	Estimation robuste de Dikin								Estimation robuste de Dikin – Estimation par moindres carrés							
-oduite (m)	Tx(cm)	Ty(cm)	Tz(cm)	D(ppb)	Rx(mas)	Ry(mas)	Rz(mas)	e_H(m)	Tx(cm)	Ty(cm)	Tz(cm)	D(ppb)	Rx(mas)	Ry(mas)	Rz(mas)	e_H(m)
0,001	0,01	0,01	-0,01	-0,03	0,02	0,01	0,01	0,0008	0,03	0,01	0,01	-0,07	0,01	0,00	0,00	0,0006
0,005	-0,04	0,02	-0,01	-0,04	0,01	0,01	-0,01	0,0007	-0,01	0,01	0,01	-0,04	0,01	0,00	-0,02	0,0008
0,010	-0,06	-0,01	-0,01	0,01	0,02	0,00	-0,02	0,0009	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
0,015	-0,07	-0,02	0,00	0,03	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
0,020	-0,07	-0,02	-0,01	0,04	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
0,040	-0,07	-0,02	-0,01	0,04	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
0,060	-0,07	-0,03	-0,01	0,04	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
0,100	-0,07	-0,02	-0,02	0,04	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
0,150	-0,06	-0,02	-0,02	0,04	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
0,200	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
0,400	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
0,600	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
0,800	-0,08	-0,01	-0,02	0,02	0,02	0,00	-0,02	0,0009	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
1,000	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
2,000	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
5,000	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
10,000	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
15,000	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
20,000	-0,07	-0,02	-0,02	0,02	0,02	0,00	-0,02	0,0009	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005
30,000	-0,07	-0,02	-0,02	0,03	0,02	0,00	-0,02	0,0010	0,02	-0,02	0,03	-0,01	0,01	0,00	0,00	0,0005

Table 1.5: Valeurs estimées des 7 paramètres de la transformation de Helmert en fonction de l'amplitude de l'erreur introduite. e_H est définie par 1.42.

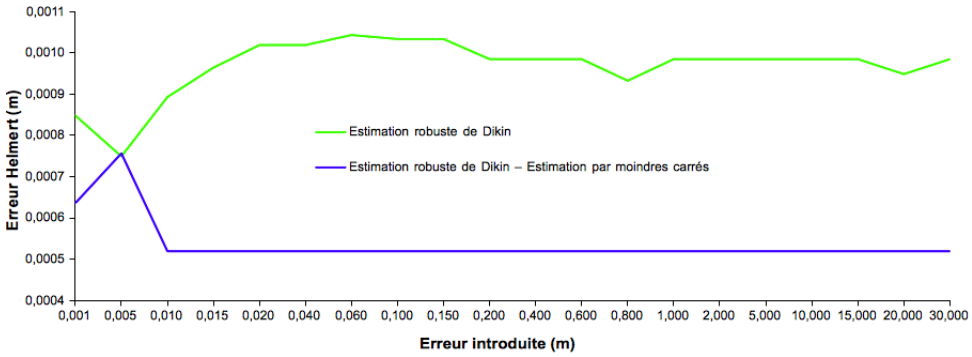


Figure 1.22: Diagramme de l'évolution des erreurs Helmert des 7 paramètres en fonction de l'amplitude de l'erreur introduite

Résultats

On voir que cette stratégie (la détection des outlier par l'estimation robuste de Dikin et l'estimation finale par moindres carrés sur les données validées) est bonne et finalement peu coûteuse en temps du calcul. Elle permet d'obtenir les meilleurs résultats avec une précision plus élevée et est finalement plus stable que l'estimation robuste de Dikin seule.

1.6.4 Conclusions

Les tests réalisés nous permettent de proposer une stratégie optimale, robuste et rapide pour l'estimation des 7 paramètres de la transformation de Helmert :

- On utilise une solution de référence de grande qualité et à jour : la solution IGS la plus actuelle

- On utilise une algorithme rapide utilisant la norme L1 pour la détection d'outlier. L'utilisation la transformation de Helmert en coordonnée locale permet de retirer séparément les outliers présents sur la composante verticale.

- On réalise une estimation finale par moindres carrés sur données validées.

Nos tests ont aussi mis en évidence des limitations de notre approche, en particulier :

- Les deux estimateurs, l'estimateur par moindres carrés et l'estimation robuste de Dikin, sont influencés par l'effet d'asymétrie de réseau.

- L'estimateur robuste de Dikin ne détecte bien pas toujours plusieurs erreurs dans les observations (ici, on a testé avec 5 erreurs sur 30 observations, c'est-à-dire 17% d'observation).

Cependant, notre approche permet de gagner en robustesse et en temps par rapport aux approches proposées en défaut dans le logiciel comme Globk.

1.7 Implémentation: le programme pyacs_make_time_series.py

1.7.1 Le framework PYACS

PYACS (Yet Another Combination Software) est un ensemble de programmes développés dans le cadre de ma thèse, par moi et le directeur de thèse Jean-Mathieu NOCQUET. Il est écrit en langage Python. Il consiste des programmes permettant des analyses géodésiques et une bibliothèque de programmation objet.

pyacs_make_time_series.py, un programme du pyacs, implémente les méthodes décrites précédemment. Il permet d'estimer rapidement et de manière robuste une transformation de Helmert entre la solution libre journalière et les positions prédites le même jour par une solution de référence en positions et vitesses. La solution journalière est produite par Gamit 10.04 (Herring et al. [2010b], Herring et al. [2010a], Herring et al. [2010c]).

De plus, les autres programmes développés dans un "framework" PYACS permettent d'un calcul GPS sur Cluster, d'une modélisation cinématique en blocs, tenseurs de déformation et de la modélisation élastique inter- et co-sismique.

1.7.2 Le langage Python et ses bibliothèques scientifiques

Le langage de programmation Python est développé depuis 1989 par Guido Van Rossum et de nombreux contributeurs bénévoles. Il est considéré comme un langage dynamique,

extensible, et puissant en calcul. Il permet (sans l'imposer) d'une approche modulaire et orientée objet de la programmation. Il fonctionne sur la plupart des plates-formes informatiques, des supercalculateurs aux ordinateurs centraux, de Windows à Unix en passant par Linux, Mac OS, ou encore Android, et aussi avec Java ou encore .NET. Il est distribué sous forme de logiciel libre sur le lien http://www.python.org/.

Python est un langage ouvert avec une communauté de développeurs nombreuse. Python permet de profiter des bibliothèques scientifiques par exemple NumPy et SciPy pour le calcul matriciel, Maplotlib et Pydot pour la visualisation.

En programmation du pyacs nous utilisons principalement les bibliothèques NumPy, SciPy pour les aspects d'algèbres et Maplotlib pour la visualisation.

NumPy (Numeric Python, http://www.numpy.org/) fournit des routines de base pour manipuler des matrices.

SciPy (Scientifique Python, http://www.scipy.org/) étend les fonctionnalités de NumPy avec une collection d'algorithmes utiles, comme la minimisation, la transformation de Fourier, de régression et d'autres techniques de mathématiques appliquées. Nous utilisons en particulier ses fonctionnalité de factorisation par la méthode de Cholesky pour l'estimation par moindres carrés.

Enfin, Matplotlib (http://matplotlib.org/) est une bibliothèque graphique qui permet de tracer une variété importante de graphiques. Un point important est que l'on peut utiliser matplotlib par l'intermédiaire d'un GUI (Graphic User Interface) ou en ligne de commande

1.7.3 Programmation objet

Nous avons développé une bibliothèque de classes pour le framework pyacs, qui contient :

- Sinex.py : ce module lit une solution SINEX (Solution INdependent EXchange Format, ftp://ftp.kiam1.rssi.ru/pub/gps/lib/formats/sinex05.txt). Il permet aussi de manipuler la solution SINEX, par exemple trouver les sites communs entre la solution de référence et celle libre, détecter les sites communs erronés dans la solution libre.

- Discontinuities.py : ce module lit un fichier de discontinuités de la solution IGS.

- AstroTime.py : ce module contient un ensemble de routines pour les conversions de la base de temps, telles que la conversion entre des dates de calendrier et jour Julien, la conversion de UT à temps sidéral local,...

- Coordinates.py : ce module assure les conversions de coordonnées, permettant par exemple de transformer les coordonnées géocentriques en coordonnées géodésiques, ou de passer des coordonnées générales en coordonnées locales.

- Gpoint.py : ce module génère les paramètres du point, par exemple : les coordonnées et la matrice variance-covariance, la vitesse,...

- Helmert.py : ce module estime une transformation de Helmert. Il crée les équations de la transformation à partir de points communs, réalise l'estimation robuste de Dikin pour la détection des outliers et l'estimation finale par moindres carrés pour les 7 paramètres de la transformation.

- RobustEstimators.py : une classe d'estimateur, incluant l'estimation robuste de Dikin, que nous avons implémentée.

1.7.4 Structure de calcul du programme pyacs_make_time_series.py

Les données en entrée du programme sont la solution de référence et les solutions libres. Pour la transformation journalière en l'ITRF, on prend la solution libre journalière générée par Gamit transformée en format SINEX, et la solution de référence IGS en format SINEX ou en format apr de Globk.

La figure 1.23 présente la structure du calcul de pyacs_make_time_series.py.

Lors de l'estimation des paramètres de transformation, pyacs_make_time_series.py inspecte tout d'abord tous les sites communs entre la solution de référence et la solution libre pour détecter les erreurs grossières (ce sont les sites dont la distance entre la position de la solution libre P_i^{lib} et celle de la solution de référence P_i^{ref} est plus de 10m)

$$|P_i^{ref} - P_i^{lib}| > 10m \tag{1.45}$$

Simultanément il vérifie le fichier de discontinuités de IGS qui est mis à jour hebdomadairement et disponible sur `ftp://igs-rf.ign.fr/pub/discontinuities/soln.snx`.

Il passe ensuite des coordonnées globales (x, y, z) aux coordonnées locales (e, n, u).

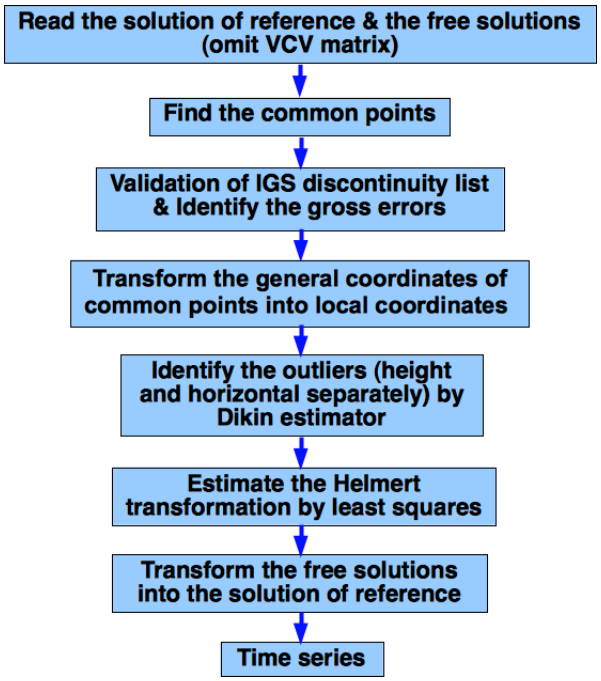


FIGURE 1.23: Schéma de calcul du programme pyac_make_time_series.py

Puis il estime les 7 paramètres de la transformation par l'estimateur robuste Dikin pour identifier les composantes des points qui dégradent la qualité de l'estimation. Dans cette étape, les "outliers" sont les composantes qui correspondent à ses résidus plus grands que 3 fois de la médiane des résidus (formule 1.44).

Grâce à l'équation de transformation en coordonnées locales, le rejet des outliers est effectué séparément sur les deux composantes horizontales et la composante verticale. On peut donc retenir les deux composantes horizontales, si la composante verticale est outlier. Par contre, si une de deux composantes horizontales est outlier, il rejette toutes les 3 composantes du point du calcul.

A l'étape suivante, il estime finalement les 7 paramètres de la transformation de Helmert par moindres carrés à partir des données validées.

Enfin il transmet les solutions libres à la solution de référence, on obtient donc les séries temporelles de positions dans l'ITRF en coordonnées géocentriques (x, y, z) et les séries temporelles résiduelles en coordonnées locales (e, n, u) sous le format mb_file de Gamit (http://www-gpsg.mit.edu/~tah/GGMatlab/).

1.7.5 Limitations de l'estimation robuste et rapide

Dans notre approche, on néglige les matrices variance-covariances à la fois de la solution référence et des solutions libres dans un souci à la fois de rapidité du développement et

de vitesse de calcul.

En effet, la taille de matrice variance-covariance est égale m$_x$m (m est nombre d'observations) et augmente suivant le carré du nombre d'observations. La lecture, la manipulation et l'inversion de la matrice sont donc très longs et prennent plus de mémoire. Tous les réseaux GPS que l'on étudie sont permanents avec des centaines de stations de plusieurs années des données à des dizaines d'années des données. La matrice variance-covariance rend le calcul lent.

Négliger la matrice variance-covariance revient à négliger les poids relatifs entre stations et la corrélation entre stations. En effet, on constate en pratique que, pour des réseaux permanents et des stations calculées par sessions de 24h, l'écart-type est sensiblement le même entre les stations et stable d'un jour sur l'autre. De plus, la corrélation dans la solution libre est "manquée" par la faible définition du repère de référence. L'impact au final est assez faible.

1.7.6 Exemples pratiques

On présente quelques résultats de pyacs_make_time_series.py. Des analyses plus complètes sont aussi présentées dans le chapitre 4.

Nous avons estimé 2 réseaux GPS permanents qui sont la solution Géoazur I du réseau RENAG et le réseau GPS Taiwain par le programme pyacs_make_time_series.py sur un iMac avec un processeur Intel Core 2 Duo cadencé à 2.66 Ghz. Les calculs sont très rapides, concrètement les temps d'estimations ne sont que de quelques minutes (voir le tableau 1.6).

Réseau	No de sites	Période	Temps d'estimation
Geoazur I - RENAG	102	2006.0 - 2010.8	3 mn
Taiwan GPS Network	424	1994.8 - 2011.6	9 mn

TABLE 1.6: Temps de calcul de la réalisation des séries temporelles par pyacs_make_time_series.py pour les réseaux RENAG et Taiwan

En estimation des paramètres de la transformation de l'époque 2006.001 de la solution Géoazur I, le programme a trouvé 34 points communs. Lors que l'estimation, la validation des outliers par l'estimation robuste de Dikin a rejeté 6 points et la composante verticale du point PENC. Il a donc utilisé 83 observations "nettoyées" de 28 points communs en estimation finale par moindres carrés. Les résidus pour deux estimateurs sont présentées sur le tableau 1.7.

Epoch: 2006.00136939 #File sinex: h0601011200_alp1.snx.ss

No	Points	RMSe(mm)	RMSn(mm)	RMSu(mm)	3D(or 2D)(mm)	Outlier_rejected
with outliers. Dikin estimation						
1	ABER	-0.8	-1.3	0.4	1.6	
2	ALAC	-3.9	-1.2	3.9	5.7	
3	BOGO	-1.1	-0.6	3.1	3.4	
4	BRST	-3.4	-8.9	12.1	15.4	East,North,Up
5	BRUS	1.5	-1.1	-4.3	4.7	
6	BZRG	-0.2	0.4	0.4	0.6	
7	CAGL	-0.4	1.8	-1.2	2.2	
8	CHIZ	3.5	1.6	-1.4	4.1	
9	EBRE	-5.2	5.0	8.8	11.4	East,North,Up
10	GENO	-3.6	-0.1	-1.3	3.9	
11	GRAS	1.9	1.3	-4.9	5.4	
12	GRAZ	16.5	5.6	-9.4	19.8	East,North,Up
13	JOZE	0.9	4.7	58.4	58.6	East,North,Up
14	KOSG	0.0	-0.3	0.1	0.4	
15	LAMP	-0.8	1.2	0.3	1.5	
16	LROC	2.0	-0.4	-1.2	2.4	
17	MALL	5.1	-6.1	-1.2	8.1	East,North,Up
18	MARS	0.8	0.9	-4.8	4.9	
19	MATE	-1.1	-0.5	4.1	4.3	
20	MEDI	5.9	4.0	1.4	7.3	East,North,Up
21	NOT1	-1.9	-0.7	0.7	2.1	
22	ONSA	-0.3	-1.2	-3.1	3.3	
23	ORID	1.8	1.1	2.3	3.1	
24	PENC	5.2	-3.8	-19.2	20.3	Up
25	POTS	0.3	-0.7	0.2	0.8	
26	SFER	-1.8	0.1	-2.5	3.1	
27	SJDV	-0.7	1.3	-2.6	3.0	
28	SOFI	2.2	0.1	-3.5	4.1	
29	TLSE	-3.0	0.5	0.4	3.1	
30	TORI	-4.5	-0.5	1.1	4.7	
31	VILL	3.1	-0.9	1.4	3.5	
32	WSRT	0.4	-0.7	1.8	2.0	
33	WTZR	0.4	-1.1	-1.0	1.5	
34	ZIMM	-2.9	0.4	2.4	3.8	

No	Points	RMSe(mm)	RMSn(mm)	RMSu(mm)	3D(or 2D)(mm)
without outliers. Least square estimation					
1	ABER	-1.3	-1.5	1.5	2.4
2	ALAC	-3.5	-1.3	4.5	5.8
3	BOGO	-1.1	-0.0	3.2	3.4
4	BRUS	1.3	-1.0	-3.7	4.1
5	BZRG	-0.0	0.6	0.6	0.9
6	CAGL	0.0	1.8	-1.2	2.2
7	CHIZ	3.5	1.5	-0.7	3.9
8	GENO	-3.4	0.0	-1.0	3.6
9	GRAS	2.1	1.4	-4.6	5.2
10	KOSG	-0.2	-0.2	0.7	0.8
11	LAMP	-0.2	1.3	0.1	1.3
12	LROC	2.1	-0.6	-0.5	2.2
13	MARS	1.0	0.9	-4.4	4.6
14	MATE	-0.6	-0.2	3.9	3.9
15	NOT1	-1.3	-0.5	0.4	1.5
16	ONSA	-0.7	-0.9	-2.5	2.7
17	ORID	2.3	1.5	1.9	3.4
18	PENC	5.4	-3.4		6.3
19	POTS	0.2	-0.4	0.5	0.7
20	SFER	-1.4	-0.1	-1.7	2.2
21	SJDV	-0.6	1.4	-2.1	2.6
22	SOFI	2.7	0.6	-3.9	4.8
23	TLSE	-2.9	0.4	1.0	3.1
24	TORI	-4.4	-0.4	1.5	4.6
25	VILL	3.4	-1.2	2.1	4.2
26	WSRT	0.2	-0.5	2.4	2.5
27	WTZR	0.4	-0.8	-0.7	1.2
28	ZIMM	-2.8	0.5	2.8	4.0

3D rms: 2.0 mm
mean rms: East: 1.3 mm; North: 0.7 mm; Up: 2.4 mm
rms max: 6.3 mm at PENC
rms min: 0.7 mm at POTS
Number of reference sites: 28
Number of outliers: 19 over 80 (23.8%)

TABLE 1.7: Résidus des points communs en estimation des 7 paramètres de la transformation (époque 2006.001, Géoazur I)

On montre sur la figure 1.24 ensuite la série temporelle brute du site IESG du réseau Géoazur I directement dérivée par pyacs_make_time_series.py.

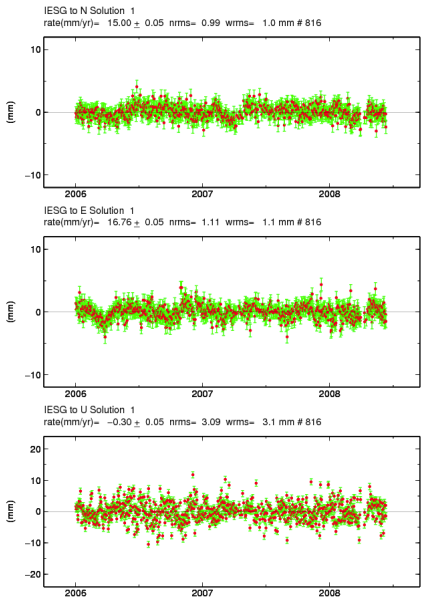


FIGURE 1.24: Série temporelle brute (pente retirée) du site IESG (Géoazur I) obtenue par pyacs_make_time_series.py

1.8 Conclusions

En réalisant les séries temporelles dans l'ITRF, la solution IGS la plus actuelle nous donne l'accès à une solution de référence de haute qualité, mise à jour et meilleure que la solution de l'ITRF courante. De plus, on peut prendre la liste de discontinuités IGS pour limiter les erreurs (outliers) en estimation des paramètres de la transformation.

La transformation de Helmert est exprimée en coordonnées locales, ce qui nous permet d'isoler la composante verticale et la retirer de l'estimation si une erreur est détectée.

L'estimation robuste et rapide des paramètres de la transformation de Helmert est obtenue en (1) négligeant la matrice variance-covariance, (2) détectant des erreurs (outliers) suivant la norme L1 par l'estimation robuste de Dikin, et (3) estimant les finalement par moindres carrés sur données validées.

Notre stratégie est appliquée dans le programme pyacs_make_time_series.py pour estimer rapidement et robustement les paramètres de la transformation de Helmert et réaliser les séries temporelles de positions GPS dans l'ITRF.

Cette approche permet de générer des séries temporelles de haute qualité que nous analysons ensuite, avec les méthodes présentées dans le chapitre suivant.

CHAPTER 2

ANALYSES DES SÉRIES TEMPORELLES DE POSITION GPS

2.1 Introduction

L'analyse des séries temporelles est un domaine complexe qui fait l'objet de recherches actives, non seulement en géodésie/géophysique mais dans toutes les sciences observationnelles et en sciences sociales. Il était impossible dans le cadre d'une thèse d'avoir une idée précise de l'ensemble des méthodes d'analyse des séries temporelles existant dans les différentes disciplines intéressées par ce domaine. Nous avons plutôt choisi de définir les spécificités des séries temporelles GPS et de proposer des méthodes pour résoudre les problèmes les plus communément rencontrés.

Les informations généralement extraites des séries temporelles GPS sont des applications les plus classiques en tectonique :

- La détermination de la vitesse

Il s'agit du plus grand nombre d'applications. C'est ce que l'on souhaite obtenir pour les applications tectoniques classiques, l'étude du rebond post-glaciaire, l'étude du couplage inter-sismique [Segall and Davis, 1997].

- La détermination des termes saisonniers

Les variations saisonnières de la position verticale de stations GPS sont utiles pour l'étude des effets de chargement de masse (atmosphérique, océanique, ou hydrologique) et pour la réalisation du repère de référence terrestre [Nicolas et al., 2006]. Suivant Dong et al. [2002], deux avantages de l'étude des termes saisonniers sont : (1) pour mieux comprendre les champs de déformation saisonnière causée par les contributeurs géophysiques et tectoniques, et (2) pour identifier le spectre d'erreur du GPS en cours d'analyse aux échelles saisonnières. En outre, la détermination des termes saisonniers est nécessaire pour améliorer la précision de l'estimation de la vitesse.

- La détermination des mouvements co-sismiques, post-sismiques

Les mouvements co-sismiques, post-sismiques sont nécessaire pour étudier les séismes et leur dynamiques. Par exemple, la détermination des déplacements co-sismiques peut être inversée afin de déterminer la distribution de rupture du séisme [Segall and Davis, 1997].

- La détermination des mouvements transitoires

Les mouvements transitoires nous permettent de contraindre la nature des déformations crustales et potentiellement d'obtenir des aperçus sur les risques naturels [Ji and Herring, 2013]. La connaissance précise des évènements de glissements lents, un des mouvements transitoires, permet de mieux comprendre du processus de chargement et de déchargement sur les failles [Vergnolle et al., 2010].

Plus rarement, on pourra souhaiter déterminer :

- L'accélération

Jiang et al. [2010] étudie le mouvement vertical des marges rocheuses du Groenland, l'Islande et le Svalbard à partir des accélérations verticales pour déterminer la perte en glace de ces zones.

- Le changement de vitesses long terme

Ce résultat est extrêmement rare, mais a été proposé pour une zone des USA par Davis et al. [2006].

L'extraction de chaque signal doit surmonter ou minimiser des problèmes particuliers, et donc suivre une stratégie propre que nous essaierons de définir par la suite.

Un autre point spécifique aux séries temporelles GPS, est qu'un nombre limité de problèmes a été rencontré dans l'ensemble des réseaux traités. Nous pouvons donc les lister et tester le comportement des algorithmes sur ces cas spécifiques. Nous les listons avec une description dans le premier paragraphe de ce chapitre.

Par ailleurs, le développement des réseaux GPS et autres GNSS continus, comprenant des centaines de stations, rend difficile de regarder attentivement chaque série temporelle individuelle et rend nécessaire l'analyse automatique pour les grands ensembles de données. Par exemple, les offsets et les outliers trouvés dans les séries temporelles de positions journalières, que ce soit en raison du changement d'équipement ou de processus géophysiques, conduisent à un biais de l'estimation de vitesse de station GPS si elle n'est pas correctement identifiée ou traitée dans l'analyse. L'identification visuelle puis la correction manuelle de ces erreurs est pratiquement impossible. L'objectif du travail que j'ai réalisé et que je présente dans cette partie est l'amélioration de la robustesse et de la précision des séries temporelles de positions GPS. Nous avons vu dans le chapitre 1, une stratégie pour obtenir une implémentation robuste du système de référence, ce qui constitue la première étape nécessaire pour réaliser des " bonnes " séries temporelles. Ce chapitre présente

l'étape suivante : il décrit les méthodes qui identifient et estiment automatiquement, semi-automatiquement les outliers, les offsets et les déformations post-sismiques dans les séries temporelles. Ces méthodes sont implémentées dans le programme pygps_time_series.py qui enveloppe " wrap " des fonctions individuelles que je vais décrire.

2.2 Les séries temporelles de positions GPS

2.2.1 Les séries temporelles de position GPS

Dans sa définition la plus générale, une série temporelle est une suite de valeurs numériques représentant l'évolution d'une quantité spécifique au cours du temps. De telles suites de valeurs peuvent être exprimées mathématiquement afin d'en analyser le comportement, généralement pour comprendre son évolution passée et pour éventuellement en prévoir le comportement futur. Une telle transposition mathématique utilise le plus souvent des concepts de probabilités et de statistique.

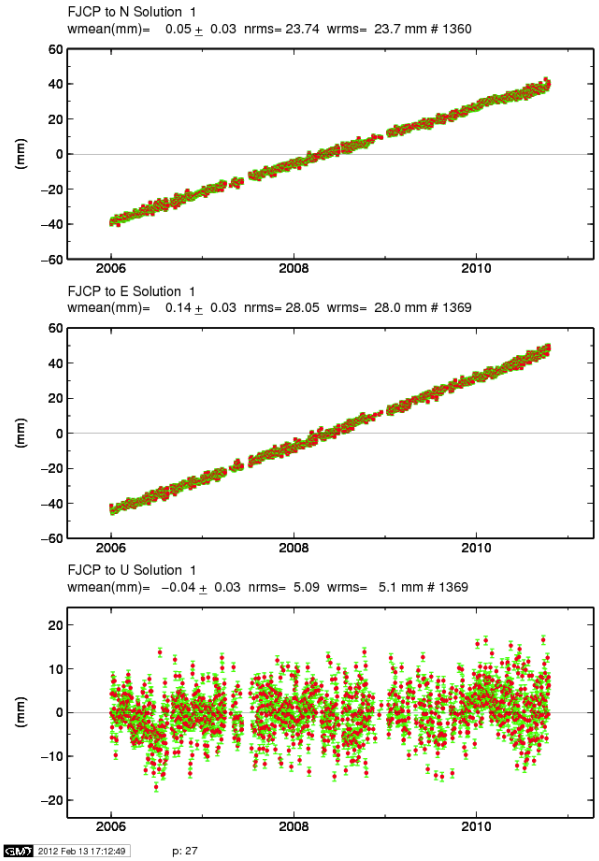


Figure 2.1: Série temporelle du site FJCP (réseau RENAG)

Les séries temporelles de positions de stations GPS (en plus court " séries temporelles GPS ") sont une représentation de l'évolution temporelle d'un point dans un même espace

tridimensionnel en coordonnée géocentriques cartésiennes (x, y, z) ou en coordonnée locale (e, n, u) . Par exemple, la figure 2.1 présente la série temporelle de résidus de position GPS du site FJCP (réseau RENAG) en coordonnées locales sur trois composantes. Ainsi, ce point a trois coordonnées évoluant dans le temps de façon associée, c'est-à-dire que le mouvement du point se répercute sur l'ensemble des trois coordonnées. Or, les coordonnées cartésiennes (x, y, z) des points considérés à une date t_0 sont issues d'une estimation complexes, en général réalisée par moindres carrés, et les résultats obtenus sont donnés avec une matrice variance-covariance associée. Aussi, pour une station donnée, on a une information sur les corrélations des trois coordonnées entre elles.

Pour estimer la qualité de la précision des séries temporelles obtenues, on utilise la répétabilité journalière, un indicateur moyen de la dispersion des positions. La répétabilité peut être calculée sur les lignes de base et les positions. La répétabilité sur les positions est définie comme le résidu quadratique moyen pondéré (wrms) des positions journalières autour de la position moyennée, soit pour chaque composante Est, Nord et Vertical de chaque site [Nocquet, 2011]

$$\text{wrms} = \sqrt{\frac{n\sum_{i=1}^{n}\frac{(y_i-\bar{y})^2}{\sigma_i^2}}{(n-1)\sum_{i=1}^{n}\sigma_i^2}} \tag{2.1}$$

où n est le nombre de positions, $\bar{y}$ est la moyenne pondérée des solutions journalières y_i ou la valeur prédite par un modèle d'évolution, σ_i est la variance de position journalière y_i.

2.2.2 Inventaire des cas à traiter lors de l'analyse des séries temporelles

En pratique pour l'analyse des séries temporelles des réseaux RENAG, GEONET-Japan et le réseau GPS du Taiwan, nous rencontrons fréquemment les problèmes suivants :

1. Outlier isolé

Le cas d'outlier est un cas très fréquent, présents dans la plupart des séries temporelles de position. Les outliers sont des points aberrants qui peuvent être dus à une mauvaise mesure, un manque de données, un problème de matériel, un problème dans la stratégie de calcul (changement de stratégie en cours de série), un effet de réseau (stations de référence changeant ou en nombre différent en fonction des époques), des conditions météorologiques atypiques.

Les outliers apparaissent dans les séries soit comme outlier isolé, séries d'outliers groupés. Des augmentations du bruit moyen ne sont pas à proprement parlé des outliers mais l'analyste peut choisir de considérer les points de ces séquences de forts bruits comme outliers et vouloir les retirer de l'analyse.

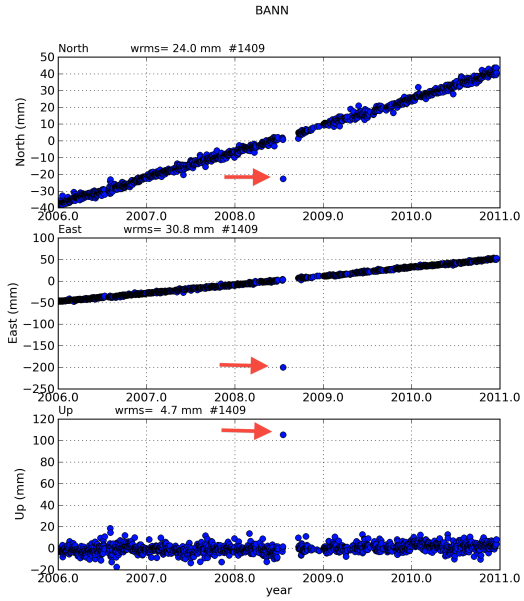


FIGURE 2.2: Exemple d'outlier isolé (site BANN - RENAG)

La figure 2.2 montre la série temporelle du site BANN du réseau RENAG. On peut voir très clairement un outlier isolé dans la série (montré par la flèche rouge). Il est un point aberrant dû à des mauvaises mesures ou un manque de données.

2. Outliers groupés

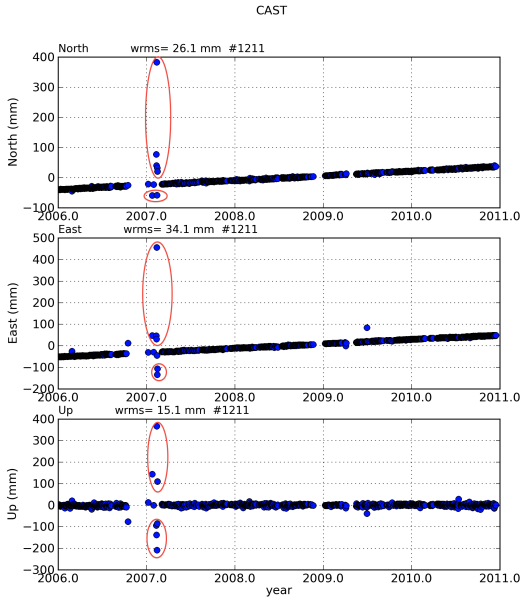


FIGURE 2.3: Exemple d'outliers groupés (site CAST - RENAG)

Un autre cas aussi très fréquent est celui d'outliers groupés par période(s), par exemple

dans la série temporelle du site CAST dans la solution RENAG (figure 2.3). Dans ce cas là, on peut voir des outliers groupés et successifs (les outliers dans l'ellipse rouge). Ils peuvent être dus à des mauvaises mesures ou des problèmes de matériel durant plusieurs jours.

3. Augmentation du bruit

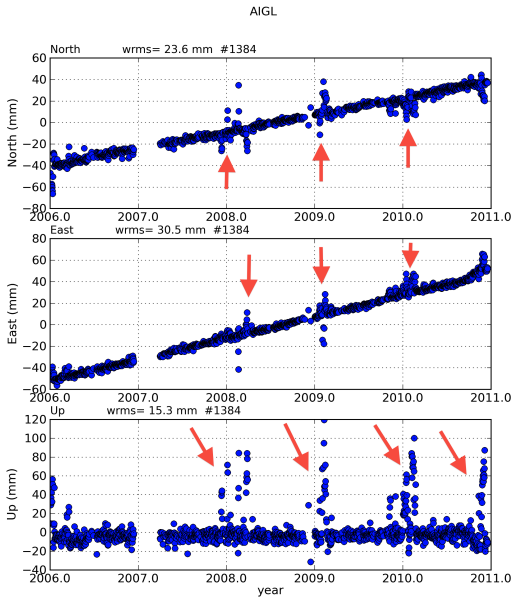


FIGURE 2.4: Exemple d'augmentation du bruit dans une série temporelle (site AIGL - RENAG)

Dans ce cas-là, le nombre d'outliers augmente dans le temps en raison du disfonctionnement d'antenne, du changement de la végétation près de l'antenne, de la couverture de neige sur l'antenne en hiver. La série du site AIGL du réseau RENAG (figure 2.4) montre que les outliers paraissent dans les hivers en 2008, 2009, 2010, 2011 et qu'ils ont tendance à augmenter. Le rejet d'outliers dans ces séries est plus difficile que dans les deux cas précédents.

4. Offset

L'autre problème que l'on rencontre très souvent est la présence d'offsets dans la série. Les offsets sont des sauts causés généralement pour raison matérielle telle que le changement d'antenne, pour des raisons des phénomènes liés à l'environnement de la station comme un choc sur l'antenne ou la coupe d'un arbre à proximité de l'antenne, ou des processus géophysiques comme le pompage rapide d'eau ou de pétrole dans les nappes souterraines et bien sûr les séismes. Enfin, certains sauts peuvent être liés au calcul par exemple lors de la modification des modèles utilisés (atmosphère), le changement d'algorithme ou de stratégie d'analyse.

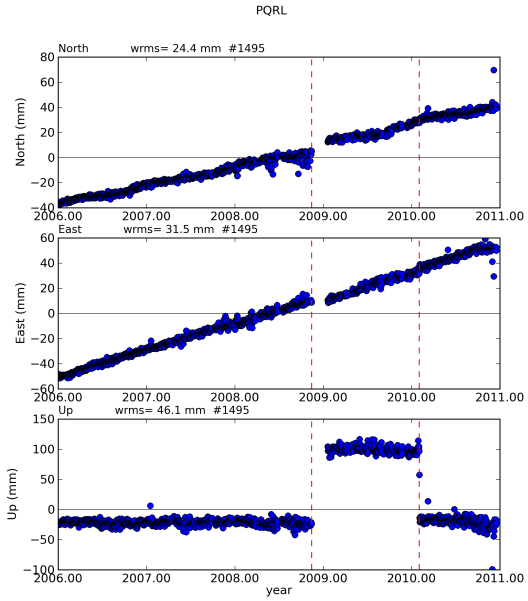


FIGURE 2.5: Exemple d'offsets (site PQRL - réseau RGP)

La figure 2.5 montre deux offsets clairement visibles sur la composante verticale (lignes verticales rouges). Ces deux offsets sont causés par la changement d'antenne, ils n'influencent donc pas fortement la composante horizontale.

5. Déplacement co-sismique simple

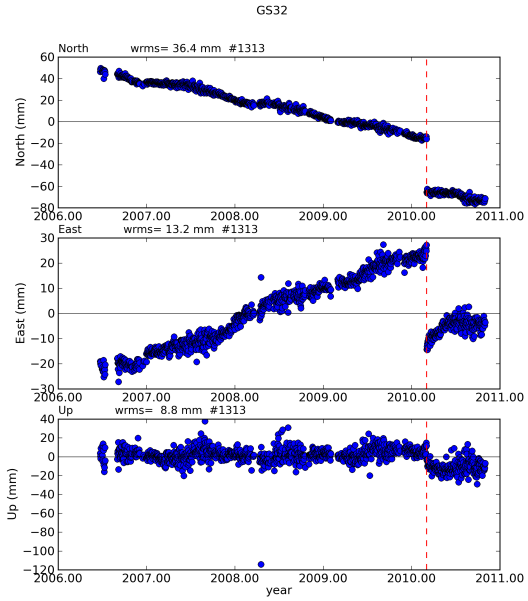


FIGURE 2.6: Exemple de déplacement co-sismique simple (site GS32 - réseau GPS du Taiwan)

Les stations GPS permanentes "voient" généralement un séisme par des offsets dans les

séries temporelles, s'ils sont suffisamment proches de la rupture. La figure 2.7 montre la série temporelle du site GS32 (réseau du Taiwan), où l'on voit le déplacement co-sismique (causé par le séisme Kaohsiung au 4 Mars 2010) sur deux directions Nord et Est.

6. Déplacement co-sismique et post-sismique

On rencontre parfois les offsets co-sismiques et les relaxations post-sismiques à la fois dans les séries temporelles. En particulier, le réseau GPS du Japon (GEONET) a enregistré de nombreux séismes suivies de mouvements non-linéaires.

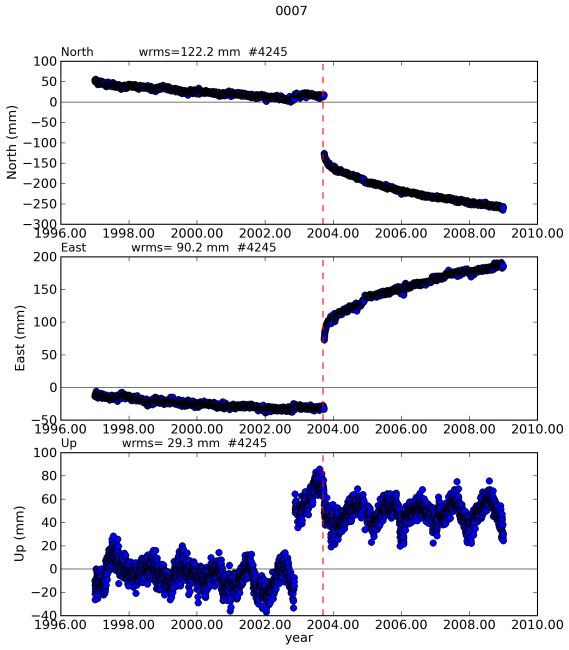


FIGURE 2.7: Exemple de déplacement co-sismique et post-sismique (site 0007 - GEONET Japon)

La figure 2.10 est la série temporelle de position du site 0007 (GEONET - Japon), où l'on voit le déplacement co-sismique (ligne verticale rouge, séisme de Tokachi-Oki au 25 Septembre 2003) et la relaxation post-sismique ensuite décrite comme une courbe exponentielle sur la composante horizontale. Ce cas-là est plus complexe à modéliser dans la série temporelle car il faut estimer simultanément l'offset co-sismique et les paramètres de relaxation post-sismique.

7. Présence de termes saisonniers

Le problème suivant est la présence de termes saisonniers dans les séries temporelles. Ce signal périodique peut être généré par des effets de surcharge (hydrologie continentale par exemple) ou des phénomènes tels que des chutes de neige sur l'antenne.

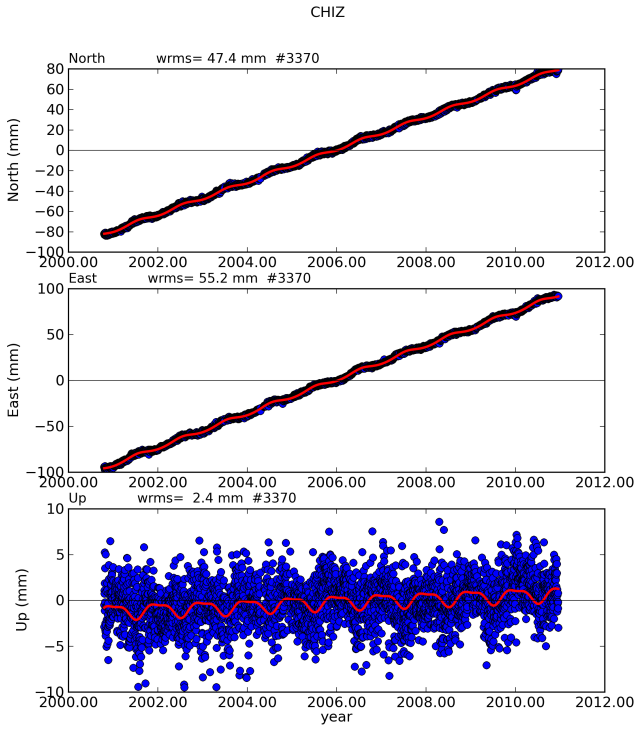


FIGURE 2.8: Exemples de termes saisonniers forts dans une série temporelle (site CHIZ - réseau RENAG)

On voit la signification des termes saisonniers (annuel et semi-annuel) qui est décrite comme la sinusoïde, un signal périodique, sur les composantes horizontales de la série temporelle du site CHIZ du réseau RENAG (figure 2.8). Un des problèmes est que la présence de ces termes peut biaiser l'estimation de la vitesse lorsque les séries temporelles sont courtes. Blewitt and Lavallée [2002] indique que, pour les séries temporelles au-delà de 4.5 ans, ce biais devient négligeable.

8. Signaux transitoires

Un cas, qui reste rare, est celui de signaux transitoires. Ils peuvent être causés par des phénomènes tectoniques tells que le glissement lent sur des failles, mais aussi par des pompages à proximité des points.

La figure 2.9 montre un signal transitoire sur les composantes horizontales de la série temporelle du site CABP (réseau ADN-Equateur, source : J.-M Nocquet), sans doute associé à un séisme lent en Equateur.

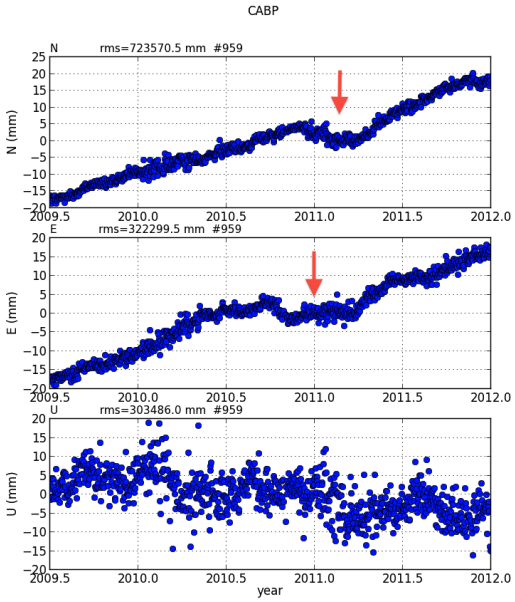


FIGURE 2.9: Exemple de signal transitoire (site CABP - projet ANR ADN, source : J.-M Nocquet)

Quelques auteurs ont étudié cette problématique du point de vue analyse du signal, Ji and Herring [2013] présente une méthode d'amélioration du rapport signal-sur-bruit basée sur le filtrage de Kalman pour analyser les signaux transitoires dans les séries temporelles de position GPS. Vergnolle et al. [2010] applique une interpolation linéaire pour détecter la signal transitoire lié à l'évènement lent dans la série temporelle du site ACAP (de la zone Guerro, Mexique). Cette problématique est difficile et n'a pas été traitée dans ma thèse.

9. Le bruit corrélé des séries temporelles GPS

Le spectre des résidus des séries temporelles GPS montrent une énergie croissante avec la fréquence, indiquant que celui-ci n'est pas blanc [Mao et al., 1999]. L'origine de ce bruit corrélé n'est toujours pas compris de manière claire. Le bruit lié aux petits déplacements du monument supportant l'antenne GPS peut être une marche aléatoire [Langbein, 2008]. Amiri-Simkooei et al. [2007] proposent une méthodologie pour évaluer les caractéristiques de bruit dans les séries temporelles en utilisant l'estimation de la variance de composante par moindres carrés (LS-VCE). Les autres sources de bruit coloré peuvent être une mauvaise modélisation des orbites de satellite, des effets de réseau de référence (orientation de la Terre), une mauvaise modélisation des effets atmosphériques, une mauvaise modélisation des effets de centre de phase d'antenne (qui peut varier avec l'élévation du satellite, l'azimut, les facteurs environnementaux localement) [Mao et al., 1999]. La principale implication de ce type de bruit est une décroissance plus lente de l'incertitude de la vitesse avec la durée d'observation que dans le cas du bruit blanc. Un effet secondaire est la

difficulté du GPS a voir des déplacement transitoires sur de longues périodes.

2.3 Analyse des séries temporelles GPS

Dans le paragraphe précédent, 2.2.2, nous avons listé les problèmes rencontrés dans les séries temporelles GPS. Dans mon travail de thèse, nous analysons automatiquement/semi-automatiquement 3 de ces problèmes rencontrés fréquemment dans les série temporelles : les points aberrants (outliers), les offsets et les dérives post-sismique. Dans le paragraphe suivant, nous proposons une méthode automatique qui permet de détecter et supprimer les outliers, ainsi que de corriger et d'estimer les offsets, et semi-automatiquement les déformations post-sismiques.

2.3.1 Modèle général des séries temporelles GPS

Nous reprenons ici le modèle générale proposé par [Nikolaidis, 2002] : une série temporelle de position GPS est la combinaison de la tendance linéaire, des termes saisonniers (annuels et semi-annuels), éventuellement d'offsets, de déformations post-sismiques. Cette dernière peut être modélisée par une dérive en exponentielle décroissante et un changement de vitesse. Nous utilisons donc le modèle général $y(t)$ pour la série temporelle de n positions GPS :

$$\begin{aligned} y(t_i) = [a + bt_i] + [csin(2\pi t_i) + dcos(2\pi t_i)] + [esin(4\pi t_i) + fcos(4\pi t_i)] \\ + [\sum_{j=1}^{n_g} g_j H(t_i - T_{gj})] + [\sum_{j=1}^{n_h} h_j H(t_i - T_{hj}) t_i \\ + \sum_{j=1}^{n_k} k_j e^{-\frac{t_i - T_{kj}}{\tau_j}} H(t_i - T_{kj})] + v_i \end{aligned} \quad (2.2)$$

Où $t_i (i \in [1; n])$ est l'époque de position i de solution

a est la position à une époque de référence

b est la vitesse

(c, d) et (e, f) sont les paramètres des termes saisonniers aux périods annuels et semi-annuels respectivement

g_j désigne l'amplitude de l'offset au temps T_{gj} (n_{gj} est le nombre d'offset)

H est la fonction d'échelon Heaviside

$$H(t_i - T) = 0 : t_i \leq T$$
$$H(t_i - T) = 1 : t_i > T$$

Enfin, on peut aussi chercher à évaluer les déformations post-sismiques, où h_j est le changement de vitesse au temps T_{hj}, k est l'amplitude de post-séisme au temps T_{kj} (τ est le temp de relaxation exponentielle de post-séisme).

En supposant que les époques d'offset, de changement de vitesse et de relaxation exponentielle sont connus, on obtient un modèle linéaire dont les paramètres inconnus décrivant la série temporelle sont

$$x = [\ a \quad b \quad c \quad d \quad e \quad f \quad g_1 \quad ... \quad g_{n_g} \quad h_1 \quad ... \quad h_{n_h} \quad k_1 \quad ... \quad k_{n_k}\ ]^T \tag{2.3}$$

de sorte que le modèle 2.2 est écrit sous la forme de matrice

$$y = Ax + v \tag{2.4}$$

avec la matrice A:

$$A = \begin{bmatrix} 1 & ... & 1 & ... & 1 & ... & 1 \\ t_1 & ... & t_{n_g} & ... & t_{n_h} & ... & t_{n_k} \\ sin(2\pi t_1) & ... & sin(2\pi t_{n_g}) & ... & sin(2\pi t_{n_h}) & ... & sin(2\pi t_{n_k}) \\ cos(2\pi t_1) & ... & cos(2\pi t_{n_g}) & ... & cos(2\pi t_{n_h}) & ... & cos(2\pi t_{n_k}) \\ sin(4\pi t_1) & ... & sin(4\pi t_{n_g}) & ... & sin(4\pi t_{n_h}) & ... & sin(4\pi t_{n_k}) \\ cos(4\pi t_1) & ... & cos(4\pi t_{n_g}) & ... & cos(4\pi t_{n_h}) & ... & cos(4\pi t_{n_k}) \\ 0 & ... & 1 & ... & 1 & ... & 1 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & ... & 1 & ... & 1 \\ 0 & ... & t_{n_g} & ... & t_{n_h} & ... & t_{n_k} \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & ... & t_{n_h} & ... & t_{n_k} \\ 0 & ... & 1 & ... & 1 & ... & 1 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & ... & 1 & ... & 1 \end{bmatrix}^T$$

Si nous faisons une hypothèse de bruit blanc, la matrice variance-covariance des observations C est simplement

$$C = \begin{bmatrix} \sigma_1^2 & 0 & ... & 0 \\ 0 & \sigma_2^2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \sigma_n^2 \end{bmatrix} \tag{2.5}$$

Où σ_i^2 est la variance de position.

Le système d'équations 2.4 est résolu par moindres carrés donc le vecteur de paramètres inconnus est

$$\hat{x} = (A^T C^{-1} A)^{-1} (A^T C^{-1} y) \tag{2.6}$$

A partir de l'équation 2.4, les positions modélisées sont

$$\hat{y} = A\hat{x} \tag{2.7}$$

et les résidus de positions sont

$$v = y - A\hat{x} \tag{2.8}$$

Le facteur de variance a posteriori est

$$\hat{\sigma}_0^2 = \frac{v^T C^{-1} v}{n - dim(\hat{x})} \tag{2.9}$$

L'amplitude, la phase des termes annuels et semi-annuels sont obtenus par

$$\begin{aligned} A_{ann} &= \sqrt{c^2 + d^2} \\ \phi_{ann} &= \frac{arctan2(c/d)}{2\pi} \\ A_{semi-ann} &= \sqrt{e^2 + f^2} \\ \phi_{semi-ann} &= \frac{arctan2(e/f)}{4\pi} \end{aligned} \tag{2.10}$$

En réalité, on ne rencontre pas beaucoup de séries qui incluent toutes les composantes décrites ci-dessus (heureusement !).

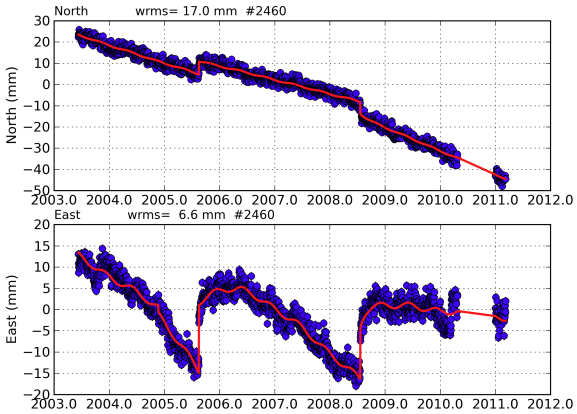


Figure 2.10: Série temporelle du site P106 (GEONET - Japon) et sa modélisation

La figure 2.10 est une des séries temporelles les plus complexes que nous avons rencontrées est celle site P106 du réseau GEONET au Japon, montrant les deux composantes

horizontales. Les solutions de position journalière (cercle bleu) sont modélisées avec trois tendances linéaires (ligne rouge), la sinusoïde des termes annuels et semi-annuels, deux offsets co-sismiques en 2005 et 2008, et deux relaxations exponentielle post-sismiques.

2.3.2 Détection automatique des outliers et des offsets dans les séries temporelles

2.3.2.1 Proposition

Une difficulté pratique est que l'estimation des paramètres décrivant les séries temporelles par moindres carrés nécessitent des séries temporelles "propres". En effet, les outliers peuvent biaiser l'estimation des autres paramètres. La première étape est donc l'identification et la suppression des outliers dans les séries temporelles. En pratique, comme nous allons le voir en présentant les algorithmes de détection des outliers, outliers et offsets présentent des caractéristiques communes. Par exemple, un offset non corrigé peut induire des résidus important qui risquent d'être (mal) interprétés comme outliers. Les algorithmes devront d'abord identifier les anomalies et les classer ensuite soit comme outliers ou comme offset. Dans la suite, je présente l'algorithme de notre méthode, et son comportement sur plusieurs séries temporelles.

2.3.2.2 Première identification par rms

Cette approche est basée sur les variations de rms (root mean square) calculées par rapport à la position moyenne locale dans le temps. Elle permet d'identifier les fenêtres de temps présentant un comportement anormal. Une fois identifiées les fenêtres temporelles "anormales", la différenciation des positions successives dans la fenêtre permet d'isoler les outliers et les offsets. La détection des outliers et des offsets se fait automatiquement sur chacune des composantes e, n, u.

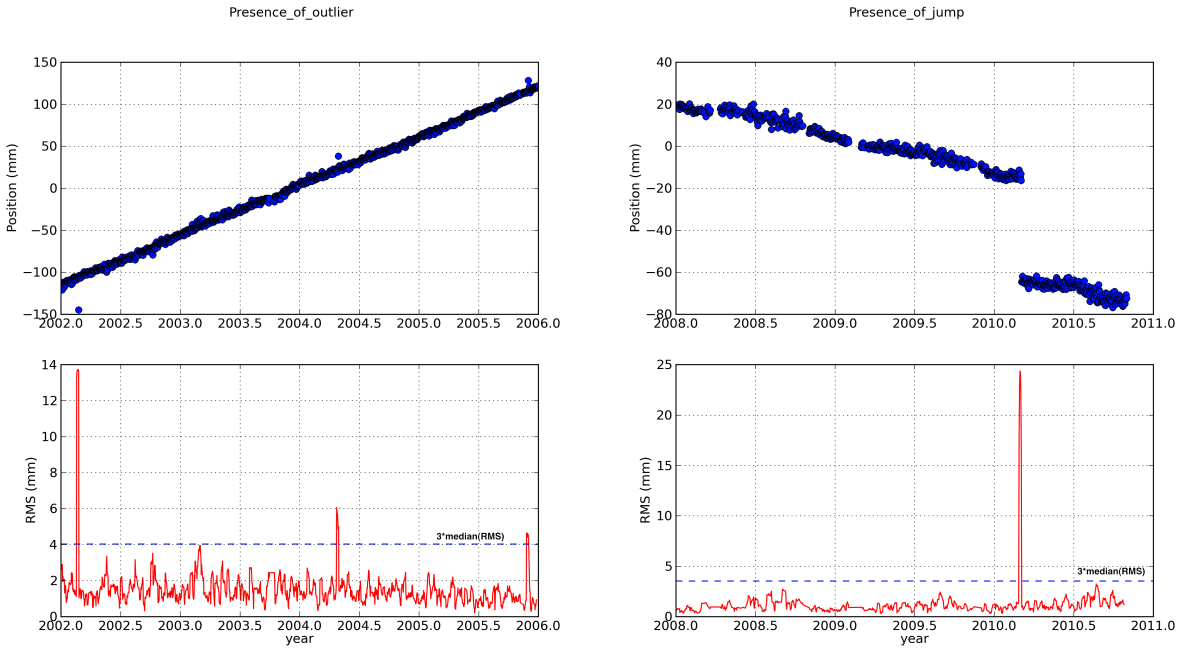
Pour cela, nous divisons la série temporelle de position en fenêtres glissantes de n positions successives, calculons ensuite ses rms par la formule 2.11. Nous obtenons la série temporelle de rms de fenêtres glissantes.

$$rms_i = (\sqrt{\frac{\sum_{j=1}^{n}(P_j - \bar{P}_i)^2}{n-1}})_i \tag{2.11}$$

Où P_j est la position j dans la fenêtre i ; $\bar{P}_i$ est la position moyenne de la fenêtre i. n doit être choisi suffisamment grand pour obtenir une statistique significative, mais

suffisamment petit pour qu'un outlier influe significativement sur la valeur du rms. Une valeur de 7 est celle ensuite proposée par défaut dans la recherche automatique.

Un exemple de cette approche est illustré dans la figure 2.11. On voit que les rms anormaux correspondent aux fenêtres qui contiennent soit des outliers (figure 2.11 gauche), soit des offsets (figure 2.11 droit).



En présence d'ouliers En présence d'offset

FIGURE 2.11: Série temporelle de position (haut) et série temporelle de rms calculé par fenêtre glissante de 7 jours (bas)

Nous définissons une valeur de seuil 2 (ou 3) fois la médiane de série temporelle de rms pour détecter les rms anormaux. Sur la figure 2.11 (bas), cette valeur de seuil est indiquée par le trait horizontal bleu. Cette identification des fenêtres anormales va être ensuite analysée par d'autres approches.

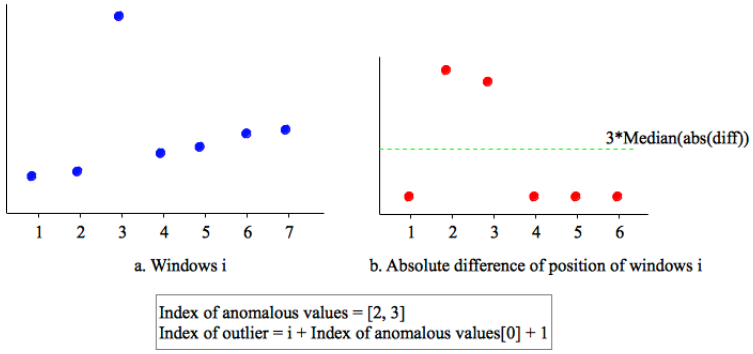
2.3.2.3 Détection des outliers

On cherche maintenant les valeurs aberrantes, basées sur la différence entre 2 positions glissantes successives dans la fenêtre anormale identifiée. Pour cela, on calcule la valeur absolue de la différence entre 2 positions glissantes successives. Pour $n = 7$, on a par exemple:

$$diff = [|P_2 - P_1|, |P_3 - P_2|, ..., |P_7 - P_6|] \tag{2.12}$$

On choisit une valeur de seuil pour la détection. On choisit une valeur de 2 (ou 3) fois la médiane des valeurs absolues des différences.

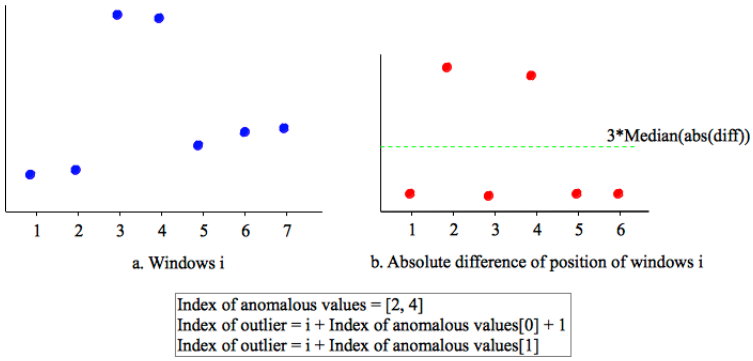
A ce stade, il faut remarquer qu'il existe un nombre maximum d'outliers qui peuvent être détectés dans une fenêtre de longeur n. Par exemple, pour une fenêtre de 7 jours, on ne peut détecter que 1 (figure 2.12) ou 2 outliers consécutifs (figure 2.13). Dans ces deux cas-là, on trouvera toujours deux valeurs aberrantes dans les valeurs absolues de différence de 2 positions successives.



Cercle bleu: la position
Cercle rouge: la valeur absolue de la différence de 2 positions

Figure 2.12: Détection d'un outlier dans la fenêtre anormale i

Le cas d'un outlier dans la fenêtre anormale (figure 2.12 gauche), les deux valeurs anormales sont successives (figure 2.12 droite). L'indice de l'outlier est défini comme l'indice de première valeur anormale + 1.

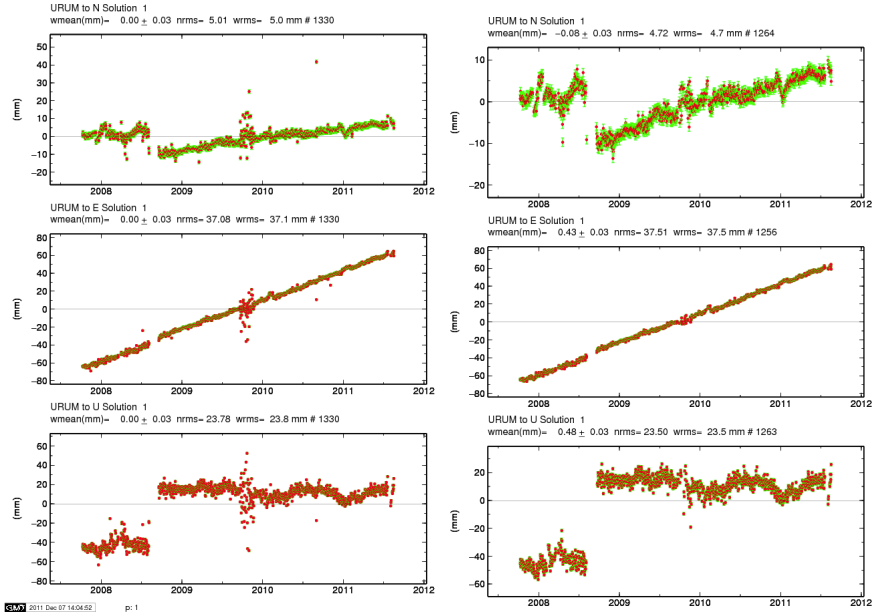


Cercle bleu: la position
Cercle rouge: la valeur absolue de la différence de 2 positions

Figure 2.13: Détection de deux outliers continus dans la fenêtre anormale i

Le deuxième cas est le cas de deux outliers successifs dans la fenêtre anormale (figure 2.13 gauche). Dans ce cas là, deux valeurs anormales ne sont donc pas successives (figure 2.13 droite), puisque l'indice du premier outlier est l'indice de première valeur anormale + 1, et l'indice de deuxième outlier est l'indice de deuxième valeur anormale.

Cette méthode permet de détecter très bien les outliers, surtout dans le cas où coexistent les outliers et les offsets. On voit dans la figure 2.14 un exemple de la détection d'outliers dans la série temporelle du site URUM, une série temporelle complexe avec un offset et les outliers.



a. Série temporelle bruitée b. Série temporelle nettoyée

Figure 2.14: Détection d'outliers dans la série temporelle du site URUM

La limitation de cette méthode est sa sensibilité. Lors de l'application de cette méthode, nous constatons qu'elle se trompe parfois dans la date correcte. Nous la réalisons donc avant la détection des offsets, comme l'étape de l'identification " préliminaire " d'outliers.

L'identification finale d'outliers est réalisée après la détection des offsets (et l'estimation de temps des relaxations post-sismiques s'il est estimé) à partir du modèle complet.

Pour cela, on estime le modèle complet 2.2 sur la série temporelle intégrale avec les époques d'offsets, les τ_i,... que l'on a défini. On calcule ensuite les résidus de positions (formule 2.8) et élimine les outliers si leurs résidus sont supérieurs au seuil d'élimination des outliers suivant la formule 1.44 :

$$|v_i^e| \geq 3 * median(|v^e|)$$
$$|v_i^n| \geq 3 * median(|v^n|)$$
$$|v_i^u| \geq 3 * median(|v^u|)$$

Cette méthode est illustrée par la figure 2.15. Les cercles bleus sont les résidus de positions, la ligne rouge est le modèle de la série, les lignes vertes sont le seuil d'élimination des outliers. On détecte 3 outliers qui sont 3 résidus en dehors des lignes vertes.

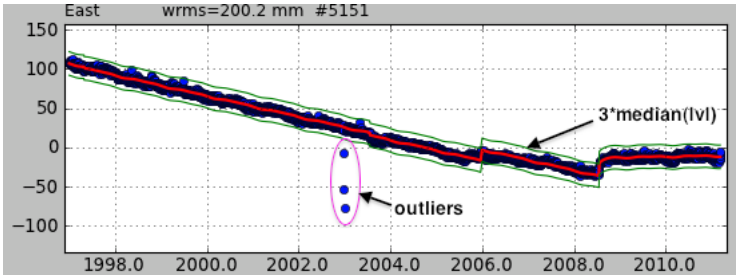
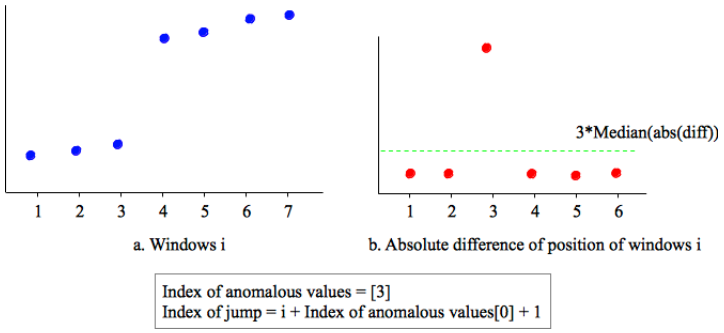


Figure 2.15: Exemple de détection d'outliers par le test 1.44

Le processus d'identification et de rejet des outliers est réalisé itérativement jusqu'il n'y ait plus outliers ou que le pourcentage des outliers sur les positions soit supérieure ou égal à 5% . Dans ce dernier cas, la série présente d'autres problèmes qui sont analysés par d'autres algorithmes. Bien que cette dernière méthode qui nécessite l'ensemble de la série, soit plus lente que la méthode basée sur rms, elle s'est révélée assez robuste dans la plupart des cas rencontrés.

2.3.2.4 Détection des offsets

Dans le cas des offsets, on ne trouvera toujours qu'une seule valeur aberrante sur les différences absolues de positions successives dans la fenêtre anormale identifiée. L'indice d'offset est défini comme l'indice de valeur anormale + 1 (figure 2.16).



Cercle bleu: la position
Cercle rouge: la valeur absolue de la différence de 2 positions

Figure 2.16: Détection d'offset dans la fenêtre i

Cette méthode est rapide et efficiente, surtout une fois les séries temporelles nettoyées des outliers évidents. Son défaut est d'être très sensible. Elle identifie donc des petits offsets qui ne sont pas significatifs. On voit un exemple de la détection d'offsets pour le site BRST, un site de la solution Géoazur - I du RENAG (figure 2.17). L'algorithme trouve trois offsets, alors que deux sont corrects (correct : ligne verticale bleu, incorrect : ligne verticale rouge).

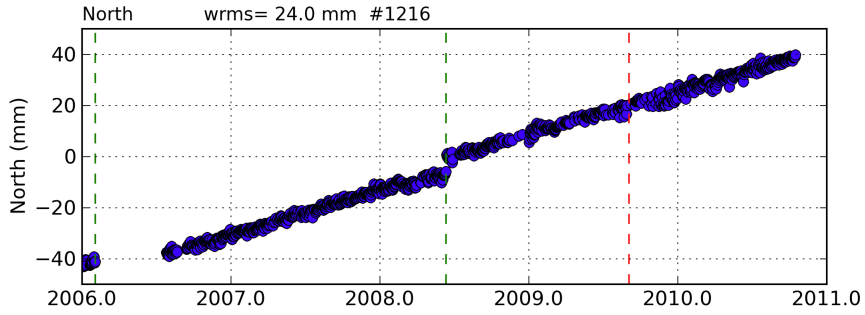


FIGURE 2.17: Série temporelle de la composante nord du site BRST (Geoazur I) avec les offsets détectés par la méthode 2.3.2

Pour surmonter cette limitation, nous effectuons a posteriori le test statistique de Fisher (F ratio) qui permet de vérifier si les offsets détectés sont bien significatifs.

2.3.2.5 Vérification des offsets détectés

Le test "F ratio" est présenté pour exemple dans la thèse HDR de Nocquet [2011] pour tester les modèles cinématiques globaux. Il permet à la comparaison de deux estimations par moindres carrés, avec de décider lequel des deux modèles explique le mieux les données.

Pour un ensemble d'observations, rajouter des paramètres estimés va diminuer le χ^2. Le F ratio indique si la diminution de χ^2 d'un modèle à p_2 vs p_1 ($p_1 > p_2$) degrés de liberté est significatif. La quantité $\chi^2 = v^T C^{-1} v$ suit une loi de probabilité du χ^2 à p degrés de liberté.

La quantité F ratio

$$F = \frac{(\chi^2(p_1) - \chi^2(p_2))p_2}{(p_1 - p_2)\chi^2(p_2)} \tag{2.13}$$

suit une distribution de Fisher-Snedecor à (p_1, p_2) degrés de liberté. La valeur calculée empiriquement à partir des données est donc comparée à la prédiction de la loi $F(p_2, p_1 - p_2)$ pour un niveau de risque α donné ($\alpha = 1\% \div 5\%$, correspondant à un niveau de confiance de $99\% \div 95\%$).

L'hypothèse nulle dans le test est que l'ensemble de la série temporelle peut être modélisé avec le modèle en cours, et l'hypothèse non-nulle que l'ensemble de la série temporelle peut être modélisé avec le modèle en cours + 1 offset. On teste donc

$$F \leq f_\alpha^{p_2, p_1 - p_2} \tag{2.14}$$

où $f_\alpha^{p_2, p_1 - p_2}$ est le fractile $\alpha\%$ de la loi de Fisher-Snedecor à (p_1, p_2) degrés de liberté.

Dans notre étude, on a donc

$$p_1 = n - t$$

où t est le nombre d'inconnues dans le modèle actuel, n est le nombre d'observations. En réalité, nous limitons le nombre d'observations n pour tester un offset donné, par le nombre d'observation entre l'offset précèdent et le nombre d'observation entre l'offset à tester et celui qui le suit immédiatement.

La seconde est simplement :

$$p_2 = n - t - 1$$

Si (2.14) n'est pas vérifiée, on peut en conclure avec $\alpha\%$ de risque de se tromper que l'offset testé n'est pas significatif.

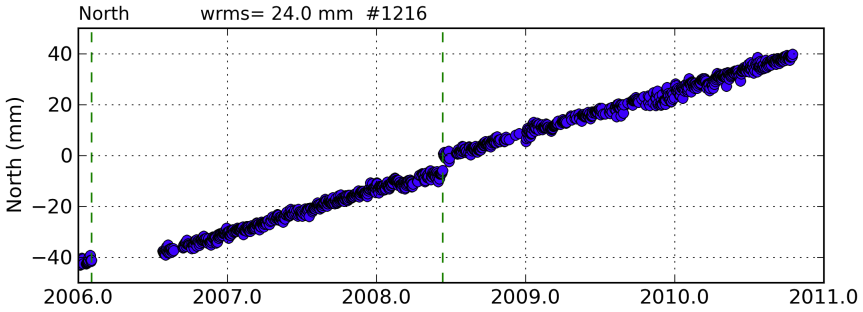


FIGURE 2.18: Série temporelle de la composante nord du site BRST (Geoazur I - RENAG) avec les offsets détectés et testés

Nous réalisons le test F ratio pour vérifier ces 3 offsets de la composante nord de BRST dans le paragraphe 2.3.2.4. Par rapport à la figure 2.17, on voit que le test a rejeté le troisième offset (figure 2.18).

2.3.3 Estimation du temps de relaxation post-sismique

Dans le modèle 2.2 de la série temporelle, les paramètres de la déformation post-sismique sont le temps de relaxation post-sismique Tau (τ) et l'amplitude de celle-ci [**Nikolaidis, 2002**]. Dans notre étude, τ est déterminé par la résolution approximative itérative du modèle 2.2. A τ fixé, le modèle est linéaire et peut donc être estimé par moindres carrés.

Ainsi, pour un processus itératif de la résolution, on résout l'équation 2.2 en deux étapes. La première étape, le τ_j initial est donné, on résout ensuite le système d'équation linéaire 2.4 et obtient le vecteur des paramètres $\hat{x}$. Le vecteur $\hat{x}$ est pour les valeurs initiales (a^0, b^0, c^0, d^0, e^0, f^0, g_j^0, h_j^0, k_j^0) à l'étape suivante.

La deuxième étape est la définition des temps de relaxation post-sismique τ_j. Pour cela, on isole l'équation 2.2 en deux portions : une portion causée par les séismes contenant

les offsets sismiques, les relaxation post-sismique et peut-êtres les changements de vitesse ; l'autre portion est le reste de l'équation qu'il s'agit la tendance linéaire, les termes saisonniers et les offsets matériels. On a l'équation équivalente :

$$y(t_i) = [a + bt_i + csin(2\pi t_i) + dcos(2\pi t_i) + esin(4\pi t_i) + fcos(4\pi t_i) + \sum_{j=1}^{(n_g - n_k)} g_j H(t_i - T_{gj})] + [\sum_{j=1}^{n_k} [g_j H(t_i - T_{kj}) + h_j H(t_i - T_{kj}) t_i + k_j e^{-\frac{t_i - T_{kj}}{\tau_j}} H(t_i - T_{kj})]] + v_i \quad (2.15)$$

Puis cette équation est dérivée avec les paramètres inconnus g_j, h_j, k_j, τ_j . On a

$$y'(t_i) = \sum_{j=1}^{n_k} [g_j^0 H(t_i - T_{kj}) \partial g_j + h_j^0 H(t_i - T_{kj}) t_i \partial h_j + k_j^0 e^{-\frac{t_i - T_{kj}}{\tau_j^0}} H(t_i - T_{kj}) \partial k_j + k_j^0 \frac{(t_i - T_{kj})}{(\tau_j^0)^2} e^{-\frac{(t_i - T_0)}{\tau_j^0}} H(t_i - T_{kj}).\partial \tau] \quad (2.16)$$

où

$$y'(t_i) = y(t_i) - [a^0 + bt_i^0 + c^0 sin(2\pi t_i) + d^0 cos(2\pi t_i) + e^0 sin(4\pi t_i) + f^0 cos(4\pi t_i) + \sum_{j=1}^{n_g} g_j^0 H(t_i - T_{gj}) + \sum_{j=1}^{n_h} h_j^0 H(t_i - T_{hj}) t_i + \sum_{j=1}^{n_k} k_j^0 e^{-\frac{t_i - T_{kj}}{\tau_j^0}} H(t_i - T_{kj})] \quad (2.17)$$

On écrit le système d'équations 2.16 sous la forme de matrice

$$A\partial = y,\ C^{-1} \quad (2.18)$$

où le vecteur des paramètres inconnus ∂ est :

$$\partial = [\ \partial g_1 \quad ... \quad \partial g_{n_k} \quad \partial h_1 \quad ... \quad \partial h_{n_k} \quad \partial k_1 \quad ... \quad \partial k_{n_k} \quad \partial \tau_1 \quad ... \quad \partial \tau_{n_k}\ ]^T$$

la matrice des dérivés partielles A soit

$$A = \begin{bmatrix} g_1^0 & \dots & g_1^0 & \dots & g_1^0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & g_{n_k}^0 & \dots & g_{n_k}^0 \\ h_1^0 t_i & \dots & h_1^0 t_{i+j} & \dots & h_1^0 t_n \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & h_{n_k}^0 t_{i+j} & \dots & h_{n_k}^0 t_n \\ k_1^0 e^{-\frac{t_i - T_{k1}}{\tau_1^0}} & \dots & k_1^0 e^{-\frac{t_{i+j} - T_{k1}}{\tau_1^0}} & \dots & k_1^0 e^{-\frac{t_n - T_{k1}}{\tau_1^0}} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & k_{n_k}^0 e^{-\frac{t_{i+j} - T_{kn_k}}{\tau_{n_k}^0}} & \dots & k_{n_k}^0 e^{-\frac{t_n - T_{kn_k}}{\tau_{n_k}^0}} \\ k_1^0 \frac{(t_i - T_{k1})}{(\tau_1^0)^2} e^{-\frac{(t_i - T_{k1})}{\tau_1^0}} & \dots & k_1^0 \frac{(t_{i+j} - T_{k1})}{(\tau_1^0)^2} e^{-\frac{(t_{i+j} - T_{k1})}{\tau_1^0}} & \dots & k_1^0 \frac{(t_n - T_{k1})}{(\tau_1^0)^2} e^{-\frac{(t_n - T_{k1})}{\tau_1^0}} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & k_{n_k}^0 \frac{(t_{i+j} - T_{kn_k})}{(\tau_{n_k}^0)^2} e^{-\frac{(t_{i+j} - T_{kn_k})}{\tau_{n_k}^0}} & \dots & k_{n_k}^0 \frac{(t_n - T_{kn_k})}{(\tau_{n_k}^0)^2} e^{-\frac{(t_n - T_{kn_k})}{\tau_{n_k}^0}} \end{bmatrix}^T$$

le vecteur de coefficients libres y soit

$$y = \begin{bmatrix} y'_i & \dots & y'_n \end{bmatrix}^T$$

L'estimateur de ∂ est déterminé par moindres carrés :

$$\widehat{\partial} = (A^T C^{-1} A)^{-1} (A^T C^{-1} y) \tag{2.19}$$

De sorte que le temps de relaxation à l'itération $i+1$ devient :

$$\tau_j^{i+1} = \tau_j^i + \partial \tau_j \tag{2.20}$$

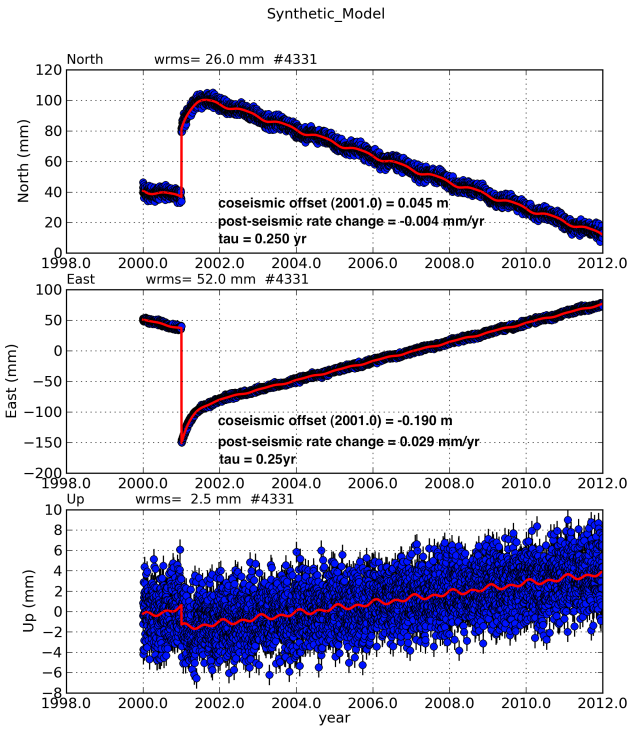
Le processus itératif s'arrête quand le maximum de $\partial \tau_j$ est inférieur à la valeur de seuil et que donc la convergence est réalisée. En pratique, nous choisissons cette valeur à 0.003 an, qui correspond au temps entre deux observations de position journalière.

L'estimation du temps de relaxation post-sismique requiert la date de séisme T_k. Nous avons tenté de détecter automatiquement les dates T_k et signaux post-sismiques pour analyser automatiquement les déformations post-sismiques dans les séries temporelles, mais nous n'avons pas réussi à obtenir des résultats robustes. Il faut donc fournir les dates des séismes susceptibles de générer un signal post-sismique significatif comme l'information a priori pour la analyse, une information a priori généralement disponible, car elle est associée à la magnitude du séisme.

Concrètement, les époques de séisme de la solution de référence IGS sont trouvés dans la liste des discontinuités publiés et mise à jour hebdomadairement sur le lien `ftp://igs-rf.ign.fr/pub/IGS08/soln_IGS08.snx`.

L'estimation du temps de relaxation post-sismique τ est simultanément estimée sur 2 composantes horizontales, puis τ est appliqué pour 3 composantes.

On examine maintenant la performance de cet algorithme par un test simple. Dans ce test, nous produisons tout d'abord un modèle synthétique de la série temporelle de positions journalières pour la période 2000.0-2012.0. Nous introduisons un offset "co-séisme" en 2001.0 et un effet "post-séisme" sur les composantes horizontales. L'amplitude de l'offset co-séisme est 0.045 m du Nord, -0.190 m de l'Est. L'effet post-séisme est modélisé par un changement de vitesse (-4 mm/yr du Nord, 29 mm/yr de l'Est) et une relaxation exponentielle $\tau = 0,25yr$. Ensuite les positions du modèle synthétique sont perturbées par la loi Normale $\sigma = 0.002m$. Nous estimons ensuite les paramètres co- et post-sismiques (figure 2.19) et les comparons aux valeurs réelles.



Cercle bleu: position jounalière
Ligne rouge: modèle

Figure 2.19: Série temporelle synthétique et sa modélisation automatique réalisée par pygps_time_series.py

Utilisons cet algorithme pour estimer la série temporelle synthétique, nous obtenons le très bon résultat que les paramètres estimés sont presque à ceux donnés (voir le tableau 2.1, la figure 2.20).

Composante	Paramètre	Amplitude du offset (m)	Changement de vitesse (m/yr)	Tau (yr)
Nord	Synthétique	0.045	- 0.004	0.25
	Estimé	0.0447	- 0.0042	0.250
		0.0003	0.0002	0
Est	Synthétique	- 0,190	- 0.029	0.25
	Estimé	- 0.1892	- 0.0293	0.250
		-0.0008	0.0003	0

TABLE 2.1: Comparaison des paramètres synthétiques et estimés

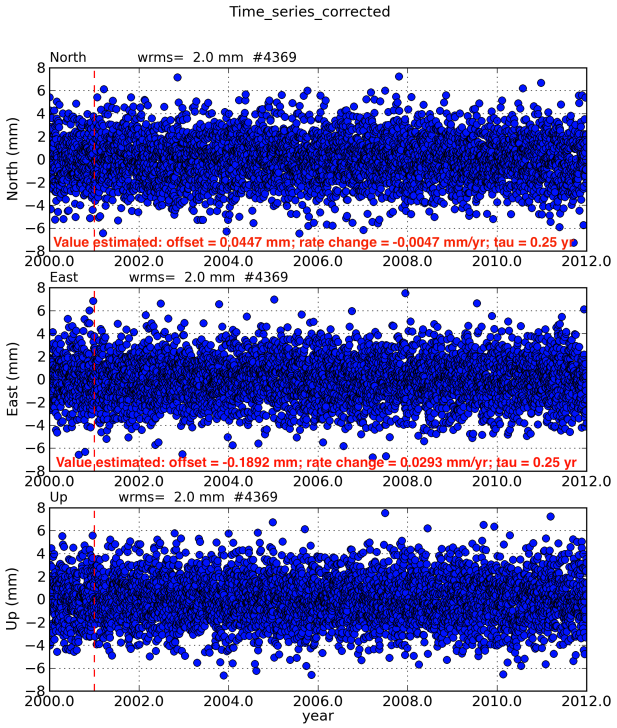


Figure 2.20: Série temporelle résiduelle, une fois le modèle estimé par pygps_time_series.py.

L'estimation est très rapide avec quatre itérations, la valeur estimée du Tau est 0.25 yr, est égale à celle synthétique. On voit les séries temporelles corrigées sur la figure 2.20, tous les paramètres sont bien estimés avec un $wrms = 2.0mm$ sur les 3 composante de la série temporelle, égaux à l'amplitude de la perturbation introduite.

2.4 Conclusions

La méthode de la détection des outliers et des offsets que nous développons est basée sur les variations de rms calculées par rapport à la position moyenne locale. Cette approche

nous permet d'analyser automatiquement et rapidement les séries temporelles de positions GPS dans les cas simples.

L'algorithme de détermination du temps de relaxation post-sismique τ requiert a priori l'époque du séisme. Cet algorithme a d'abord été testé sur un cas synthétique. Nous l'avons ensuite testé sur plusieurs séries temporelles réelles du réseau GEONET et du réseau GPS du Taiwan. Dans quelques cas, la détermination de τ n'a pas été possible, l'algorithme ne convergeant pas. En effet, comme déjà notée par Nikolaidis [2002], il existe une corrélation entre τ et le changement de vitesse post-séisme. Suivant la durée de l'observation et la forme du signal, il n'est pas possible d'obtenir ces deux paramètres simultanément.

Dans le chapitre suivant, nous décrivons l'implémentation pratique des méthodes vues précédemment.

CHAPITRE 3

IMPLÉMENTATION : LA CLASSE GTS ET LE PROGRAMME PYGPS_TIME_SERIES.PY

3.1 Philosophie

Les réseaux GPS continus comportent régulièrement plusieurs centaines de stations sur maintenant une décennie d'observation. Si l'analyse d'une série temporelle est toujours possible manuellement, une approche manuelle n'est plus envisageable lorsqu'il est question d'analyser plusieurs centaines de séries temporelles. De manière idéale, on souhaiterait disposer d'un programme capable de détecter les problèmes, de les résoudre, d'identifier tous les signaux d'intérêt géophysique et de fournir l'ensemble des paramètres décrivant la série temporelle. Pour l'instant, bien que l'inventaire des problèmes présents dans les séries temporelles que nous avons réalisés montre un nombre finis de cas, l'analyse en aveugle des séries temporelles GPS n'est possible que pour des cas simples. Devant ce constat, le but de notre développement doit tenter d'atteindre 3 objectifs :

1. Permettre une utilisation souple dans un mode entièrement manuel

L'objectif est de fournir à l'utilisateur un ensemble de fonctions permettant de réaliser des opérations primitives par un opérateur, en mode interactif. Une opération primitive est par exemple une extraction de la série temporelle sur un ou des périodes de temps, la suppression manuelle d'un point de la série, ou le retrait d'une pente donnée.

2. Fournir des routines qui réalisent certaines manipulations sur les séries temporelles de manière paramétrée

Ici, on s'appuie sur les fonctions primitives pour réaliser des opérations plus complexes. Par exemple, il s'agit de la recherche automatique d'outliers qui doit être contrôlée par un petit nombre de paramètres. Cette étape correspond à une démarche d'analyse assistée, où l'on recherche les meilleures paramètres pour contrôler les résultats des algorithmes.

3. Fournir un ou des programmes "wrapper"

Le "wrapper" enveloppe la succession d'opérations primitives ou de manipulations intermédiaires pour traiter de manière automatique un grand nombre de séries temporelles.

En matière d'analyse automatique de séries temporelles, peu de travaux ayant pour but l'analyse de signaux complexes ont été menés. En revanche, il existe des méthodes d'analyse manuelle, ou assistée, telles que le module Tsview Gamit/Globk (http://www-gpsg.mit.edu/~tah/GGMatlab/#_tsview), reposant sur le logiciel commercial Matlab. Tsview peut être qualifiée de méthode assistée puisque, si elle permet d'éliminer automatiquement les outliers, il faut indiquer un certain nombre d'informations, dates et durée d'évènements par exemple, pour permettre l'estimation des offsets ou des paramètres du mouvement post-sismique.

3.2 Implémentation sous Python

Le langage Python est bien adapté pour réaliser un développement modulaire permettant l'emboîtement des différentes couches, ainsi que la visualisation des résultats. Le principe que nous avons retenu est celui d'un développement orienté-objet. Dans Python, l'approche objet repose sur des classes. Une classe peut être vue comme un objet abstrait sur lequel on va appliquer des méthodes. La classe peut être naturellement une série temporelle, mais aussi un objet plus abstrait comme un estimateur. Une méthode peut être le fait de retirer un point de la série temporelle ou de la visualiser. Python utilise la syntaxe du "." pour l'application d'une méthode à une instance d'une classe. Ainsi pour visualiser une série temporelles à laquelle on a retiré la pente, il suffira d'écrire : ts.detrend().plot().

3.2.1 La Classe Gts

La première Classe (Geodetic time series) que nous devons définir est celle d'une série temporelle. La structure que nous avons choisie est la suivante :

Une série temporelle est d'abord un tableau (en fait un numeric array sous Numpy) à 9 colonnes, où chaque colonne contient :

- la date en année décimale
- les composantes Nord, East, Vertical
- les écart-types sur les trois composantes

Par ailleurs, la Classe Gts inclu aussi les informations suivantes :

- le code : il s'agit généralement de l'acronyme codé classiquement par 4 lettres.
- les coordonnées géographiques approchées longitude, latitude, exprimées en degrés décimaux.
- la liste des outliers : il s'agit d'une liste d'indices correspondant à la date de l'outlier.

- les valeurs des offsets : il s'agit d'un tableau numérique (Numpy array), contenant 7 colonnes (date en année décimale, valeurs des offsets sur les composantes Nord, East, Verticale, valeurs des écart-types des offsets)
- les paramètres des mouvements post-sismiques : il s'agit d'un tableau sur 8 colonnes contenant la date du séisme, les amplitudes du déplacements post-sismiques, leur écart-type et le temps de relaxation.
- les termes annuels : il s'agit d'une liste comprenant 6 termes : les amplitudes du terme annuel pour les trois composantes et la phase associée.
- les termes semi-annuels : il s'agit d'une liste comprenant 6 termes : les amplitudes du terme semi-annuel pour les trois composantes et la phase associée.
- la vitesse : il s'agit d'une liste comprenant les valeurs de la vitesse sur les 3 composantes et les écart-types associés.

3.2.2 Les méthodes primitives

Les méthodes primitives permettent les opérations élémentaires sur les séries temporelles. Ce sont :

- read_mb_file : lit les fichiers des séries temporelles résiduelles sous le format mb_file pour générer les classes Gts des séries temporelles.
- copy : effectue une copie de la classe de la série temporelle.
- differentiate : différencie une série temporelle.
- remove_outliers : retire les outliers de la série temporelle et fournit la série temporelle nettoyée.
- make_model : estime le modèle de la série temporelle par moindres carrés en utilisant l'option du modèle contenant la pente (detrend), ou du modèle contenant la pente avec les termes annuels (detrend_annual), ou du modèle contenant la pente avec les termes annuels, semi-annuels (detrend_seasonal). Cette fonction retourne la nouvelle classe Gts de la série temporelle qui est maintenant la série temporelle résiduelle et ses valeurs associées (la vitesse, les termes annuel, semi-annuel,...).
- rms : calcule les rms des fenêtres glissantes de n positions successives de la série temporelle.

3.2.3 Les méthodes de second niveau

Les méthodes de second niveau encapsulent des fonctions primitives et utilisent des algorithmes itératif pour réaliser des opérations complexes. Ce sont :

- find_outliers_simple : trouve rapidement les outliers dans la série temporelle en comparant les positions avec leur médiane.
- find_outliers_and_offsets_through_differentiation : trouve rapidement les outliers et les offsets à la fois en utilisé la différenciation des 2 positions successives dans la série.
- find_outliers_by_RMS_ts : trouve les outliers dans une série temporelle par notre méthode basée sur les rms des fenêtres glissantes, décrite dans le paragraphe 2.3.2.3.
- find_outliers_by_residuals : trouve les outliers en comparant les résidus obtenus à partir de l'estimation du modèle complet de la série avec l'écart type. Cette méthode est aussi décrite dans le paragraphe 2.3.2.3.
- find_time_offsets : trouve les date des offsets suspects dans la série temporelle par notre méthode basée sur les rms des fenêtres glissantes (méthode décrite dans le paragraphe 2.3.2.4).
- Ftest_4_offsets : réalise un test "F_ratio" de significativité d'un offset trouvé automatiquement dans la série temporelle.
- find_tau_4_neu : estime le temps de la relaxation post-sismique en utilisant le fichier de discontinuités pour la date du séisme susceptible de générer un signal post-sismique significatif. La méthode utilisée est décrite dans le paragraphe 2.3.3.

3.2.4 La méthode de visualisation plot

La visualisation des résultats est un point qui se révèle très important dans l'analyse des séries temporelles. Nous avons développé une méthode permettant de visualiser les informations d'une instance de la Classe Gts. La méthode plot permet de :

- visualiser les dates des offsets par une ligne verticale
- visualiser les outliers
- choisir une fenêtre de temps
- fixer l'intervalle des ordonnées
- choisir les composantes à représenter
- visualiser des périodes par transparence
- visualiser une date données
- sauver la figure en *.jpeg, *.png ou *.pdf

La méthode plot a été développée par J.-M. Nocquet.

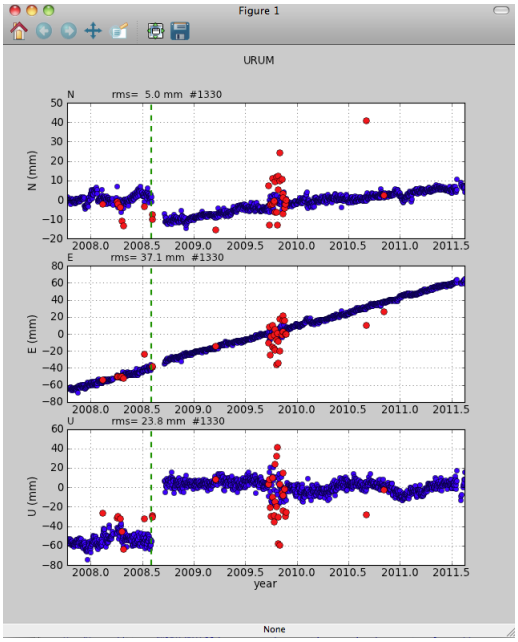


FIGURE 3.1: Exemple de visualisation d'une série temporelle, à travers la méthode plot

La figure 3.1 présente la fenêtre de visualisation de la séries temporelle du site URUM. Dans cette fenêtre, on peut effectuer des opérations telles que le zoom, l'enregistrement de la figure. Les positions sont exprimées par les cercles bleus dans le cadre avec l'axe vertical de coordonnées (ou résidus de coordonnées) et l'axe horizontal du temps. Le résultat de l'analyse de la série temporelle est aussi visualisé dans la fenêtre de visualisation, par exemple les outliers (cercles rouges), les offsets (ligne verticale verte). Plusieurs figures d'illustration dans cette thèse sont réalisées par cette méthode.

3.2.5 Le "wrapper" pygps_time_serie.py

Le "wrapper" pygps_time_series.py développé en utilisant les méthodes primitives et les méthodes de second niveau pour analyser les problèmes plus courants rencontrés dans les séries temporelles, de manière automatique ou semi-automatique un grand nombre de séries temporelles.

Les données en entrée sont les fichiers de série temporelle de positions (ou celles de résidus des positions) de station GPS. Pour l'instant, pour les résultats produits par Gamit, le "wrapper" (programme) pygps_time_series.py utilise le format mb_file de Gamit (http://www-gpsg.mit.edu/~tah/GGMatlab/), mais la méthode genfromtxt du module Numpy permet le chargement d'autre formats que l'utilisateur peut définir par lui même.

L'analyse de pygps_time_series.py est réalisée successivement pour chaque station du réseau. Le wrapper a 2 modes d'analyse : automatique et non-automatique. L'analyse

automatique des séries temporelles par pygps_time_series.py est montrée sur la figure 3.2.

Lors de l'analyse automatique de séries temporelles, la première étape est l'identification préliminaire de outliers. Elle est réalisée simultanément sur les séries temporelles brutes des 3 composantes e, n, u de positions du point, suivant l'algorithme dans le paragraphe 2.3.2.3. Le programme identifie ensuite des offsets suivant la méthode présentée sur le paragraphe 2.3.2.4. En fin de cette étape, le test F-ratio est réalisé pour tous les offsets détectés automatiquement. Il permet d'éliminer les offsets non significatifs.

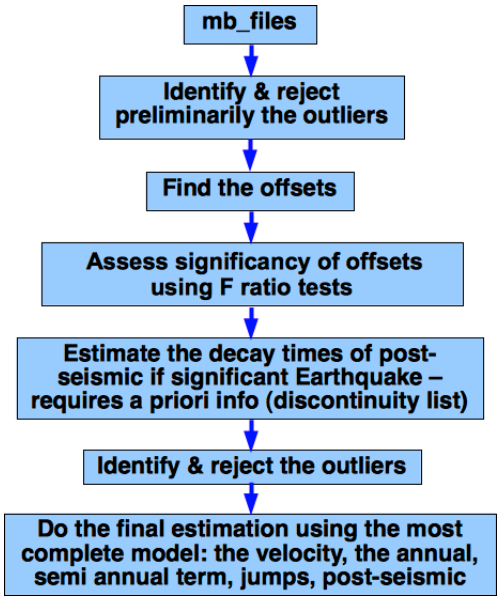


FIGURE 3.2: Diagramme des opérations élémentaires pour l'analyse automatique des séries temporelles GPS

La deuxième étape est l'estimation des temps de relaxation post-sismique selon l'algorithme décrite dans le paragraphe 2.3.3. Pour cela, il est obligé de fournir les époques des séismes. Ces informations a prioris sont automatiquement prises en compte à partir de la liste de discontinuités (produite par IGS sur `ftp://igs-rf.ign.fr/pub/IGS08/soln_IGS08.snx` ou ITRF sur `http://itrf.ensg.ign.fr/ITRF_solutions/2008/computation_strategy.php?page=2`) ou un fichier de discontinuités crée au format :

```
CHEN   1999.203   offset
CHEN   2010.170   Earthquake
 ...
```

Chaque ligne contient 3 colonnes correspondant au nom de station, l'époque et la raison de discontinuité.

Le programme estimera le temps de relaxation post-sismique pour la discontinuité s'il rencontre le mot-clé "Earthquake".

S'il ne se trouve pas les informations sismiques pour la série à analyser, la deuxième étape est négligée, le programme réalise tout de suite l'étape suivante.

Lors de la troisième étape, le programme re-identifie finalement les outliers avec le modèle complet et toutes les données de la séries temporelles.

La dernière étape est l'estimation finale sur les séries temporelles nettoyées. A cette étape la vitesse, les termes annuels et semi-annuels, les offsets et les déformations post-sismiques sont estimés simultanément par moindres carrés.

Par ailleurs, le programme a aussi un mode "non-automatic" d'analyse qui permet d'analyser les séries temporelles à partir des informations fournies par l'utilisateur, qui sont les époques des offsets et les époques de séismes dans le fichier de discontinuités pygps. Dans ce mode, le programme ne réalise que 3 étapes : le rejet des outliers passés en paramètres, l'estimation du temps de relaxation post-sismique si on fournit l'époque de séisme et l'estimation finale pour la vitesse, les terme annuels et semi-annuels, les offsets, les déformations post-sismiques.

Les résultats obtenus de l'analyse des séries temporelles d'un point sont : la vitesse, les termes annuels et semi-annuels, les amplitudes des offsets, les amplitudes des mouvements post-sismiques, les temps de relaxations post-sismiques, la série temporelle nettoyée des outliers, la série temporelle du modèle, la série temporelle de résidus de positions.

3.3 Test de validité

Plusieurs tests sur séries réelles ont été réalisés. Nous constatons que la plupart des séries temporelles est bien analysée. C'est en particulier le cas pour les séries temporelles ne comportant que des outliers et des offsets. Les cas plus complexes fournissent des résultats plus mitigés, en particulier dans le cas de successions de déplacements co-sismiques/post-sismiques.

Le temps de calcul requis dépend de la longueur de la série temporelle, mais aussi de sa complexité. Par exemple pour l'analyse du réseau de la Solution Geoazur I - RENAG (102 stations) dans la période 2006 - 2010, qui ne comporte que des outliers ou des offsets, le temps moyen d'analyse de série temporelle d'une station est 5 seconds d'environs sur un Macintosh avec un processeur Intel Core 2 Duo cadencé à 2.1 Ghz. En revanche, pour le réseau GEONET - Japon, un réseau des séries temporelles complexes et bruitées le temps moyen d'analyse d'une série temporelle dans la période 1997 - 2011 est plus long 30 secondes d'environs.

Cependant, même si l'algorithme entièrement automatique ne résout que 70% des problèmes, il s'agit déjà d'une aide précieuse pour l'analyse. Je présente dans la suite de ce

paragraphe, quelques exemples obtenus avec les paramètres standards optimaux que nous avons déterminés.

Le premier cas est l'analyse de série temporelle du site PQRL de la solution Geoazur I - RENAG (figure 3.3). Lors d'analyse, pygps_time_series.py a rejeté 72 outliers et détecté 2 offsets (lignes verticales rouges) qui sont clairement visibles sur la composante Up.

Composante	Vitesse (mm/an)	Terme annuel		Terme semi-annuel		offset (m)		wrms (mm)
		A(mm)	ϕ(an)	A(mm)	ϕ(an)	2008.9	2010.1	
N	15.40 ± 0.05	0.57	-0.28	0.14	-0.15	0.004	0.001	1.2
E	20.88 ± 0.05	1.44	0.17	0.27	-0.03	-0.003	0.003	1.2
U	-0.28 ± 0.16	0.88	0.27	0.48	0.16	0.121	-0.116	3.9

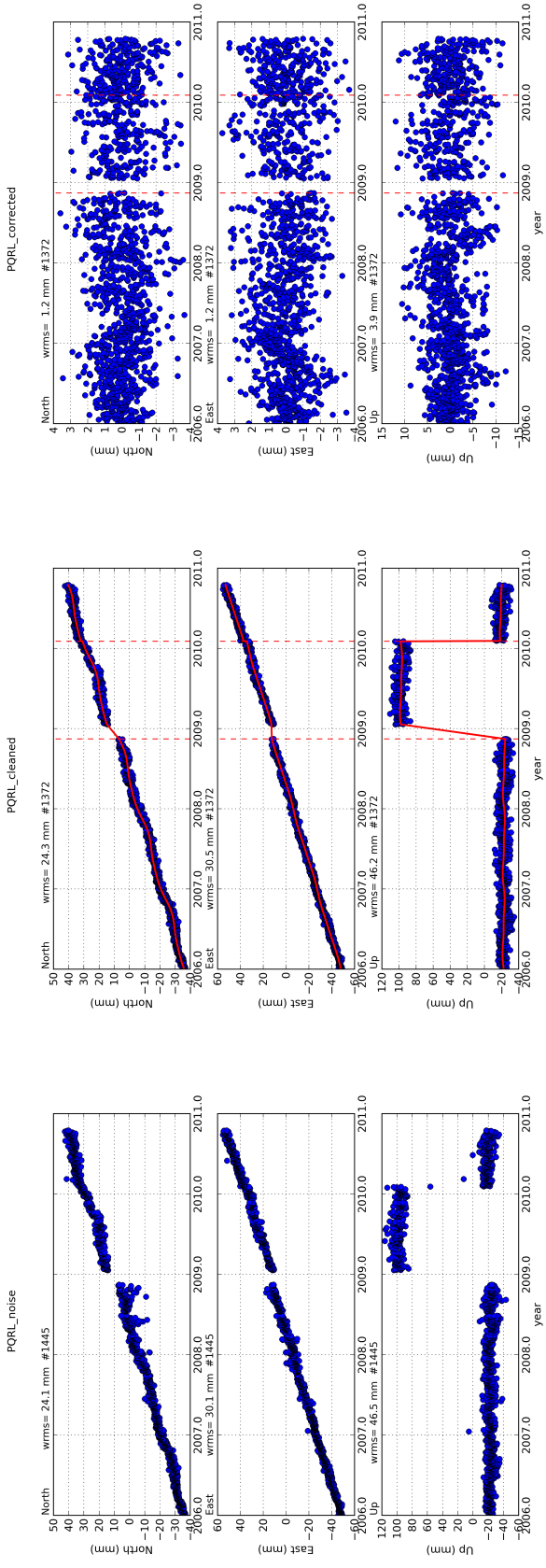
Table 3.1: Paramères décrivant la série temporelle du site PQRL (Géoazur I - RENAG)

On voit clairement que les outliers sont bien identifiés et rejetés (figure 3.3b) ; deux offsets sont bien détectés (figure 3.3b) et corrigés (figure 3.3c). On obtient une répétabilité finale de la série temporelle de résidus de 1.2 mm pour la composante Nord et Est, 3.9 mm pour la composante verticale. Les vitesses sur les 3 composantes indiquent des précisions très élevées, mais qui ne sont pas réelles car la corrélation temporelle est négligée (voir le tableau 3.1).

Le deuxième cas typique est la série temporelle du site GS32 du réseau GPS permanent Taiwan (figure 3.4). Outre les outliers visibles, on peut voir un offset induit par le séisme de Kaohsiung au 4 Mars 2010 (Mw = 6.7) et le signal post-sismique qui suit.

pygps_time_series.py a identifié et rejeté 50 outliers sur les 3 composantes ; il a aussi automatiquement détecté la date et l'offset correspondant au séisme (ligne verticale rouge sur la figure 3.4b). La série temporelles des résidus indiquent que les séries temporelles brutes sont bien corrigées du déplacement co-sismique et de la relaxation post-sismique (figure 3.4c).

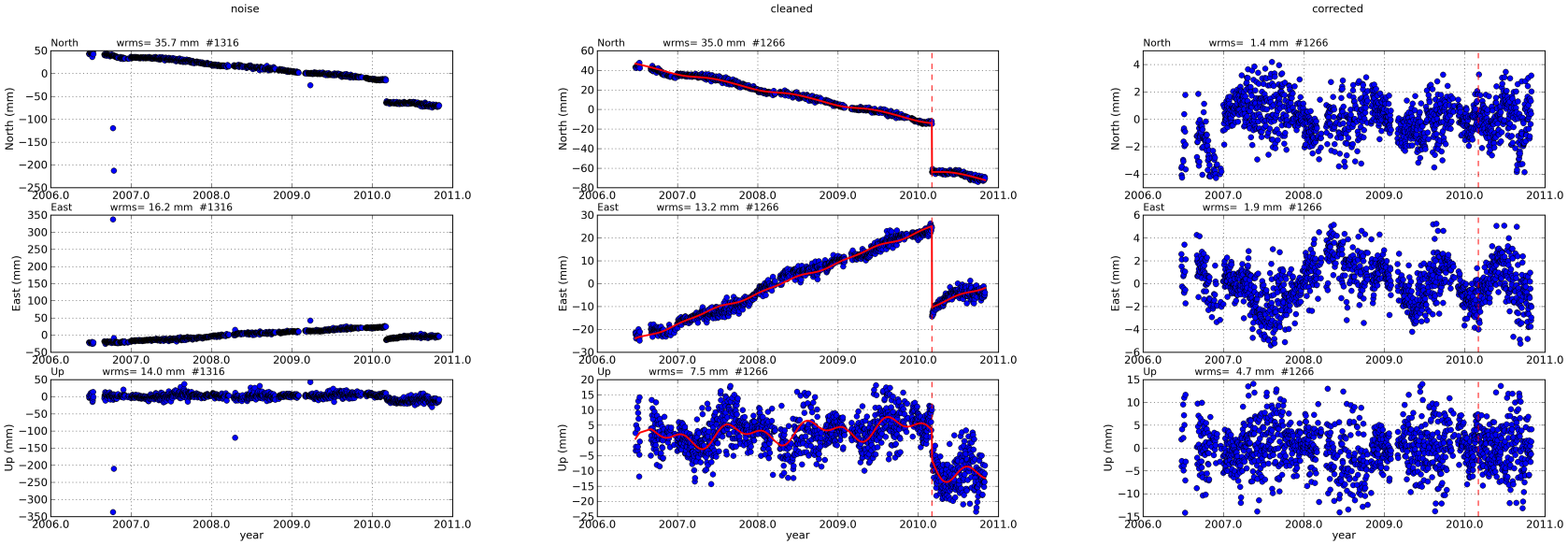
Pour estimer le temps de relaxation τ du post-séisme qui suit le co-séisme Kaohsiung, j'ai fourni l'époque du séisme dans le fichier de discontinuités en format pygps (la même époque trouvée automatique). Lorsque l'analyse, le τ estimé sur composante horizontale est de 0.211 an. On voir les paramètres d'estimation finale d'analyse sur le tableau 3.2.



(a) (b) (c)

(a) série brute, wrms Nord/Est/Vertical 24.1/30.1/46.5 mm
(b) série nettoyée, wrms Nord/Est/Vertical 24.3/30.5/46.2 mm, et son modèle (ligne rouge)
(c) série de résidus, wrms Nord/Est/Vertical 1.2/1.2/3.9 mm

Figure 3.3: Série temporelle du site PQRL (Géoazur I - RENAG)



(a) (b) (c)

a : série brute, wrms Nord/Est/Vertical 35.7/16.2/14.0 mm
b : série nettoyée, wrms Nord/Est/Vertical 35.0/13.2/7.5 mm, et son modèle (ligne rouge)
c : série de résidus, wrms Nord/Est/Vertical 1.4/1.9/4.7 mm

Figure 3.4: Série temporelle du site GS32 (réseau GPS Taiwan)

Comp-	Vitesse (mm/an)	Terme annuel		Terme semi-annuel		offset (m)	Post-séisme		wrms (mm)
osante		A(mm)	ϕ(an)	A(mm)	ϕ(an)		A(m)	Tau(an)	
N	-15.85 ± 0.06	0.55	0.38	0.33	0.03	-0.050	-0.002	0.211	1.4
E	13.47 ± 0.04	1.41	-0.49	0.14	-0.16	-0.036	-0.000	0.211	1.9
U	1.21 ± 0.14	2.53	-0.26	1.96	0.08	-0.010	0.009	#	4.7

Table 3.2: Paramètres décrivant la série temporelle du site GS32 (réseau GPS Taiwan)

Comme la série temporelle du site PQRL, on peut dire que cette série est bien analysée (voir la figure figure 3.4c, le tableau 3.2) avec une répétabilité de 1.4 mm pour la composante Nord, 1.9 mm pour la composante Est et 4.7 mm pour la verticale.

Limitations et discussion

Au cours du développement des outils pygps_time_series.py, nous avons testé nos approches sur un très grand nombre de séries temporelles réelles provenant des réseaux EUREF, RENAG, GEONET - Japon, Taiwan GPS Network. Ce programme les analyse relativement bien (environ 80% des cas) et très rapidement. Le 20% des cas où des problèmes persistent sont des séries temporelles particulièrement complexes. Les principales difficultés que nous avons notées sont :

- Pygps_time_series.py détecte très bien tous les grands offsets mais dans quelques cas, les petits offsets ne sont pas détectés. C'est surtout le cas pour des séries temporelles bruitées.
- L'estimation du temps de relaxation post-sismique τ ne converge pas dans quelques cas. Elle peut aussi converger, le τ estimé ne se semble pas correct. Une des raisons possible est la corrélation entre τ et le changement de vitesse lorsque celui-ci est estimé. Ce problème est déjà souligné par Nikolaidis [2002].
- Nous n'étudions pas encore le cas de mouvement transitoire tels que des séismes lents. Une perspective possible à mettre en oeuvre serait d'indiquer les périodes des séries temporelles où aucuns des algorithmes testés ne fournit de "bon" résidus.

En conclusion, les difficultés que l'on rencontre principalement est l'estimation de τ sur les séries temporelles complexes, par exemple le réseau GEONET - Japon. Dans ces cas là, il est nécessaire de vérifier et regarder les séries temporelles avant et après ses analyses.

3.4 Conclusions

Nous avons présenté des méthodes d'analyse rapide et robuste des séries temporelles que nous avons implémenté dans une série d'outils. La Classe Gts (Geodetic time series)

contient les méthodes primitives et de second niveau, la méthode de visualisation plot qui nous donne une utilisation souple dans un mode entièrement manuel et certaines manipulations sur les séries temporelles (l'aide de la Classe Gts est montrée dans l'annexe C). Le programme pygps_time_series.py nous permet d'analyser automatiquement et semi-automatiquement les problèmes des outliers, des offsets et des déformations post-sismiques sur un grand nombre de séries temporelles.

Les difficultés rencontrées en analyse des séries temporelles de positions GPS sont que :

- Quelques petits offsets ne sont pas automatiquement, surtout les petites offsets dans les séries complexes. Dans ce cas-là, on doit ajuster les paramètres spécifiques en analysant les séries.
- L'analyse de la déformation dans les séries temporelles requiert les époques co-sismiques. Pour les sites dont l'époque co-sismique n'est pas trouvée dans la liste de discontinuités IGS, il faut regarder ses séries temporelles et/ou prendre par exemple le fichier provenant de www.globalcmt.org pour ajouter les informations manquantes dans la liste de discontinuités en format pygps. Cela peut prendre du temps, surtout les réseaux dans des zones très sismiques telle que le réseau GEONET - Japon.

Malgré ces défauts, les résultats obtenus en test restent satisfaisant. Nous présentons dans le chapitre suivant comment nous avons utilisé ces outils pour l'analyse de deux cas concrets.

CHAPITRE 4

ETUDES DE CAS : LES SOLUTIONS DES RÉSEAUX RENAG ET TAIWAN

4.1 La solution RENAG

4.1.1 Présentation du réseau RENAG

Le réseau RENAG (Réseau National GPS permanent, http://www.renag.fr, https://gpscope.dt.insu.cnrs.fr/spip/spip.php?rubrique51) est le réseau GPS permanent français pour les applications scientifiques. Il est issu du réseau REGAL [Calais et al., 2000] installé à partir de 1997 pour étudier la déformation tectonique dans les Alpes occidentales. Ce réseau, étendu à l'ensemble du territoire métropolitain est devenu Service d'Observation de l'Institut National des Sciences de l'Univers (INSU) du CNRS en 2006 et est aujourd'hui une composante du Service d'Observation (SO) Géodésie-Gravimétrie.

Ses objectifs scientifiques sont (1) la mesure de la faible déformation tectonique en France métropolitaine (2) l'étude du contenu en vapeur d'eau de l'atmosphère pour les applications météorologiques (3) l'étude des variations du niveau de la mer par co-localisation de stations GPS et marégraphes (4) l'étude des phénomènes de surcharge et de l'interaction entre la Terre solide et les enveloppes superficielles. De manière plus marginale, les données du réseau RENAG servent comme station de référence pour les étude d'instabilités de versants, les systèmes de référence à travers la contribution de RENAG au Réseau GPS permanent RGP géré par l'institut géographique national (IGN) et le réseau européen EUREF, la détection des ondes de surface des grands séismes ou l'étude des variations du contenu total électronique (TEC) de l'ionosphère.

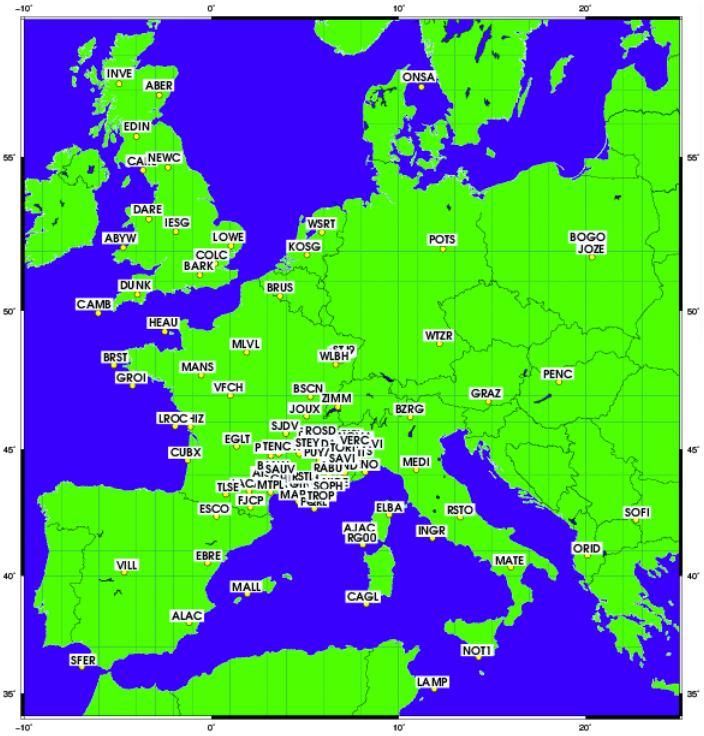


FIGURE 4.1: Carte du réseau RENAG

Le réseau RENAG comporte une soixantaine de stations gérées par une dizaine de laboratoires en France. Le centre de données national, hébergé à l'Observatoire de la Côte d'Azur assure la collecte, le contrôle-qualité, la diffusion, l'archivage, le suivi des méta-données et le calcul régulier des solutions journalières du réseau. Les données du réseau RENAG sont accessibles gratuitement et publiquement sur le site `ftp://renag.unice.fr`. Par ailleurs, à côté du réseau RENAG, le réseau RGP (`http://rgp.ign.fr`) est le réseau GPS permanent géré par l'IGN aussi sur la base du regroupement de plusieurs fournisseurs de données, comme en particulier le réseau Teria de l'ordre des géomètres-experts, comportant environ 200 stations sur le territoire français. Enfin, un partenariat entre l'INSU et la société Geodata Diffusion, permet d'avoir accès aux données du réseau privé Orphéon (`http://reseau-orpheon.fr/la-societe-geodata-diffusion`).

4.1.2 Présentation de la solution étudiée

La solution étudiée dans la suite de cette section est issue de l'analyse réalisée avec le logiciel Gamit/Globk développé par le MIT (Massachusetts Institute of Technology). En plus des sites propres du réseau RENAG, cette solution inclut des sites du RGP, ainsi que des sites du réseau de la région Piémont en Italie (`http://gnss.regione.piemonte.it`) qui permettent d'encadrer les Alpes occidentales. Enfin, il inclut des sites fournit par le

centre de données BIGF (http://www.bigf.ac.uk/) pour le Royaume-Uni. La solution étudiée s'étend sur la période 1998.0-2010.9 soit 4678 jours et comprend 102 stations.

La stratégie de calcul est celle classiquement utilisée pour les calculs à but géodynamique. Les positions journalières sont estimées par sessions de 24 heures. Le logiciel Gamit utilisent les doubles-différences de la combinaison linéaire L3 permettant de supprimer le premier ordre des retards ionosphériques pour estimer simultanément les coordonnées des sites du réseau. Dans ce traitement, les orbites issues de la campagne de reprocessing IGS (http://acc.igs.org/reprocess.html) sont fixées dans le calcul. La troposphère est modélisée suivant la stratégie décrite par Tregoning and Watson [2012]. Les effets de la surcharge océanique sont corrigées suivant les prédiction du modèle FES2004 [Lyard et al., 2006] et les variations du centre de phase des antennes sont corrigées en utilisant les fichiers de calibration absolue recommandées par l'IGS [Schmid et al., 2007]. Le calcul Gamit suit la stratégie décrite par Feigl et al. [1993], où la solution finale, une fois les ambiguïtés fixées, est obtenue en relâchant les contraintes sur tous les paramètres estimés. A l'issue du calcul Gamit, on obtient des solutions journalières quasi-libres ("loosely constrained"), c'est à dire des solutions pour lesquelles les coordonnées relatives sont bien déterminées à partir des doubles-différences de phase, mais pour lesquelles les coordonnées absolues sont approximatives. On appelle "libres" ces solutions dans la suite du document.

4.1.3 Expression des solutions libres issues de Gamit dans l'ITRF2008

Les séries temporelles ont été obtenues par le module pyacs_make_time_series.py en estimant une transformation à 7 paramètres entre la solution journalière libre issue de Gamit et la solution combinée cumulée produite par le centre IGN de combinaison IGS, mise à jour au 13 Aout 2012, IGS12P33.SNX. Le fichier de discontinuités associé à été pris en compte. La solution inclut 4678 jours, 102 stations soit une moyenne de 12 années par stations. A titre d'illustration des performances, la réalisation des séries temporelles pour l'ensemble des sites a pris environ 12 minutes sur un iMac avec un processeur Intel Core 2 Duo cadencé à 2.66 Ghz.

La solution Gamit inclut une moyenne de 29 sites avec la solution de l'IGS pendant la période considérée. Le niveau d'accord moyen entre les solutions journalières libres et la solution IGS12P33.SNX est de 1.8 mm pour la composante Est, 1.0 mm pour la composante Nord et 3.2 mm pour le vertical. Sans l'estimation avec la norme L1 et la détection d'erreur, on obtient 3.3 mm pour la composante Est, 1.9 mm pour la composante Nord et 11.84 mm pour le vertical. La différence entre les deux provient généralement d'un ou plusieurs sites de la solution IGS12P33 qui présentent un niveau de désaccord avec la position estimée par Gamit. Il serait difficile de séparer une erreur provenant du calcul

Gamit (par exemple une information de changement d'équipement mal pris en compte) ou dans la solution IGS. L'estimateur de Dikin permet donc de maintenir un excellent niveau de stabilité de la référence pour l'ensemble de la solution.

On regarde ensuite plus précisément la répétabilité, qui est un indicateur moyen de la dispersion des positions. Celle-ci est calculée à partir des solutions journalières en retirant de la pente (répétabilité long-terme). On obtient des répétabilités médianes du réseau de 2.1 mm pour la composante Est, 1.8 mm pour la composante Nord, 4.9 mm pour la composante verticale. On constate que la plupart des sites présentent des répétabilités inférieures à 6 mm sur les composantes horizontales et une répétabilité de l'ordre de 10 mm sur la composante verticale.

La répétabilité moyenne est de 69.6 mm pour la composante Est, 71.0 mm pour la composante Nord, 21.1 mm pour la composante verticale. Ces valeurs élevées et très différentes de la répétabilité médiane indiquent immédiatement des problèmes sur quelques sites. Par exemple, on trouve que les sites VERC, ALES et ZIMM ont des répétabilités anormalement grandes : VERC : 5173.9 mm Est, 1967.1 mm Nord, 10.0 mm pour le Vertical, ALES : 616.1 mm, 4826.4 mm, 260.9 mm et ZIMM : 833.6 mm, 131.3 mm, 970.3 mm. Elle provient des grands offsets dans la série (voir la figure 4.5). Pour VERC et ALES, il s'agit d'une ré-installation de l'antenne qui n'est pas documentée. Ce sont des sites du réseau topographique de la région Piémont en Italie. L'offset pour ZIMM est plus surprenant, cette station faisant partie du réseau IGS. Enfin, des répétabilités élevées pour des points comme AIGL (3.3 mm, 4.2 mm, 15.2 mm), MODA (25.5 mm, 6.7 mm, 27.4 mm) sont dues à des outliers ou des offsets plus petits dans la série temporelle (voir les figures 4.4, 4.5).

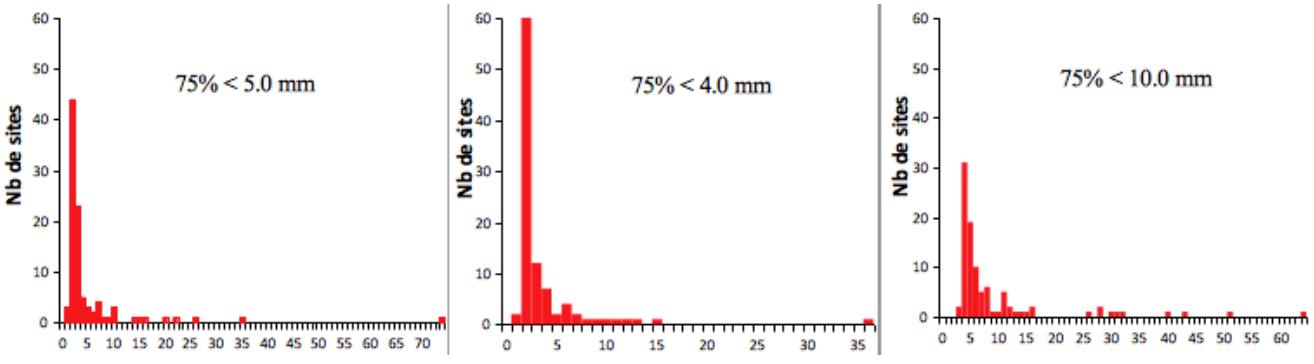


FIGURE 4.2: Histogrammes des répétabilités obtenues par pyacs_make_time_series.py pour le réseau RENAG

Les sites ALES, VERC, ZIMM dont la répétablilité trop élevée est retirée. La figure 4.2 représente l'histogramme des répétabilités du réseau obtenues par pyacs_make_time_series.py. Les points ALES, VERC et ZIMM ne sont pas inclus. La figure montre que 75% des sites

dont la répétabilité est inférieure à 5.0 mm sur l'Est, 4.0 mm sur le Nord et 10.0 mm sur le Vertical. Ces nombres, plus élevés que la précision naturelle d'un traitement GPS indiquent que plus de 25% des sites présentent des problèmes dans les séries temporelles qu'il va falloir analyser.

4.1.4 Typologie des problèmes rencontrés dans les séries temporelles

Une inspection visuelle des séries temporelles de position montrent que la plupart de ces séries présentent des outliers, soit isolés, soit groupés par périodes. Même dans les séries temporelle présentant des répétabilités bonnes (<1.5 mm sur les composantes horizontales), il existe encore des positions que nous pouvons considérer comme les outliers (4.3, les outliers sont indiqués par les cercles rouges).

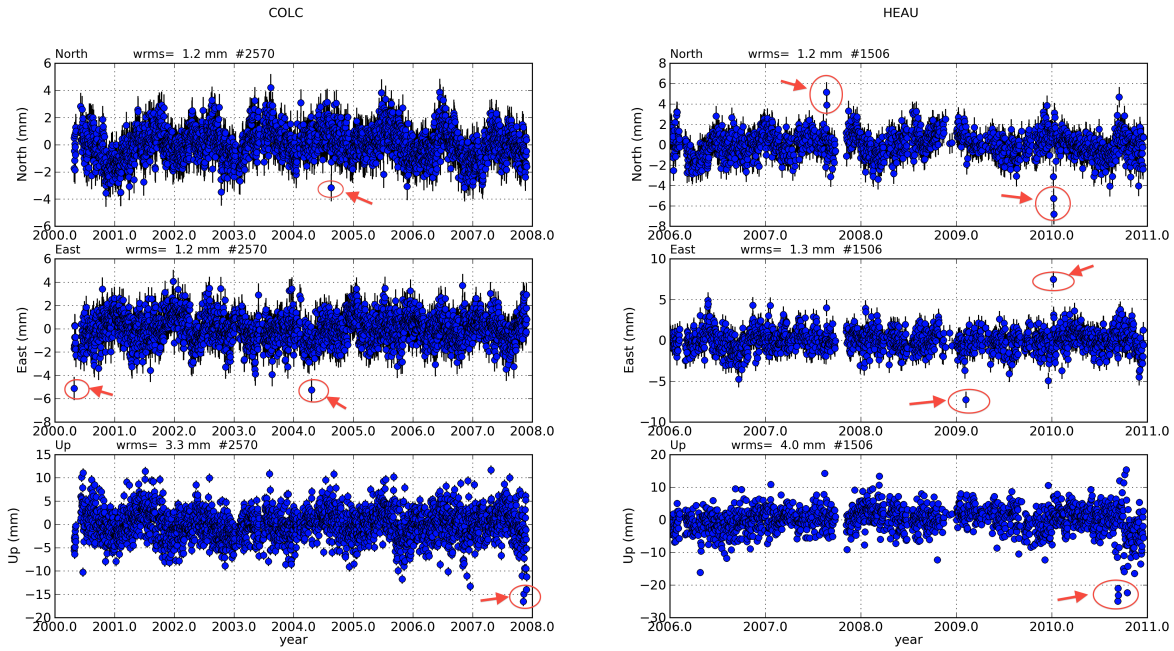


FIGURE 4.3: Série temporelle (pente retirée) des sites COLC et HEAU

Ainsi, le type de bruit le plus répandu, est la présence d'outliers dans la série temporelle. La figure suivante 4.4 montre les séries temporelles brutes de Nord, Est et Vertical des sites MODA et MEDI. La série temporelle du site MEDI (figure 4.4 droit) montre une série très bruitée, comprenant à la fois des outliers isolés et des outliers groupés par périodes. L'origine de ces outliers resterait à étudier en revenant aux résultats du calcul Gamit, mais nous évaluons ici la capacité de notre approche à traiter ces problèmes.

On note par ailleurs l'existence de variations périodiques, composées de variations annuelles et semi-annuelles (les lignes sinusoïdal en rouges dans la figure 4.4 droit). L'existence du terme semi-annuel est indiquée par l'asymétrie du signal périodique annuel. Ces termes saisonniers peuvent biaiser l'estimation de la vitesse estimée à partir de celle, surtout les séries temporelles courtes (au-deçà de 2.5 ans et jusqu'à 4.5 ans, Blewitt and Lavallée [2002]).

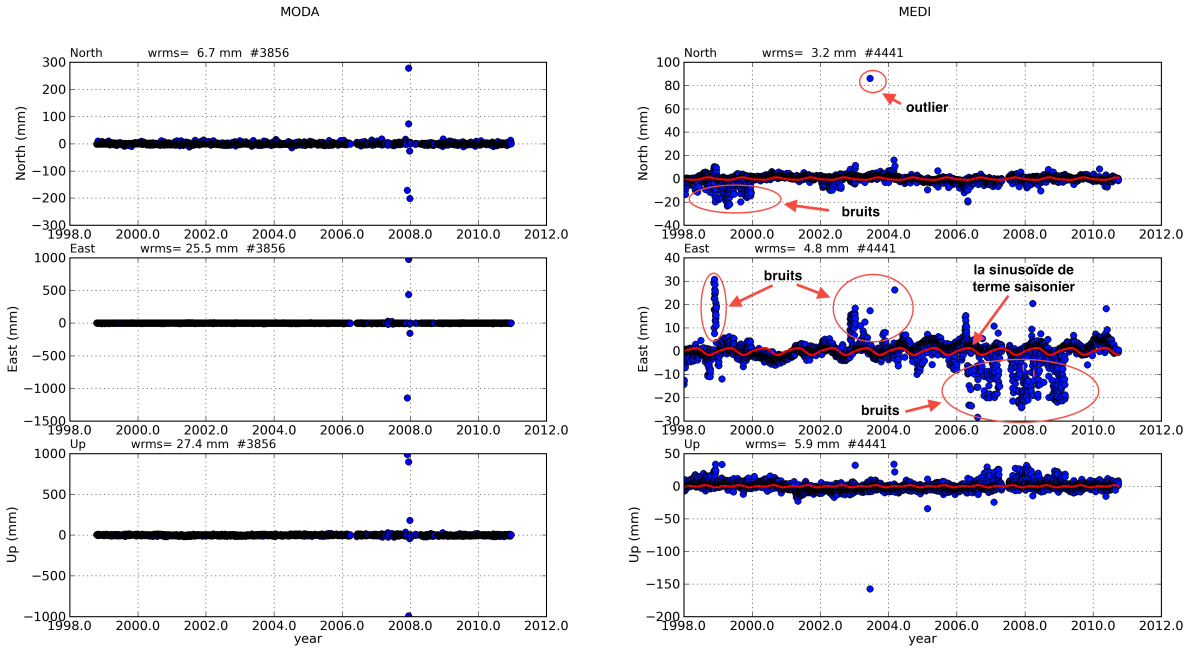


FIGURE 4.4: Exemples de séries temporelles (pente retirée) avec outliers isolés (gauche, site MODA) et groupés (droite, site MEDI).

Le troisième problème fréquent dans les séries du RENAG est la présence d'offsets. Leur fréquence varie d'une série à l'autre, quelque fois d'une composante à l'autre. Ils sont aussi assez variables en amplitude, de quelques millimètres à plusieurs centimètres. Une série caractéristique de ce problème est celle d'AXPV (figure 4.5).

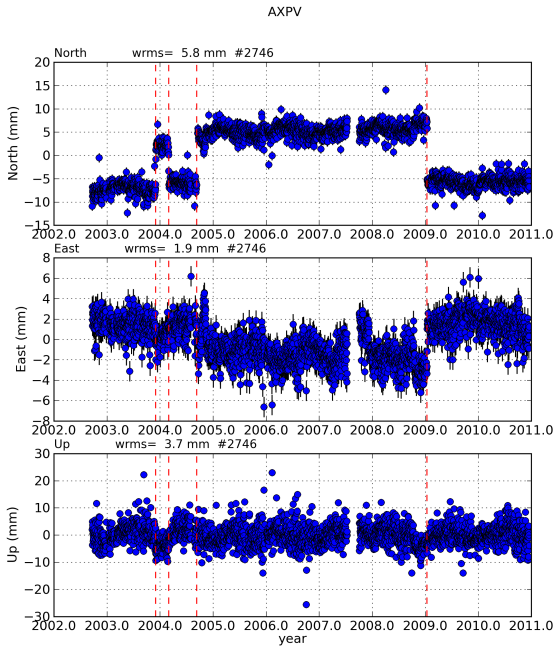


FIGURE 4.5: Série temporelle (pente retirée) avec offsets (Site AXPV)

Les problèmes rencontrés dans les séries temporelles du réseau RENAG sont classiques : outliers, offsets et des termes saisonniers. On s'intéresse maintenant à la performance de nos algorithmes pour résoudre automatiquement ces problèmes.

4.1.5 Résultat de l'analyse automatique avec les paramètres par défaut

Les séries temporelles sous la format mb_file généré par Globk ou Glred sont ensuite analysés automatiquement.

Les paramètres par défaut du programme pygps_time_series.py sont les suivants :

- nombre de positions dans une fenêtre glissante pour la détection préliminaire d'outliers et la détection d'offsets par rms : 7

Les autres paramètres choisis par défaut sont (décrits concrètement dans le paragraphe 2.3.2) :

- le seuil de détection de rms anormal : 3

- le seuil de détection d'outlier : 3

- le seuil d'outliers : 5%

- le seuil de détection d'offset : 3.

Les séries temporelles du réseau RENAG sont automatiquement analysées par le programme en utilisant les paramètres par défaut et le fichier de discontinuités IGS08. L'analyse automatique a pris environ 31 minutes sur le même iMac dans 4.1.3.

Pour une évaluation globale des performances, on regarde l'histogramme des répétabilités des stations (figure 4.6). On voit qu'après ce traitement, 95% des sites ont une répétabilité inférieure à 2.2 mm sur la composante Est, 2.0 mm sur la composante Nord et 5.5 mm sur le Vertical. Par rapport aux histogrammes des répétabilités donnés par pyacs_make_time_series.py (4.2), on note donc une amélioration très importante, pour seulement 30 minutes de traitement supplémentaire !

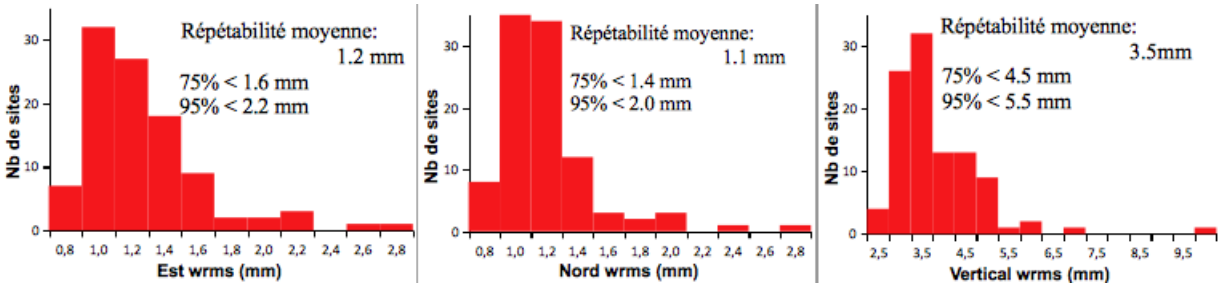
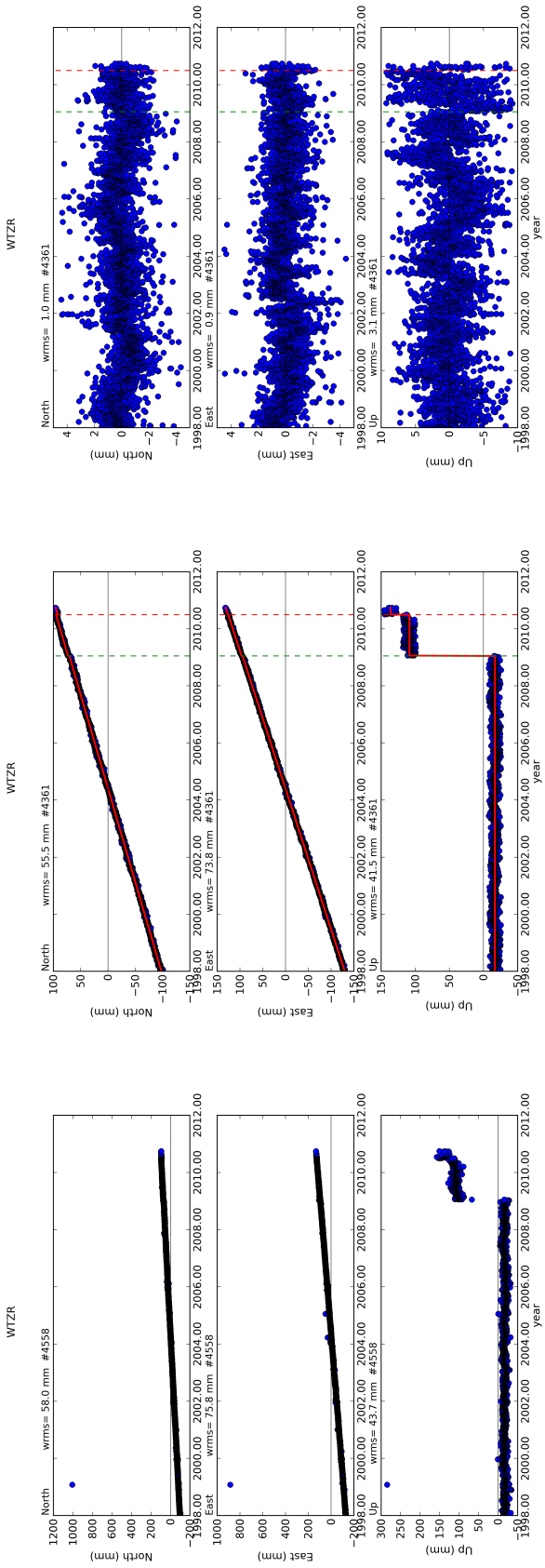


FIGURE 4.6: Histogrammes des répétabilités, après traitement par pygps_time_series.py

La figure 4.7 montre un cas typique du comportement de mes algorithmes avec les paramètres par défaut pour un site du réseau global IGS WTZR (Wettzel, Allemagne). On voit que les outliers les plus importants sont bien détectés et rejetés. Les deux offsets sont aussi correctement détectés. Le second est documenté dans le fichier de discontinuités IGS08 (ligne verticale rouge sur la figure 4.7b) et la date correcte a été retrouvée par mon programme. La figure 4.7b est la séries des résidus. La répétabilité est maintenant de 0.9 mm pour la composante Est, 1.0 mm pour la composante Nord, 3.1 mm pour la composante verticale, ce qui est un très bon résultat pour des séries temporelles journalières. La figure 4.7b suggère pourtant que d'autres outliers pourraient être détectés et que les paramètres par défaut que j'ai proposés restent relativement "conservateurs".

Sur les résultats de l'analyse des séries temporelle du réseau RENAG, en regardant les séries temporelles de résidus de positions par l'outil de visualisation plot de la Classe Gts, nous évaluons que le taux de réussite de l'analyse automatique des séries temporelles du réseau avec les paramètres par défaut est d'environ 85%-90%. Les 10-15% de problèmes restants se trouvent dans les séries temporelles complexes, en particulier celles présentant des outliers groupés (par exemple le site MEDI, figure 4.4 droite) ou des petits offsets (BRST, BRUS, BSCN, PUYV,...).

Cependant, nous pouvons conclure que l'analyse automatique fonctionne très bien dans le cas-là du réseau RENAG qui présente un grand nombre de séries temporelles classiques.



(a) (b) (c)

(a) Série temporelle brute
(b) Série temporelle nettoyée et son modèle (ligne rouge)
Un offset détecté automatiquement (ligne verticale verte), un offset détecté dans le fichier de discontinuités IGS (ligne verticale rouge).
(c) Série de résidus

FIGURE 4.7: Analyse des séries temporelles du site WTZR

4.1.6 Ajustement des paramètres de l'analyse automatique

La re-analyse des séries temporelles où des problèmes persistent requiert l'ajustement de paramètres du programme pygps_time_series.py. Pour le cas des outliers restants, la modification la plus simple a été simplement d'ajuster le seuil de outliers, en le passant de 5% à 10% (MEDI). Une inspection visuelle permet d'ajuster ce paramètre.

Pour le cas de petits offsets que le programme ne détecte pas en utilisant les paramètres par défaut, il faut ajuster le seuil de détection de rms anormal, ou le seuil de détection d'offset, ou les deux seuils à la fois. Dans le cas de la série temporelle du site PUYV (figure 4.8a) a un offset visible sur la composante verticale, causé par le changement d'antenne, que le programme ne détecte pas avec les paramètres par défaut (figure 4.8b). En re-analyse de cette série temporelle, j'ai ajusté la paramètre de seuil de rms anormal de 3 (par défaut) à 2, il a donc bien identifié l'offset (ligne vert à la figure 4.8c) et le rejette correctement de l'analyse.

4.1.7 Résultat final

En regardant les séries temporelles finales obtenues, nous constatons qu'elles ont une très bonne qualité, indiquée par la répétabilité moyenne de 1.2 mm pour la composante Est, 1.1 mm pour la composante Nord et 3.4 mm pour la composante verticale. Après ajustement de paramètres spécifique pour les séries temporelles présentant encore des problèmes, nous pouvons aboutir à une solution de très bonne qualité sans intervention manuelle.

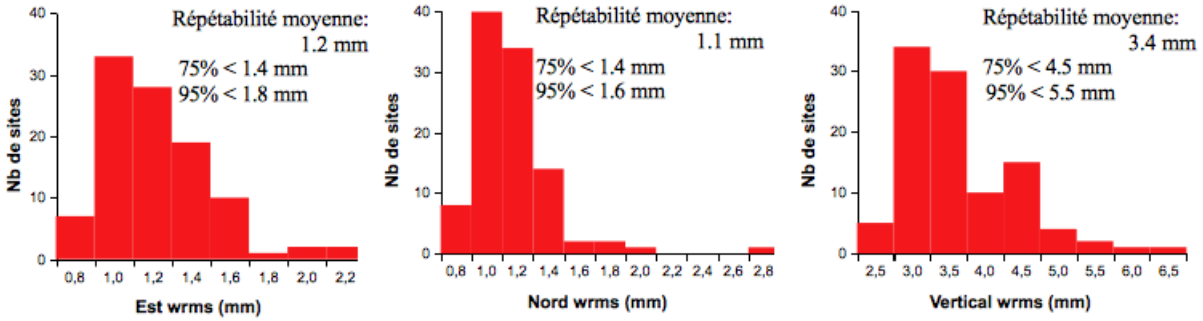
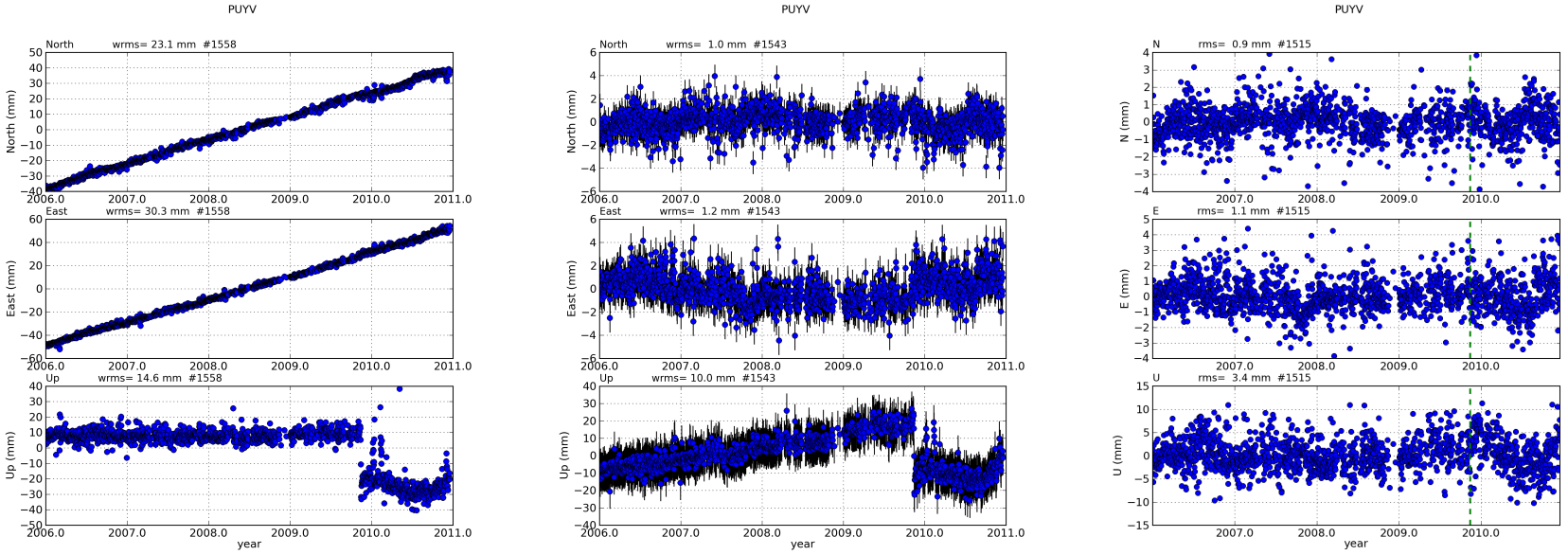


FIGURE 4.9: Histogrammes des répétabilités issues du traitement final pour le réseau RENAG

La figure 4.9 présente l'histogramme des répétabilités finales. Il montre que 95% des sites ont une répétabilité inférieure à 1.8 mm sur l'Est et 1.6 mm sur le Nord. La répétabilité sur la composante Verticale n'est pas améliorée, ce qui suggère que la répétabilité obtenue avec les paramètres par défaut était déjà proche de la variabilité intrinsèque à cette composante.



(a) (b) (c)

(a) Série temporelle brute
(b) Série temporelle corrigée en utilisant des paramètres par défaut, le offset n'est pas détecté. La vitesse a été retirée
(c) Série temporelle corrigée en ajustant des paramètres. La vitesse a été retirée

FIGURE 4.8: Re-analyse de la série temporelle du site PUYV

En utilisant les paramètres par défaut, le programme pygps_time_series.py n'identifie pas l'offset dans la série de la composante verticale (4.8b). En ajustant la paramètre de seuil de rms anormal de 3 à 2, il identifie et corrige bien cet offset (4.8c).

4.1.8 Analyse rapide de la qualité du champ de vitesses obtenus

Vitesses horizontales

L'analyse précédente nous permet d'obtenir les vitesses estimées dans l'ITRF2008 à partir ces séries temporelles de très bonne qualité (tableaux A.1, A.2 et A.3). Pour évaluer la qualité du champ de vitesse obtenu, nous calculons un pôle de rotation pour l'Europe stable, telle que définie dans Nocquet [2012]. Nous utilisons 28 sites de la solution Géoazur-I- RENAG. Le rms obtenu est de 0.19 mm/an avec des vitesses résiduelles maximum de 0.7 mm/an (tableau A.4). En ne traitant que les Alpes et leur pourtour immédiat, le résultat obtenu est un wrms de 0.14 mm/an.

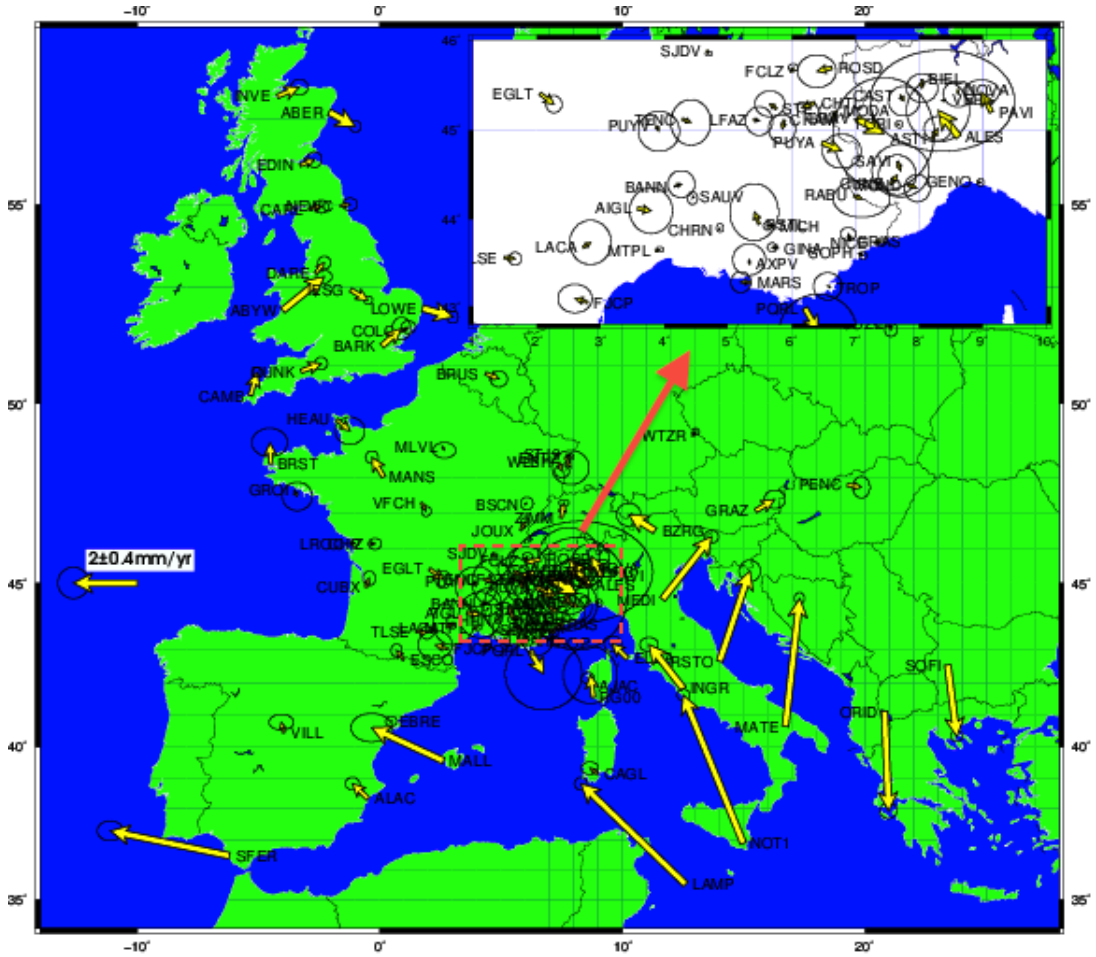


FIGURE 4.10: Champ de vitesses horizontal résiduel du réseau RENAG (après retrait du pôle d'Euler)

Dans l'état actuel des méthodes développées au cours de ma thèse, aucune analyse sur les incertitudes associées aux vitesses n'est faite, ce qui empêche toute analyse statistique sur la significativité des vitesses résiduelles obtenues. Cependant, on voit que la qualité du

résultat obtenu avec une analyse rapide est équivalente à celle obtenue avec des méthodes plus complètes et plus complexes, par exemple utilisées dans Nocquet [2012].

Vitesses verticales

La cartes des vitesses verticales du réseau RENAG dans l'ITRF2008 est présentée dans la figure 4.11 et le tableau A.5.

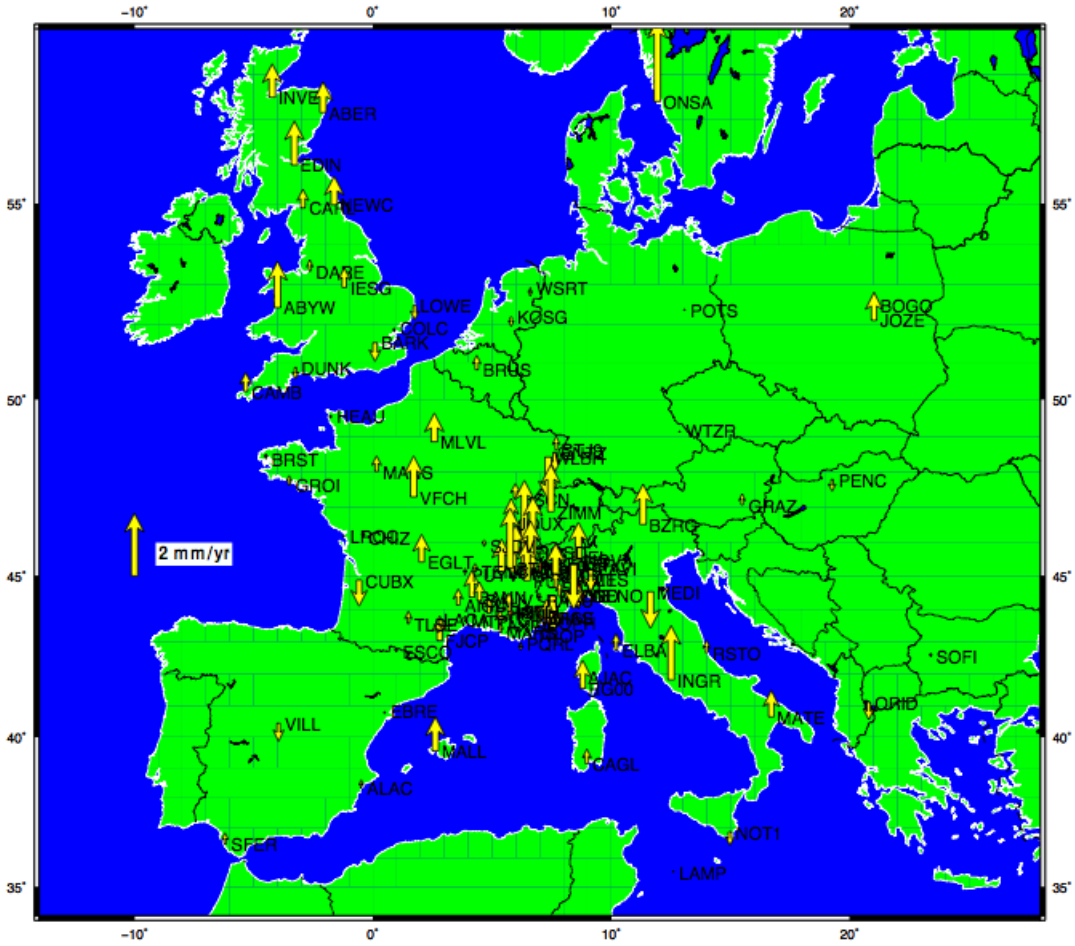


FIGURE 4.11: Champ de la vitesses vertical du réseau RENAG dans l'ITRF2008

Les extremas trouvés sont de -1.43 mm/an (VERC) et +2.71 mm/an (CHTL). La figure met en évidence plusieurs tendances. Tout d'abord, des vitesses verticales très faibles sont trouvées en Europe centrales. Les vitesses verticales positives (surrection) sont trouvée en France, les Alpes et le sud de l'Europe, ainsi que dans le nord de l'Angleterre et l'Ecosse. Enfin, les Alpes présentent une surrection différentielle par rapport à leur périphérie. Sans entrer dans la discussion sur les processus responsables de ces vitesses verticales, on peut dire que la cohérence des vitesses obtenues est aussi un indicateur indirect de la qualité du champ de vitesse obtenu. Ce champ est de plus cohérent avec les résultats obtenu par Nocquet [2011] avec une analyse plus complexe.

Dans le cas du réseau RENAG, les problèmes qui affectent la détermination de la vitesse verticale des sites sont les changements d'antenne qui génèrent des offsets, parfois petits dans les séries temporelles et les mouvements de termes saisonniers qui sont plus difficiles à estimer que sur les composantes horizontales en raison du bruit propre de la composante verticale.

4.1.9 Conclusions

L'analyse du réseau RENAG de 102 sites GPS permanents avec une moyenne de 12 ans de données par site est réalisée par notre programme en trois étapes : (1) l'expression des solutions libres issues de Gamit dans l'ITRF2008 en utilisant la solution de référence IGS12P33.SNX, (2) l'analyse des problèmes dans les séries temporelles en mode automatique avec les paramètres par défaut du programme et (3) la re-analyse des séries temporelles présentant des problèmes résiduels en ajustant les paramètres du programme.

A titre d'illustration des performances, cette analyse prend environ 12 minutes pour la première étape, 30 minutes pour la deuxième étape, sur un iMac avec un processeur Intel Core 2 Duo cadencé à 2.66 Ghz. Le taux de réussite de l'analyse automatique est environ 85% des cas du réseau RENAG.

Dans son utilisation pratique, le module de visualisation "plot" des séries temporelles est très pratique, en particulier pour vérifier les choix et les corrections estimées par l'analyse automatique.

Les résultats de l'analyse du réseau RENAG montrent des vitesses résiduelles horizontales faibles lors de l'estimation d'un pôle de rotation définie à partir de 26 sites en Europe. Le champ de vitesses verticales montrent aussi des vitesses verticales faibles de -1.43 mm/an à 2.71 mm/an dans l'ITRF2008, mais présentant une cohérence spatiale.

4.2 La solution Réseau GPS du Taiwan

4.2.1 Présentation

Le réseau GPS du Taiwan a été développé progressivement à partir de 1993, avec tout d'abord 8 stations installées pour le repère de référence par le Ministère de l'Intérieur du Taiwan (MOI). Après le tremblement de terre destructeur de ChiChi (Mw=7.6) du 21 Septembre 1999, plusieurs stations ont été mises en place par le MOI, l'Academia Sinica, le Centre de service météorologique et le Centre de la Géologie [Yeh et al., 2007]. Aujourd'hui, il comporte plus 400 stations avec une couverture spatiale dense à Taiwan.

Les objectifs du réseau GPS du Taiwan sont la mesure de la déformation crustale et l'activité sismique au Taiwan, le repère de référence et les études météorologiques. Ce réseau est traité dans le cadre de la thèse de Lin Kuan-Chuan avec lequel j'ai collaboré pour cette étude.

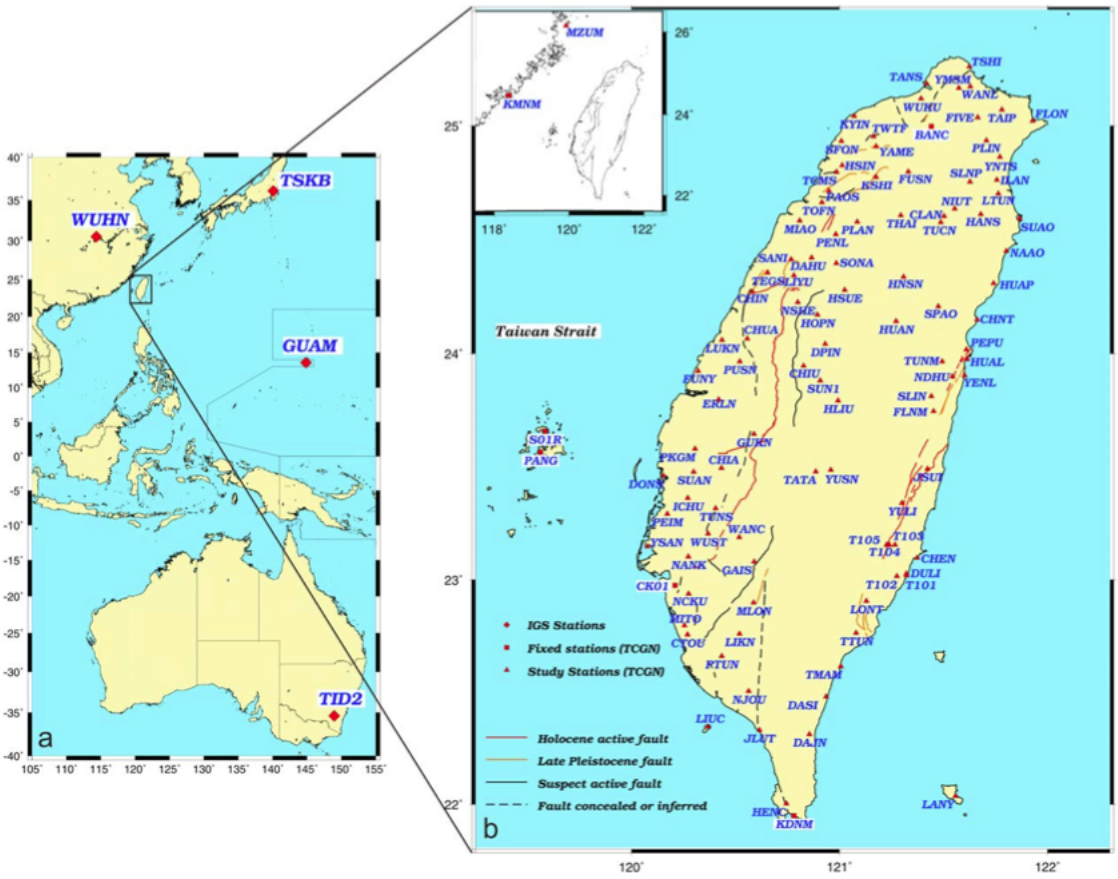


FIGURE 4.12: Cartes des 114 sites du réseau GPS du Taiwan [Lin et al., 2010]

4.2.2 Expression des solutions libres dans l'ITRF2008

Dans l'étude de cas du réseau GPS du Taiwan, nous continuons à utiliser la solution de référence IGS12P33.SNX et le fichier de discontinuité associé pour exprimer les solutions libres du réseau dans l'ITRF2008 par pyacs_make_time_series.py.

La solution journalière libre issue de Gamit, réalisé par Lin Kuan Chuan dans le carde de sa thèse, inclut 3986 jours, 424 stations soit une moyenne de 7 années par stations. La réalisation des séries temporelles par pyacs_make_time_series.py pour l'ensemble des sites, à titre d'illustration des performances, a pris environ 9 minutes sur un Macintosh avec un processeur Intel Core 2 Duo cadencé à 2.66 Ghz.

La solution journalière libre inclut une moyenne de 11 sites communs avec la solution de l'IGS pendant la période considérée. Le niveau d'accord moyen entre les solutions journa-

lières libres et la solution IGS12P33.SNX est de 1.1 mm pour la composante Est, 1.1 mm pour la composante Nord et 4.7 mm pour le vertical en utilisant notre stratégie pour estimer des 7 paramètres de la transformation de Helmert, qui maintient un excellent niveau de stabilité de la référence pour l'ensemble de la solution. Cependant, sans l'estimation avec la norme L1 et la détection d'erreur, on obtient 1.9 mm pour la composante Est, 1.8 mm pour la composante Nord et 8.8 mm pour le vertical.

La répétabilité médiane des séries temporelles du réseau est de 3.0 mm pour la composante Est, 2.5 mm pour la composante Nord et 7.8 mm pour la composante verticale. La répétabilité moyenne est de 95.6 mm pour la composante Est, 58.6 mm pour la composante Nord, 170.4 mm pour la composante verticale. La raison de cette grande différence entre les deux est que la répétabilité de 10 sites (CHNL, TAPO, CHYN, DNSH, KMNM, S058, S092, S106, SUNM, TACH) est de plusieurs mètres, en raison des sauts co-sismiques. L'histogramme des répétabilité pour les sittes non affectés par des sauts co-sismiques forts est présentée dans la figure 4.13.

En outre, nous constatons que la répétabilité de deux sites (DANA et LNCH) est très élevée et que leur série temporelle est très bruitée, et que le nombre de jour données est faible (le site DANA a 38 positions journalières dans 3 ans, pour LNCH il a 25 positions journalières dans 4 ans). Nous avons donc décidé de les retirer de l'analyse.

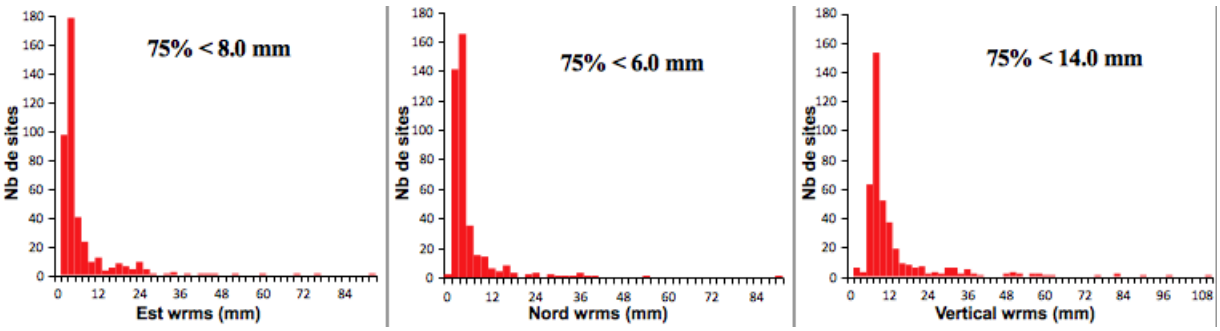


FIGURE 4.13: Histogrammes des répétabilités obtenues par pyacs_make_time_series.py

L'histogramme montre que 75% des sites ont une répétabilité inférieure à 9.0 mm sur les composantes horizontales, et 14.0 mm sur le Vertical, indiquant qu'une analyse plus poussée est nécessaire.

4.2.3 Typologie des problèmes rencontrés dans les séries temporelles

Les séries temporelles du réseau GPS du Taiwan sont beaucoup plus complexes que celles du réseau RENAG. Outre des problèmes similaires à ceux des séries temporelles du réseau RENAG (outliers, offsets, termes saisonniers), nous rencontrons aussi de nombreux offsets

co-sismiques et déformations post-sismiques, ainsi que des signaux transitoires et des séries sans vitesse constante apparente.

Dans la période de 1993 à 2011, on recense 13 tremblements de terre à Taiwan et sa région (http://earthquake.usgs.gov/earthquakes/world/historical_country.php#taiwan). Au final, nous trouvons qu'environ 120 sites (30% des sites) présentent des déplacements des co-séismes et/ou post-séismes dans les séries temporelles.

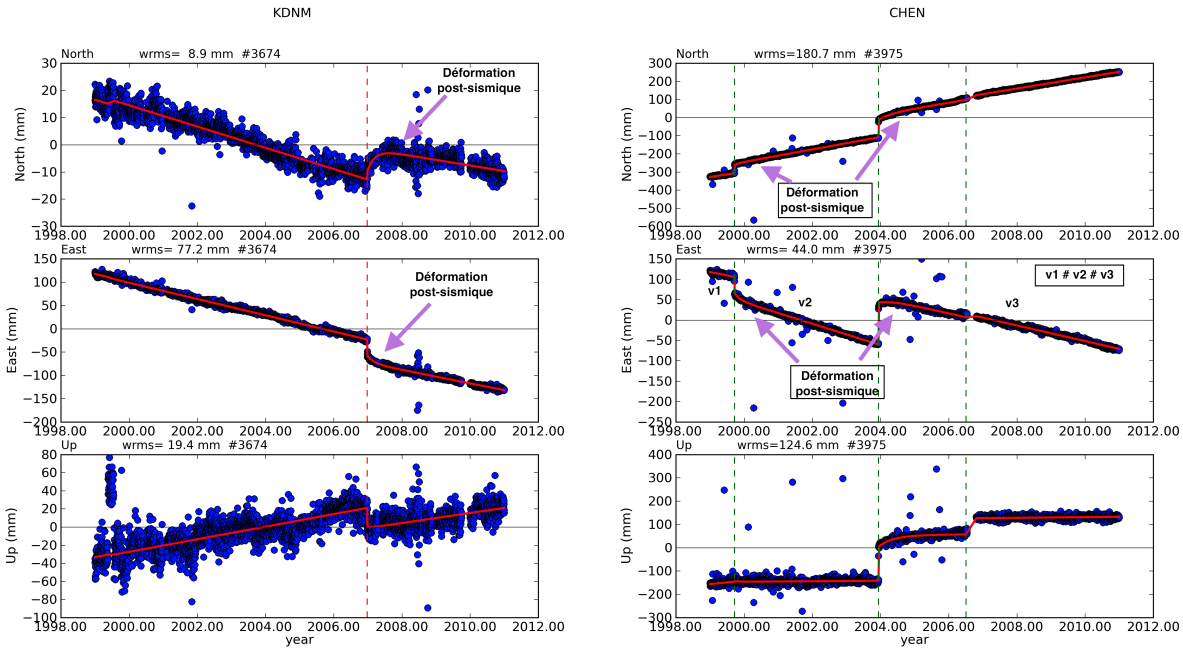


FIGURE 4.14: Exemple de série temporelle complexe (sites KDNM et CHEN). La courbe rouge indique le modèle déterminé par mon programme pygps_time_series.py.

La figure 4.14 gauche est une des séries temporelles présentant une relaxation post-sismique (site KDNM).

La figure 4.14 droite est la série temporelle plus complexe du site CHEN. Elle comporte trois offsets (deux causés par séisme et un causé par changement d'antenne) montrés par les lignes verticales vertes, deux relaxations post-sismiques (flèches violettes) et le changement de vitesse. Le changement de vitesse est le plus clairement visible sur la composante Nord.

Les signaux transitoires sont rencontrés sur quelques séries temporelles, par exemple celle du site AKND, CENT, CHIH, CHIN, FUSI, GS02, S016. Ce problème n'est pas étudié dans ma thèse, ces sites sont donc retirés de l'analyse. La figure 4.15 montre deux signaux transitoires typiques dans les séries temporelles du réseau du Taiwan. A droite est le signal post-sismique lent du site GS02, à gauche est le mouvement non-linéaire du site S167.

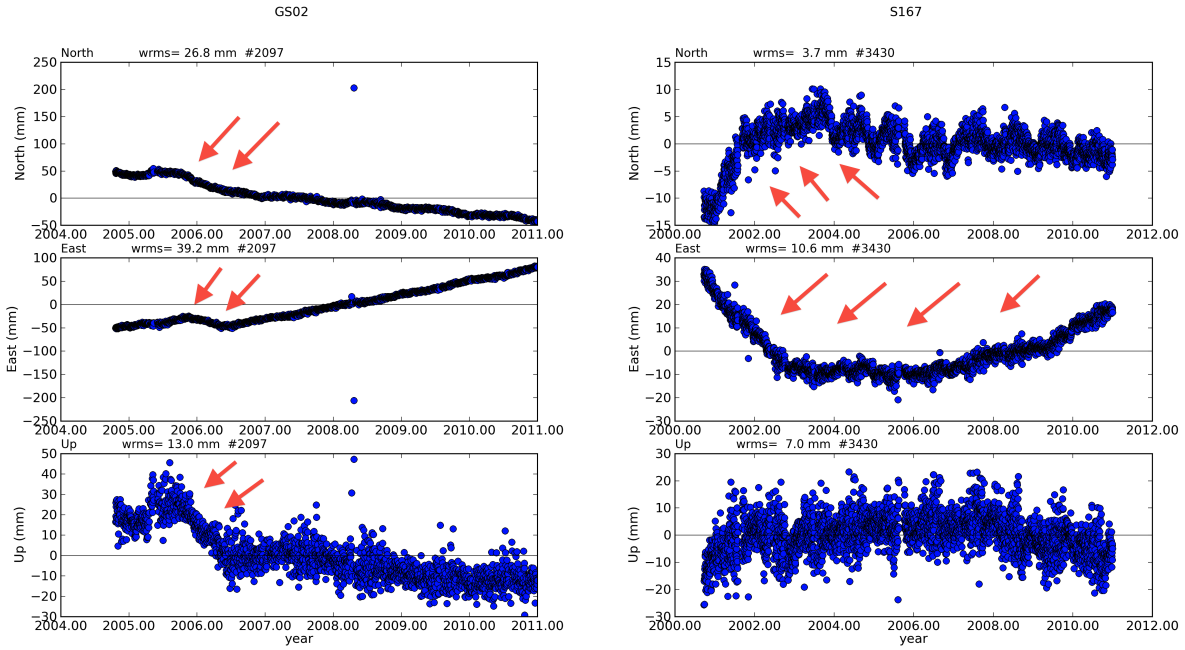


FIGURE 4.15: Série temporelle avec le séisme lent du site GS02 et celle non-linéaire du site S106

De plus, nous trouvons que la série du site DNSH est trop bruite et 15 sites (CHIK, CIME, DONY, HUSI, JIBE, JNSA, JYGU, LEYU, LSBA, MINK, T107, T108, T109, T110, WIAN) dont le nombre de positions dans la série temporelle n'est pas suffisant pour estimer précisément la vitesse. Ces sites sont aussi retirés de l'analyse des séries temporelle du réseau. Au total, le nombre des sites retirés d'analyse des séries temporelles est 49 sur 424.

4.2.4 Résultat de l'analyse automatique avec les paramètres par défaut

L'étape suivante est l'analyse des séries temporelles du réseau GPS du Taiwan produites par pygps_time_series.py. Comme pour le réseau RENAG, tout d'abord les 385 séries temporelles sont analysées automatiquement avec les paramètres par défaut et le fichier de discontinuités IGS08. L'analyse automatique prend environ 57 minutes par le même ordinateur.

Pour les offsets dans les séries temporelles, le programme pygps_time_series.py fournit de très bon résultats, avec un taux de réussite de 85%. 15% des cas mal analysés sont les petits offsets ou les offsets dans les séries complexes. Pour le cas des déformations post-sismiques, le programme requiert la date du séisme pour estimer la déformation post-sismique. Cependant, les seuls 8 sites communs avec le réseau IGS (COCO, GUAM, IISC, KUNM, PIMO, TSKB, TWTF, USUD) qui sont listés dans le fichier de discontinuité de

IGS08 sont analysés. Il y a donc environ 112 sites en présence de signal post-sismique qui requiert l'ajout d'information extérieures. Au total, nous évaluons que le taux de réussite de l'analyse automatique est de 60% - 70%.

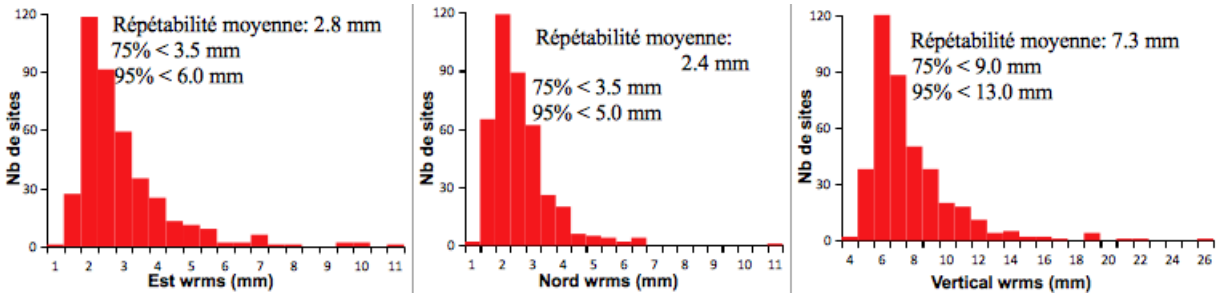


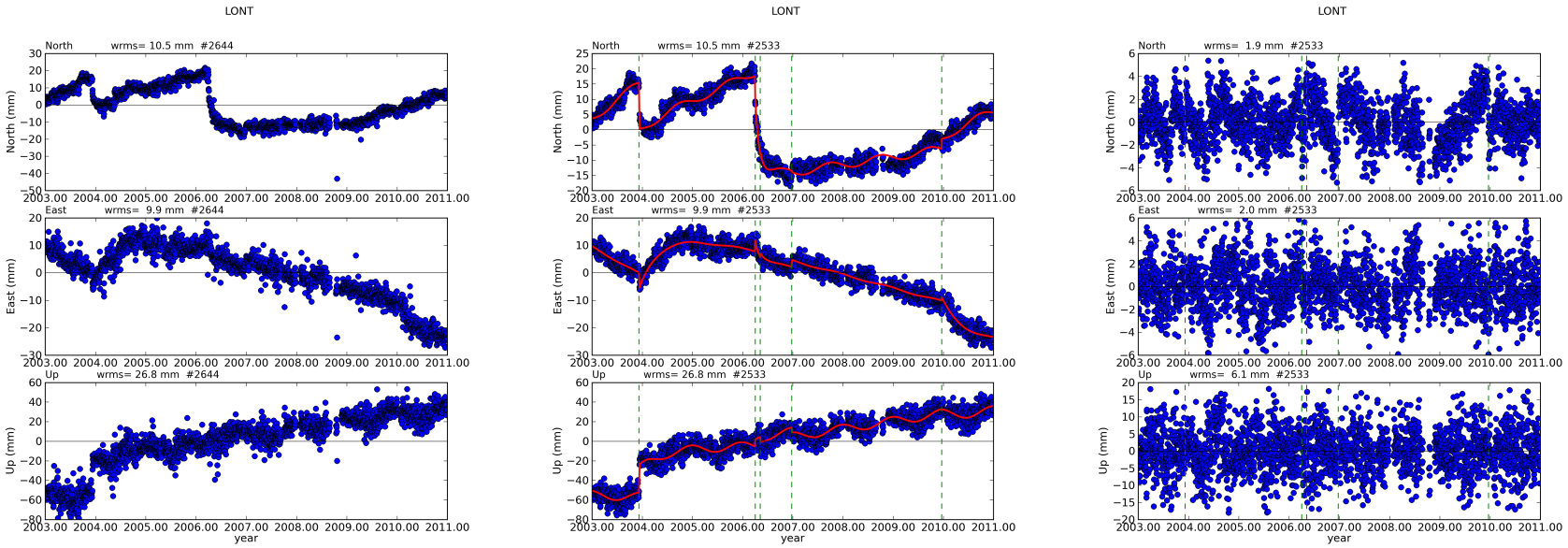
FIGURE 4.16: Histogrammes des répétabilités, après traitement par pyacs_make_time_series.py

Les histogrammes des répétabilités des stations traitement par pygps_time_series.py (figure 4.6) montrent que 75% des sites ont une répétabilité inférieure à 4.9 mm sur les composantes horizontales et 9.0 mm sur la Verticale. La précision des séries temporelles est bien améliorée par rapport aux histogrammes des répétabilités issues directement de pyacs_make_time_series.py (figure 4.13). Néanmoins, les séries temporelles de 112 sites nécessite une analyse supplémentaires.

4.2.5 Analyse semi-automatique des déformations post-sismiques

Lors du développement des outils, nous avons tenté de détecter automatiquement les dates et les caractéristiques des déplacements post-sismiques. Nous n'avons pas réussi à obtenir un algorithme robuste. Aussi, la stratégie proposée est "d'aider" l'analyse en fournissant les dates des séismes susceptibles de générer un signal post-sismique significatif. Pour re-analyser les séries temporelles existantes présentant des signaux post-sismiques et des offsets après l'analyse automatique avec les paramètres par défaut, nous créons un fichier de discontinuités en format pygps (paragraphe 3.2.5) qui liste le nom, et l'époque de la discontinuité. Ensuite nous re-analysons les séries temporelles de ces sites par pygps_time_series.py avec les paramètres par défaut en utilisant le fichier de discontinuités pygps. La validation des séries temporelles pour détecter les dates des séismes susceptibles de générer un signal post-sismique significatif et les signaux post-sismiques est faite manuellement et cette étape est assez longue.

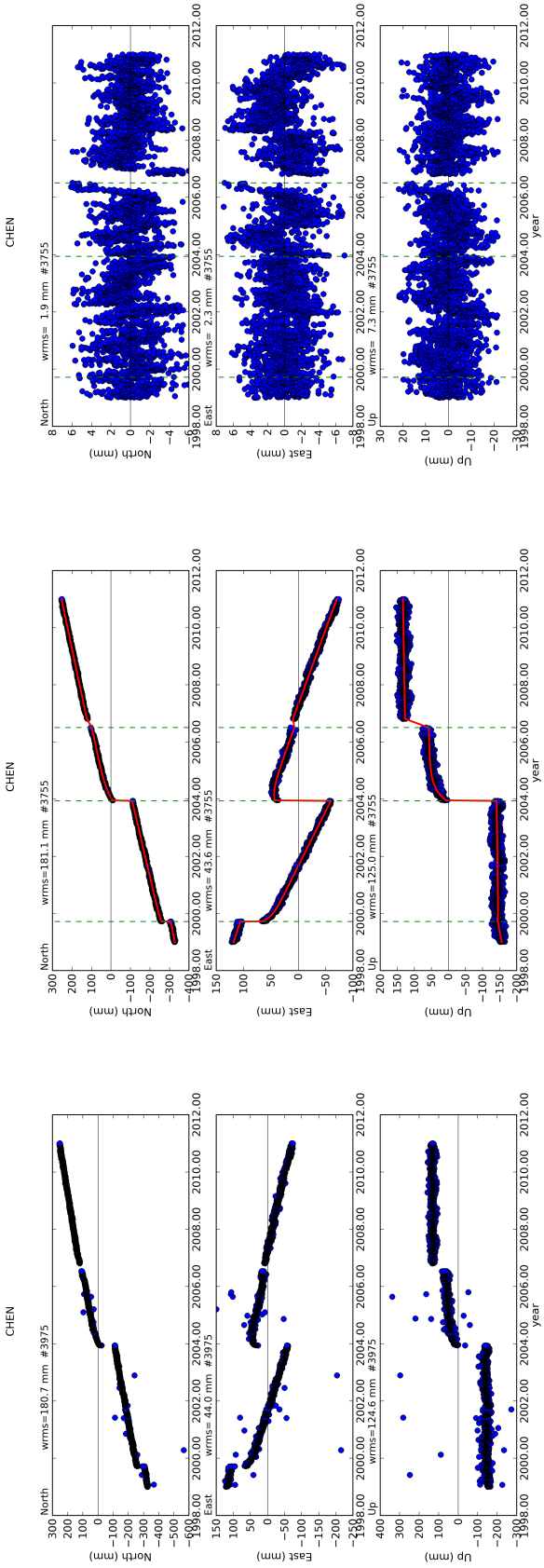
Les figures 4.17 et 4.18 montrent deux cas typiques des résultats à l'issue de la re-analyse des séries temporelles avec les paramètres par défaut. Ici, pour estimer la déformation du site LONT, nous avons ajouté dans le fichier de discontinuités en format pygps 4 époques



(a) (b) (c)

a : série brute
b : série nettoyée et son modèle avec les termes annuels (ligne rouge)
c : série de résidus
Les lignes verticales vertes sont les déplacements co-sismiques trouvés dans le fichiers de discontinuités pygps

Figure 4.17: Re-analyse de la série temporelle du site LONT



(a) (b) (c)

a : série brute
b : série nettoyée et son modèle avec prise en compte des changements de vitesse (ligne rouge)
c : série de résidus
Les lignes verticales vertes sont les déplacements co-sismiques trouvés dans le fichiers de discontinuités pygps

Figure 4.18: Re-analyse de la série temporelle du site CHEN

de co-séisme : 2003.938 (séisme Chengkung du 10 Décembre 2003), 2006.25, 2006.975 (séisme Hengchun du 26 Décembre 2006) et 2009.966 (séisme Hualien du 19 Décembre 2009). Pour le site CHEN, 2 époques de co-séisme ajoutés : 1999.712 (séisme Chichi du 20 Septembre 1999) et 2003.938 (séisme Chengkung du 10 Décembre 2003).

Le tableau 4.1 montre le résultat de la re-analyse de la série temporelle des site LONT et CHEN estimé automatiquement en utilisant du fichier de discontinuités par le programme pygps_time_series.py. Ici, je présente seulement les paramètres décrivant le déplacement co-sismique (époque, amplitude) et de la déformation post-sismique (amplitude, temps de relaxation τ).

Site	Co-séisme	Amplitude de co-séisme			Amplitude de post-séisme			τ
	(dec.year)	Nord (m)	Est (m)	Verticale (m)	Nord (m)	Est (m)	Verticale (m)	(dec.year)
LONT	2003.938	0.009	-0.011	0.032	-0.0043	-0.0224	-0.0113	0.3929
	2006.250	0.081	0.023	-0.030	0.0158	0.0063	0.0033	0.0869
	2006.975	-0.078	-0.023	-0.009	0.0008	0.0030	0.0103	0.2427
	2009.970	-0.037	-0.079	0.002	-0.0137	0.2209	0.4628	0.4256
CHEN	1999.712	0.054	-0.058	-0.005	-0.0057	0.0138	0.0029	0.2133
	2003.938	0.017	0.009	0.070	-0.0302	-0.0158	-0.0508	0.3268

TABLE 4.1: Paramètres des déplacements co- et post-sismique issus de la re-analyse de la série temporelle des sites LONT et CHEN

La figure 4.17a présente les mouvements post-sismiques dans la série du site LONT qui sont modélisés par une relaxation exponentielle avec ses termes annuels (ligne rouge dans la figure 4.17b). Le programme pygps_time_series.py produit un résultat satisfaisant. En effet, la série des résidus de positions ne présente plus de déformations post-sismiques résiduelles et les terme annuels sont aussi bien modélisés (figure 4.17c). De manière similaire, pour le site CHEN (figure 4.18), le programme re-analyse très bien la série temporelle en modélisant correctement les offsets co-sismiques, les relaxations post-sismique et les changements de vitesse.

Cette étapes soulignent cependant quelques difficultés. D'abord, l'époque de début de mouvement post-sismique doit être déterminée exactement. L'analyse n'est donc pas automatique. Nous devons introduire manuellement l'époque du séisme susceptible de générer un signal post-sismique significatif. Deuxièment, l'algorithme du programme pygps_time_series.py ne peut pas estimer des temps de relaxation post-sismique τ à partir des données de séries temporelles lorsque l'époque est manquante, par exemple pour le déplacement post-sésmique en 2009.970 du site YUSN.

4.2.6 Résultat final

Le résultat final obtenu montre une bonne qualité. Les séries temporelles du réseau ont la répétabilité moyenne de 2.1 mm pour la composante Est, 1.9 mm pour la composante Nord et 6.7 mm pour la composante verticale.

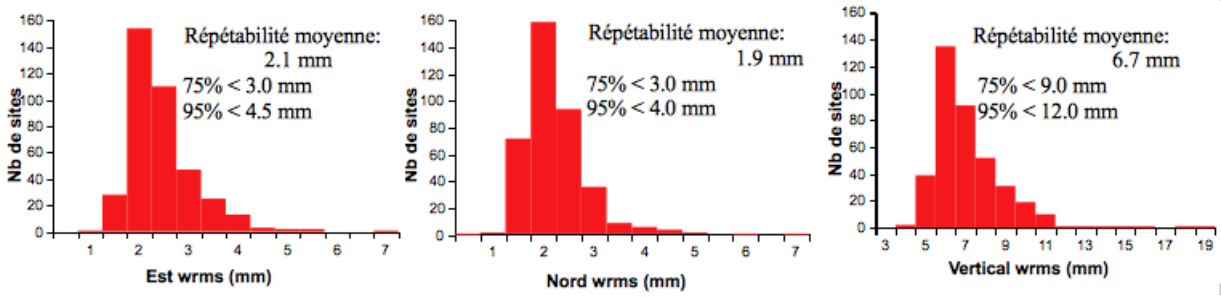


FIGURE 4.19: Histogrammes des répétabilités obtenus pour le traitement final par pyacs_make_time_series.py

La figure 4.19 présente les histogrammes des répétabilités du traitement final. Il montre que 75% des sites ont une répétabilité inférieure à 3.0 mm pour la composante Nord et Est, à 9.0 mm pour la verticale.

Le champ de vitesses obtenus

On obtient finalement le champ de vitesses horizontales (figure 4.20) et le champ de vitesses verticales (figure 4.21) dans l'ITRF2008. Les vitesses sont détaillé dans les tableaux B.1, B.2, B.3, B.4, B.5.

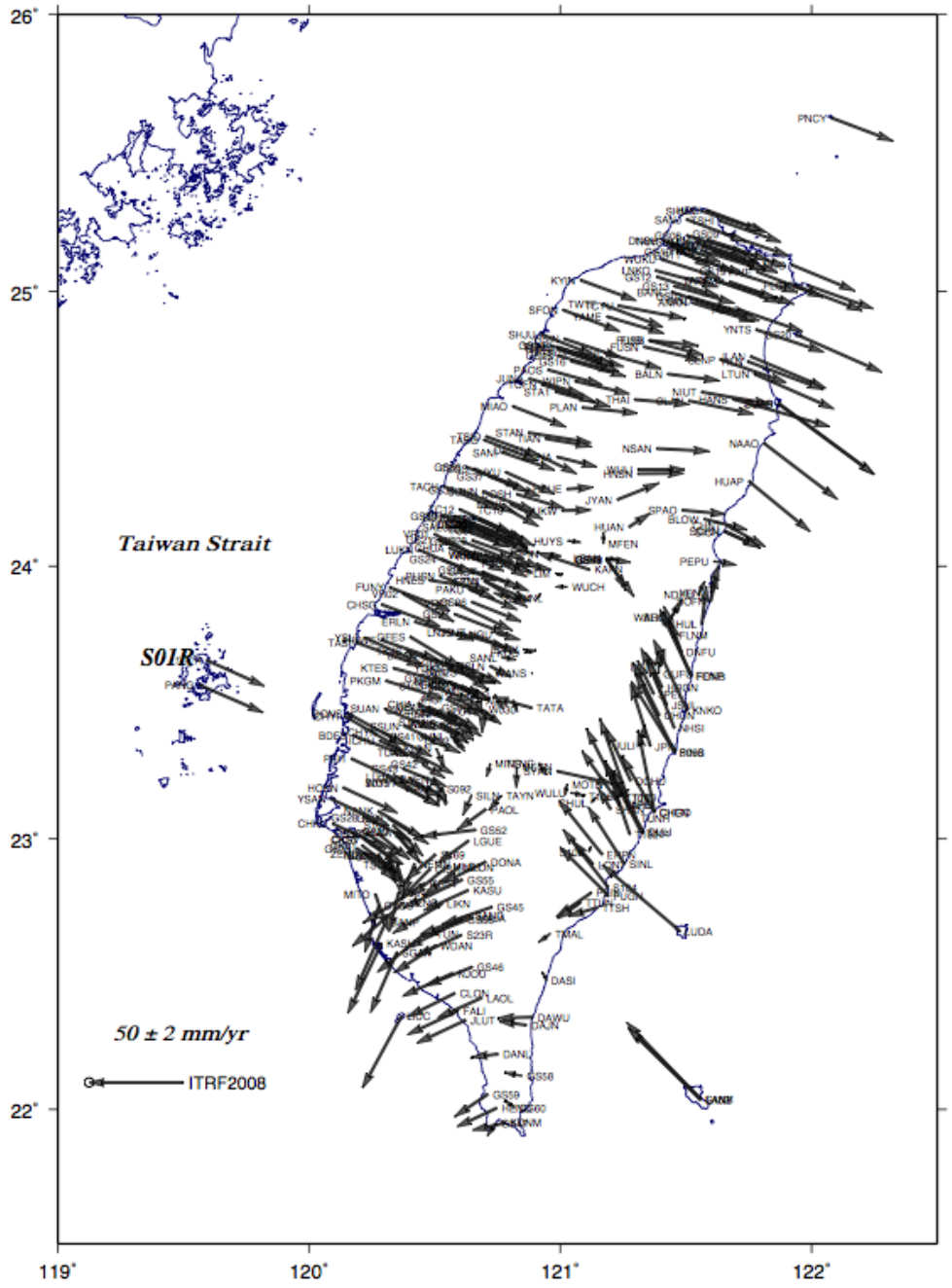


FIGURE 4.20: Champ de vitesses horizontal dans l'ITRF2008 du réseau GPS du Taiwan

Le champ de vitesses horizontales dans l'ITRF2008 présente des vitesses allant de 0.5 à 71.9 mm/an vers les directions différentes. Les vitesses des stations dans la zone du Nord et de l'Ouest-Nord montrent le mouvement vers le Sud-Est. Cependant, celles des stations du centre de l'Est montrent le mouvement vers le Nord-Est et celles des stations de la zone du Sud montrent le mouvement vers l'Est. En comparaison avec les résultats des analyses plus complètes des auteurs comme Lin et al. [2010], Shin et al. [2011], notre résultat de l'analyse rapide est généralement similaire.

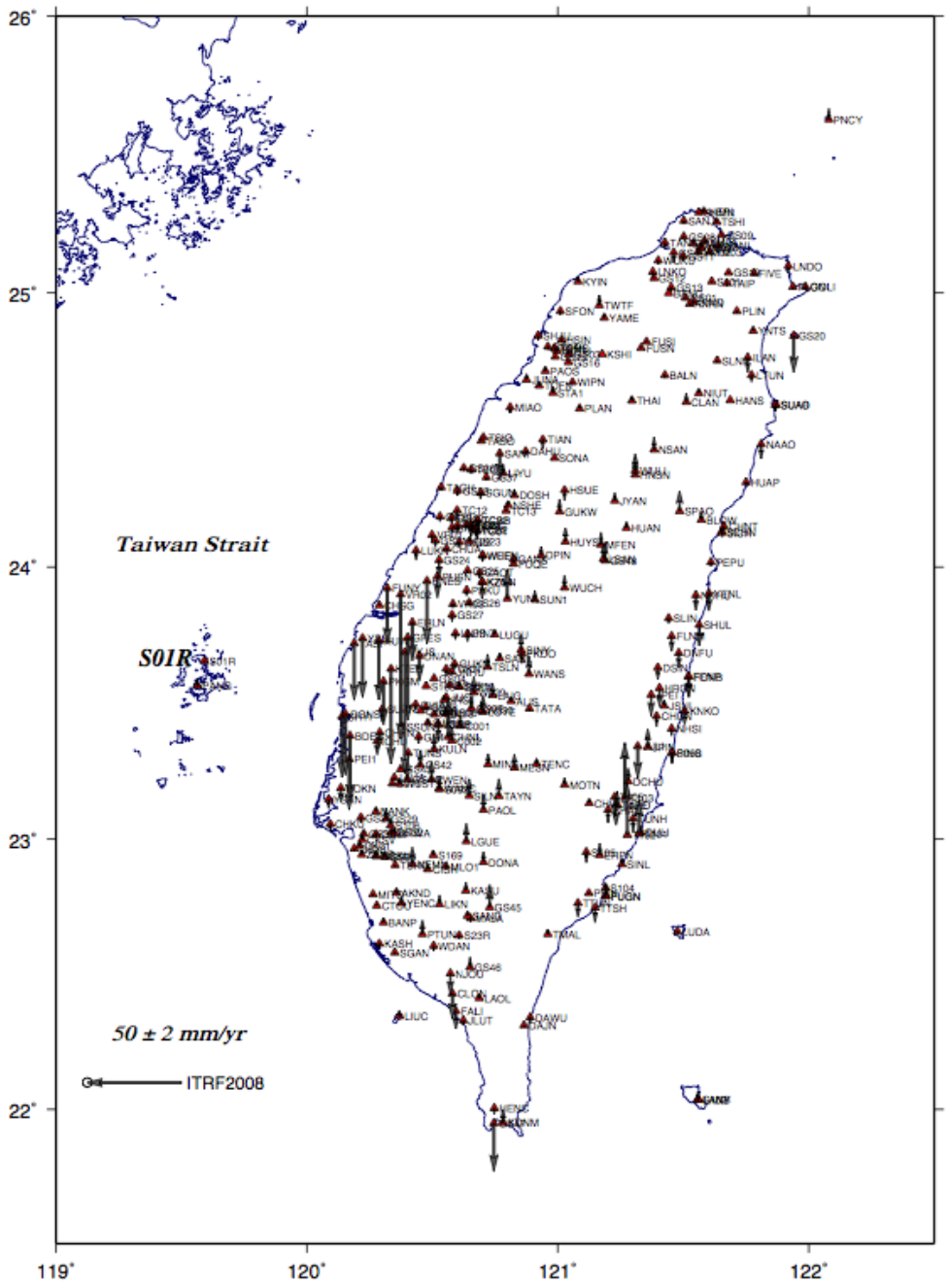


FIGURE 4.21: Champ de vitesses vertical dans l'ITRF2008 du réseau GPS du Taiwan

La figure 4.21 montre les vitesses verticales de plusieurs tendances. Certaines vitesses sont positives et fortes. Cependant, les vitesses négatives (subsidence) sont fortes, trouvées sur Beigang, Puzi (zone Sud-Ouest). La vitesse négative la plus forte est de -76.4 mm/an (VR02).

Dans le cas du réseau GPS du Taiwan, bien que nos approches rapides analysent relativement bien les problèmes des offsets (causés par les changements d'antenne, par les co-séismes) et des déformations post-sismiques dans les séries temporelles. Mais ces pro-

blèmes affectent encore la détermination de la vitesse verticale de quelques sites, par exemple du site CHEN, GS52,... (dont la vitesse est retirée de la carte 4.21). Nous trouvons donc la petite différence dans le résultat de l'analyse du champ de vitesses verticales par rapport des résultats des autres auteurs comme Chen et al. [2012].

4.2.7 Conclusions

L'analyse du réseau de 424 stations GPS permanents du Taiwan est réalisée par les 3 mêmes étapes que celle du réseau RENAG. Cependant, elle a pris beaucoup plus du temps, en particulier, à l'étape de la re-analyse des séries temporelles pour 112 stations des signaux post-sismiques. Nous avons retiré 49 stations de l'analyse qui ont des signaux transitoires ou des données insuffisantes. De plus, le temps de relaxation post-sismique τ n'est pas résolu par notre méthode dans quelques cas.

Le champ de vitesse dans l'ITRF2008 du réseau GPS du Taiwan montrent les vitesses horizontales fortes de 0.5mm/an à 71.9 mm/an dans l'ITRF2008. Il montre aussi certaines vitesses verticales positives, faibles. Les vitesses verticales négatives sont fortes et trouvées dans la zone Sud-Ouest.

CONCLUSION GÉNÉRALE

Dans le contexte du développement de grands réseaux GPS permanents, et de séries temporelles longues maintenant disponibles, j'ai développé au cours de ma thèse un ensemble d'approches et de programmes généraux permettant de réaliser rapidement une analyse géodésique pour les applications en tectonique. Les programmes développés s'insèrent plus généralement dans un "framework" PYACS qui inclut en amont de mon travail outils de calcul GPS sur Cluster, et en aval, une modélisation cinématique en blocs, tenseurs de déformation et la modélisation élastique inter- et co-sismique.

Dans ce contexte, mon travail aboutit à une méthode et des outils permettant la réalisation rapide et robuste des séries temporelles exprimée dans le repère de référence et l'obtention de champs de vitesses. Mon travail aboutit aussi à un ensemble d'approches, implémentés dans des modules (la Classe Geodetic time series Gts et les méthodes qui s'y rapportent), permettant de manipuler les séries temporelles afin d'extraire rapidement les signaux d'intérêt géophysique. Les applications possibles sont nombreuses avec la réalisation de champ de vitesses fiables et de grande qualité pour l'étude de la tectonique, la quantification des déplacements co- et post-sismiques, la recherche des signaux transitoires, l'étude des surcharges. Par sa conception modulaire et l'approche orientée objet de Python, les outils permettent une utilisation manuelle où l'utilisateur peut choisir précisément son analyse. Les modules de second niveau et les programmes "wrapper" permettent une analyse automatique des problèmes les plus courants pour traiter de grands volumes de données. Concrètement, l'analyse du réseau RENAG n'a pris qu'une dizaine de minutes pour la réalisation les séries temporelles dans l'ITRF2008, et 31 minutes pour l'analyse automatique des problèmes dans les séries. L'analyse manuelle des séries temporelles présentant encore des problèmes est résolue par un ajustement des paramètres de traitement automatique, et finalement les problèmes résiduels et l'extraction de signaux spécifiques peut être réalisée avec les outils de la Classe "Gts". Le résultat obtenu de très bonne qualité, par exemple pour la solution RENAG est indiqué par la répétabilité moyenne de 1.6 mm pour les composantes horizontales et 3.4 mm pour la composante verticale, ou encore le rms de 0.15 mm/an obtenu lors de l'estimation d'un pôle d'Euler dans les Alpes et en Europe.

A titre personnel, le travail réalisé dans la cadre de ma thèse m'a permis d'apprendre les étapes de l'analyse des réseaux GPS permanents pour la tectonique et ses problèmes. Les études, les connaissances obtenues me seront très utiles pour mes études futures et pour mes travaux au Vietnam. J'espère surtout que les outils et le prototype auquel j'arrive à la fin de cette thèse sera utile pour les collègues, étudiants, ou utilisateurs souhaitant travailler en géodésie pour la tectonique. Plusieurs étudiants du laboratoires ont d'ailleurs utilisé ces outils pour leur propres études.

Les limitations

Le travail présenté correspond à celui réalisé sur une période de 3 ans, où une partie du temps a été consacré à la formation, en particulier à l'apprentissage du langage Python. Les outils, bien qu'utilisables en l'état sont amenés à évoluer. Les perspectives d'amélioration concernent soit la performance (vitesse, robustesse, taux de réussite) des algorithmes, soit l'ajout de nouvelles fonctionnalités. Parmi les limitations principales de mes outils, les pistes d'améliorations à envisager sont :

a. Le traitement des incertitudes

Dans le but à la fois de rapidité du développement et de vitesse de calcul, les matrices variances-covariances du jeu de référence et des solutions libres ont été négligées. On néglige donc les écart-type sur les coordonnées elles-mêmes, les corrélations inter-composantes et les corrélations inter-sites. Bien que nous n'ayons pas testé l'influence de cette simplification, dans le cas des réseaux permanents calculés par session journalière de 24 heures, l'écart-type d'un jour sur l'autre est similaire et les corrélations sont faibles. On anticipe donc que l'influence de négliger les matrices variance-covariance doit être faible sur le résultat final, à la fois en terme de valeur et en terme d'incertitude.

Un soucis plus important est que la corrélation temporelle entre positions est négligée et que donc, le bruit des séries temporelles GPS est supposé être blanc (par opposition à coloré/corrélé). Cette simplification que l'on sait être fausse depuis la fin des années 1990 impacte les incertitudes sur les vitesses. La conséquence la plus directe et la plus problématique est que les incertitudes dérivées sur les vitesses issues de mes outils n'est pas utilisable, sauf à la redilater avant utilisation par ajout des critères externes.

b. Un pas vers les solutions complètes

Dans les logiciels Globk et CatRef, la solution finale est obtenue par cumul des solutions libres après étude et correction des problèmes à partir de l'analyse des séries temporelles. Ils nous permettent donc de prendre en compte la corrélation temporelle entre paramètres et pour Globk, une forme de prise en compte de la corrélation temporelle du bruit. Une perspective serait que la solution issue de nos approches rapides pourrait servir à générer

des fichiers d'outliers, d'offsets utilisables ensuite pour les calculs complets sous Globk et Catref.

c. Les améliorations de la robustesse des algorithmes

En traitement automatique d'un grand nombre de séries temporelles, les problèmes d'outliers et d'offsets sont en général bien analysés. Cependant, l'analyse du déplacement post-sismique n'est pas automatique. Elle requiert les dates des séismes susceptibles de générer un signal post-sismique préalablement à l'analyse. Il reste difficile de détecter et d'analyser automatiquement les mouvements post-sismiques dans les séries temporelles en aveugle. La perspective ne serait donc pas la détection automatique des signaux transitoires, mais l'amélioration de la performance et de la précision de l'algorithme d'estimation du temps et de l'amplitude de la relaxation post-sismique en utilisant les informations a priori.

d. La détection et quantification des phénomènes transitoires

Lors de l'analyse des séries temporelles de positions GPS du réseau GEONET-Japon, du réseau GPS du Taiwan, nous constatons que plusieurs séries présentent des mouvements transitoires, non lié à un séisme. Bien que ce phénomène ne soit pas courant, la détermination des mouvements transitoires est importante pour les études des phénomènes géophysiques. Dans mon travail de thèse, je n'ai pas du temps d'aborder l'analyse des signaux transitoires.

Cependant, les travaux menés permettent d'envisager des pistes de détection de ces mouvements. Lors de l'analyse des séries temporelles, nous avons constaté que les positions correspondant aux périodes transitoires sont généralement rejetées comme outliers par nos méthodes de détection, mais dans ce cas, les outliers sont toujours nombreux et groupés dans le temps. Pour détecter les signaux transitoires dans les séries temporelles, une perspective serait d'analyser les séries temporelles d'outliers avec quelques patterns simples en exponentielle ou fonctions tangentes, puis de les injecter dans le modèle générale des séries temporelles décrit dans le chapitre 3.

Pour conclure, je pense que, bien qu'il s'agisse d'un prototype, les outils que j'ai développés devraient être mise à disposition publiquement et peuvent constituer une base pour les développements futurs. Ces outils sont utilisés dans plusieurs études en cours et participeront à plusieurs articles en préparation : Ils ont été utilisés par Jean-Mathieu Nocquet pour la réalisation d'un champ de vitesse à l'échelle du Nord des Andes (Geodetic evidence for a giant aseismic segment and a new Inca micro-plate along the Nazca/South America subduction zone, Nocquet et al., en préparation), Juan Carlos Villegas Lanza pour l'étude d'un séisme lent au Nord Pérou (Anatomy of a seismic swarm synchronous to a slow slip event in a subduction zone : The Bayovar (northern Peru) 2009 sequence, Villegas et al., en préparation). Enfin, ils ont été utilisés pour une étude préliminaire du

système de référence géodésique national de l'Equateur (Stability of National Geodetic Reference Frame in a Actively Deforming Area : the case study of Ecuador, Cisneros et al., en préparation). La mise à disposition publique des codes que j'ai développés est prévue dans le cadre d'un article : Rapid prototyping of GPS solutions for tectonic analysis (Tran and Nocquet, en préparation).

BIBLIOGRAPHY

Z. Altamimi. Systèmes de référence terrestres: définition, réalisation, application à l'ITRF, état actuel et perspectives. Habilitation à diriger des recherches. Université Pièrre et Marie Curie (Paris 6), France, 2006.

Z. Altamimi, X. Collilieux, and L. Metivier. ITRF2008: an improved solution of the international terrestrial reference frame. *Journal of geodesy*, 85:457–473, 2011. doi: 10.1007/s00190-011-0444-4.

A. R. Amiri-Simkooei, C. C. J. M. Tiberius, and P. J. G. Teunissen. Assessment of noise in GPS coordinate time series: Methodology and results. *Journal of geophysical research*, 112(B07413), 2007. doi: 10.1029/2006JB004913.

G. Blewitt and D. Lavallée. Effect of annual signals on geodetic velocity. *Journal of geophysical research*, 107(B7), 2002. doi: 10.1029/2001JB000570.

C. Boucher and Z. Altamimi. ITRS, PZ-90 and WGS84: current realizations and the related transformation parameters. *Journal of geodesy*, 75:613–619, 2001.

E. Calais, R. Bayer, J. Chéry, F. Cotton, E. Doerflinger, M. Flouzat, F. Jouanne, M. Kasser, M. Laplanche, D. Maillard, J. Martinod, F. Mathieu, P. Nicolon, J.-M. Nocquet, O. Scotti, L. Serrurier, M. Tardy, and C. Vigny. Regal : réseau GPS permanent dans les Alpes occidentales. Configuration et premiers résultats. *Académie des sciences*, (331):435–442, 2000.

H.-Y. Chen, J.-C. Lee, H. Tung, S.-B. Yu, Y.-J. Hsu, and H. Lee. Determination of Vertical Velocity Field of Southernmost Longitudinal Valley AbSTrAcT in Eastern Taiwan: A Joint Analysis of Leveling and GPS Measurements. *Terr. Atmos. Ocean. Sci.*, 23(4): 355–376, 2012. doi: 10.3319/TAO.2012.02.29.01(TT).

G. B. Dantzig and M. N. Thapa. *Linear programming 1: introdution*. Springer Verlag, 1997.

J. L. Davis, B. P. Wernicke, S. Bisnath, N. A. Niemi, and P. Elosegui. Subcontinental-scale crustal velocity changes along the Pacific–North America plate boundary. *Nature*, 441, 2006. doi: 10.1038/nature04781.

D. Dong, P. Fang, Y. Bock, M. K. Cheng, and S. Miyazaki. Anatomy of apparent seasonal variations from GPS-derived site position time series. *Journal of geophysical research*, 107(B4), 2002. doi: 101029/2001JB000573.

K. L. Feigl, D. C. Agnew, Y. Bock, D. Dong, A. Donnellan, B. H. Hager, T. A. Herring, D. D. Jackson, T. H. Jordan, I. W. King, S. Larsen, K. M. Larson, M. H. Murray, Z. Shen, and F. H. Webb. Space Geodetic Measurement of Crustal Deformationin Central and Southern California, 1984-1992. *Journal of geophysical research*, 98(B12): 21677–21712, 1993.

Q. Gui and J. Zhang. Robust biased estimation and its applications in geodetic adjustments. *Journal of geodesy*, 72:430–435, 1998.

T. A. Herring, R. W. King, and S. C. McClusky. *GAMIT reference manual. GPS analysis at MIT. Release 10.04*. Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, October 2010a.

T. A. Herring, R. W. King, and S. C. McClusky. *Introdution to GAMIT/GLOBK release 10.04*. Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, October 2010b.

T. A. Herring, R. W. King, and S. C. McClusky. *GLOBK reference manual. Global Kalman filter VLBI and GPS analysis program. Release 10.04*. Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, December 2010c.

Kang Hyeun Ji and Thomas A. Herring. A method for detecting transient signals in gps position time-series: smoothing and principal component analysis. *Geophysical Journal International*, 2013. doi: 10.1093/gji/ggt003.

Y. Jiang, T. H. Dixon, and S. Wdowinski. Accelerating uplift in the North Atlantic region as an indicator of ice loss. *Nature geoscience*, 16 May 2010. doi: 10.1038/NGEO854.

A. Khodabandeh and A. R. Amiri-Simkooei. Recursive algorithm for L1 norm estimation in linear models. *Journal of serveying engineering*, 2011. doi: 10.1061/(ASCE)SU.1943-5428.0000031.

K.-R Koch. *Parameter estimation and hypothesis testing in linear models*. Springer Verlag, 1999.

J. Langbein. Noise in GPS displacement measurements from Southern California and Southern Nevada. *Journal of geophysical research*, 113(B05405), 2008. doi: 10.1029/2007JB005247.

R. B. Langley. The almanac. *GPS World*, 2011. URL http://www.gpsworld.com/GNSS20System/Almanac/almanac-4265.

K.-C Lin, J.-C. Hu, K.-E. Ching, J. Angelier, R.-J. Rau, S.-B. Yu, C.-H. Tsai, T.-C. Shin, and M.H. Huang. GPS crustal deformation, strain rate, and seismic activity after the 1999 Chi-Chi earthquake in Taiwan. *Journal of geophysical research*, 115(B07404), 2010. doi: 10.1029/2009JB006417.

F. Lyard, F. Lefevre, T. Letellier, and O. Francis. Modelling the global ocean tides: modern insights from FES2004. *Ocean Dynamics*, 56:394–415, 2006. doi: 10.1007/s10236-006-0086-x.

A. Mao, C. G. A. Harrison, and T. H. Dixon. Noise in GPS coordinate time series. *Journal of geophysical research*, 104(B2):2797–2816, 1999. doi: 10.1029/1998JB900033.

D. D. McCarthy and G. Petit. IERS conventions. IERS Technical Note 32, IERS Convention Centre, 2003.

J. Nicolas, J.-M. Nocquet, M. Van Camp, T. van Dam, J.-P. Boy, J. Hinderer, P. Gegout, E. Calais, , and M. Amalvict. Seasonal effect on vertical positioning by Satellite Laser Ranging and Global Positioning System and on absolute gravity at the OCA geodetic station, Grasse, France. *Geophysical Journal International*, 167:1127–1137, 2006. doi: 10.1111/j.1365-246X.2006.03205.x.

R. Nikolaidis. *Observation of geodetic and seismic deformation with the GPS*. PhD thesis, University of California, San Diego, USA, 2002.

J.-M Nocquet. Apport de la géodésie spatiale à la connaissance de la géodynamique en Europe et en Méditerrannée. Habilitation à diriger des recherches. Université de Nice Sophia Antipolis, France, 2011.

J.-M. Nocquet. Present-day kinematics of the Mediterranean: A comprehensive overview of GPS results. *Tectonophysics*, (579):220–242, 2012. doi: 10.1016/j.tecto.2012.03.037.

W. H. Press, S. A. Teukolskey, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in Fortran 77: The art of scientific computing*, volume 1 of *ISBN 0-521-43064-X*. Press Syndicate of the University of Cambridge, 1992.

J. Ray, D. Dong, and Z. Altamimi. IGS reference frames: status and future improvements. *GPS solution*, (8):251–266, 2004. doi: 10.1007/s10291-004-0110-x.

P. Rebischung, J. Griffiths, J. Ray, R. Schmid, X. Collilieux, and B. Garayt. IGS08: the IGS realization of ITRF2008. *GPS solution*, 16:483–494, 2012. doi: 10.1007/s10291-011-0248-2.

R. Schmid, P. Steigenberger, G. Gendt, M. Ge, and M. Rothacher. Generation of a consistent absolute phase center correction model for GPS receiver and satellite antennas. *Journal of geodesy*, 81:781–798, 2007. doi: 10.1007/s00190-007-0148-y.

P. Segall and J. L. Davis. GPS applications for geodynamics and earthquake studies. *Annual reviews of Earth and planetary sciences*, 25:301–336, 1997.

T.-C. Shin, K.-W. Kuo, P.-L. Leu, C.-H. Tsai, and J.-S. Jiang. Continuous CWB GPS Array in Taiwan and Applications to Monitoring Seismic Activity. *Terr. Atmos. Ocean. Sci.*, 22(5):521–533, 2011. doi: 10.3319/TAO.2011.05.18.01(T).

P. Sillard. *Estimation par moindres carreés*. ISBN 2-7462-03339-1. Hermès Science Publications, 2001.

P. Tregoning and C. Watson. Atmospheric effects and spurious signals in GPS analyses. *Journal of geophysical research*, 114(B09403), 2012. doi: 10.1029/2009JB006344.

M. Vergnolle, A. Walpersdorf, V. Kostoglodov, P. Tregoning, J. A. Santiago, N. Cotte, and S. I. Franco. Slow slip events in mexico revised from the processing of 11 year GPS observations. *Journal of geophysical research*, 115(B08403), 2010. doi: 10.1029/2009JB006852.

Y. Yang. Robust estimation of geodetic datum transformation. *Journal of geodesy*, 73: 268–274, 1999.

T. K. Yeh, C. S. Wang, B. F. Chao, C. S. Chen, and C. W. Lee. Automatic data-quality monitoring for continuous GPS tracking stations in Taiwan. *Metrologia*, (44):393–401, 2007. doi: 10.1088/0026-1394/44/5/016.

ANNEXES

A Résultats du réseau RENAG

A.1 Vitesses dans l'ITRF2008 et après estimation d'un pôle de rotation

				Vitesse				Ecart-type	
		Position		ITRF2008		Pôle de l'Europe stable			
No	Site	long.	lat.	Ve	Vn	Ve	Vn	Sve	Sve
1	ABER	-2.08022	57.14401	15.28	15.49	0.82	-0.45	0.14	0.14
2	ABYW	-4.00438	52.42408	16.96	17.03	1.30	1.06	0.19	0.16
3	AIGL	3.58126	44.12140	20.17	15.69	0.42	-0.07	0.52	0.56
4	AJAC	8.76261	41.92746	21.12	15.68	-0.14	0.22	0.16	0.14
5	ALAC	-0.48123	38.33892	20.01	16.39	-0.49	0.47	0.17	0.16
6	ALES	8.61633	44.92315	19.86	16.30	-0.64	0.83	0.77	0.77
7	ASTI	8.20320	44.90570	20.60	15.81	0.18	0.31	0.33	0.30
8	AXPV	5.33320	43.49121	20.24	15.81	-0.01	0.13	0.40	0.38
9	BANN	4.15632	44.36920	20.01	15.81	0.22	0.07	0.33	0.33
10	BARK	0.09690	51.51538	17.46	16.44	0.61	0.55	0.23	0.28
11	BIEL	8.04806	45.56075	20.22	15.24	0.00	-0.27	0.38	0.35
12	BOGO	21.03534	52.47595	20.70	13.92	-0.25	-0.31	0.24	0.23
13	BRST	-4.49659	48.38049	16.85	16.65	-0.03	0.67	0.44	0.35
14	BRUS	4.35922	50.79782	18.41	15.55	0.42	-0.17	0.21	0.19
15	BSCN	5.98939	47.24688	19.42	15.67	0.06	0.04	0.16	0.16
16	BZRG	11.33680	46.49902	19.79	15.77	-0.82	0.51	0.30	0.28
17	CAGL	8.97275	39.13591	21.69	15.63	-0.24	0.18	0.21	0.16
18	CAMB	-5.32737	50.21844	16.33	16.71	0.22	0.73	0.14	0.12
19	CARL	-2.93827	54.89543	15.46	16.04	0.41	0.08	0.17	0.16
20	CAST	7.70930	45.38671	20.31	15.35	0.11	-0.18	0.38	0.37
21	CHAM	5.88108	45.11071	19.86	15.37	-0.07	-0.27	0.33	0.35
22	CHIZ	-0.40770	46.13346	18.56	15.91	0.16	0.00	0.16	0.14
23	CHRN	4.86177	43.88137	20.08	15.75	0.03	0.05	0.09	0.10
24	CHTL	6.35856	45.30414	19.56	15.47	-0.41	-0.15	0.12	0.12
25	COLC	0.89725	51.89437	17.12	15.99	0.21	0.13	0.17	0.16
26	CUBX	-0.56639	44.86886	18.84	16.20	0.10	0.28	0.16	0.17
27	CUNE	7.55357	44.39499	20.64	15.83	0.21	0.29	0.45	0.49
28	DARE	-2.64049	53.34481	15.92	16.29	0.27	0.34	0.16	0.16
29	DUNK	-3.23989	50.86046	16.97	16.21	0.62	0.25	0.16	0.16
30	EBRE	0.49236	40.82089	20.03	15.88	-0.01	-0.01	0.14	0.12

TABLE A.1: Vitesse du réseau RENAG
Vitesses (Ve, Vn) et écart-types (Sve,Svn) en mm/an. Longitude, latitude en degrés décimaux.

				Vitesse				Ecart-type	
		Position		ITRF2008		Pôle de l'Europe stable			
No	Site	long.	lat.	Ve	Vn	Ve	Vn	Sve	Sve
31	EDIN	-3.29479	55.92479	15.07	16.15	0.46	0.19	0.17	0.19
32	EGLT	2.05200	45.40335	19.53	15.50	0.43	-0.33	0.19	0.21
33	ELBA	10.21110	42.75290	20.86	15.81	-0.47	0.46	0.21	0.19
34	ENTZ	7.63990	48.54936	19.50	15.22	0.17	-0.31	0.45	0.42
35	ESCO	0.97567	42.69358	19.42	16.18	-0.22	0.31	0.14	0.16
36	FCLZ	5.98568	45.64300	19.88	15.78	0.08	0.14	0.12	0.09
37	FJCP	2.79490	43.04816	19.49	15.92	-0.40	0.12	0.42	0.35
38	GENO	8.92114	44.41939	20.77	15.43	0.09	-0.02	0.09	0.09
39	GINA	5.78698	43.67572	20.10	15.66	-0.18	0.01	0.12	0.10
40	GRAS	6.92057	43.75474	20.41	15.79	-0.07	0.21	0.19	0.16
41	GRAV	7.01659	45.12772	20.99	15.10	0.85	-0.47	1.24	1.37
42	GRAZ	15.49348	47.06713	21.86	15.23	0.60	0.36	0.26	0.23
43	GROI	-3.50798	47.64801	17.40	15.78	0.08	-0.19	0.37	0.35
44	HEAU	-1.77796	49.58066	17.49	15.59	0.42	-0.35	0.37	0.33
45	IESG	-1.19228	52.94069	16.65	15.60	0.54	-0.33	0.10	0.10
46	INGR	12.51480	41.82808	20.87	16.54	-1.08	1.38	0.23	0.17
47	INVE	-4.21926	57.48625	14.53	16.27	0.69	0.30	0.21	0.19
48	JOUX	5.79566	46.52857	19.68	15.90	0.16	0.25	0.10	0.10
49	JOZE	21.03154	52.09727	21.07	14.03	0.02	-0.20	0.16	0.17
50	KOSG	5.80964	52.17843	18.05	15.87	0.17	0.23	0.14	0.17
51	LACA	2.72758	43.68099	19.93	15.95	0.23	0.15	0.50	0.54
52	LAMP	12.60566	35.49977	19.95	18.32	-3.29	3.16	0.16	0.16
53	LERW	-1.18494	60.13914	13.74	16.58	0.14	0.66	0.17	0.17
54	LFAZ	5.39846	45.11663	20.01	15.63	0.18	-0.04	0.33	0.35
55	LOWE	1.75020	52.47322	17.86	15.56	0.95	-0.27	0.14	0.16
56	LROC	-1.21932	46.15894	18.39	15.94	0.16	0.01	0.09	0.09
57	MALL	2.62455	39.55263	18.48	16.86	-2.26	1.05	0.50	0.35
58	MANS	0.15528	48.01862	17.59	16.51	-0.37	0.62	0.16	0.16
59	MARS	5.35379	43.27877	20.00	15.68	-0.30	0.01	0.24	0.26
60	MATE	16.70446	40.64913	23.37	18.77	0.45	4.02	0.12	0.14
61	MEDI	11.64681	44.51996	22.71	17.20	1.54	1.97	0.17	0.16
62	MICH	5.71735	43.92416	20.02	15.65	-0.19	0.00	0.10	0.10
63	MLVL	2.58731	48.84106	18.27	15.70	0.06	-0.10	0.28	0.19
64	MODA	6.71009	45.21378	19.93	15.40	-0.13	-0.19	0.10	0.12
65	MOND	7.82720	44.38894	20.77	15.41	0.29	-0.11	0.31	0.33
66	MTPL	3.86484	43.63744	20.06	15.80	0.13	0.05	0.09	0.07

TABLE A.2: Vitesse du réseau RENAG (continues)
Vitesses (Ve, Vn) et écart-types (Sve,Svn) en mm/an. Longitude, latitude en degrés décimaux.

		Position		Vitesse ITRF2008		Vitesse Pôle de l'Europe stable		Ecart-type	
No	Site	long.	lat.	Ve	Vn	Ve	Vn	Sve	Sve
67	NEWC	-1.61657	54.97913	15.62	15.98	0.30	0.05	0.17	0.16
68	NICE	7.29998	43.72550	20.79	15.61	0.24	0.05	0.10	0.10
69	NOT1	14.98979	36.87584	21.47	19.59	-1.89	4.66	0.16	0.14
70	NOVA	8.61397	45.44723	20.36	15.34	0.00	-0.13	0.37	0.31
71	ONSA	11.92552	57.39530	17.22	14.26	-0.37	-0.94	0.07	0.07
72	ORID	20.79405	41.12731	23.63	11.14	0.13	-3.13	0.21	0.19
73	PAVI	9.13614	45.20298	20.21	16.05	-0.32	0.62	0.44	0.30
74	PENC	19.28153	47.78960	22.23	14.38	0.43	-0.07	0.21	0.26
75	POTS	13.06609	52.37930	19.25	14.77	-0.11	-0.33	0.07	0.07
76	PQRL	6.20609	42.98328	20.95	14.91	0.41	-0.72	0.94	0.89
77	PUYA	6.47895	44.85767	20.72	15.35	0.61	-0.26	0.47	0.42
78	PUYV	3.87887	45.04362	19.63	15.59	0.07	-0.16	0.50	0.49
79	RABU	6.97706	44.26782	20.60	15.47	0.25	-0.11	0.70	0.45
80	RG00	8.79393	41.55884	21.26	16.16	-0.10	0.70	0.66	0.70
81	ROSD	6.62823	45.69149	19.45	15.49	-0.47	-0.11	0.47	0.40
82	RSTL	5.48370	43.94096	20.01	16.05	-0.15	0.38	0.59	0.70
83	RSTO	14.00148	42.65839	22.98	17.86	0.95	2.84	0.26	0.26
84	SAUV	4.46692	44.25545	19.83	15.67	-0.05	-0.05	0.12	0.16
85	SAVI	7.66067	44.64764	20.48	15.25	0.09	-0.28	0.70	0.64
86	SFER	-6.20565	36.46435	16.27	16.78	-3.74	0.78	0.31	0.24
87	SJDV	4.67658	45.87908	19.52	15.62	0.04	-0.09	0.09	0.05
88	SOFI	23.39473	42.55609	23.94	11.68	0.29	-2.24	0.16	0.12
89	SOPH	7.05410	43.61143	20.67	15.50	0.13	-0.07	0.09	0.09
90	STEY	5.76178	45.23522	19.66	15.81	-0.21	0.16	0.37	0.31
91	STJ9	7.68382	48.62166	19.47	15.47	0.15	-0.06	0.07	0.07
92	TENC	4.28747	45.12457	19.87	15.62	0.26	-0.11	0.49	0.56
93	TLSE	1.48089	43.56069	19.85	15.82	0.35	-0.03	0.16	0.16
94	TORI	7.66128	45.06337	20.33	15.54	0.05	0.01	0.09	0.09
95	TROP	6.60101	43.21948	20.48	15.64	-0.07	0.04	0.35	0.33
96	VERC	8.42090	45.33129	20.24	15.47	-0.12	-0.01	1.76	1.24
97	VFCH	1.71967	47.29420	18.64	15.58	0.15	-0.26	0.12	0.12
98	VILL	-3.95198	40.44359	19.26	16.32	-0.08	0.34	0.30	0.19
99	WLBH	7.35130	48.41517	19.43	15.27	0.12	-0.28	0.17	0.16
100	WSRT	6.60451	52.91461	17.73	16.03	-0.10	0.44	0.05	0.07
101	WTZR	12.87891	49.14420	20.33	15.24	0.11	0.12	0.09	0.09
102	ZIMM	7.46528	46.87710	19.85	16.01	0.09	0.47	0.12	0.09

TABLE A.3: Vitesse du réseau RENAG (continues)
Vitesses (Ve, Vn) et écart-types (Sve,Svn) en mm/an. Longitude, latitude en degrés décimaux.

A.2 Résidus de vitesse des sites pour définir la pôle de rotation

		Position		Résidu		Ecart-type	
No	Site	long.	lat.	Ve	Vn	Sve	Svn
1	AXPV	5.33320	43.49121	-0.00	0.13	0.37	0.37
2	BANN	4.15632	44.36920	0.22	0.07	0.37	0.37
3	CHAM	5.88108	45.11071	-0.06	-0.27	0.37	0.37
4	CHRN	4.86177	43.88137	0.03	0.05	0.18	0.18
5	CHTL	6.35856	45.30414	-0.40	-0.15	0.18	0.18
6	FCLZ	5.98568	45.64300	0.08	0.14	0.18	0.18
7	GINA	5.78698	43.67572	-0.18	0.01	0.18	0.18
8	GRAS	6.92057	43.75474	-0.06	0.21	0.18	0.18
9	JOUX	5.79566	46.52857	0.16	0.25	0.18	0.18
10	LFAZ	5.39846	45.11663	0.18	-0.04	0.37	0.37
11	MARS	5.35379	43.27877	-0.30	0.00	0.18	0.37
12	MICH	5.71735	43.92416	-0.18	-0.00	0.18	0.18
13	MODA	6.71009	45.21378	-0.13	-0.19	0.18	0.18
14	NICE	7.29998	43.72550	0.24	0.05	0.18	0.18
15	PUYA	6.47895	44.85767	0.61	-0.26	0.55	0.37
16	PUYV	3.87887	45.04362	0.07	-0.16	0.55	0.55
17	RABU	6.97706	44.26782	0.25	-0.11	0.73	0.55
18	ROSD	6.62823	45.69149	-0.46	-0.11	0.55	0.37
19	RSTL	5.48370	43.94096	-0.15	0.38	0.55	0.73
20	SAUV	4.46692	44.25545	-0.05	-0.05	0.18	0.18
21	SJDV	4.67658	45.87908	0.04	-0.09	0.00	0.00
22	SOPH	7.05410	43.61143	0.13	-0.07	0.18	0.18
23	STEY	5.76178	45.23522	-0.21	0.16	0.37	0.37
24	TENC	4.28747	45.12457	0.26	-0.11	0.55	0.55
25	TORI	7.66128	45.06337	0.05	0.01	0.18	0.18
26	TROP	6.60101	43.21948	-0.07	0.04	0.37	0.37
		rms = 0.19 mm/yr				wrms = 0.14 mm/yr	

TABLE A.4: Résidus de vitesse des sites pour définir la pôle de rotation
Vitesses (Ve, Vn) et écart-types (Sve, Svn) en mm/an. Longitude, latitude en degrés décimaux.

A.3 Vitesses verticales du réseau RENAG dans l'ITRF2008

No	Site	Position long.	Position lat.	Vitesse ITRF2008 Vup
1	ABER	-2.08022	57.14401	0.98
2	ABYW	-4.00438	52.42408	1.48
3	AIGL	3.58126	44.12140	0.51
4	AJAC	8.76261	41.92746	0.34
5	ALAC	-0.48123	38.33892	0.25
6	ALES	8.61633	44.92315	-0.97
7	ASTI	8.20320	44.90570	-0.07
8	AXPV	5.33320	43.49121	0.48
9	BANN	4.15632	44.36920	0.87
10	BARK	0.09690	51.51538	-0.60
11	BIEL	8.04806	45.56075	0.10
12	BOGO	21.03534	52.47595	0.36
13	BRST	-4.49659	48.38049	0.15
14	BRUS	4.35922	50.79782	0.46
15	BSCN	5.98939	47.24688	0.48
16	BZRG	11.33680	46.49902	1.29
17	CAGL	8.97275	39.13591	0.45
18	CAMB	-5.32737	50.21844	0.60
19	CARL	-2.93827	54.89543	0.60
20	CAST	7.70930	45.38671	0.61
21	CHAM	5.88108	45.11071	1.14
22	CHIZ	-0.40770	46.13346	0.18
23	CHRN	4.86177	43.88137	0.47
24	CHTL	6.35856	45.30414	2.71
25	COLC	0.89725	51.89437	-0.12
26	CUBX	-0.56639	44.86886	-0.83
27	CUNE	7.55357	44.39499	-1.15
28	DARE	-2.64049	53.34481	0.36
29	DUNK	-3.23989	50.86046	-0.35
30	EBRE	0.49236	40.82089	-0.07
31	EDIN	-3.29479	55.92479	1.41
32	EGLT	2.05200	45.40335	0.88
33	ELBA	10.21110	42.75290	0.53
34	ENTZ	7.63990	48.54936	-0.61
35	ESCO	0.97567	42.69358	0.12
36	FCLZ	5.98568	45.64300	0.93
37	FJCP	2.79490	43.04816	0.78
38	GENO	8.92114	44.41939	-0.09
39	GINA	5.78698	43.67572	0.38
40	GRAS	6.92057	43.75474	0.37
41	GRAV	7.01659	45.12772	0.34
42	GRAZ	15.49348	47.06713	0.36
43	GROI	-3.50798	47.64801	0.28
44	HEAU	-1.77796	49.58066	-0.11
45	IESG	-1.19228	52.94069	0.63
46	INGR	12.51480	41.82808	1.72
47	INVE	-4.21926	57.48625	1.06
48	JOUX	5.79566	46.52857	0.81
49	JOZE	21.03154	52.09727	0.89
50	KOSG	5.80964	52.17843	-0.29
51	LACA	2.72758	43.68099	-0.44
52	LAMP	12.60566	35.49977	0.05
53	LERW	-1.18494	60.13914	0.78
54	LFAZ	5.39846	45.11663	1.04
55	LOWE	1.75020	52.47322	-0.46
56	LROC	-1.21932	46.15894	0.05
57	MALL	2.62455	39.55263	1.08
58	MANS	0.15528	48.01862	0.49
59	MARS	5.35379	43.27877	0.29
60	MATE	16.70446	40.64913	0.85
61	MEDI	11.64681	44.51996	-1.16
62	MICH	5.71735	43.92416	0.59
63	MLVL	2.58731	48.84106	0.91
64	MODA	6.71009	45.21378	2.29
65	MOND	7.82720	44.38894	0.42
66	MTPL	3.86484	43.63744	0.15
67	NEWC	-1.61657	54.97913	0.92
68	NICE	7.29998	43.72550	0.61
69	NOT1	14.98979	36.87584	-0.43
70	NOVA	8.61397	45.44723	1.22
71	ONSA	11.92552	57.39530	2.63
72	ORID	20.79405	41.12731	-0.56
73	PAVI	9.13614	45.20298	-0.75
74	PENC	19.28153	47.78960	-0.38
75	POTS	13.06609	52.37930	-0.03
76	PQRL	6.20609	42.98328	-0.23
77	PUYA	6.47895	44.85767	1.29
78	PUYV	3.87887	45.04362	0.18
79	RABU	6.97706	44.26782	0.24
80	RG00	8.79393	41.55884	0.89
81	ROSD	6.62823	45.69149	1.00
82	RSTL	5.48370	43.94096	0.19
83	RSTO	14.00148	42.65839	0.39
84	SAUV	4.46692	44.25545	0.60
85	SAVI	7.66067	44.64764	0.51
86	SFER	-6.20565	36.46435	0.35
87	SJDV	4.67658	45.87908	0.21
88	SOFI	23.39473	42.55609	0.10
89	SOPH	7.05410	43.61143	0.24
90	STEY	5.76178	45.23522	1.91
91	STJ9	7.68382	48.62166	0.42
92	TENC	4.28747	45.12457	0.26
93	TLSE	1.48089	43.56069	0.39
94	TORI	7.66128	45.06337	0.97
95	TROP	6.60101	43.21948	-0.08
96	VERC	8.42090	45.33129	-1.43
97	VFCH	1.71967	47.29420	1.31
98	VILL	-3.95198	40.44359	-0.58
99	WLBH	7.35130	48.41517	-1.37
100	WSRT	6.60451	52.91461	-0.25
101	WTZR	12.87891	49.14420	-0.03
102	ZIMM	7.46528	46.87710	1.50

TABLE A.5: Vitesses verticales dans l'ITRF2008 du réseau RENAG
Vitesses en mm/an. Longitude, latitude en degrés décimaux.

B Résultats du réseau GPS du Taiwan

		Position		Vitesse dans l'ITRF2008					Position		Vitesse dans l'ITRF2008		
No	Site	Long.	Lat.	Ve	Vn	Vu	No	Site	Long.	Lat.	Ve	Vn	Vu
1	8118	120.55298	23.46298	17.43	-12.23	3.66	41	DASI	120.94442	22.47843	-1.97	3.84	5.92
2	AKND	120.35726	22.80331	-5.02	-20.77	2.20	42	DAWU	120.89005	22.34060	-18.21	-0.45	-2.06
3	ALIC	133.88552	-23.67012	32.17	59.53	-0.50	43	DCHU	121.28057	23.21318	-13.08	27.61	6.67
4	ALIS	120.81330	23.50817	2.04	-5.80	1.81	44	DNAN	120.44799	23.67381	28.09	-9.60	-12.94
5	ANKN	121.52473	24.95910	32.99	-10.17	-0.78	45	DNFU	121.48229	23.68513	-5.57	15.73	-7.48
6	BALN	121.42612	24.69949	27.34	-3.52	-1.36	46	DONA	120.70352	22.91562	-22.63	-11.24	5.42
7	BANC	121.44211	24.99766	31.76	-10.53	-0.02	47	DONS	120.15378	23.46049	31.26	-17.13	-32.30
8	BANP	120.30541	22.69314	-15.97	-35.88	1.02	48	DOSH	120.82688	24.26269	25.11	-7.45	2.78
9	BDES	120.17189	23.38058	29.24	-11.05	-33.97	49	DPIN	120.93281	24.04308	9.24	-1.04	4.48
10	BLOW	121.57125	24.17176	22.80	-5.72	3.89	50	DSIN	121.39804	23.63122	2.11	9.31	-6.11
11	C001	120.61240	23.41795	5.38	-5.79	4.61	51	DULI	121.33060	23.02567	-17.45	32.54	3.28
12	C002	120.57719	23.36175	5.67	-5.78	7.31	52	ERLN	120.41956	23.79759	31.11	-9.84	-19.98
13	CAOT	120.68874	23.97939	24.60	-10.17	-2.92	53	ERPN	121.16612	22.94218	-22.75	28.46	5.81
14	CHEN	121.37358	23.09741	-16.96	29.49	14.83	54	FALI	120.59360	22.36526	-26.77	-12.54	-9.25
15	CHGO	121.37454	23.09833	-8.66	42.73	4.01	55	FCU1	120.64952	24.17834	27.11	-11.79	-0.53
16	CHIA	120.43320	23.49598	22.13	-10.78	-2.62	56	FENP	121.51942	23.59845	-14.90	33.92	-9.04
17	CHKU	120.09276	23.05585	28.27	-13.24	-3.45	57	FIVE	121.78108	25.07105	35.04	-13.49	-1.19
18	CHNL	120.56347	23.37715	7.07	-5.97	19.89	58	FKDO	120.85628	23.68356	6.22	1.26	-4.92
19	CHNT	121.66190	24.14921	20.51	-11.31	-2.77	59	FLNM	121.45336	23.74631	-5.38	15.30	-7.85
20	CHSG	120.28915	23.86035	31.24	-11.80	0.28	60	FLON	121.93748	25.02037	35.20	-13.38	-1.18
21	CHUA	120.55730	24.06605	29.12	-12.29	-0.99	61	FNGU	120.72445	24.03517	20.16	-7.35	4.45
22	CHUK	120.60472	23.43871	14.01	-9.42	-7.89	62	FONB	121.52204	23.59817	-16.59	35.04	-10.49
23	CHUL	121.12569	23.13237	2.11	2.77	1.29	63	FONL	121.45348	23.74631	1.61	8.87	-11.02
24	CHUN	121.39310	23.45286	-14.08	28.31	-5.13	64	FUGN	121.19217	22.79075	-25.54	25.42	4.47
25	CHYI	120.14016	23.45077	30.03	-17.41	-31.68	65	FUNY	120.32017	23.92232	32.07	-12.40	-29.29
26	CHYN	120.29080	23.39327	26.35	-12.99	-7.01	66	FUQE	120.82331	24.01116	18.53	-3.52	0.53
27	CISH	120.48123	22.88959	-16.02	-16.49	1.44	67	FUSB	121.35285	24.82121	25.66	-2.68	-2.78
28	CK01	120.21047	22.97589	18.94	-14.70	2.63	68	FUSI	121.35285	24.82121	28.43	-7.69	0.28
29	CKST	120.22001	22.99885	18.42	-13.40	-3.26	69	FUSN	121.33147	24.79905	30.27	-7.03	0.70
30	CKSV	120.22001	22.99885	20.31	-13.18	2.72	70	GANK	120.82257	24.02840	13.14	-5.18	-1.15
31	CLAN	121.51201	24.60225	34.19	-7.02	4.54	71	GFES	120.40249	23.74141	26.99	-16.14	-43.65
32	CLON	120.57960	22.43005	-24.99	-12.42	-13.32	72	GLES	120.60776	23.61215	13.51	-9.32	2.79
33	COCO	96.83397	-12.18835	44.16	50.58	-1.11	73	GOLI	121.98743	25.02039	34.30	-12.22	-1.37
34	CTOU	120.27785	22.75469	-7.46	-24.76	0.35	74	GS01	121.50816	24.98200	32.88	-9.72	-0.80
35	CWEN	120.45276	23.47305	21.53	-8.39	0.11	75	GS03	121.04407	24.77660	27.54	-7.18	-0.11
36	DACU	120.43730	23.21893	7.68	-3.49	7.36	76	GS04	120.50679	23.59152	25.08	-9.26	0.30
37	DAEJ	127.37448	36.39943	26.56	-12.16	0.96	77	GS05	120.56840	23.56711	17.91	-8.17	-0.12
38	DAHU	120.87184	24.42288	25.53	-12.37	-1.67	78	GS06	120.55422	23.46561	15.56	-9.12	1.93
39	DAJN	120.86498	22.31130	-14.36	1.95	-0.76	79	GS07	120.65483	23.48291	12.05	-7.91	5.93
40	DANL	120.75191	22.20413	-13.65	-1.67	14.02	80	GS08	121.50149	25.20373	32.14	-10.00	-0.91

TABLE B.1: Vitesses verticales dans l'ITRF2008 du réseau GPS du Taiwan Vitesses (Ve, Vn, Vn) en mm/an. Longitude, latitude en degrés décimaux.

		Position		Vitesse dans l'ITRF2008					Position		Vitesse dans l'ITRF2008		
No	Site	Long.	Lat.	Ve	Vn	Vu	No	Site	Long.	Lat.	Ve	Vn	Vu
81	GS09	121.65193	25.20861	33.97	-11.62	0.37	121	GS49	121.18450	24.02599	10.47	-15.88	-15.82
82	GS10	121.46155	25.14615	30.93	-10.89	-5.36	122	GS50	121.18335	24.02315	14.21	-19.00	8.04
83	GS11	121.49877	25.13357	31.00	-10.35	-2.77	123	GS51	120.54815	22.99846	-16.78	-17.97	6.56
84	GS12	121.38276	25.05303	32.64	-10.23	-0.28	124	GS52	120.66281	23.03233	-28.71	-3.44	30.40
85	GS13	121.45224	25.02001	31.85	-9.06	-2.35	125	GS53	120.40639	22.83823	-8.73	-16.90	11.18
86	GS14	120.95945	24.80322	30.25	-9.37	-1.39	126	GS54	120.46018	22.83535	-32.19	-20.63	5.40
87	GS15	120.99050	24.76665	29.61	-8.42	-0.23	127	GS55	120.61033	22.84893	-30.94	-14.73	12.04
88	GS16	121.04161	24.74787	28.99	-8.19	-0.24	128	GS56	120.60981	22.70213	-29.35	-15.02	3.39
89	GS17	120.60584	23.56115	13.93	-7.20	2.24	129	GS57	120.77289	22.12576	-15.45	-5.37	2.60
90	GS18	120.47378	23.48498	21.15	-9.89	-0.91	130	GS58	120.84696	22.12461	-8.68	1.66	0.66
91	GS19	121.67982	25.07226	35.06	-11.12	0.33	131	GS59	120.71242	22.05564	-17.31	-11.63	0.14
92	GS20	121.93998	24.84251	45.06	-19.86	-19.02	132	GS60	120.81776	22.00577	-4.79	3.44	1.23
93	GS21	120.51238	24.09731	29.36	-11.86	-3.29	133	GUAM	144.86836	13.58933	-9.13	5.75	0.16
94	GS22	120.60628	24.09087	27.48	-10.78	-2.55	134	GUFU	121.39095	23.60740	-0.50	11.12	-9.91
95	GS23	120.64769	24.09459	27.37	-9.74	1.82	135	GUKN	120.58878	23.64587	23.12	-10.61	-1.78
96	GS24	120.52609	24.02523	28.98	-11.59	-8.59	136	GUKW	121.00655	24.20220	15.37	0.34	6.03
97	GS25	120.64022	23.98651	26.59	-9.45	-1.16	137	HANS	121.68714	24.60951	51.33	-14.55	0.33
98	GS26	120.64602	23.86984	26.16	-10.38	-1.54	138	HENC	120.74645	22.00393	-21.03	-9.85	-3.21
99	GS27	120.57788	23.82475	27.70	-10.64	-3.38	139	HERI	121.58092	25.29444	31.73	-11.76	0.99
100	GS28	120.21436	23.08099	25.44	-12.42	-2.92	140	HNES	120.47914	23.94836	31.60	-4.97	-32.86
101	GS29	120.31582	23.07513	21.24	-13.58	5.23	141	HNSN	121.30807	24.33770	24.15	0.21	8.97
102	GS30	120.22629	23.02048	19.84	-12.22	0.73	142	HOKN	120.13488	23.18839	28.77	-13.39	-7.06
103	GS31	120.27577	23.01890	16.61	-13.68	-2.13	143	HRGN	121.40511	23.55529	-8.89	18.56	-9.32
104	GS32	120.33709	23.02579	13.67	-15.64	1.20	144	HSIN	121.01426	24.82775	30.08	-9.92	-0.76
105	GS33	120.18782	22.96443	20.38	-12.34	-1.46	145	HSUE	121.02645	24.28060	13.78	1.04	-5.08
106	GS34	120.27514	22.93923	10.29	-13.53	1.30	146	HUAN	121.27263	24.14348	10.96	7.03	2.85
107	GS35	120.30938	22.93552	9.09	-14.52	2.14	147	HUAP	121.74945	24.30900	32.34	-26.52	-1.13
108	GS36	120.62533	24.36205	29.22	-10.90	-1.44	148	HUWE	120.28663	23.72940	30.11	-10.32	-31.23
109	GS37	120.71544	24.32858	27.11	-10.81	-1.52	149	HUYS	121.02941	24.09233	6.56	-0.75	7.23
110	GS38	120.60059	24.27981	29.66	-12.24	-2.65	150	ICHU	120.27931	23.36066	25.81	-12.24	-6.97
111	GS39	120.52960	24.18441	29.69	-11.51	-2.52	151	IISC	77.57038	13.02117	41.27	36.19	0.96
112	GS40	120.57631	24.18124	29.29	-10.63	-2.63	152	ILAN	121.75662	24.76404	39.68	-16.70	-7.52
113	GS41	120.44493	23.37756	20.75	-8.80	-3.47	153	JHCI	120.54742	23.51373	17.69	-8.49	1.82
114	GS42	120.45201	23.27323	11.62	-7.00	7.26	154	JLUT	120.62279	22.32996	-25.96	-12.20	-2.54
115	GS43	120.37361	23.25716	22.57	-9.18	0.43	155	JNHU	120.57672	23.61526	21.70	-6.55	-0.73
116	GS44	120.40040	23.22225	15.65	-9.50	10.13	156	JONP	120.52401	23.42297	17.51	-8.52	2.65
117	GS45	120.72817	22.74911	-26.20	-9.12	11.16	157	JPEI	121.37140	23.53160	-8.47	17.38	-10.57
118	GS46	120.64946	22.52753	-24.81	-10.03	4.49	158	JPIN	121.35889	23.34108	-10.44	32.60	9.10
119	GS47	120.74525	21.94875	-4.59	-6.39	-24.78	159	JSUI	121.42389	23.49198	-13.58	30.01	2.95
120	GS48	121.18858	24.02361	11.66	-1.38	8.87	160	JUCI	120.55421	23.52236	18.51	-7.17	-0.18

TABLE B.2: Vitesses verticales dans l'ITRF2008 du réseau GPS du Taiwan (continues) Vitesses (Ve, Vn, Vn) en mm/an. Longitude, latitude en degrés décimaux.

		Position		Vitesse dans l'ITRF2008					Position		Vitesse dans l'ITRF2008		
No	Site	Long.	Lat.	Ve	Vn	Vu	No	Site	Long.	Lat.	Ve	Vn	Vu
161	JULI	121.31817	23.34167	2.12	10.67	-16.45	201	MFEN	121.17248	24.08216	-0.03	5.71	6.10
162	JUNA	120.87537	24.68395	28.33	-8.28	2.95	202	MIAO	120.81027	24.58345	28.09	-10.90	-3.04
163	JYAN	121.22635	24.24247	21.96	8.74	4.08	203	MINS	120.72091	23.27540	-1.83	-6.46	5.34
164	KAFN	121.11649	23.98759	-27.69	8.94	9.82	204	MITO	120.26316	22.79586	6.68	-18.88	-0.65
165	KASH	120.28835	22.61450	-17.80	-36.87	-2.28	205	MLO1	120.55382	22.89996	-19.71	-16.01	-0.27
166	KASU	120.63298	22.81019	-23.17	-10.55	3.94	206	MLON	120.59584	22.89666	-17.88	-14.87	0.73
167	KDNM	120.78202	21.94943	-17.43	-3.91	6.95	207	MOTN	121.02685	23.20055	-0.95	-5.58	3.46
168	KMNM	118.38858	24.46382	32.03	-12.66	-0.14	208	MSPK	120.55731	23.57794	21.14	-9.25	0.28
169	KNKO	121.50574	23.47222	-17.86	37.44	-6.81	209	MZUM	119.93318	26.15711	32.30	-12.46	0.32
170	KOKB	-159.66493	22.12626	-62.77	34.98	0.98	210	NAAO	121.81021	24.44935	38.70	-29.55	-7.45
171	KSHI	121.17604	24.77667	29.76	-8.23	-0.08	211	NANK	120.27440	23.10199	23.66	-14.23	1.66
172	KTES	120.33429	23.62662	28.75	-7.96	-49.97	212	NCKU	120.27581	22.93845	9.43	-14.60	2.41
173	KULN	120.50703	23.33099	4.21	-9.86	7.53	213	NCTU	120.99732	24.78883	30.95	-11.26	1.07
174	KUNM	102.79720	25.02954	31.19	-17.70	1.33	214	NDHU	121.55082	23.89724	3.27	7.00	-7.67
175	KYIN	121.08042	25.04105	29.71	-10.89	1.45	215	NEMN	120.42008	22.90812	-6.22	-16.30	8.56
176	KZN1	120.69798	23.94626	23.82	-10.56	-2.52	216	NHSI	121.45302	23.40622	-13.78	35.33	-3.44
177	KZON	120.69750	23.94569	22.46	-13.69	-8.74	217	NIUT	121.56158	24.63478	36.61	-6.36	1.75
178	LAIG	120.73989	23.52940	4.18	-7.55	5.84	218	NJOU	120.57141	22.50390	-26.62	-12.18	-10.76
179	LANB	121.55815	22.03721	-37.68	36.89	-0.82	219	NSAN	121.38281	24.42820	27.81	-1.41	6.28
180	LANY	121.55810	22.03731	-38.96	40.49	1.22	220	NSHE	120.80091	24.22580	17.02	-7.84	2.23
181	LAOL	120.68730	22.41192	-23.10	-10.35	1.16	221	PAK1	120.63604	23.91568	27.12	-4.12	-33.75
182	LGUE	120.63538	22.99291	-16.45	-11.86	8.91	222	PAKU	120.63604	23.91568	25.42	-10.71	-2.79
183	LIKN	120.52791	22.75861	-25.88	-15.39	4.37	223	PANG	119.56374	23.56521	33.60	-14.61	3.17
184	LIUC	120.36907	22.34668	-20.69	-38.02	2.61	224	PAOL	120.70287	23.10862	-14.95	-10.83	7.38
185	LIYU	120.78181	24.34306	26.07	-11.70	3.69	225	PAOS	120.95028	24.71493	29.01	-7.99	1.65
186	LNDO	121.91808	25.09738	33.77	-13.94	-3.34	226	PEI1	120.16859	23.29384	28.83	-13.39	-26.47
187	LNJS	120.59216	23.75745	26.13	-8.12	-2.79	227	PEIM	120.17284	23.29217	32.97	-11.83	-14.16
188	LNKO	121.37818	25.07637	32.80	-8.71	-3.17	228	PEIN	121.12313	22.80108	-18.30	-12.53	-0.47
189	LONT	121.13056	22.90632	-10.45	11.97	-1.92	229	PEPU	121.61034	24.01789	11.77	-1.95	-2.11
190	LOYE	120.70120	23.46863	8.01	-6.65	8.17	230	PERT	115.88525	-31.80197	39.11	58.03	-2.75
191	LSAN	121.18217	24.02935	5.62	-0.74	7.73	231	PIMO	121.07773	14.63572	-28.97	7.33	2.05
192	LSB0	120.63423	24.15337	27.53	-11.01	-1.57	232	PING	121.45434	23.31946	-22.73	36.83	-4.87
193	LTUN	121.77163	24.70003	42.08	-18.44	-3.40	233	PINT	120.64057	23.75607	25.48	-10.59	-3.99
194	LUDA	121.47589	22.65810	-41.05	35.01	-0.19	234	PKGM	120.30550	23.57989	29.94	-9.39	-22.71
195	LUGU	120.74769	23.75227	12.08	-3.22	1.76	235	PLAN	121.08661	24.57896	29.36	-3.07	1.45
196	LUJA	120.34878	23.22703	23.00	-10.11	-0.64	236	PLIM	120.98200	23.97389	2.72	-0.84	5.18
197	LUKN	120.43513	24.06001	29.11	-12.69	-5.01	237	PLIN	121.71395	24.93363	33.42	-11.48	-1.33
198	MAJA	120.65206	22.70765	-26.69	-9.81	-1.64	238	PNCY	122.07944	25.62749	32.18	-12.16	4.65
199	MATZ	119.92307	26.16940	29.38	-12.93	-1.25	239	PTUN	120.45969	22.64985	-25.06	-20.30	7.05
200	MESN	120.82626	23.26364	-0.27	-10.73	6.73	240	PUSN	120.52011	23.96494	29.00	-12.06	-11.04

TABLE B.3: Vitesses verticales dans l'ITRF2008 du réseau GPS du Taiwan (continues) Vitesses (Ve, Vn, Vn) en mm/an. Longitude, latitude en degrés décimaux.

		Position		Vitesse dans l'ITRF2008					Position		Vitesse dans l'ITRF2008		
No	Site	Long.	Lat.	Ve	Vn	Vu	No	Site	Long.	Lat.	Ve	Vn	Vu
241	RENI	120.50857	23.45894	18.37	-9.95	0.49	281	STA1	120.97926	24.63560	28.89	-7.99	3.19
242	S011	120.33945	23.20542	23.58	-11.66	-0.81	282	STAN	120.87307	24.48728	33.17	-5.36	-8.77
243	S012	120.48827	23.05948	-4.64	-11.11	15.06	283	SUAB	121.86790	24.59391	50.65	-37.65	-3.11
244	S01R	119.59237	23.65527	30.56	-13.70	-0.48	284	SUAN	120.29988	23.47758	26.92	-11.00	-6.57
245	S058	121.45436	23.31943	-20.57	41.34	-5.13	285	SUAO	121.86708	24.59239	50.87	-37.41	-3.23
246	S092	120.52847	23.18490	-2.49	-8.93	6.64	286	SUN1	120.90835	23.88124	1.60	2.63	4.54
247	S101	121.61386	25.04045	34.12	-11.87	-0.76	287	SUNM	120.90825	23.88117	-37.92	30.09	50.71
248	S102	121.55815	22.03721	-38.26	37.57	3.84	288	SYAN	120.98675	23.24800	29.56	-6.53	-20.53
249	S103	120.47519	23.56438	23.61	-8.43	-0.33	289	T101	121.32362	23.02034	-18.57	36.54	17.30
250	S104	121.18939	22.82076	-23.76	30.52	1.95	290	T102	121.27680	23.01601	-13.06	38.72	26.76
251	S105	121.11290	22.95166	1.02	2.61	-5.14	291	T103	121.26739	23.15501	-14.41	40.90	28.35
252	S106	120.33410	23.05080	15.87	-16.39	3.29	292	T104	121.24018	23.15482	-17.97	35.79	22.20
253	S169	120.50331	22.94230	-15.41	-13.04	2.15	293	T105	121.22735	23.15546	5.85	3.48	-10.07
254	S23R	120.60618	22.64499	-24.01	-10.99	-1.56	294	TACH	120.53513	24.29076	29.26	-12.32	-0.87
255	SALU	120.57831	24.14445	28.67	-10.80	-2.35	295	TAIP	121.67317	25.03453	33.93	-11.79	-1.34
256	SAND	120.64065	22.71727	-26.08	-9.97	1.52	296	TANS	121.42690	25.18150	27.68	-10.92	-3.39
257	SANI	120.76873	24.41435	27.86	-9.43	-13.61	297	TAPE	121.23088	23.12557	-3.07	7.22	-10.45
258	SANJ	121.50087	25.26080	31.34	-11.77	-1.39	298	TAPN	120.60685	23.56131	14.27	-7.25	1.63
259	SANL	120.76856	23.66448	7.19	-1.10	2.92	299	TAPO	121.23742	23.12706	-19.97	27.99	6.37
260	SCHN	121.65160	24.12779	18.00	-8.07	1.06	300	TAPP	121.61398	25.04032	32.29	-9.34	-8.28
261	SFON	121.01016	24.93290	30.11	-11.38	-2.03	301	TASI	120.18882	23.72027	34.67	-12.63	-29.70
262	SGAN	120.34966	22.58127	-13.98	-31.81	-0.78	302	TASO	120.69508	24.46127	28.55	-11.08	0.45
263	SGUN	120.69189	24.27163	27.54	-10.26	-3.07	303	TATA	120.88705	23.48140	-20.55	5.65	-1.05
264	SHAN	121.19949	23.10886	-6.33	10.93	-6.21	304	TAYN	120.76420	23.15935	-6.33	-7.62	8.74
265	SHJU	120.92195	24.84508	28.58	-11.26	-2.64	305	TC01	120.67828	24.13231	25.36	-9.90	-3.72
266	SHMN	121.56216	25.29146	31.63	-10.56	-0.55	306	TC11	120.59922	24.15279	26.59	-11.81	-4.16
267	SHUL	121.56274	23.78762	2.40	19.91	-10.86	307	TC12	120.59794	24.20752	27.60	-12.48	-4.83
268	SHWA	120.34781	23.02143	12.32	-15.89	2.34	308	TC13	120.79342	24.20585	15.70	-8.69	0.61
269	SICH	121.65442	24.12573	18.43	-8.51	-1.06	309	TC21	120.67377	24.13629	25.64	-9.55	-3.32
270	SILN	120.64604	23.16039	-4.43	-10.46	8.58	310	TC22	120.68064	24.17087	24.45	-11.60	-4.22
271	SIND	121.54238	24.96680	33.11	-10.63	-0.70	311	TC23	120.65324	24.15374	27.24	-11.21	-5.25
272	SINL	121.25463	22.90829	-19.14	29.89	3.07	312	TCB2	120.67377	24.13629	28.52	-13.56	-2.46
273	SINY	120.85321	23.69649	4.20	-1.55	7.73	313	TCBB	120.68064	24.17087	28.29	-13.82	-4.85
274	SJPU	120.48126	23.42854	19.43	-9.92	1.03	314	TCBC	120.65324	24.15374	28.66	-14.64	-5.69
275	SLIN	121.44141	23.81186	4.50	7.11	-2.40	315	TCMS	120.98739	24.79799	29.87	-9.20	-0.05
276	SLNP	121.63564	24.75312	35.27	-12.38	-0.93	316	TCYU	121.23082	24.94740	35.16	-6.78	-13.58
277	SOFN	121.59817	23.87026	3.27	14.65	-13.64	317	TEGS	120.65499	24.35621	28.79	-12.11	-2.08
278	SONA	120.98585	24.39782	21.07	-5.25	0.25	318	TENC	120.91567	23.27779	3.92	-5.87	2.56
279	SPAO	121.48487	24.20502	23.07	-2.11	10.42	319	THAI	121.29560	24.60712	29.02	-2.53	2.41
280	SSUN	120.37778	23.41415	22.23	-10.50	-1.11	320	TIAN	120.94072	24.46504	24.41	-4.24	-4.53

TABLE B.4: Vitesses verticales dans l'ITRF2008 du réseau GPS du Taiwan (continues) Vitesses (Ve, Vn, Vn) en mm/an. Longitude, latitude en degrés décimaux.

		Position		Vitesse dans l'ITRF2008					Position		Vitesse dans l'ITRF2008		
No	Site	Long.	Lat.	Ve	Vn	Vu	No	Site	Long.	Lat.	Ve	Vn	Vu
321	TIDB	148.98000	-35.39920	18.30	55.99	-1.01	349	WFEN	120.69947	24.04195	24.19	-9.38	-3.16
322	TKJS	120.38983	23.68800	30.41	-11.13	-63.86	350	WIPN	121.05856	24.67464	28.77	-6.75	-0.44
323	TMAL	120.95988	22.64892	-6.04	-4.34	2.14	351	WUCH	121.02723	23.92377	-5.88	0.06	6.32
324	TNML	120.98735	24.79795	29.82	-9.46	-0.64	352	WUFN	120.69939	24.04278	21.88	-11.72	-2.42
325	TNSM	116.72456	20.70264	32.00	-12.43	1.32	353	WUHN	114.35726	30.53165	32.13	-11.90	0.19
326	TOFN	120.92481	24.66204	28.41	-9.08	1.17	354	WUKU	121.40067	25.11728	31.68	-10.18	-1.97
327	TSHI	121.63277	25.25691	32.74	-11.44	0.94	355	WULI	121.30846	24.35228	24.80	-0.01	8.45
328	TSHO	120.35186	22.90422	2.86	-18.09	-0.68	356	WULU	121.04148	23.16926	6.77	-1.55	17.57
329	TSIO	120.70410	24.47280	29.18	-10.67	0.39	357	WUST	120.36817	23.20524	19.69	-12.28	2.54
330	TSKB	140.08750	36.10568	-3.85	-7.34	1.16	358	YAME	121.18530	24.90846	30.18	-8.90	1.69
331	TSLN	120.71939	23.63433	11.12	-5.86	6.53	359	YENC	120.37615	22.76352	-27.20	-22.99	3.59
332	TTSH	121.14763	22.74714	-15.24	-4.64	-6.89	360	YENL	121.60185	23.90350	3.70	14.57	-9.06
333	TTUN	121.08071	22.76455	-12.12	-8.38	-6.85	361	YILN	121.74565	24.74608	39.80	-14.78	-6.33
334	TUNH	121.30022	23.07516	-11.47	24.78	-6.54	362	YM01	121.53955	25.17888	32.38	-10.53	0.82
335	TUNS	120.40405	23.31724	22.55	-11.24	-1.77	363	YM02	121.56069	25.14838	33.66	-10.58	-1.39
336	TWTF	121.16450	24.95357	32.35	-12.25	5.35	364	YM03	121.60491	25.14662	34.52	-11.44	-0.85
337	USUD	138.36204	36.13311	0.02	-7.13	-1.82	365	YM04	121.58547	25.18761	32.37	-10.96	-1.21
338	VR01	120.49824	24.11717	27.41	-14.10	-2.07	366	YM05	121.56548	25.16401	32.10	-11.02	-0.85
339	VR02	120.37348	23.89994	26.89	-14.84	-76.43	367	YMSM	121.57406	25.16571	32.15	-12.70	1.47
340	VR03	120.58152	23.86443	29.96	-14.36	-1.48	368	YNTS	121.77894	24.86174	35.77	-11.85	1.00
341	W025	120.52401	23.42264	14.30	-10.04	-7.93	369	YOKN	120.55419	23.62610	24.74	-7.22	-3.57
342	W029	120.66425	23.54080	2.70	-11.15	8.05	370	YSAN	120.08598	23.14655	30.76	-14.25	-4.60
343	W030	120.69553	23.47411	-0.90	-12.05	4.51	371	YSLL	120.22234	23.73917	30.15	-13.92	-30.97
344	WANC	120.52634	23.18684	2.24	-9.17	2.68	372	YUNL	120.79858	23.88389	-8.35	2.58	15.16
345	WANL	121.63756	25.16939	34.21	-11.63	-0.47	373	YUSN	120.95914	23.48730	-61.38	25.15	-2.08
346	WANS	120.88520	23.60751	0.38	-0.35	9.21	374	ZEND	120.21757	22.94328	16.55	-13.52	2.35
347	WARO	121.44093	23.81203	5.24	8.81	-1.71	375	ZWEN	120.49735	23.21974	6.85	-12.52	9.24
348	WDAN	120.50432	22.60606	-21.96	-14.30	-2.71							

TABLE B.5: Vitesses verticales dans l'ITRF2008 du réseau GPS du Taiwan (continues) Vitesses (Ve, Vn, Vn) en mm/an. Longitude, latitude en degrés décimaux.

C Méthode de la Classe Gts "Geodetic time series"

Help on class Gts in module pygps_ts_class_15 :

class Gts

##

define the Gts

##

Methods defined here :

Ftest_4_offsets(self, data, t0)

This method makes a test of significancy of an offset in a time series

The time series is assumed to have ONLY ONE offset provided at provided offset using an F_ratio test

The chi2 of two estimation are compared :

- one without offset

- one with one offset

The F_ratio is then formed to decide wether the offset is significant or not

__init__(self, code=None, lat=None, lon=None, data=None, offsets=None, offsets_values=None, earthquakes=None, outliers=None, post_seismics=None, annual=None, semi_annual=None, velocity=None, loption=None, tminmax=None, rate_change=None, H_conf=None)

data is a numpy array with 7 columns : dec.year, N, E, U, S_N, S_E, S_U (m)

code station 4-letters code

lon,lat approximate longitude and latitude (geodetic, deg.dec)

outliers list of index of outliers

offsets_values 7 columns : dec.year offset_N, offset_E, offset_U, S_offset_N, S_offset_E, S_offset_U (m)

offsets list of dates for offsets

earthquakes 7 columns : dec.year offset_N, offset_E, offset_U, S_offset_N, S_offset_E, S_offset_U (m)

post_seismics 8 columns : seismic_epoch Amplitude_N, Amplitude_E, Amplitude_U, S_Amplitude_N, S_Amplitude_E, S_Amplitude_U, Tau (seismic_epoch in dec.year ; Tau in dec.year ; amplitude in m)

annual 6 columns : Amplitude_N, Phase_N, Amplitude_E, Phase_E, Amplitude_U, Phase_U

semi_annual 6 columns : Amplitude_N, Phase_N, Amplitude_E, Phase_E, Amplitude_U, Phase_U

velocity 6 columns : vel_N, vel_E, vel_U, S_vel_N, S_vel_E, S_vel_U

apply_loffset(self, loffset, in_place=False)

Applies given offsets to a times series

if in_place = True then replace the current time series

copy(self)

makes a copy of the time series

delete_small_offsets(self, offsets, del_by_pricise=False)

The aim for test_offset modul.

Estimate the offsets with clean data.

Then delete the offsets which their values are so small.

input : list of time offsets

output : list of time offsets tested

detrend(self, in_place=False)

detrend a time series

if in_place = True then replace the current time series

detrend_annual(self, in_place=False)

estimates a trend + annual terms in a time series and removes them

if in_place = True then replace the current time series

detrend_seasonal(self, in_place=False)

estimates a trend + annual + semi-annual terms in a time series and removes them

if in_place = True then replace the current time series

differentiate(self)

differentiate time series

estimate_local_offset(self, window_length=4)

Estimate the local offset, just used window_length positions before & window_length positions behind of offset

The Velocity, annual/semi-annual term are removed

output: amplitude of local offsets

extract_lperiod(self, lperiod, in_place=False)

extract periods

return a new Gts for the selected periods

if in_place = True then replace with the extracted time series

find_outliers_and_offsets_through_differentiation(self, th=100)

find outliers and offsets using differenciation

find_outliers_by_RMS_ts(self, ndays=7, th_detection=5, th_rejection=2)

Find index of outliers in a time series, populate self.outliers.

- rms time series are first calculated over ndays

- time windows are kept for further inspection if rms(t) > th_detection * median(rms(ts))

- for each anomalous time windows, differentiate positions, find the max

- test whether it is a true outlier (differentiated(t) > th_rejection * median(differentiated))

output: None

find_outliers_by_residuals(self, option=None, th_percent=2.5, th_rejection=3)

Find index of outliers by trendline/trendline_annual/trendline_seasonal (the complete model)

Then the outliers are detected if their residuals are greater than th_rejection*standard_deviation

Add the outliers into self.outliers

output: Add the list of outlier index into self.outliers

find_outliers_simple(self, threshold=100, window_length=10)
Find the outliers by comparing the positions with their median

find_tau_4_neu(self)

Estimate the relaxation time of post-seismic movement for time series

find_time_offsets(self, option=None, ndays=7, th_detection_rms=3, th_detection_offset=3)

Find the time of suspected offsets by rms time series calculated over ndays

Then check the time of offsets: if one offset is too small/None then it is removed

input:

- ndays: number of positions in the windows. rms time series are calculated over ndays.

- th_detection_rms: the threshold for detecting the anomalous windows rms(t) > th_detection_rms*median(rms(ts)).

- th_detection_offset: the threshold for detecting the offsets, for each anomalous time windows, differentiate positions then test whether it is a suspected offset (differentiated(t) > th_rejection * median(differentiated))

output: add the time of offsets in to self.offsets

fit_exponential(self, t0=None, lperiod=None, in_place=False)

fits an exponential decay function to a time series estimated the time windows provided in lperiod, using non-linear least-squares y(t)=a.exp(-(t-t0)/tau)

Uses scipy.optimize.curve_fit

t0 : the epoch of reference (usually time of the earthquake)

a : amplitude of the post-seismic final displacement

tau : relaxation time

returns a list of 3 tuples (a,tau,s_a,s_tau) for each component typical use for estimating post-seismic displacement returns

info(self, info=2)

print various informations about the time series

make_model(self, option=None, loutlier=None)

Estimate the model of time series by least square estimation

input: data:

GTS format

option are: 'detrend'/'detrend_annual'/'detrend_seasonal'

output:

new Gts object: time series is now the residuals wrt to the model and its associated values (vel, annual, semi-annual etc)

plot(self, loffset=True, loutliers=True, verbose=True, date=[], yaxis=[], lcomponent=['N', 'E', 'U'], error_scale=0.0, lperiod=[], lvline=[], save=None, show=True, unit='mm')

Make plots of a time series and related info (offsets, outliers) using Matplotlib

coordinates are assumed to be in meters

default plots units will be mm; Use unit='m' to get meter

read_mb_file(self, directory)

Read GAMIT/GLOBK mb_file then return self.data

input: directory of mb_file

remove_outliers(self, in_place=False)

removes outliers provided in self.outliers

return a new Gts without the outliers if

in_place = True then self has the outliers

removed as well (in _place)

remove_velocity(self, vel_neu, in_place=False)

remove velocity from a time series vel_neu is a 1D array of any arbitrary length, but with the velocities (NEU) to be removed in the first 3 columns

if in_place = True then replace the current time series

rms(self, ndays)

Returns a rms time series calculated over ndays

test_offsets(self, verbose=False, debug=True, window_length=None)

Test the offset:

- delete the small offset (1mm for East/North, 2mm for Up)

- then make a F ratio test

- re-check to delete the small offsets

TABLE DES FIGURES

1.1 Réseau utilisé pour l'estimation des paramètres de transformation entre ITRF2008 et ITRF2005 (http://itrf.ensg.ign.fr) 7

1.2 Histogrammes de l'erreur sphériques des positions et des vitesses de l'ITRF [Altamimi et al., 2011] 8

1.3 Réseau GPS permanent de l'IGS (http://igscb.jpl.nasa.gov/network/netindex.html) 8

1.4 Schéma décrivant la transformation de Helmert. T : la translation d'origine. R_X, R_Y, R_Z : les rotations suivant les 3 axes X,Y,Z. La transformation de Helmert comprend en plus l'application d'un facteur d'échelle. 10

1.5 Coordonnées géocentriques cartésiennes et coordonnées locales (http://en.academic.ru/dic.nsf/enwiki/498677) 12

1.6 Exemple illustrant le comportement d'un estimateur robuste et de l'estimateur des moindres carrés. En présence d'outlier, pour le cas de la régression linéaire [Press et al., 1992] 17

1.7 Temps de calcul pour une estimation robuste par la méthode simplex et la méthode Dikin [Khodabandeh and Amiri-Simkooei, 2011] 21

1.8 Carte des sites du réseau global utilisé pour le test 1.6.3.1 23

1.9 Diagramme de l'évolution des erreurs de la transformation de Helmert en fonction de l'amplitude de l'erreur introduite pour les estimateurs de Dikin et des moindres carrés 24

1.10 Carte des sites du réseau local utilisé pour le test 1.6.3.2 25

1.11 Diagramme de l'évolution des erreurs de la transformation en fonction de l'amplitude de l'erreur introduite pour les estimateurs de Dikin et des moindres carrés . . . 26

1.12 Carte des sites du réseau asymétrique utilisé pour le test 1.6.3.3 27

1.13 Diagramme de l'évolution des erreurs de la transformation en fonction de l'amplitude de l'erreur introduite pour les estimateurs de Dikin et des moindres carrés . . . 28

1.14 Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 1cm est introduite sur la composante X du site X000 au 3ème jour (ligne verticale) . 30

1.15 Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 10 cm est introduite sur la composante X du site X000 au 3ème jour (ligne verte verticale). Cette erreur est détectée, calculée et montrée pour le site X000 (carré rouge). 31
1.16 Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 1 m est introduite sur la composante X du site X000 au 3ème jour (ligne verte verticale). Cette erreur est détectée, calculée approximativement et montrée pour le site X000 (carré rouge). 32
1.17 Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 1 cm est introduite sur la composante X des 5 sites X000, X001, X003, X006 et X009 au 3ème jour (ligne verte verticale). 34
1.18 Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 10 cm est introduite sur la composante X des 5 sites X000, X001, X003, X006 et X009 au 3ème jour (ligne verte verticale). Cette erreur est détectée, calculée approximativement et montrée pour le site X000 (carré rouge). 35
1.19 Séries temporelles des sites X000 (en haut) et X004 (en bas). Une erreur de 1 m est introduite sur la composante X des 5 sites X000, X001, X003, X006 et X009 au 3ème jour (ligne verte verticale). Cette erreur est détectée, calculée approximativement et montrée pour le site X000 (carré rouge). 36
1.20 Evolution du temps de calcul des estimateurs suivant l'augmentation du nombre de transformation de Helmert (nombre de jours) 38
1.21 Evolution du temps de calcul de l'estimateur robuste de Dikin en fonction du nombre d'observations 38
1.22 Diagramme de l'évolution des erreurs Helmert des 7 paramètres en fonction de l'amplitude de l'erreur introduite 40
1.23 Schéma de calcul du programme pyac_make_time_series.py 44
1.24 Série temporelle brute (pente retirée) du site IESG (Géoazur I) obtenue par pyacs_make_time_series.py 46

2.1 Série temporelle du site FJCP (réseau RENAG) 51
2.2 Exemple d'outlier isolé (site BANN - RENAG) 53
2.3 Exemple d'outliers groupés (site CAST - RENAG) 53
2.4 Exemple d'augmentation du bruit dans une série temporelle (site AIGL - RENAG) 54
2.5 Exemple d'offsets (site PQRL - réseau RGP) 55
2.6 Exemple de déplacement co-sismique simple (site GS32 - réseau GPS du Taiwan) 55
2.7 Exemple de déplacement co-sismique et post-sismique (site 0007 - GEONET Japon) 56

2.8 Exemples de termes saisonniers forts dans une série temporelle (site CHIZ - réseau RENAG) . . . 57
2.9 Exemple de signal transitoire (site CABP - projet ANR ADN, source : J.-M Nocquet) . . . 58
2.10 Série temporelle du site P106 (GEONET - Japon) et sa modélisation . . . 61
2.11 Série temporelle de position (haut) et série temporelle de rms calculé par fenêtre glissante de 7 jours (bas) . . . 63
2.12 Détection d'un outlier dans la fenêtre anormale i . . . 64
2.13 Détection de deux outliers continus dans la fenêtre anormale i . . . 64
2.14 Détection d'outliers dans la série temporelle du site URUM . . . 65
2.15 Exemple de détection d'outliers par le test 1.44 . . . 66
2.16 Détection d'offset dans la fenêtre i . . . 66
2.17 Série temporelle de la composante nord du site BRST (Geoazur I) avec les offsets détectés par la méthode 2.3.2 . . . 67
2.18 Série temporelle de la composante nord du site BRST (Geoazur I - RENAG) avec les offsets détectés et testés . . . 68
2.19 Série temporelle synthétique et sa modélisation automatique réalisée par pygps_time_series.py 71
2.20 Série temporelle résiduelle, une fois le modèle estimé par pygps_time_series.py. 72

3.1 Exemple de visualisation d'une série temporelle, à travers la méthode plot . . . 79
3.2 Diagramme des opérations élémentaires pour l'analyse automatique des séries temporelles GPS . . . 80
3.3 Série temporelle du site PQRL (Géoazur I - RENAG) . . . 83
3.4 Série temporelle du site GS32 (réseau GPS Taiwan) . . . 84

4.1 Carte du réseau RENAG . . . 88
4.2 Histogrammes des répétabilités obtenues par pyacs_make_time_series.py pour le réseau RENAG . . . 90
4.3 Série temporelle (pente retirée) des sites COLC et HEAU . . . 91
4.4 Exemples de séries temporelles (pente retirée) avec outliers isolés (gauche, site MODA) et groupés (droite, site MEDI). . . . 92
4.5 Série temporelle (pente retirée) avec offsets (Site AXPV) . . . 93
4.6 Histogrammes des répétabilités, après traitement par pygps_time_series.py . . . 94
4.7 Analyse des séries temporelles du site WTZR . . . 95
4.9 Histogrammes des répétabilités issues du traitement final pour le réseau RENAG 96
4.8 Re-analyse de la série temporelle du site PUYV . . . 97
4.10 Champ de vitesses horizontal résiduel du réseau RENAG (après retrait du pôle d'Euler) . . . 98

4.11 Champ de la vitesses vertical du réseau RENAG dans l'ITRF2008 99
4.12 Cartes des 114 sites du réseau GPS du Taiwan [Lin et al., 2010] 101
4.13 Histogrammes des répétabilités obtenues par pyacs_make_time_series.py . . . 102
4.14 Exemple de série temporelle complexe (sites KDNM et CHEN). La courbe rouge indique le modèle déterminé par mon programme pygps_time_series.py. 103
4.15 Série temporelle avec le séisme lent du site GS02 et celle non-linéaire du site S106 104
4.16 Histogrammes des répétabilités, après traitement par pyacs_make_time_series.py 105
4.17 Re-analyse de la série temporelle du site LONT 106
4.18 Re-analyse de la série temporelle du site CHEN 107
4.19 Histogrammes des répétabilités obtenus pour le traitement final par pyacs_make_time_series.py 109
4.20 Champ de vitesses horizontal dans l'ITRF2008 du réseau GPS du Taiwan . . . 110
4.21 Champ de vitesses vertical dans l'ITRF2008 du réseau GPS du Taiwan 111

LISTE DES TABLEAUX

1.1 Critères utilisés pour la sélection des stations IGS08 à partir des station IGS de l'ITRF2008 [Rebischung et al., 2012] 9
1.2 Evolution de la valeur des 7 paramètres de la transformation de Helmert en fonction de l'amplitude de l'erreur introduite. e_H est définie par 1.42. 24
1.3 Evolution de la valeur des 7 paramètres de la transformation de Helmert en fonction de l'amplitude de l'erreur introduite. e_H est définie par 1.42 et 1.43. . 26
1.4 Evolution de la valeur des 7 paramètres de la transformation de Helmert en fonction de l'amplitude de l'erreur introduite. e_H est définie par 1.42 et 1.43. . 28
1.5 Valeurs estimées des 7 paramètres de la transformation de Helmert en fonction de l'amplitude de l'erreur introduite. e_H est définie par 1.42. 40
1.6 Temps de calcul de la réalisation des séries temporelles par pyacs_make_time_series.py pour les réseaux RENAG et Taiwan 45
1.7 Résidus des points communs en estimation des 7 paramètres de la transformation (époque 2006.001, Géoazur I) 46

2.1 Comparaison des paramètres synthétiques et estimés 72

3.1 Paramères décrivant la série temporelle du site PQRL (Géoazur I - RENAG) . 82
3.2 Paramètres décrivant la série temporelle du site GS32 (réseau GPS Taiwan) . . 85

4.1 Paramètres des déplacements co- et post-sismique issus de la re-analyse de la série temporelle des sites LONT et CHEN 108
A.1 Vitesse du réseau RENAG 122
A.2 Vitesse du réseau RENAG (continues) 123
A.3 Vitesse du réseau RENAG (continues) 124
A.4 Résidus de vitesse des sites pour définir la pôle de rotation 125
A.5 Vitesses verticales dans l'ITRF2008 du réseau RENAG 126
B.1 Vitesses verticales dans l'ITRF2008 du réseau GPS du Taiwan 127
B.2 Vitesses verticales dans l'ITRF2008 du réseau GPS du Taiwan (continues) . . . 128
B.3 Vitesses verticales dans l'ITRF2008 du réseau GPS du Taiwan (continues) . . . 129
B.4 Vitesses verticales dans l'ITRF2008 du réseau GPS du Taiwan (continues) . . . 130

B.5 Vitesses verticales dans l'ITRF2008 du réseau GPS du Taiwan (continues) . . . 131