Statistical analysis of vertical land motions and sea level measurements at the coast

Kévin Gobron

LA ROCHELLE UNIVERSITÉ
LIÈGE UNIVERSITÉ

THÈSE EN COTUTELLE

Laboratoire Littoral Environnement et Sociétés (LIENSs)
Ecole Doctorale n°618 Euclide | La Rochelle Université

Département de Géographie | Liège Université

présentée par :

Kevin GOBRON

soutenue le : **04 février 2021**
pour obtenir le grade de : **Docteur de La Rochelle Université**
et de **Docteur en Sciences de Liège Université**

Discipline : **Terre solide et enveloppes superficielles**

Statistical analysis of vertical land motions and sea level measurements at the coast

---

JURY :

Mélanie BECKER	Directrice de recherche CNRS, La Rochelle Université, Présidente du jury
Éric CALAIS	Professeur, Ecole Normale Supérieure de Paris, Rapporteur
Simon D. P. WILLIAMS	Senior Research Scientist, National Oceanography Centre of Liverpool, Rapporteur
Paul REBISCHUNG	Chargé de recherche, Institut National de l'Information Géographique et Forestière (IGN), Examinateur
Pierre MAGAIN	Professeur, Liège Université, Examinateur
Olivier DE VIRON	Professeur, La Rochelle Université, Directeur de thèse
Michel VAN CAMP	Chef du service séismologie/gravimétrie, Observatoire Royal de Belgique, Co-directeur
Alain DEMOULIN	Professeur, Liège Université, Co-directeur

Abstract

Statistical analysis of vertical land motions and sea level measurements at the coast

by Kevin GOBRON

Ensuring the long-term stability of sea-level measurements at the coast is essential for the sustainable management of coastal territories. To do this, scientists mainly rely on two complementary measurement techniques: tide-gauges, which measure changes in sea-level with respect to the coast, and space geodesy, which allows measuring the vertical movements of the coast itself, with respect to the centre of the Earth. These two techniques are complementary because the correction of the vertical land motions estimated using geodetic techniques allows estimating geocentric sea-level change at the coast, and thereby to understand the origin of sea-level change better.

To estimate and improve the stability and quality of these two sources of information, this thesis proposes methodological developments dedicated to the estimation of the precision of each type of observation and investigates the potential sources of errors.

Remerciements

Tout d'abord j'aimerais exprimer toute ma gratitude envers Olivier de Viron et Michel Van Camp de m'avoir offert cette opportunité de réaliser une thèse de doctorat. Vu mon parcours de technicien, ce n'était pas gagné ! Cette aventure a été l'expérience professionnelle la plus enrichissante que j'ai vécue. Aussi, vous avez été d'une patience et d'une disponibilité exemplaires à mon égard. Cette thèse vous doit énormément ! Je vous remercie également d'avoir eu la bonté de m'avoir autorisé à dormir entre 00:10 et 00:15 tout en sachant que c'était deux fois plus que nécessaire.

J'aimerais également remercier Alain Demoulin d'avoir accepté de prendre le train de la géodésie en marche, mais aussi de m'accompagner dans les aventures administratives de cette cotutelle entre La Rochelle Université et l'Université de Liège. Cette thèse aura finalement pris une tournure très « statistique », mais j'espère qu'elle aura su retenir votre intérêt !

I want to thank Eric Calais and Simon Williams for accepting to review this manuscript. I am also grateful toward Mélanie Becker, Paul Rebischung, and Pierre Magain for agreeing to be part of the jury members.

En outre, je tiens à remercier Paul Rebischung, qui m'a rendu visite en avril 2019, et avec qui j'ai eu le plaisir d'explorer les plus obscures (ou les plus colorés) recoins de l'analyse de séries temporelles GNSS ! Nos discussions et tes conseils ont eu un impact déterminant sur la plupart des chapitres de cette thèse, merci beaucoup !

I would like to thank Jim Ray, who visited me in November 2019. Talking about GNSS time series analysis with you (sometimes around a glass of wine) have been a great pleasure. The last chapters of this thesis would have been very different without our discussions !

Un grand merci à l'équipe SONEL, en particulier à Guy Woppelmann, Laurent Testut, Valérie Ballu, Médéric Gravelle, Étienne Poirier, Mikael Guichard et Philippe Téchiné pour m'avoir fait découvrir le monde de la marégraphie.

Merci également à l'équipe du service sismologie/gravimétrie d'Observatoire Royal de Belgique pour m'avoir accueilli à l'observatoire, et en particulier à Thomas, Koen, Arnaud, Corentin, Baptiste, et Aurélie pour les entraînements de foot de précision et les afterworks pendant mon passage au plat pays !

Merci aussi à Laurent Morel et à Jérôme Verdun de l'École Supérieure de Géomètres et Topographes, pour m'avoir permis d'enseigner dans l'établissement où j'ai moi-même été étudiant.

J'aimerais demander pardon à Caroline, Heitea, et Sophie pour vous avoir certainement déconcentré durant la cohabitation.

Big up à tous ceux avec qui j'ai pu passer du temps, discuter, et/ou boire des coups tant au LIENSs qu'à l'ORB ! ... allez, je tente une liste : Clémence, Treden, Marc, Jamal, Brieuc, Baptiste, Cyril, Mannaïg, Laura, Camille, Raphaël, Yannick, Aurélie, Amélie, Anaïs, Mélanie, Guillaume, Mery, Thomas, Vincent, Monia,

Victoria, Natascha, Clément, Céline, Yann, Jean-Remy, Raphaël, Sabrina, Julia, Alfonso, Alexandro, Amadeo, Greg, Elena, Aymeric, Alexia, Pierre, ... ? Si j'ai oublié quelqu'un, je suis désolé, là il est 03:47. Vous pouvez néanmoins faire une réclamation via ce lien : réclamation.

Aussi, j'aimerais faire une petite dédi-smash aux irréductibles du badminton, Nathalie et Nicolas, j'espère que l'on pourra se refaire une partie à l'occasion !

J'aimerais également remercier toute ma famille, en particulier mes parents et leur soutien inconditionnel. C'est promis, maintenant j'arrête avec les diplômes.

Enfin, j'aimerais remercier, ma chère et tendre, Lucie, pour avoir su entrecouper mes efforts avec des bouffées d'air purs, tout en me laissant assez de liberté pour finaliser cette tâche un tout petit peu prenante.

Ces travaux de thèse ont été financés par la Direction Générale de l'Armement (DGA), la région Nouvelle Aquitaine, et le Centre National des Etudes Spatiales (CNES) dans le cadre des missions de géodésie.

This PhD thesis has been financially supported by the Direction Générale de l'Armement (DGA), the Nouvelle-Aquitaine region, and by the Centre National des Études Spatiales (CNES) as an application of the geodesy missions.

Contents

Abstract i

Remerciements iii

1 Introduction 1
- 1.1 Measuring sea-level at the coast . . . 1
- 1.2 Related metrological requirements . . . 2
  - 1.2.1 Tide gauge accuracy . . . 2
  - 1.2.2 GNSS time series accuracy . . . 3
- 1.3 Research objectives . . . 4
- 1.4 Outline of the thesis . . . 4

2 Estimation of deterministic parameters from observations 7
- 2.1 Functional and stochastic models . . . 7
  - 2.1.1 Scalar case . . . 7
  - 2.1.2 Multivariate case . . . 8
- 2.2 Model of observation equations . . . 9
  - 2.2.1 Linear case . . . 10
  - 2.2.2 Nonlinear case . . . 10
  - 2.2.3 Estimation principle . . . 11
- 2.3 Weighted Least-Squares Estimation . . . 11
  - 2.3.1 Weighted Least-Squares Estimator . . . 11
  - 2.3.2 Uncertainty of the weighted least-squares estimator . . . 12
- 2.4 Maximum Likelihood Estimation (MLE) . . . 12
  - 2.4.1 Likelihood function . . . 12
  - 2.4.2 Maximum Likelihood Estimator . . . 13
  - 2.4.3 Uncertainty of the maximum likelihood estimator . . . 13
- 2.5 Best Linear Unbiased Estimation . . . 14
- 2.6 Measures of inconsistency . . . 14
  - 2.6.1 Residuals . . . 15
  - 2.6.2 Misclosures . . . 16
- 2.7 Misspecifications . . . 17

3 Estimation of stochastic parameters from observations 19
- 3.1 Parametric stochastic model . . . 19
- 3.2 Maximum Likelihood Estimation (MLE) . . . 20
- 3.3 Restricted Maximum Likelihood Estimation (RMLE) . . . 20
- 3.4 Least-Squares Variance Component Estimation (LS-VCE) . . . 21
  - 3.4.1 Linear stochastic model . . . 21
  - 3.4.2 Vectorizing the problem . . . 22
  - 3.4.3 Best Linear Unbiased Estimator . . . 23
  - 3.4.4 Iterative algorithm . . . 24
  - 3.4.5 Uncertainty of the variance components . . . 25

3.4.6 Nonlinear stochastic model . . . 27
3.4.7 Possible estimation issues . . . 29
Divergence . . . 29
Negative estimates . . . 29
3.4.8 Numerical examples . . . 29
Accuracy of the method . . . 30
Speed of the optimized algorithm . . . 31

4 Elements of geodetic time series analysis **33**
4.1 Trajectory model . . . 33
4.1.1 Trend model . . . 33
4.1.2 Periodic oscillations model . . . 34
4.1.3 Discontinuity models . . . 34
4.1.4 Transient models . . . 35
4.2 Stochastic processes . . . 36
4.2.1 White noise . . . 36
4.2.2 Time-correlated processes . . . 37
First-Order Gauss-Markov (FOGM) processes . . . 38
Power-Law (PL) processes . . . 40
Generalized Gauss Markov (GGM) processes . . . 41
4.2.3 Combination of stochastic processes . . . 44
4.2.4 Missing values . . . 45
4.3 Functional and stochastic model selection . . . 45

5 Event detection using hypothesis testing in linear models **47**
5.1 Hypothesis testing . . . 47
5.1.1 The null hypothesis . . . 47
5.1.2 The alternative hypothesis . . . 48
5.1.3 The Generalized Likelihood Ratio Test (GLRT) . . . 48
5.2 Practical detection examples . . . 50
5.2.1 Spectral analysis . . . 50
5.2.2 Detection of discontinuities . . . 51
Offsets detection . . . 51
Outlier detection . . . 53
Velocity change detection . . . 55

6 Assessment of tide gauge sensor offsets in the presence of scale errors **57**
6.1 Introduction . . . 57
6.2 Influence of a sensor offset and a scale error on sea-level differences . . . 58
6.3 Impact of the scale error on mean difference interpretation. . . . 62
6.4 How long should a calibration experiment last ? . . . 64
6.5 Conclusion . . . 65

7 Assessment of tide gauges biases and precisions by the combination of multiple co-located time series **67**
7.1 Introduction . . . 67
7.2 The Aix Island experiment . . . 69
7.3 Calibration methods . . . 71
7.3.1 Sea level error model . . . 72
7.3.2 The difference-based calibration method (DIFF) . . . 73
7.3.3 The combination-based calibration method (COMB) . . . 74

Functional model . . . 74
Stochastic model . . . 76
Least-Squares estimation . . . 77
Least-Squares Variance Component Estimation . . . 77
7.4 Results . . . 79
7.4.1 Calibration with the combination (COMB) method . . . 79
7.4.2 Comparison with the difference (DIFF) method . . . 83
7.5 Discussion . . . 85
7.5.1 Performance of the tide gauges . . . 85
7.5.2 Nature of the biases . . . 86
7.5.3 Improvement over difference based methods . . . 87
7.6 Conclusion . . . 87
7.7 Acknowledgments . . . 88

8 Investigating the potential of the nonlinear LS-VCE method for the stochastic modelling of GNSS position time series 89
8.1 Introduction . . . 89
8.2 Functional and stochastic model . . . 91
8.3 Least-Squares Variance Component Estimation . . . 92
8.3.1 Nonlinear LS-VCE . . . 93
8.3.2 Non-negative LS-VCE . . . 93
8.4 Partial derivatives of the stochastic model . . . 93
8.4.1 Partial derivatives with respect to σ_w^2 and σ_{pl}^2 . . . 93
8.4.2 Partial derivative with respect to κ . . . 94
8.4.3 Partial derivative with respect to ϕ . . . 95
8.5 Performance on synthetic time series . . . 97
8.5.1 WN+PL model . . . 97
8.5.2 WN+GGM model . . . 99
8.6 Contributions and perspectives . . . 101

9 Impact of offsets on the stochastic modelling of GNSS position time series. 103
9.1 Introduction . . . 103
9.2 Functional and stochastic model of GNSS time series . . . 105
9.2.1 Functional model . . . 105
9.2.2 Stochastic model . . . 106
9.2.3 Properties of position residuals . . . 106
9.3 Influence of offsets on spectral analyses . . . 108
9.3.1 Least-Squares Harmonic Estimation . . . 108
Definition . . . 108
Expectation . . . 109
9.3.2 Lomb Scargle periodogram . . . 109
Computation . . . 109
Expectation . . . 109
Lomb-Scargle periodogram bias resulting from residual estimation . . . 109
9.3.3 Influence of offsets on periodogram interpretation . . . 110
9.4 Influence of offsets on stochastic model selection . . . 113
9.4.1 Log-likelihood . . . 113
Computation . . . 113
Expectation . . . 113
9.4.2 Restricted log-likelihood . . . 114

Computation . . . 114
Expectation . . . 115
9.4.3 Influence on the noise model selection . . . 115
9.5 Uncertainty of variance components . . . 120
9.5.1 Influence on the variance component uncertainties . . . 121
Influence on a WN+GGM model . . . 121
Influence on a WN+FL+RW model . . . 121
9.6 Discussion . . . 125
9.7 Conclusion . . . 125

10 Spatial and temporal properties of the stochastic variability in the position time series of over 10,000 stations. 127
10.1 Introduction . . . 127
10.2 NGL Solutions . . . 129
10.3 Variance Component Estimation . . . 130
10.3.1 Least-Squares Variance Component Estimation . . . 130
10.3.2 Selection of stochastic models . . . 131
10.4 Spatial correlation . . . 132
10.4.1 Spatial correlation estimation strategy . . . 132
10.4.2 Estimation . . . 132
Change of frame . . . 132
Pairwise correlation estimation . . . 133
Average spatial correlation estimation . . . 134
10.4.3 Time dependence analysis . . . 134
10.5 Time series analysis . . . 135
10.5.1 Selection of observations . . . 135
10.5.2 Stochastic model . . . 135
10.5.3 Initial functional model . . . 136
10.5.4 Improving the functional model . . . 137
10.5.5 Variance component estimation strategy . . . 138
10.6 Results . . . 138
10.6.1 Spatial variability of the stochastic parameters . . . 138
10.6.2 Spatial correlation . . . 139
10.7 Discussion . . . 152
10.7.1 Spatial dependence of noise parameters . . . 152
10.7.2 Time-variable spatial correlation . . . 156
10.7.3 Influence of the reprocessing . . . 156
10.8 Conclusion . . . 157

11 Influence of non-tidal atmospheric and oceanic loading on the stochastic properties of vertical land motion time series 159
11.1 Introduction . . . 159
11.2 Computing loading deformations . . . 160
11.2.1 Loading deformation predicted by the EMSGFZ-Postdam . . . 160
11.3 Preliminary change in repeatability analysis . . . 161
11.4 Influence of NTAOL on noise properties . . . 161
11.4.1 Impact on temporal correlation . . . 164
11.4.2 Impact on spatial correlation . . . 167
11.5 Discussion . . . 169
11.6 Conclusion . . . 170

12 Conclusions & perspectives **175**
12.1 Tide gauge observations 175
12.1.1 Contributions 175
12.1.2 Perspectives 175
12.2 GNSS-based vertical land motion estimates 176
12.2.1 Contributions 176
12.2.2 Perspectives 177

Bibliography **179**

List of Figures

3.1 Example of simulated observations. 30
3.2 Normalized distributions of variance component estimates for 5000 simulations (*blue*). True variance components (*black*). Normal distribution approximation using the standard deviation predicted by Equation (3.30). Result for 50 observations per time series (*top*). Result for 100 observations per time series (*bottom*) 31
3.3 Evolution of the run-time of the standard and optimized LS-VCE algorithm 32

4.1 Example of simulated first-order Gauss-Markov processes generated from the same white noise time series. 39
4.2 Theoretical spectra of discrete first-order Gauss-Markov (FOGM) processes. 40
4.3 Example of simulated discrete power-law processes generated from the same white noise time series. 42
4.4 Theoretical power-spectra of discrete power-law processes. 42
4.5 Example of simulated discrete generalized Gauss Markov processes generated from the same white noise time series. 43
4.6 Theoretical spectra of discrete generalized Gauss Markov processes . . 44

5.1 Synthetic time series generated for the examples. 50
5.2 Example of a spectral analysis in a synthetic time series using the GLRT. Simulated time series (*top*). T_2 statistic (*bottom*). The grey line is the detection threshold for a confidence level $\alpha = 0.99$. The vertical red lines are the true frequencies. 52
5.3 Example of offset detection in a synthetic time series using the GLRT. Simulated time series (*top*). T_1 statistic (*bottom*). The grey line is the detection threshold for a confidence level $\alpha = 0.99$. The vertical red line is the true date of the offset. 54
5.4 Example of outlier detection in a synthetic time series using the GLRT. Simulated time series (*top*). T_1 statistic (*bottom*). The grey line is the detection threshold for a confidence level $\alpha = 0.99$. The vertical red line is the true date of the outlier. 55
5.5 Example of velocity change detection in a synthetic time series using the GLRT. Simulated time series (*top*). T_1 statistic (*bottom*). The grey line is the detection threshold for a confidence level $\alpha = 0.99$. The vertical red line is the true date of the velocity change. 56

6.1 Configuration with no offset and no scale error. RSM: Reference Survey Mark. PZR: Point of Zero Range. 59
6.2 Configuration with an offset but no scale error. RSM: Reference Survey Mark. PZR: Point of Zero Range. 59

6.3 Configuration with an offset and a scale error. RSM: Reference Survey Mark. PZR: Point of Zero Range. 60
6.4 Examples of the deterministic influence of common measurement biases on the Van de Casteele diagram. $H_T(t)$ is the sea-level recorded by the tested gauge at time t. $\Delta_T(t) = H_T(t) - H_R(t)$ is the difference of sea-level between the tested gauge and the reference gauge. 62
6.5 Visual identification of the offset on the Van de Casteele diagram of the tested gauge. 63
6.6 Evolution of the interpretation error resulting from the assumption that $o_T = -\bar{\Delta}_T$ in the presence of scale error. 64
6.7 Evolution of the standard deviation of the sensor offset σ_{o_T} with the cumulative observation time and the tide amplitude A. 66

7.1 The four ground-based tide gauges: RADAR, POLE, PROBE and LASER. 70
7.2 The two GNSS buoys: BUOY1 and BUOY2. 70
7.3 Sea level time series y_i recorded by all tide gauges. 72
7.4 Synthetic examples of Van de Casteele diagrams for the most common types of range measurement errors: (a) random measurement errors only; (b) random measurement errors and a height reference error; (c) random measurement errors and a scale error; (d) random measurement errors and a clock error. 74
7.5 Residual time series of each tide gauge for the estimated linear combination model. 81
7.6 Combined solution (*top*), the standard deviation of the combined solution (*middle*) and available observations for each gauge (*bottom*). . . . 82
7.7 Van de Casteele diagram of BUOY1. The sea level bias model estimated with the COMB method is displayed in blue, and the one estimated with the DIFF method is displayed in red. 85
7.8 Representation of the bias parameter estimates in the parameter space for both COMB and DIFF methods. 86

8.1 Distribution of variance component estimates obtained using the unconstrained nonlinear LS-VCE method for the WN+PL stochastic model (*blue histogram*). True variance component (*dashed red line*). Normal distribution centred on the true value, with variance computed using Equation (8.13) based on the true variance components in Table (8.1) (*solid red line*). 98
8.2 Distribution of variance component estimates obtained using the non-negative nonlinear LS-VCE method for the WN+PL stochastic model (*blue histogram*). True variance component (*dashed red line*). Normal distribution centred on the true value, with variance computed using Equation (8.13) based on the true variance components in Table (8.1) (*solid red line*). 98
8.3 Runtime of the unconstrained nonlinear LS-VCE method with the time span of the time series. 99
8.4 Distribution of the WN+GGM model variance component estimates obtained using the unconstrained nonlinear LS-VCE. 100
8.5 Distribution of the WN+GGM model variance component estimates obtained using the nonlinear LS-VCE with non-negativity constraints. 100

9.1 Influence of the functional model on the expected LS-HE periodogram $P(f_k)$. The expected LS-HE periodograms in the cases (a), (b), and (c) overlap. . . . 111
9.2 Influence of the functional model on the expected Lomb-Scargle periodogram $P'(f_k)$. . . . 112
9.3 Expected log-likelihood (*top*) and restricted log-likelihood (*bottom*) for WN+GGM models with different values of ϕ. For a better legibility, the (restricted) log-likelihood of the true WN+PL model is systematically subtracted. . . . 116
9.4 Influence of the parameter ϕ on the power-spectrum. . . . 117
9.5 Expected log-likelihood (*top*) and restricted log-likelihood (*bottom*) for WN+GGM models with different values of τ. For a better legibility, the (restricted) log-likelihood of the true WN+PL model is systematically subtracted. . . . 118
9.6 Expected log-likelihood (*top*) and restricted log-likelihood (*bottom*) for WN+PL models with different values of κ. For a better legibility, the (restricted) log-likelihood of the WN+PL model with $\kappa = -0.8$ is systematically subtracted. . . . 119
9.7 Evolution of the standard deviation of variance components with offset frequency for the WN+GGM model. . . . 122
9.8 Evolution of the standard deviation of variance components with offset frequency for the WN+FL+RW model. . . . 124

10.1 Illustration of the enh-frames and uvh-frames for two stations S_i and S_j. 133
10.2 Histogram of the number of stations used as a function of the number of observations for the NGL-IGS14 solution and the Up coordinate. . . . 135
10.3 Global variability of white noise amplitudes (*top*), spectral indices (*middle*); and power-law amplitudes (*bottom*) for the North coordinates of the NGL-IGS14 solution. . . . 140
10.4 Global variability of white noise amplitudes (*top*), spectral indices (*middle*); and power-law amplitudes (*bottom*) for the East coordinates of the NGL-IGS14 solution. . . . 141
10.5 Global variability of white noise amplitudes (*top*), spectral indices (*middle*); and power-law amplitudes (*bottom*) for the Up coordinates of the NGL-IGS14 solution. . . . 142
10.6 Latitude dependence of white noise amplitudes $\hat{\sigma}_w$ for the North coordinate (*top*), for the East coordinate (*middle*), for the Up coordinate (*bottom*), and for both the NGL-IGS08 solution (*left*) and NGL-IGS14 solution (*right*). The dots in cyan represent the raw estimates. The continuous blue lines represent running medians over centered 10° latitude windows. The shaded areas represent the corresponding interquartile ranges. The dashed black lines correspond to the running median for NGL-IGS08, reported on the NGL-IGS14 results to ease the comparison. . . . 143

10.7 Latitude dependence of spectral indices $\hat{\kappa}$ for the North coordinate (*top*), for the East coordinate (*middle*), for the Up coordinate (*bottom*), and for both the NGL-IGS08 solution (*left*) and NGL-IGS14 solution (*right*). The dots in cyan represent the raw estimates. The continuous blue lines represent running medians over centered 10° latitude windows. The shaded areas represent the corresponding interquartile ranges. The dashed black lines correspond to the running median for NGL-IGS08, reported on the NGL-IGS14 results to ease the comparison. 144
10.8 Latitude dependence of power-law process amplitudes $\hat{\sigma}'_{pl}$ for the North coordinate (*top*), for the East coordinate (*middle*), for the Up coordinate (*bottom*), and for both the NGL-IGS08 solution (*left*) and NGL-IGS14 solution (*right*). The dots in cyan represent the raw estimates. The continuous blue lines represent running medians over centered 10° latitude windows. The shaded areas represent the corresponding interquartile range. The dashed black lines correspond to the running median for NGL-IGS08, reported on the NGL-IGS14 results to ease the comparison. 145
10.9 Latitude dependence of the Root Mean Square Error (RMSE), for the North coordinate (*top*), for the East coordinate (*middle*), for the Up coordinate (*bottom*), and for both the NGL-IGS08 solution (*left*) and NGL-IGS14 solution (*right*). The dots in cyan represent the raw estimates. The continuous blue lines represent running medians over centered 10° latitude windows. The shaded areas represent the corresponding interquartile ranges. For each time series, RMSE are based on residual estimates computed using the best stochastic model. The dashed black lines correspond to the running median for NGL-IGS08, reported on the NGL-IGS14 results to ease the comparison. 146
10.10Spatial correlation estimates $\hat{\rho}(t_k, d_l)$ for the NGL-IGS08 solution. Extension/compression direction (*top*), Shear direction (*middle*) and Up direction *(bottom*). The blue curve represents the median correlogram. The dashed black line represents the Gaussian correlation function. . . 147
10.11Spatial correlation estimates $\hat{\rho}(t_k, d_l)$ for the NGL-IGS14 solution. Extension/compression direction (*top*), Shear direction (*middle*) and Up direction *(bottom*). The blue curve represents the median correlogram. The dashed black line represents the Gaussian correlation function. . . 148
10.12Time-variable spatial correlation estimates $\hat{\rho}(t_k, d_l)$ for 4 classes of distances of the Up coordinate of the NGL-IGS14 solution. The coloured lines represent the estimates at given class of distance d_l. The dashed black lines represent the corresponding time-variable Gaussian correlation function. 151
10.13Spatial variability of the standard deviation oscillation for the North coordinates of NGL-IGS14. Annual amplitudes of monthly standard deviations (*top*). Annual phase of monthly standard deviations. The phase is expressed as the month corresponding to the maximum of dispersion (*bottom*). for time series presenting more than 4 years of cumulative observations 153
10.14Spatial variability of the standard deviation oscillation for the East coordinates of NGL-IGS14. Annual amplitudes of monthly standard deviations (*top*). Annual phase of monthly standard deviations. The phase is expressed as the month corresponding to the maximum of dispersion (*bottom*). 154

10.15Spatial variability of the standard deviation oscillation for the Up coordinates of NGL-IGS14. Annual amplitudes of monthly standard deviations (*top*). Annual phase of monthly standard deviations. The phase is expressed as the month corresponding to the maximum of dispersion (*bottom*). 155

11.1 Standard deviation of the predicted vertical land motions resulting from surface mass loading. Non-Tidal Atmospheric Loading (NTAL) (*top*). Non-Tidal Oceanic Loading (NTOL) (*middle*). Hydrological loading (HYDL) (*bottom*). 162
11.2 Global distribution of RMSE reduction percentages after removing deformation predictions. Non-tidal atmospheric loading (NTAL) (*top*). Non-tidal oceanic loading (NTOL) (*middle*). Hydrological loading (HYDL) (*bottom*). 163
11.3 Global variability of white noise amplitudes (top), spectral indices (middle); and power-law amplitudes (bottom) for the Up coordinates with NTAOL deformations removed. 165
11.4 Latitude dependence of RMSE (*top*), white noise amplitudes (*upper middle*), spectral indices (*lower middle*), power-law amplitudes (*bottom*) in the presence of NTAOL deformations (*left*) and with NTAOL deformations removed (*right*). The dots in cyan represent the raw estimates. The continuous blue lines represent running medians over centered 10° latitude windows. The shaded areas represent the corresponding interquartile ranges. The dashes black lines are running medians before NTAOL corrections reported on that after corrections. 166
11.5 Global variability of the restricted log-likelihood differences. 167
11.6 Global variability of the trend uncertainties for a 8 years long time series. Without NTAOL corrections (*top*). With NTAOL corrections (*bottom*). 168
11.7 Spatial correlation estimates. Without NTAOL corrections (*top*). With NTAOL corrections (*bottom*). The blue curve represents the mean correlogram. The dashed black lines represent the Gaussian correlation model. 171
11.8 Time variable spatial correlation estimates for 4 classes of distances with NTAOL corrections. The colored lines represent the correlation estimates with NTAOL corrections. The light grey lines represent the correlation estimates without NTAOL corrections. The dashed black lines represent the time variable Gaussian correlation model. 172
11.9 Spatial variability of the standard deviation oscillation with NTAOL corrections. Annual amplitudes of monthly standard deviations (*top*). Annual phase of monthly standard deviations. The phase is expressed as the month corresponding to the maximum of dispersion (*bottom*). . 173

List of Tables

7.1 Tide gauge cross-calibration results obtained using the COMB method. PROBE scale error and intercept are conventionally set to zero. 80
7.2 Alternative tide gauge cross-calibration results obtained using the COMB method and by defining BUOY1 as the conventionally unbiased gauge. BUOY1 scale error and intercept are conventionally set to zero. 83
7.3 Tide gauge calibration results obtained using the DIFF method. PROBE is the reference gauge. 83
7.4 Difference between DIFF and COMB calibration results. 84
7.5 Reduction of the standard deviations of the bias parameters obtained using the COMB method with respect to the DIFF method. 84

8.1 True and approximate (initial) variance components used for the simulation 97
8.2 Average variance component estimates for the unconstrained LS-VCE (LS-VCE) and the non-negative LS-VCE (NNLS-VCE). 97
8.3 True and approximate (initial) variance components used for the simulation 100

9.1 Considered variance components for the WN+PL model. 110
9.2 Considered variance components for the WN+GGM model. 121
9.3 Percentage of standard deviation increase for the WN+GGM variance components when considering functional models with 1.0 opy compared to 0.0 opy (that is, no offsets). 122
9.4 Considered variance components for the WN+FL+RW model. 123
9.5 Percentage of standard deviation increase for the WN+FL+RW variance components when considering functional models with 1.0 opy compared to 0.0 opy (that is, no offsets). 123

10.1 Parameter estimates for the average Gaussian correlation function (Equation 10.27). 149
10.2 Parameter estimates for the time-variable Gaussian correlation function (Equation 10.28). The $\hat{r}$ are reported values from Table (10.1). . . . 150

11.1 Parameter estimates of the Gaussian model adjustment on the mean correlograms. 167
11.2 Parameter estimates of the time variable Gaussian model adjustment. The $\hat{r}$ are reported values from Table (11.1). 169

À la mémoire de ma petite sœur, Sabrina.

Chapter 1

Introduction

1.1 Measuring sea-level at the coast

Over the past two centuries, the overproduction of greenhouse gases by industrial societies has resulted in a long-lasting increase in the surface temperature. One of the main consequences, but also one of the most indisputable evidence of this global change, is the increase in the global mean sea-level (GMSL), resulting from both sea-water expansion and continental ice-melting. In particular, based on the analysis of multidecadal time series provided by tide-gauges scattered around the world, the rate of sea-level change over the 1901-2010 period is estimated at 1.7 mm/yr (Church et al., 2013). With the advent of the satellite era, it has become possible to estimate the change in GMSL from the combination of globally distributed altimeter satellite observations. From these new observations, the rate of sea-level change is estimated at 3.2 mm/yr for the period 1993-2017 (Cazenave et al., 2018).

Although symptomatic of global transformations, the current GMSL change is not at all representative of the sea-level change that is and will be observed at the world coasts. One of the main reasons is that sea-level change shows a strong regional variability due to the heterogeneous distribution of continental ice-melting, sea-water temperature, salinity, and ocean dynamics (Cazenave and Cozannet, 2014). For instance, the rate of the regional sea-level exceeds 10.0 mm/yr in the west-pacific, which is several times the rate of the GMSL (Becker et al., 2012).

In addition to the spatial variability of absolute sea-level change, the evolution of sea-level is also influenced by vertical motions of the coast itself (Wöppelmann and Marcos, 2016). The subsidence or the uplift of the coast directly translates into an increase or a decrease in sea-level relative to the ground, which is the relevant quantity to consider for coastal management. Vertical land motions can be caused by a wide variety of natural and anthropogenic processes (Pfeffer et al., 2017), including the viscoelastic response to past and current continent ice-melting, known as glacial isostatic adjustment (GIA) (Peltier and Andrews, 1976; King et al., 2010), tectonic activity (Ballu et al., 2011; Ballu et al., 2019), sediment compaction (Syvitski et al., 2009; Galloway and Burbey, 2011) or surface mass loading (Van Dam, Blewitt, and Heflin, 1994; Mémin, Boy, and Santamaria-Gomez, 2020).

Forecasting relative sea-level change at the coast is a crucial issue in sea risk management given the anthropization of the coast and the vulnerability of both private and public infrastructure (Nicholls and Cazenave, 2010). Given the spatial variability of both absolute sea-level change and vertical land motion, to address this challenge, it is necessary to benefit from reliable in-situ measurements of both sea-level and vertical land motion.

Due to intrinsic difficulties in distinguishing water from land, satellite altimetry observations are less accurate near the coasts, ground-based tide-gauges are still considered as the most accurate source of coastal sea-level measurements. For the vertical land motion, the most reliable technique to monitor long-term deformations in a geocentric frame is currently the Global Navigation Satellite Systems (GNSS), and the Global Positioning System (GPS) in particular (Wöppelmann et al., 2007).

Studying and understanding the temporal evolution of the sea-level along the coasts requires long time series with high metrological quality, and not only for sea-level but also for the vertical land motions. The quality and stability of the two main references, that is, tide gauges and GNSS time series are still a matter of research (Wöppelmann and Marcos, 2016) and providing observations with such high standards is a significant metrological challenge.

1.2 Related metrological requirements

1.2.1 Tide gauge accuracy

The simplest and oldest types of sea-level measurements originate from graduated poles placed against a vertical structure at the coast, requiring human intervention to perform observations. Since then, the art of measuring sea-level evolved towards self-recording tide gauges, first with mechanical gauges (IOC, 1985), and then with pressure gauges (IOC, 2002), acoustic gauges (IOC, 2006), and, more recently, radar gauges (IOC, 2016). Today, they are clustered into networks of continuously operating stations, and are key components of storm surge or tsunami warning systems and climate-related monitoring programs, such as the Global Sea Level Observing System (GLOSS) (IOC, 2012).

Among the scientific and technical applications of tide gauges, climate studies require the highest accuracy. For instance, a tide gauge complying with GLOSS standards should be capable of measuring instantaneous sea-level with an accuracy better than 1.0 cm, in all conditions of tide, waves, currents and weather (IOC, 2016). In practice, this level of accuracy cannot be ensured by laboratory experiments, so tide-gauges must be calibrated and regularly controlled in real measuring conditions. Hence, considering the multiple sources of error that can affect tide gauge measurements (Míguez Martín, Testut, and Wöppelmann, 2008) and the effects of aggressive marine environment, maintaining the stability of sea-level measurement over several decades constitutes an important financial, technical and scientific challenge.

The benefit in maintaining tide gauge accuracy goes beyond the analysis of tide gauge time series for climatic studies. In practice, coastal tide gauge are also involved in the calibration of several other sea-level sensors, including GNSS buoys (André et al., 2013), kinematic systems (Chupin et al., 2020), GNSS reflectometry (Larson, Ray, and Williams, 2017; Williams et al., 2020), and altimeter satellites (Bonnefond, Haines, and Watson, 2011; Watson et al., 2011). Therefore, ensuring the quality of tide-gauge measurement is beneficial for many oceanographic studies.

1.2.2 GNSS time series accuracy

According to Wöppelmann and Marcos (2016), vertical land motions should be determined with standard errors that are one order of magnitude lower than the contemporary climate signals of 1 to 3 millimetres per year observed on average in sea-level records, either by using tide gauges or satellites.

Over the years, analyses of globally distributed networks of GNSS stations have proved capable to assess secular VLMs at below millimetre per year level (Wöppelmann et al., 2007; Santamaría-Gómez et al., 2012). However, estimating VLMs with an accuracy below the 0.1 millimetres per year level remains a considerable challenge because the final uncertainty on the vertical component of a station depends on two main limitations.

The first and most direct limitation is the precision GNSS positioning itself. In particular, VLM accuracy depends on many factors, which includes, for instance: the accuracy of orbits and clock products, of the modelling of tropospheric and ionospheric delays of GNSS signals propagating through the atmosphere, of the accuracy on the phase centres of both satellites and antennas, of the number of satellites visible by the antenna, of possible multi-path effects, and the stability of the origin and the scale of the reference frame used to express the estimated motion (Altamimi et al., 2016). Therefore, any improvement in one of these aspects of GNSS positioning would benefit to the accuracy of VLM estimates.

The second and more indirect limitation is the difficulty of extracting reliable secular trends from GNSS position time series. From the statistical inference point of view, the uncertainty of trend estimates depends on the span and completeness of the time series, but also on the statistical model used to describe its motion. In particular, two main effects limit the precision of trend estimates, namely the presence of *position offsets* and the influence of *time-correlated stochastic processes*.

To model the deterministic motion visible in GNSS potion time series, one relies on a trajectory model. This trajectory model usually accounts for a linear trend, periodic oscillations, and position offsets (Bevis and Brown, 2014). Position offsets are abrupt changes in position that typically result from instrumental changes or large earthquakes (Gazeaux et al., 2013). If not accounted for in the trajectory model, offsets bias systematically trend estimates (Williams, 2003b). However, even when accounted for, offsets remain a nuisance as it increases the uncertainty of all estimates, and especially the trend uncertainties (Williams, 2003b; Griffiths and Ray, 2016; Wang and Herring, 2019).

To model the part of the time series that is not described by the trajectory model, referred to as stochastic variability, we use a stochastic model. The stochastic variability was first described using uncorrelated white noise. However, several studies demonstrated the stochastic variability in GNSS position time series is time-correlated, and that a combination of white noise and power-law stochastic processes more realistic (Zhang et al., 1997; Mao, Harrison, and Dixon, 1999; Calais, 1999; Williams et al., 2004; Santamaría-Gómez et al., 2011). In particular, modelling the stochastic variability using time-correlated processes is crucial for the estimation of trends, because assuming only white noise yields to far too optimistic results. However, the origin of such a time-correlation of the GNSS stochastic variability remains unknown. In particular, it is unclear whether such a time correlation results

from positioning errors or from non-modelled geophysical processes. Significant progress in this research field could dramatically increase the precision of VLM estimates.

1.3 Research objectives

The scientific goal of this thesis is to improve our understanding of the biases and uncertainties of tide gauges and GNSS products, both by the development of statistical methods and by the analyses of real observations. Through these statistical analyses, we aim at identifying potential areas of improvements that could lead to an enhanced metrologic quality of these geodetic products. Due to the fundamental differences between the two quantities of interest, our research work is divided into two main research axes, one for tide-gauge observations, and one for GNSS position time series.

Concerning the study of the tide gauge calibration procedure, we aim at providing a complete characterization of the performance of tide gauges during calibration experiments and evaluate whether Intergovernmental Oceanographic Commission (IOC) requirements can be met. To this end, we present a combination method dedicated to the assessment of the biases and uncertainties of different sea-level sensor technologies from inter-comparison experiments involving multiple instruments.

The second research axis is dedicated to the study of the time-correlation in GNSS vertical land motion time series. Here, we aim at identifying possible sources of time-correlated processes in vertical land motions. To this end, we first study the statistical methods for determining the stochastic model of GNSS position time series. Then, we examine the global variability of time-correlation properties in globally distributed GNSS stations. Finally, we investigate whether the deformation induced by geophysical processes could explain the observed variability.

1.4 Outline of the thesis

This thesis presents six contributions to the assessment and the improvement of the quality of tide-gauge sea-level measurements and GNSS-based vertical land motions. Because these contributions tend to address separate research questions, we chose to present them in 6 independent chapters that can be read independently of one another. In particular, each contribution chapter is presented in an article format with its introduction, a recall of the methods, and its results, so the reader is not required to read the thesis entirely.

Although they address different research questions, these six contributions tend to rely on similar statistical tools. Considering that the article format employed for the contribution chapters cannot provide a beginner-friendly introduction to these methods, the six contribution chapters are preceded by four methodological chapters providing a comprehensive introduction to the main statistical tools used in this thesis.

In chapter 2, we recall the main methods dedicated to the determination of *functional* parameters from observations. In particular, we present the weighted least-squares and the maximum likelihood estimation methods, which are extensively

used throughout the thesis. We also present commonly used accuracy metrics and introduce the misclosures, used in chapters 8-11.

In chapter 3, we provide an introduction to the determination of *stochastic* parameters from observations. We present three variance component estimation methods, namely, the maximum likelihood estimation (MLE), the restricted maximum likelihood estimation (RMLE), and the Least-Squares Variance Component Estimation (LS-VCE). These methods are used in chapters 7-11 and compared in chapter 9.

In chapter 4, we introduce the basics of geodetic time series analysis. We present functional models allowing to describe trends, periodic oscillations, discontinuities and transients in time series. We also introduce three frequent stochastic models, namely first-order Gauss-Markov, power-law, and generalised Gauss-Markov processes. These models are employed in chapters 8-11.

The chapter 5 presents a method dedicated to the detection of non-modelled signal in geodetic time series. Based on hypothesis testing in linear models, and the Generalised Likelihood Ratio Test (GLRT) in particular, the method allows testing for the significance of signals in the presence of time-correlated processes. We illustrate applications of the GLRT to the detection of periodic signals and discontinuities. This method is used in chapters 9 and 10.

In chapter 6, we study a tide-gauge calibration experiment. We assess the influence of a scale error on different aspects of the determination of the tide-gauge sensor offset. Although estimating the sensor offset is a mandatory step, in other studies, it is unclear how this step is handled in the presence of a scale error. Hence, this chapter addresses this lack of information and provides indications for the design of calibration experiments.

In chapter 7, we focus on the assessment of the biases and uncertainties of tide-gauges. We propose a combination method based on the analysis of multiple collocated time series. The proposed method provides a complete description of the performances of each tested gauge. We apply it to a tide-gauge comparison experiment involving six collocated tide-gauges, realized in 2016 on the Aix island off the mid-Atlantic coast of France.

In chapter 8, we study the estimation of stochastic parameters from GNSS position time series. We investigate the potential of the nonlinear LS-VCE method for the estimation of the parameters of both power-law and generalized Gauss-Markov stochastic models. Because this method not only provides unbiased and minimum-variance stochastic parameter estimates, but also their uncertainties, it could reveal to be useful for future stochastic model investigations.

In chapter 9, we assess the influence of offsets on the estimation of the stochastic parameters in GNSS time series. The effect of offsets on functional parameters is already well established. However, their impact in the stochastic modelling remains poorly understood. In this chapter, we show that offsets affect various aspects of the stochastic modelling, namely spectral analysis, stochastic model selection, and variance component uncertainty.

In chapter 10, we examine the global-scale variability of stochastic model parameters in global GNSS solutions. In particular, we analyse the north, east, and vertical

time series of over 10,000 stations computed by the Nevada Geodetic Laboratory at the University of Nevada Reno. To better understand the potential source of stochastic variability, we also study the influence of a recent reprocessing. We show that, even in the most recent solution, the vertical land motions reveal a spatial variability of stochastic parameter estimates that stands out from horizontal displacements.

In chapter 11, we further examine the spatial variability of stochastic parameters of vertical land motions. In particular, we investigate whether surface mass loading effects can explain their spatial variability. To this end, we use both the most recent NGL time series and the non-tidal atmospheric loading, non-tidal oceanic loading and hydrological loading deformation time series computed by the Earth System Modelling team of the GeoForschungsZentrum Potsdam. We show that non-tidal atmospheric and oceanic loading deformations explain a large part of the stochastic variability observed in vertical land motions.

Finally, chapter 12 synthesises the different contributions of this thesis and discusses future research perspectives.

Chapter 2

Estimation of deterministic parameters from observations

Whether analysing vertical land motions, water level fluctuations, or any other measurement of a geophysical quantity, an essential step of data analysis is to convert raw observations into interpretable parameters. The main challenge is that, in experimental sciences, observations never can be modelled using a purely deterministic model. At the very least, one must account for the uncertainties of the experimental protocol used to obtain measurements. But more generally, complex physical phenomena failing to benefit from physical modelling must also be considered as a stochastic variability.

The presence of stochastic variability in the observations has two main consequences on data analysis. The first one is that there exists no perfect way to estimate parameters, and the analyst must make a choice among possible estimation methods. The second is that parameters can only be estimated with limited precision, which must be evaluated for physical interpretations.

To estimate parameters and their uncertainties, we rely on *inferential statistics*. In this chapter, we review the basics of geodetic *weighted least-squares estimation* and *maximum likelihood estimation*. These two estimation methods, and their underlying mathematical foundations, are extensively used throughout this thesis. This chapter should provide a comprehensive description of the essential statistical concepts.

2.1 Functional and stochastic models

2.1.1 Scalar case

A common practice is to assume that an observation, noted y_i, can be decomposed as the sum of deterministic component, noted μ_i, and a stochastic component, noted e_i, so that

$$y_i = \mu_i + e_i. \tag{2.1}$$

The term e_i, also referred to as *residual* or *noise*, represents the effect of the stochastic variability on the observation y_i. In practice, e_i is assumed to be the realization of a continuous stochastic variable, noted $\underline{e_i}$, following a centred normal distribution. The dispersion of this distribution is quantified by a parameter, the variance, noted σ_i^2. These assumptions for $\underline{e_i}$ are more compactly denoted as $\underline{e_i} \sim \mathcal{N}(0, \sigma_i^2)$.

The normal distribution assumption for $\underline{e_i}$ is motivated by its omnipresence in experimental sciences. This is supported theoretically by the central limit theorem,

which explains how the distribution of a sum of independent and identically distributed stochastic variables converges to the normal distribution when the number of considered variables increases.

A direct implication of the above distribution assumption for $\underline{e_i}$ is that the observation y_i can itself be described as a realization of a continuous stochastic variable, noted $\underline{y_i}$, following a normal distribution with the same variance σ_i^2, but centred on μ_i, that is $\underline{y_i} \sim \mathcal{N}(\mu_i, \sigma_i^2)$.

Hence, in some studies, Equation (2.1) is often directly written in terms of the stochastic variables $\underline{y_i}$ and $\underline{e_i}$, namely

$$\underline{y_i} = \mu_i + \underline{e_i}, \tag{2.2}$$

so that, by "observation", we often refer to the underlying stochastic variable $\underline{y_i}$ and not specifically each scalar measurement y_i.

In order to precise a bit more these distribution assumptions for $\underline{y_i}$, we can define its probability density function (PDF), noted $p_{\underline{y_i}}(y_i)$, as

$$p_{\underline{y_i}}(y_i) = \frac{1}{\sqrt{2\pi \cdot \sigma_i^2}} \exp\left(-\frac{1}{2}\frac{(y_i - \mu_i)^2}{\sigma_i^2}\right). \tag{2.3}$$

An important property of $p_{\underline{y_i}}(y_i)$ that it is fully described using only two parameters, namely its *first raw moment* and its *second central moment*. The first raw moment is the so-called **expectation**, noted $\mathsf{E}\{.\}$, and defined by

$$\mathsf{E}\{\underline{y_i}\} = \mu_i, \tag{2.4}$$

and its second central moment is the so-called **variance**, noted $\mathsf{var}\{.\}$, and defined by

$$\mathsf{var}\{\underline{y_i}\} = \mathsf{E}\{(\underline{y_i} - \mu_i)^2\} = \sigma_i^2. \tag{2.5}$$

In other words, providing a description for the deterministic component μ_i of an observation corresponds to the study of $\mathsf{E}\{\underline{y_i}\}$, whereas providing a description the stochastic component e_i corresponds to the study of $\mathsf{var}\{\underline{y_i}\}$.

In the following, this distinction is crucial as it leads to the definition of two complementary models, namely the **functional** and the **stochastic** models. The **functional model** defines the expectation of $\underline{y_i}$, whereas the **stochastic model** defines the variance of $\underline{y_i}$. These models are complementary in the sense that a correct description of both the **functional** and the **stochastic model** is required to *fully* describe the distribution of the observations.

2.1.2 Multivariate case

So far, we considered a single observation y_i. For a set of m observations, we can re-write the relation (2.2) using vector notation so that

$$\begin{bmatrix} \underline{y_1} \\ \underline{y_2} \\ \vdots \\ \underline{y_m} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{bmatrix} + \begin{bmatrix} \underline{e_1} \\ \underline{e_2} \\ \vdots \\ \underline{e_m} \end{bmatrix}, \tag{2.6}$$

which can be more compactly written as

$$\boxed{\underline{\mathbf{y}} = \boldsymbol{\mu} + \underline{\mathbf{e}}.} \tag{2.7}$$

In the multivariate case, $\underline{\mathbf{y}}$ is a stochastic vector following a *multivariate normal distribution*, whose multivariate probability density function, noted $p_{\underline{\mathbf{y}}}(\mathbf{y})$, is defined by

$$p_{\underline{\mathbf{y}}}(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^m \cdot \det(\mathbf{Q}_{\mathbf{y}})}} \exp\left(-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T \mathbf{Q}_{\mathbf{y}}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right), \tag{2.8}$$

in which $\mathbf{Q}_{\mathbf{y}}$ is the so-called $m \times m$ covariance matrix of $\underline{\mathbf{y}}$, $\det(.)$ denotes the determinant, $(.)^{-1}$ denotes the inverse, and $(.)^T$ the transpose.

In the multivariate case, the functional model of the observation vector $\underline{\mathbf{y}}$ is noted $\mathsf{E}\{\underline{\mathbf{y}}\}$ and follows

$$\boxed{\mathsf{E}\{\underline{\mathbf{y}}\} = \boldsymbol{\mu}.} \tag{2.9}$$

and the stochastic model of the observation vector $\underline{\mathbf{y}}$, denoted $\mathsf{var}\{\underline{\mathbf{y}}\}$, follows

$$\boxed{\mathsf{var}\{\underline{\mathbf{y}}\} = \mathbf{Q}_{\mathbf{y}}.} \tag{2.10}$$

The diagonal elements of the covariance matrix $\mathbf{Q}_{\mathbf{y}}$ models the variance of each observation, whereas the non-diagonal elements model the covariance between observations.

In this chapter, it is assumed that the covariance matrix $\mathbf{Q}_{\mathbf{y}}$ and, thereby, the stochastic model of the observations, is *completely known*. In practice, this is almost never the case, so in chapter 3, we will relax this hypothesis.

2.2 Model of observation equations

We now assume that the deterministic vector $\boldsymbol{\mu}$ in Equation (2.9) can be modelled with a set of n unknown parameters $(x_1, x_2, \cdots, x_n)$, through a multivariate function $\boldsymbol{f}(.)$ so that

$$\mathsf{E}\{\underline{\mathbf{y}}\} = \boldsymbol{\mu} = \boldsymbol{f}(\mathbf{x}), \tag{2.11}$$

in which $\mathbf{x}$ is the unknown $n \times 1$ parameter vector, defined as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

The model $\mathsf{E}\{\underline{\mathbf{y}}\} = f(\mathbf{x})$ is referred to as the **model of observation equations** (Teunissen, 2000a). For such functional models, the goal of parameter estimation methods is to derive an estimator $\underline{\hat{\mathbf{x}}}$ of the unknown parameter vector $\mathbf{x}$ based on the observation vector $\underline{\mathbf{y}}$ and the function $f(.)$.

2.2.1 Linear case

When the function $f(.)$ is linear with respect to $\mathbf{x}$, the *linear* model of observation equations can be written as

$$\mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x} \tag{2.12}$$

in which $\mathbf{A}$ is the so-called $m \times n$ design matrix, that is, the linear function $f(.)$ expressed in matrix notation.

Although it may seem like an oversimplification, a wide range of estimation problems in geodesy turns out to be linear with respect to the unknown parameters. For instance, this thesis exclusively deals with linear functional models. Besides, linear models are convenient because they are easy to solve, such that nonlinear problems are turned into linear ones whenever it is possible.

2.2.2 Nonlinear case

When the function $f(.)$ is nonlinear, a common strategy is to linearize $f(.)$ around an approximate estimate of $\mathbf{x}$, noted $\mathbf{x}_0$. The first-order Taylor series expansion of $f(\mathbf{x})$ at $\mathbf{x}_0$ reads

$$\begin{aligned} f(\mathbf{x}) &\approx f(\mathbf{x_0}) + \left.\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_0} (\mathbf{x} - \mathbf{x}_0), \\ &\approx f(\mathbf{x_0}) - \left.\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_0} \mathbf{x}_0 + \left.\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_0} \mathbf{x}. \end{aligned} \tag{2.13}$$

Because $\mathsf{E}\{\underline{\mathbf{y}}\} = f(\mathbf{x})$, and because $\mathbf{x}_0$ is known, one can recover a system of linear equations as

$$\mathsf{E}\{\underline{\mathbf{y}}\} - f(\mathbf{x_0}) + \left.\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_0} \mathbf{x}_0 \approx \left.\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}_0} \mathbf{x}. \tag{2.14}$$

Hence, it is possible to turn nonlinear models of observation equations into linear ones. This approach is not perfect because the linearized function is just a local approximation of the nonlinear function $f(.)$. This approximation issue can be handled using the so-called Newton-Raphson algorithm. The principle of this algorithm is to start with an initial parameter $\mathbf{x}_0$, and then solve Equation (2.14) to obtain a new estimate $\mathbf{x}_1$. By repeating this step, one can obtain new estimates $\mathbf{x}_{i+1}$ using previous estimates $\mathbf{x}_i$. The algorithm stops when the difference between two estimates become negligible. The application of this method to least-squares estimation is described in more details in (Teunissen, 2000a). Note that this strategy can only work if the guess $\mathbf{x}_0$ is relatively close to the true solution. Hence, in some problems, the choice of $\mathbf{x}_0$ may be important.

Because solving a nonlinear problem can come down to solving linear ones, in the following, we exclusively focus on parameter estimation for linear models.

2.2.3 Estimation principle

A basic requirement for the *unconstrained* estimation of n parameters from m observations in a linear equation system is that there are at least n observations to avoid an infinite number of solutions. In practice, it is preferable to derive only a few parameters n from many observations m, that is $m >> n$, to benefit from a redundancy of information.

Since there is usually no exact solution for $m > n$, one must instead find the estimator $\hat{\mathbf{x}}$ that optimizes a certain criterion. In the following, methods optimizing two different criteria are presented: the **Weighted Least-Squares Estimation** method and the **Maximum Likelihood Estimation** method.

2.3 Weighted Least-Squares Estimation

2.3.1 Weighted Least-Squares Estimator

The objective of Weighted Least-Squares Estimation is to derive an estimator $\hat{\mathbf{x}}$ that minimizes a quadratic *error function* such that

$$\hat{\underline{\mathbf{x}}} = \underset{\mathbf{x}\in\mathbb{R}^n}{\operatorname{argmin}} (\underline{\mathbf{y}} - \mathbf{A}\mathbf{x})^T \mathbf{W} (\underline{\mathbf{y}} - \mathbf{A}\mathbf{x}), \tag{2.15}$$

in which $\mathbf{W}$ is the so-called $m \times m$ weight matrix. In principle, the weight matrix $\mathbf{W}$ can be any symmetric and positive-definite matrix. In the particular case in which the weight matrix is the identity matrix, that is $\mathbf{W} = \mathbf{I}$, the expression is equivalent to minimizing the sum of the squares of the residuals, hence the name *Least-Squares*.

The standard strategy to minimize a function is to search the argument for which its derivative is null. In other words, the goal is to solve

$$\frac{\partial}{\partial \mathbf{x}} \{ (\underline{\mathbf{y}} - \mathbf{A}\mathbf{x})^T \mathbf{W} (\underline{\mathbf{y}} - \mathbf{A}\mathbf{x}) \} = \mathbf{0}, \tag{2.16}$$

which leads to

$$\mathbf{A}^T \mathbf{W} \mathbf{A} \underline{\mathbf{x}} - \mathbf{A}^T \mathbf{W} \underline{\mathbf{y}} = \mathbf{0}, \tag{2.17}$$

which can be rewritten the so-called *normal equations system*

$$\boxed{\mathbf{A}^T \mathbf{W} \mathbf{A} \underline{\mathbf{x}} = \mathbf{A}^T \mathbf{W} \underline{\mathbf{y}}.} \tag{2.18}$$

The $n \times n$ matrix $\mathbf{A}^T\mathbf{W}\mathbf{A}$ in the left hand-side of Equation (2.18) is known as the *normal matrix*, and is usually denoted $\mathbf{N}$. The $n \times 1$ vector in the right hand-side of Equation (2.18) does not have a specific name, but is sometimes denoted $\underline{\mathbf{l}} = \mathbf{A}^T \mathbf{W} \underline{\mathbf{y}}$, such that

$$\mathbf{N} \underline{\mathbf{x}} = \underline{\mathbf{l}}. \tag{2.19}$$

If the design matrix $\mathbf{A}$ is of full column rank, – that is, if the columns of $\mathbf{A}$ are linearly independent – then the normal matrix $\mathbf{N}$ is positive-definite. It can be shown that $\mathbf{N}$ is the second-order partial derivative of the error function. Hence, its positive-definiteness indicates that the solution of Equation (2.16) is indeed the argument of a (local) minimum.

Finding $\underline{\hat{\mathbf{x}}}$ comes down to solving a linear system of normal equations. Since $\mathbf{N}$ is positive definite, it is invertible, and the weighted least-squares estimator $\underline{\hat{\mathbf{x}}}$ can be obtained as $\underline{\hat{\mathbf{x}}} = \mathbf{N}^{-1}\underline{\mathbf{l}}$, which is the short notation for

$$\boxed{\underline{\hat{\mathbf{x}}} = (\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}\underline{\mathbf{y}}.} \tag{2.20}$$

Noting $\mathbf{A}^+$ the weighted pseudo-inverse of $\mathbf{A}$ defined as

$$\mathbf{A}^+ = (\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}, \tag{2.21}$$

the Equation (2.20) can be more compactly written as

$$\boxed{\underline{\hat{\mathbf{x}}} = \mathbf{A}^+\underline{\mathbf{y}},} \tag{2.22}$$

which illustrates that the estimator $\underline{\hat{\mathbf{x}}}$ is a linear transformation of $\underline{\mathbf{y}}$, that is to say, $\underline{\hat{\mathbf{x}}}$ is a *linear estimator* of $\mathbf{x}$.

Moreover, by substituting $\underline{\mathbf{y}} = \mathbf{A}\mathbf{x} + \underline{\mathbf{e}}$ into Equation (2.20) and applying linear propagation law of means, it can be shown that

$$\begin{aligned} \mathsf{E}\{\underline{\hat{\mathbf{x}}}\} &= (\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}\mathbf{A}\mathbf{x} + \mathbf{A}^+\mathsf{E}\{\underline{\mathbf{e}}\} \\ &= \mathbf{x} + \mathbf{A}^+\mathbf{0} \\ &= \mathbf{x}, \end{aligned} \tag{2.23}$$

which demonstrates that the weighted least squares estimator $\underline{\hat{\mathbf{x}}}$ is an *unbiased* estimator of $\mathbf{x}$. This property is true for any choice of positive-definite weight matrix $\mathbf{W}$.

2.3.2 Uncertainty of the weighted least-squares estimator

To evaluate how the variance of the observations and the chosen functional model affects the precision of the estimates, one can compute the $n \times n$ covariance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$ of the estimator $\hat{\mathbf{x}}$. Since the weighted least-squares estimator is expressed as a linear transformation of the observation vector $\underline{\mathbf{y}}$, one can apply the linear propagation law of variance, and show that the covariance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$ follows

$$\boxed{\mathbf{Q}_{\hat{\mathbf{x}}} = \mathbf{A}^+\mathbf{Q}_{\mathbf{y}}\mathbf{A}^{+T},} \tag{2.24}$$

in which, the i-th diagonal element of $\mathbf{Q}_{\hat{\mathbf{x}}}$ corresponds to the estimated variance $\sigma^2_{\hat{x}_i}$ of the i-th parameter estimate $\hat{x}_i$. The non diagonal elements denote the covariance between the estimated parameters.

2.4 Maximum Likelihood Estimation (MLE)

2.4.1 Likelihood function

In the case of linear models, the multivariate PDF of $\underline{\mathbf{y}}$ depends on the parameter vector $\mathbf{x}$, and is usually denoted $p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x})$ so that

$$p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^m \cdot \det(\mathbf{Q}_\mathbf{y})}} \exp\left(-\frac{1}{2}(\underline{\mathbf{y}} - \mathbf{A}\mathbf{x})^T \mathbf{Q}_\mathbf{y}^{-1}(\underline{\mathbf{y}} - \mathbf{A}\mathbf{x})\right). \tag{2.25}$$

In practice, the observation vector $\mathbf{y}$ and the covariance matrix $\mathbf{Q}_\mathbf{y}$ in Equation (2.25) are known. When $p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x})$ is taken as a function of $\mathbf{x}$ and not $\mathbf{y}$, it is referred to as the *likelihood* function.

2.4.2 Maximum Likelihood Estimator

The objective of Maximum Likelihood Estimation (MLE) is to find the parameter vector $\hat{\underline{\mathbf{x}}}$ for which the observation vector $\mathbf{y}$ is the most likely to be observed. This is done by maximizing the likelihood function $p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x})$ of observing $\mathbf{y}$ for a given parameter vector $\mathbf{x}$, that is

$$\hat{\underline{\mathbf{x}}} = \underset{\mathbf{x}\in\mathbb{R}^n}{\operatorname{argmax}}\, p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}). \tag{2.26}$$

The logarithm function $\ln(.)$ being monotonous, it is actually equivalent, and more convenient, to maximize the log-likelihood function $\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}))$ defined as

$$\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x})) = -\frac{1}{2}[m\ln(2\pi) + \ln(\det(\mathbf{Q}_\mathbf{y})) + (\underline{\mathbf{y}} - \mathbf{A}\mathbf{x})^T \mathbf{Q}_\mathbf{y}^{-1}(\underline{\mathbf{y}} - \mathbf{A}\mathbf{x})]. \tag{2.27}$$

Note that maximizing Equation (2.27) is equivalent to minimizing the quadratic term

$$\hat{\underline{\mathbf{x}}} = \underset{\mathbf{x}\in\mathbb{R}^n}{\operatorname{argmin}}\, (\underline{\mathbf{y}} - \mathbf{A}\mathbf{x})^T \mathbf{Q}_\mathbf{y}^{-1}(\underline{\mathbf{y}} - \mathbf{A}\mathbf{x}). \tag{2.28}$$

Hence, under the normal distribution assumption, maximizing the likelihood function is equivalent to minimizing the norm of the residual vector $\mathbf{e}$ using the matrix $\mathbf{Q}_\mathbf{y}^{-1}$ as a metric. The general solution to the above minimization problem is addressed by the Weighted Least-Squares Estimation method presented in the previous section.

Thus, using Equation (2.20) with $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-1}$, one can show that the Maximum Likelihood Estimator of $\mathbf{x}$ reads

$$\boxed{\hat{\underline{\mathbf{x}}} = (\mathbf{A}^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_\mathbf{y}^{-1} \underline{\mathbf{y}}.} \tag{2.29}$$

2.4.3 Uncertainty of the maximum likelihood estimator

The uncertainty of the maximum likelihood estimator follows from that of the weighted least-squares estimator in Equation (2.24). By taking $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-1}$, after a few simplifications, it follows that

$$\boxed{\mathbf{Q}_{\hat{\mathbf{x}}} = (\mathbf{A}^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{A})^{-1}.} \tag{2.30}$$

2.5 Best Linear Unbiased Estimation

Intuitively, there must be some limit to the precision of the parameters estimated from a limited amount of observations. In the case of linear estimators, this limit is formalized by the so-called Cramér-Rao lower bound. This lower bound provides the variance of the *Best Linear Unbiased Estimator* (BLUE) when it exists.

The Cramér-Rao lower bound, is given as

$$\mathcal{I}(\mathbf{x})^{-1}, \tag{2.31}$$

in which, $\mathcal{I}(\mathbf{x})$ is the so-called $n \times n$ Fisher information matrix defined as

$$\mathcal{I}(\mathbf{x}) = -\mathsf{E}\left\{\frac{\partial^2}{\partial \mathbf{x}\mathbf{x}^T}\{\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}))\}\right\}, \tag{2.32}$$

where $\frac{\partial^2}{\partial \mathbf{x}\mathbf{x}^T}\{.\}$ denotes the second order partial derivative, or Hessian matrix, of a multivariate function with respect to $\mathbf{x}$.

Since the only differences between minimizing the least-squares criterion and maximizing the log-likelihood are a change of sign and a change of weight matrix, it follows from Equation (2.17) that the Hessian matrix of $\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}))$ follows

$$\frac{\partial^2}{\partial \mathbf{x}\mathbf{x}^T}\{\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}))\} = -(\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{A}). \tag{2.33}$$

Given that $(\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{A})$ is positive-definite, the Hessian matrix is negative-definite, as expected from a local maximum. Substituting Equation (2.33) into Equation (2.32) gives

$$\mathcal{I}(\mathbf{x}) = (\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{A}), \tag{2.34}$$

which, using Equation (2.31), leads to the Cramér-Rao lower bound, namely

$$\mathcal{I}(\mathbf{x})^{-1} = (\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{A})^{-1}. \tag{2.35}$$

One important observation is that the covariance matrix of the maximum likelihood estimator in Equation (2.30) equals the Cramér-Rao lower bound (2.35), that is $\mathbf{Q}_{\hat{\mathbf{x}}} = \mathcal{I}(\mathbf{x})^{-1}$. Hence, by choosing $\mathbf{W} = \mathbf{Q}_{\mathbf{y}}^{-1}$ as a weight matrix, we obtain an *unbiased* and *minimum-variance* least-squares estimator of $\mathbf{x}$. In other words we obtain the best possible Linear Unbiased Estimator, which is often referred to as the *Best Linear Unbiased Estimator*, or *BLUE*.

Another useful observation is that the Hessian matrix is a measure of the curvature of a multivariate function, hence the minimum possible variance of an estimator is closely related to the sharpness of the log-likelihood function around $\mathbf{x}$ (Bos et al., 2020). This observation helps in understanding the results in chapter 9.

2.6 Measures of inconsistency

One benefit of having more observations than unknowns is that it allows quality control. In practice, there exist several measures of inconsistency that can be used to investigate the quality of a functional model and identify possible miss-specifications.

In the following, we present two of these measures: the estimator of the residuals, and the misclosure vector.

2.6.1 Residuals

The most intuitive measure of inconsistency is probably the estimator of the residual vector, denoted $\underline{\hat{\mathbf{e}}}$. It is defined as the difference between the observations $\mathbf{y}$ and the best fitting model $\mathbf{A}\underline{\hat{\mathbf{x}}}$, that is

$$\underline{\hat{\mathbf{e}}} = \underline{\mathbf{y}} - \mathbf{A}\underline{\hat{\mathbf{x}}} \tag{2.36}$$

In some applications, the vector $\underline{\hat{\mathbf{e}}}$ is directly expressed as a linear function of the observations $\underline{\mathbf{y}}$ such as

$$\boxed{\underline{\hat{\mathbf{e}}} = \mathbf{P}_{\mathbf{A}}^{\perp}\underline{\mathbf{y}}}, \tag{2.37}$$

in which the $m \times m$ matrix $\mathbf{P}_{\mathbf{A}}^{\perp}$ is the least-squares orthogonal projector defined by

$$\boxed{\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}}. \tag{2.38}$$

The weighted least-squares estimator being unbiased, the expectation of the estimator of the residuals reads

$$\mathsf{E}\{\underline{\hat{\mathbf{e}}}\} = \mathbf{0}, \tag{2.39}$$

meaning that variability of the estimated residuals should be purely stochastic.

By applying the linear propagation law of variance to Equation (2.37), one can show that the covariance matrix of the vector $\underline{\hat{\mathbf{e}}}$, noted $\mathbf{Q}_{\hat{\mathbf{e}}}$, reads

$$\mathsf{var}\{\underline{\hat{\mathbf{e}}}\} = \mathbf{Q}_{\hat{\mathbf{e}}} = \mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{Q}_{\mathbf{y}}\mathbf{P}_{\mathbf{A}}^{\perp T}. \tag{2.40}$$

Since the vector $\underline{\hat{\mathbf{e}}}$ has the same structure as the observations, analysing residuals allows a quick identification of possible modelling problems and observation outliers. Besides, residuals are often used to compute other accuracy metrics, such as the Root Mean Square Error (RMSE)

$$\mathsf{RMSE} = \sqrt{\frac{1}{m}\underline{\hat{\mathbf{e}}}^T\underline{\hat{\mathbf{e}}}}, \tag{2.41}$$

or the Weighted RMSE (WRMSE)

$$\mathsf{WRMSE} = \sqrt{\frac{1}{m}\underline{\hat{\mathbf{e}}}^T\mathbf{W}\underline{\hat{\mathbf{e}}}}. \tag{2.42}$$

Although useful to evaluate the goodness of fit of a chosen model, or to get an insight about the stochastic variability in the observations, the estimated residuals $\underline{\hat{\mathbf{e}}}$ cannot fully represent the true stochastic variability $\underline{\mathbf{e}}$ because the estimated parameters inevitably absorb part of it.

This is visible when analyzing the variance of Equation (2.36), which reads

$$\mathbf{Q}_{\hat{\mathbf{e}}} = \mathbf{Q}_{\mathbf{y}} - \mathbf{A}\mathbf{Q}_{\hat{\mathbf{x}}}\mathbf{A}^T. \tag{2.43}$$

Consequently, quality metrics exclusively based on the estimated residuals $\underline{\hat{\mathbf{e}}}$, such as RMSE or WRMSE, systematically underestimate the stochastic variability of

the observations. Note that the degree of underestimation depends on the chosen functional model, but also on the weight matrix $\mathbf{W}$.

2.6.2 Misclosures

Let's denote $b = m - n$ the degree of redundancy of the functional model. An alternative measure of inconsistency is the so-called misclosure $b \times 1$ vector $\underline{\mathbf{t}}$ defined as

$$\mathbf{B}^T \underline{\mathbf{y}} = \underline{\mathbf{t}}, \tag{2.44}$$

in which, $\mathbf{B}$ is a $m \times b$ matrix, whose columns span the null space of the matrix $\mathbf{A}^T$, so that

$$\mathbf{B}^T \mathbf{A} = \mathbf{0}. \tag{2.45}$$

The use of the matrix $\mathbf{B}$ has the effect of removing the influence of the functional model from the observations, so that $\underline{\mathbf{t}}$ only reflects the redundancy.

Given that $\underline{\mathbf{t}}$ is defined as a linear transformation of $\underline{\mathbf{y}}$, it also follows a multivariate normal distribution. Besides, using Equation (2.45), one can show that its expectation, noted $\mathsf{E}\{\underline{\mathbf{t}}\}$, reads

$$\mathsf{E}\{\mathbf{B}^T \underline{\mathbf{y}}\} = \mathsf{E}\{\underline{\mathbf{t}}\} = \mathbf{0} \tag{2.46}$$

and that its variance, noted $\mathsf{var}\{\underline{\mathbf{t}}\}$, reads

$$\mathsf{var}\{\underline{\mathbf{t}}\} = \mathbf{Q}_{\mathbf{t}} = \mathbf{B}^T \mathbf{Q}_{\mathbf{y}} \mathbf{B}. \tag{2.47}$$

The model (2.46), is referred to as the *model of condition equations*. Even though the vector $\underline{\mathbf{t}}$ cannot be used for a direct visual inspection of outliers because it does not have same structure as the observations, it is extremely used in quality control theory (Teunissen, 2000b) and Variance Component Estimation (Teunissen, 1988; Koch, 1986).

To understand why, we can study the following transformation (Teunissen and Amiri-Simkooei, 2008)

$$\begin{bmatrix} \underline{\hat{\mathbf{x}}} \\ \underline{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} (\mathbf{A}^T \mathbf{Q}_{\mathbf{y}}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_{\mathbf{y}}^{-1} \\ \mathbf{B}^T \end{bmatrix} \underline{\mathbf{y}} \tag{2.48}$$

First, one can show that this transformation in invertible, so that

$$\underline{\mathbf{y}} = \begin{bmatrix} \mathbf{A} & \mathbf{Q}_{\mathbf{y}} \mathbf{B} \mathbf{Q}_{\mathbf{t}}^{-1} \end{bmatrix} \begin{bmatrix} \underline{\hat{\mathbf{x}}} \\ \underline{\mathbf{t}} \end{bmatrix}, \tag{2.49}$$

which implies that all the information that is not absorbed by $\underline{\hat{\mathbf{x}}}$ is contained in $\underline{\mathbf{t}}$.

Then, from Equations (2.23) and (2.46), it follows that the expectation of Equation (2.48) reads

$$\mathsf{E}\left\{ \begin{bmatrix} \underline{\hat{\mathbf{x}}} \\ \underline{\mathbf{t}} \end{bmatrix} \right\} = \begin{bmatrix} \mathbf{x} \\ \mathbf{0} \end{bmatrix} \tag{2.50}$$

and that its variance reads

$$\mathsf{var}\left\{\begin{bmatrix}\underline{\hat{\mathbf{x}}}\\ \underline{\mathbf{t}}\end{bmatrix}\right\} = \begin{bmatrix}\mathbf{Q}_{\underline{\hat{\mathbf{x}}}} & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}_{\mathbf{t}}\end{bmatrix} \tag{2.51}$$

which shows that $\underline{\hat{\mathbf{x}}}$ and $\underline{\mathbf{t}}$ are independent.

Hence, unlike the variance of $\underline{\hat{\mathbf{e}}}$ in Equation (2.43), the variance of $\underline{\mathbf{t}}$ is independent from the variance of the estimator $\underline{\hat{\mathbf{x}}}$.

2.7 Misspecifications

So far, we assumed that the functional model (2.12) is correct. However, when dealing with actual measurements, it is frequent that a lack of knowledge about the measured process results in misspecifications of the functional model.

The most problematic kind of misspecification is certainly the *under-parametrization* of the functional model. When a functional model does not account for all deterministic effects, the weighted least-squares estimator is biased. Under-parametrization is an important issue in data analysis in general, as biased estimates can seriously affect conclusions drawn from a scientific study. The best strategy to overcome under-parametrization is probably to perform a careful analysis of the observation, sometimes with the help of hypothesis testing methods. We present this part in detail chapter 5.

A somewhat less dramatic kind of misspecification is the *over-parametrization* of the functional model. Over-parametrization implies that although all deterministic effects are considered, there exist unnecessary parameters in the functional model. The negative impact of such useless parameters is a systematic increase in uncertainty of the estimators. One way to overcome this is to remove non-significant parameters from the functional model.

The last possible issue is a misspecification in the stochastic model. This kind of misspecification has two main consequences. The first one is that the covariance matrix of the estimator obtained using Equation (2.24) or Equation (2.30) is inaccurate. Incorrect covariance matrix causes the estimated uncertainties to be either too optimistic or too pessimistic. The second consequence is that the best linear unbiased estimator loses the minimum variance property, meaning that the estimates are not as close as the true values as they could be.

Chapter 3

Estimation of stochastic parameters from observations

So far, we only addressed parameter estimation in the case the stochastic model is fully known. In other words, we assumed a perfect knowledge about the uncertainty of the measurements, and about the nature and the amplitude of the variability present in the observations. If it is even possible, such conditions may only be met in laboratory experiments. Hence, for sea-level and vertical land motion observations, it is necessary to relax this hypothesis and to be able to assess part of the stochastic properties from observations.

In practice, it is particularly important to be able to adapt the stochastic model based on observations because both parameter uncertainties and the minimum-variance property of the BLUE requires a realistic stochastic model.

In this chapter, we present how to derive stochastic parameters from observations, provided that a parametric stochastic model can be formulated. The main content of this chapter comes from the work of Amiri-Simkooei (2007), Teunissen (1988), and Teunissen and Amiri-Simkooei (2008).

3.1 Parametric stochastic model

In this chapter, we consider that the covariance matrix of the observation $\mathbf{Q_y}$ is partially unknown, and can be described using p unknown stochastic parameters $(\sigma_1, \sigma_2 \cdots \sigma_p)$ called *variance components*. Hence, hereafter, it is implicitly assumed that the matrix $\mathbf{Q_y}$ is a parametric function of these variance components, that is

$$\mathbf{Q_y} = \mathbf{Q_y}(\boldsymbol{\sigma}) \tag{3.1}$$

where $\mathbf{Q_y}(\boldsymbol{\sigma})$ is a symmetric and positive-definite $m \times m$ matrix and $\boldsymbol{\sigma}$ is the $p \times 1$ variance component vector, defined such that

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_p \end{bmatrix}. \tag{3.2}$$

The statistical methods devoted to the estimation of the variance components are referred to as *Variance Component Estimation (VCE)* methods. In the last decades, numerous VCE methods have been developed. Among the ones that are commonly used in geodesy, we can mention the Maximum Likelihood Estimation (MLE) (Harville, 1977), the Restricted Maximum Likelihood Estimation (RMLE) (Koch, 1986), the

Minimum Quadratic Unbiased Estimation (MINQUE) (Rao, 1971), the Best Invariant Quadratic Unbiased Estimation (BIQUE), the Helmert method, and the Least-Squares Variance Component Estimation (LS-VCE) (Teunissen and Amiri-Simkooei, 2008).

These methods were developed based on different estimation criteria. However, except the MLE method, for normally distributed observations, the RMLE, MINQUE, BIQUE, Helmert and LS-VCE methods are mathematically equivalent (Teunissen and Amiri-Simkooei, 2008). For a review of these VCE methods, we refer to Fotopoulos (2003) and Amiri-Simkooei (2007). In this chapter, we restrict our presentation to three different VCE methods discussed or used at some point in this thesis, namely the MLE, the RMLE and LS-VCE methods.

3.2 Maximum Likelihood Estimation (MLE)

This VCE method is the direct extension of the MLE principles introduced in chapter 2 to the determination of the variance components. It consists in adding the variance component vector $\hat{\sigma}$ as an additional parameter to the likelihood maximisation problem, that is

$$\hat{\mathbf{x}}, \hat{\sigma} = \underset{\mathbf{x}\in\mathbb{R}^n, \sigma\in\mathbb{R}^p}{\operatorname{argmax}} \ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}, \sigma)), \tag{3.3}$$

in which, assuming a normal distribution, $\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}, \sigma)$ is given by

$$\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}, \sigma)) = -\frac{1}{2}\left[m\ln(2\pi) + \ln(\det(\mathbf{Q}_{\mathbf{y}})) + (\underline{\mathbf{y}} - \mathbf{A}\mathbf{x})^T\mathbf{Q}_{\mathbf{y}}^{-1}(\underline{\mathbf{y}} - \mathbf{A}\mathbf{x})\right]. \tag{3.4}$$

For a given variance component vector σ, the optimal functional parameter vector $\hat{\mathbf{x}}$ can be directly obtained using the Best Linear Unbiased Estimator (BLUE). Hence, the likelihood $\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}, \sigma))$ is mostly a function of σ. However, unlike for $\hat{\mathbf{x}}$, there is no direct expression of the optimal variance components $\hat{\sigma}$ as a function of $\underline{\mathbf{y}}$. Hence, the assessment of $\hat{\sigma}$ relies on the use of nonlinear optimization methods. In some applications, the Nelder–Mead optimization method – also known as the downhill simplex method – is used, because it does not require to compute the partial derivatives of the likelihood function (Williams, 2008) with respect to the variance components.

Although the MLE method can provide unbiased and minimum-variance estimates of the functional parameters $\mathbf{x}$, this method is known to be a *biased estimator* of variance components (Harville, 1977). This bias is caused by the loss of redundancy resulting from the estimation of the functional parameters. In particular, the method is influenced by the uncertainty of the estimator $\underline{\mathbf{x}}$. Nonetheless, when the redundancy increases, the uncertainty of the estimator $\underline{\mathbf{x}}$ systematically reduces, and so is the bias on the variance components. Hence, the method is considered as *asymptotically unbiased*.

3.3 Restricted Maximum Likelihood Estimation (RMLE)

An *unbiased* alternative to the MLE method is the so-called Restricted Maximum Likelihood Estimation (RMLE) method as defined in Patterson and Thompson (1971)

and Koch (1986). To avoid the bias caused by the loss of redundancy, the RMLE maximizes the log-likelihood of the misclosure vector $\underline{\mathbf{t}}$, presented in chapter 2, instead of the observation vector $\mathbf{y}$. Hence, the estimator of the variance components is defined as

$$\hat{\sigma} = \underset{\sigma \in \mathbb{R}^p}{\operatorname{argmax}} \ \ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\sigma)) \tag{3.5}$$

in which, assuming a normal distribution, $\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\sigma))$ is given by

$$\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\sigma)) = -\frac{1}{2}\left[(m-n)\ln(2\pi) + \ln(\det(\mathbf{Q}_{\mathbf{t}})) + \underline{\mathbf{t}}^T\mathbf{Q}_{\mathbf{t}}^{-1}\underline{\mathbf{t}}\right] \tag{3.6}$$

with $\mathbf{Q}_{\mathbf{t}} = \mathbf{B}^T\mathbf{Q}_{\mathbf{y}}(\sigma)\mathbf{B}$ and $\underline{\mathbf{t}}$ is obtained from $\underline{\mathbf{y}}$ using Equation (2.44).

As for the MLE, the RMLE does not provide a direct expression for the variance components, and the assessment of $\hat{\sigma}$ also relies on the use of a nonlinear optimisation method.

3.4 Least-Squares Variance Component Estimation (LS-VCE)

The Least-Squares Variance Component Estimation method was first introduced in 1988 by Teunissen (1988), and further developed in Amiri-Simkooei (2007) and Teunissen and Amiri-Simkooei (2008).

The method consists in reformulating the variance component estimation problem into a linear system of observation equations identical to that defined in chapter 2. The LS-VCE method then applies the theory of weighted least-squares estimation to provide *unbiased* estimators of the variance components, without making any distribution assumptions. But, under the additional assumption of elliptically contoured distributions, such as the multivariate normal distribution, the LS-VCE method provides *unbiased* and *minimum-variance* estimators of the variance components.

The LS-VCE method also provides a direct derivation of the *covariance matrix of the variance components* as well as an *iterative algorithm* for the computation of the variance components, which will both reveal to be extremely useful for our investigations.

3.4.1 Linear stochastic model

So far, we did not make any assumption about the linear or the nonlinear nature of the stochastic model. However, in its original form, the LS-VCE method considers a linear functional model

$$\mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x}, \tag{3.7}$$

and a linear stochastic model such that

$$\mathsf{var}\{\underline{\mathbf{y}}\} = \mathbf{Q}_{\mathbf{y}} = \mathbf{Q}_0 + \sum_{i=1}^{p} \sigma_i \mathbf{Q}_i. \tag{3.8}$$

in which the $m \times m$ matrices $\mathbf{Q}_0$ and $\mathbf{Q}_i$ are the so-called co-factor matrices. The co-factor matrices correspond to the known components of the stochastic model. They

usually represent the covariances of groups of observations or the covariances of specific stochastic processes known up to a scaling factor.

The overall method used to transform the stochastic model (3.8) into a system of observation equations relies on the properties of the misclosure vector $\underline{\mathbf{t}} = \mathbf{B}^T\underline{\mathbf{y}}$ presented in chapter 2, namely

$$\mathsf{E}\{\underline{\mathbf{t}}\} = \mathbf{0}, \tag{3.9}$$

and

$$\mathsf{var}\{\underline{\mathbf{t}}\} = \mathbf{Q}_\mathbf{t} = \mathsf{E}\{\underline{\mathbf{t}}\underline{\mathbf{t}}^T\} = \mathbf{B}^T\mathbf{Q}_\mathbf{y}\mathbf{B}. \tag{3.10}$$

Since the vector $\underline{\mathbf{y}}$ is observable, so is the vector $\underline{\mathbf{t}}$, and the $b \times b$ matrix $\underline{\mathbf{t}}\underline{\mathbf{t}}^T$ can be viewed as an observation of the covariance matrix $\mathbf{Q}_\mathbf{t}$. Hence, by substituting Equation (3.8) into Equation (3.10), the matrix $\underline{\mathbf{t}}\underline{\mathbf{t}}^T$ can be directly linked to the unknown variance components σ_i as

$$\mathsf{E}\{\underline{\mathbf{t}}\underline{\mathbf{t}}^T\} = \mathbf{B}^T\mathbf{Q}_0\mathbf{B} + \sum_{i=1}^{p}\sigma_i\mathbf{B}^T\mathbf{Q}_i\mathbf{B}, \tag{3.11}$$

or, by moving the known component $\mathbf{B}^T\mathbf{Q}_0\mathbf{B}$ on the left hand-side of Equation (3.11) one obtains

$$\mathsf{E}\{\underline{\mathbf{t}}\underline{\mathbf{t}}^T - \mathbf{B}^T\mathbf{Q}_0\mathbf{B}\} = \sum_{i=1}^{p}\sigma_i\mathbf{B}^T\mathbf{Q}_i\mathbf{B}. \tag{3.12}$$

In chapter 2, the functional model of observation equations (Equation 2.12) defines the expectation of the observation vector as a linear combination of unknown functional parameters x_i with known vectors $\mathbf{a}_i$ corresponding to the columns of the design matrix $\mathbf{A}$. In this chapter, Equation (3.12) similarly defines the expectation the observation matrix $\underline{\mathbf{t}}\underline{\mathbf{t}}^T - \mathbf{B}^T\mathbf{Q}_0\mathbf{B}$ as a linear combination of unknown stochastic parameters σ_i with known matrices $\mathbf{B}^T\mathbf{Q}_i\mathbf{B}$.

3.4.2 Vectorizing the problem

The matrices $\underline{\mathbf{t}}\underline{\mathbf{t}}^T$, $\mathbf{B}^T\mathbf{Q}_0\mathbf{B}$, and $\mathbf{B}^T\mathbf{Q}_i\mathbf{B}$ being symmetric, one can consider that all their non-redundant entries are concentrated in their lower-triangular part. Hence, the strategy adopted in Teunissen and Amiri-Simkooei (2008) is to vectorize the problem using the so-called *half-vectorization* operator, noted $\mathsf{vh}\{.\}$, which only vectorizes the lower-triangular part of each matrix. For instance, given a symmetric $b \times b$ matrix $\mathbf{M}$, $\mathsf{vh}\{\mathbf{M}\}$ returns a $\frac{1}{2}b(b+1) \times 1$ vector such that

$$\mathsf{vh}\{\mathbf{M}\} = \mathsf{vh}\left\{\begin{bmatrix} m_{11} & m_{12} & \dots & m_{1b} \\ m_{12} & m_{22} & \dots & m_{2b} \\ \vdots & \vdots & \ddots & \vdots \\ m_{1b} & m_{2b} & \dots & m_{bb} \end{bmatrix}\right\} = \begin{bmatrix} m_{11} \\ m_{12} \\ \vdots \\ m_{1b} \\ m_{22} \\ \vdots \\ m_{2b} \\ \vdots \\ m_{bb} \end{bmatrix} \tag{3.13}$$

The half-vectorization $\mathsf{vh}\{.\}$ operator depicted in Equation (3.13) is linear. Hence, when applied to Equation (3.12), it follows that

$$\mathsf{E}\{\mathsf{vh}\{\underline{\mathbf{t}}\underline{\mathbf{t}}^T - \mathbf{B}^T\mathbf{Q}_0\mathbf{B}\}\} = \sum_{i=1}^{p} \sigma_i \mathsf{vh}\{\mathbf{B}^T\mathbf{Q}_i\mathbf{B}\}, \tag{3.14}$$

which, by noting the $\frac{1}{2}b(b+1) \times 1$ observation vector $\mathbf{y}_{\mathsf{vh}}$ as

$$\underline{\mathbf{y}}_{\mathsf{vh}} = \mathsf{vh}\{\underline{\mathbf{t}}\underline{\mathbf{t}}^T - \mathbf{B}^T\mathbf{Q}_0\mathbf{B}\}, \tag{3.15}$$

and the $\frac{1}{2}b(b+1) \times p$ design matrix $\mathbf{A}_{\mathsf{vh}}$ as

$$\mathbf{A}_{\mathsf{vh}} = \begin{bmatrix} \mathsf{vh}\{\mathbf{B}^T\mathbf{Q}_1\mathbf{B}\} & \cdots & \mathsf{vh}\{\mathbf{B}^T\mathbf{Q}_p\mathbf{B}\} \end{bmatrix}, \tag{3.16}$$

gives a more familiar system of observation equations

$$\boxed{\mathsf{E}\{\underline{\mathbf{y}}_{\mathsf{vh}}\} = \mathbf{A}_{\mathsf{vh}}\sigma.} \tag{3.17}$$

Thanks to this vectorization of the VCE problem, it becomes possible to apply the body of knowledge of weighted least-squares estimation, presented in chapter 2, to derive an unbiased and minimum-variance estimator $\hat{\sigma}$ of the variance components.

3.4.3 Best Linear Unbiased Estimator

Searching for the weighted least-squares solution to the linear system in Equation (3.17) leads to the formulation of the following normal equations system

$$(\mathbf{A}_{\mathsf{vh}}^T\mathbf{W}_{\mathsf{vh}}\mathbf{A}_{\mathsf{vh}})\underline{\hat{\sigma}} = \mathbf{A}_{\mathsf{vh}}^T\mathbf{W}_{\mathsf{vh}}\underline{\mathbf{y}}_{\mathsf{vh}} \tag{3.18}$$

in which $\mathbf{W}_{\mathsf{vh}}$ can be virtually any $\frac{1}{2}b(b+1) \times \frac{1}{2}b(b+1)$ symmetric and positive-definite weight matrix.

Noting $\mathbf{N}_{\mathsf{vh}} = (\mathbf{A}_{\mathsf{vh}}^T\mathbf{W}_{\mathsf{vh}}\mathbf{A}_{\mathsf{vh}})$ and $\underline{\mathbf{l}}_{\mathsf{vh}} = \mathbf{A}_{\mathsf{vh}}^T\mathbf{W}_{\mathsf{vh}}\underline{\mathbf{y}}_{\mathsf{vh}}$, Equation (3.18) becomes

$$\mathbf{N}_{\mathsf{vh}}\underline{\hat{\sigma}} = \underline{\mathbf{l}}_{\mathsf{vh}}. \tag{3.19}$$

If the p matrices $\mathbf{B}^T\mathbf{Q}_1\mathbf{B}, \cdots, \mathbf{B}^T\mathbf{Q}_p\mathbf{B}$ are linearly independent, the design matrix $\mathbf{A}_{\mathsf{vh}}$ is of full column rank and the normal matrix $\mathbf{N}_{\mathsf{vh}}$ is invertible. Thus, a necessary but *not sufficient* condition to the invertibility of $\mathbf{N}_{\mathsf{vh}}$ is that the p co-factor matrices $\mathbf{Q}_1, \cdots, \mathbf{Q}_p$ are linearly independent. This condition is not sufficient because, through the span of the matrix $\mathbf{B}^T$, the design matrix $\mathbf{A}$ influences the estimability of variance components.

Assuming that $\mathbf{N}_{\mathsf{vh}}$ is invertible, an *unbiased* estimator of the variance components $\underline{\hat{\sigma}}$ can be obtain as $\underline{\hat{\sigma}} = \mathbf{N}_{\mathsf{vh}}^{-1}\underline{\mathbf{l}}_{\mathsf{vh}}$, that is

$$\underline{\hat{\sigma}} = (\mathbf{A}_{\mathsf{vh}}^T\mathbf{W}_{\mathsf{vh}}\mathbf{A}_{\mathsf{vh}})^{-1}\mathbf{A}_{\mathsf{vh}}^T\mathbf{W}_{\mathsf{vh}}\underline{\mathbf{y}}_{\mathsf{vh}}. \tag{3.20}$$

Hence, noting $\mathsf{var}\{\mathbf{y}_{\mathsf{vh}}\} = \mathbf{Q}_{\mathsf{vh}}$, the unbiased and *minimum-variance* estimator $\hat{\sigma}$ of σ is obtained by taking $\mathbf{W}_{\mathsf{vh}} = \mathbf{Q}_{\mathsf{vh}}^{-1}$ so that

$$\boxed{\underline{\hat{\sigma}} = (\mathbf{A}_{\mathsf{vh}}^T\mathbf{Q}_{\mathsf{vh}}^{-1}\mathbf{A}_{\mathsf{vh}})^{-1}\mathbf{A}_{\mathsf{vh}}^T\mathbf{Q}_{\mathsf{vh}}^{-1}\underline{\mathbf{y}}_{\mathsf{vh}}} \tag{3.21}$$

in which the weight matrix $\mathbf{Q}_{\mathsf{vh}}^{-1}$ can be directly computed from $\mathbf{Q}_{\mathbf{t}}^{-1}$ using

$$\mathbf{Q}_{\mathsf{vh}}^{-1} = \frac{1}{2}\mathbf{D}^T(\mathbf{Q}_{\mathbf{t}}^{-1} \otimes \mathbf{Q}_{\mathbf{t}}^{-1})\mathbf{D} \tag{3.22}$$

where $\{.\} \otimes \{.\}$ denotes the Kroenecker product, and $\mathbf{D}$ denotes the so-called duplication matrix.

Written out in full, Equation (3.21) reads

$$\boxed{\begin{bmatrix} \underline{\hat{\sigma}_1} \\ \underline{\hat{\sigma}_2} \\ \vdots \\ \underline{\hat{\sigma}_p} \end{bmatrix} = \begin{bmatrix} n_{11} & n_{12} & \dots & n_{1p} \\ n_{12} & n_{22} & \dots & n_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ n_{1p} & n_{2p} & \dots & n_{pp} \end{bmatrix}^{-1} \begin{bmatrix} \underline{l_1} \\ \underline{l_2} \\ \vdots \\ \underline{l_p} \end{bmatrix}} \tag{3.23}$$

where the entries n_{ij} and $\underline{l_i}$ are given by

$$n_{ij} = \mathsf{vh}\{\mathbf{B}^T\mathbf{Q}_i\mathbf{B}\}\mathbf{Q}_{\mathsf{vh}}^{-1}\mathsf{vh}\{\mathbf{B}^T\mathbf{Q}_j\mathbf{B}\}, \tag{3.24}$$

and

$$\underline{l_i} = \mathsf{vh}\{\mathbf{B}^T\mathbf{Q}_i\mathbf{B}\}\mathbf{Q}_{\mathsf{vh}}^{-1}\mathsf{vh}\{\underline{\mathbf{t}}\underline{\mathbf{t}}^T - \mathbf{B}^T\mathbf{Q}_0\mathbf{B}\}. \tag{3.25}$$

In Equations (3.24) and (3.25), the entries n_{ij} and l_i are defined as inner products between large vectors with the $\mathbf{Q}_{\mathsf{vh}}^{-1}$-metric, which is unpractical. Using the properties of the $\mathsf{vh}\{.\}$ operator, Amiri-Simkooei (2007) and Teunissen and Amiri-Simkooei (2008) provided more convenient formulations for n_{ij} and l_i.

In the case of a linear stochastic model and a functional model of observation equations, alternative formulations for n_{ij} and l_i are given by

$$\boxed{n_{ij} = \frac{1}{2}\mathrm{tr}(\mathbf{Q}_i\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{Q}_j\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{P}_{\mathbf{A}}^{\perp}).} \tag{3.26}$$

and

$$\boxed{\underline{l_i} = \frac{1}{2}\underline{\hat{\mathbf{e}}}^T\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{Q}_i\mathbf{Q}_{\mathbf{y}}^{-1}\underline{\hat{\mathbf{e}}} - \frac{1}{2}\mathrm{tr}(\mathbf{Q}_i\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{Q}_0\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{P}_{\mathbf{A}}^{\perp}).} \tag{3.27}$$

where $\underline{\hat{\mathbf{e}}}$ is obtained as $\underline{\hat{\mathbf{e}}} = \mathbf{P}_{\mathbf{A}}^{\perp}\underline{\mathbf{y}}$ with $\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1}$.

Note that the entries n_{ij} and l_i are defined using the unknown covariance matrix $\mathbf{Q}_{\mathbf{y}}^{-1}$. In other words, Equation (3.21) defines the estimator of σ as a function of σ itself. Such type of circular problems can be solved numerically through an iterative procedure.

3.4.4 Iterative algorithm

To solve Equation (3.23), (Amiri-Simkooei, 2007) proposes the use of the Newton-Raphson algorithm. This optimization algorithm starts with initial values for the variance component vector, noted $\sigma^{(0)}$

$$\boldsymbol{\sigma}^{(0)} = \begin{bmatrix} \sigma_1^{(0)} \\ \sigma_2^{(0)} \\ \vdots \\ \sigma_p^{(0)} \end{bmatrix} \tag{3.28}$$

which, depending on the study, can be obtained from a simple guess, from sensor specifications, from previous publications, or from approximate VCE methods (Amiri-Simkooei, 2007).

The initial variance component vector $\boldsymbol{\sigma}^{(0)}$ is used to solve Equation (3.21) and obtain an improved estimate for the variance component vector, noted $\boldsymbol{\sigma}^{(1)}$. Then, by iterating this procedure, each time a variance component $\boldsymbol{\sigma}^{(k)}$ is estimated, a more precise estimate $\boldsymbol{\sigma}^{(k+1)}$ can be obtained. The iterations stop when the difference between two successive variance component estimates becomes negligible.

In practice, the iterative determination of variance components involves a large number of algebraic operations. At each iteration, the most computationally expensive operation is the inversion of the covariance matrix. This inversion can only be optimized in particular cases. Another expensive operation in the iterative process is the evaluation of n_{ij} and l_i from Equations (3.24) and (3.25), which involves multiplying possibly large matrices.

To reduce the computational burden of each iteration without loss in generality, in Algorithm (1), we propose an implementation of LS-VCE different from (Amiri-Simkooei, 2007; Amiri-Simkooei, 2020) to avoid several unnecessary matrix multiplications, and we optimize the evaluation of the trace in Equations (3.24) and (3.25).

This optimized algorithm evaluates the vector $\mathbf{Q}_\mathbf{y}^{-1}\hat{\mathbf{e}}$, the matrix $\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp$, and the $p+1$ matrices $\mathbf{Q}_0\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp$, $\mathbf{Q}_1\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp$, $\cdots$, $\mathbf{Q}_p\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp$ prior to the computation of n_{ij} and l_i. Then the trace in Equations (3.26) and (3.27) is computed from those $p+1$ matrices using the following relation

$$\operatorname{tr}(\mathbf{UV}) = \sum_{k=1}^{m}\sum_{l=1}^{m} u_{kl}v_{lk}, \tag{3.29}$$

which avoids large matrix multiplications.

We illustrate the gain in speed resulting from this optimization in a numerical example at the end of this chapter.

3.4.5 Uncertainty of the variance components

When the iterative algorithm converges to a final variance component estimate $\hat{\boldsymbol{\sigma}}$, its covariance matrix $\mathbf{Q}_{\hat{\sigma}}$ can be obtained by inversion of the normal matrix $\mathbf{N}_{\mathsf{vh}}$, that is

$$\boxed{\mathbf{Q}_{\hat{\sigma}} = \mathbf{N}_{\mathsf{vh}}^{-1}.} \tag{3.30}$$

The i-th diagonal elements of $\mathbf{Q}_{\hat{\sigma}}$ correspond to the variance of the i-th variance components, noted $\sigma^2_{\sigma_i}$. It follows that $\sigma_{\sigma_i} = \sqrt{\sigma^2_{\sigma_i}}$ is the standard deviation of the i-th variance component.

Input

1. Observation vector $\mathbf{y}$
2. Design matrix $\mathbf{A}$
3. Co-factor matrices $\mathbf{Q}_0, \mathbf{Q}_1, \dots \mathbf{Q}_p$
4. Approximate variance components vector $\boldsymbol{\sigma}^{(0)}$
5. Small convergence criterion $\epsilon \sim 10^{-4}$
6. Stop $\leftarrow$ False

while *Stop = False* **do**

Evaluate the covariance matrix $\mathbf{Q}_\mathbf{y} = \mathbf{Q}_0 + \sum_{i=1}^{p} \sigma_i \mathbf{Q}_i$
Evaluate the weight matrix as $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-1}$
Evaluate projector matrix as $\mathbf{P}_\mathbf{A}^{\perp} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}$
Evaluate normalized projector as $\mathbf{K} = \mathbf{W}\mathbf{P}_\mathbf{A}^{\perp}$
Evaluate normalized residual vector as $\hat{\mathbf{q}} = \mathbf{K}\mathbf{y}$
Evaluate the $p+1$ matrices $\mathbf{R}_i = \mathbf{Q}_i\mathbf{K}$
Evaluate the entries n_{ij} and l_i of $\mathbf{N}_{\mathsf{vh}}$ and $\mathbf{l}_{\mathsf{vh}}$ as
for $i \leftarrow 1$ **to** p **do**
  $l_i \leftarrow \frac{1}{2}\hat{\mathbf{q}}^T\mathbf{Q}_i\hat{\mathbf{q}} - \frac{1}{2}\mathrm{tr}(\mathbf{R}_i\mathbf{R}_0)$
  for $j \leftarrow i$ **to** p **do**
    $v \leftarrow \frac{1}{2}\mathrm{tr}(\mathbf{R}_i\mathbf{R}_j)$
    $n_{ij} \leftarrow v$
    $n_{ji} \leftarrow v$
  end
end
Solve for $\hat{\boldsymbol{\sigma}}$ the system $\mathbf{N}_{\mathsf{vh}}\boldsymbol{\sigma} = \mathbf{l}_{\mathsf{vh}}$
if $(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_0)^T\mathbf{N}_{\mathsf{vh}}(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_0) \geq \epsilon$ **then**
  $\boldsymbol{\sigma}_0 \leftarrow \hat{\boldsymbol{\sigma}}$
else
  Evaluate $\mathbf{Q}_{\hat{\sigma}} = \mathbf{N}_{\mathsf{vh}}^{-1}$
  Stop $\leftarrow$ True
end

end

Result: $\hat{\boldsymbol{\sigma}}$ and $\mathbf{Q}_{\hat{\sigma}}$

Algorithm 1: Algorithm for the iterative estimation of $\hat{\boldsymbol{\sigma}}$ and $\mathbf{Q}_{\hat{\sigma}}$ for linear stochastic models. Each evaluation of the trace is done using Equation (3.29).

Although the observations are normally distributed, due to the specific structure of $\underline{\mathbf{y}}_{\text{vh}}$, the estimator of the variance components is not normally distributed. According Equations (3.28) and (3.29) in Teunissen (1988), the estimator of a variance component $\hat{\underline{\sigma}}_i$ is distributed as

$$\hat{\underline{\sigma}}_i \sim \sum_{k=1}^{r} \lambda_k \chi^2_k(1,0), \tag{3.31}$$

where $\chi^2_k(1,0)$ denote mutually independent chi-squared variables, and r is the number of positive eigenvalues λ_k of the following matrix

$$\mathbf{Q}_{\mathbf{y}}^{-1} \left(\frac{1}{2} \sum_{j=1}^{p} \sigma_{\sigma_i} \sigma_{\sigma_j} \mathbf{Q}_j \right) \mathbf{Q}_{\mathbf{y}}^{-1} \mathbf{Q}_{\hat{\mathbf{e}}}. \tag{3.32}$$

Except for very specific cases, such as $p = 1$, the probability density function of $\hat{\underline{\sigma}}_i$ does not have a closed form expression. Therefore, in this thesis, we illustrate the distribution of the variance component estimates using the normal distribution approximation.

3.4.6 Nonlinear stochastic model

We now consider that the stochastic model $\mathbf{Q}_{\mathbf{y}}(\boldsymbol{\sigma})$ is a nonlinear function of the variance components vector $\boldsymbol{\sigma}$. Just like the classical weighted least-squares estimation method, the LS-VCE method can handle nonlinear stochastic model models by linearizing the problem around the approximate variance components $\boldsymbol{\sigma}^{(0)}$ using the first-order Taylor series expansion, so that (Amiri-Simkooei, 2007)

$$\mathbf{Q}_{\mathbf{y}}(\boldsymbol{\sigma}) \approx \mathbf{Q}_{\mathbf{y}}(\boldsymbol{\sigma}^{(0)}) - \sum_{i=1}^{p} \frac{\partial \mathbf{Q}_{\mathbf{y}}(\boldsymbol{\sigma})}{\partial \sigma_i}\bigg|_{\boldsymbol{\sigma}^{(0)}} \sigma_i^{(0)} + \sum_{i=1}^{p} \frac{\partial \mathbf{Q}_{\mathbf{y}}(\boldsymbol{\sigma})}{\partial \sigma_i}\bigg|_{\boldsymbol{\sigma}^{(0)}} \sigma_i \tag{3.33}$$

A more familiar linear stochastic model of the form

$$\mathbf{Q}_{\mathbf{y}}(\boldsymbol{\sigma}) \approx \mathbf{Q}_0^{(k)} + \sum_{i=1}^{p} \sigma_i \mathbf{Q}_i^{(k)} \tag{3.34}$$

can be obtained, for each iteration (k), by noting

$$\mathbf{Q}_0^{(k)} = \mathbf{Q}_{\mathbf{y}}(\hat{\boldsymbol{\sigma}}^{(k)}) - \sum_{i=1}^{p} \frac{\partial \mathbf{Q}_{\mathbf{y}}(\boldsymbol{\sigma})}{\partial \sigma_i} \hat{\sigma}_i^{(k)}\bigg|_{\hat{\boldsymbol{\sigma}}^{(k)}} \tag{3.35}$$

and

$$\mathbf{Q}_i^{(k)} = \frac{\partial \mathbf{Q}_{\mathbf{y}}(\boldsymbol{\sigma})}{\partial \sigma_i}\bigg|_{\hat{\boldsymbol{\sigma}}^{(k)}}, \tag{3.36}$$

in which $\mathbf{Q}_0^{(k)}$ and $\mathbf{Q}_i^{(k)}$ now depends on the variance components estimates.

The non-linearity of the stochastic model affects the iterative procedure used to obtain variance component estimates. Hence, in Algorithm (2), we present a modified algorithm suited for nonlinear stochastic models.

Input

1. Observation vector $\mathbf{y}$
2. Design matrix $\mathbf{A}$
3. covariance function $\mathbf{Q}_\mathbf{y}(\boldsymbol{\sigma})$ and its p partial derivatives $\frac{\partial \mathbf{Q}_\mathbf{y}(\boldsymbol{\sigma})}{\partial \sigma_i}$
4. Approximate variance components vector $\boldsymbol{\sigma}^{(0)}$
5. Small convergence criterion $\epsilon \sim 10^{-3}$
6. Stop $\leftarrow$ False
7. $k \leftarrow 0$

while *Stop* = *False* **do**
  Evaluate $\mathbf{Q}_0^{(k)}$ and $\mathbf{Q}_i^{(k)}$ from Equations (3.35) and (3.36)
  Evaluate the covariance matrix $\mathbf{Q}_\mathbf{y} = \mathbf{Q}_0^{(k)} + \sum_{i=1}^{p} \hat{\sigma}_i^{(k)} \mathbf{Q}_i^{(k)}$
  Evaluate the weight matrix as $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-1}$
  Evaluate projector matrix as $\mathbf{P}_\mathbf{A}^{\perp} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}$
  Evaluate normalized projector as $\mathbf{K} = \mathbf{W}\mathbf{P}_\mathbf{A}^{\perp}$
  Evaluate normalized residual vector as $\hat{\mathbf{q}} = \mathbf{K}\mathbf{y}$
  Evaluate the $p+1$ matrices $\mathbf{R}_i = \mathbf{Q}_i^{(k)}\mathbf{K}$
  Evaluate the entries n_{ij} and l_i of $\mathbf{N}_{\mathsf{vh}}$ and $\mathbf{l}_{\mathsf{vh}}$ as
  for $i \leftarrow 1$ **to** p **do**
    $l_i \leftarrow \frac{1}{2}\hat{\mathbf{q}}^T\mathbf{Q}_i^{(k)}\hat{\mathbf{q}} - \frac{1}{2}\mathrm{tr}(\mathbf{R}_i\mathbf{R}_0)$
    for $j \leftarrow i$ **to** p **do**
      $v \leftarrow \frac{1}{2}\mathrm{tr}(\mathbf{R}_i\mathbf{R}_j)$
      $n_{ij} \leftarrow v$
      $n_{ji} \leftarrow v$
    end
  end
  Solve for $\hat{\boldsymbol{\sigma}}$ the system $\mathbf{N}_{\mathsf{vh}}\boldsymbol{\sigma} = \mathbf{l}_{\mathsf{vh}}$
  if $(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_0)^T\mathbf{N}_{\mathsf{vh}}(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_0) \geq \epsilon$ **then**
    $\boldsymbol{\sigma}^{(k+1)} \leftarrow \hat{\boldsymbol{\sigma}}$
    $k \leftarrow k+1$
  else
    Evaluate $\mathbf{Q}_{\hat{\sigma}} = \mathbf{N}_{\mathsf{vh}}^{-1}$
    Stop $\leftarrow$ True
  end
end
Result: $\hat{\boldsymbol{\sigma}}$ and $\mathbf{Q}_{\hat{\sigma}}$

Algorithm 2: Algorithm for the iterative estimation of $\hat{\boldsymbol{\sigma}}$ and $\mathbf{Q}_{\hat{\sigma}}$ for nonlinear stochastic models. Each evaluation of the trace is done using Equation (3.29).

3.4.7 Possible estimation issues

Divergence

The LS-VCE being iterative, the algorithm used to derive variance component estimates may diverge in some occasions. There are three main possible reasons for this problem. The first and most frequent one is an *incorrect stochastic model*. Intuitively, if one attempts to fit a stochastic model that is not representative of the observations, the LS-VCE has a higher chance to fail at providing estimates. The second reason is a *lack of redundancy*. Some stochastic model requires more observations than others to obtain reliable estimates. The last reason is a *badly chosen set of approximate variance components*. If the approximate values are far from the true values, the algorithm has more difficulty to converge.

Negative estimates

Another occasional problem is the negativity of some variance component estimates. The LS-VCE being unconstrained, it can provide negative values, even when it makes no physical sense. This issue is usually caused by lack of redundancy, which results in an imprecise estimator and random fluctuations of the estimates can reach negative values, especially when the true variance component amplitude is relatively close to zero. Another possibility is that variance component estimates are biased toward negative values. This usually indicates that the stochastic model poorly represents the observations.

To systematically overcome the negativity issue, Amiri-Simkooei (2016) proposes to use the non-negative sequential coordinate-wise algorithm developed by (Franc, Hlaváč, and Navara, 2005) to solve the LS-VCE normal equation system (3.21) at each iteration.

Note that the non-negativity constraint conflicts with the unbiased property of the LS-VCE estimator (Amiri-Simkooei, 2016). Consequently, the use of the non-negative LS-VCE method should ideally be used in the case the unconstrained LS-VCE failed to provide positive estimates.

3.4.8 Numerical examples

To illustrate the application of the LS-VCE method, we consider a synthetic experiment consisting in the estimation of the variance and covariance of two time series.

The two time series are measurements of the *same* linear trend, of intercept a and rate b, obtained from two correlated measurement devices. In this case, the functional model follows

$$\mathsf{E}\left\{\begin{bmatrix}\underline{\mathbf{y}}_1\\ \underline{\mathbf{y}}_2\end{bmatrix}\right\} = \begin{bmatrix}\mathbf{A}\\ \mathbf{A}\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix} \tag{3.37}$$

with

$$\mathbf{A} = \begin{bmatrix}1 & t_1\\ 1 & t_2\\ \vdots & \vdots\\ 1 & t_m\end{bmatrix}. \tag{3.38}$$

in which, t_i is the time of measurement.

The stochastic model of this synthetic experiment can be written as

$$\begin{aligned} \text{var}\left\{\begin{bmatrix}\underline{\mathbf{y}}_1\\ \underline{\mathbf{y}}_2\end{bmatrix}\right\} &= \sigma_1^2\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} + \sigma_2^2\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} + \sigma_1\sigma_2\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{I} & \mathbf{0}\end{bmatrix} \\ &= \sigma_1^2\mathbf{Q}_1 + \sigma_2^2\mathbf{Q}_2 + \sigma_1\sigma_2\mathbf{Q}_3 \end{aligned} \tag{3.39}$$

in which σ_1^2, σ_2^2, and $\sigma_1\sigma_2$ correspond to the variance components and $\mathbf{Q}_1$, $\mathbf{Q}_2$ and $\mathbf{Q}_3$ correspond to the co-factor matrices.

Accuracy of the method

For the simulation, the variances and the covariance are chosen as $\sigma_1^2 = 5$ mm^2, $\sigma_2^2 = 10$ mm^2, and $\sigma_1\sigma_2 = 3$ mm^2. A set of simulated observations following these properties is visible in Figure (3.1). By taking $\sigma_1^{2(0)} = 15$ mm^2, $\sigma_2^{2(0)} = 15$ mm^2, and $\sigma_1\sigma_2{}^{(0)} = 0$ mm^2 as approximate values, we applied the LS-VCE method (1) for two different sample sizes $m_1 = 50$ and $m_2 = 100$, and so, for 5000 simulations.

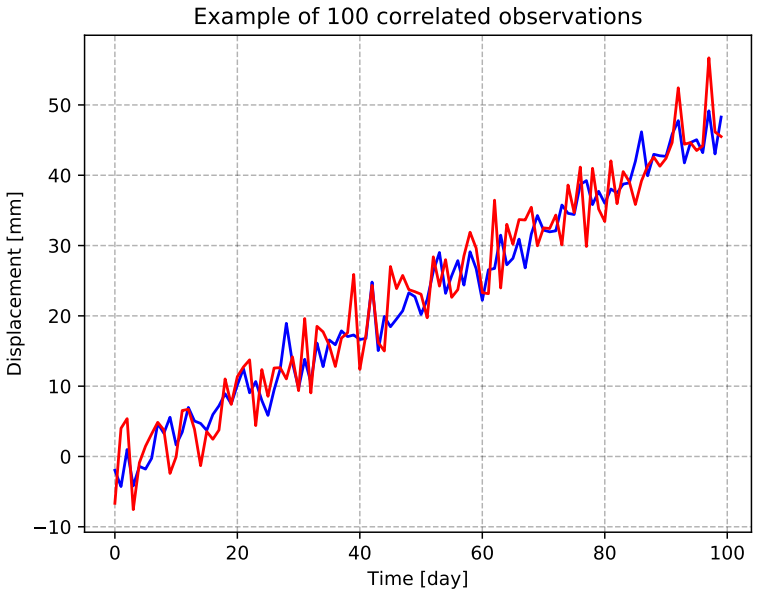

FIGURE 3.1: Example of simulated observations.

The distributions of the estimated variance components $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$ and $\hat{\sigma}_1\hat{\sigma}_2$ are presented in Figure (3.2). The LS-VCE estimates are centred on the true variance components. This results from the unbiased property of the estimator. Note that the uncertainty of the variance component estimator induces a spread of estimates distribution. When increasing the sample size, this spread reduces.

Although the empirical distributions of the variance component estimates are slightly skewed due to the special distribution of the variance component estimates, the normal distributions drawn in red remain good approximations. The standard deviations used to draw the normal probability density functions in Figure (3.2) are

that predicted by the LS-VCE method. Hence, the method also provides reliable uncertainties on its estimates.

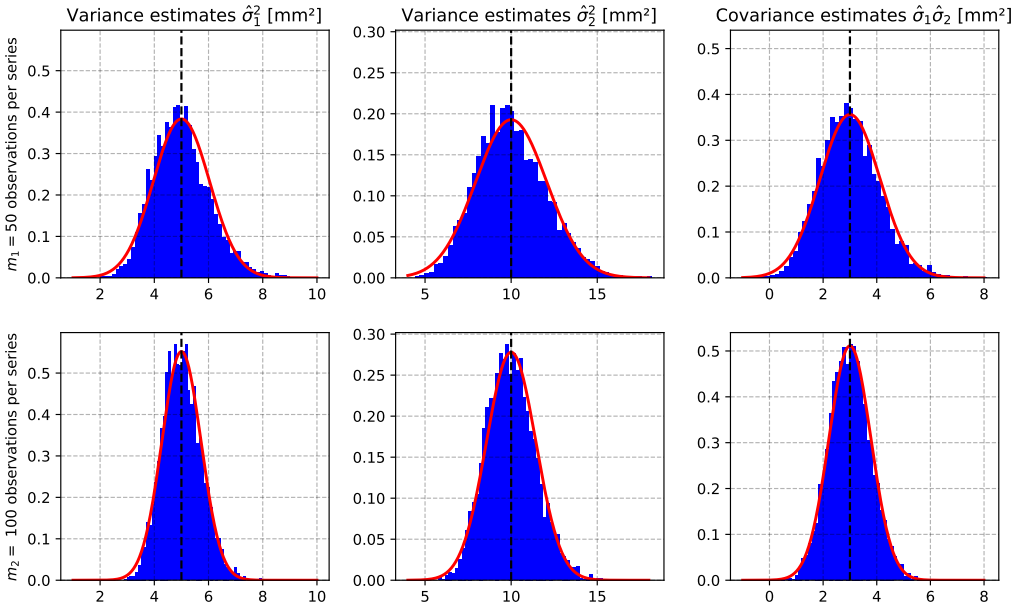

FIGURE 3.2: Normalized distributions of variance component estimates for 5000 simulations (*blue*). True variance components (*black*). Normal distribution approximation using the standard deviation predicted by Equation (3.30). Result for 50 observations per time series (*top*). Result for 100 observations per time series (*bottom*)

Speed of the optimized algorithm

To illustrate the gain in speed of the proposed algorithm over the standard implementation presented in (Amiri-Simkooei, 2007; Amiri-Simkooei, 2020), we tested the evolution of both algorithm runtimes with the number of observation per time series.

The results obtained for a number of observations per time series ranging from 50 to 3000 are presented in Figure (3.3). Our simulation shows that the runtime of the optimized algorithm is more than 3 times faster than that of the standard one for 3000 observation per time series (that is, 6000 observations in total).

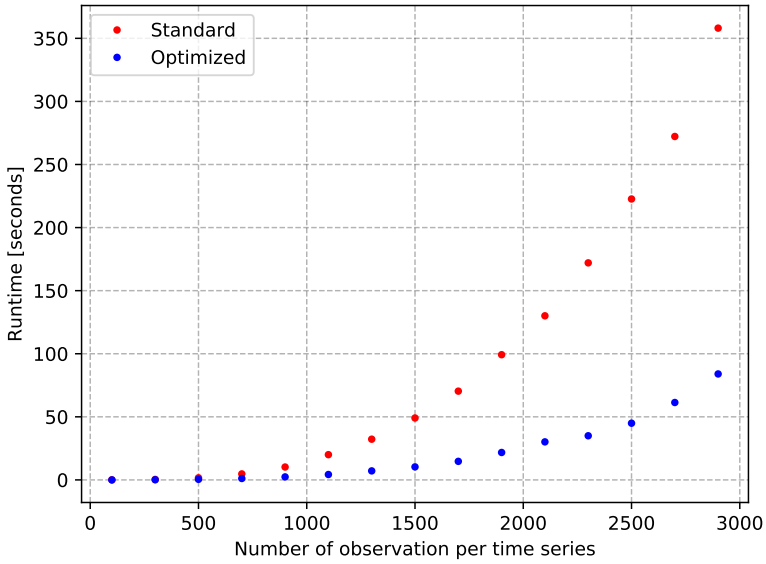

FIGURE 3.3: Evolution of the run-time of the standard and optimized LS-VCE algorithm

Chapter 4

Elements of geodetic time series analysis

Whether analysing ground deformations or sea-levels, it is often necessary to model how a series of observations measured at a specific geographic location evolves as a function of time. In such cases, the statistical analysis comes under the field of time series analysis, which has specific ways to describe both the functional and the stochastic models.

This chapter reviews the basic principles of geodetic time series analysis. In particular, we present parametric functional models and parametric stochastic models frequently used by the geodetic community. Hence, this chapter should also provide a more in-depth description of the statistical models used in chapters 8-11.

The content of this methodological chapter is mostly derived from Bevis and Brown (2014) and Bos et al. (2020).

4.1 Trajectory model

When the considered time variable observations, noted $\underline{y}(t_i)$, correspond to position coordinates in one or more spatial dimensions, the functional model of geodetic time series is often referred to as *trajectory model* (Bevis and Brown, 2014). In most applications, this trajectory model can be defined as a linear combination of elementary behaviours, including a general trend, periodic oscillations, step discontinuities and possible transients.

Hence, in the following, the functional model of the time variable observations, noted $\mathsf{E}\{\underline{y}(t_i)\}$, is defined such that

$$\boxed{\mathsf{E}\{\underline{y}(t_i)\} = \mu_{\text{trend}}(t_i) + \mu_{\text{periodic}}(t_i) + \mu_{\text{discontinuities}}(t_i) + \mu_{\text{transients}}(t_i),} \tag{4.1}$$

in which $\mu_{\text{trend}}(t_i)$, $\mu_{\text{periodic}}(t_i)$, $\mu_{\text{discontinuities}}(t_i)$, and $\mu_{\text{transients}}(t_i)$ are the expected values of each elementary behaviour at time t_i.

4.1.1 Trend model

In practice, the trend $\mu_{\text{trend}}(t_i)$ is modelled using low-degree polynomials. Therefore, $\mu_{\text{trend}}(t_i)$ can be written as

$$\boxed{\mu_{\text{trend}}(t_i) = \alpha + \beta t_i \; (+\frac{\gamma}{2} t_i^2),} \tag{4.2}$$

in which the intercept α, the velocity β, and the acceleration γ are parameters that must be estimated from the observations. The acceleration term is left in parentheses because a very few phenomena actually show accelerations (this is the case, for example, of ice mass changes, not investigated in this study).

4.1.2 Periodic oscillations model

At first approximation, periodic oscillations induced by tides and by climatic processes can be described by a linear combination of pure sine waves, that is

$$\mu_{\text{periodic}}(t_i) = \sum_k A_k \sin(\omega_k t_i + \varphi_k), \tag{4.3}$$

in which A_k denotes the amplitude, ω_k the angular frequency, and φ_k the phase offset. ω_k is related to time frequency f_k as $\omega_k = 2\pi f_k$.

We often know the frequencies of the oscillations f_k from a prior frequency analysis or the physical understanding of the measured process, like the frequencies of tidal constituents measured by tide gauges or the seasonality in surface mass distribution. Hence, the parameters to be estimated from observation are the amplitude A_k and the phase offset φ_k, for each considered frequency f_k. Equation (4.3) depends non-linearly on A_k and φ_k, so their direct estimation requires using a nonlinear estimation method, which is not convenient when dealing with numerous frequencies. Instead, Equation (4.3) can be linearized using the Fourier series decomposition so that

$$\boxed{\mu_{\text{periodic}}(t_i) = \sum_k a_k \cos(\omega_k t_i) + b_k \sin(\omega_k t_i),} \tag{4.4}$$

in which a_k and b_k are referred to as Fourier coefficients.

For each considered frequency, f_k, the amplitude A_k can be computed from the Fourier coefficients a_k and b_k as

$$\boxed{A_k = \sqrt{a_k^2 + b_k^2},} \tag{4.5}$$

and the phase offset φ_k as

$$\boxed{\varphi_k = \arctan\left(\frac{a_k}{b_k}\right).} \tag{4.6}$$

Because Equation (4.4) assumes a constant amplitude A_k and phase offset φ_k, it may be too simple for some processes. In such cases, more advanced methods (not used in this thesis) have been developed (Davis, Wernicke, and Tamisiea, 2012; Klos et al., 2019).

4.1.3 Discontinuity models

When measuring a signal continuously over years or decades, instrumental changes and large earthquakes may introduce discontinuities into the time series (Gazeaux et al., 2013). These discontinuities can be modelled using the so-called Heaviside function $\mathcal{H}(.)$, which models a unit step discontinuity at time t_k as

$$\mathcal{H}(t_i - t_k) = \begin{cases} 0, & t_i - t_k < 0 \\ 1, & t_i - t_k \geq 0 \end{cases}. \tag{4.7}$$

A collection of step discontinuities, or *offsets*, with unknown amplitudes, can be modelled as a linear combination of Heaviside functions, that is

$$\boxed{\mu_{\text{offsets}}(t_i) = \sum_k \zeta_k \mathcal{H}(t_i - t_k),} \tag{4.8}$$

in which ζ_k is the amplitude of an offset occurring at time t_k.

The occurrence time t_k of an offset is usually given by the date of earthquakes, by the instrument metadata, or by automatic change detection algorithms. Therefore, the amplitude parameters ζ_k are the only one to be estimated from the observations, so that the trajectory model remains linear when accounting for offsets.

Occasionally, geodetic time series also exhibit abrupt changes in velocity triggered by geophysical events or human activities in the vicinity of the sensors. In such cases, a set of velocity changes occurring at times t_k can also be modelled using the Heaviside function as

$$\boxed{\mu_{\text{velocities}}(t_i) = \sum_k \eta_k (t_i - t_k) \mathcal{H}(t_i - t_k)} \tag{4.9}$$

in which η_k correspond to the velocity increment at time t_k.

As for the offsets, the occurrence time t_k of each velocity discontinuity is determined prior to the design of the trajectory model, using either metadata or detection algorithms. Therefore, the velocity increments η_k are the only parameters to be estimated from the observations, and the trajectory model remains linear.

When considering both type of discontinuities, the term $\mu_{\text{discontinuities}}(t_i)$ in Equation (4.1) is defined as

$$\boxed{\mu_{\text{discontinuities}}(t_i) = \mu_{\text{offsets}}(t_i) + \mu_{\text{velocities}}(t_i).} \tag{4.10}$$

4.1.4 Transient models

Another phenomenon occasionally present in geophysical time series are transients signals. It occurs, for instance, when the sensor is affected by post-seismic deformations (Bevis and Brown, 2014) or instrumental drifts (Watts and Kontoyiannis, 1990; Van Camp and Francis, 2007).

Such signals are usually modelled using a linear combination of exponential functions (Savage and Prescott, 1978)

$$\boxed{\mu_{\text{exp}}(t_i) = \sum_k \phi_k \left(1 - \exp\left(-\frac{t_i - t_k}{\tau_k}\right)\right)} \tag{4.11}$$

or logarithmic functions (Marone, Scholtz, and Bilham, 1991)

$$\boxed{\mu_{\log}(t_i) = \sum_k \psi_k \ln\left(1 + \frac{t_i - t_k}{\tau_k}\right)} \tag{4.12}$$

in which the parameters ϕ_k and ψ_k are unknown amplitudes and τ_k are the time-constants of the transient signals starting at time t_k.

Sometimes, both $\mu_{\exp}(t_i)$ and $\mu_{\log}(t_i)$ models must be combined to really fit the transient signal observed in some time series (Altamimi et al., 2016) so that

$$\boxed{\mu_{\text{transients}}(t_i) = \mu_{\exp}(t_i) + \mu_{\log}(t_i).} \tag{4.13}$$

In practice, the time-constants τ_k are unknown and must be estimated along with the amplitudes ϕ_k and ψ_k. Thus, the introduction of such transients in the trajectory model makes the functional model *nonlinear*.

4.2 Stochastic processes

In addition to the trajectory model, which only models the part of the observations considered as deterministic, the remaining stochastic variability must be described by a stochastic model, using stochastic processes. In the context of time series analysis, the considered stochastic processes are random time variable functions existing in both continuous and discrete forms. Since we usually deal with regularly sampled observations, with possible gaps, this chapter exclusively focuses on the presentation of *discrete stochastic processes*.

In the following, the distinction is made between *white noise* processes, that is *uncorrelated* noise, and *time correlated* processes. Although white noise processes were used as default hypotheses for the stochastic model, time-correlated processes have shown to be much more realistic when dealing with geophysical time series (Agnew, 1992; Press, 1978). In particular, numerous studies demonstrated that including time possible correlated processes in the stochastic model is essential to obtain realistic functional parameter uncertainties from geodetic time series (Zhang et al., 1997; Williams, 2003a).

In practice, continuous stochastic processes are mathematically defined through their auto-covariance functions, or, equivalently, through their power-spectra. For discrete stochastic processes, auto-covariance functions can be written in the form of covariance matrices. Being able to express the auto-correlation of a process in the form of a covariance matrix is especially important when using a Variance Component Estimation method, as it is the basic element of a stochastic model (see chapter 3). Therefore, besides presenting each process, we also specify the computation of their covariance matrix for practical applications.

4.2.1 White noise

The most simple type of discrete stochastic process is probably the discrete white noise process. Noting $\underline{v}_i$ independent and identically normally distributed errors with unit variance – that is, $\underline{v}_i \sim \mathcal{N}(0,1)$, a white noise process with variance σ^2, noted $\underline{w}_i$, is defined as

$$\underline{w}_i = \sigma \underline{v}_i \tag{4.14}$$

where σ denotes the standard deviation of the process.

The auto-covariance between two samples of a white noise process, noted $\underline{w}_i$ and $\underline{w}_j$ is given by

$$\mathsf{cov}\{\underline{w}_i, \underline{w}_j\} = \begin{cases} \sigma^2 & , \text{ if } i = j \\ 0 & , \text{ if } i \neq j \end{cases}, \tag{4.15}$$

where σ^2 denotes the variance of the process.

Hence, for m observations of a white noise process starting at t_0, the Equation (4.14) can be re-written using matrix notation as

$$\underline{\mathbf{w}} = \sigma \underline{\mathbf{v}} \tag{4.16}$$

noting

$$\underline{\mathbf{w}} = \begin{bmatrix} \underline{w}_0 \\ \underline{w}_1 \\ \vdots \\ \underline{w}_{m-1} \end{bmatrix}, \underline{\mathbf{v}} = \begin{bmatrix} \underline{v}_0 \\ \underline{v}_1 \\ \vdots \\ \underline{v}_{m-1} \end{bmatrix}. \tag{4.17}$$

with $\underline{\mathbf{v}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.

Applying the linear variance propagation law to Equation (4.16), we obtain the covariance matrix of the white noise process $\underline{\mathbf{w}}$, which is simply

$$\mathsf{var}\{\underline{\mathbf{w}}\} = \sigma^2 \mathbf{I} \tag{4.18}$$

with $\mathbf{I}$ the cofactor matrix associated to the white noise process, and σ^2 its amplitude.

4.2.2 Time-correlated processes

The hypothesis that the variability considered as stochastic in geophysical time series can be modelled with only white noise is a strong assumption in general. If incorrect, it ultimately leads to an underestimation of functional parameter uncertainties, and therefore to an over-interpretation of measurements (Williams, 2003a; Bos et al., 2014; Van Camp, Williams, and Francis, 2005). To avoid this problem, it has become a standard practice to consider possible time-correlation of the stochastic component of geophysical observations.

A wide variety of discrete and time-correlated stochastic processes are in use in geodesy. However, their definitions usually come down to filtering a white noise process using a process-specific filter h_k introducing time correlation. In this thesis, all the time-correlated processes, noted $\underline{r}_i$, are defined through the following convolution

$$\underline{r}_i = \sigma \sum_{k=-\infty}^{+\infty} h_k \underline{v}_{i-k}, \tag{4.19}$$

in which, $\underline{v}_{i-k}$ are lagged values of a white noise process with unit variance, h_k is the value of the process-specific filter at lag k, and σ is the amplitude of the time-correlated process.

In general, it is assumed that the stochastic process $\underline{r}_i$ starts at t_0, therefore the maximum lag is i and the summation in Equation (4.19) stops at i. Besides, it is

always assumed that $h_k = 0$ for $k < 0$, and the summation actually starts at 0. Under these assumptions, Equation (4.19) can be re-written as

$$\underline{r}_i = \sigma \sum_{k=0}^{i} h_k \underline{v}_{i-k}, \tag{4.20}$$

The covariance between two samples of a discrete time-correlated process defined from Equation (4.20), noted $\underline{r}_i$ and $\underline{r}_j$, is given by (Bos et al., 2008)

$$\mathsf{cov}\{\underline{r}_i, \underline{r}_j\} = \sigma^2 \sum_{k=0}^{i} h_k h_{k+(j-i)}, \tag{4.21}$$

For m observations of a time-correlated process starting at t_0, the Equation (4.20) can be re-written using matrix notation as

$$\underline{\mathbf{r}} = \sigma \mathbf{L} \underline{\mathbf{v}} \tag{4.22}$$

by noting

$$\underline{\mathbf{r}} = \begin{bmatrix} \underline{r}_0 \\ \underline{r}_1 \\ \vdots \\ \underline{r}_{m-1} \end{bmatrix}, \mathbf{L} = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{m-1} & \cdots & h_1 & h_0 \end{bmatrix}, \underline{\mathbf{v}} = \begin{bmatrix} \underline{v}_0 \\ \underline{v}_1 \\ \vdots \\ \underline{v}_{m-1} \end{bmatrix}. \tag{4.23}$$

with $\underline{\mathbf{v}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.

Applying the linear propagation law of variances to Equation (4.22), one obtains the covariance matrix of the time-correlated process $\underline{\mathbf{r}}$ as

$$\mathsf{var}\{\underline{\mathbf{r}}\} = \sigma^2 \mathbf{Q} = \sigma^2 \mathbf{L}\mathbf{L}^T. \tag{4.24}$$

with $\mathbf{Q}$ the cofactor matrix associated to the considered stochastic processes, and σ^2 its amplitude.

From the definition of the filter h_k, it is possible to generate a wide range of time-correlated stochastic processes. In the following, we present three types of stochastic processes, namely *first-order Gauss-Markov*, *power-law*, and *generalized Gauss-Markov* processes.

First-Order Gauss-Markov (FOGM) processes

As a first example of time-correlated process, we consider the so-called *First-Order Gauss-Markov processes* (FOGM), also called *first-order Auto-Regressive processes* (AR1). These processes are often used to model the time correlation observed in sea-level and climatic time series (Allen and Smith, 1996; Hughes and Williams, 2010; Camargo et al., 2020).

FOGM processes are defined through the convolution of a white noise process with an exponentially decaying function of the time delay k. In their discrete form, FOGM processes can be obtained from Equation (4.19) by defining the filter h_k as

$$\begin{aligned} h_0 &= 1 \\ h_k &= \phi h_{k-1} \end{aligned} \tag{4.25}$$

where ϕ is a parameter defined so that $0 < \phi < 1$.

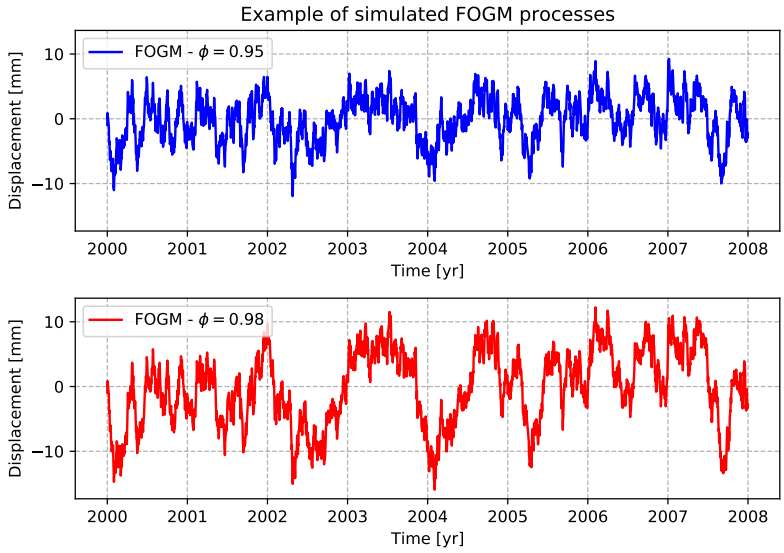

FIGURE 4.1: Example of simulated first-order Gauss-Markov processes generated from the same white noise time series.

To obtain the auto-covariance and the covariance matrix of a FOGM process, one can substitute the filter definition (4.25) in Equations (4.21) and (4.24). Because h_k decay exponentially fast with the time delay k, the FOGM only exhibits *short-memory effects*. Two examples of simulated FOGM processes with the same variance, but with different ϕ, generated using Equation (4.22) are presented in Figure (4.1).

For some numerical application, the use of the unit-less parameter ϕ may be unpractical. Instead, one can transform the parameter ϕ into the time-constant of a continuous FOGM process, noted τ, expressed in years, and defined as

$$\tau = -\frac{\Delta_t}{\ln(\phi)} \tag{4.26}$$

in which Δ_t is the sampling period in years.

To identify the nature of the stochastic process, a common strategy is to analyse the shape of power-spectrum of the time series by comparing it to the theoretical power spectrum of a given process. The theoretical power-spectrum $P(f)$ of a discrete FOGM process sampled at a frequency f_s reads

$$P(f) = 2\frac{\sigma^2}{f_s}[1 + \phi^2 - 2\phi\cos(2\pi f/f_s)]^{-1} \tag{4.27}$$

The theoretical power-spectra of FOGM processes for different values of ϕ are depicted in Figure (4.2). Due to the time correlation, the energy of the process increases at low frequencies, but, as the correlation between two samples decays exponentially fast, the spectrum eventually flattens at very low frequencies. In log-log scales, the power spectrum (4.2), show a slope of -2.0 at high frequencies and a systematic flattening at low frequencies. The crossover frequency of this flattening f_c is controlled by the time-constant τ as

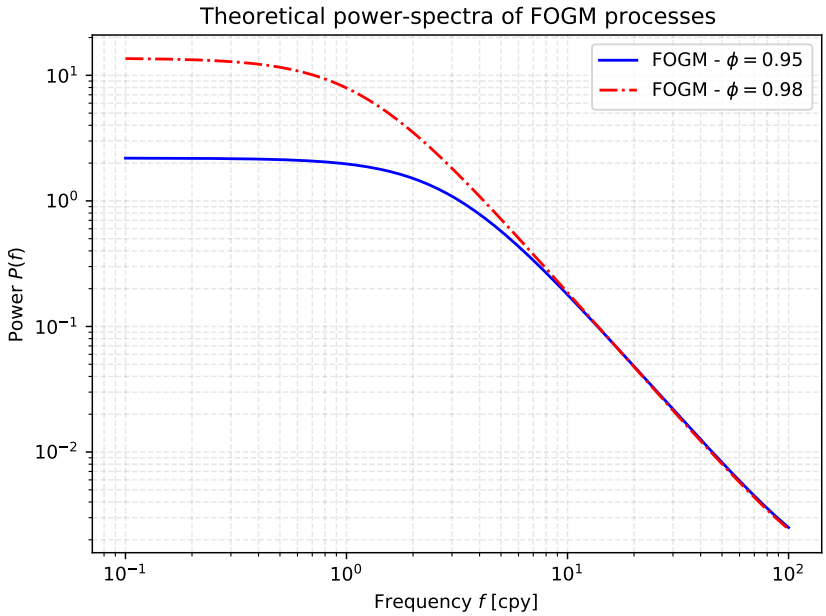

FIGURE 4.2: Theoretical spectra of discrete first-order Gauss-Markov (FOGM) processes.

$$f_c = \frac{1}{2\pi\tau}. \tag{4.28}$$

Note that, for $\phi = 0$, the FOGM process is a simple homogeneous white noise. In this case, the expression power spectrum (4.2) does not depend on the frequency f anymore, and it looks flat in a graphic representation.

FOGM processes depend on two stochastic parameters: the parameter ϕ and the amplitude parameter σ^2. If ϕ is known, σ^2 is the only unknown variance component. In this case the stochastic model is linear, and σ^2 can be estimated using the linear LS-VCE method. However, if ϕ is to be estimated, the stochastic model becomes nonlinear and σ^2 and ϕ must be estimated using the nonlinear LS-VCE method (Khazraei and Amiri-Simkooei, 2019).

Power-Law (PL) processes

Since the work of Hurst (1951) and Hurst (1957), who evidenced a power-law dependence between the cumulative flow of rivers and the observation time, more recent studies – e.g., Press (1978) and Agnew (1992) – showed that many geophysical and astronomical time series present similar power-spectra $P(f)$ following power-law functions of the time-frequency f, that is

$$P(f) \approx P_0 \left(\frac{f}{f_0}\right)^{\kappa}, \tag{4.29}$$

where both P_0 and f_0 are normalization constants, and κ is the so-called spectral index (Agnew, 1992).

This discovery had a significant impact on statistical time series analysis as it implied that some natural phenomenon showed long-term dependencies, or *long*

memory effects, which could not be described by the stochastic processes known at the time. This issue led Mandelbrot and Van Ness (1968) to develop the so-called *fractional Brownian motions*, which formalized the concept of continuous power-law (PL) processes. Later, Granger and Joyeux (1980) and Hosking (1981) developed fractional differencing which provided a statistical framework for the analysis of discrete PL processes.

PL processes are defined through the convolution of a white noise process with a power-law decaying function of the time delay. In their discrete form, PL processes can be obtained from Equation (4.19) by defining the filter h_k as (Kasdin, 1995)

$$\begin{aligned} h_0 &= 1 \\ h_k &= a_k h_{k-1} \end{aligned} \tag{4.30}$$

with

$$a_k = 1 - \frac{\kappa}{2k} - \frac{1}{k}. \tag{4.31}$$

To obtain the auto-covariance and the covariance matrix of a PL process, one can substitute the filter definition (4.30) in Equations (4.21) and (4.24). The difference with the definition of a FOGM process is that a_k is a more slowly decaying function, which introduces a longer memory effect. Three examples of PL time series with different spectral indices simulated using Equation (4.22) are presented in Figure (4.3).

The theoretical power spectrum of a discrete PL process is given by

$$P(f) = \frac{\sigma^2}{f_s}[2\sin(\pi f / f_s)]^{\kappa} \tag{4.32}$$

which, for $f/f_s \to 0$, satisfies the power-law property defined in Equation (4.29).

The theoretical power-spectra of three PL processes with different spectral indices are presented in Figure (4.4). When plotted in log-log scales, the theoretical power-spectrum of a PL process is a linear trend of which the slope is the spectral index κ.

PL processes depend on two stochastic parameters: the spectral index κ and the variance σ^2. If κ is known, σ^2 is the only variance component and the stochastic model is linear. In such a case, σ^2 can be estimated using the linear LS-VCE method. However, if κ is to be estimated, the stochastic model becomes nonlinear. In chapter 8 we discuss the estimation of κ using the nonlinear LS-VCE method.

Generalized Gauss Markov (GGM) processes

In some cases, the geophysical time series exhibits power-law properties at high frequencies and white noise properties at low frequencies (Bos et al., 2014; He et al., 2019). To account for this behaviour, Langbein (2004) introduced Generalized Gauss Markov (GGM) processes, which is a generalization of both FOGM and PL processes.

Generalized Gauss-Markov processes are defined through the convolution of a PL process with an exponentially decaying function of the time delay k. In their

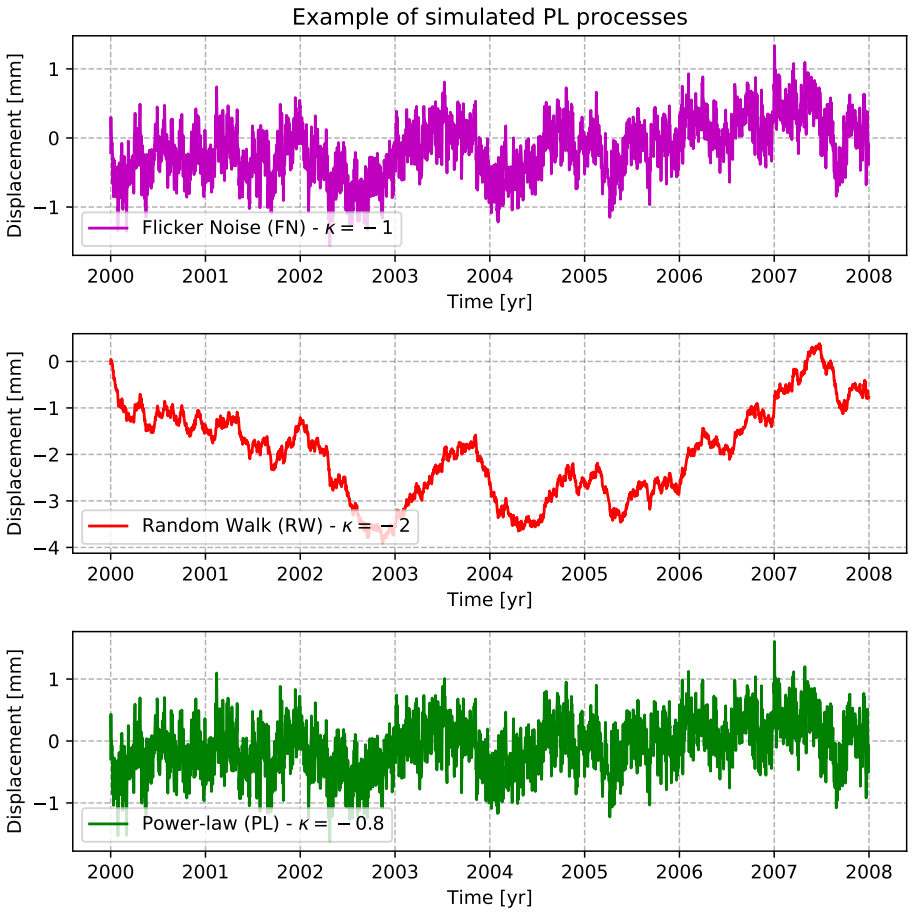

FIGURE 4.3: Example of simulated discrete power-law processes generated from the same white noise time series.

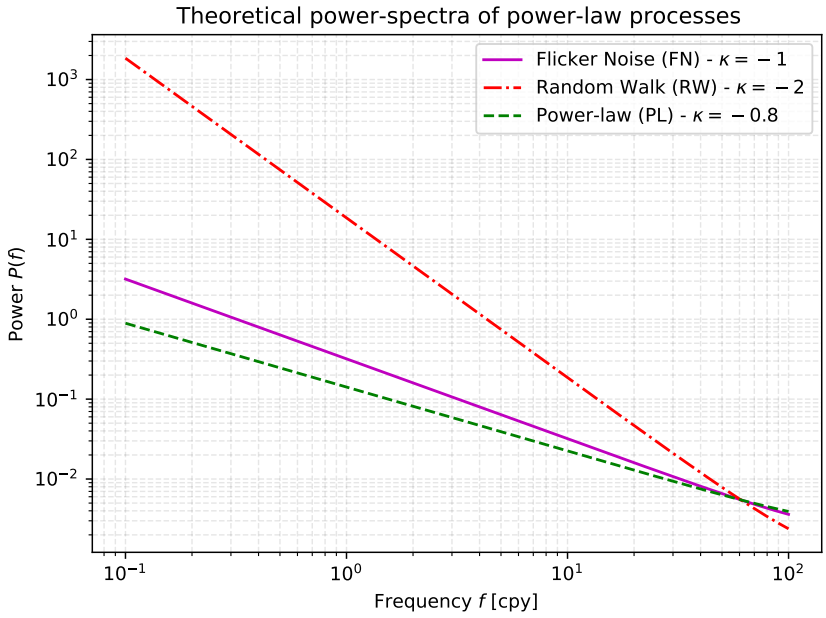

FIGURE 4.4: Theoretical power-spectra of discrete power-law processes.

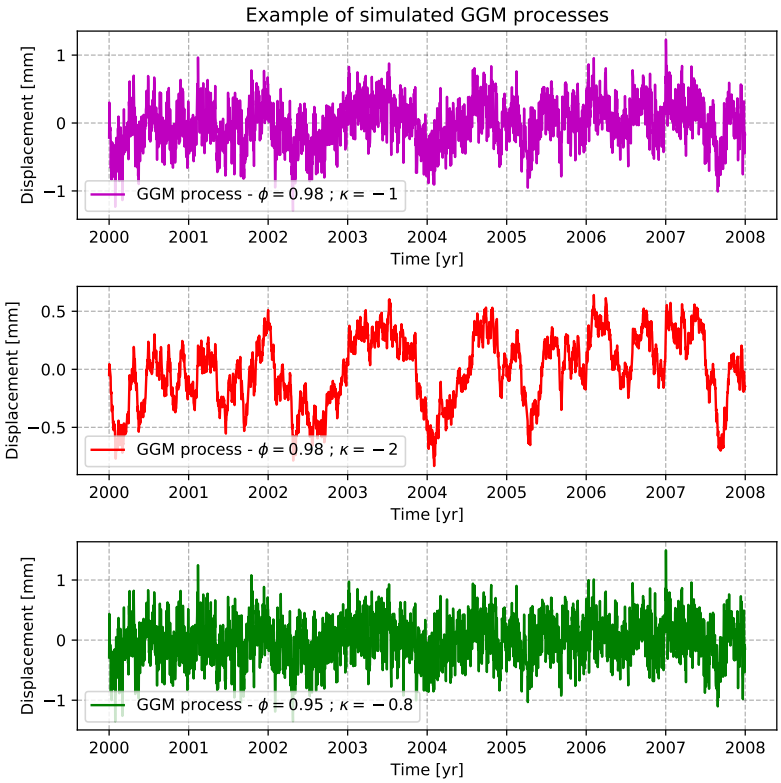

FIGURE 4.5: Example of simulated discrete generalized Gauss Markov processes generated from the same white noise time series.

discrete form, GGM processes can be obtained from Equation (4.19) by defining the filter h_k as (Langbein, 2004; Bos et al., 2020)

$$\begin{aligned} h_0 &= 1 \\ h_k &= \phi a_k h_{k-1} \end{aligned} \tag{4.33}$$

in which ϕ is the parameter of the FOGM process $0 < \phi \leq 1$) and a_k is the same power-law parameter as defined in Equation (4.31).

To obtain the auto-covariance and the covariance matrix of a GGM process, one can substitute the filter definition (4.33) in Equations (4.21) and (4.24). The difference with the definition of PL processes is the influence of the parameter ϕ, which removes the long term dependencies of the power-law. Three examples of GGM time series with different parameters κ and ϕ simulated using Equation (4.22) are presented in Figure (4.5).

The theoretical power-spectrum of a discrete GGM process is given by

$$P(f) = 2\frac{\sigma^2}{f_s}[1 + \phi^2 - 2\phi\cos(2\pi f/f_s)]^{\kappa/2}. \tag{4.34}$$

It follows that, for $\kappa = -2$, the GGM power-spectrum (4.34) becomes that of the FOGM process (4.27). Besides, using trigonometric identities, it can be shown that, when $\phi = 1$, the power-spectrum (4.34) is that of a pure PL process (4.32).

The theoretical power-spectra of three GGM processes with different spectral indices and parameters ϕ are presented in Figure (4.6). When plotted in log-log scales, the theoretical power-spectrum of GGM processes shows a linear trend at high frequencies and a flattening at low frequencies.

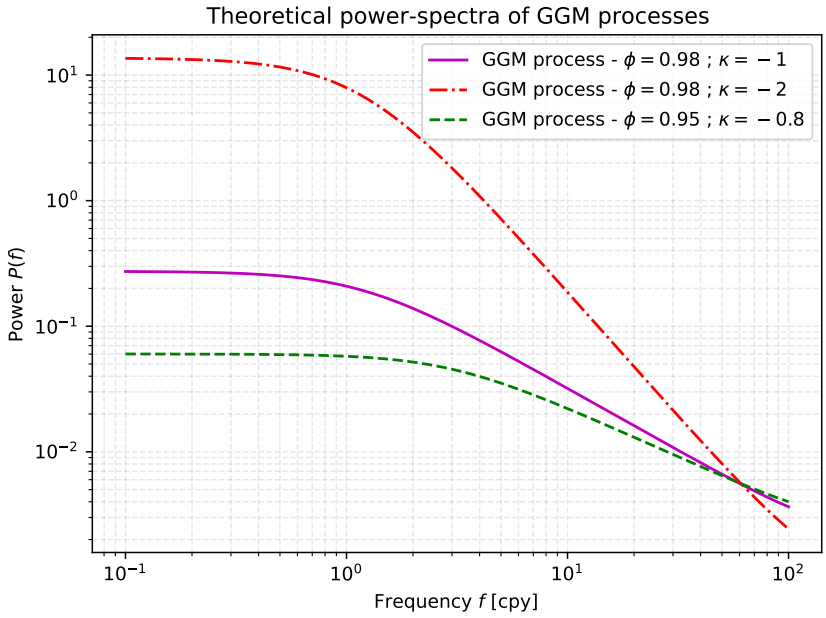

FIGURE 4.6: Theoretical spectra of discrete generalized Gauss Markov processes

GGM processes depend on three stochastic parameters, namely the parameter ϕ, the spectral index κ, and the variance σ^2. If both ϕ and κ are known, σ^2 is the only variance component and the stochastic model is linear. In such a case, σ^2 can be estimated using the linear LS-VCE method. However, if ϕ or κ is to be estimated, the stochastic model becomes nonlinear. In chapter 8, we discuss the estimation of ϕ and κ using nonlinear LS-VCE method.

4.2.3 Combination of stochastic processes

The stochastic variability observed in geodetic time series is sometimes too complex to be realistically modelled using only one stochastic process. In such cases, one can define the stochastic model as a linear combination of several stochastic processes. In particular, it is frequent to combine white noise with one or two types of correlated process. The white noise processes aim at describing high-frequency variability whereas the time-correlated process rather describes the low-frequency variability. When doing so, the stochastic model $\mathsf{var}\{\underline{\mathbf{y}}\}$ becomes a linear combination of covariance matrices of each process, which increases the number of unknown variance components.

For instance, for a white noise process $\underline{\mathbf{w}}$ combined with a PL process $\underline{\mathbf{r}}$, the stochastic model reads

$$\mathsf{var}\{\underline{\mathbf{y}}\} = \mathsf{var}\{\underline{\mathbf{w}}\} + \mathsf{var}\{\underline{\mathbf{r}}\} = \sigma_w^2\mathbf{I} + \sigma_{pl}^2\mathbf{Q}_{pl}, \tag{4.35}$$

where the variance of each process, σ_w^2 and σ_{pl}^2, but also the spectral index κ of the PL process can be unknown.

4.2.4 Missing values

In practice, long-lasting observations are rarely perfectly continuous. Hence, one must consider missing values during the analysis. For the trajectory model, missing observations are not an issue since it is possible to evaluate each function at arbitrary times t_i. For the stochastic model, this is less trivial because we consider discrete stochastic processes defined using recurrence relations.

The most simple strategy to handle missing values is to first assume continuous measurements when constructing the covariance matrix of each process using Equation (4.24), and then to remove the lines and columns that correspond to the missing observations in the covariance matrix $\mathbf{Q}_\mathbf{y}$. Note that it is important not to delete lines of $\mathbf{L}$ before computing $\mathbf{L}\mathbf{L}^T$ as it would erroneously assume that the process stopped during the unobserved epochs.

4.3 Functional and stochastic model selection

When searching for the best functional or stochastic model for a geodetic time series, it is common to test several possible models and then chose the one that optimizes a given selection criterion.

If the considered models have the same number of functional and stochastic parameters, the likelihood can be used as a criterion to be maximized. However, when the number of parameters differs, using the likelihood is not ideal because it tends to promote models with more parameters.

To avoid over-parametrisation issues, one can perform hypothesis testing to assess whether the additional parameters significantly improve alternative models given a considered type of noise. This methodology is presented in more details in chapter 5.

Another strategy for model selection is to define a new criterion that penalizes models with more parameters in order to introduce parsimony. Two famous examples of such criterion are the Akaike Information Criterion (AIC) (Akaike, 1974) and the Bayesian Information Criterion (BIC) (Schwarz, 1978).

They are defined as functions of the log-likelihood $\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\underline{\hat{\mathbf{x}}}, \underline{\hat{\sigma}}))$, the number of observations m, the number of functional parameters n, and the number of stochastic parameters p such that

$$\mathrm{AIC} = 2(n+p) - 2\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\underline{\hat{\mathbf{x}}}, \underline{\hat{\sigma}})), \tag{4.36}$$

and

$$\mathrm{BIC} = 2\ln(m)(n+p) - 2\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\underline{\hat{\mathbf{x}}}, \underline{\hat{\sigma}})). \tag{4.37}$$

Note that, unlike the log-likelihood which is a criterion to maximize, both AIC and BIC are criteria to minimize. The main practical difference between the two criteria is that the BIC tends to penalize more models with many parameters.

Chapter 5

Event detection using hypothesis testing in linear models

In chapter 2, we mention that a necessary condition to the unbiased estimation of unknown functional parameters is an appropriate choice for the functional model. In particular, the model must not be under-parametrized. Unfortunately, the functional model is rarely known prior to the data analysis step. Although some general patterns are expected from a given set of observations, each time series may present specific deviations from these expectations, which must be taken into account.

An appropriate strategy to avoid miss-specifications in the functional model is to analyse each time series and select an optimal trajectory model, for example, using a Bayesian or Akaike information criterion. However, for large data-sets, this step can be cumbersome, and one may rely on automatic model selection methods. Automatic segmentation, that is, the automatic identification of the trajectory model of time series, is a wide topic, and the review of this field is beyond the scope of this chapter.

In this chapter, we consider that we have a reliable prior idea about the functional and stochastic model for the observations, as well as a mathematical description of the possible undetected signals. In particular, we show how *hypothesis testing in linear models* is helpful to construct automatic detectors for commonly observed patterns in geodetic time series.

The methods presented here result from Teunissen (2000b) and Amiri-Simkooei (2020). They are extensively used in chapters 9 and 10.

5.1 Hypothesis testing

The goal of hypothesis testing is to quantify whether extending the prior functional model to include additional effects is justified considering the observations and a given stochastic model. To do so, it is necessary to define a so-called *null hypothesis*, that is, a prosaic hypothesis in which the current functional model is correct, an *alternative hypothesis*, in which the functional model includes additional effects, and finally, a *statistical test* to decide between the two.

5.1.1 The null hypothesis

Under the null hypothesis, noted H_0, the functional and stochastic models of the observation vector $\underline{\mathbf{y}}$ are known and can be written as

$$H_0 : \mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x} \quad ; \quad \mathsf{var}\{\underline{\mathbf{y}}\} = \mathbf{Q}_\mathbf{y}, \tag{5.1}$$

in which $\mathbf{A}$ is the $m \times n$ design matrix and $\mathbf{x}$ is unknown the $n \times 1$ parameter vector, and $\mathbf{Q}_\mathbf{y}$ is the covariance matrix of the observations, which is assumed to be known.

5.1.2 The alternative hypothesis

Under the alternative hypothesis, the functional model must account for q additional parameters so that

$$H_a : \mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x} + \mathbf{C}_\mathbf{y}\boldsymbol{\nabla} \quad ; \quad \mathsf{var}\{\underline{\mathbf{y}}\} = \mathbf{Q}_\mathbf{y}, \tag{5.2}$$

in which $\mathbf{C}_\mathbf{y}$ is $m \times q$ the design matrix of the functional model extension, $\boldsymbol{\nabla}$ is the unknown $q \times 1$ vector of additional parameters, and $\mathbf{Q}_\mathbf{y}$ is unchanged.

5.1.3 The Generalized Likelihood Ratio Test (GLRT)

The statistical test considered in deciding between the null and the alternative hypothesis is the so-called *Generalized Likelihood Ratio Test (GRLT)*. It consists in comparing the maximum of likelihood under each hypothesis to define whether the likelihood under H_a is significantly higher than under H_0.

The maximum of likelihood under H_0 can be written as

$$\max_{\mathbf{x}_0 \in \mathbb{R}^n} \quad p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x}_0), \tag{5.3}$$

which corresponds to the likelihood of the observations $p_{\underline{\mathbf{y}}}(\mathbf{y}|\underline{\hat{\mathbf{x}}}_0)$ for a model based on the Best Linear Unbiased Estimator (BLUE) of the functional parameters $\underline{\hat{\mathbf{x}}}_0$ under H_0.

Likewise, the maximum of likelihood under H_a can be written as

$$\max_{x_a \in \mathbb{R}^n, \boldsymbol{\nabla} \in \mathbb{R}^q} \quad p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x_a}, \boldsymbol{\nabla}), \tag{5.4}$$

which corresponds to the likelihood $p_{\underline{\mathbf{y}}}(\mathbf{y}|\hat{\mathbf{x}}_\mathbf{a}, \hat{\boldsymbol{\nabla}})$ of a model based on the BLUE estimators of the functional parameters $\underline{\hat{\mathbf{x}}}_a$ and $\underline{\hat{\boldsymbol{\nabla}}}$ under H_a.

From the maximum likelihood under each hypothesis, the generalized likelihood ratio, noted GLR, is defined as

$$\underline{\mathsf{GLR}} = \frac{\max\limits_{x_0 \in \mathbb{R}^n} \quad p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x_0})}{\max\limits_{x_a \in \mathbb{R}^n, \nabla \in \mathbb{R}^q} \quad p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{x_a}, \nabla)}. \tag{5.5}$$

The observations can only be better modelled with additional parameters, therefore, the maximum likelihood under H_a is always greater or equal than that under H_0, such that $0 \leq \underline{\mathsf{GLR}} \leq 1$.

Assuming the normal distribution for both the H_0 and H_a hypotheses, the expression for the GLR becomess

$$\underline{\mathsf{GLR}} = \exp\left[-\frac{1}{2}(\hat{\mathbf{e}}_\mathbf{0}{}^T \mathbf{Q}_\mathbf{y}^{-1} \hat{\mathbf{e}}_\mathbf{0} - \underline{\hat{\mathbf{e}}}_a^T \mathbf{Q}_\mathbf{y}^{-1} \underline{\hat{\mathbf{e}}}_a)\right], \tag{5.6}$$

in which $\underline{\hat{\mathbf{e}}}_0$ is the estimator of the residuals under H_0 and $\underline{\hat{\mathbf{e}}}_a$ is that under H_a. Both residual vectors are obtained using the BLUE under their respective hypotheses.

Instead of analyzing the GLR, Teunissen (2000b) proposes to analyze the so-called T-statistic $\underline{T}_q$ defined as

$$\underline{T}_q = -2\ln(\underline{\text{GLR}}) = \underline{\hat{\mathbf{e}}}_0^T \mathbf{Q}_\mathbf{y}^{-1} \underline{\hat{\mathbf{e}}_0} - \underline{\hat{\mathbf{e}}}_a^T \mathbf{Q}_\mathbf{y}^{-1} \underline{\hat{\mathbf{e}}}_a, \tag{5.7}$$

that is, the difference between the quadratic forms of the residuals, under each hypothesis.

Equation (5.7) implies that deriving the T-statistic requires the use of the BLUE under each hypothesis. However, Teunissen (2000b) showed that Equation (5.7) can be re-written as

$$\boxed{\underline{T}_q = \underline{\hat{\mathbf{e}}}_0^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{C}_\mathbf{y} [\mathbf{C}_\mathbf{y}^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{P}_\mathbf{A}^\perp \mathbf{C}_\mathbf{y}]^{-1} \mathbf{C}_\mathbf{y}^T \mathbf{Q}_\mathbf{y}^{-1} \underline{\hat{\mathbf{e}}}_0,} \tag{5.8}$$

in which $\mathbf{P}_\mathbf{A}^\perp$ is the orthogonal projector of the BLUE under the null hypothesis, that is $\mathbf{P}_\mathbf{A}^\perp = \mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_\mathbf{y}^{-1}$.

Unlike Equation (5.7), Equation (5.8) does not explicitly require the calculation of $\underline{\hat{\mathbf{e}}}_a$.

To test whether the null hypothesis should be rejected based on the value of $\underline{T}_q$, it is necessary to know the distribution of $\underline{T}_q$ under H_0. Using the fact that $\underline{T}_q$ is expressed as a quadratic form of the residuals, it can be shown that, under H_0, the statistic $\underline{T}_q$ follows a centred chi-squared distribution with q degrees of freedom (Teunissen, 2000b), that is

$$\boxed{H_0 : \underline{T}_q \sim \chi^2(q, 0).} \tag{5.9}$$

Hence, for a given confidence level α, one can accept the alternative hypothesis if $\underline{T}_q > \chi_\alpha^2(q, 0)$.

In the case in which the alternative model only requires one additional parameter, that is $q = 1$, the design matrix of the functional model extension $\mathbf{C}_\mathbf{y}$ becomes a vector noted $\mathbf{c}_\mathbf{y}$. In this case, Equation (5.8) becomes

$$\underline{T}_1 = \underline{\hat{\mathbf{e}}}_0^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{c}_\mathbf{y} [\mathbf{c}_\mathbf{y}^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{P}_\mathbf{A}^\perp \mathbf{c}_\mathbf{y}]^{-1} \mathbf{c}_\mathbf{y}^T \mathbf{Q}_\mathbf{y}^{-1} \underline{\hat{\mathbf{e}}}_0, \tag{5.10}$$

which can be simplified into a simple ratio of two scalars such that (Amiri-Simkooei et al., 2019)

$$\boxed{\underline{T}_1 = \frac{(\mathbf{c}_\mathbf{y}^T \mathbf{Q}_\mathbf{y}^{-1} \underline{\hat{\mathbf{e}}}_0)^2}{\mathbf{c}_\mathbf{y}^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{P}_\mathbf{A}^\perp \mathbf{c}_\mathbf{y}}.} \tag{5.11}$$

In conclusion, as long as an event detection problem can be formulated in the form of hypothesis testing in linear models, one can use Equation (5.8) and (5.11) to build any kind of customized detector. In the following section, we demonstrate the construction of such detectors for the identification of periodic oscillations and different types of discontinuities.

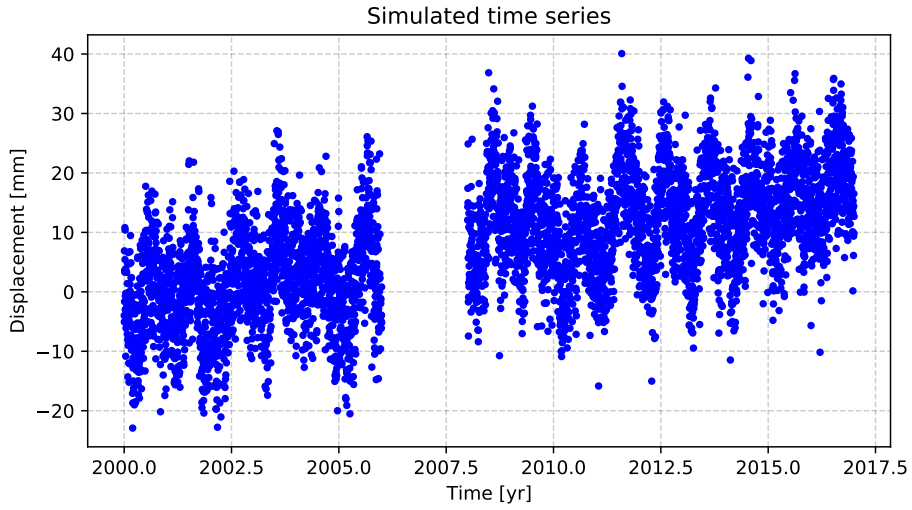

FIGURE 5.1: Synthetic time series generated for the examples.

5.2 Practical detection examples

In this section, we present some practical applications of the GLRT for the detection of specific patterns in geodetic time series, namely periodic oscillations, offsets, outliers and velocity changes.

For the sake of the demonstration, we simulated a daily displacement time series presenting a linear trend of 1 mm/yr along with an annual signal with a 5 mm amplitude and a semi-annual signal with a 3 mm amplitude. The simulated stochastic variability is a linear combination of a white noise process with 5 mm amplitude and a power-law process with spectral index $\kappa = -1.0$ and a 15 mm $\text{yr}^{-0.25}$ amplitude. The simulated time series is presented in Figure (5.1).

Hereafter, the stochastic properties of the time series are considered as fully known. In particular, the covariance matrix $\mathbf{Q_y}$ used in Equation (5.8) and Equation (5.11) is exactly that of the simulated stochastic processes. Using a realistic *stochastic* model is important when trying to detect some patterns because, otherwise, there is a risk of spurious detection (Amiri-Simkooei et al., 2019).

In practice, the stochastic model may depend on unknown variance components. In this case, one has to satisfy with the stochastic model estimated from the observations, keeping in mind that the variance component estimates can be biased.

5.2.1 Spectral analysis

The detection of non-modelled periodic oscillations in a geodetic time series can be formulated within the framework of hypothesis testing by testing whether adding a periodic oscillation with a frequency f_k induces a significant increase in the maximum of likelihood given a considered stochastic model.

When testing for periodic oscillations with arbitrary phase offsets and amplitudes, the design matrix extension $\mathbf{C_y}$ of the alternative hypothesis H_a reads

$$\mathbf{C_y} = \begin{bmatrix} \cos(2\pi f_k t_1) & \sin(2\pi f_k t_1) \\ \cos(2\pi f_k t_2) & \sin(2\pi f_k t_2) \\ \vdots & \vdots \\ \cos(2\pi f_k t_m) & \sin(2\pi f_k t_m) \end{bmatrix}, \tag{5.12}$$

and the additional parameters correspond to the two Fourier coefficients, noted a_k and b_k, of the considered frequency f_k.

It follows from Equation (5.8), that the application of the GLRT comes down to evaluating a statistic $\underline{T}_{q=2}$ distributed as $\chi^2(2,0)$ under H_0. In practice, the frequency of a possible undetected oscillation is usually unknown. Hence, one does not want to test a single frequency f_k, but rather a whole range of frequencies to draw a detection spectrum. When analysing $\underline{T}_2$ as a function of the frequency f_k, one obtains the so-called *Least-Squares Harmonic Estimation* methods (LS-HE) introduced in Amiri-Simkooei, Tiberius, and Teunissen (2007) and used for various geodetic applications (Amiri-Simkooei and Asgari, 2012; Amiri-Simkooei, 2013; Amiri-Simkooei, Zaminpardaz, and Sharifi, 2014).

The choice of the frequencies f_k, or the periods p_k, at which to test for possible periodic oscillations is completely arbitrary. However, in practice, to draw a spectrum, it is preferable to have a higher resolution at high frequency than at low frequency. In this case, the periods p_k can be sampled using the following recurrence expression (Amiri-Simkooei and Tiberius, 2007)

$$\begin{aligned} p_0 &= p_{\mathsf{min}}, \\ p_{k+1} &= p_k \left(1 + \alpha \frac{p_k}{p_{\mathsf{max}}}\right), \end{aligned} \tag{5.13}$$

in which, p_{min} is the minimum period considered, p_{max} is the maximum period considered, and $\alpha = 0.1$ is a parameter controlling the amplitude of the step between two sampled periods p_k and p_{k+1}.

Assuming a linear trend as the null hypothesis, we used Equation (5.8) to compute $\underline{T}_2$ for frequencies ranging between 0.193 cycle per year (cpy) and 182.625 cpy. The detection results are presented in Figure (5.2). The bottom plot of Figure (5.2) is referred to as LS-HE power-spectrum. As expected, $\underline{T}_2$ shows two distinct peaks at the 1 cpy and 2 cpy frequencies. Each peak is clearly above the significance threshold obtained by evaluating $\chi^2_\alpha(2,0)$ with confidence level of $\alpha = 0.99$, that is $\chi^2_{0.99}(2,0) = 9.21$. The LS-HE power-spectrum is intrinsically normalized by the combined influence of the functional and the stochastic models on the residual vector $\underline{\hat{\mathbf{e}}}_0$. As a result, the value of the statistic is unit-less and the background variability looks flat even in the presence of a power-law process.

5.2.2 Detection of discontinuities

Offsets detection

The detection of offsets in geodetic time series can be formulated within the framework of hypothesis testing by testing whether adding a Heaviside function $\mathcal{H}(t_i - t_k)$ at a time t_k induces a significant increase in the maximum of likelihood given a considered stochastic model.

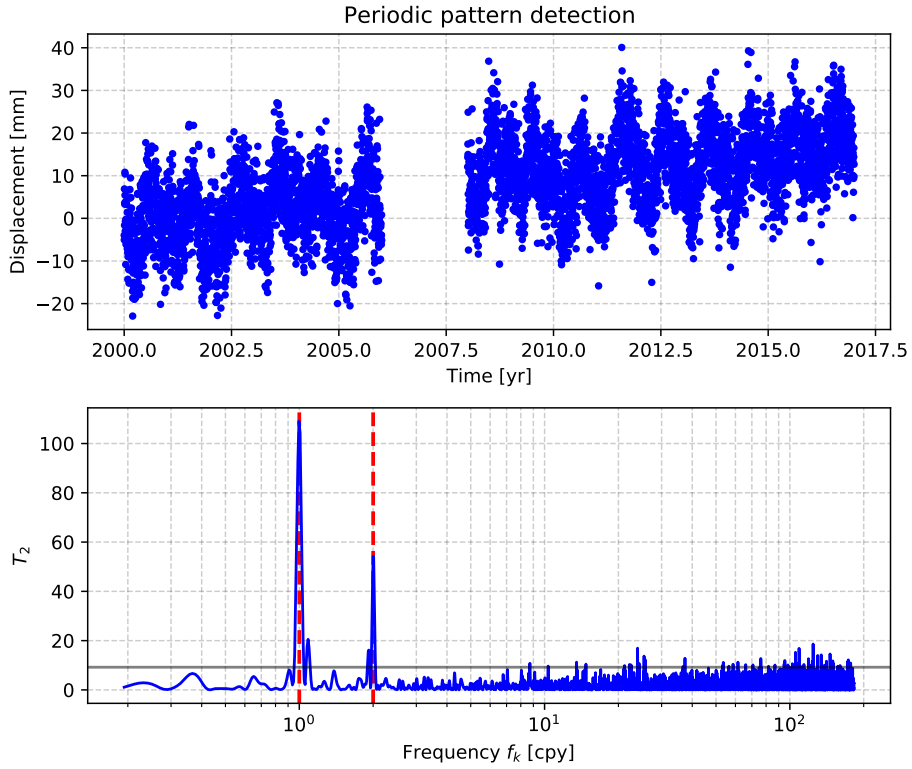

FIGURE 5.2: Example of a spectral analysis in a synthetic time series using the GLRT. Simulated time series (*top*). T_2 statistic (*bottom*). The grey line is the detection threshold for a confidence level $\alpha = 0.99$. The vertical red lines are the true frequencies.

When testing for an offset with unknown amplitude, the design matrix extension $\mathbf{C_y}$ of the alternative hypothesis H_a is a vector $\mathbf{c_y}$ defined as

$$\mathbf{c_y} = \begin{bmatrix} \mathcal{H}(t_1 - t_k) \\ \mathcal{H}(t_2 - t_k) \\ \vdots \\ \mathcal{H}(t_m - t_k) \end{bmatrix} \tag{5.14}$$

and the additional parameter is the offset amplitude.

It follows from Equation (5.11) that the application of the GLRT to offset detection comes down to evaluating a $\underline{T}_1$ statistic distributed as $\chi^2(1,0)$ under H_0. Often, the time of occurrence of an offset is unknown. Hence, it is necessary to test for offsets at multiple time t_i to identify the date of a possibly undetected offset. When analysing the evolution of $\underline{T}_1$ as a function of the time t_k one obtains the univariate offset detection method proposed in (Amiri-Simkooei et al., 2019).

To illustrate the offset detection method, we added a 15 mm offset in 2010.00 to the simulated time series in Figure (5.1). Assuming a linear trend, an annual and a semi-annual signal as the null hypothesis, we used Equation (5.11) to compute the $\underline{T}_1$ statistic for each day between 2000 and 2017. The detection results are presented in Figure (5.3). The $\underline{T}_1$ statistic shows a maximum at the date of the undetected offset. Moreover, the magnitude of the maximum is above the significance threshold obtained by evaluating $\chi^2_\alpha(1,0)$ with confidence level of $\alpha = 0.99$, that is $\chi^2_{0.99}(1,0) = 6.23$. Hence, in this example, the GLRT correctly detects the non-modelled offset.

Outlier detection

The detection of outliers in geodetic time series can be formulated within the framework of hypothesis testing by testing whether adding a Dirac function $\delta(t_i - t_k)$ at a time t_k induce a significant increase in the maximum of likelihood given a considered stochastic model.

When testing for an outlier with unknown amplitude, at a time t_k, the design matrix extension $\mathbf{C_y}$ of the alternative hypothesis H_a is a vector $\mathbf{c_y}$ defined as

$$\mathbf{c_y} = \begin{bmatrix} \delta(t_1 - t_k) \\ \delta(t_2 - t_k) \\ \vdots \\ \delta(t_m - t_k) \end{bmatrix} \tag{5.15}$$

and the additional parameter is the outlier amplitude.

As for offset detection, the application of the GLRT to outlier detection comes down to evaluating a $\underline{T}_1$ statistic distributed as $\chi^2(1,0)$ under H_0. Since the occurrence time of an undetected outlier is also unknown, it is necessary to test for outliers at multiple time t_i.

To illustrate a case of outlier detection, we added a 50 mm outlier, occurring in 2004.00, to the simulated time series in Figure (5.1). Assuming a linear trend, an annual and a semi-annual signal as the null hypothesis, we used Equation (5.11) to compute the $\underline{T}_1$ statistic for each day between 2000 and 2017. The detection results are presented in Figure (5.4). In Figure (5.4), the $\underline{T}_1$ statistic shows a maximum at the

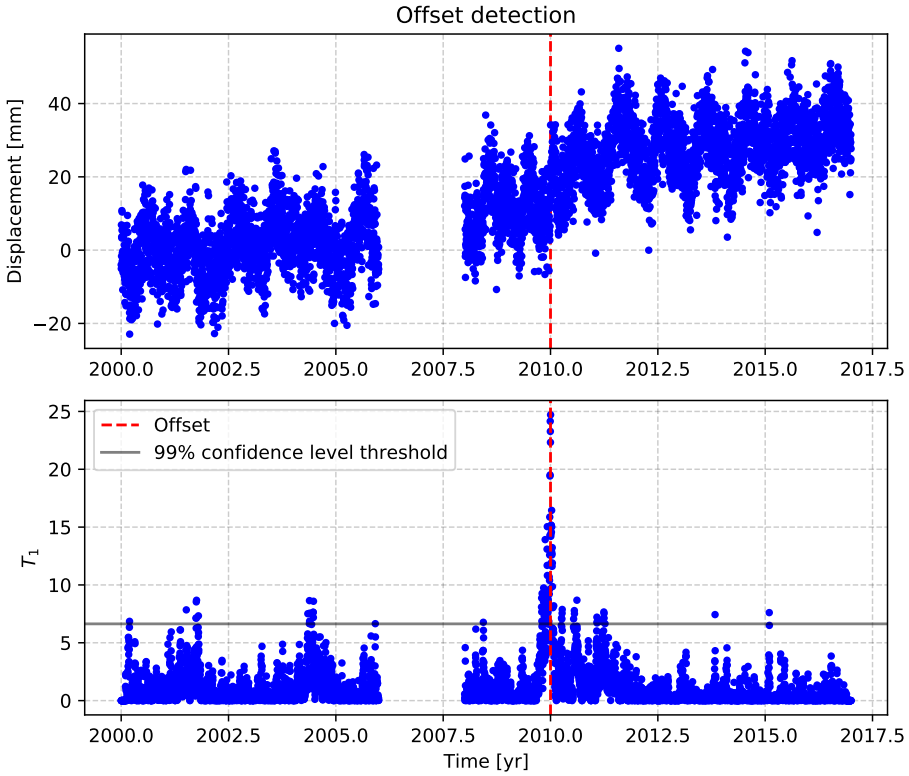

FIGURE 5.3: Example of offset detection in a synthetic time series using the GLRT. Simulated time series (*top*). T_1 statistic (*bottom*). The grey line is the detection threshold for a confidence level $\alpha = 0.99$. The vertical red line is the true date of the offset.

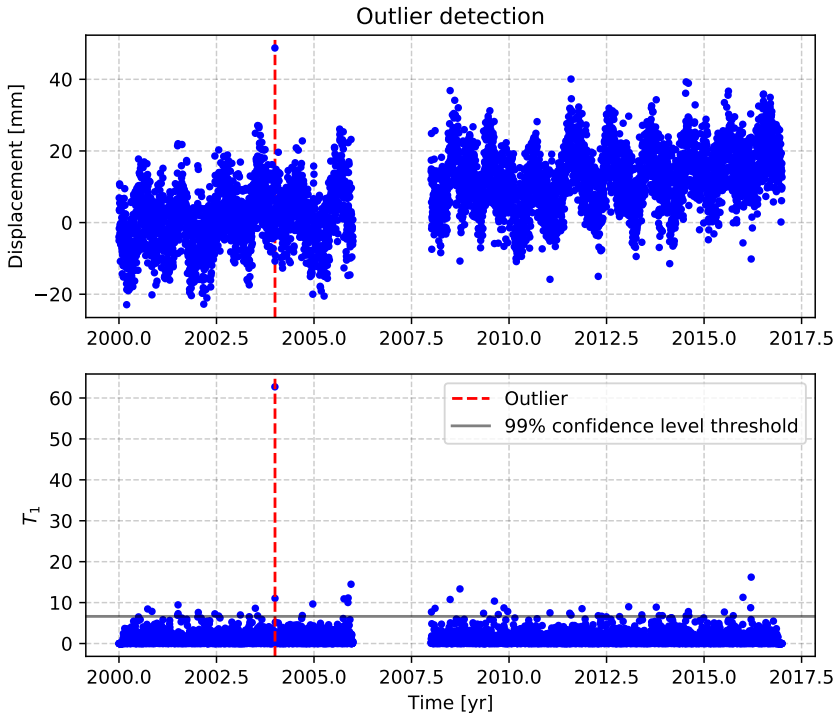

FIGURE 5.4: Example of outlier detection in a synthetic time series using the GLRT. Simulated time series (*top*). T_1 statistic (*bottom*). The grey line is the detection threshold for a confidence level $\alpha = 0.99$. The vertical red line is the true date of the outlier.

date of the undetected offset. Moreover, the magnitude of the maximum is above the significance threshold obtained by evaluating $\chi^2_\alpha(1,0)$ with confidence level of $\alpha = 0.99$, that is $\chi^2_{0.99}(1,0) = 6.23$. Hence, in this example, the GLRT correctly detects the outlier.

Velocity change detection

By slightly modifying the offset detection method, one can also define a velocity change detector. In particular, when testing for a velocity change with unknown amplitude, the design matrix extension $\mathbf{C_y}$ of the alternative hypothesis H_a is a vector $\mathbf{c_y}$ defined as

$$\mathbf{c_y} = \begin{bmatrix} (t_1 - t_k)\mathcal{H}(t_1 - t_k) \\ (t_2 - t_k)\mathcal{H}(t_2 - t_k) \\ \vdots \\ (t_m - t_k)\mathcal{H}(t_m - t_k) \end{bmatrix}, \tag{5.16}$$

and the additional parameter is amplitude of the velocity change.

To illustrate a velocity change detection, we added a 5 mm/yr velocity change, occurring in 2012.00, to the simulated time series in Figure (5.1). Assuming a linear trend, an annual and a semi-annual signal as the null hypothesis, we used Equation (5.11) to compute the $\underline{T}_1$ statistic for each day between 2000 and 2017. The detection results are presented in Figure (5.5). In Figure (5.5), the $\underline{T}_1$ statistic does not

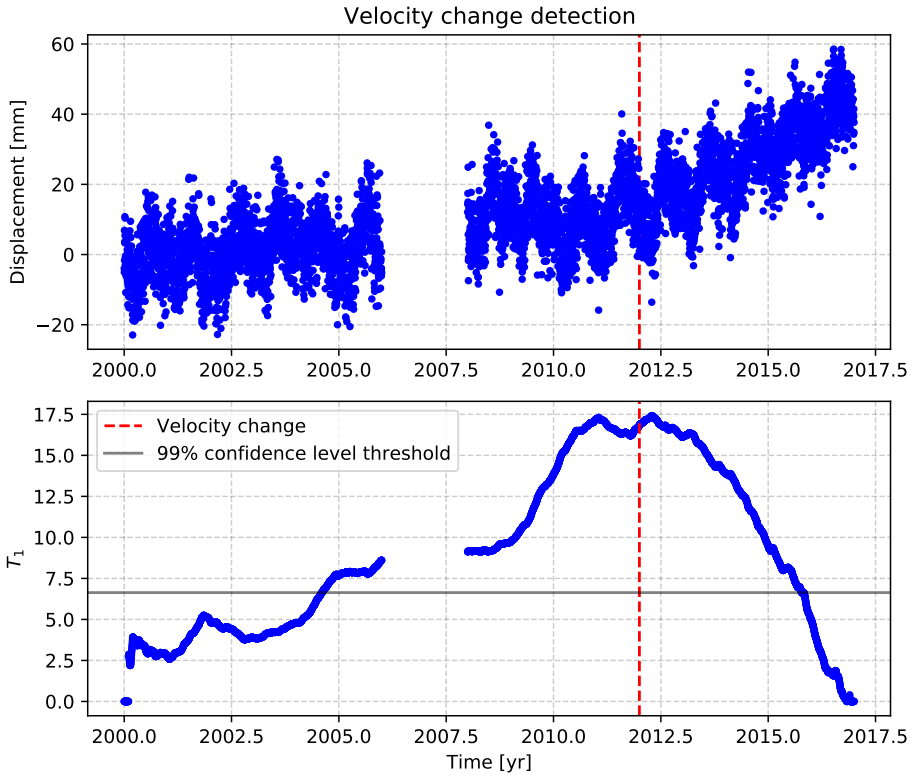

FIGURE 5.5: Example of velocity change detection in a synthetic time series using the GLRT. Simulated time series (*top*). T_1 statistic (*bottom*). The grey line is the detection threshold for a confidence level $\alpha = 0.99$. The vertical red line is the true date of the velocity change.

show a clear maximum at the exact date of the undetected velocity change. This implies that identifying the date of this velocity change is challenging with the considered stochastic model. However, the maximum $\underline{T}_1$ statistic is above the significance threshold obtained by evaluating $\chi^2_\alpha(1,0)$ with a confidence level of $\alpha = 0.99$, that is $\chi^2_{0.99}(1,0) = 6.23$. Hence, the detector does identify a velocity change, but with a few month error on the date.

Chapter 6

Assessment of tide gauge sensor offsets in the presence of scale errors

In this first contribution chapter, we address the calibration of tide gauges in its simplest form: the comparison of two collocated gauges. We specifically focus on a mandatory step in the installation of a tide-gauge: the determination of so-called the sensor offset. In the presence of a scale error, it is unclear how to estimate the sensor offset. Here, we provide a thorough description of the link between the measurement biases and sea-level biases observed between two gauges, which we then use to provide indications about the estimation of the sensor offset.

6.1 Introduction

With the advent of climate monitoring from space, oceanographic quantities, such as sea level, started to be investigated globally, making it necessary to combine data from different sensors. To ensure the long term stability and interoperability of coastal sea level time series, all sensors must provide comparable and consistent quantities. In practice, several technical problems can affect sea-level observations, and the accuracy of sea-level time series must be regularly controlled during calibration campaigns.

This chapter focuses on the calibration of ground-based coastal tide gauges, such as pressure gauges, acoustic gauges (IOC, 1994) and radar gauges (IOC, 2016). These types of tide gauge are the main constituent of sea-level program such as the Global Sea Level Observing System (GLOSS) (IOC, 2012). Besides, coastal tide gauges are commonly used as references – or ground-truth – to assess the performance of other technologies. They have been involved in the characterization of systems based on Global Navigation Satellite System (GNSS), such as GNSS buoys (André et al., 2013; Chupin et al., 2020) or GNSS reflectometry (Larson, Ray, and Williams, 2017), and they played a role in the calibration and validation of satellite altimeters (Bonnefond, Haines, and Watson, 2011; Watson et al., 2011).

Several types of biases can affect tide gauge measurements. During calibration campaigns, the most commonly observed ones are the sensor offsets, the scale errors and the time delays (Míguez Martín, Testut, and Wöppelmann, 2008). Sensor offsets are constant biases induced by a wrong assumption on the position of the sensor's point of zero range (PZR) (IOC, 2016). Scale factors induce an error dependent on the measured sea level. They result from a sensor tilts or errors in the analog to

digital conversion. Time delays result from either lag between the actual locations of two gauges, or from clock synchronization issues.

Unlike sensor offsets and scale errors, time delays do not depend linearly on the time or the sea level. As the statistical tools used in this study assumes linearity, time delays must be estimated and removed prior to further analyses. This can be done, for instance, by finding the delay that maximizes the cross-correlation between the tested gauge and the reference gauge time series.

In this chapter, we present a tide gauge calibration procedure in the presence of a sensor offset and of a scale error. The influence of the scale error on the analysis of sea-level differences have been studied by Pérez et al. (2014) in the context of datum offset determination. In this study, we more specifically focus on the determination of the sensor offset, both for its computation and its uncertainty.

To this end, we first demonstrate the exact expression of the sea-level difference as a function of the sensor offset and the scale factor. Using this expression, we present a methodology allowing a quick interpretation of sea-level differences in terms of tide gauge measurement biases. Finally, we evaluate the experiment duration required to reach a sensor offset determination with a few millimetres uncertainty.

6.2 Influence of a sensor offset and a scale error on sea-level differences

The sea-level time series $H_T(t)$ for a given gauge is computed from $m_T(t)$, the measured distance between the sensor Point of Zero Range (PZR) and the water surface, with

$$H_T(t) = c_T + m_T(t), \tag{6.1}$$

where c_T is the reference level, given at a reference point named Reference Survey Mark (RSM) (IOC, 2016). To simplify notations, it is assumed that the RSM marks the position of the expected PZR. Note that $H_T(t)$, c_T and $m_T(t)$ can be positive or negative depending on the chosen reference level – or datum – and sign convention. Hereafter, vertical distances are taken as positive when measured upwards.

In the absence of sensor offset, the RSM coincides with the sensor PZR, as illustrated in Figure (6.1). In this ideal case, the measurement $m_T(t)$ equals the true distance between the RSM and the water surface, noted $d_T(t)$, that is

$$m_T(t) = d_T(t).$$

However, in reality, the RSM provided by a tide gauge manufacturer may differ from the sensor PZR. This would result in a sensor offset, noted o_T, defined as the distance from RSM and pointing toward the gauge's PZR, as depicted in Figure (6.2). This sensor offset o_T will affect the measured range $m_T(t)$ such that

$$m_T(t) = d_T(t) - o_T.$$

in which the sign of o_T comes from the fact that it is defined in the opposite direction.

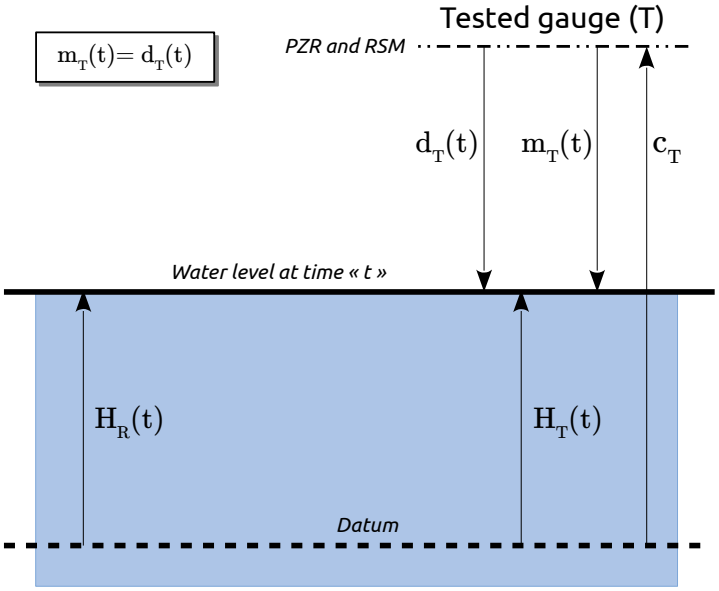

FIGURE 6.1: Configuration with no offset and no scale error. RSM: Reference Survey Mark. PZR: Point of Zero Range.

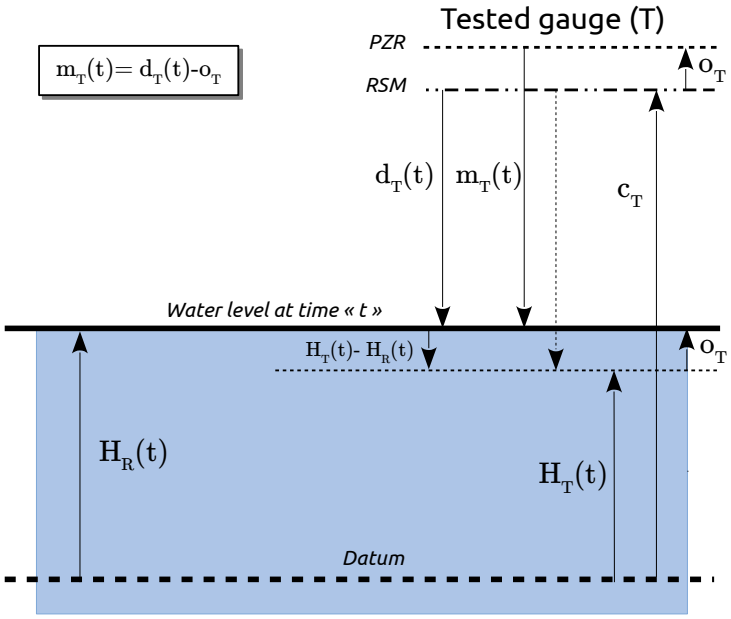

FIGURE 6.2: Configuration with an offset but no scale error. RSM: Reference Survey Mark. PZR: Point of Zero Range.

The tested gauge may, in addition, be affected by a scale error, noted s_T, introducing a distance-dependent measurement error (Figure 6.3). In practice, this error is expected to be small compared to the measured distance, so that the overall distance is scaled by a factor close to 1.0 and noted $(1 + s_T)$. The scale factor $(1 + s_T)$ affects the distance between the sensor PZR and the water, that is $d_T(t) - o_T$, such that

$$m_T(t) = (1 + s_T)(d_T(t) - o_T). \tag{6.2}$$

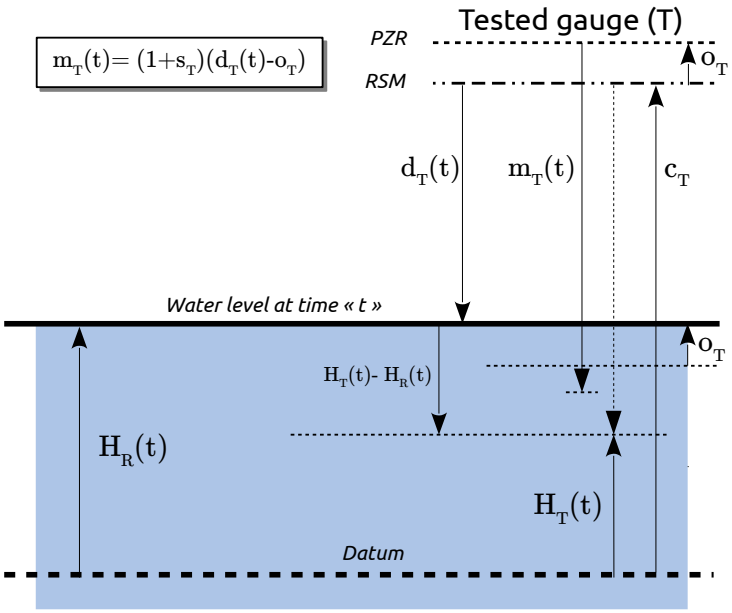

FIGURE 6.3: Configuration with an offset and a scale error. RSM: Reference Survey Mark. PZR: Point of Zero Range.

Substituting Equation (6.2) into Equation (6.1) leads to an expression for the sea-level $H_T(t)$ as a function of the scale, the sensor offset o_T, the measured distance d_T, and reference level c_T:

$$H_T(t) = c_T - o_T - s_T o_T + d_T(t) + s_T d_T(t). \tag{6.3}$$

During a calibration experiment, the sea-level $H_T(t)$ is compared to the sea-level provided by a reference gauge $H_R(t)$, assumed to be unbiased. $H_R(t)$ can be directly linked to the tested gauge reference level c_T and distance $d_T(t)$ as (Figure 6.1)

$$H_R(t) = c_T + d_T(t). \tag{6.4}$$

The difference $\Delta_T(t) = H_T(t) - H_R(t)$ corresponds to a sea-level bias that must be modelled and corrected. The differences $\Delta_T(t)$ can be obtained by subtracting Equation (6.4) to Equation (6.3) such that

$$\Delta_T(t) = H_T(t) - H_R(t) = -(1 + s_T)o_T + s_T d_T(t). \tag{6.5}$$

Because the reference level $H_R(t)$ is only temporarily measured and the true distance $d_T(t)$ is unknown, the sea-level bias $\Delta_T(t)$ must be expressed as a function of $H_T(t)$. Equation (6.3) can be re-arranged into

$$d_T(t) = \frac{1}{(1 + s_T)} H_T(t) - \frac{1}{(1 + s_T)} c_T + o_T, \tag{6.6}$$

and substituting this expression in Equation (6.5), so that

$$\Delta_T(t) = -o_T - \frac{s_T}{(1+s_T)}c_T + \frac{s_T}{(1+s_T)}H_T(t). \tag{6.7}$$

Finally, Equation (6.7) can be simplified by noting

$$\boxed{\beta_T = \frac{s_T}{(1+s_T)}}, \tag{6.8}$$

and

$$\boxed{\alpha_T = -o_T - \beta_T c_T}, \tag{6.9}$$

so that

$$\boxed{H_T(t) - H_R(t) = \alpha_T + \beta_T H_T(t)}. \tag{6.10}$$

which demonstrates that the sea-level bias is a linear function of $H_T(t)$ in the presence of a sensor offset and a scale factor.

The Equation (6.10) points out to an inconsistency in the Equation (2) of Míguez Martín, Testut, and Wöppelmann (2008), in which the intercept is absent. Equation (2) in Míguez Martín, Testut, and Wöppelmann (2008) assumes that $H_T(t) - H_R(t) = 0$ for $H_T(t) = 0$, which is, in general, not verified.

Comparing two tide gauges, the parameters α_T and β_T can be estimated by a linear regression in the so-called Van de Casteele diagram. Proposed in 1962 for mechanical gauges by Lennon (1968) and revisited for modern technologies in Míguez Martín, Testut, and Wöppelmann (2008), this diagram displays the sea-level (H) difference between a tested (T) and a reference (R) instrument, $\Delta_T(t) = H_T(t) - H_R(t)$ as a function of $H_T(t)$ or $H_R(t)$. However, for a more intuitive reading, $\Delta_T(t)$ is used for the horizontal axis and $H_T(t)$ for the vertical axis . This tool has been applied in numerous studies (Míguez Martín, Le Roy, and Wöppelmann, 2008; André et al., 2013; Pérez et al., 2014; Larson, Ray, and Williams, 2017; Pytharouli, Chaikalis, and Stiros, 2018). Figure (6.4) illustrates the patterns visible in a Van de Casteele diagram in the presence of the most common biases. A precaution that should be taken when estimating α_T and β_T from the Van de Casteele diagram, is that the diagram must display the difference $H_T(t) - H_R(t)$ as a function of the tested gauge sea-level $H_T(t)$ and not reference gauge sea-level $H_R(t)$.

Once the estimated values $\hat{\alpha}_T$ and $\hat{\beta}_T$ are obtained, a corrected sea-level, that is, an estimate of the sea-level of the reference gauge noted $\hat{H}_R(t)$, can be obtained from $H_T(t)$ as

$$\hat{H}_R(t) = H_T(t) - (\hat{\alpha}_T + \hat{\beta}_T H_T(t)) \tag{6.11}$$

Although useful to obtain $\hat{H}_R(t)$, the parameters $\hat{\alpha}_T$ and $\hat{\beta}_T$ do not directly provide information about the bias on the measurement $m_T(t)$. To derive estimates of the scale error ($\hat{s}_T$) and the sensor offset ($\hat{o}_T$) from $\hat{\alpha}_T$ and $\hat{\beta}_T$, one can show from Equations (6.8) and (6.9) that

$$\boxed{\hat{s}_T = \frac{\hat{\beta}_T}{1-\hat{\beta}_T}}, \tag{6.12}$$

and

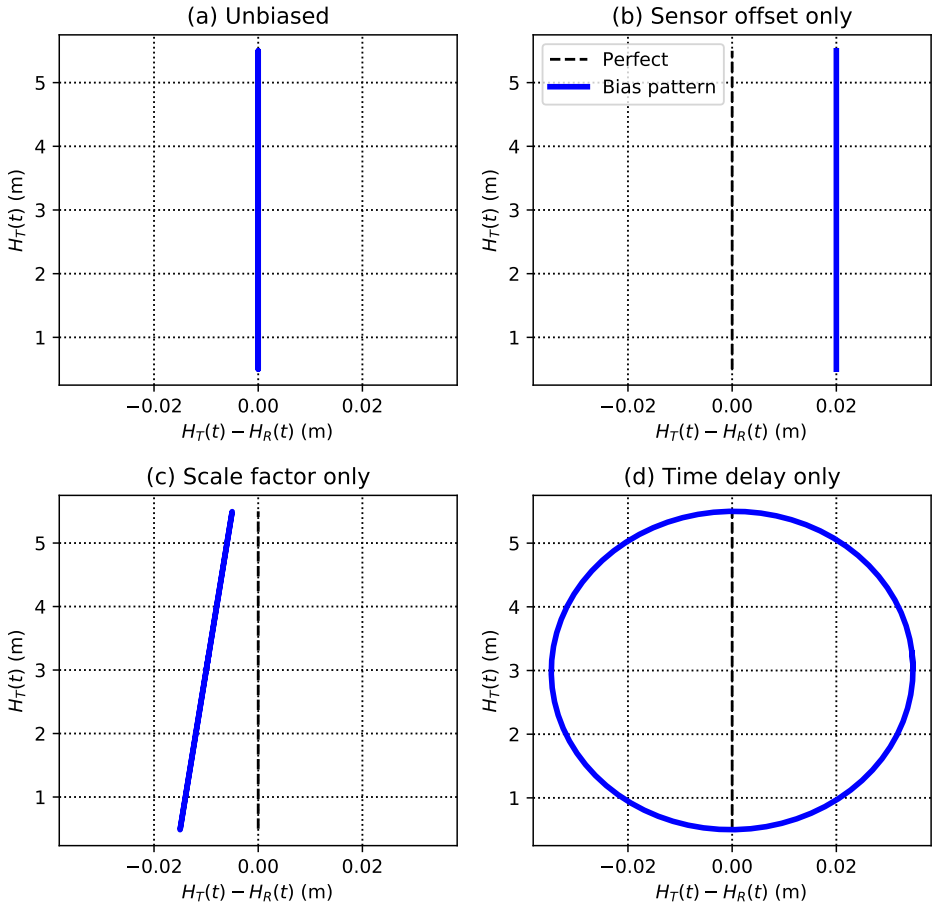

FIGURE 6.4: Examples of the deterministic influence of common measurement biases on the Van de Casteele diagram. $H_T(t)$ is the sea-level recorded by the tested gauge at time t. $\Delta_T(t) = H_T(t) - H_R(t)$ is the difference of sea-level between the tested gauge and the reference gauge.

$$\boxed{\hat{o}_T = -\hat{\alpha}_T - \hat{\beta}_T c_T.} \tag{6.13}$$

$\hat{s}_T$ and $\hat{o}_T$ can either be used for a direct diagnosis of instrumental issues affecting $m_T(t)$ or for the computation of parameters $\alpha_T^{'}$ and $\beta_T^{'}$ in the case the tested gauge has to be moved from the reference level c_T to a new level $c_T^{'}$.

Note that $-o_T$ is the value of the sea level bias model (6.10) for $H_T(t) = c_T$. Hence, as illustrated on Figure (6.5), by drawing the intersection between the bias model and reference level, the sensor offset o_T has a simple geometrical interpretation on the Van de Casteele diagram.

6.3 Impact of the scale error on mean difference interpretation.

When analysing sea-level differences between two tide-gauges, it is easy to assume that the mean sea-level difference, noted $\bar{\Delta}_T$, should correspond to the sensor offset,

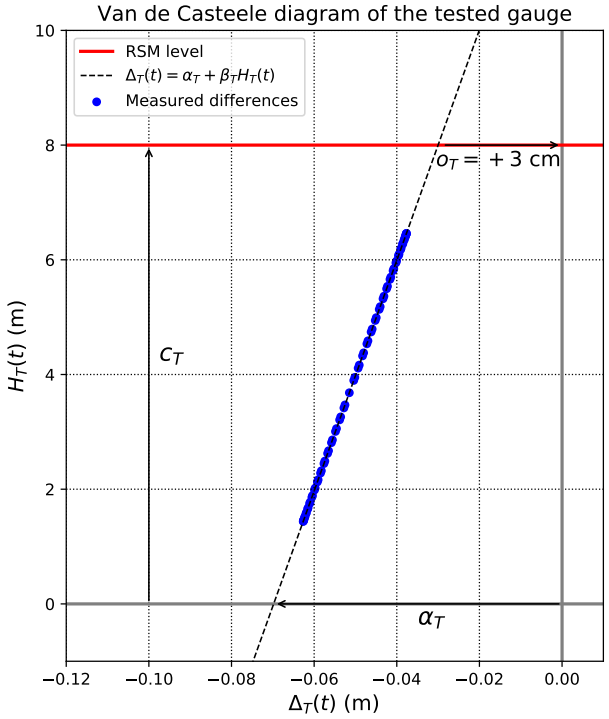

FIGURE 6.5: Visual identification of the offset on the Van de Casteele diagram of the tested gauge.

up to a conventional change of sign. When $s_T = 0$, Equation (6.7) gives $o_T = -\Delta_T(t)$, therefore, o_T can be estimated from $\bar{\Delta}_T$.

However, Equation (6.7) also shows that $-\bar{\Delta}_T \neq o_T$ as long as $s_T \neq 0$. Hence, neglecting the influence of the scale error when interpreting the mean difference would lead to an error $(-\bar{\Delta}_T - o_T)$, that can be quantified using Equations (6.10) and (6.13).

From Equation (6.10), one can show that

$$\bar{\Delta}_T = \alpha_T + \beta_T \bar{H}_T. \tag{6.14}$$

where $\bar{H}_T$ denotes the average sea-level measured by the tested gauge. If the calibration experiment is done over one or more tidal cycles, then $\bar{H}_T$ is approximately the mean sea-level.

From Equations (6.13) and (6.14) the mean difference error $(-\bar{\Delta}_T - o_T)$ can be expressed as

$$\boxed{(-\bar{\Delta}_T - o_T) = \frac{s_T}{1 + s_T}(c_T - \bar{H}_T),} \tag{6.15}$$

where $(c_T - \bar{H}_T)$ is approximately the sensor height above mean sea-level.

Based on Equation (6.15), the Figure (6.6) shows the amplitude of the mean difference interpretation error $(-\bar{\Delta}_T - o_T)$ as a function of the scale error s_T and the sensor height above mean sea-level $(c_T - \bar{H}_T)$. In Figure (6.6), the interpretation error quickly reaches a few centimetres. Hence, the influence of a scale factor cannot be neglected while estimating the sensor offset o_T.

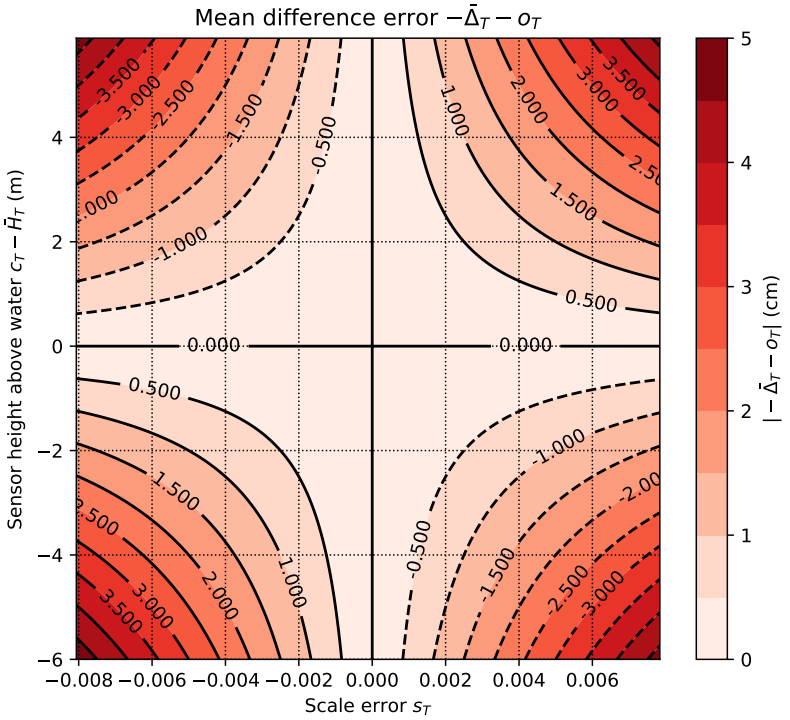

FIGURE 6.6: Evolution of the interpretation error resulting from the assumption that $o_T = -\bar{\Delta}_T$ in the presence of scale error.

6.4 How long should a calibration experiment last ?

When designing a calibration experiment, one may aim at assessing the sensor offset $\hat{o}_T$ with a few millimetres precision for later applications. In that case, it is important to design a calibration experiment long enough to reach the desired level of precision.

Given that $\hat{o}_T$ is a function of $\hat{\alpha}_T$ and $\hat{\beta}_T$ in Equation (6.13), its precision depends on the variance and the covariance of $\hat{\alpha}_T$ and $\hat{\beta}_T$. In particular, let's denote $\underline{\hat{\mathbf{x}}}$ the estimator of these parameters such that

$$\underline{\hat{\mathbf{x}}} = \begin{bmatrix} \underline{\hat{\alpha}}_T \\ \underline{\hat{\beta}}_T \end{bmatrix}. \tag{6.16}$$

The 2×2 covariance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$ of the estimator $\underline{\hat{\mathbf{x}}}$ can be denoted as

$$\mathbf{Q}_{\hat{\mathbf{x}}} = \begin{bmatrix} \sigma^2_{\hat{\alpha}_T} & \sigma_{\hat{\alpha}_T \hat{\beta}_T} \\ \sigma_{\hat{\alpha}_T \hat{\beta}_T} & \sigma^2_{\hat{\beta}_T} \end{bmatrix} \tag{6.17}$$

and, by applying the linear propagation law of variances to Equation (6.13), one can show that the standard deviation of $\underline{\hat{o}}_T$, noted $\sigma_{\hat{o}_T}$, reads

$$\boxed{\sigma_{\hat{o}_T} = \sqrt{\sigma^2_{\hat{\alpha}_T} + 2c_T \sigma_{\hat{\alpha}_T \hat{\beta}_T} + c_T^2 \sigma^2_{\hat{\beta}_T}}.} \tag{6.18}$$

Hence, to assess the evolution of $\sigma_{\hat{\delta}_T}$ with the calibration experiment duration, one must simulate the covariance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$ under theoretical conditions. When estimating $\hat{\alpha}_T$ and $\hat{\beta}_T$ from a linear regression on a Van de Casteele diagram, and under the assumption of both independent and identically distributed observation uncertainties, the covariance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$ reads

$$\mathbf{Q}_{\hat{\mathbf{x}}} = \sigma^2_{\Delta_T} (\mathbf{A}^T \mathbf{A})^{-1} \tag{6.19}$$

in which $\sigma^2_{\Delta_T}$ is the considered variance for the measured differences Δ_T, and $\mathbf{A}$ is the $m \times 2$ design matrix of the linear regression defined as

$$\mathbf{A} = \begin{bmatrix} 1 & H_T(t_1) \\ 1 & H_T(t_2) \\ \vdots & \vdots \\ 1 & H_T(t_m) \end{bmatrix}, \tag{6.20}$$

with m the considered number of observation of the experiment.

Given the definition $\mathbf{A}$, the sensor offset uncertainty depends on the measured sea-level $H_T(t)$, which is not known before the experiment. However, one can assess the evolution of $\sigma_{\hat{\delta}_T}$ based on a tide model.

To illustrates the dependence of $\sigma_{\hat{\delta}_T}$ with both the experiment duration and the expected tide amplitude, we use the following simplified tide model

$$H_T(t) = M + A \cdot \cos(2\pi f_{\mathsf{M}_2} t) \tag{6.21}$$

where M is the mean sea-level, f_{M_2} is the frequency of the M_2 tidal constituent, and A is the tide amplitude.

Hence, by considering sea-level differences $\Delta_T(t)$ measured every 10 min with a standard deviation $\sigma_{\Delta_T} = 1$ cm, we simulated the standard deviation of the sensor offset $\sigma_{\hat{\delta}_T} = \sqrt{\sigma^2_{\hat{\delta}_T}}$ for experiments with duration ranging from 8h to 48h and tide models with amplitudes A ranging from 0.5 m to 2.5 m. The considered mean sea level is $M = 3.0$ m. The simulation results are presented in Figure (6.7).

In Figure (6.7), we observe that to reach a given level of precision, locations with small tidal ranges requires longer experiments than that with large tidal ranges. This is because a precise determination of the scale parameter $\hat{\beta}_T$ requires comparing tide gauges over a large tidal range. Besides, we note that for a few millimetres accuracy on the sensor offset, a few days of continuous observation are quickly necessary. Therefore, for practical reasons, it seems very difficult to achieve a precise tide gauge calibration with a reference gauge requiring continuous human intervention.

6.5 Conclusion

Calibration experiment is a crucial step for the installation and control of coastal tide gauges. The two most frequent types of tide gauge measurement biases observed in such experiments are the sensor offset and the scale error. In this chapter, we studied how a sensor offset and scale error influence the measured sea-level. We also provided expressions allowing to assess these parameters from the sea-level differences measured between two tide gauges.

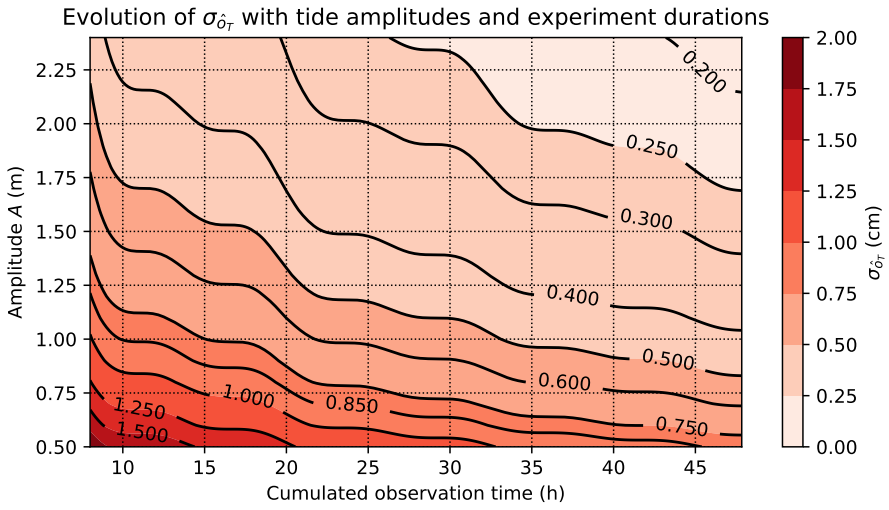

FIGURE 6.7: Evolution of the standard deviation of the sensor offset σ_{o_T} with the cumulative observation time and the tide amplitude A.

Using these expressions, we showed that the effect of the scale error cannot be neglected when interpreting the mean difference between two sensors. In particular, we showed that assuming the mean difference corresponds to the sensor offset in the presence of a scale error can easily result in a few centimetres interpretation error.

Finally, based on simulations, we estimated the dependence of the sensor offset uncertainty with the calibration experiment duration and tidal amplitude. We showed that estimating the sensor offset with a few millimetres uncertainty quickly requires a few days experiment, especially for a small tidal range. This observation may influence the choice of sensor used for such calibrations by promoting self-recording devices over human-dependent measurements.

Chapter 7

Assessment of tide gauges biases and precisions by the combination of multiple co-located time series

This chapter proposes a new method for the cross-calibration of tide gauges. Based on the combination of at least three co-located sea level time series, it takes advantage of the Least-Squares Variance Component Estimation method to assess both sea-level biases and uncertainties in real conditions. The method was applied to a multi-instrument experiment carried out on Aix island, France, in 2016. Six tide gauges were deployed to carry out simultaneous sea level recordings for 11 hours.

The content of this chapter has been published as Kevin Gobron et al. (2019). "Assessment of Tide Gauge Biases and Precision by the Combination of Multiple Collocated Time Series". In: *Journal of Atmospheric and Oceanic Technology* 36.10, pp. 1983–1996

7.1 Introduction

Tide gauges aim at measuring the vertical distance between the sea level and a reference level (or datum). Historically, tide gauges were first used for tide prediction and navigation (Cartwright, 2000); today, their applications have been extended (Pugh and Woodworth, 2014). Clustered into networks of continuously operating stations, they are the key components of storm surge or tsunami warning systems and climate-related monitoring programs, such as the Global Sea Level Observing System (GLOSS) (IOC, 2012).

A wide range of distance meter technologies can serve to implement a tide gauge, as long as it can resolve both sea level and datum along the vertical. The datum of a sea level station is a local and conventional reference level, independent from any instruments. It enables the construction of long time series with successive or overlapping tide gauges. The datum is defined through a network of benchmarks grounded around the sea level station, some of them can also act as benchmarks for leveling networks (IOC, 1985; Pugh and Woodworth, 2014). Thus, a preliminary step in field calibrations consists of tying the reference gauge to the station datum or controlling whether it is properly tied.

The simplest and oldest types of tide gauge are graduated poles or tide poles placed against a vertical structure at the coast (Cartwright, 2000). Tide poles requiring human-made measurements are still in-use, along with electric tape probes for on-site field calibration of more elaborated self-recording tide gauges. Since 1985,

the manuals of the Intergovernmental Oceanographic Commission (IOC) have covered the basic principles of the main types of tide gauges in use across the world, ranging from mechanical float gauges (IOC, 1985) to radar technologies (IOC, 2016), including pressure and acoustic gauges (IOC, 2002; IOC, 2006).

Over the past decade, radar-based technologies appeared as the preferred ones (IOC, 2016). However, new technologies are emerging, based on Global Navigation Satellite System (GNSS) buoys (André et al., 2013), GNSS reflectometry (GNSS-R) (Larson, Ray, and Williams, 2017), or laser distance measurement (MacAulay, O'Reilly, and Thompson, 2008). A tide gauge complying with GLOSS standards should be capable of measuring instantaneous sea level with an accuracy better than 1.00 cm, in all conditions of tide, waves, currents and weather (IOC, 2016). As laboratory testings do not ensure those performances, the practice has evolved towards field experiments (Míguez Martín, Testut, and Wöppelmann, 2008; Míguez Martín, Le Roy, and Wöppelmann, 2008; Park, Heitsenrether, and Sweet, 2014; Pérez et al., 2014).

When dealing with accuracy requirements, it is useful to distinguish random and systematic errors. The random error is the error component that, in replicate measurements, varies in an unpredictable manner, whereas the systematic error is the error component that, in replicate measurements, changes in a predictable manner (BIPM et al., 2008).

Given the crucial role of tide gauges in coastal sea level observation, the increasing number of available technologies and the evolution of accuracy requirements, this chapter aims at providing a cross-calibration method that quantifies both systematic errors – the biases – and random errors of sea level time series.

Determining the errors of given time series can be achieved through three approaches : (a) the observed time series can be compared with that from a more precise instrument, (b) it can be compared with theory in cases where the observed phenomena can be very precisely modelled, and (c) observed time series of three or more instruments can be analyzed to obtain enough information to determine the uncertainty of each.

Approach (a), also known as buddy-checking, is routinely used during calibration campaigns where a pair of tide gauges are compared over a tidal cycle, sometimes with the help of the so-called Van de Casteele diagram (Lennon, 1968; IOC, 1985). During the last decade, several studies have investigated the performances of radar gauges, pressures gauges, GNSS buoys or GNSS reflectometry based on this approach (Watson, Coleman, and Handsworth, 2008; Míguez Martín, Le Roy, and Wöppelmann, 2008; Míguez Martín, Testut, and Wöppelmann, 2012; Pérez et al., 2014; Larson, Ray, and Williams, 2017; Pytharouli, Chaikalis, and Stiros, 2018). Even if this approach can provide bias estimates and general accuracy metrics, such as mean error or root mean square error (RMSE), it cannot rigorously separate the uncertainties of each gauge.

Approach (b) would correspond to removing a tide model from the measured sea level time series. But, because of the complexity of meteorological and ocean dynamics involved in sea level fluctuations, these models are not precise enough to assess the performance of tide gauges at the targeted centimeter level.

Approach (c) is classically used in metrology (Pálinkáš et al., 2017) and has often been used in geodesy through the three-cornered-hat (TCH) estimation method (Gray and Allan, 1974), for example, to determine the stability of reference station positions (Feissel-Vernier, Viron, and Le Bail, 2007; Abbondanza et al., 2015) or the precision of space gravity model (Viron et al., 2008; Valty et al., 2013). The TCH is not the only possible implementation of the approach (c): the more general framework of Variance Component Estimation (VCE) can similarly address this problem, as shown by the theoretical example (4.10) of Amiri-Simkooei (2007). The TCH and VCE examples can separate the uncertainty of each gauge, but assume the absence of sea level biases.

To take advantage of both approaches (a) and (c), this chapter proposes a combination model that extends the use of approach (c) to the analysis of potentially biased time series. Assessing the tide gauge uncertainties in addition to the sea level bias parameters is made possible by the use the Least-Squares Variance Component Estimation (LS-VCE) method (Teunissen and Amiri-Simkooei, 2008). As the model can handle an arbitrary number of time series, it is suited for multi-instrument experiments.

The method is applied to an on-site field calibration experiment carried out at Aix Island, mid-Atlantic coast of France, where a permanent radar gauge has operated for several years (Gouriou, Míguez, and Wöppelmann, 2013), and various types of tide gauges (including some emerging technologies) were temporarily deployed during the experiment within meters from each other over a tidal cycle in 2016.

7.2 The Aix Island experiment

This experiment was carried out on June 7, 2016, by a team of scientists (see Acknowledgment section). For 11 hours, they recorded one semi-diurnal spring tidal cycle with a tidal range of 5.22 m using 6 different instruments:

- a permanent radar gauge (hereafter, RADAR),
- a permanent tide pole (hereafter, POLE),
- an electrical contact probe (hereafter, PROBE),
- a laser distance-meter (hereafter, LASER),
- 2 GNSS buoys (hereafter, BUOY1 and BUOY2).

RADAR, POLE, PROBE, and LASER are shown in Figure (7.1) and the two GNSS buoys in Figure (7.2). All tide gauges and the reference GNSS station were referenced to the station datum by leveling. Each gauge record is defined as an average over a 2 minutes acquisition window every 10 minutes.

The radar gauge (RADAR) is the primary tide gauge of the permanent sea level observatory of the Aix Island. This station contributes to the French sea level observation network (RONIM) operated by the French hydrographic service (SHOM). It is a Krohne Optiwave 7300C gauge that measures the air range between the transmitter fixed above the sea surface and the sea surface with a sampling frequency of 1 Hz using a frequency modulated continuous wave technology (IOC, 2016).

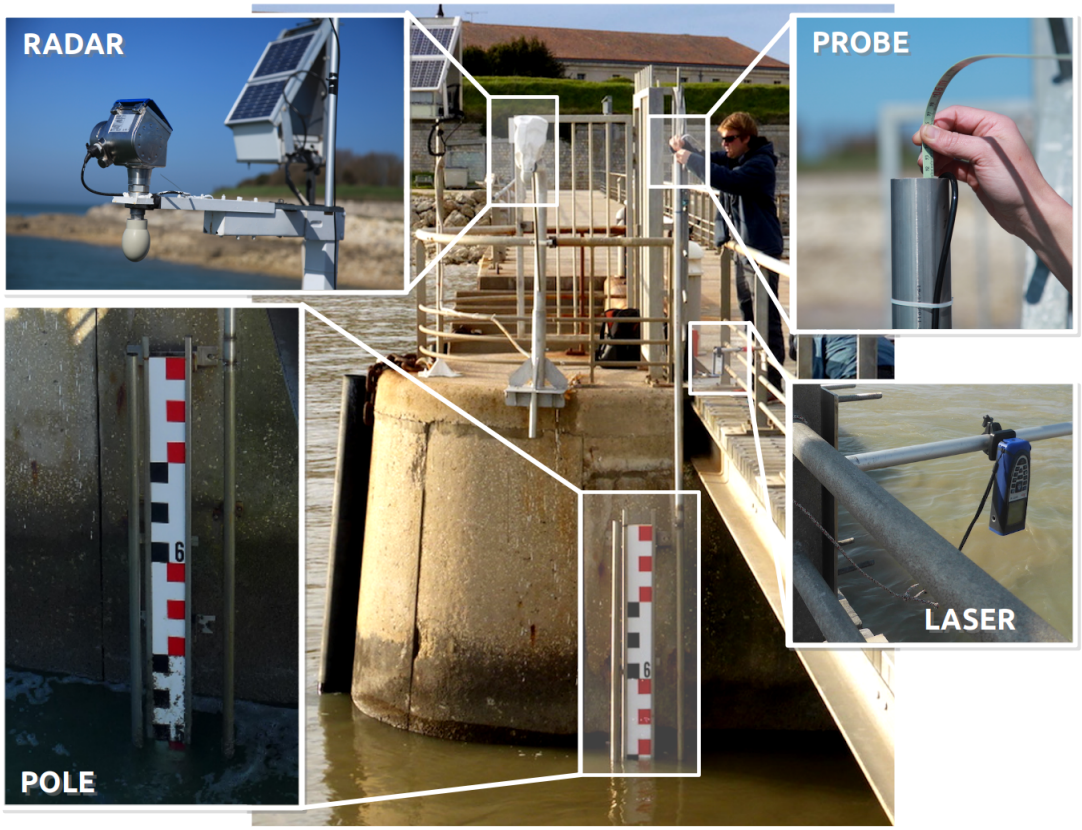

FIGURE 7.1: The four ground-based tide gauges: RADAR, POLE, PROBE and LASER.

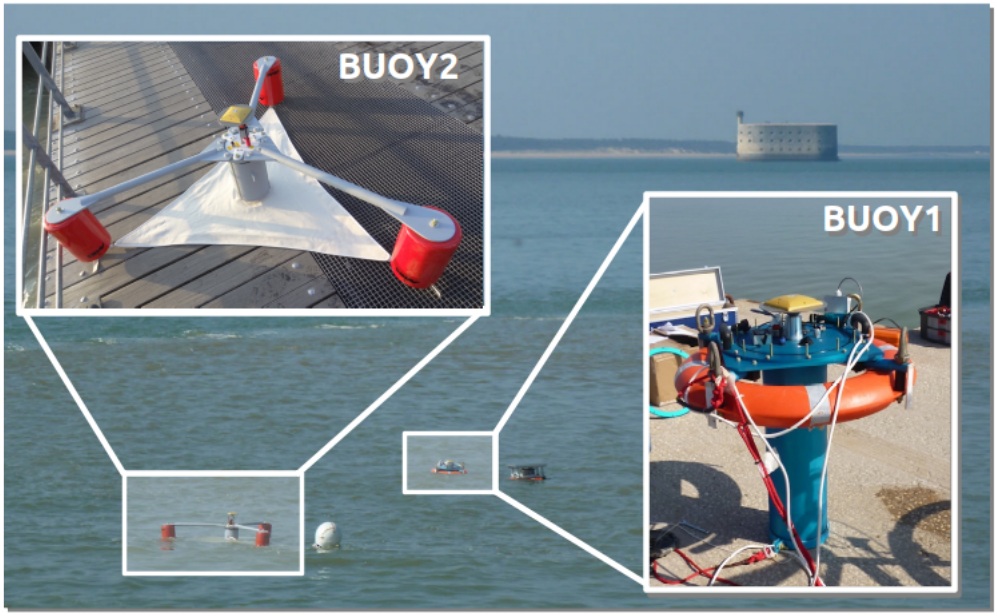

FIGURE 7.2: The two GNSS buoys: BUOY1 and BUOY2.

The tide pole (POLE) is a permanent instrument made of a plastic staff with graduations every 10 centimeters fixed vertically by a stainless steel structure (Figure 7.1). The operator estimates the sea level visually over the pre-defined 2 minutes acquisition period.

The electrical contact probe (PROBE) is a measuring tape with millimeter graduations ended by an electrical device that emits a short signal when detecting the seawater surface. We used a Schill probe installed within a stilling pipe anchored along the tide pole (Figure 7.1). A sea level record from PROBE is an average over the 2 minutes of human-made readings every 15 seconds. Electric probes are typically used as the reference gauge in tide gauge calibrations, so was it in our study. The stilling pipe was too short to allow measurements at the lowest sea levels, which resulted in a gap between 10:00 and 12:10 A.M.

The first GNSS buoy (BUOY1), designed at the Institut de Physique du Globe de Paris (IPGP), is a GNSS antenna installed above a lifebuoy and protected from the water by a radome (Figure 7.2). The second one (BUOY2), designed by the Division Technique de l'Institut National des Sciences de l'Univers (DT INSU), is a GNSS antenna housed in the center of a tripod floating structure (Figure 7.2). The receivers and batteries of the buoys are located inside a metallic cylinder under each antenna. These two buoys (BUOY1 and BUOY2) were already used in previous campaigns (André et al., 2013). The heights between their phase centers and the water surface are known at the sub-centimeter level thanks to previous testings carried out under calm conditions.

The buoy vertical positions, i.e., ellipsoidal heights, from GPS were assessed by post-processing, using a double-differences strategy with a baseline of about 300 m from the ILDX GNSS reference station. Only satellites with elevation angles above 15 degrees were used, with a combination of both L1 and L2 frequencies. The centimeter level accuracy was achieved, using full ambiguity resolution with the RTKlib software suite with RTKPOST v2.4.2 program (Takasu, 2013).

LASER is a reflector-free distance-meter Leica DISTO A6. This type of instrument is built for solid surface ranging but showed fair to good performances during this experiment. This instrument uses an optical laser beam with a wavelength of 635 nm. Each LASER record corresponds to an average of measurements done every 4 seconds.

All instruments time series are presented in Figure (7.3). Due to data transmission loss and GNSS recording issues during the experiment, some records from the LASER, BUOY1, and BUOY2 instruments are missing.

7.3 Calibration methods

This chapter proposes a combination method to go beyond the classical difference methods, allowing a better determination of the biases and their uncertainties. For comparison, we processed the time series using both the combination method and the classical difference method used by the hydrography community, the so-called Van de Casteele (VdC) diagram (Lennon, 1968).

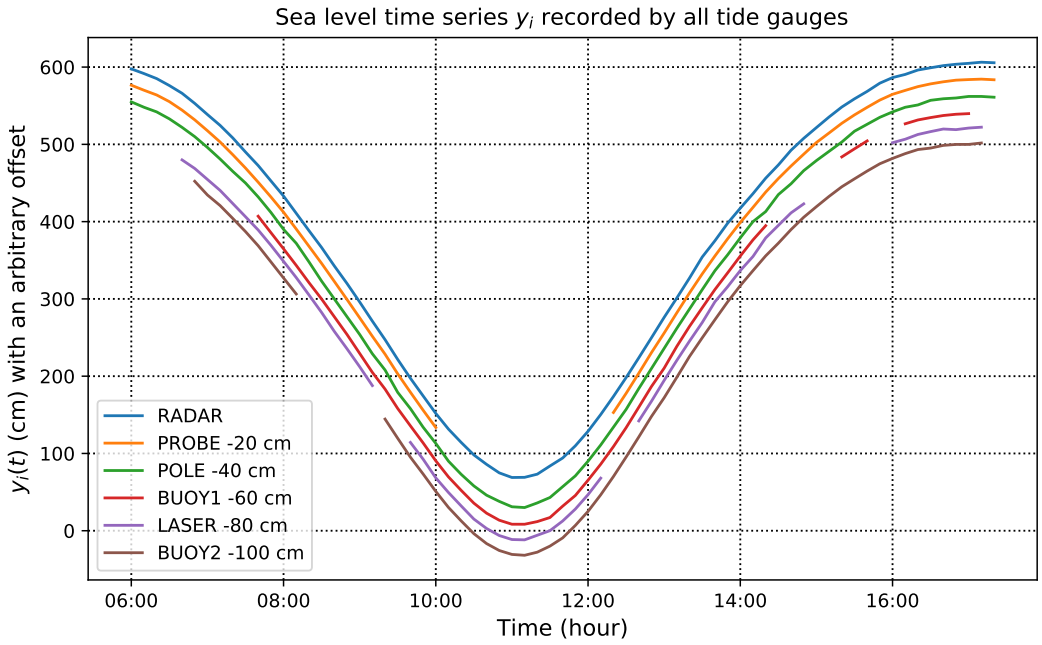

FIGURE 7.3: Sea level time series y_i recorded by all tide gauges.

7.3.1 Sea level error model

Due to the short recording period, this chapter only considers the influence of the three most common types of range measurement biases on the resulting sea level time series, namely: height reference, scale, and clock synchronization errors (Watson, Coleman, and Handsworth, 2008; Míguez Martín, Testut, and Wöppelmann, 2008).

While converting original range measurements into sea level time series, range biases turn into sea level biases that must be quantified and removed. This chapter proposes a linear sea level bias model, which expresses the sea level bias as a function of the measured sea level itself. More precisely, the model links the i-th sea level time series $\underline{y}_i(t)$ to the real sea level $h(t)$ through

$$\underline{y}_i(t) = h(t - \tau_i) + \beta_i \cdot \underline{y}_i(t) + \alpha_i + \underline{e}_i(t), \tag{7.1}$$

where $\beta_i \times \underline{y}_i(t) + \alpha_i$ is the linear sea level bias model, and $\underline{e}_i(t)$ is a random error modelled by a centered normal distribution of unknown variance σ_i^2.

In Equation (7.1), α_i corresponds to the intercept: a constant term representing the sea level bias when $\underline{y}_i(t) = 0$. It may result from a height reference error, but also from the influence of a scale error, as mentioned shown in chapter 6. β_i corresponds to the scale error: a multiplying factor that causes a sea level bias proportional to the tidal range. It can result from both instrument or installation defaults. Finally, τ_i is the time delay between different tide gauges: it results from clock synchronization issues.

The measured sea-level $\underline{y}_i(t)$ depends non-linearly on the time delay τ_i, which makes linear determinations, like the one proposed in this paper, impossible. However, it can be corrected before the other bias estimations, e.g., by computing the

delay that maximizes the cross-correlation between a tested signal and a reference signal. Obtaining τ_i by cross-correlation avoids any assumptions on the periodicity of the measured signal. In our case, the time delay estimation showed that the best correlation was achieved with no delays added i.e., $\tau_i = 0, \forall i$.

The sea-level bias model directly quantifies the amplitude of the bias associated with the measurement $\underline{y}_i(t)$. The correction of the sea level time series can be done after the calibration experiment by subtracting the estimated bias model from the measurements. This linear model can be adapted to other types of biases. For example, longer time series analysis (several days, months, or years) may require to consider time-dependent biases such as trends and jumps (Pytharouli, Chaikalis, and Stiros, 2018).

7.3.2 The difference-based calibration method (DIFF)

Difference-based methods (DIFF) consist in analyzing the differences $\underline{\Delta}_{y_i}(t)$

$$\underline{\Delta}_{y_i}(t) = \underline{y}_i(t) - \underline{y}_{\text{ref}}(t),$$

in which $\underline{y}_i(t)$ is the time series of a tested instrument

$$\underline{y}_i(t) = h(t) + \beta_i \cdot \underline{y}_i(t) + \alpha_i + \underline{e}_i(t),$$

and $\underline{y}_{\text{ref}}(t)$ is that of a reference instrument

$$\underline{y}_{\text{ref}}(t) = h(t) + \underline{e}_{\text{ref}}(t)$$

A commonly used tool for DIFF methods is the Van de Casteele (VdC) diagram, which represents the sea level difference $\Delta_{y_i}(t)$ as a function of $\underline{y}_i(t)$. Initially developed in 1962, for mechanical tide gauges (IOC, 1985), the VdC diagram is nonetheless still applicable for modern sea level measurement technologies (Míguez Martín, Testut, and Wöppelmann, 2008). The most attractive feature of this diagram is a fast, visual, detection of possible biases with only one tidal cycle. Figure (7.4) shows the sea level error patterns resulting from the most common range measurement errors (IOC, 1985).

In the presence of the linear biases mentioned before, $\underline{\Delta}_{y_i}(t)$ follows

$$\underline{\Delta}_{y_i}(t) = \beta_i \cdot y_i(t) + \alpha_i + \underline{e}_i(t) - \underline{e}_{\text{ref}}(t). \tag{7.2}$$

In other words, estimates of the sea level bias parameters α_i and β_i of Equation (7.1) can be obtained by a linear regression of $\underline{\Delta}_{y_i}(t)$ on $\underline{y}_i(t)$, which corresponds to fitting a linear model on a VdC diagram.

Assuming that both random errors $\underline{e}_i(t)$ and $\underline{e}_{\text{ref}}(t)$ are uncorrelated, the term $\underline{e}_i(t) - \underline{e}_{\text{ref}}(t)$ in Equation (7.2) follows a centered normal distribution with an unknown variance $\sigma_i^2 + \sigma_{\text{ref}}^2$. The merge of the random errors $\underline{e}_i(t)$ and $\underline{e}_{\text{ref}}(t)$ in the differences $\Delta_{y_i}(t)$ implies that, without assumption, the DIFF methods can only assess the variance $\sigma_i^2 + \sigma_{\text{ref}}^2$, which is just an upper bound to the tested gauge variance σ_i^2 (Lentz, 1993; Míguez Martín, Testut, and Wöppelmann, 2008; Pytharouli, Chaikalis, and Stiros, 2018). To separate σ_i^2 and σ_{ref}^2, an additional piece of information is needed, such as a third time series.

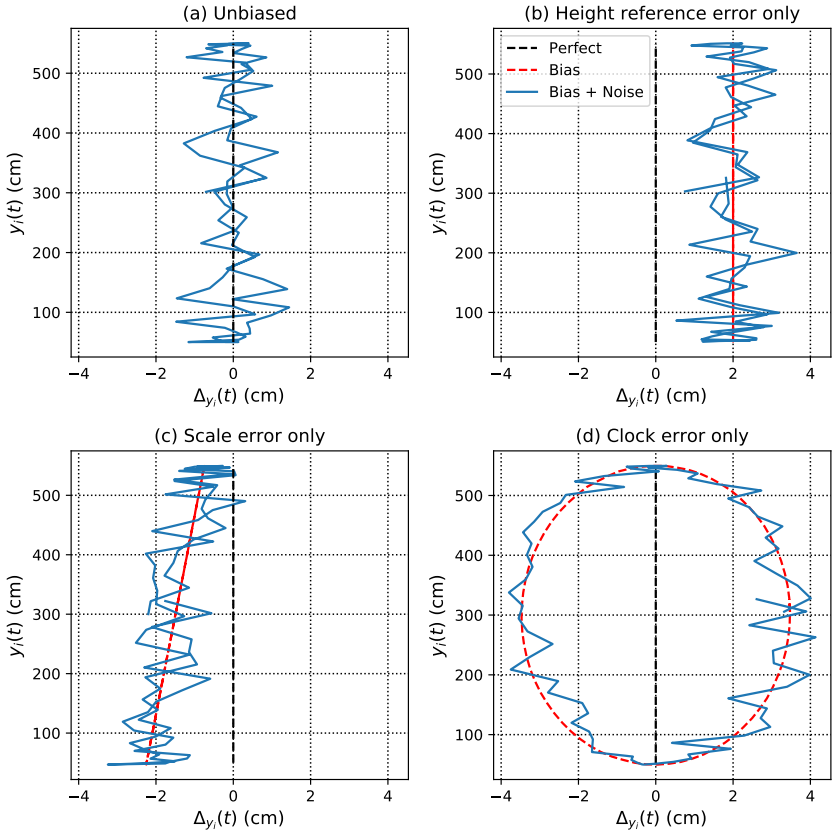

FIGURE 7.4: Synthetic examples of Van de Casteele diagrams for the most common types of range measurement errors: (a) random measurement errors only; (b) random measurement errors and a height reference error; (c) random measurement errors and a scale error; (d) random measurement errors and a clock error.

7.3.3 The combination-based calibration method (COMB)

When more than 2 time series are available, it becomes possible to assess the random errors and biases from each tide gauge by estimating a weighted combination of all the time series, using a variance component estimation method. In the following, the acronym COMB refers to the combination method.

Functional model

Noting $\underline{\mathbf{y}}_i$ the i-th gauge $k \times 1$ observation vector (or time series), the full $pk \times 1$ stacked vector $\underline{\mathbf{y}}$, containing all observations from the p instruments, can be written as

$$\underline{\mathbf{y}} = \begin{bmatrix} \underline{\mathbf{y}}_1 \\ \underline{\mathbf{y}}_2 \\ \vdots \\ \underline{\mathbf{y}}_p \end{bmatrix}.$$

The functional model links the expectation $\mathsf{E}\{.\}$ of the $pk \times 1$ observations vector $\underline{\mathbf{y}}$ to q unknown parameters using a model of observation equations . When there

is no theoretical model for the observed signal, one can estimate a $k \times 1$ combined solution $\mathbf{h}$ from the p measured time series, so that

$$\mathbf{h} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_k \end{bmatrix}.$$

In the case of unbiased gauges, the functional model would be $\mathsf{E}\{\underline{\mathbf{y}}_i\} = \mathbf{h}$ for each gauge. In the case of the cross-calibration of possibly biased time series, the functional model should also account for the biases:

$$\mathsf{E}\{\underline{\mathbf{y}}_i\} = \begin{cases} \mathbf{h}, & \text{if the } i\text{-th gauge is unbiased} \\ \mathbf{h} + \beta_i \cdot \underline{\mathbf{y}}_i + \alpha_i, & \text{otherwise} \end{cases}. \tag{7.3}$$

Because biases are always defined with respect to a conventional reference, at least one time series must be considered as conventionally unbiased to avoid an ill-posed equations system. Hence, in the following, the first time series $\mathbf{y}_1$ will be considered as conventionally unbiased.

This functional model can be written using matrix algebra, so that

$$\mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x} = \begin{bmatrix} \mathbf{A}_\mathbf{h} & \mathbf{A}_{\boldsymbol{\alpha}} & \mathbf{A}_{\boldsymbol{\beta}} \end{bmatrix} \begin{bmatrix} \mathbf{h} \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix}, \tag{7.4}$$

where $\mathbf{h} = \begin{bmatrix} h_1 \cdots h_k \end{bmatrix}^T$ is the combined solution vector, $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_2 \cdots \alpha_p \end{bmatrix}^T$ is the intercept parameter vector, and $\boldsymbol{\beta} = \begin{bmatrix} \beta_2 \cdots \beta_p \end{bmatrix}^T$ is the scale error parameter vector.

In Equation (7.4), the combination design $pk \times k$ matrix $\mathbf{A}_h$ corresponds to p stacked identity matrices $\mathbf{I}_{k\times k}$ such as

$$\mathbf{A}_h = \begin{bmatrix} \mathbf{I}_{k\times k} \\ \vdots \\ \mathbf{I}_{k\times k} \end{bmatrix},$$

and both the intercept design $pk \times (p-1)$ matrix $\mathbf{A}_\alpha$ and the scale error design $pk \times (p-1)$ matrix $\mathbf{A}_\beta$ are constituted with block non-zeros vectors so that

$$\mathbf{A}_\alpha = \begin{bmatrix} \mathbf{0}_{k\times 1} & \cdots & \cdots & \mathbf{0}_{k\times 1} \\ \mathbf{1}_{k\times 1} & \ddots & & \vdots \\ \mathbf{0}_{k\times 1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{k\times 1} \\ \mathbf{0}_{k\times 1} & \cdots & \mathbf{0}_{k\times 1} & \mathbf{1}_{k\times 1} \end{bmatrix},$$

and

$$\mathbf{A}_\beta = \begin{bmatrix} \mathbf{0}_{k\times 1} & \cdots & \cdots & \mathbf{0}_{k\times 1} \\ \mathbf{y}_2 & \ddots & & \vdots \\ \mathbf{0}_{k\times 1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{k\times 1} \\ \mathbf{0}_{k\times 1} & \cdots & \mathbf{0}_{k\times 1} & \mathbf{y}_p \end{bmatrix},$$

where $\mathbf{0}_{k\times 1}$, $\mathbf{1}_{k\times 1}$ refer to $k \times 1$ vectors respectively filled with zeros and ones.

Stochastic model

The stochastic model describes the variance $\mathrm{var}\{.\}$ of the observation vector $\mathbf{y}$. Considering that all measurements are statistically independent and that the uncertainty of the i-th instrument follows a multivariate normal distribution with a variance σ_i^2, the $pk \times pk$ covariance matrix of the observations $\mathrm{var}\{\mathbf{y}\} = \mathbf{Q}_\mathbf{y}$ reads

$$\mathbf{Q}_\mathbf{y} = \begin{bmatrix} \sigma_1^2\mathbf{I}_{k\times k} & \mathbf{0}_{k\times k} & \cdots & \cdots & \mathbf{0}_{k\times k} \\ \mathbf{0}_{k\times k} & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \sigma_i^2\mathbf{I}_{k\times k} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \mathbf{0}_{k\times k} \\ \mathbf{0}_{k\times k} & \cdots & \cdots & \mathbf{0}_{k\times k} & \sigma_p^2\mathbf{I}_{k\times k} \end{bmatrix}, \tag{7.5}$$

where $\mathbf{I}_{k\times k}$ and $\mathbf{0}_{k\times k}$ are respectively the $k \times k$ identity and null matrices.

To use the LS-VCE method, $\mathbf{Q}_\mathbf{y}$ needs to be expressed as a linear combination of cofactor matrices $\mathbf{Q}_i$ such as

$$\mathbf{Q}_\mathbf{y} = \sigma_1^2\mathbf{Q}_1 + \ldots + \sigma_p^2\mathbf{Q}_p = \sum_{i=1}^{p} \sigma_i^2\mathbf{Q}_i, \tag{7.6}$$

where the σ_i^2 are referred to as variance components, and correspond to the instrument uncertainties.

In this chapter, the $\mathbf{Q}_i$ are known $pk \times pk$ cofactor matrices that follow

$$\mathbf{Q}_1 = \begin{bmatrix} \mathbf{I}_{k\times k} & & & \\ & \mathbf{0}_{k\times k} & & \\ & & \ddots & \\ & & & \mathbf{0}_{k\times k} \end{bmatrix}$$

$$\mathbf{Q}_2 = \begin{bmatrix} \mathbf{0}_{k\times k} & & & \\ & \mathbf{I}_{k\times k} & & \\ & & \mathbf{0}_{k\times k} & \\ & & & \ddots \end{bmatrix}$$

$$\vdots$$

$$\mathbf{Q}_p = \begin{bmatrix} \mathbf{0}_{k\times k} & & & \\ & \ddots & & \\ & & \mathbf{0}_{k\times k} & \\ & & & \mathbf{I}_{k\times k} \end{bmatrix}$$

Least-Squares estimation

According to the least-squares estimation theory (Teunissen, 2000a), for normally distributed observations, an unbiased and minimum variance estimator $\hat{\mathbf{x}}$ of the $n \times 1$ parameter vector $\mathbf{x}$ can be obtained as

$$\underline{\hat{\mathbf{x}}} = (\mathbf{A}^T \mathbf{Q}_{\mathbf{y}}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_{\mathbf{y}}^{-1} \underline{\mathbf{y}}, \tag{7.7}$$

and its covariance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$ as

$$\mathbf{Q}_{\hat{\mathbf{x}}} = (\mathbf{A}^T \mathbf{Q}_{\mathbf{y}}^{-1} \mathbf{A})^{-1}. \tag{7.8}$$

In the case of a lack of knowledge on the on-site precision of the tide gauges, that is, on stochastic model $\mathbf{Q}_{\mathbf{y}}$, a variance component estimation method can be used to assess the uncertainty of each gauge. As the minimum variance property of least-squares estimates requires a realistic weighting between sea level time series, the use of a variance component estimation method also allows for more realistic estimates of the parameter vector $\underline{\hat{\mathbf{x}}}$ and its covariance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$.

Least-Squares Variance Component Estimation

A review of most variance component estimation methods can be found in Fotopoulos (2003) and Amiri-Simkooei (2007). Here, we consider the application of the Least-Squares Variance Component Estimation (LS-VCE), which is based on the same least-squares estimation principles used in section 7.3.3. LS-VCE was first introduced in 1988 by Teunissen (1988) and further developed by Amiri-Simkooei (2007) and Teunissen and Amiri-Simkooei (2008). Under the assumption of the multivariate normal distribution considered in section 7.3.3, the method provides an unbiased and minimum variance estimator of the variance components. It can also be shown that the LS-VCE estimates maximize the restricted likelihood function of the considered normal distribution (Amiri-Simkooei, 2007). This property is common to most rigorous VCE methods. However, the LS-VCE is more generally applicable and offers additional features, including a direct derivation of the uncertainty of each variance component estimate (Teunissen and Amiri-Simkooei, 2008).

The LS-VCE method consists in using the redundancy of information of a system to infer the variance of the observations. In the case of a linear parametric functional model, one can compute a residual $pk \times 1$ vector $\underline{\hat{\mathbf{e}}}$ such as

$$\begin{aligned} \underline{\hat{\mathbf{e}}} &= \underline{\mathbf{y}} - \mathbf{A}\underline{\hat{\mathbf{x}}} \\ &= \mathbf{P}_{\mathbf{A}}^{\perp} \underline{\mathbf{y}}, \end{aligned} \tag{7.9}$$

where $\mathbf{P}_{\mathbf{A}}^{\perp}$ is a projector matrix defined by

$$\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{Q}_{\mathbf{y}}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_{\mathbf{y}}^{-1}. \tag{7.10}$$

The residual vector $\hat{\mathbf{e}}$ gives pieces of information about observation uncertainties, potential model miss-specifications, and the presence of outliers. By assuming the absence of outliers and model miss-specifications, the LS-VCE provides an estimator of the observation uncertainties using $\underline{\hat{\mathbf{e}}}$ and $\mathbf{P}_{\mathbf{A}}^{\perp}$.

As for the standard least-squares estimation, the LS-VCE method estimates the unknown variance components $p \times 1$ vector

$$\sigma = \begin{bmatrix} \sigma_1^2 \\ \vdots \\ \sigma_p^2 \end{bmatrix}$$

by solving a normal equations system:

$$\underline{\hat{\sigma}} = \mathbf{N}^{-1}\underline{\mathbf{l}} = \begin{bmatrix} n_{11} & \cdots & \cdots & n_{1p} \\ \vdots & \ddots & & \vdots \\ & & n_{ij} & \\ \vdots & & \ddots & \vdots \\ n_{p1} & \cdots & \cdots & n_{pp} \end{bmatrix}^{-1} \begin{bmatrix} \underline{l}_1 \\ \vdots \\ \underline{l}_i \\ \vdots \\ \underline{l}_p \end{bmatrix}, \tag{7.11}$$

where the normal matrix $\mathbf{N}$ and the vector $\underline{\mathbf{l}}$ are specific to the stochastic model, and thus different from the normal matrix $\mathbf{N}$ and vector $\underline{\mathbf{l}}$ in Equation (7.7).

For the stochastic model defined in section 7.3.3, for which all variance components are to be estimated, the elements n_{ij} and $\underline{l}_i$ of $\mathbf{N}$ and $\mathbf{l}$ are defined by (Amiri-Simkooei, 2007) :

$$n_{ij} = \frac{1}{2}\mathrm{tr}\{\mathbf{Q}_i\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp\mathbf{Q}_j\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp\} \tag{7.12}$$

$$\underline{l}_i = \frac{1}{2}\underline{\hat{\mathbf{e}}}^T\mathbf{Q}_\mathbf{y}^{-1}\mathbf{Q}_i\mathbf{Q}_\mathbf{y}^{-1}\underline{\hat{\mathbf{e}}} \tag{7.13}$$

where $\mathrm{tr}\{.\}$ stands for the trace operator.

Note that $\underline{\hat{\sigma}}$ is involved in the definition of n_{ij} and $\underline{l}_i$ through $\mathbf{Q}_\mathbf{y}^{-1}$. Hence, Equation (7.11) expresses $\underline{\hat{\sigma}}$ as a function of $\mathbf{Q}_\mathbf{y}$, which is already a function of $\underline{\hat{\sigma}}$ in Equation (7.6). Such system of equations, where the equations for the unknowns include functions of the unknowns, can be numerically solved using an iterative procedure starting with an initial guess on the unknowns: the prior variance component vector $\sigma^{(0)}$.

The first iteration consists in using the prior vector $\sigma^{(0)}$ and cofactor matrices $\mathbf{Q}_i$ to compute $\mathbf{Q}_\mathbf{y}$ and $\mathbf{P}_\mathbf{A}^\perp$, which are necessary to build the normal equations system (7.11). Solving this normal equations system (7.11) leads to the estimation of an updated variance component vector $\underline{\sigma}^{(1)}$. The next k iterations consist in successively updating the variance component vector $\underline{\sigma}^{(k)}$ by solving the normal Equations system (7.11) built using the previously estimated variance component vector $\underline{\sigma}^{(k-1)}$. The iterations stop when the difference between two estimated variance component vectors becomes negligible. To obtain more details on the implementation of the LS-VCE method, we refer to chapter 3.

Once convergence is achieved, an insight into the quality the variance component estimates $\hat{\sigma}^2$ – the covariance matrix of the variance component estimates – can be obtained by inverting the normal matrix $\mathbf{N}$:

$$\mathbf{Q}_{\hat{\sigma}} = \mathbf{N}^{-1}, \tag{7.14}$$

The i-th diagonal element of $\mathbf{Q}_{\hat{\sigma}}$ corresponds to the variance of the i-th variance component $\underline{\hat{\sigma}}$. As for $\mathbf{Q}_{\hat{\mathbf{x}}}$, the uncertainties of variance component estimates depend on the system redundancy and the precision of the observations.

To get interpretable variance component estimates, one can change variance components $\hat{\sigma}_i^2$ into standard deviation components $\hat{\sigma}_i = \sqrt{\hat{\sigma}_i^2}$. To obtain variance component uncertainties with interpretable units, one can follow Amiri-Simkooei, Teunissen, and Tiberius (2009), and approximate the new variance of the standard deviation component $\sigma_{\hat{\sigma}_i}^2$ by applying variance propagation law through the linearized square root function:

$$\sigma_{\hat{\sigma}_i}^2 \approx \sigma_{\hat{\sigma}_i^2}^2 \cdot \left(\frac{1}{2\sqrt{\hat{\sigma}_i^2}} \right)^2, \tag{7.15}$$

The more interpretable standard deviation of the standard deviation component $\sigma_{\hat{\sigma}_i} = \sqrt{\sigma_{\hat{\sigma}_i}^2}$ can then be derived by taking the square root of both sides of Equation (7.15), which gives:

$$\sigma_{\hat{\sigma}_i} \approx \frac{\sigma_{\hat{\sigma}_i^2}}{2\hat{\sigma}_i}, \tag{7.16}$$

where $\sigma_{\hat{\sigma}_i^2}$ is the standard deviation of the i-th variance component $\sigma_{\hat{\sigma}_i^2} = \sqrt{\sigma_{\hat{\sigma}_i^2}^2}$.

Hence, one can express the uncertainty estimate of the i-th tide gauge as $\hat{\sigma}_i \pm \sigma_{\hat{\sigma}_i}$ [cm].

7.4 Results

To compare COMB and DIFF methods on a similar basis, the PROBE time series has been considered conventionally unbiased for both methods.

To remove the influence of potential outliers, residuals time series were computed using Equation (7.9) before the actual processing of both methods. The functional model (7.4) and the covariance matrix $\mathbf{Q_y} = \mathbf{I}$ were considered in Equation (7.10). Observations that showed residuals above five times the median absolute deviation of the gauge residual time series were removed from the data-set. In practice, it concerned less than 2 observations by time series.

7.4.1 Calibration with the combination (COMB) method

Before the assessment of the unknown bias parameters and the combined solution, a realistic covariance matrix $\mathbf{Q_y}$ was first computed using LS-VCE. An arbitrary standard deviation of 0.80 cm for all time series was used to build the prior variance component vector. Starting with $\boldsymbol{\sigma}^{(0)}$, the iterative procedure, summarized in section 7.3.3 and fully described in (Amiri-Simkooei, 2007), provided the final variance components vector estimate $\hat{\boldsymbol{\sigma}}$ and its covariance matrix $\mathbf{Q}_{\hat{\sigma}}$. As the elements of both $\hat{\boldsymbol{\sigma}}$ and $\mathbf{Q}_{\boldsymbol{\sigma}}$ are not directly interpretable, the Equation (7.16) was used to express each tide gauge uncertainty estimate as $\hat{\sigma}_i \pm \sigma_{\hat{\sigma}_i}$ (cm).

The bias parameters and the combined solution were estimated by solving the functional model (7.4) using the final variance component estimates: $\hat{\boldsymbol{\sigma}}$ was substituted in Equations (7.7) and (7.8) through Equation (7.6) to obtain the vector $\hat{\mathbf{x}}$ and its covariance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$.

Both estimated sea level bias parameters and uncertainties for 10 min records are given, in centimeter, in Table (7.1). The electrical PROBE is found to be the most

precise gauge in this experiment, with an uncertainty of 0.30 cm. The least precise gauges are the tide pole POLE (1.23 cm) and the BUOY1 (1.25 cm). BUOY1 is nearly two times less precise than BUOY2 (0.74 cm).

TABLE 7.1: Tide gauge cross-calibration results obtained using the COMB method. PROBE scale error and intercept are conventionally set to zero.

Gauges	$\hat{\alpha}_i \pm \sigma_{\hat{\alpha}_i}$ (cm)	$\hat{\beta}_i \pm \sigma_{\hat{\beta}_i}$ (cm m^{-1})	$\hat{\sigma}_i \pm \sigma_{\hat{\sigma}_i}$ (cm)
RADAR	-1.87 $\pm$ 0.30	0.52 $\pm$ 0.07	0.81 $\pm$ 0.08
PROBE	.	.	0.31 $\pm$ 0.10
POLE	-0.13 $\pm$ 0.39	-0.32 $\pm$ 0.09	1.23 $\pm$ 0.12
BUOY1	-4.30 $\pm$ 0.41	0.00 $\pm$ 0.11	1.25 $\pm$ 0.14
LASER	-3.42 $\pm$ 0.35	0.13 $\pm$ 0.08	0.90 $\pm$ 0.10
BUOY2	-3.53 $\pm$ 0.30	0.17 $\pm$ 0.07	0.74 $\pm$ 0.09

In Table (7.1), 4 time series – RADAR, LASER, BUOY1, and BUOY2 – show intercept estimates $\hat{\alpha}_i$ significant at the $3\sigma_{\hat{\alpha}_i}$ – or 99% – confidence level. Their amplitudes range from -1.87 cm (RADAR) to -4.30 cm (BUOY1). For the scale errors $\hat{\beta}_i$, only RADAR and POLE show estimates above $3\sigma_{\hat{\beta}_i}$, with about 0.50 cm m^{-1} and -0.30 cm m^{-1} respectively.

The residual time series of each tide gauge are presented in Figure (7.5). BUOY1 exhibits a mean shift of about -2.00 cm between 07:20 and 09:40. This artifact appears in the residual time series because it cannot result from the combination model. It means that the other gauges did not observe such a shift, otherwise, it would have been modelled by the combined solution. The presence of this artifact in the BUOY1's residual time series lowers its precision in Table (7.1). For the other gauges, no clear pattern appears in the residual time series, which suggests that their biases are correctly modelled.

The combined solution $\hat{h}$ and its uncertainty $\sigma_{\hat{h}}$ are presented in Figure (7.6). Each missing value in one of the time series increases the uncertainty of the combined solution to an extent proportional to its precision. The available measurements are displayed for each tide gauge, in the bottom of Figure (7.6). When the most precise tide gauge (PROBE) is not recording, between 10:00 and 12:10, the uncertainty $\sigma_{\hat{h}}$ of the combined solution increases by almost a factor of two. Despite the missing values of PROBE, the combined solution is estimated for the entire experiment period because all available observations are taken into account.

To investigate whether PROBE is found to be the most precise gauge because it is the conventionally unbiased gauge, the calibration has been reprocessed by instead considering BUOY1 as conventionally unbiased. The alternative calibration results are presented in Table (7.2). The choice of another conventionally unbiased gauge does not change uncertainty estimates but changes bias parameter estimates and their uncertainties. Bias parameters are the most affected because they intrinsically

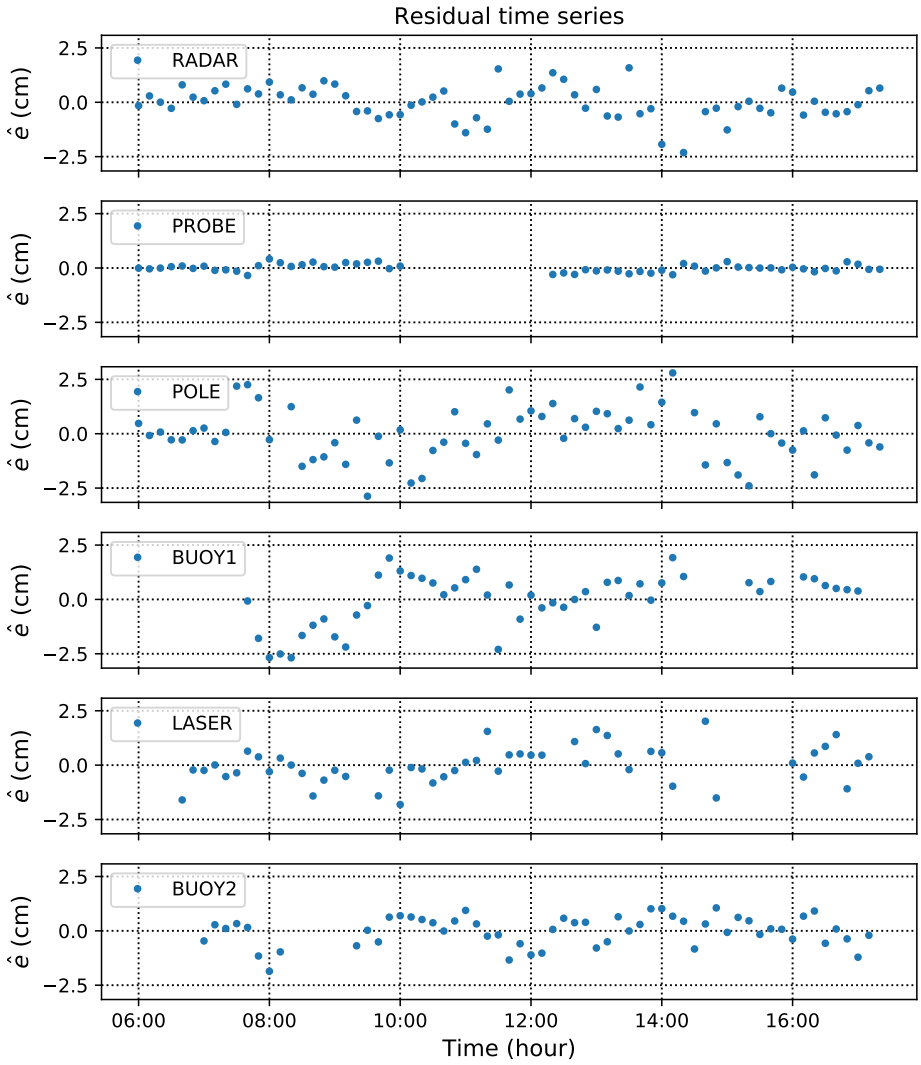

FIGURE 7.5: Residual time series of each tide gauge for the estimated linear combination model.

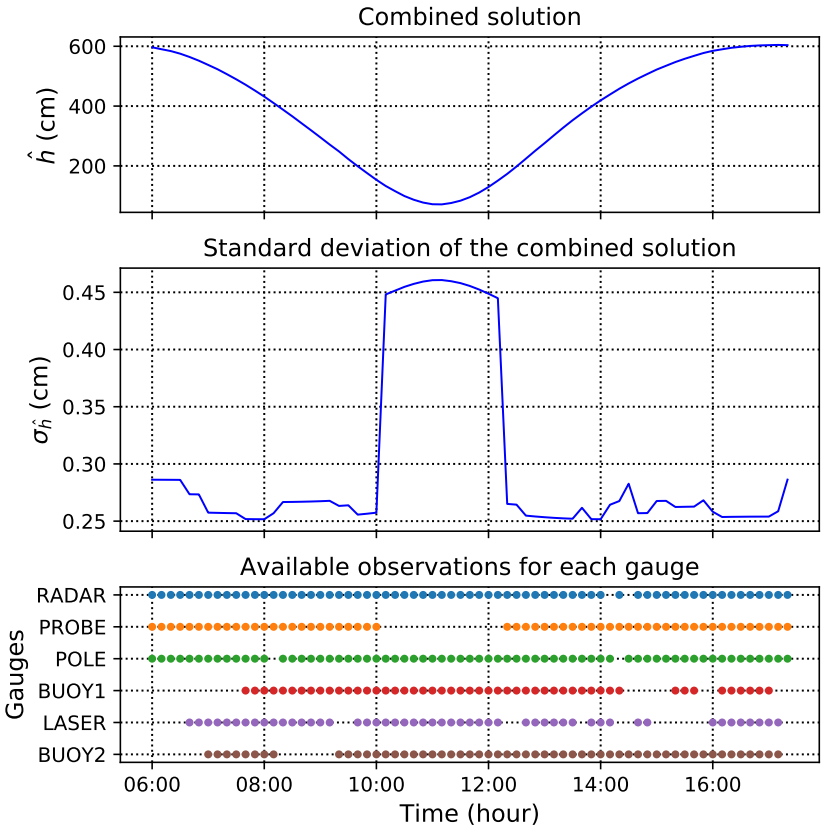

FIGURE 7.6: Combined solution (*top*), the standard deviation of the combined solution (*middle*) and available observations for each gauge (*bottom*).

depend on the definition of a convention. As BUOY1 does not exhibit any scale error in Table (7.1), the changes in scale error estimates in Table (7.2) are not dramatic. The sea level time series uncertainty estimates are identical in both alternatives because all biases are considered in each case. An alternative functional model ignoring an existing bias would not have provided identical results.

TABLE 7.2: Alternative tide gauge cross-calibration results obtained using the COMB method and by defining BUOY1 as the conventionally unbiased gauge. BUOY1 scale error and intercept are conventionally set to zero.

Gauges	$\hat{\alpha}_i \pm \sigma_{\hat{\alpha}_i}$ (cm)	$\hat{\beta}_i \pm \sigma_{\hat{\beta}_i}$ (cm m^{-1})	$\hat{\sigma}_i \pm \sigma_{\hat{\sigma}_i}$ (cm)
RADAR	2.34 $\pm$ 0.42	0.55 $\pm$ 0.11	0.81 $\pm$ 0.08
PROBE	4.18 $\pm$ 0.42	0.03 $\pm$ 0.11	0.31 $\pm$ 0.10
POLE	4.07 $\pm$ 0.49	-0.29 $\pm$ 0.13	1.22 $\pm$ 0.12
BUOY1	.	.	1.25 $\pm$ 0.14
LASER	0.72 $\pm$ 0.45	0.15 $\pm$ 0.12	0.90 $\pm$ 0.10
BUOY2	0.68 $\pm$ 0.42	0.19 $\pm$ 0.11	0.74 $\pm$ 0.09

7.4.2 Comparison with the difference (DIFF) method

Using PROBE as the reference gauge, we plotted the VdC diagram for RADAR, POLE, LASER, BUOY1, and BUOY2. A linear regression on each diagram provided intercept and scale error estimates for each gauge. The DIFF method estimates are presented in Table (7.3). The differences with the COMB method estimates are summarized in Table (7.4).

TABLE 7.3: Tide gauge calibration results obtained using the DIFF method. PROBE is the reference gauge.

Gauges	$\hat{\alpha}_i \pm \sigma_{\hat{\alpha}_i}$ (cm)	$\hat{\beta}_i \pm \sigma_{\hat{\beta}_i}$ (cm m^{-1})
RADAR	-1.54 $\pm$ 0.47	0.42 $\pm$ 0.10
PROBE	.	.
POLE	0.09 $\pm$ 0.66	-0.36 $\pm$ 0.14
BUOY1	-5.05 $\pm$ 0.72	0.15 $\pm$ 0.18
LASER	-3.07 $\pm$ 0.77	0.12 $\pm$ 0.17
BUOY2	-3.42 $\pm$ 0.47	0.18 $\pm$ 0.10

The discrepancies between COMB and DIFF methods reach 0.75 cm for the intercepts (BUOY1) and 0.15 cm m^{-1} for the scale errors (BUOY1). In Table (7.5),

TABLE 7.4: Difference between DIFF and COMB calibration results.

Gauges	$\Delta\hat{\alpha}_i$ (cm)	$\Delta\sigma_{\hat{\alpha}_i}$ (cm)	$\Delta\hat{\beta}_i$ (cm m^{-1})	$\Delta\sigma_{\hat{\beta}_i}$ (cm m^{-1})
RADAR	-0.33	-0.17	0.10	-0.03
PROBE	.	.	.	.
POLE	-0.22	-0.27	0.04	-0.05
BUOY1	0.75	-0.31	-0.15	-0.07
LASER	-0.35	-0.42	0.01	-0.09
BUOY2	-0.11	-0.17	-0.01	-0.03

changes in bias uncertainties between methods are expressed in terms of uncertainty reduction percentages. The DIFF method provides slightly different results from the COMB method because it only considers a smaller subset of the data-set for each pair of gauge and because it does not take into account the precision of each time series. In this chapter, the DIFF method can only take into account the overlapping observations between PROBE and the tested gauges. Given that PROBE has no observation between 10:00 and 12:10, the DIFF method ignores several observations, which deteriorates the precision of bias estimates. As a consequence, Table (7.5) shows that the COMB method provides 30% to 55% smaller uncertainties than the DIFF method for bias parameter estimates.

TABLE 7.5: Reduction of the standard deviations of the bias parameters obtained using the COMB method with respect to the DIFF method.

Gauge	$\Delta\sigma_{\hat{\alpha}_i}$ (%)	$\Delta\sigma_{\hat{\beta}_i}$ (%)
RADAR	-36	-30
PROBE	.	.
POLE	-41	-36
BUOY1	-43	-39
LASER	-55	-53
BUOY2	-36	-30

The presence of the scale error induces a height-dependency of the sea level bias models and their confidence intervals. To illustrate this, Figure (7.7) displays the estimated sea level bias models and their uncertainties, obtained with both methods, on the VdC diagram for BUOY1, which is the time series with the most substantial differences between the two models. At the lowest tide, sea level bias models obtained with COMB and DIFF method differs of about 3 millimeters. Besides, both

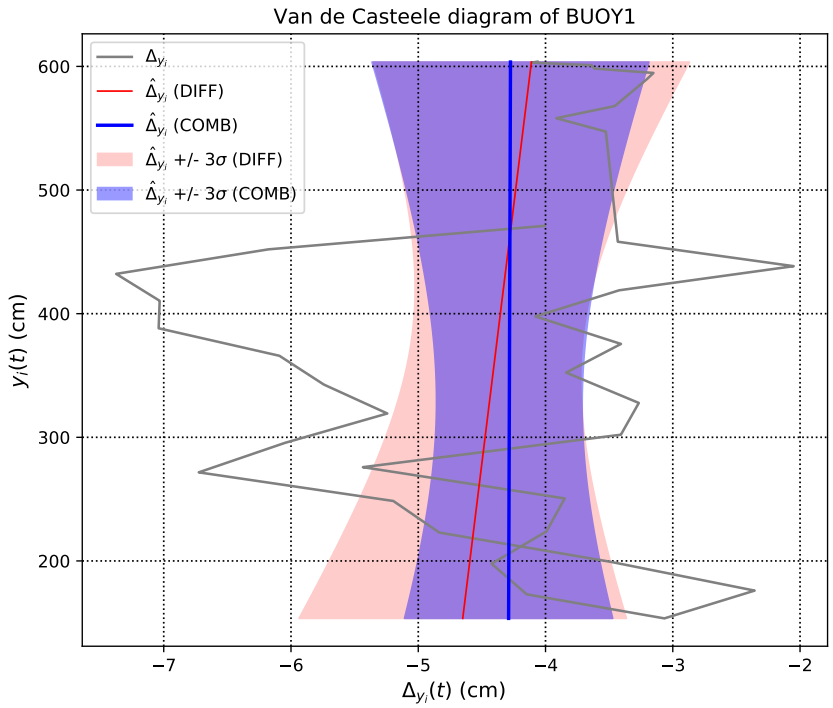

FIGURE 7.7: Van de Casteele diagram of BUOY1. The sea level bias model estimated with the COMB method is displayed in blue, and the one estimated with the DIFF method is displayed in red.

sea level bias models are more precise around the mean tide than near the tidal extrema. As a consequence, the combined solution of the COMB method is also less precise near the tidal extrema, which results in the few millimeter changes for $\sigma_{\hat{h}}$ that appears in Figure (7.6) at the lowest tide: between 10:00 and 12:10.

A representation of all bias estimates obtained with both DIFF and COMB methods is given in Figure (7.8). Bias estimates are shown as points in the bias parameter space - intercept vs scale error. Their uncertainties appear as 1σ confidence ellipses. The correlations between bias parameters, always around -0.9, induce an inclination of the ellipses. As the cause of the correlation is the same – same signal and same bias model – for every time series, so are the inclinations in Figure 7.8. Figure (7.8) also shows that, while providing more precise estimates, the COMB method still globally agrees with the DIFF method for bias detection.

7.5 Discussion

7.5.1 Performance of the tide gauges

The PROBE time series is twice more precise than that of BUOY1, the next most precise tide gauge. Its good performance results probably from the use of the stilling pipe, which stabilizes the water level and allows accurate readings on the measuring tape. This result comforts the use of electrical contact probes as references in tide gauge calibration campaigns.

The results also show that RADAR, LASER, and BUOY2 uncertainty estimates are below the centimeter level, which confirms that they could provide sea level

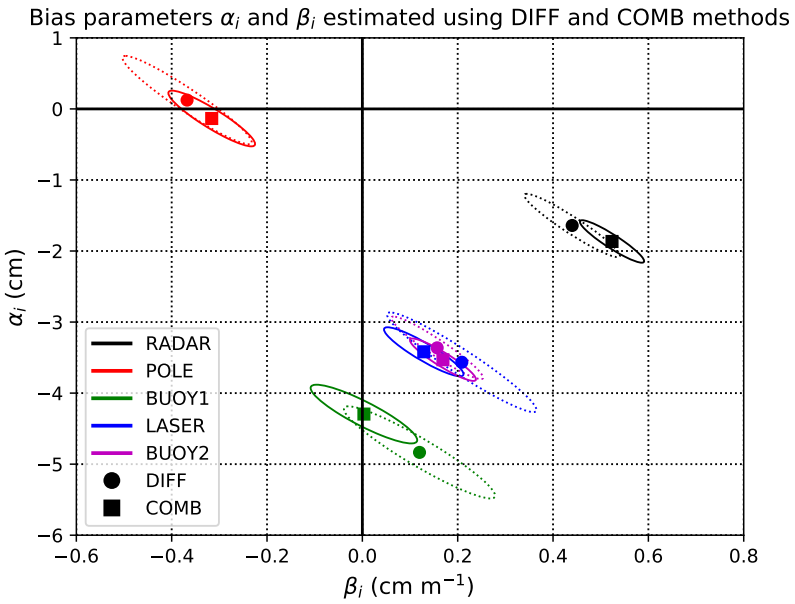

FIGURE 7.8: Representation of the bias parameter estimates in the parameter space for both COMB and DIFF methods.

records with the level of accuracy specified by the IOC with a confidence level of more than 67% if they were not affected by biases.

Among the 6 tested gauges in this work, only two, of which one automatic gauge, present an uncertainty above 1.00 cm: POLE (1.23 cm) and BUOY1 (1.25 cm). The 1.23 cm uncertainty of POLE might result from the limitation of human eye reading on the 10.00 cm graduations. The lower performance of BUOY1 compared to BUOY2 is assigned to the presence of the artifact between 07:20 and 09:40. Considering its floating structure is less stable than the more recent model BUOY2, this artifact could be due to the buoy instability in the presence of currents during the ebb tide. BUOY2 did not measure when BUOY1 observed the artifact; one cannot exclude that the artifact is due to a miss-modeling of the GNSS data.

7.5.2 Nature of the biases

Separating instrumental and environmental contributions in bias estimates is difficult, especially when the gauges are not fully collocated. We can nonetheless draw some hypotheses for bias attribution.

Usually, significant intercept estimates are due to instrumental height errors. But in this experiment, other explanations are plausible for BUOY1, BUOY2, LASER, and RADAR.

BUOY1 and BUOY2 show similar intercept estimates while being deployed a few tens of meters away from the ground-based instruments. Hence, changes in the dynamic topography due to currents likely impacted their intercept estimates (Pérez et al., 2014). In that case, an environmental effect is detected, not an instrumental bias.

As LASER is not dedicated to water surface measurements, the intercept estimate is certainly due to the penetration of the laser beam into the water. More appropriate laser systems have already been developed, using floating mirrors (MacAulay, O'Reilly, and Thompson, 2008).

For RADAR, the significant intercept estimate is influenced by the strong scale error. Theoretically, LASER, RADAR, and POLE could show scale error in the case of vertical alignment defaults. This cause is plausible for RADAR and LASER. However, the vertical alignment of POLE can be considered as reliable and the human-reading is the most likely source of its scale error.

Even though the nature of significant bias parameters α_i and β_i could remain unclear, one can still obtain corrected sea level time series by subtracting the bias model $\beta_i \times y_i(t) + \alpha_i$ to the measured sea level $y_i(t)$.

7.5.3 Improvement over difference based methods

The proposed calibration method provides an unbiased and minimum variance estimate of the tide gauge uncertainties, their sea level biases, and the combined solution from all time series. The variance of all estimates, including tide gauge uncertainties, are also determined. Thus, the COMB method leads to a more complete tide gauge calibration than the DIFF method.

The application to the Aix Island experiment revealed that the proposed methodology also leads to more precise bias estimates. This improvement is attributed to the combination of all available observations along with the realistic weighting between each gauge. The drastic precision improvement, from 30% to 55% on the uncertainty of the bias parameters, mostly shows that this method is more robust to the missing values of the most precise time series (PROBE), which is used as a reference to build the VdC diagrams.

For comparison purposes, this chapter considers only one conventionally unbiased time series. However, the COMB method allows using several unbiased time series and partially unbiased time series at the same time, which is not possible with the DIFF method. Adding unbiased time series should further improve the results of the COMB method.

7.6 Conclusion

The present contribution proposes a method for the cross-calibration of tide gauges. Based on the combination of multiple collocated time series, it takes advantage of the Least-Squares Variance Component Estimation method to assess both instrumental biases and measurement uncertainties in real conditions. The method was applied to a multi-instrument experiment carried out at Aix Island in 2016. Six instruments were deployed and performed simultaneous sea level recordings for 11 hours, with a 10 minutes sampling.

The electrical probe was found to be two to four times more precise than the other gauges. RADAR, LASER, and BUOY2 uncertainty estimates are below the centimeter level, which confirms that, in those conditions, they could provide sea level records with the level of accuracy specified by the IOC if they were not affected

by biases. We showed that, within our time series, significant bias parameters were found for all the tested gauges. Hence, this chapter shows that it is possible to assess both the biases and the precision – i.e. the full accuracy – for each gauge.

The results obtained with the combination method have been compared to that of a difference based method. It showed that the combination of all time series also provides more precise bias estimates.

Because this chapter is based on an 11 hours experiment, time-dependent biases and random errors have not been considered. Further studies using the COMB methods are necessary to investigate the time-dependency of sea level bias parameters and tide gauge precisions.

7.7 Acknowledgments

This study would not have been possible without the fieldwork carried out on the Aix island in 2016 by Laurent Testut, Guy Woppelmann, Valérie Ballu, Etienne Poirier, Médéric Gravelle, Álvaro Santamaría-Gómez, Mikaël Guichard, Elizabeth Prouteau, Pierre Chasseloup, Cyril Poitevin, Fabien Durand, Ronan Le Gall, and Pascal Le Dû.

Chapter 8

Investigating the potential of the nonlinear LS-VCE method for the stochastic modelling of GNSS position time series

The two previous chapters focused on the assessment of tide-gauge accuracy. In the next four chapters, we focus on the stochastic properties of GNSS position time series. The first step in the statistical modelling of GNSS position time series is the choice of the statistical method. In practice, the estimation of the functional parameters is usually done with the Best Linear Unbiased Estimator (BLUE) presented in chapter 2. On the other hand, the estimation of the stochastic parameters, or variance components, is usually done with the Maximum Likelihood Estimator (MLE) presented in chapter 3. Although the MLE method is easy to implement, it is not unbiased. Hence, is in this chapter, we investigate whether the nonlinear Least-Squares Variance Component Estimation (LS-VCE) method, also presented in chapter 3, which is both unbiased and minimum-variance, could be an interesting alternative to the MLE method for GNSS time series analysis.

The theoretical developments presented in this chapter to apply the nonlinear LS-VCE method are used in the chapter 9 for the study of the influence of offsets on the uncertainty of variance component estimates.

8.1 Introduction

The study of geodetic position time series plays a central role in the understanding of geophysical processes such as plate tectonics (Argus and Heflin, 1995; Kreemer, Blewitt, and Klein, 2014), glacial isostatic adjustment (Nocquet, Calais, and Parsons, 2005; Schumacher et al., 2018), surface mass distribution (Van Dam, Blewitt, and Heflin, 1994; Van Dam et al., 2001), and vertical land motions (Pfeffer and Allemand, 2016; Wöppelmann and Marcos, 2016).

To convert raw observations into usable parameters, the statistical modelling of geodetic time series is divided into two complementary models, namely the functional model and the stochastic model. The functional model describes deterministic behaviours identified, or expected, in the observations. This usually includes a trend, periodic patterns, discontinuities and possible transient deformations (Bevis and Brown, 2014). On the other hand, the stochastic model describes identified, or expected, stochastic variability of observations. It usually accounts for white noise

(WN) and time-correlated stochastic processes such as power-law (PL) (Press, 1978; Agnew, 1992; Zhang et al., 1997) or generalized Gauss-Markov processes (Langbein, 2004; Bos et al., 2020).

Power-law processes (PL) are stochastic processes characterized by their power-spectrum $P(f_k)$, which follows a power-law function of the time frequency f_k, such that

$$P(f_k) \approx P_0 \left(\frac{f_k}{f_0}\right)^{\kappa} \tag{8.1}$$

in which P_0 and f_0 are normalization constants, and κ is the spectral index, a scalar usually ranging between $[-3, 1]$ (Agnew, 1992) defining the nature of the power-law process. In GNSS position time series, several studies showed that the preferred stochastic model is a combination of a WN and a PL noise with a spectral index close to -1, that is close to a flicker noise (Zhang et al., 1997; Williams et al., 2004; Santamaría-Gómez et al., 2011).

Generalized Gauss-Markov (GGM) processes are stochastic processes introduced in Langbein (2004) showing power-law properties at high-frequencies and white noise properties at low frequencies. The transition between power-law and white noise behaviours is parametrized by an additional parameter denoted ϕ, which limits the energy of the process at low frequencies and makes the GGM process stationary. PL processes are special cases of GGM processes in the sense that a PL process is a GGM process whose transition towards white noise properties never occurs.

Stochastic model parameters, namely the amplitude of each process, the spectral index, and the parameter ϕ are generally referred to as the variance components. If unknown, they can be estimated from the observations using a Variance Component Estimation (VCE) method. The most widely used VCE method in time series analysis is the maximum likelihood estimation method, which benefits from several software implementations such as CATS (Williams, 2008), HECTOR (Bos et al., 2013) and est_noise (Langbein, 2017). Although the MLE method is widely used, it can only give asymptotically unbiased estimator of variance components because it does not account for the influence of the functional model on variance components estimates (Harville, 1977).

There also exist other VCE methods that directly provides unbiased estimates. We can mention the Restricted Maximum Likelihood Estimation (Koch, 1986), the Minimum Quadratic Unbiased Estimation (Rao, 1971), the Best Invariant Quadratic Unbiased Estimation, the Helmert method, and the Least-Squares Variance Component Estimation (Teunissen and Amiri-Simkooei, 2008). These VCE methods were developed based on different estimation criteria. However, in the case of normally distributed observations, they are mathematically equivalent (Teunissen and Amiri-Simkooei, 2008). For a review of existing VCE methods we refer to Fotopoulos (2003) and Amiri-Simkooei (2007).

In this chapter, we investigate the potential of the nonlinear Least-Squares Variance Component Estimation for the estimation of variance component in geodetic time series. In previous studies, the linear LS-VCE has shown to be useful for the unbiased and minimum-variance estimation of noise amplitudes of power-law processes with fixed spectral indices κ (Amiri-Simkooei, Tiberius, and Teunissen, 2007;

Amiri-Simkooei, 2009; Amiri-Simkooei, 2016). Here, by providing an expression for the partial derivatives of the stochastic model with respect to κ and ϕ, we extend the application of the LS-VCE to the estimation of the spectral index κ and the GGM parameter ϕ. This results in an algorithm providing unbiased and minimum-variance estimates of all variance components, along with their uncertainties.

8.2 Functional and stochastic model

The functional model describes the expectation $\mathsf{E}\{.\}$ of the $m \times 1$ observation vector $\underline{\mathbf{y}}$. For linear(-ized) models of observation equations, the functional model can be written as

$$\mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x}, \tag{8.2}$$

in which $\mathbf{A}$ is the known $m \times n$ design matrix, and $\mathbf{x}$ is the $n \times 1$ functional parameter vector.

The stochastic model describes the variance $\mathsf{var}\{.\}$ of the observation vector $\underline{\mathbf{y}}$. Power-law processes being a particular case of generalized Gauss-Markov processes, if we consider either a WN+PL or WN+GGM combination, the stochastic model follows:

$$\mathbf{Q}_{\mathbf{y}} = \sigma_w^2 \mathbf{I} + \sigma_{pl}^2 \mathbf{Q}_{\kappa\phi}, \tag{8.3}$$

in which σ_w^2 is the white noise amplitude, $\mathbf{I}$ is the $m \times m$ identity matrix, σ_{pl}^2 is the time-correlated process amplitude, and $\mathbf{Q}_{\kappa\phi}$ is the $m \times m$ covariance matrix associated to a GGM process of spectral index κ and parameter ϕ.

The covariance matrix $\mathbf{Q}_{\kappa\phi}$ in Equation (8.3) can be obtained as

$$\mathbf{Q}_{\kappa\phi} = \Delta_t^{-\kappa/2} \mathbf{L}_{\kappa\phi} \mathbf{L}_{\kappa\phi}^T \tag{8.4}$$

in which Δ_t is the sampling period in years, and $\mathbf{L}_{\kappa\phi}$ is the $m \times m$ transformation matrix defined as

$$\mathbf{L}_{\kappa\phi} = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{m-1} & \cdots & h_1 & h_0 \end{bmatrix}, \tag{8.5}$$

where the entries h_i are defined by the following recurrence expression (Langbein, 2004; Bos et al., 2020):

$$\begin{aligned} h_0 &= 1, \\ h_i &= \phi a_i h_{i-1}. \end{aligned} \tag{8.6}$$

with

$$a_i = 1 - \frac{\kappa}{2i} - \frac{1}{i} \tag{8.7}$$

When $\phi = 1$, the covariance matrix $\mathbf{Q}_{\kappa\phi}$ is that of a pure power-law process. When $\phi < 1$, it corresponds to that of a Generalized Gauss-Markov process.

8.3 Least-Squares Variance Component Estimation

Introduced by Teunissen (1988) and further developed by Teunissen and Amiri-Simkooei (2008) and Amiri-Simkooei (2007), under the normal distribution assumption the LS-VCE provides an unbiased and minimum-variance estimator of the variance components. The method also allows for a direct assessment of the variance components uncertainty.

Noting $\mathsf{E}\{.\}$ and $\mathsf{var}\{.\}$ the expectation and variance operators, the linear LS-VCE considers both a linear functional model $\mathsf{E}\{\underline{\mathbf{y}}\}$ and a linear stochastic model $\mathsf{var}\{\underline{\mathbf{y}}\} = \mathbf{Q}_\mathbf{y}$ such that

$$\mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x}, \quad \mathbf{Q}_\mathbf{y} = \mathbf{Q}_0 + \sum_{i=1}^{p} \sigma_i^2 \cdot \mathbf{Q}_i \tag{8.8}$$

where $\mathbf{Q}_0$ and $\mathbf{Q}_i$ are known $m \times m$ co-factor matrices, σ_i^2 are p unknown variance components, $\mathbf{A}$ is the $m \times n$ design matrix, and and $\mathbf{x}$ is the unknown $n \times 1$ parameter vector of the functional model.

The LS-VCE provide an estimator of the $p \times 1$ variance component vector

$$\underline{\hat{\sigma}} = \begin{bmatrix} \underline{\hat{\sigma}_1}^2 \\ \vdots \\ \underline{\hat{\sigma}_p}^2 \end{bmatrix} \tag{8.9}$$

by solving a normal equation system $\mathbf{N}\underline{\hat{\sigma}} = \underline{\mathbf{l}}$, in which the entries n_{ij} and $\underline{l}_i$ of the $p \times p$ normal matrix $\mathbf{N}$ and the $p \times 1$ vector $\underline{\mathbf{l}}$ are given by (Amiri-Simkooei, 2007)

$$n_{ij} = \frac{1}{2}\mathrm{tr}(\mathbf{Q}_i\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp\mathbf{Q}_j\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp), \tag{8.10}$$

$$\underline{l}_i = \frac{1}{2}(\underline{\hat{\mathbf{e}}}^T\mathbf{Q}_\mathbf{y}^{-1}\mathbf{Q}_i\mathbf{Q}_\mathbf{y}^{-1}\underline{\hat{\mathbf{e}}}) - \frac{1}{2}\mathrm{tr}(\mathbf{Q}_i\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp\mathbf{Q}_0\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp). \tag{8.11}$$

In Equations (8.10) and (8.11), $\underline{\hat{\mathbf{e}}}$ denotes the residual vector, $\mathbf{P}_\mathbf{A}^\perp$ is an orthogonal projector defined by $\mathbf{P}_\mathbf{A}^\perp = \mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{Q}_\mathbf{y}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{Q}_\mathbf{y}^{-1}$ so that $\underline{\hat{\mathbf{e}}} = \mathbf{P}_\mathbf{A}^\perp\underline{\mathbf{y}}$, and $\mathbf{I}$ is the $m \times m$ identity matrix.

Assuming that the cofactor matrices $\mathbf{Q}_0$ and $\mathbf{Q}_i$ are linearly independent, the unbiased and minimum-variance estimator of the variance components can be obtained by inverting the normal matrix $\mathbf{N}$ such that

$$\underline{\hat{\sigma}} = \mathbf{N}^{-1}\underline{\mathbf{l}}. \tag{8.12}$$

Since the unknown covariance matrix $\mathbf{Q}_\mathbf{y}$ is involved in the definition of n_{ij} and $\underline{l}_i$, the estimation is done in an iterative way. One starts with a first guess on the variance component vector $\underline{\hat{\sigma}}_0$ to approximate $\mathbf{Q}_\mathbf{y}$ and obtain n_{ij} and $\underline{l}_i$. σ_0 can be obtained from previous publications, or simplified VCE methods. By solving the normal equation system, a new variance component vector $\underline{\hat{\sigma}}_1$ can be obtained. Then, by iterating this procedure, each time a variance component $\underline{\hat{\sigma}}_i$ is estimated, a new estimate $\underline{\hat{\sigma}}_{i+1}$ can be obtained. The iterations stop when the difference between two successive variance component estimates become negligible. For more details about the implementation of the LS-VCE, we refer to Amiri-Simkooei, 2007 and chapter 3.

When the iterative process converges to a final variance component vector estimate $\underline{\hat{\sigma}}$, its covariance matrix $\mathbf{Q}_{\hat{\sigma}}$ can be obtained as

$$\mathbf{Q}_{\hat{\sigma}} = \mathbf{N}^{-1}. \tag{8.13}$$

8.3.1 Nonlinear LS-VCE

When the stochastic model depends non-linearly on one of the variance components, the estimation procedure must be adapted. In particular, the matrices $\mathbf{Q}_0$ and $\mathbf{Q}_i$ must be updated at each iteration k to handle the non-linearity of the problem.

The expression for the matrix $\mathbf{Q}_0^{(k)}$ at the k-th iteration can be obtained as (Amiri-Simkooei, 2007)

$$\mathbf{Q}_0^{(k)} = \mathbf{Q}_\mathbf{y}(\hat{\boldsymbol{\sigma}}^{(k)}) - \sum_{i=1}^{p} \frac{\partial \mathbf{Q}_\mathbf{y}(\boldsymbol{\sigma})}{\partial \sigma_i} \hat{\sigma}_i^{(k)} \Big|_{\hat{\boldsymbol{\sigma}}^{(k)}}, \tag{8.14}$$

and that of the matrices $\mathbf{Q}_i^{(k)}$ as

$$\mathbf{Q}_i^{(k)} = \frac{\partial \mathbf{Q}_\mathbf{y}(\boldsymbol{\sigma})}{\partial \sigma_i} \Big|_{\hat{\boldsymbol{\sigma}}^{(k)}}. \tag{8.15}$$

(Amiri-Simkooei, 2007) and chapter 3 provide a detailed description of the algorithm to implement the nonlinear LS-VCE. In the nonlinear case, variance components can represent virtually any stochastic parameter, and not necessarily process amplitudes. Besides, Equations (8.14) and (8.15) depend on the partial derivatives of the stochastic model. Hence, to apply the nonlinear LS-VCE method, it is necessary to express the partial derivatives of the stochastic model with respect to all variance components.

8.3.2 Non-negative LS-VCE

The LS-VCE method being unconstrained, it can occasionally provide negative estimates for some variance components. When the concerned variance component is a noise amplitude, this is problematic because the result cannot be interpreted physically. To overcome the negativity issue, Amiri-Simkooei (2016) proposes to use the non-negative sequential coordinate-wise algorithm developed by (Franc, Hlaváč, and Navara, 2005) to solve the LS-VCE normal equation system at each iteration.

The non-negativity constraint conflicts with the unbiased property of the LS-VCE estimator. Consequently, the use of the non-negative LS-VCE method should ideally be used only in the case in which the unconstrained LS-VCE failed to provide positive estimates.

8.4 Partial derivatives of the stochastic model

8.4.1 Partial derivatives with respect to σ_w^2 and σ_{pl}^2

As the covariance matrix $\mathbf{Q}_\mathbf{y}$ depends linearly on the stochastic parameters σ_w^2 and σ_{pl}^2, one can show that the partial derivatives of the stochastic model with respect to σ_w^2 and σ_{pl}^2 follows

$$\frac{\partial}{\partial \sigma_w^2}\{\mathbf{Q_y}\} = \mathbf{I}, \tag{8.16}$$

and

$$\frac{\partial}{\partial \sigma_{pl}^2}\{\mathbf{Q_y}\} = \mathbf{Q}_{\kappa\phi}. \tag{8.17}$$

Since the stochastic parameters κ and ϕ are involved in the definition of the matrix $\mathbf{Q}_{\kappa\phi}$ in Equation (8.6), the computation of the corresponding partial derivatives is more intricate.

8.4.2 Partial derivative with respect to κ

To study the partial derivative of the stochastic model with respect to κ and ϕ, we denote $\underline{r}_i$ the element of a discrete GGM process at a time index i. For $j \geq i$, the auto-covariance between two elements $\underline{r}_i$ and $\underline{r}_j$ can be written as (Bos et al., 2008)

$$\mathrm{cov}\{\underline{r}_i, \underline{r}_j\} = \sigma_{pl}^2 \Delta_t^{-\kappa/2} \sum_{k=0}^{i} h_k h_{k+(j-i)}. \tag{8.18}$$

The partial derivative of $\mathrm{cov}\{\underline{r}_i, \underline{r}_j\}$ with respect to κ can thus be written as

$$\frac{\partial}{\partial \kappa}\{\mathrm{cov}\{\underline{r}_i, \underline{r}_j\}\} = \sigma_{pl}^2 \frac{\partial}{\partial \kappa}\{\Delta_t^{-\kappa/2} \sum_{k=0}^{i} h_k h_{k+(j-i)}\} \tag{8.19}$$

and developed into

$$\frac{\partial}{\partial \kappa}\{\mathrm{cov}\{\underline{r}_i, \underline{r}_j\}\} = \sigma_{pl}^2 \left[\frac{\partial}{\partial \kappa}\{\Delta_t^{-\kappa/2}\} \sum_{k=0}^{i} h_k h_{k+(j-i)} + \Delta_t^{-\kappa/2} \frac{\partial}{\partial \kappa}\{\sum_{k=0}^{i} h_k h_{k+(j-i)}\}\right] \tag{8.20}$$

The two elements to be computed are $\frac{\partial}{\partial \kappa}\{\Delta_t^{-\kappa/2}\}$ and $\frac{\partial}{\partial \kappa}\{\sum_{k=0}^{i} h_k h_{k+(j-i)}\}$. For the computation of $\frac{\partial}{\partial \kappa}\{\Delta_t^{-\kappa/2}\}$, we obtain

$$\frac{\partial}{\partial \kappa}\{\Delta_t^{-\kappa/2}\} = -\frac{1}{2}\ln(\Delta_t)\Delta_t^{-\kappa/2}. \tag{8.21}$$

For the computation of $\frac{\partial}{\partial \kappa}\{\sum_{k=0}^{i} h_k h_{k+(j-i)}\}$, using the linearity of the derivative one can write

$$\frac{\partial}{\partial \kappa}\{\sum_{k=0}^{i} h_k h_{k+(j-i)}\} = \sum_{k=0}^{i} \frac{\partial}{\partial \kappa}\{h_k h_{k+(j-i)}\} \tag{8.22}$$

which can be developed into

$$\frac{\partial}{\partial \kappa}\{\sum_{k=0}^{i} h_k h_{k+(j-i)}\} = \sum_{k=0}^{i} \frac{\partial}{\partial \kappa}\{h_k\} h_{k+(j-i)} + \sum_{k=0}^{i} h_k \frac{\partial}{\partial \kappa}\{h_{k+(j-i)}\} \tag{8.23}$$

.

To simplify the notation, we denote

$$u_k = \frac{\partial}{\partial \kappa}\{h_k\} \tag{8.24}$$

so that

$$\frac{\partial}{\partial \kappa}\{\sum_{k=0}^{i} h_k h_{k+(j-i)}\} = \sum_{k=0}^{i} u_k h_{k+(j-i)} + \sum_{k=0}^{i} h_k u_{k+(j-i)}. \tag{8.25}$$

By substituting Equation (8.21) and Equation (8.25) in Equation (8.20) one obtains

$$\frac{\partial}{\partial \kappa}\{\mathrm{cov}\{\underline{r}_i, \underline{r}_j\}\} = \sigma_{pl}^2 \Delta_t^{-\kappa/2} \times \left[-\frac{1}{2}\ln(\Delta_t)\sum_{k=0}^{i} h_k h_{k+(j-i)} + \sum_{k=0}^{i} u_k h_{k+(j-i)} + \sum_{k=0}^{i} h_k u_{k+(j-i)}\right], \tag{8.26}$$

Equation (8.26) can be re-written using matrix notation as

$$\boxed{\frac{\partial}{\partial \kappa}\{\mathbf{Q_y}\} = \sigma_{pl}^2 \Delta_t^{-\kappa/2}\left[-\frac{1}{2}\ln(\Delta_t)\mathbf{L}_{\kappa\phi}\mathbf{L}_{\kappa\phi}^T + \mathbf{U}_{\kappa\phi}\mathbf{L}_{\kappa\phi}^T + \mathbf{L}_{\kappa\phi}\mathbf{U}_{\kappa\phi}^T\right]}, \tag{8.27}$$

in which $\mathbf{U}_{\kappa\phi}$ is a lower triangular $m \times m$ matrix defined as

$$\mathbf{U}_{\kappa\phi} = \begin{bmatrix} u_0 & 0 & \cdots & 0 \\ u_1 & u_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ u_{m-1} & \cdots & u_1 & u_0 \end{bmatrix}. \tag{8.28}$$

The entries u_k in $\mathbf{U}_{\kappa\phi}$ can be obtained by computing

$$u_k = \frac{\partial}{\partial \kappa}\{\phi a_k h_{k-1}\} \tag{8.29}$$

which leads to the recurrence expression

$$\begin{aligned} u_0 &= 0, \\ u_k &= -\frac{\phi}{2k} h_{k-1} + \phi a_k u_{k-1}. \end{aligned} \tag{8.30}$$

Note that, when using non-negativity constraints, the use of the spectral index κ is problematic because it is negative in geodetic applications. To overcome this issues, one can estimate $\alpha = -\kappa$ instead of κ. In this case, the partial derivative with respect to α is simply

$$\frac{\partial}{\partial \alpha}\{\mathbf{Q_y}\} = -\frac{\partial}{\partial \kappa}\{\mathbf{Q_y}\} \tag{8.31}$$

8.4.3 Partial derivative with respect to ϕ

The derivation of partial derivative with respect to ϕ is similar to that with respect to κ. We start by writing the partial derivative of $\mathrm{cov}\{\underline{r}_i, \underline{r}_j\}$ with respect to ϕ as

$$\frac{\partial}{\partial \phi}\{\mathrm{cov}\{\underline{r}_i, \underline{r}_j\}\} = \sigma_{pl}^2 \Delta_t^{-\kappa/2} \frac{\partial}{\partial \phi}\{\sum_{k=0}^{i} h_k h_{k+(j-i)}\} \tag{8.32}$$

and develop it into

$$\frac{\partial}{\partial \phi}\{\mathrm{cov}\{\underline{r}_i, \underline{r}_j\}\} = \sigma_{pl}^2 \Delta_t^{-\kappa/2} \left[\sum_{k=0}^{i} \frac{\partial}{\partial \phi}\{h_k\} h_{k+(j-i)} + \sum_{k=0}^{i} h_k \frac{\partial}{\partial \phi}\{h_{k+(j-i)}\} \right]. \tag{8.33}$$

To simplify the notation, we denote

$$v_k = \frac{\partial}{\partial \phi}\{h_k\} \tag{8.34}$$

so that

$$\frac{\partial}{\partial \phi}\{\mathrm{cov}\{\underline{r}_i, \underline{r}_j\}\} = \sigma_{pl}^2 \Delta_t^{-\kappa/2} \left[\sum_{k=0}^{i} v_k h_{k+(j-i)} + \sum_{k=0}^{i} h_k v_{k+(j-i)} \right]. \tag{8.35}$$

Equation (8.35) can be re-written using matrix notation so that

$$\boxed{\frac{\partial}{\partial \phi}\{\mathbf{Q}_\mathbf{y}\} = \sigma_{pl}^2 \Delta_t^{-\kappa/2} \left[\mathbf{V}_{\kappa\phi} \mathbf{L}_{\kappa\phi}^T + \mathbf{L}_{\kappa\phi} \mathbf{V}_{\kappa\phi}^T \right]}, \tag{8.36}$$

in which $\mathbf{V}_{\kappa\phi}$ is a lower triangular $m \times m$ matrix defined as

$$\mathbf{V}_{\kappa\phi} = \begin{bmatrix} v_0 & 0 & \cdots & 0 \\ v_1 & v_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ v_{m-1} & \cdots & v_1 & v_0 \end{bmatrix}. \tag{8.37}$$

The entries v_k of the matrix $\mathbf{V}_{\kappa\phi}$ can be obtained by computing

$$v_k = \frac{\partial}{\partial \phi}\{\phi a_k h_{k-1}\} \tag{8.38}$$

which leads to the following recurrence expression:

$$\begin{aligned} v_0 &= 0, \\ v_k &= a_k h_{k-1} + \phi a_k v_{k-1}. \end{aligned} \tag{8.39}$$

The parameter ϕ is already a non-negative parameter, however it has an upper-bound at $\phi = 1$. The LS-VCE being unconstrained, this upper-bound can be problematic for numerical evaluation. Hence, instead of estimating ϕ, it is numerically more convenient to estimate another non-negative parameter, noted τ, which is the time-constant of the generalized Gauss Markov process (Langbein, 2004). The parameters τ and ϕ are linked through

$$\tau = -\frac{\Delta_t}{\ln \phi}. \tag{8.40}$$

The partial derivative of the stochastic model with respect to τ is obtained from Equation (8.36) as

$$\frac{\partial}{\partial \tau}\{\mathbf{Q}_\mathbf{y}\} = \frac{\partial}{\partial \phi}\{\mathbf{Q}_\mathbf{y}\} \frac{\partial}{\partial \tau}\{\phi\}, \tag{8.41}$$

with

$$\frac{\partial}{\partial \tau}\{\phi\} = \exp\left\{-\frac{\Delta_t}{\tau}\right\}\frac{\Delta_t}{\tau^2}. \tag{8.42}$$

8.5 Performance on synthetic time series

In this section, we test the performance of the nonlinear LS-VCE method for the variance component estimation of both WN+PL and WN+GGM models.

8.5.1 WN+PL model

To investigate whether the LS-VCE is able to retrieve both reliable process amplitudes and the spectral index estimates for WN+PL models, we evaluate the method on 2000 synthetic time series with daily observations and a span of 8 years. For each synthetic time series, both the unconstrained and the non-negative LS-VCE are applied.

The simulated deterministic component is a linear trend with annual and semi-annual periodic signals. The simulated stochastic component is a combination of a white noise and a power-law process. The true values as well as the chosen approximate – that is, initial – variance components used for the simulations are presented in Table (8.1).

TABLE 8.1: True and approximate (initial) variance components used for the simulation

Parameters	True	Initial
σ_w^2 (mm^2)	9	100
σ_{pl}^2 (mm^2 yr$^{-0.4}$)	225	100
κ	-0.8	-1

The distributions of the variance component estimates obtained with the unconstrained LS-VCE method are presented in Figure (8.1). The unconstrained LS-VCE algorithm converged toward a solution for 1960 time series out of 2000, which correspond to as success rate of 98%, although the initial values were far from the true values.

TABLE 8.2: Average variance component estimates for the unconstrained LS-VCE (LS-VCE) and the non-negative LS-VCE (NNLS-VCE).

Parameters	True	LS-VCE	NNLS-VCE
σ_w^2 (mm^2)	9	8.56	7.25
σ_{pl}^2 (mm^2 yr$^{\kappa/2}$)	225	224.21	218.54
κ	-0.8	-0.80	-0.77

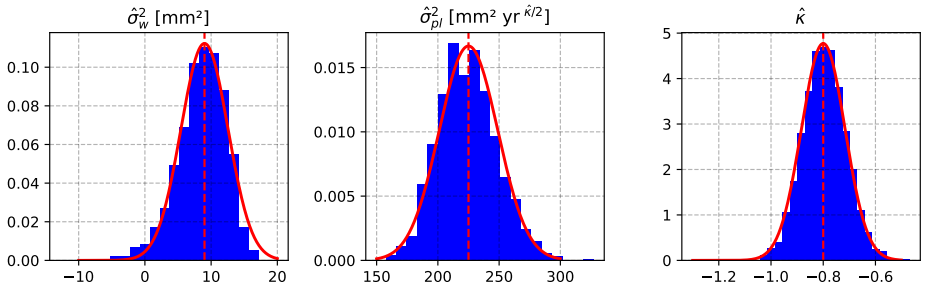

FIGURE 8.1: Distribution of variance component estimates obtained using the unconstrained nonlinear LS-VCE method for the WN+PL stochastic model (*blue histogram*). True variance component (*dashed red line*). Normal distribution centred on the true value, with variance computed using Equation (8.13) based on the true variance components in Table (8.1) (*solid red line*).

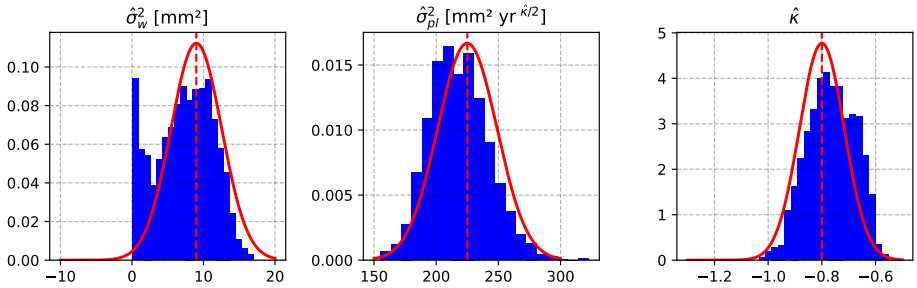

FIGURE 8.2: Distribution of variance component estimates obtained using the non-negative nonlinear LS-VCE method for the WN+PL stochastic model (*blue histogram*). True variance component (*dashed red line*). Normal distribution centred on the true value, with variance computed using Equation (8.13) based on the true variance components in Table (8.1) (*solid red line*).

The distributions of the variance component estimates obtained with the NNLS-VCE method are presented in Figure (8.2). The constrained LS-VCE algorithm converged toward a solution for 1981 time series out of 2000, which corresponds to a success rate of 99.05 %. This success rate is slightly better than that of the unconstrained algorithm. The average estimates for each variance component are presented in Table (8.2). In this example, the nonlinear NNLS-VCE slightly underestimates process amplitudes and overestimates the spectral index. This illustrates that the non-negativity introduces a bias. This bias is also visible in the distributions presented in Figure (8.2).

Our simulations show that both the unconstrained and the non-negative nonlinear LS-VCE method can provide reliable estimates of the variance components. To consider their use in other studies, it is also useful to investigate the speed of the method for time series with different time spans.

Therefore, we tested the runtime of the unconstrained nonlinear LS-VCE algorithm for time series with daily observations and a time span ranging between 4 and 20 years. The simulated noise properties and initial variance components are

the same as in Table (8.1). The algorithm was implemented in the Python 3 programming language. Algebraic operations were processed using the Numpy library (Harris et al., 2020) and parallelized on 4 CPUs Intel® Core™ i7-6500U (2.50 GHz).

The evolution of the algorithm runtime as a function of the span of the daily time series is presented in Figure (8.3). The nonlinear LS-VCE method provides variance components estimates in about 8 seconds for 4 years long time series, and in about 10 minutes for 20 years long time series, which is reasonably fast. As a result, the speed nonlinear LS-VCE method does not seem prohibitive for practical applications. More investigations are required to define whether the nonlinear LS-VCE method provides a significant increase in speed compared to other software.

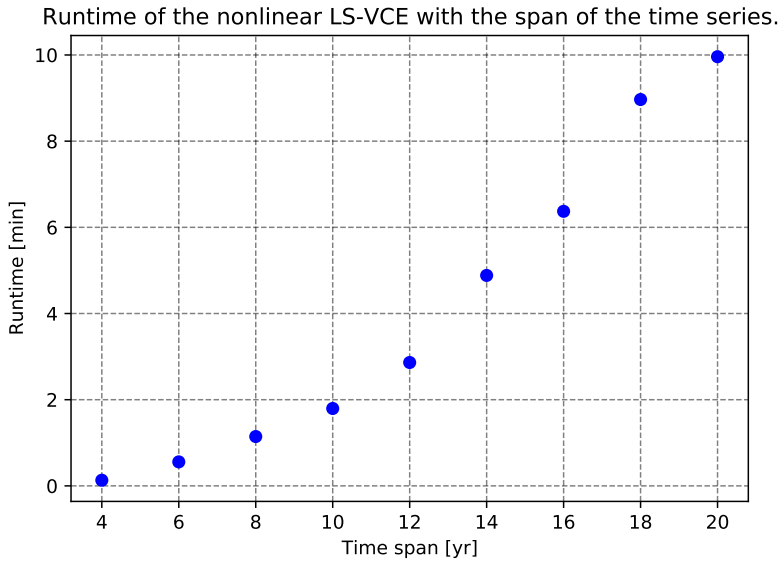

FIGURE 8.3: Runtime of the unconstrained nonlinear LS-VCE method with the time span of the time series.

8.5.2 WN+GGM model

The differences between a WN+GGM and a WN+PL model are expected to be noticeable when analyzing long time series (He et al., 2019). Therefore, to investigate whether the nonlinear LS-VCE method can retrieve reliable process amplitude, spectral index, and Gauss-Markov parameter estimates for WN+GGM models, in this section, we test the method on 200 synthetic time series with a span of 20 years. For each synthetic time series, we apply both the unconstrained and the non-negative nonlinear LS-VCE method. The true values, as well as the chosen approximate variance components used for the simulations, are presented in Table (8.3).

The distributions of the variance component estimates obtained using the unconstrained nonlinear LS-VCE method are presented in Figure (8.4). The unconstrained LS-VCE algorithm converged toward a solution for only 96 time series out of 200, which corresponds to a success rate of only 48%. Hence, although it provides estimates within the expected uncertainty in case of convergence, the unconstrained nonlinear LS-VCE method cannot offer stable results with the chosen initial values.

TABLE 8.3: True and approximate (initial) variance components used for the simulation

Parameters	True	Initial
σ_w^2 (mm^2)	25	100
σ_{ggm}^2 (mm^2 yr$^{\kappa/2}$)	400	225
κ	-1.50	-1.20
τ	0.09	0.15
(ϕ)	(0.970)	(0.981)

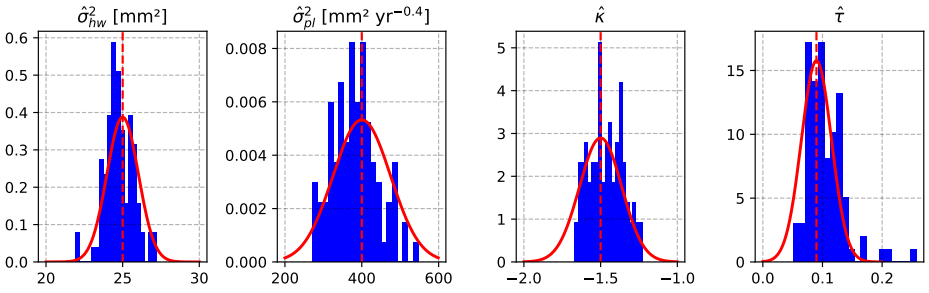

FIGURE 8.4: Distribution of the WN+GGM model variance component estimates obtained using the unconstrained nonlinear LS-VCE.

The distribution of the variance component estimates obtained with the nonlinear LS-VCE method with non-negativity constraints are presented in Figure (8.5). Here again, the algorithm only converged toward a solution for 90 time series out of 200, which correspond to as success rate of 45%. Hence, the non-negative nonlinear LS-VCE method is not more stable than the unconstrained one.

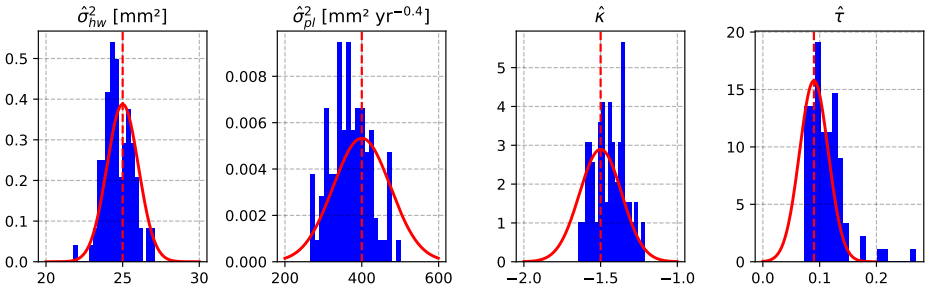

FIGURE 8.5: Distribution of the WN+GGM model variance component estimates obtained using the nonlinear LS-VCE with non-negativity constraints.

Given its poor rate of convergence, more studies are needed to try adapting the nonlinear LS-VCE method to the variance component estimation of WN+GGM models. Indeed, the inability of the method to converge under synthetic cases tends to indicate that this optimization problem is numerically challenging. In particular,

it is possible that the simultaneous estimation of $\sigma_w^2 \approx 25$, $\sigma_{pl}^2 \approx 225$, $\kappa \approx -1.5$, and $\tau \approx 0.1$ is an ill-posed problem with the considered functional model, time span, and initial values. Considering a new formulation of the estimation problem, for instance by estimating $\ln(\tau)$ instead of τ, or optimizing the choice of initial values, could improve the success rate.

8.6 Contributions and perspectives

In this chapter, we investigated the potential of the nonlinear LS-VCE method for time-correlated stochastic process parameter estimation. Stochastic models combining white noise processes with either generalized Gauss-Markov processes (WN+GGM) or power-law processes (WN+PL) have been considered. To apply the nonlinear LS-VCE method, we first presented how to derive the partial derivatives of the WN+GGM model with respect to the spectral index κ and the parameter ϕ. We then tested the accuracy, but also the speed of the method on synthetic time series.

Our results show that the nonlinear LS-VCE method provides accurate estimates of all variance components in the case of the WN+PL model, both for the unconstrained and the non-negative cases. However, in the non-negative case, variance component estimates tend to be slightly biased due to the use of constraints. To consider possible further applications of the method, we also assessed its speed. Our runtime analyses show that, with a 10 minutes runtime for 20 years long daily time series, the nonlinear LS-VCE method provides reasonably fast estimates, and is therefore not prohibitive for future applications. Further comparative studies are considered to investigate whether the method provides a real gain in speed and precision over other existing software.

Unfortunately, for the WN+GGM model, the nonlinear LS-VCE method fails to converge more than half of the time for both the unconstrained and the non-negative cases, which makes it completely unsuited for practical applications. This result suggests that the estimation of all variance components for the WN+GGM process considered in this study is a challenging or ill-posed estimation problem. Further studies are needed to investigate whether a re-parametrization of the problem, or better initial values, could lead to a better success rate in the estimation of WN+GGM parameters.

In addition to estimating variance components, the nonlinear LS-VCE method also computes the standard deviations of the spectral index and the parameter ϕ. This by-product of the nonlinear LS-VCE method is exploited in the next chapter, to evaluate the influence of offsets on low-frequency stochastic parameters such as the parameter ϕ.

Chapter 9

Impact of offsets on the stochastic modelling of GNSS position time series.

In addition to the presence of time-correlated processes in GNSS time series, another frequent issue is the presence of position offsets. Even when detected, offsets remain a nuisance as they increase the uncertainty of all functional model parameters, and especially that of trend estimates. Although their influence on functional parameters has already been investigated, their impact on the estimation of stochastic parameters remains poorly understood.

In this study, we examine how offsets influence the stochastic modelling of GNSS position time series. In particular, we study their impact on various parts of the stochastic modelling, including the spectral analysis, the variance component estimation – both when using the Maximum Likelihood Estimation (MLE) or Restricted Maximum Likelihood Estimation (RMLE) methods –, and the uncertainty of variance components. For the latter part, we use the partial derivatives presented in the previous chapter.

9.1 Introduction

Investigating the time correlation of the stochastic variability observed in GNSS position time series is now a standard practice in many geodetic studies (Bos et al., 2020). In particular, a realistic stochastic model, that is a model realistic of the stochastic variability, is crucial to derive reliable parameter uncertainties, especially for trend estimation (Williams, 2003a; Santamaría-Gómez et al., 2011). Besides, as we shall see in sections 10 and 11, assessing the stochastic variability also helps identify possible non-modelled processes in position time series. This could eventually lead to methodological improvements in the modelling of GNSS observations and therefore to enhance the precision of GNSS products.

Over almost three decades, numerous studies demonstrated the presence of power-law processes in stochastic variability of various geophysical time series (Agnew, 1992; Langbein and Johnson, 1997; Langbein, 2004; Van Camp, Williams, and Francis, 2005), and in GNSS position time series in particular (Zhang et al., 1997; Mao, Harrison, and Dixon, 1999; Calais, 1999; Williams et al., 2004). Power-law processes are stochastic processes characterized by their power-spectrum $P(f)$, which follows a power-law function of the time-frequency f, that is $P(f) \propto f^{\kappa}$, the parameter κ being referred to as the spectral index (Agnew, 1992). When $\kappa = 0.0$, the power-law

process is a simple uncorrelated white noise. When $\kappa < 0.0$, the power-law process is time-correlated. In GNSS position time series analysis, most studies reported that the position residuals tend to follow power-law processes with $\kappa \approx -1.0$, corresponding to the so-called flicker noise (FL) (Williams et al., 2004; Santamaría-Gómez et al., 2011; Rebischung et al., 2016).

The most commonly selected stochastic model for global GNSS solutions is a linear combination of a white noise (WN) and a power-law power process (PL), noted (WN+PL), with possibly unknown amplitudes and spectral index (Williams et al., 2004; Santamaría-Gómez et al., 2011). In the future, with the increasing span of position time series, this stochastic model could evolve. In particular, long time series could allow assessing low-frequency properties of the stochastic variability (He et al., 2019). Two main alternatives to the usual (WN+PL) model are considered. The first one is the detection of random walk (RW) processes, and the second one is the detection of Generalized Gauss Markov (GGM) processes. RW processes are power-law processes with $\kappa = -2.0$, which appear as an increase in the power spectrum of the noise at low frequencies. RW processes can result from ground instabilities around the station (Langbein, 2012). GGM processes, introduced by (Langbein, 2004), are not power-law processes and should appear as a flattening spectrum at low frequencies. Identifying either RW or GGM would improve the understanding of the stochastic variability in GNSS time series and affect parameter uncertainties.

In this chapter, we study the biases and uncertainties that can affect the estimation of low-frequency properties of GNSS position time series. To do so, we investigate the properties of the main statistical tools used for such investigations, namely the spectral estimation methods and the Variance Component Estimation (VCE) methods.

In stochastic modelling, spectral estimation methods aim at identifying the nature of the noise from the shape of the power spectrum of position residuals. Since position time series have unevenly spaced observations, one usually relies on a Least-Squares Spectrum Analysis (LSSA) method, in particular, the Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982). This periodogram is not the only example of LSSA method: Amiri-Simkooei and Tiberius (2007) showed that it could be generalized by the Least-Squares Harmonic Estimation (LS-HE), a method derived from hypothesis testing in linear models (Teunissen, 2000b). The main difference between the Lomb-Scargle and LS-HE periodograms is that the LS-HE takes into account the functional and stochastic models used to obtain position residuals.

Complementary, VCE methods aim at estimating unknown stochastic parameters, named variance components, for a predefined stochastic model. The most widely used VCE method in geodetic time series analysis is the Maximum Likelihood Estimation method (MLE), which is implemented in several statistical modelling software (Williams, 2008; Bos et al., 2013). This VCE method is conceptually simple and provides asymptotically unbiased variance component estimates. In other words, the estimation bias gets smaller when increasing the redundancy of observations. Since one usually model thousands of observations with a few tens parameters only, this bias is assumed to be negligible. Nonetheless, there are unbiased alternatives to the MLE method. We can mention the Restricted Maximum

Likelihood (REML) (Koch, 1986), the Minimum Norm Quadratic Unbiased Estimation (MINQUE) (Rao, 1971), and the Least-Squares Variance Component Estimation (LS-VCE) (Teunissen and Amiri-Simkooei, 2008). Under the normal distribution assumption, these VCE methods yield identical results (Teunissen and Amiri-Simkooei, 2008).

In previous studies, using the MLE and the Lomb-Scargle periodogram method, Dmitrieva, Segall, and Bradley (2017) and Santamaría-Gómez and Ray (2020) showed that the investigation of low-frequency properties of GNSS position times could be biased by the estimation of a linear trend and offsets. Offsets are abrupt discontinuities in time series, usually caused by instrumental changes or large earthquakes (Gazeaux et al., 2013). Undetected, offset induce a bias on functional parameters, and especially station velocities, to an extent related to their magnitude and the span of the time series (Williams, 2003b). Moreover, undetected offsets correlate with random walk processes, such that it also biases variance component estimates. For these reasons, not accounting for offsets is a modelling error that should be avoided at all cost.

In this work, we only consider the case in which offsets are detected and modelled. In particular, even when detected, the presence of offsets is not harmless because they systematically weaken the functional model and increases parameter uncertainties (Williams, 2003b; Griffiths and Ray, 2016; Wang and Herring, 2019).

In this chapter, we assess the influence of trend and offsets estimation on the estimation of low-frequency stochastic properties of geodetic time series. To this end, we investigate 3 complementary aspects of the stochastic modelling: the periodogram computation, variance component estimation biases and, finally, variance component estimation uncertainties. Unlike Dmitrieva, Segall, and Bradley (2017) and Santamaría-Gómez and Ray (2020), we also assess the influence of the chosen LSSA method, Lomb-Scargle and LS-HE, and VCE methods, MLE and RMLE, on the result, which leads to recommendations for future studies.

9.2 Functional and stochastic model of GNSS time series

9.2.1 Functional model

The functional model describes the expectation $\mathsf{E}\{.\}$ of the $m \times 1$ observation vector $\underline{\mathbf{y}}$. For linear(-ized) models of observation equations, the functional model can be written as

$$\mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x}, \tag{9.1}$$

in which $\mathbf{A}$ is the known $m \times n$ design matrix, and $\mathbf{x}$ is the $n \times 1$ functional parameter vector.

The design matrix $\mathbf{A}$ models the identified deterministic effects in the time series. For GNSS time series analysis, the functional model usually accounts for a linear trend, periodic oscillations, offsets, and eventual transients. Hereafter, for comparison purpose, two types of functional models are considered: a simple linear trend model (trend) and models accounting for additional offsets (trend+offsets).

9.2.2 Stochastic model

The stochastic model describes the variance var{.} of the observation vector $\underline{\mathbf{y}}$. Power-law processes being a particular case of generalized Gauss-Markov processes, the preferred (WN+PL) stochastic model follows:

$$\mathbf{Q}_{\mathbf{y}} = \sigma_w^2 \mathbf{I} + \sigma_{pl}^2 \mathbf{Q}_{\kappa\phi}, \tag{9.2}$$

in which σ_w^2 is the white noise amplitude, $\mathbf{I}$ is the $m \times m$ identity matrix, σ_{pl}^2 is the power-law process amplitude, and $\mathbf{Q}_{\kappa\phi}$ is the $m \times m$ covariance matrix associated to a GGM process of spectral index κ and parameter ϕ.

The covariance matrix $\mathbf{Q}_{\kappa\phi}$ can be obtained as

$$\mathbf{Q}_{\kappa\phi} = \Delta_t^{-\kappa/2} \mathbf{L}_{\kappa\phi} \mathbf{L}_{\kappa\phi}^T \tag{9.3}$$

in which Δ_t is the sampling period in years and $\mathbf{L}_{\kappa\phi}$ is a $m \times m$ transformation matrix defined as

$$\mathbf{L}_{\kappa\phi} = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{m-1} & \cdots & h_1 & h_0 \end{bmatrix}, \tag{9.4}$$

where the entries h_i are defined by the following recurrence expression (Langbein, 2004; Bos et al., 2020):

$$\begin{aligned} h_0 &= 1, \\ h_i &= \phi(1 - \frac{\kappa}{2i} - \frac{1}{i}) h_{i-1}. \end{aligned} \tag{9.5}$$

Note that, for $\phi = 1$, the covariance matrix $\mathbf{Q}_{\kappa\phi}$ is that of a power-law process (Williams, 2003a). When $0 < \phi < 1$, the covariance matrix corresponds to that of a Generalized Gauss-Markov process. The parameter ϕ in Equation (9.5) is a unit-less parameter. For interpretation purpose, ϕ can be translated into the time-constant τ of the GGM process, expressed in years, and defined as (Langbein, 2004)

$$\tau = -\frac{\Delta_t}{\ln(\phi)}. \tag{9.6}$$

9.2.3 Properties of position residuals

The primary observational evidence for the presence of time-correlated stochastic processes in GNSS position time series comes from the properties of position residuals (Zhang et al., 1997; Mao, Harrison, and Dixon, 1999). Hence, to understand how offsets may influence the stochastic modelling, we explain in some details how residuals are computed.

From Equation (9.1), the $m \times 1$ residual vector $\underline{\mathbf{e}}$ is defined from the observation vector $\underline{\mathbf{y}}$, as

$$\underline{\mathbf{e}} = \underline{\mathbf{y}} - \mathbf{A}\mathbf{x}. \tag{9.7}$$

Assuming no under-parametrization of the functional model, the expectation and the variance of the residual vector $\underline{\mathbf{e}}$ read

$$\mathsf{E}\{\underline{\mathbf{e}}\} = \mathbf{0}, \tag{9.8}$$

and

$$\mathsf{var}\{\underline{\mathbf{e}}\} = \mathsf{var}\{\underline{\mathbf{y}}\} = \mathbf{Q}_{\mathbf{y}} = \mathbf{Q}_{\mathbf{e}}. \tag{9.9}$$

which shows that, when the functional parameter vector $\mathbf{x}$ is known, the covariance matrices of the observations and the residuals are identical.

However, in most geodetic applications, the functional parameter vector $\mathbf{x}$ is unknown and must be estimated from the observations. To do so, one usually uses the weighted least-squares estimator (Teunissen, 2000a) to obtain an estimator of $\mathbf{x}$, noted $\hat{\underline{\mathbf{x}}}$, such that

$$\hat{\underline{\mathbf{x}}} = (\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}\underline{\mathbf{y}}, \tag{9.10}$$

where $\mathbf{W}$ can be virtually any $m \times m$ symmetric and positive-definite weight matrix. When the weight matrix is chosen to be inverse of the covariance matrix of the observation $\mathbf{W} = \mathbf{Q}_{\mathbf{y}}^{-1}$, one obtains the Best Linear Unbiased Estimator (BLUE) of $\mathbf{x}$.

Using the estimator of the functional parameters $\hat{\underline{\mathbf{x}}}$, one can define the *estimator* of the residuals, noted $\hat{\underline{\mathbf{e}}}$, such that

$$\hat{\underline{\mathbf{e}}} = \underline{\mathbf{y}} - \mathbf{A}\hat{\mathbf{x}}, \tag{9.11}$$

Using Equations (9.10) and (9.11), this estimator can be written as a linear function of $\underline{\mathbf{y}}$ as

$$\hat{\underline{\mathbf{e}}} = \mathbf{P}_{\mathbf{A}}^{\perp}\underline{\mathbf{y}}, \tag{9.12}$$

in which the $m \times m$ matrix $\mathbf{P}_{\mathbf{A}}^{\perp}$ is the weighted least-squares orthogonal projector defined by

$$\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}. \tag{9.13}$$

As the estimated residuals vector $\hat{\underline{\mathbf{e}}}$ is expressed as a linear function of the observation vector $\underline{\mathbf{y}}$, using the linear propagation law of means and variances, one can show that

$$\mathsf{E}\{\hat{\underline{\mathbf{e}}}\} = \mathbf{0}, \tag{9.14}$$

and that

$$\mathsf{var}\{\hat{\underline{\mathbf{e}}}\} = \mathbf{Q}_{\hat{\mathbf{e}}} = \mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{Q}_{\mathbf{y}}\mathbf{P}_{\mathbf{A}}^{\perp T}. \tag{9.15}$$

The crucial difference between the residual vector $\underline{\mathbf{e}}$ and its *estimator* $\hat{\underline{\mathbf{e}}}$ follows from Equation (9.15), which shows that $\mathbf{Q}_{\hat{\mathbf{e}}} \neq \mathbf{Q}_{\mathbf{y}}$. Moreover, $\mathbf{Q}_{\hat{\mathbf{e}}}$ depends on both the chosen design matrix $\mathbf{A}$ and the chosen weight matrix $\mathbf{W}$ through the definition of $\mathbf{P}_{\mathbf{A}}^{\perp}$ in Equation (9.13). In other words, adding offsets to the functional model, whether necessary or not, will systematically affect the covariance of the estimated

residuals. In the sequel, we more specifically assess the influence of offsets on spectral properties of the residuals and on variance component estimation.

9.3 Influence of offsets on spectral analyses

In this section, we study the influence of trend and offsets estimation on the computation of both the LS-HE and Lomb-Scargle periodograms.

9.3.1 Least-Squares Harmonic Estimation

Definition

The first step to LS-HE is to define a null hypothesis, noted H_0, in which the stochastic model is known, and the functional model already describes all the deterministic effects in the observations, that is

$$H_0 : \mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x} \quad ; \quad \mathsf{var}\{\underline{\mathbf{y}}\} = \mathbf{Q}_\mathbf{y} \tag{9.16}$$

The following step is to consider an alternative hypothesis H_a, in which an additional periodic signal with a frequency f_k should be added to the functional model, keeping the stochastic model unchanged, that is

$$H_a : \mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x} + \mathbf{C}_\mathbf{y}\boldsymbol{\nabla} \quad ; \quad \mathsf{var}\{\underline{\mathbf{y}}\} = \mathbf{Q}_\mathbf{y} \tag{9.17}$$

with

$$\mathbf{C}_\mathbf{y} = \begin{bmatrix} \cos(2\pi f_k t_1) & \sin(2\pi f_k t_1) \\ \cos(2\pi f_k t_2) & \sin(2\pi f_k t_2) \\ \vdots & \vdots \\ \cos(2\pi f_k t_m) & \sin(2\pi f_k t_m) \end{bmatrix} ; \boldsymbol{\nabla} = \begin{bmatrix} a_k \\ b_k \end{bmatrix} \tag{9.18}$$

where a_k and b_k are two additional parameters corresponding to the Fourier coefficients associated to the time frequency f_k.

Following (Teunissen, 2000b), the decision to either accept or reject H_0 depends on the value of a T-statistic, noted $\underline{T}_q$, with q the number of additional parameters. In the context of periodic signal detection, $\underline{T}_{q=2}$ can also be viewed as the power $\underline{P}(f_k)$ of the frequency f_k (Amiri-Simkooei, 2020). This T-statistic $\underline{T}_2$ reads

$$\underline{T}_2 = \underline{P}(f_k) = \underline{\hat{\mathbf{e}}}^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{C}_\mathbf{y} [\mathbf{C}_\mathbf{y}^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{P}_\mathbf{A}^\perp \mathbf{C}_\mathbf{y}]^{-1} \mathbf{C}_\mathbf{y}^T \mathbf{Q}_\mathbf{y}^{-1} \underline{\hat{\mathbf{e}}} \tag{9.19}$$

in which the projector $\mathbf{P}_\mathbf{A}^\perp$ is assumed to be defined by taking $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-\mathbf{1}}$.

When designing a stochastic model, the covariance matrix $\mathbf{Q}_\mathbf{y}$ is never known. Hence, in practice, instead of using $\mathbf{Q}_\mathbf{y}^{-1}$ in Equation (9.19), one has to satisfy with a weight matrix $\mathbf{W}$ representing the prior assumptions about the stochastic property of the time series. In result, the LS-HE power spectrum is computed as

$$\underline{P}(f_k) = \underline{\hat{\mathbf{e}}}^T \mathbf{W} \mathbf{C}_\mathbf{y} [\mathbf{C}_\mathbf{y}^T \mathbf{W} \mathbf{P}_\mathbf{A}^\perp \mathbf{C}_\mathbf{y}]^{-1} \mathbf{C}_\mathbf{y}^T \mathbf{W} \underline{\hat{\mathbf{e}}}. \tag{9.20}$$

Expectation

The LS-HE periodogram in Equation (9.20) is expressed as a quadratic form of the stochastic vector $\underline{\hat{\mathbf{e}}}$. Hence, the periodogram is itself a random variable. To understand their impact on the shape of periodograms, it is necessary to study how offsets affect the expectation of $\underline{P}(f_k)$.

Assuming the functional model is not under-parametrized, then $\mathsf{E}\{\underline{\hat{\mathbf{e}}}\} = \mathbf{0}$, and one can show that the expectation of $\underline{P}(f_k)$ reads

$$\begin{aligned} \mathsf{E}\{\underline{P}(f_k)\} &= \mathrm{tr}\{\mathbf{Q}_{\hat{\mathbf{e}}}\mathbf{W}\mathbf{C}_\mathbf{y}[\mathbf{C}_\mathbf{y}^T\mathbf{W}\mathbf{P}_\mathbf{A}^\perp\mathbf{C}_\mathbf{y}]^{-1}\mathbf{C}_\mathbf{y}^T\mathbf{W}\} \\ &= \mathrm{tr}\{\mathbf{P}_\mathbf{A}^\perp\mathbf{Q}_\mathbf{y}\mathbf{P}_\mathbf{A}^{\perp T}\mathbf{W}\mathbf{C}_\mathbf{y}[\mathbf{C}_\mathbf{y}^T\mathbf{W}\mathbf{P}_\mathbf{A}^\perp\mathbf{C}_\mathbf{y}]^{-1}\mathbf{C}_\mathbf{y}^T\mathbf{W}\}. \end{aligned} \tag{9.21}$$

Hence, for given design and weight matrices $\mathbf{A}$ and $\mathbf{W}$, one can compute the expected LS-HE periodogram for any stochastic model $\mathbf{Q}_\mathbf{y}$ using Equation (9.21).

In the special case in which $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-1}$, Teunissen (2000b) shows that $\underline{P}(f_k)$ follows a chi-squared distribution, noted $\chi^2(2,0)$, with expectation $\mathsf{E}\{\underline{P}(f_k)\} = 2$. This implies that $\underline{P}(f_k)$ should appear to be constant (flat) at all frequencies f_k, even in the presence of time-correlated processes.

9.3.2 Lomb Scargle periodogram

Computation

The expression of Lomb-Scargle periodogram can be derived from that of the LS-HE periodogram. In particular, assuming that, in Equation (9.19), $\mathbf{W} = \mathbf{I}$ and that $\mathbf{e}$ is not estimated, but directly observed, so that $\mathbf{P}_\mathbf{A}^\perp = \mathbf{I}$, one obtains the Lomb-Scargle periodogram defined as (Vaníček, 1969; Lomb, 1976; Scargle, 1982)

$$\underline{P}'(f_k) = \underline{\hat{\mathbf{e}}}^T\mathbf{C}_\mathbf{y}[\mathbf{C}_\mathbf{y}^T\mathbf{C}_\mathbf{y}]^{-1}\mathbf{C}_\mathbf{y}^T\underline{\hat{\mathbf{e}}} \tag{9.22}$$

In practice, the Lomb-Scargle periodogram $\underline{P}'(f_k)$ of the residuals is more frequently used than the LS-HE periodogram $\underline{P}(f_k)$. However, $\underline{P}'(f_k)$ ignores the fact the residual vector $\underline{\mathbf{e}}$ is computed using the functional model.

Expectation

As the Lomb-Scargle periodogram is also defined as a quadratic form of the residual vector $\underline{\hat{\mathbf{e}}}$, one can show that

$$\begin{aligned} \mathsf{E}\{\underline{P}'(f_k)\} &= \mathrm{tr}\{\mathbf{Q}_{\hat{\mathbf{e}}}\mathbf{C}_\mathbf{y}[\mathbf{C}_\mathbf{y}^T\mathbf{C}_\mathbf{y}]^{-1}\mathbf{C}_\mathbf{y}^T\} \\ &= \mathrm{tr}\{\mathbf{P}_\mathbf{A}^\perp\mathbf{Q}_\mathbf{y}\mathbf{P}_\mathbf{A}^{\perp T}\mathbf{C}_\mathbf{y}[\mathbf{C}_\mathbf{y}^T\mathbf{C}_\mathbf{y}]^{-1}\mathbf{C}_\mathbf{y}^T\} \end{aligned} \tag{9.23}$$

Hence, for a given design matrix $\mathbf{A}$ and weight matrix $\mathbf{W}$, one can compute the expected Lomb-Scargle periodogram for any stochastic model $\mathbf{Q}_\mathbf{y}$ using Equation (9.23).

Lomb-Scargle periodogram bias resulting from residual estimation

In geodetic time series analysis, it is frequent to choose a stochastic model from the shape of the power spectrum of the estimated residuals $\underline{P}'(f_k)$. It is also common

Parameter	Value
σ_w^2 (mm^2)	25
σ_{pl}^2 (mm^2 yr$^{-0.4}$)	225
κ	-0.8

TABLE 9.1: Considered variance components for the WN+PL model.

to demonstrate the relevance of a given stochastic model by showing how its theoretical power-spectrum, noted $\underline{P}'_{\text{true}}(f_k)$, fits the estimated one $\underline{P}'(f_k)$. When doing so, one assumes the Lomb-Scargle periodogram of the estimated residuals should be similar to that of the true residuals $\underline{P}'_{\text{true}}(f_k)$.

This assumption can be tested by analysing the expected difference between $\underline{P}'(f_k)$ and $\underline{P}'_{\text{true}}(f_k)$

$$\begin{aligned}
\mathsf{E}\{\underline{P}'(f_k) - \underline{P}'_{\text{true}}(f_k)\} &= \mathsf{E}\{\underline{P}'(f_k)\} - \mathsf{E}\{\underline{P}'_{\text{true}}(f_k)\} \\
&= \mathsf{tr}\{\mathbf{Q}_{\hat{\mathbf{e}}}\mathbf{C}_\mathbf{y}[\mathbf{C}_\mathbf{y}^T\mathbf{C}_\mathbf{y}]^{-1}\mathbf{C}_\mathbf{y}^T\} - \mathsf{tr}\{\mathbf{Q}_\mathbf{y}\mathbf{C}_\mathbf{y}[\mathbf{C}_\mathbf{y}^T\mathbf{C}_\mathbf{y}]^{-1}\mathbf{C}_\mathbf{y}^T\} \\
&= \mathsf{tr}\{(\mathbf{Q}_{\hat{\mathbf{e}}} - \mathbf{Q}_\mathbf{y})\mathbf{C}_\mathbf{y}[\mathbf{C}_\mathbf{y}^T\mathbf{C}_\mathbf{y}]^{-1}\mathbf{C}_\mathbf{y}^T\} \\
&= \mathsf{tr}\{(\mathbf{P}_\mathbf{A}^\perp\mathbf{Q}_\mathbf{y}\mathbf{P}_\mathbf{A}^{\perp T} - \mathbf{Q}_\mathbf{y})\mathbf{C}_\mathbf{y}[\mathbf{C}_\mathbf{y}^T\mathbf{C}_\mathbf{y}]^{-1}\mathbf{C}_\mathbf{y}^T\}
\end{aligned} . \tag{9.24}$$

From Equation (9.24), it follows that, as long as $\mathbf{P}_\mathbf{A}^\perp \neq \mathbf{I}$, the Lomb-Scargle periodogram $\underline{P}'(f_k)$ is systematically biased from $\underline{P}'_{\text{true}}(f_k)$. Because the bias depends on $\mathbf{P}_\mathbf{A}^\perp$ and f_k (through $\mathbf{C}_\mathbf{y}$), both the design matrix $\mathbf{A}$ and the weight matrix $\mathbf{W}$ introduce a frequency dependent bias and, therefore, engender a distortion of the spectrum. Note that, using Equation (9.24), one can actually study the bias for any design matrix $\mathbf{A}$ or weight matrix $\mathbf{W}$. However, in the following numerical simulations, we more specifically study the distortion introduced by a linear trend and offsets, assuming different types of time-correlated processes.

9.3.3 Influence of offsets on periodogram interpretation

So far, we only derived expressions allowing to assess the expectations of both the LS-HE $\underline{P}(f_k)$ and Lomb-Scargle $\underline{P}'(f_k)$ periodograms, for a given stochastic model $\mathbf{Q}_\mathbf{y}$ and based on residuals $\underline{\hat{\mathbf{e}}}$ estimated using a design matrix $\mathbf{A}$ and a weight matrix $\mathbf{W}$. To evaluate to which extent the choice for $\mathbf{A}$ and $\mathbf{W}$ may influence the shape of periodograms, we now compute $\mathsf{E}\{\underline{P}(f_k)\}$ and $\mathsf{E}\{\underline{P}'(f_k)\}$ under different synthetic cases.

For the sake of simulation, we consider a fully continuous vertical land motion time series with daily observations over a time span of 20 years. The true stochastic model of this time series is assumed to be a combination of a white noise process of amplitude σ_w^2 and a power-law process with a spectral index κ and an amplitude σ_{pl}^2. The numerical values of the considered variance components are presented in Table (9.1).

Using Equations (9.21) and (9.23), we computed $\mathsf{E}\{\underline{P}(f_k)\}$ and $\mathsf{E}\{\underline{P}'(f_k)\}$ under the following cases:

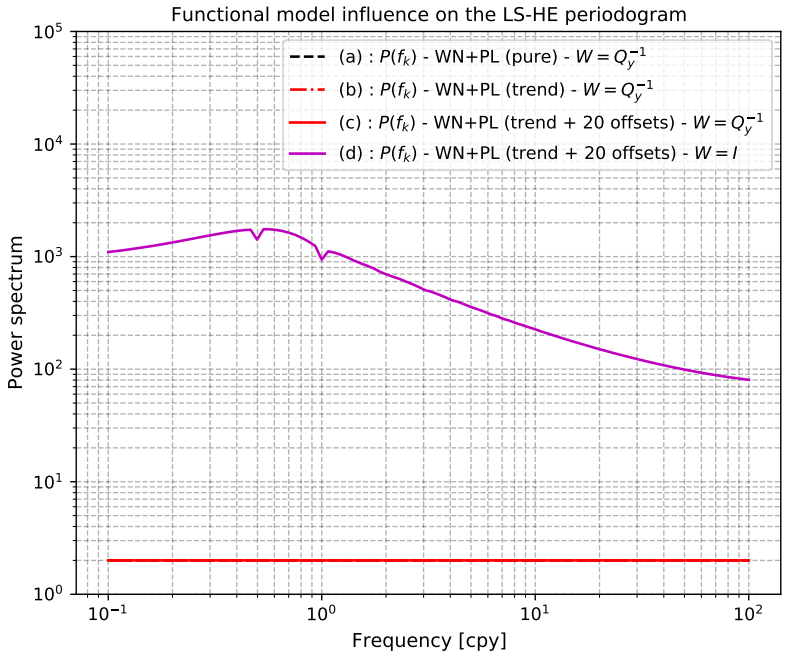

FIGURE 9.1: Influence of the functional model on the expected LS-HE periodogram $P(f_k)$. The expected LS-HE periodograms in the cases (a), (b), and (c) overlap.

- (a) in the case of a direct observation of the noise content (pure),
- (b) in the case of residuals obtained by removing a trend (trend) using the BLUE least-squares projector (that is, using $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-1}$ in Equation (9.13)),
- (c) in the case of residuals obtained by removing a trend and 20 evenly spaced offsets (trend+20 offsets) using the BLUE least-squares projector,
- (d) in the case of residuals obtained by removing a trend and 20 evenly spaced offsets (trend+20 offsets) using the unweighted least-squares projector (that is, using $\mathbf{W} = \mathbf{I}$ in Equation (9.13)).

For $\mathsf{E}\{\underline{P}(f_k)\}$, we assume that $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-1}$ in Equation (9.21), except for the case (d). For both periodograms, the ideal power-spectrum is that of the (pure) stochastic model, which does not involve any functional model.

Figure (9.1) presents the expected LS-HE periodograms $\mathsf{E}\{\underline{P}(f_k)\}$ for each configuration. In the cases (a), (b) and (c), we can see that $\mathsf{E}\{\underline{P}(f_k)\} = 2$ because $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-1}$. This illustrates that the LS-HE periodogram is unaffected by offsets as long as there is no under-parametrization of the functional model and that the correct weight matrix $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-1}$ is used. However, in the case (d), the estimated power spectrum is biased from the true spectrum (a). In this case, $\mathsf{E}\{\underline{P}(f_k)\}$ indicates that there is an anomalous power at low frequency that is not consistent with the chosen weight matrix $\mathbf{W} = \mathbf{I}$. Since $\mathbf{W}$ should ideally be taken as $\mathbf{W} = \mathbf{Q}_\mathbf{y}^{-1}$, in a real analysis, this would typically suggest considering time-correlation in the stochastic model.

Figure (9.2) presents the expected Lomb-Scargle periodograms $\mathsf{E}\{\underline{P}'(f_k)\}$ for each configuration. In the case (a), $\mathsf{E}\{\underline{P}'(f_k)\}$ follows the periodogram expected from a

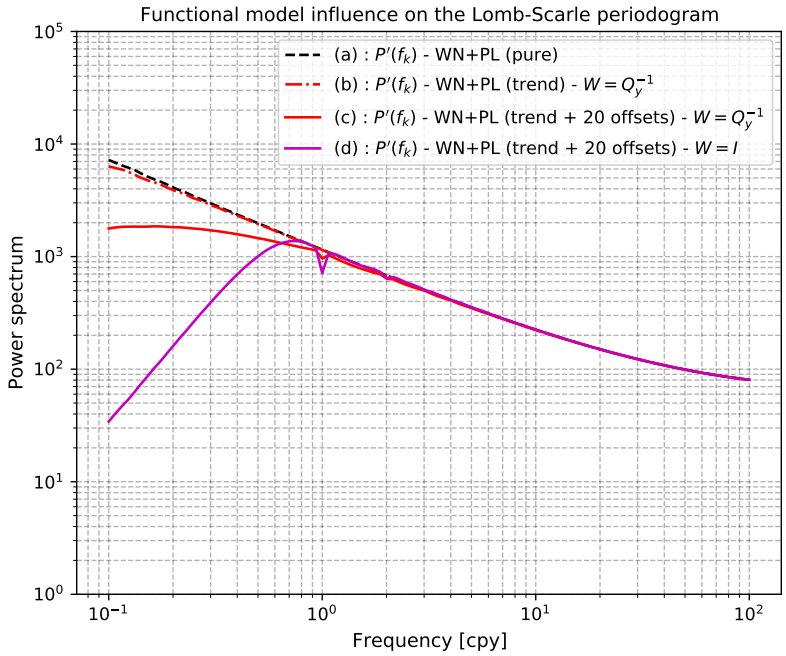

FIGURE 9.2: Influence of the functional model on the expected Lomb-Scargle periodogram $P'(f_k)$.

WN+PL process, that is, a slope at low frequencies and a flattening at high frequencies. But, when a functional model is used, $\mathsf{E}\{\underline{P}'(f_k)\}$ shows distortions from the true spectrum (a). These departures are mostly visible at low frequencies, in particular below the considered offset occurrence frequency of 1 cpy, in agreement with the observations of Santamaría-Gómez and Ray (2020). In the case (b), the effect of the trend estimation is relatively small. However, in the cases (c) and (d), the distortion caused by offsets is dramatic, especially for (d), which uses an unweighted least-square projector. As a result, the Lomb-Scargle periodogram is strongly influenced by offsets, but also by the chosen weight matrix $\mathbf{W}$.

Although we consider a WN+PL process, in cases (c) and (d) the expected Lomb-Scargle periodograms show either flattening or a collapse of the power spectrum at low-frequency. In particular, from the shape of the Lomb-Scargle periodogram, in the case (c) a WN+GGM model may erroneously appear as more appropriate. Hence, if the stochastic characterization of a time series is exclusively based on the shape of the Lomb-Scargle periodogram, in the presence of several offsets, there is a risk of incorrect stochastic model identification.

Note that, in the case (c), the LS-HE periodogram is flat, which would indicate a good agreement between the stochastic model and the observations, and would not promote a WN+GGM model. Hence, in the presence of offsets, one may prefer the use of the LS-HE periodogram for the stochastic model investigations. However, as the LS-HE periodogram only indicates a disagreement with a noise model and observations, one can start with a first guess about the stochastic model, and then adapt this model to account for possible deviations observed in the LS-HE periodogram.

9.4 Influence of offsets on stochastic model selection

The use of periodograms for the identification of the stochastic model is frequent but not systematic. Some studies directly tries different known stochastic models and select the best one using a likelihood-based criterion (Williams et al., 2004; Santamaría-Gómez et al., 2011; He et al., 2019). In such cases, it is necessary to investigate whether offsets can impact the selection criterion used.

The most frequently used method for stochastic model selection in geodetic time series analysis is the Maximum Likelihood Estimation (MLE) (Bos et al., 2020). This method proposes to chose among possible stochastic models, noted $\bar{\mathbf{Q}}_{\mathbf{y}}$, by selecting the one maximizing the log-likelihood $\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\bar{\mathbf{Q}}_{\mathbf{y}}))$ of the observations $\mathbf{y}$. When the observations $\mathbf{y}$ depend on functional parameters $\mathbf{x}$, the MLE is biased. However, this bias reduces for when increasing the redundancy $m - n$, hence, the MLE method is asymptotically unbiased. Since in geodetic time series analysis, the $m - n >> 0$, the bias of the MLE method is assumed negligible.

Nonetheless, there exists an unbiased alternative to the MLE method, namely the restricted maximum likelihood estimation method (RMLE). This method proposes to choose among possible stochastic models $\bar{\mathbf{Q}}_{\mathbf{y}}$ by selecting the one maximizing the log-likelihood $\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\bar{\mathbf{Q}}_{\mathbf{y}}))$ of a misclosure vector $\mathbf{t}$ constructed from the observations (Koch, 1986).

To understand the influence of offsets on the selection of a stochastic model, we study how offsets affect both the log-likelihood $\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\bar{\mathbf{Q}}_{\mathbf{y}}))$ and the restricted log-likelihood $\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\bar{\mathbf{Q}}_{\mathbf{y}}))$.

9.4.1 Log-likelihood

Computation

For an observation vector $\underline{\mathbf{y}}$ following a multivariate normal distribution, the log-likelihood $\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\bar{\mathbf{Q}}_{\mathbf{y}}))$ of $\underline{\mathbf{y}}$ given a stochastic model $\bar{\mathbf{Q}}_{\mathbf{y}}$ follows

$$\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\bar{\mathbf{Q}}_{\mathbf{y}})) = -\frac{1}{2}[m\ln(2\pi) + \ln(\det(\bar{\mathbf{Q}}_{\mathbf{y}})) + \underline{\hat{\mathbf{e}}}^T\bar{\mathbf{Q}}_{\mathbf{y}}^{-1}\underline{\hat{\mathbf{e}}}]. \tag{9.25}$$

in which, $\underline{\hat{\mathbf{e}}}$ is obtained by taking $\mathbf{W} = \bar{\mathbf{Q}}_{\mathbf{y}}^{-1}$ in Equation (9.13).

Note that, in Equation (9.25), the vector $\underline{\hat{\mathbf{e}}}$ is used regardless of the number parameters or the type of functional model used to compute it. In other words, the log-likelihood does not account for the variance absorbed by the estimator of the functional parameter vector $\underline{\hat{\mathbf{x}}}$, which systematically biases variance component estimation.

Expectation

As the log-likelihood in Equation (9.25) depends on a quadratic form of the stochastic vector $\underline{\hat{\mathbf{e}}}$, the log-likelihood $\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\bar{\mathbf{Q}}_{\mathbf{y}}))$ is itself a stochastic variable. Hence, to assess the systematic influence of offsets on stochastic model selection using the MLE method, we propose to study the expectation of the log-likelihood and not directly the log-likelihood.

One can show that the expected log-likelihood of the vector $\underline{\mathbf{y}}$ for given stochastic model $\bar{\mathbf{Q}}_{\mathbf{y}}$ and true stochastic model $\mathbf{Q}_{\mathbf{y}}$ reads

$$\begin{aligned} \mathsf{E}\{\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\bar{\mathbf{Q}}_{\mathbf{y}}))\} &= -\frac{1}{2}[m\ln(2\pi) + \ln(\det(\bar{\mathbf{Q}}_{\mathbf{y}})) + \mathrm{tr}\{\mathbf{Q}_{\hat{\mathbf{e}}}\bar{\mathbf{Q}}_{\mathbf{y}}^{-1}\}]. \\ &= -\frac{1}{2}[m\ln(2\pi) + \ln(\det(\bar{\mathbf{Q}}_{\mathbf{y}})) + \mathrm{tr}\{\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{Q}_{\mathbf{y}}\mathbf{P}_{\mathbf{A}}^{\perp T}\bar{\mathbf{Q}}_{\mathbf{y}}^{-1}\}] \end{aligned} \tag{9.26}$$

Using Equation (9.26), one can evaluate the expected log-likelihood of $\underline{\mathbf{y}}$ for a given stochastic model $\bar{\mathbf{Q}}_{\mathbf{y}}$ and compare it to that of the true stochastic model $\mathbf{Q}_{\mathbf{y}}$. In particular, when $\bar{\mathbf{Q}}_{\mathbf{y}}$ depends on one or more variance components, one can test whether the argument of the maximum of $\mathsf{E}\{\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\bar{\mathbf{Q}}_{\mathbf{y}}))\}$ is close to the true variance components. This allows exploring how the MLE method is biased by offsets when estimating variance components.

9.4.2 Restricted log-likelihood

Computation

Unlike the log-likelihood, the restricted log-likelihood is not based on the $m \times 1$ observation vector $\underline{\mathbf{y}}$, but on the $(m-n) \times 1$ misclosure vector $\underline{\mathbf{t}}$. This misclosure vector $\underline{\mathbf{t}}$ is derived from the following transformation (Koch, 1986)

$$\begin{bmatrix} \underline{\hat{\mathbf{x}}} \\ \underline{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} (\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1} \\ \mathbf{B}^T \end{bmatrix} \underline{\mathbf{y}} \tag{9.27}$$

in which $\mathbf{B}$ is a $m \times (m-n)$ matrix that spans the null space of $\mathbf{A}^T$, such that $\mathbf{B}^T\mathbf{A} = \mathbf{0}$.

The linear application in Equation (9.27) is invertible (Teunissen, 2000a). This implies that all the variance that is not absorbed by the functional parameter estimator $\underline{\hat{\mathbf{x}}}$ is fully absorbed by the misclosure vector $\underline{\mathbf{t}}$. Besides, by studying the expectation and the variance of Equation (9.27), one can show that

$$\mathsf{E}\left\{\begin{bmatrix} \underline{\hat{\mathbf{x}}} \\ \underline{\mathbf{t}} \end{bmatrix}\right\} = \begin{bmatrix} \mathbf{x} \\ \mathbf{0} \end{bmatrix} \quad ; \quad \mathsf{var}\left\{\begin{bmatrix} \underline{\hat{\mathbf{x}}} \\ \underline{\mathbf{t}} \end{bmatrix}\right\} = \begin{bmatrix} (\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{A})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^T\mathbf{Q}_{\mathbf{y}}\mathbf{B} \end{bmatrix}, \tag{9.28}$$

which indicates that $\underline{\mathbf{t}}$ and $\underline{\hat{\mathbf{x}}}$ are independent.

The fact the misclosure vector $\underline{\mathbf{t}}$ contains all the redundancy of the system while being independent of $\underline{\hat{\mathbf{x}}}$ makes it better suited for variance component estimation than the estimator of the residuals $\underline{\hat{\mathbf{e}}}$. For these reasons, the restricted log-likelihood $\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\bar{\mathbf{Q}}_{\mathbf{y}}))$ is defined as a function of $\underline{\mathbf{t}}$ as (Koch, 1986)

$$\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\bar{\mathbf{Q}}_{\mathbf{y}})) = -\frac{1}{2}[(m-n)\ln(2\pi) + \ln(\det(\mathbf{B}^T\bar{\mathbf{Q}}_{\mathbf{y}}\mathbf{B})) + \underline{\mathbf{t}}^T(\mathbf{B}^T\bar{\mathbf{Q}}_{\mathbf{y}}\mathbf{B})^{-1}\underline{\mathbf{t}}], \tag{9.29}$$

with

$$\underline{\mathbf{t}} = \mathbf{B}^T\underline{\mathbf{y}}. \tag{9.30}$$

Expectation

As Equation (9.29) depends on a quadratic form of the stochastic vector $\underline{\mathbf{t}}$, the restricted log-likelihood $\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\bar{\mathbf{Q}}_\mathbf{y}))$ is itself a stochastic variable. As for the log-likelihood, we propose to study the influence of offsets on the RMLE estimation by analysing its expectation

$$\mathsf{E}\{\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\bar{\mathbf{Q}}_\mathbf{y}))\} = -\frac{1}{2}[(m-n)\ln(2\pi) + \ln(\det(\mathbf{B}^T\bar{\mathbf{Q}}_\mathbf{y}\mathbf{B})) + \mathrm{tr}\{\mathbf{B}^T\mathbf{Q}_\mathbf{y}\mathbf{B}(\mathbf{B}^T\bar{\mathbf{Q}}_\mathbf{y}\mathbf{B})^{-1}\}] \tag{9.31}$$

Using Equation (9.31), one can evaluate the expected restricted log-likelihood of $\underline{\mathbf{t}}$ for a given stochastic model $\bar{\mathbf{Q}}_\mathbf{y}$ and compare it to that of the true stochastic model $\mathbf{Q}_\mathbf{y}$. By doing so, one can analyze the influence of offsets on the shape of the restricted log-likelihood function and compare with that of the log-likelihood function.

9.4.3 Influence on the noise model selection

As shown in section 9.3.3, under the influence of offsets, the Lomb-Scargle periodogram can promote WN+GGM models with $\phi \neq 1$, although the true stochastic model is WN+PL model (that is, a WN+GGM model with $\phi = 1$). In this section, we investigate whether the same effect is possible for the selection of stochastic models based on either the MLE or the RMLE method. To do so, we consider the observation of a WN+PL process and investigate the expected log-likelihood and restricted log-likelihood of a WN+GGM model for different values of ϕ. The parameters of the considered WN+PL process are given in Table (9.1). The expected log-likelihood and restricted log-likelihood are computed for functional models with a varying number of offsets. In the ideal case, the argument of the maximum of the expected log-likelihood and the expected restricted log-likelihood should be $\phi \approx 1$ whatever the number of offsets.

Using Equations (9.26) and (9.31), we computed $\mathsf{E}\{\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\bar{\mathbf{Q}}_\mathbf{y}))\}$ and $\mathsf{E}\{\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\bar{\mathbf{Q}}_\mathbf{y}))\}$ for $\bar{\mathbf{Q}}_\mathbf{y}$ corresponding to WN+GGM models with ϕ ranging from 0.975 to 1. The other variance components, namely σ_w^2, σ_{pl}^2, and κ, were kept fixed to their true values. We computed each expected log-likelihood function for 3 different functional models, namely, a trend, a trend with 7 evenly spaced offsets, and a trend with 20 evenly spaced offsets. The graph of the expected log-likelihood functions are presented in Figure (9.3). For a better legibility, we subtracted the expected log-likelihood and the expected restricted log-likelihood of the true stochastic model, that is $\mathsf{E}\{\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\mathbf{Q}_\mathbf{y}))\}$ and $\mathsf{E}\{\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\mathbf{Q}_\mathbf{y}))\}$.

In Figure (9.3), we can see that the log-likelihood is maximum for $\phi < 1$ for all functional models. In particular, the most likely ϕ value is 0.9980 when accounting for a simple trend, 0.9935 in the case of 7 additional offsets, and 0.9805 in the case of 20 additional offsets. The impact of these different values of ϕ in terms of power-spectra are depicted in Figure (9.4). Hence, the log-likelihood estimate of ϕ is biased for all functional models, but the bias gets larger in the presence of additional offsets. This demonstrates that, just like the Lomb-Scargle periodogram, the MLE method can erroneously promote WN+GGM models over the WN+PL model in the presence of offsets. In Figure (9.3), the expected restricted log-likelihood is always maximum for $\phi = 1$. Hence, unlike the MLE method, the RMLE method does not promote WN+GGM models over the WN+PL model, even in the presence of 20 offsets.

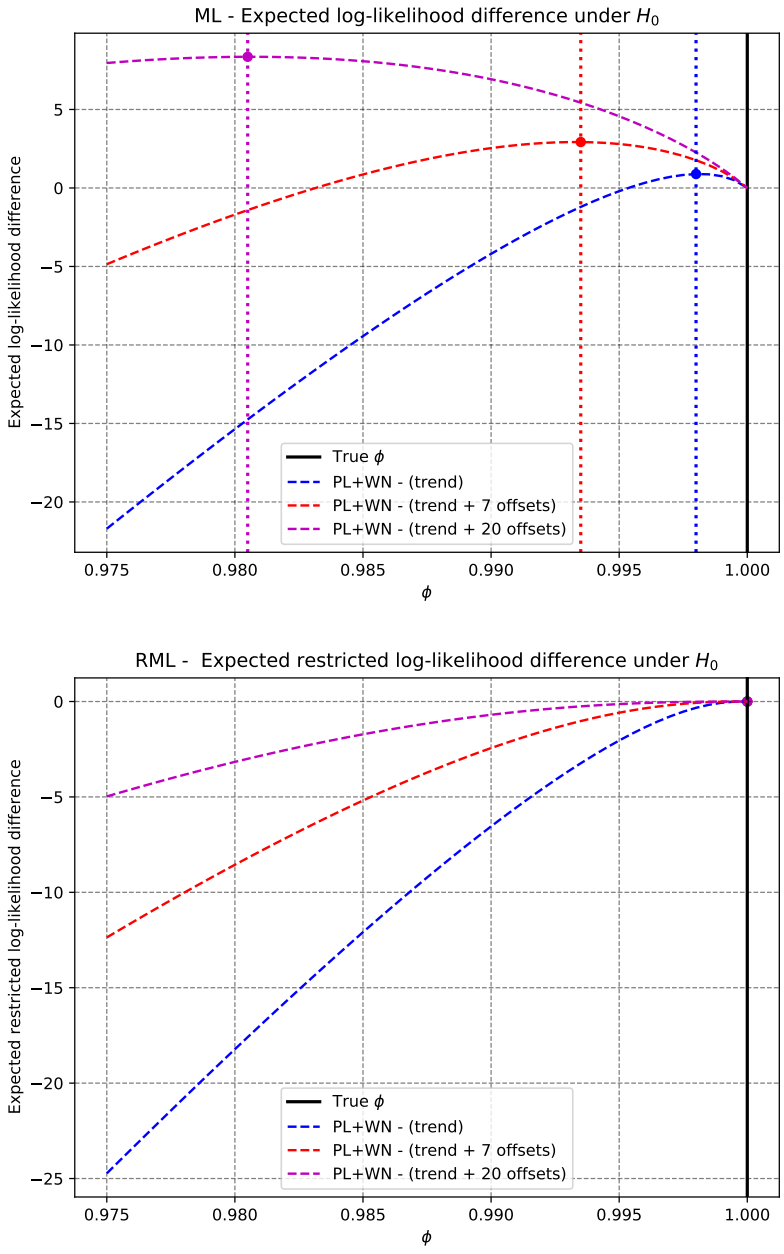

FIGURE 9.3: Expected log-likelihood (*top*) and restricted log-likelihood (*bottom*) for WN+GGM models with different values of ϕ. For a better legibility, the (restricted) log-likelihood of the true WN+PL model is systematically subtracted.

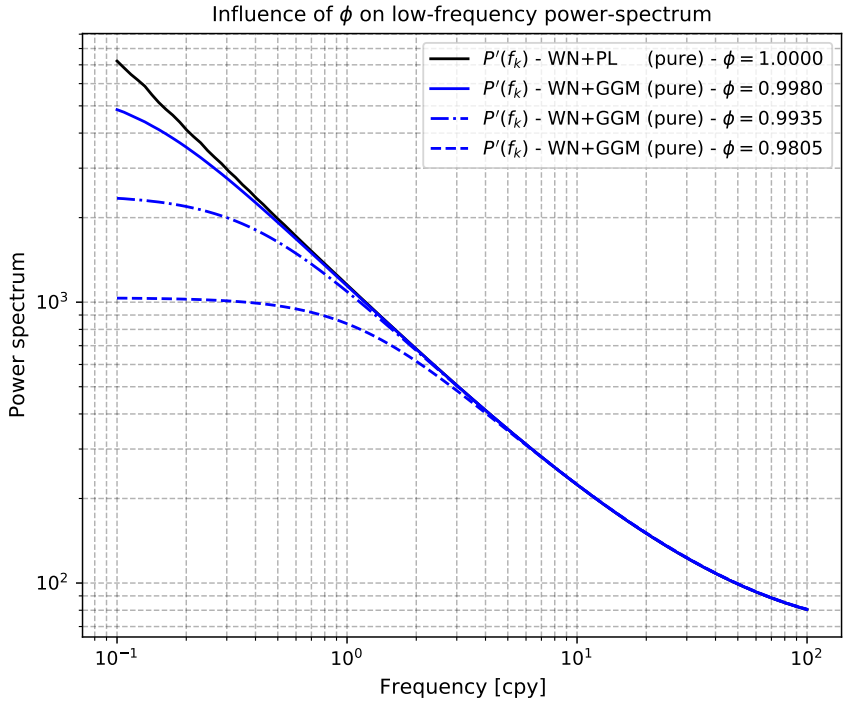

FIGURE 9.4: Influence of the parameter ϕ on the power-spectrum.

Nonetheless, in Figure (9.3), the presence of offsets results in a flattening of both the log-likelihood and the restricted log-likelihood functions. This flattening is even more apparent in Figure (9.5) where the likelihood functions are presented as a function of the time-constant τ instead of the parameter ϕ. For large τ, that is $\tau > 1$, the expected restricted log-likelihood function in Figure (9.5) is almost constant. This flattening has two main consequences. This first is that the number of offsets necessarily reduces the precision of variance components because the likelihood function curvature is linked to the minimum uncertainty of the considered parameter through the Cramér-Rao lower bound. Hence, when the curvature of the likelihood function decreases, so is the precision of the variance components. In Figure (9.5), the flattening of the expected restricted log-likelihood for large τ implies the RMLE method is unable to distinguish between a WN+PL model and WN+GGM models when $\phi \approx 1$. The second consequence is that the flattening may result in a reduction of the rate of convergence of the nonlinear optimization algorithm used for variance component estimation.

Given the impact of offsets on the assessment of ϕ, we also investigated whether similar biases can occur when estimating the spectral index κ. Using Equations (9.26) and (9.31), we computed both $\mathrm{E}\{\ln(p_{\underline{\mathbf{y}}}(\mathbf{y}|\bar{\mathbf{Q}}_{\mathbf{y}}))\}$ and $\mathrm{E}\{\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\bar{\mathbf{Q}}_{\mathbf{y}}))\}$ for $\bar{\mathbf{Q}}_{\mathbf{y}}$ corresponding to WN+PL models with a parameter κ ranging from -0.79 to -0.81. The other variance components, namely σ_w^2, σ_{pl}^2 and ϕ, were fixed to their true values. The considered functional models are the same as for ϕ. The graph of both functions are presented in Figure (9.6). Figure (9.6) shows that the log-likelihood tend to promote biased spectral indices $\kappa > -0.8$, i.e. a whiter noise model. However, the deviation for 20 offsets is only about 0.0025, which is negligible and suggests that offsets mostly affect variance components related to the low-frequency properties of the stochastic model.

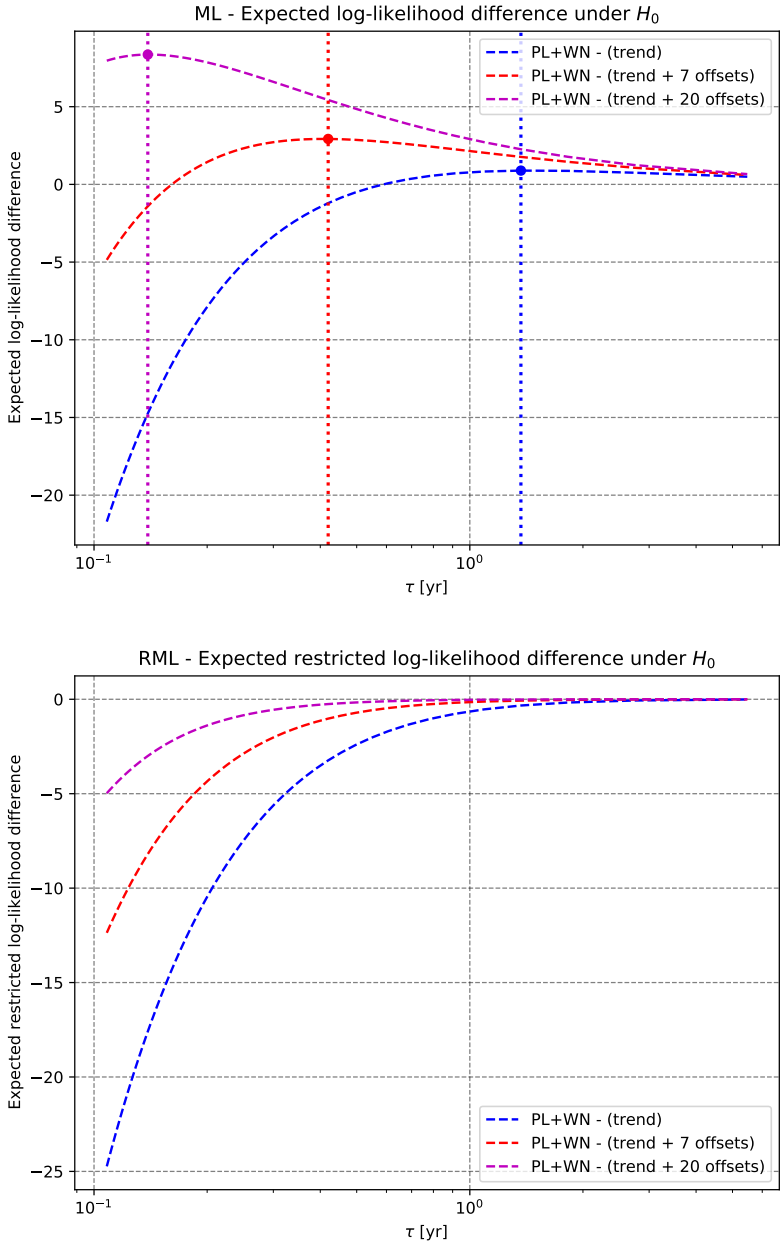

FIGURE 9.5: Expected log-likelihood (*top*) and restricted log-likelihood (*bottom*) for WN+GGM models with different values of τ. For a better legibility, the (restricted) log-likelihood of the true WN+PL model is systematically subtracted.

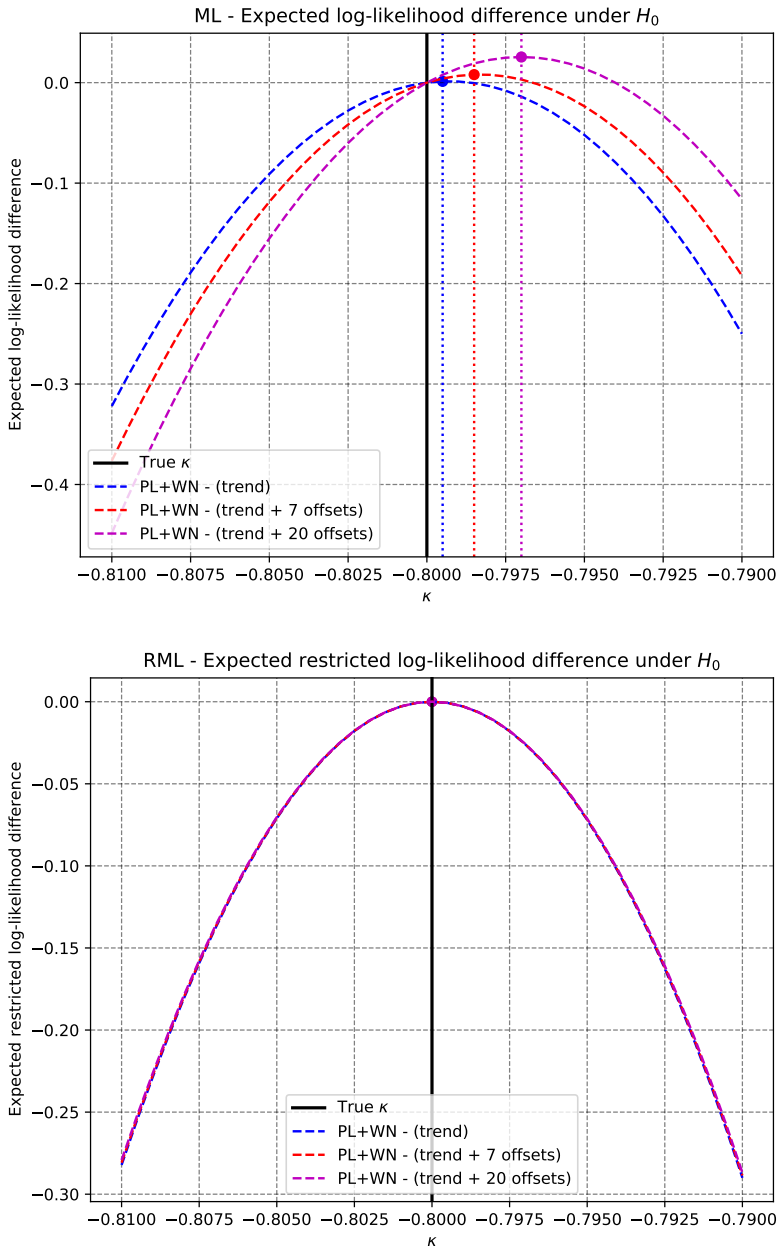

FIGURE 9.6: Expected log-likelihood (*top*) and restricted log-likelihood (*bottom*) for WN+PL models with different values of κ. For a better legibility, the (restricted) log-likelihood of the WN+PL model with $\kappa = -0.8$ is systematically subtracted.

9.5 Uncertainty of variance components

In the previous section, we suggested that the flattening of the likelihood function could cause an increase in low-frequency parameter uncertainties, even when using the RMLE method. In this section, we aim at quantifying this effect by using the Least-Squares Variance Component Estimation (LS-VCE) method. Introduced by (Teunissen, 1988) and further developed by (Teunissen and Amiri-Simkooei, 2008) and (Amiri-Simkooei, 2007), the LS-VCE provides an unbiased and minimum variance estimator of the variance components under the normal distribution assumption. Under the assumption of normally distributed observation, the LS-VCE is equivalent to the RMLE method. However, the LS-VCE allows for a direct assessment of the standard deviation of the variance components, which is ideal to investigate parameter uncertainties.

In its original form, the LS-VCE considers a linear functional model $\mathsf{E}\{\mathbf{y}\}$ and a linear stochastic model $\mathsf{var}\{\mathbf{y}\} = \mathbf{Q}_\mathbf{y}$ such that

$$\mathsf{E}\{\mathbf{y}\} = \mathbf{A}\mathbf{x}, \quad \mathbf{Q}_\mathbf{y} = \mathbf{Q}_0 + \sum_{i=1}^{p} \sigma_i \cdot \mathbf{Q}_i \tag{9.32}$$

where $\mathbf{Q}_0$ and $\mathbf{Q}_i$ are known $m \times m$ co-factor matrices, here corresponding to the covariance matrix of each stochastic process, and the σ_i are the p unknown variance components, corresponding to their amplitudes.

Noting the $p \times 1$ variance component vector as $\hat{\boldsymbol{\sigma}} = \left[\hat{\sigma}_1 \cdots \hat{\sigma}_p\right]^T$, its $p \times p$ covariance matrix $\mathbf{Q}_{\hat{\sigma}}$ can be obtained as (Amiri-Simkooei, 2007)

$$\mathbf{Q}_{\hat{\sigma}} = \mathbf{N}^{-1}. \tag{9.33}$$

in which $\mathbf{N}$ is a $p \times p$ normal matrix whose entries n_{ij} are defined by

$$n_{ij} = \frac{1}{2}\mathrm{tr}(\mathbf{Q}_i\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp\mathbf{Q}_j\mathbf{Q}_\mathbf{y}^{-1}\mathbf{P}_\mathbf{A}^\perp). \tag{9.34}$$

The diagonal elements of $\mathbf{Q}_{\hat{\sigma}}$ correspond to the variance of the variance components, noted $\sigma^2_{\sigma_i}$. Given the expression (9.34), the normal matrix $\mathbf{N}$ does not depend on the observations $\mathbf{y}$. Hence, it is possible to compute each $\sigma^2_{\sigma_i}$ without simulating surrogate time series. However, the normal matrix $\mathbf{N}$ depends on the covariance matrix $\mathbf{Q}_\mathbf{y}$, which itself depends on the variance components σ_i in Equation (9.32). Therefore, deriving $\sigma^2_{\sigma_i}$ requires an estimate of the variance component vector $\hat{\boldsymbol{\sigma}}$ so that assumptions about the amplitude of each process must be made.

When the covariance matrix $\mathbf{Q}_\mathbf{y}$ is a nonlinear function $\mathbf{Q}_\mathbf{y}(\hat{\boldsymbol{\sigma}})$ of the variance components σ_i, the matrix $\mathbf{Q}_i$ in Equation (9.34) is defined as (Amiri-Simkooei, 2007)

$$\mathbf{Q}_i = \left.\frac{\partial \mathbf{Q}_\mathbf{y}(\hat{\sigma})}{\partial \sigma_i}\right|_{\hat{\sigma}}. \tag{9.35}$$

Note that, in the nonlinear case, the variance component σ_i can be virtually any variance component, such as the spectral index κ or the parameter ϕ, and not necessarily the amplitude of a stochastic process.

Parameter	Value
σ_w^2 (mm^2)	25
σ_{ggm}^2 (mm^2 yr$^{-0.5}$)	225
κ	-1
ϕ	0.98

TABLE 9.2: Considered variance components for the WN+GGM model.

9.5.1 Influence on the variance component uncertainties

To assess the influence of offsets on the variance component uncertainties, we consider 2 stochastic models devised to the modelling of different low frequency properties, namely a combination of white noise and generalized Gauss-Markov process model (WN+GGM), and a combination of white noise, flicker noise, and random walk process (WN+FL+RW) model.

Influence on a WN+GGM model

To estimate how the presence of offsets affects the uncertainty of WN+GGM parameters, we consider the simultaneous estimation of σ_w^2, σ_{pl}^2, κ, and ϕ for the model defined in Equation (9.2). The considered parameters for σ_w^2, σ_{pl}^2, κ, and ϕ are presented in Table (9.2). Since the stochastic model depends non-linearly on the parameter κ and ϕ, to obtain (9.34) and (9.33), it is necessary to compute the partial derivatives of the stochastic model with respect to each variance component. The computation of these partial derivatives is presented in chapter 8.

Using Equation (9.33) and (9.34), we simulated the evolution of the standard deviation of each variance component estimate, noted $\sigma_{\hat{\sigma}_w^2}$, $\sigma_{\hat{\sigma}_{pl}^2}$, $\sigma_{\hat{\kappa}}$, and $\sigma_{\hat{\phi}}$ for offset frequency of occurrence ranging from 0.0 (no offsets) to 1.0 offset per year (opy). As in previous sections, we consider a continuous daily time series with a span of 20 years. The computed variations of $\sigma_{\hat{\sigma}_w^2}$, $\sigma_{\hat{\sigma}_{pl}^2}$, $\sigma_{\hat{\kappa}}$, and $\sigma_{\hat{\phi}}$ as a function of the offset frequency are presented in Figure (9.7). For each variance component, the uncertainty variation increases almost linearly with the offset frequency. However, the relative uncertainty change is not similar for all components. To illustrate this, the increase in uncertainty at 1.0 opy, compared to the uncertainty with no offset, is presented, for each component, in Table (9.3). In Table (9.3), the largest increase in uncertainty is observed for the parameter ϕ, which increased by $+31.2\%$. This is about 3 to 9 times larger than the other parameters, hence, the uncertainty on the low-frequency parameter ϕ is sensitive to offset frequency.

Influence on a WN+FL+RW model

An alternative to the WN+GGM model is the WN+FL+RW model describing a possible increase of variance at low frequency, which could result from monument instabilities. The WN+FL+RW model can be expressed as

$$\mathbf{Q}_\mathbf{y} = \sigma_w^2 \mathbf{I} + \sigma_{fl}^2 \mathbf{Q}_{fl} + \sigma_{rw}^2 \mathbf{Q}_{rw}, \tag{9.36}$$

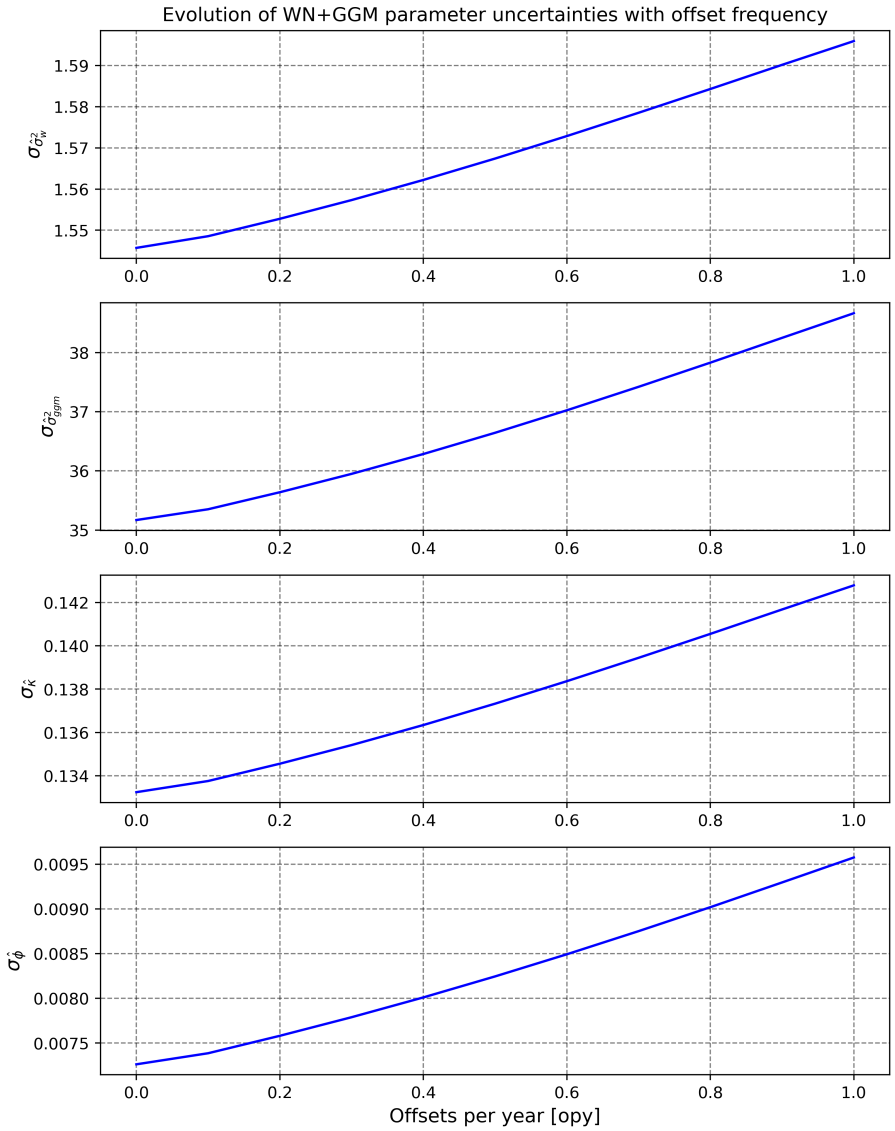

FIGURE 9.7: Evolution of the standard deviation of variance components with offset frequency for the WN+GGM model.

Parameter	Uncertainty change at 1.0 opy
$\sigma_{\sigma^2_w}$ (mm^2)	+3.3%
$\sigma_{\sigma^2_{ggm}}$ (mm^2 yr$^{-0.5}$)	+9.9%
σ_κ	+7.2%
σ_ϕ	+31.2%

TABLE 9.3: Percentage of standard deviation increase for the WN+GGM variance components when considering functional models with 1.0 opy compared to 0.0 opy (that is, no offsets).

Parameter	Value
σ_w^2 (mm^2)	25
σ_{fl}^2 (mm^2 yr$^{-0.5}$)	225
σ_{rw}^2 (mm^2 yr^{-1})	4

TABLE 9.4: Considered variance components for the WN+FL+RW model.

Parameter	Uncertainty change at 1.0 opy
$\sigma_{\sigma_w^2}$ (mm^2)	+2.9%
$\sigma_{\sigma_{fl}^2}$ (mm^2 yr$^{-0.5}$)	+6.3%
$\sigma_{\sigma_{rw}^2}$ (mm^2 yr^{-1})	+64.5%

TABLE 9.5: Percentage of standard deviation increase for the WN+FL+RW variance components when considering functional models with 1.0 opy compared to 0.0 opy (that is, no offsets).

in which $\mathbf{I}$ is the covariance matrix of the white noise process, σ_w^2 is its amplitude, $\mathbf{Q}_{fl}$ is the covariance matrix of the flicker noise process, σ_{fl}^2 is its amplitude, $\mathbf{Q}_{rw}$ is the covariance matrix of the random walk process, and σ_{rw}^2 is its amplitude. The matrices $\mathbf{Q}_{fl}$ and $\mathbf{Q}_{rw}$ can be obtained by taking $\kappa = -1.0$ and $\kappa = -2.0$ respectively, in Equation (9.4), with $\phi = 1.0$ in both cases. The considered parameters for σ_w^2, σ_{fl}^2, and σ_{rw}^2 are presented in Table (9.4).

The WN+FL+RW model depending linearly on each process amplitude, using only Equations (9.33) and (9.34), we computed the evolution of the standard deviation of each amplitude, namely $\sigma_{\hat{\sigma}_w^2}$, $\sigma_{\hat{\sigma}_{fl}^2}$, and $\sigma_{\hat{\sigma}_{rw}^2}$ in the presence of offsets with a frequency of occurrence ranging from 0.0 (no offset) to 1.0 opy. The computed variations of $\sigma_{\hat{\sigma}_w^2}$, $\sigma_{\hat{\sigma}_{fl}^2}$, and $\sigma_{\hat{\sigma}_{rw}^2}$ in functions of offset frequency are presented in Figure (9.8). As for the WN+GGM model, the uncertainty variations observed in Figure (9.8) are approximately linear for all components. As the relative variability also differs between parameters, we present the increase in uncertainty at 1.0 opy with respect to 0.0 opy, for each parameter, in Table (9.5). The largest increase in uncertainty is observed for the random walk amplitude uncertainty, which increased by +64.5%. This is about 10 times superior to that of white noise and flicker noise amplitudes. Hence, the uncertainty random walk amplitude is more sensitive to the number of offsets in the time series than other process amplitudes.

Besides, in Figure (9.8), the minimum random walk uncertainty is 5.1 mm^2 yr^{-1} whereas the random walk amplitude is 4.0 mm^2 yr^{-1}. This implies that, with the considered level of white noise and flicker noise, it is impossible to detect a random walk process with a 4.0 mm^2 yr^{-1} amplitude, even with 20 year of continuous observations.

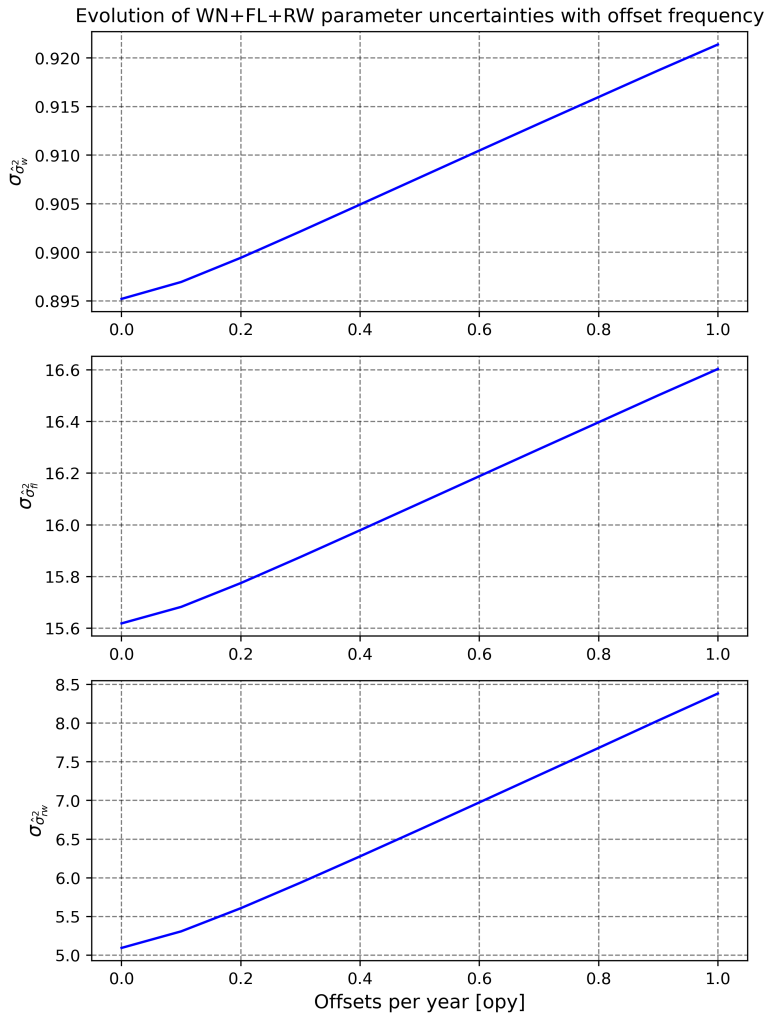

FIGURE 9.8: Evolution of the standard deviation of variance components with offset frequency for the WN+FL+RW model.

9.6 Discussion

In sections 9.3.2 and 9.4.3, we show that the systematic biases observed when using the Lomb-Scargle periodogram or the MLE method disappears when using the LS-HE periodogram and the RMLE method. Therefore, part of the problem resulting from the presence of offsets can be solved by using different statistical tools. In particular, the asymptotically unbiased property of the MLE method, which is sufficient for WN+PL models (Figure 9.6), may meet its limits for low-frequency noise investigations (Figure (9.3) and (9.5)). Hence, for such analyses, we recommend to systematically use the RMLE method and the LS-HE periodogram, to avoid a biased stochastic modelling.

The transition from the MLE method to the RMLE method is rather straightforward as it only requires the assessment of the matrix $\mathbf{B}$, which can be computed from the design matrix $\mathbf{A}$ using linear algebra subroutines. However, the transition from the Lomb-Scargle periodogram to the LS-HE periodogram is a fundamental change in methodology. Because the LS-HE periodogram can only indicate disagreements between hypotheses on the functional and stochastic models and the observations, one cannot simply identify the stochastic process from the power-spectrum of the residuals. Instead, with the LS-HE periodogram, one must start with initial hypotheses about the functional and stochastic model, and adapt either the functional or the stochastic model based on the shape of the periodogram. A flat LS-HE periodogram at all frequencies indicates a good agreement between the model and the observations.

Although systematic biases can be avoided using the RMLE method, the presence of offsets in GNSS position time series is not without consequences on variance component estimation. In Figures (9.3) and (9.5) we observe a clear flattening of the restricted log-likelihood for the parameter ϕ and τ. This flattening suggests that the uncertainty of the parameter ϕ, but possibly other low-frequency parameters, are influenced by the number of offsets. This hypothesis is confirmed in section 9.5.1, in which the uncertainty of the parameter ϕ increases by +31.2% and the uncertainty of the random walk amplitude σ^2_{rw} increases +64.5% when considering 1.0 opy compared to no offsets. Hence, an excessive number of offsets could prevent the identification of the low-frequency properties of the stochastic variability in position time series, and thereby limit our understanding of its origin.

In this study we only considered daily and continuous observations over 20 years and specific noise levels for presentation purpose. However, the overall methodology can be applied to virtually any functional and stochastic model, as well as more realistic time series. In particular, a systematic estimation of variance component uncertainties could limit the over-interpretation of RMLE estimates, especially for the parameter ϕ and the random walk noise amplitudes.

9.7 Conclusion

With the increasing span of GNSS position time series, it is possible that the most frequently used stochastic model, which consists in a combination of white noise and a power-law process, could evolve to account for new low-frequency properties. Among the considered low-frequency properties are detection of random walk processes, characterized by an increase of power spectrum of the variability

at low frequency and detection of Generalized Gauss Markov processes, characterized by power-law properties at high frequencies and white noise properties at low frequency.

In this chapter, we investigated the influence of offsets on the biases and uncertainties of statistical tools used for the determination of the stochastic model of geodetic time series. We demonstrated that adding several offsets to the trajectory model of geodetic time series systematically biases the results obtained using the Lomb-Scargle periodogram and the MLE method. In particular, we showed that both the Lomb-Scargle periodogram and the MLE method erroneously promote detection of Generalized Gauss Markov processes because of offsets. This bias is due to the fact that neither the Lomb-Scargle periodogram nor the MLE method takes into account the way residuals are estimated. We showed that these biases can be avoided by using the LS-HE periodogram and the RMLE method for stochastic model investigation, which should be considered for future studies.

Although estimation biases can be avoided, the presence of offsets is not without consequence on variance components estimation. In particular, even when using the RMLE method, offsets increase the uncertainty of low-frequency parameters more than any other variance components. Hence, in addition to being a nuisance for the precision of functional parameters, the presence of offsets also systematically limits our ability to assess the low-frequency properties of the stochastic variability in GNSS position time series.

Chapter 10

Spatial and temporal properties of the stochastic variability in the position time series of over 10,000 stations.

So far, chapters 8 and 9 only addressed methodological aspects of the stochastic modelling of GNSS time series. In this chapter, we now aim at actually imaging the spatial and temporal properties of the stochastic variability in both the North, East, Vertical displacement time series of over 10,000 globally distributed GNSS stations and computed by the Nevada Geodetic Laboratory. Through the joint analysis of the three coordinates of 10,000 stations, this chapter aims at providing new insights on the possible origin of this stochastic variability in GNSS time series in general, and in vertical land motion time series in particular.

10.1 Introduction

Thanks to the ever-increasing spatial density and longevity of Global Navigation Satellite System (GNSS) stations, it is now possible to monitor earth deformations with an unprecedented spatial resolution, which is useful to the understanding of geophysical processes such as plate tectonics (Argus and Heflin, 1995; Kreemer, Blewitt, and Klein, 2014), glacial isostatic adjustment (Nocquet, Calais, and Parsons, 2005; Schumacher et al., 2018), surface mass redistribution (Van Dam, Blewitt, and Heflin, 1994; Van Dam et al., 2001), or coastal vertical land motions (Wöppelmann et al., 2007).

To model the motion of each GNSS station, it is customary to decompose GNSS position time series as the sum of a deterministic signal and stochastic variability. The deterministic signal is described by a functional model, whereas a stochastic model describes the stochastic variability, which is often regarded as noise. The functional model – sometimes called trajectory model (Bevis and Brown, 2014) – typically consists in a set of known parametric functions – trends, step discontinuities, periodic oscillations and occasional nonlinear motions – depending on unknown functional parameters (Bos et al., 2020). The stochastic model consists of a set of stochastic processes with known covariance functions depending on unknown stochastic parameters.

Defining a realistic stochastic model is advisable for several reasons. (a) It allows deriving realistic parameter uncertainties, which is essential to assess the predictive

power of any statistical result. (b) It is required to define minimum-variance estimators, which are desirable for any application. (c) It allows performing realistic hypothesis testing, which reduces the risk of over-interpretation of the observations. And (d) it can shed some light about the nature of non-modelled processes, which could then lead to modelling improvements reducing the noise level of the observations.

In the late 90s, numerous studies showed that the stochastic variability in GNSS position time series is time-correlated, and that its power spectrum $P(f_k)$ tends to follow a power-law function of the time frequency f_k, that is (Zhang et al., 1997; Mao, Harrison, and Dixon, 1999; Calais, 1999; Williams et al., 2004)

$$P(f_k) \approx P_0 \left(\frac{f_k}{f_0} \right)^{\kappa}, \tag{10.1}$$

where both P_0 and f_0 are normalization constants, and κ is the spectral index (Agnew, 1992).

The spectral index κ is a scalar identifying the nature of the power-law correlation. Its value usually lies within $[-3, 1]$ in Earth sciences, and within $[-2, 0]$ in GNSS time series analysis. Some power-law processes with integer values of κ are known as white – that is, uncorrelated – noise ($\kappa = 0$), flicker noise ($\kappa = -1$), and random walk noise ($\kappa = -2$). Any power-law process with a $\kappa < 0$ is time-correlated. When $\kappa \leq -1$ the power-law process is also non-stationary, and its variance diverges over time. Accounting for the power-law behaviour of the noise in the stochastic model of GNSS position time series has proved to be essential to derive realistic uncertainties of time-dependent parameters, and especially for station velocities (Williams, 2003a; Santamaría-Gómez et al., 2011).

In practice, the stochastic variability observed in GNSS position time series is often not modelled with only one stochastic process, but with a combination of several stochastic processes. The most commonly used stochastic model for global solutions is a linear combination of a white noise ($\kappa = 0$) and a power-law process ($\kappa \approx -1$) with different amplitudes (Williams et al., 2004; Santamaría-Gómez et al., 2011; He et al., 2019). Note that this practice could evolve with the increasing span of GNSS time series. For instance, Generalized Gauss Markov (GGM) processes (Langbein, 2004) are sometimes used for long time series when the power-spectrum $P(f_k)$ of the stochastic variability show a flattening at low frequencies f_k.

In addition to the time correlation, the stochastic variability observed in GNSS position time series has shown to be spatially correlated (Wdowinski et al., 1997). In particular, in global solutions, significant correlations between stations are detected up to a few thousand kilometres of separation distances (Williams et al., 2004; Amiri-Simkooei, 2009; Amiri-Simkooei, 2013; Amiri-Simkooei, Mohammadloo, and Argus, 2017; Benoist et al., 2020). Just like temporal correlation affects the uncertainty of time-dependent parameters, spatial correlation impacts the uncertainty of space-dependent parameters (Razeghi, Amiri-Simkooei, and Sharifi, 2016).

Most studies acknowledge the spatial and temporal correlation of the stochastic variability, but its origin remains to be identified. By analysing different solutions, previous work showed that both the temporal and spatial correlation of the stochastic variability is influenced by the processing strategy (Rebischung et al.,

2016; Amiri-Simkooei, Mohammadloo, and Argus, 2017). This observation suggests that GNSS positioning errors play a significant role in the properties of the stochastic variability. On the other hand, other studies showed that geophysical deformations also exhibit spatial dependencies and power-law behaviours (Agnew, 1992; Santamaría-Gómez and Mémin, 2015; Rebischung et al., 2017). Hence, it also likely that a part of the stochastic variability may correspond to actual crustal deformations, and not artefacts. Besides, the large scale spatial correlations observed in global solutions also suggest that local effects, such as multi-path or monument instability, are not the primary contributor to this unidentified stochastic variability (Williams et al., 2004).

Our understanding of the stochastic variability in globally distributed GNSS position time series is presently limited to analyses considering only a few hundred stations (Amiri-Simkooei, Mohammadloo, and Argus, 2017; He et al., 2019). These studies are intrinsically limited by the sparsity of the network of stations, and investigating denser networks could help better imaging the stochastic variability. Moreover, geophysical studies now aim at combining the observations of several thousands of stations (Schumacher et al., 2018; Kreemer, Blewitt, and Klein, 2014), therefore, a realistic description of the stochastic variability in dense networks could also improve uncertainty assessment in such studies.

To provide an insight about the stochastic properties of position time series, this study investigates the temporal and spatial correlation in a network of over 10,000 globally distributed stations presenting at least 3 cumulative years of observations. To this end, we used the North, East and Up position time series provided by the Nevada Geodetic Laboratory (NGL) and processed at the University of Nevada, Reno (Blewitt, Hammond, and Kreemer, 2018).

In November 2019, the NGL released a fully reprocessed solution based on upgraded software, orbits, clocks, and models (Blewitt et al., 2019). Because such a reprocessing is a good opportunity to assess the influence of modelling improvements on the stochastic variability (Amiri-Simkooei, Mohammadloo, and Argus, 2017), in this chapter, we examine both the old and new NGL solutions to study the impact of the reprocessing on both the observed temporal and spatial correlation patterns.

10.2 NGL Solutions

The final North, East and Up daily position time series products are available at (http://geodesy.unr.edu). For both solutions, station positions are estimated using single-station precise point positioning with carrier phase ambiguity resolution (Zumberge et al., 1997). The following pieces of information about each solution are extracted from Blewitt et al. (2019).

The old solution (hereafter, NGL-IGS08) uses the Jet Propulsion Laboratory (JPL) GIPSY OASIS II software with orbits and clocks based on IGS Repro 2 standards and the IGS08 reference frame (Rebischung et al., 2012). For the tropospheric delay, the generic Global Mapping Function (GMF) is used (Böhm et al., 2006). The first-order ionospheric effects are removed using the ionospheric-free linear combination of the carrier phase and pseudorange measurements. Second-order ionospheric effects are not corrected in this solution. In 2018, the reference frame of the final products

switched from ISG08 to IGS14. Hence, for consistency, our noise estimations are restricted to station positions estimated before January 2018.

The new solution (hereafter, NGL-IGS14) uses JPL's new GipsyX 1.0 software (Bertiger et al., 2020) with orbits and clocks based on IGS Repro 3 standards and the IGS14 reference frame. For the tropospheric delay, the Vienna Mapping Function (VMF1) (Boehm, Werl, and Schuh, 2006) is used, with nominal dry and wet tropospheric delays from the same VMF1 grids that are computed from the ECMWF numerical weather model. An elevation-dependent data weighting is also used to reduce the impact of signal degradation for satellites closer to the horizon. Second-order ionospheric effects are now corrected using ionospheric TEC maps together with a model of Earth's magnetic field. For more details about the reprocessing we refer to Kreemer, Blewitt, and Davis (2020) and (http://geodesy.unr.edu/gps/ngl.acn.txt) To be consistent with the time-span of NGL-IGS08, only position estimates before January 2018 are considered.

10.3 Variance Component Estimation

Estimating unknown parameters of the stochastic model can be done using Variance Component Estimation (VCE) methods. In this study, we use the Least-Squares Variance Component Estimation (LS-VCE) method. Introduced by (Teunissen, 1988) and further developed by (Teunissen and Amiri-Simkooei, 2008) and Amiri-Simkooei, 2007, the LS-VCE is an unbiased and minimum-variance alternative to the more commonly used Maximum Likelihood Estimation (MLE) method (Zhang et al., 1997; Williams, 2008; Bos et al., 2008). The LS-VCE has already been used in its univariate and multivariate form for GNSS noise analyses (Amiri-Simkooei, Tiberius, and Teunissen, 2007; Amiri-Simkooei, 2009; Amiri-Simkooei, Mohammadloo, and Argus, 2017).

10.3.1 Least-Squares Variance Component Estimation

Noting $\mathsf{E}\{.\}$ and $\mathsf{var}\{.\}$ the expectation and variance operators, the LS-VCE considers a linear functional model $\mathsf{E}\{\underline{\mathbf{y}}\}$ and a linear stochastic model $\mathsf{var}\{\underline{\mathbf{y}}\} = \mathbf{Q}_{\mathbf{y}}$ such that

$$\mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x}, \quad \mathbf{Q}_{\mathbf{y}} = \mathbf{Q}_0 + \sum_{i=1}^{p} \sigma_i^2 \cdot \mathbf{Q}_i \tag{10.2}$$

where $\mathbf{Q}_0$ and $\mathbf{Q}_i$ are known $m \times m$ co-factor matrices, representing the covariance of each stochastic process, σ_i^2 are p unknown variance components representing their unknown amplitude, $\mathbf{A}$ is the $m \times n$ design matrix of the functional, that is, the Jacobian of the model of observation equations, and $\mathbf{x}$ is the unknown $n \times 1$ parameter vector of the functional model.

The LS-VCE provides an estimator of the $p \times 1$ variance component vector $\hat{\sigma} = \left[\hat{\sigma_1}^2 \cdots \hat{\sigma_p}^2\right]^T$ by solving a normal equations system $\mathbf{N}\hat{\sigma} = \underline{\mathbf{l}}$, in which the entries n_{ij} and l_i of the $p \times p$ normal matrix $\mathbf{N}$ and the $p \times 1$ vector $\underline{\mathbf{l}}$ are given by (Amiri-Simkooei, 2007)

$$n_{ij} = \frac{1}{2}\mathrm{tr}(\mathbf{Q}_i\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{Q}_j\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{P}_{\mathbf{A}}^{\perp}), \tag{10.3}$$

$$\underline{l}_i = \frac{1}{2}(\underline{\hat{\mathbf{e}}}^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{Q}_i \mathbf{Q}_\mathbf{y}^{-1} \underline{\hat{\mathbf{e}}}) - \frac{1}{2}\mathrm{tr}(\mathbf{Q}_i \mathbf{Q}_\mathbf{y}^{-1} \mathbf{P}_\mathbf{A}^\perp \mathbf{Q}_0 \mathbf{Q}_\mathbf{y}^{-1} \mathbf{P}_\mathbf{A}^\perp). \tag{10.4}$$

In equations (10.3) and (10.4), $\underline{\hat{\mathbf{e}}}$ is the residual vector, $\mathbf{P}_\mathbf{A}^\perp$ is an orthogonal projector defined by $\mathbf{P}_\mathbf{A}^\perp = \mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{Q}_\mathbf{y}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_\mathbf{y}^{-1}$ so that $\underline{\hat{\mathbf{e}}} = \mathbf{P}_\mathbf{A}^\perp \underline{\mathbf{y}}$, and $\mathbf{I}$ is the $m \times m$ identity matrix.

Assuming the normal matrix $\mathbf{N}$ invertible, an estimator of the variance component vector, noted $\underline{\hat{\sigma}}$, can be obtained as

$$\underline{\hat{\sigma}} = \mathbf{N}^{-1} \underline{\mathbf{l}}. \tag{10.5}$$

Since the unknown covariance matrix $\mathbf{Q}_\mathbf{y}$ is involved in the definition of n_{ij} and $\underline{l}_i$, the estimation is done in an iterative way. One starts with a first guess on the variance component vector σ_0 to approximate $\mathbf{Q}_\mathbf{y}$ and obtain n_{ij} and $\underline{l}_i$. σ_0 can be obtained from previous publications, simplified VCE methods, or a simple guess. By solving the normal equations system (Equation 10.5), a new variance component vector $\underline{\hat{\sigma}}_1$ can be obtained. Then, by iterating this procedure, each time a variance component $\underline{\hat{\sigma}}_i$ is estimated, a new estimate $\underline{\hat{\sigma}}_{i+1}$ can be obtained. The iterations stop when the difference between two successive variance component estimates become negligible. For more details about the implementation of the LS-VCE method, we refer to Amiri-Simkooei (2007).

As the LS-VCE method is unbiased and unconstrained, it can occasionally provide negative estimates of some variance components. When encountering negative estimates, the Non-Negative LS-VCE method (NNLS-VCE) introduced by Amiri-Simkooei (2016) can be applied to obtain more interpretable values. Note that the non-negativity property usually conflicts with the unbiased property of the LS-VCE estimator, so the use of the NNLS-VCE method should ideally be restricted to the case of non-negativity issues.

10.3.2 Selection of stochastic models

The linear LS-VCE and NNLS-VCE provide estimates of the unknown variance component $\hat{\sigma}_i^2$ for a given linear stochastic model. When several stochastic models are considered for a given time series, one can use the LS-VCE to obtain estimates of $\mathbf{Q}_\mathbf{y}$ for each model and select the model that maximizes the restricted log-likelihood criterion $\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\sigma))$ defined by (Koch, 1986)

$$\ln(p_{\underline{\mathbf{t}}}(\mathbf{t}|\sigma)) = -\frac{1}{2}[(m-n)\ln 2\pi + \ln(\det(\mathbf{Q}_\mathbf{t})) + \mathbf{t}^T \mathbf{Q}_\mathbf{t}^{-1} \mathbf{t}] \tag{10.6}$$

where $\underline{\mathbf{t}}$ is the $(m-n) \times 1$ misclosure vector defined by

$$\underline{\mathbf{t}} = \mathbf{B}^T \underline{\mathbf{y}}, \tag{10.7}$$

$\mathbf{Q}_\mathbf{t}$ is its $(m-n) \times (m-n)$ covariance matrix that read

$$\mathbf{Q}_\mathbf{t} = \mathbf{B}^T \mathbf{Q}_\mathbf{y} \mathbf{B}, \tag{10.8}$$

and $\mathbf{B}$ is a $m \times (m-n)$ matrix that spans the null space of $\mathbf{A}^T$.

10.4 Spatial correlation

10.4.1 Spatial correlation estimation strategy

In previous work on spatial correlation, various spatial correlation estimation strategies have been used. Williams et al. (2004) computed the variation of the cross-correlation between pairs of residual time series with the station's separation distances, for all three coordinates. Amiri-Simkooei (2009) and Amiri-Simkooei, Mohammadloo, and Argus (2017) performed a similar analysis but used a particular metric to take time correlation into account. Note that this strategy, based on the multivariate LS-VCE (Amiri-Simkooei, 2009), assumes very similar stochastic properties for all time series. Finally, Benoist et al. (2020) assumed a second-order stationary spatial distribution of residuals and investigated the spatial correlation using variograms (Wackernagel, 2013). Benoist et al. (2020) also proposed to change of frame for the analysis of pairwise correlation because it is not rigorous to compare North and East coordinates between stations, especially when stations are close to the poles. Instead of analysing position residuals in the North and East coordinates, they proposed to analyse residuals in the extension/compression and shear directions.

In this chapter, we adopt the spatial correlation estimation method of Williams et al. (2004), but adapted to the change of frame proposed by Benoist et al. (2020). By doing so, we avoid assumptions about the spatial or temporal correlation of the observations and compare consistent quantities.

10.4.2 Estimation

Change of frame

In the following, we consider a pair of stations, noted S_i and S_j, with latitudes ϕ_i and ϕ_j, and longitudes λ_i and λ_j. From the time series analysis, one obtains position residuals in East (e), North (n), and Up (h) directions, that is, in enh-frames. The goal of change of frame is, for each pair of stations, to compare the residuals in uvh-frames depending on the relative position of the stations S_i and S_j, such that the estimated residuals are expressed in the extension/compression direction (u) and shear direction (v). The change from enh-frames to uvh-frames is illustrated in Figure (10.1).

The residuals of station S_i in the enh-frame, noted $\hat{\underline{e}}_i^{\mathsf{e}}(t_k)$, $\hat{\underline{e}}_i^{\mathsf{n}}(t_k)$, and $\hat{\underline{e}}_i^{\mathsf{h}}(t_k)$ can be transformed into residuals in the uvh-frame, noted $\hat{\underline{e}}_i^{\mathsf{u}}(t_k)$, $\hat{\underline{e}}_i^{\mathsf{v}}(t_k)$, and $\hat{\underline{e}}_i^{\mathsf{h}}(t_k)$ using

$$\begin{bmatrix} \hat{\underline{e}}_i^{\mathsf{u}}(t_k) \\ \hat{\underline{e}}_i^{\mathsf{v}}(t_k) \\ \hat{\underline{e}}_i^{\mathsf{h}}(t_k) \end{bmatrix} = \begin{bmatrix} \sin(\theta_i) & \cos(\theta_i) & 0 \\ -\cos(\theta_i) & \sin(\theta_i) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\underline{e}}_i^{\mathsf{e}}(t_k) \\ \hat{\underline{e}}_i^{\mathsf{n}}(t_k) \\ \hat{\underline{e}}_i^{\mathsf{h}}(t_k) \end{bmatrix} \tag{10.9}$$

in which θ_i is an azimuth angle that can be obtained from the geographic coordinates of the two stations as

$$\theta_i = \arctan\left(\frac{\sin(\lambda_j - \lambda_i)}{\cos(\phi_i)\tan(\phi_j) - \sin(\phi_i)\cos(\phi_j - \phi_i)}\right) \tag{10.10}$$

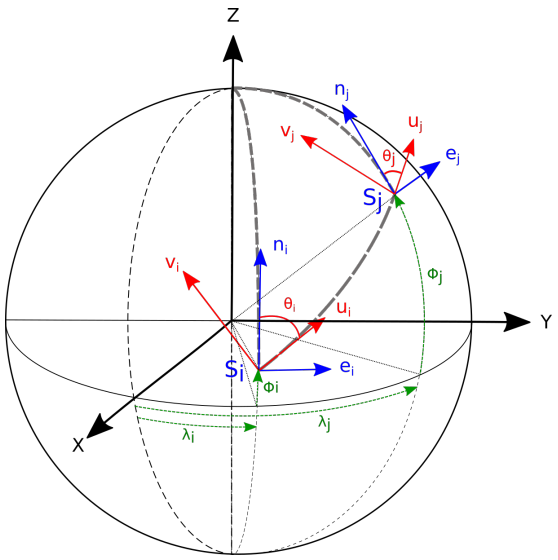

FIGURE 10.1: Illustration of the enh-frames and uvh-frames for two stations S_i and S_j.

Similarly, the residuals of station S_j in the enh-frame, noted $\underline{\hat{e}}_j^{\mathsf{e}}(t_k)$, $\underline{\hat{e}}_j^{\mathsf{n}}(t_k)$, and $\underline{\hat{e}}_j^{\mathsf{h}}(t_k)$ can be transformed into residuals in the uvh-frame, noted $\underline{\hat{e}}_j^{\mathsf{u}}(t_k)$, $\underline{\hat{e}}_j^{\mathsf{v}}(t_k)$, and $\underline{\hat{e}}_j^{\mathsf{h}}(t_k)$ using

$$\begin{bmatrix} \underline{\hat{e}}_j^{\mathsf{u}}(t_k) \\ \underline{\hat{e}}_j^{\mathsf{v}}(t_k) \\ \underline{\hat{e}}_j^{\mathsf{h}}(t_k) \end{bmatrix} = \begin{bmatrix} \sin(\theta_j) & \cos(\theta_j) & 0 \\ -\cos(\theta_j) & \sin(\theta_j) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \underline{\hat{e}}_j^{\mathsf{e}}(t_k) \\ \underline{\hat{e}}_j^{\mathsf{n}}(t_k) \\ \underline{\hat{e}}_j^{\mathsf{h}}(t_k) \end{bmatrix} \tag{10.11}$$

in which θ_j can be obtained as

$$\theta_j = \pi + \arctan\left(\frac{\sin(\lambda_i - \lambda_j)}{\cos(\phi_j)\tan(\phi_i) - \sin(\phi_j)\cos(\phi_i - \phi_j)}\right) \tag{10.12}$$

Pairwise correlation estimation

From the residuals of both stations in the uvh-frame, one can compute a pairwise correlation, noted $\underline{\hat{\rho}}_{ij}^{\mathsf{u,v,h}}$, as

$$\begin{aligned} \underline{\hat{\rho}}_{ij}^{\mathsf{u}} &= \frac{\underline{\hat{\mathrm{cov}}}_{ij}^{\mathsf{u}}}{\sqrt{\underline{\hat{\mathrm{var}}}_i^{\mathsf{hz}}\underline{\hat{\mathrm{var}}}_j^{\mathsf{hz}}}} \\ \underline{\hat{\rho}}_{ij}^{\mathsf{v}} &= \frac{\underline{\hat{\mathrm{cov}}}_{ij}^{\mathsf{v}}}{\sqrt{\underline{\hat{\mathrm{var}}}_i^{\mathsf{hz}}\underline{\hat{\mathrm{var}}}_j^{\mathsf{hz}}}} \\ \underline{\hat{\rho}}_{ij}^{\mathsf{h}} &= \frac{\underline{\hat{\mathrm{cov}}}_{ij}^{\mathsf{h}}}{\sqrt{\underline{\hat{\mathrm{var}}}_i^{\mathsf{h}}\underline{\hat{\mathrm{var}}}_j^{\mathsf{h}}}} \end{aligned} \tag{10.13}$$

with

$$\begin{aligned}
\underline{\hat{\mathrm{cov}}}_{ij}^{\mathsf{u}} &= \frac{1}{K_{ij}} \sum_{k=1}^{K_{ij}} \underline{\hat{e}}_i^{\mathsf{u}}(t_k) \cdot \underline{\hat{e}}_j^{\mathsf{u}}(t_k), \\
\underline{\hat{\mathrm{cov}}}_{ij}^{\mathsf{v}} &= \frac{1}{K_{ij}} \sum_{k=1}^{K_{ij}} \underline{\hat{e}}_i^{\mathsf{v}}(t_k) \cdot \underline{\hat{e}}_j^{\mathsf{v}}(t_k), \\
\underline{\hat{\mathrm{cov}}}_{ij}^{\mathsf{h}} &= \frac{1}{K_{ij}} \sum_{k=1}^{K_{ij}} \underline{\hat{e}}_i^{\mathsf{h}}(t_k) \cdot \underline{\hat{e}}_j^{\mathsf{h}}(t_k),
\end{aligned} \tag{10.14}$$

and

$$\begin{aligned}
\underline{\hat{\mathrm{var}}}_i^{\mathsf{hz}} &= \frac{1}{2K_{ij}} \sum_{k=1}^{K_{ij}} \underline{\hat{e}}_i^{\mathsf{u}}(t_k)^2 + \underline{\hat{e}}_i^{\mathsf{v}}(t_k)^2 \\
\underline{\hat{\mathrm{var}}}_i^{\mathsf{h}} &= \frac{1}{K_{ij}} \sum_{k=1}^{K_{ij}} \underline{\hat{e}}_i^{\mathsf{h}}(t_k)^2
\end{aligned} \tag{10.15}$$

where K_{ij} is the number of simultaneous observations of stations S_i and S_j over the considered period.

Average spatial correlation estimation

From the pairwise correlation estimates $\underline{\hat{\rho}}_{ij}^{\mathsf{u,v,h}}$, one can then assess the average spatial correlation $\underline{\hat{\rho}}^{\mathsf{u,v,h}}(d_l)$ at a given class of distance d_l as

$$\underline{\hat{\rho}}^{\mathsf{u,v,h}}(d_l) = \mathsf{mean}\{\{\underline{\hat{\rho}}_{ij}^{\mathsf{u,v,h}}\}_{d(i,j)\in\mathcal{D}_l}\} \tag{10.16}$$

in which $\{\hat{\rho}_{ij}^{\mathsf{u,v,h}}\}_{d(i,j)\in\mathcal{D}_l}$ is the set of correlation estimates for all pairs of station (i,j) whose separation distance $d(i,j)$ falls within an admissible region $\mathcal{D}_l$ defined as

$$\mathcal{D}_l = \left]d_l - \frac{\delta_l}{2}, d_l + \frac{\delta_l}{2}\right] \tag{10.17}$$

with δ_l the width of the averaging window.

10.4.3 Time dependence analysis

In the above cited studies, the spatial correlation is estimated using all observations over a span of a few years. In result, the estimated correlation is an average over the considered period. However, for shorter time scales, let's say a few weeks, one can expect the spatial correlation structure to show a time variable behaviour, especially if position residuals are related to the climatic variability.

To investigate the time dependence of the spatial correlation, we perform correlation analyses over running windows of one month. The pairwise correlation obtained from a one month window is certainly less precise than for a window of several years. However, given the number of stations analysed, the final correlation estimates is an average over a few hundred thousands pairs, which makes the result more robust.

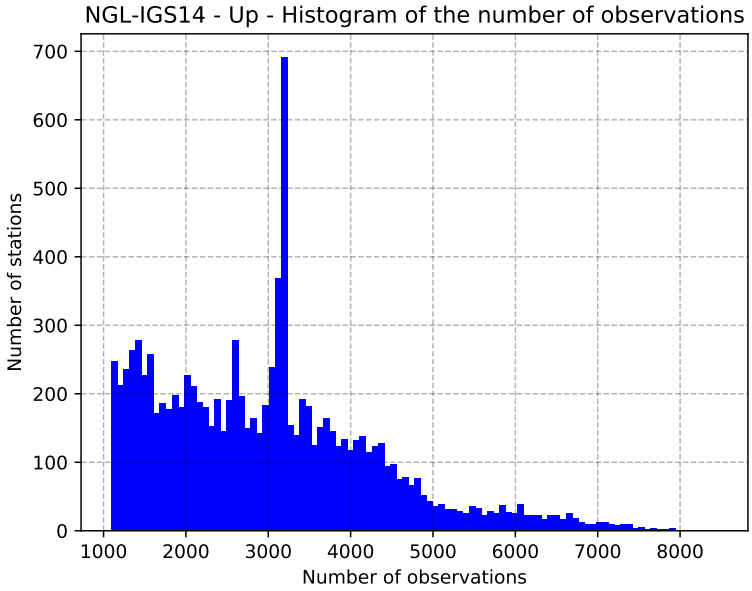

FIGURE 10.2: Histogram of the number of stations used as a function of the number of observations for the NGL-IGS14 solution and the Up coordinate.

10.5 Time series analysis

10.5.1 Selection of observations

We did not consider epochs showing a formal error above 5 mm on any of the three coordinates to avoid possibly unreliable observations. Moreover, the raw observations were compared to a running monthly median to detect large outliers. Any epoch showing a deviation from the median exceeding 5 times the median absolute deviation was also discarded.

In addition, to focus on stations that could reveal time correlation, we only considered stations presenting time series with over 1090 observations, that is, about 3 years of continuous observations. The distribution of the number of stations used as a function of the number of observations is presented for the NGL-IGS14 solution and the Up coordinate in Figure (10.2). In the end, the median cumulative span of the considered time series is 8 years.

10.5.2 Stochastic model

Assuming that the noise in each GNSS time series is a combination of white noise and a power-law process, the stochastic model reads

$$\mathbf{Q}_\mathbf{y} = \sigma_w^2 \mathbf{I} + \sigma_{pl}^2 \mathbf{Q}_\kappa. \tag{10.18}$$

where σ_w^2 is the unknown white noise amplitude, $\mathbf{I}$ is the $m \times m$ identity matrix, σ_{pl}^2 is the unknown power-law process amplitude, κ is the unknown spectral index, and $\mathbf{Q}_\kappa$ is the $m \times m$ Hosking's covariance matrix associated to a power-law process of spectral index κ.

The Hosking's covariance matrix $\mathbf{Q}_\kappa$ can be obtained as

$$\mathbf{Q}_\kappa = \Delta_t^{-\kappa/2} \mathbf{L}_\kappa \mathbf{L}_\kappa^T, \tag{10.19}$$

in which $\Delta_t = 1/365.25$ the sampling period in year and $\mathbf{L}_\kappa$ is the Hosking combination matrix. $\Delta_t^{-\kappa/2}$ is a normalization factor introduced by (Williams, 2003a) for spectral visualisation, and now used in most studies. The Hosking combination matrix $\mathbf{L}_\kappa$ is defined as

$$\mathbf{L}_\kappa = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{m-1} & \cdots & h_1 & h_0 \end{bmatrix}, \tag{10.20}$$

where the entries h_i are given by a recursive formula (Kasdin, 1995):

$$\begin{aligned} h_0 &= 1, \\ h_i &= (i - \frac{\kappa}{2} - 1)\frac{h_{i-1}}{i} \end{aligned} \tag{10.21}$$

For a fixed spectral index κ, the stochastic model (10.18) is a special case of the general model given in Equation (10.2) in which $p = 2$ and $\mathbf{Q}_0 = \mathbf{0}$. If the observations are expressed in millimetre (mm), σ_{pl} is expressed in mm yr$^{\kappa/4}$. Therefore, two σ_{pl} obtained with different spectral indices are not directly comparable.

To overcome this issue, we define a standard deviation $\hat{\sigma}'_{pl}$, expressed in mm, representing the square root of the expectation of the empirical variance of a power-law process for 8 years, which is the median cumulative span of the time series. Noting $\underline{\mathbf{r}}$ the random vector defined such that $\underline{\mathbf{r}} \sim \mathcal{N}(\mathbf{0}, \hat{\sigma}_{pl}\mathbf{Q}_\kappa)$, the amplitude $\hat{\sigma}'_{pl}$ can be obtained as

$$\hat{\sigma}'_{pl} = \sqrt{\mathsf{E}\left\{\frac{1}{m'}\sum_{i=1}^{m'}(\underline{r}_i - \bar{\underline{r}})^2\right\}} \tag{10.22}$$

where $\underline{r}_i$ is an element of the vector $\underline{\mathbf{r}}$, m' is number of observations for a 8 years long time series, and $\bar{\underline{r}}$ is the arithmetic mean of the $\underline{r}_i$.

One can show that $\hat{\sigma}'_{pl}$ reads

$$\hat{\sigma}'_{pl} = \hat{\sigma}_{pl}\sqrt{\frac{1}{m'}\mathsf{tr}\{\mathbf{Q}_\kappa\} - \frac{1}{m'^2}\mathbf{u}^T\mathbf{Q}_\kappa\mathbf{u}} \tag{10.23}$$

with $\mathbf{u}$ a $m' \times 1$ vector defined as $\mathbf{u}^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$.

10.5.3 Initial functional model

For simplification, this study is restricted to the use of linear functional models accounting for trends, step discontinuities, and periodic signals. The design matrix of functional $\mathbf{A}$ is identical for the three coordinates of a given station.

The list of periodic signals accounted for in the functional model includes the most commonly observed periodic signals in GNSS time series (Ray et al., 2008; Amiri-Simkooei, 2013), that is, the annual signal with a period of 365.25 days and

its 2 first harmonics, the draconitic signal with a period of 351.40 days and its 7 first harmonics, and the fortnightly signals with periods of 13.62, 14.17, and 14.76 days.

Also, all the offset dates indicated in the NGL master steps database have been considered by default. This database includes dates of known equipment changes and potential earthquake-related discontinuities. This database is not necessarily complete for all stations, and additional offsets must be detected to improve the functional model and obtain unbiased noise estimates.

10.5.4 Improving the functional model

The detection of offsets is crucial in stochastic parameter estimation because an under-parametrized functional model translates into time-correlated noise. It is also a key issue in GNSS time series analysis in general. Offset detection is often done manually since experimented operators can perform better than most algorithms (Gazeaux et al., 2013). However, for a few ten thousands time series, the manual detection is too cumbersome, and automatic detection is necessary.

In a recent study, Amiri-Simkooei et al. (2019) showed that one can improve the automatic detection of offsets by taking time-correlated noise into account, and by analysing the North (n), East (e) and Up (h) coordinates simultaneously. For our stochastic model, applying the multivariate LS-VCE used in Amiri-Simkooei et al. (2019) would assume a constant σ_w^2/σ_{pl}^2 ratio for all coordinates, which can be a strong assumption. In this study, we prefer adapting noise amplitudes to each coordinate and then check for discontinuities, including offsets, but also outliers and velocity changes.

The detection method used here is derived from the Detection Identification Adaptation (DIA) procedure introduced by Baarda (1968) and Teunissen (2000b). Applied to the detection of step discontinuities, outliers, or velocity changes, the DIA procedure consists in testing a null hypothesis $H_0 : \mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x}$ against alternative hypothesis $H_a : \mathsf{E}\{\underline{\mathbf{y}}\} = \mathbf{A}\mathbf{x} + \mathbf{c}_a(t_k)x_a$, in which $\mathbf{c}_a(t_k)$ is a known $m \times 1$ vector modelling a discontinuity of type a at time t_k with an unknown amplitude x_a. According to Teunissen, 2000b, accepting or rejecting H_0 depends on the value of a T-statistic defined by

$$\underline{T}_1 = \frac{(\mathbf{c}_a^T(t_k)\mathbf{Q}_{\mathbf{y}}{}^{-1}\underline{\hat{\mathbf{e}}})^2}{\mathbf{c}_a^T(t_k)\mathbf{Q}_{\mathbf{y}}{}^{-1}\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{c}_a(t_k)}, \tag{10.24}$$

Under H_0, the random variable $\underline{T}_1$ follows a central chi-squared distribution with 1 degree of freedom: $\underline{T}_1 \sim \chi^2(1,0)$. Hence, for a given confidence level α, if $\underline{T}_1 > \chi_\alpha^2(1,0)$, the null hypothesis can be rejected and $\mathbf{c}_a$ can be added to the columns of the design matrix $\mathbf{A}$.

To combine the information of all coordinates, we define the power $\underline{P}_a(t_k)$ of a discontinuity $\mathbf{c}_a(t_k)$ at time t_k as the sum of the T-statistics estimated on each coordinate, that is

$$\underline{P}_a(t_k) = \sum_{i\in\{\mathsf{n},\mathsf{e},\mathsf{h}\}} \frac{(\mathbf{c}_a^T(t_k)\mathbf{Q}_{\mathbf{y}_{(i)}}^{-1}\underline{\hat{\mathbf{e}}}_{(i)})^2}{\mathbf{c}_a^T(t_k)\mathbf{Q}_{\mathbf{y}_{(i)}}^{-1}\mathbf{P}_{\mathbf{A}_{(i)}}^{\perp}\mathbf{c}_a(t_k)}, \tag{10.25}$$

where the residual vector $\underline{\hat{\mathbf{e}}}_{(i)}$, the covariance matrix $\mathbf{Q}_{\mathbf{y}_{(i)}}$, and the orthogonal projector $\mathbf{P}_{\mathbf{A}_{(i)}}^{\perp}$ are specific to each coordinate (i).

Assuming that the coordinates are uncorrelated (Amiri-Simkooei, 2009; Benoist et al., 2020), under H_0, the power $\underline{P}_a(t_k)$ follows a central chi-squared distribution with 3 degrees of freedom, that is $\underline{P}_a(t_k) \sim \chi^2(3,0)$.

The chi-squared distribution of $\underline{T}_1$ and $\underline{P}_a(t_k)$ is based on the assumption that the covariance matrices $\mathbf{Q}_{\mathbf{y}_{(i)}}$ are known. However, in reality, the two variance components σ_w^2 and σ_{pl}^2 are unknown and must be estimated. In such a case, the exact distribution of $\underline{T}_1$ becomes complicated. Nonetheless, the chi-squared distribution remains a good approximation as the redundancy of the functional model $m - n$ is large (Amiri-Simkooei et al., 2019).

To find a good candidate for the alternative hypothesis $\mathbf{c}_a(t_k)$, offsets $\mathbf{c}_o(t_k)$, outliers $\mathbf{c}_t(t_k)$, and velocity changes $\mathbf{c}_v(t_k)$ were systematically tested for each epoch t_k. The most-likely discontinuity date $\hat{t}_k$ and type $\hat{a}$, were chosen as

$$\hat{a}, \hat{t}_k = \arg\max_{a,t_k} \underline{P}_a(t_k). \tag{10.26}$$

When the power of the most likely alternative $P_{\hat{a}}(\hat{t}_k)$ exceeded a given threshold $P_{\hat{a}}(\hat{t}_k) > 25$, the discontinuity $\mathbf{c}_{\hat{a}}(\hat{t}_k)$ was added to the functional model. The threshold of 25 is arbitrarily chosen value, intentionally large, and corresponding to a confidence level $\alpha > 0.9999$ under the $\chi^2(3,0)$ hypothesis. This detection procedure was repeated until no significant discontinuities were found.

At each iteration, the variance components of the stochastic model were (re)estimated for all coordinates. However, to reduce the computational burden, for this detection step only, we approximated κ as $\kappa \approx -0.8$ and used the NNLS-VCE method by default.

10.5.5 Variance component estimation strategy

To assess both variance components $\hat{\sigma}_w^2$, $\hat{\sigma}_{pl}^2$ and the spectral index $\hat{\kappa}$, we tested 30 stochastic models with values of κ ranging from -1.8 to -0.3 with a step of 0.05 for each time series. The best set of stochastic parameters $(\hat{\sigma}_w^2, \hat{\sigma}_{pl}^2, \hat{\kappa})$ was that maximizing the restricted log-likelihood criterion defined in section (10.3.2).

The LS-VCE was used by default to benefit from unbiased and minimum-variance estimates. But when encountering negativity issues, the NNLS-VCE was systematically applied to obtain interpretable estimates. This resulted in occasional null process amplitudes, mostly for the white noise amplitude of the Up coordinates.

Due to differences in the convergence of some time series and limitations in the allocation of the resources used in this work, the number of station analysed fluctuates depending on the considered solution and coordinates. However, there is always over 10,000 time series analysed in each case.

10.6 Results

10.6.1 Spatial variability of the stochastic parameters

The global distribution of variance component estimates are presented, for the NGL-IGS14 solution, in Figure (10.3) for the North coordinates, in Figure (10.4) for the East

coordinates and in Figure (10.5) for the Up coordinates. Each one of these figures exhibits global-scale variations in the distribution of the stochastic parameters. In a first approximation, these variations are latitude-dependent, especially for the Up coordinates. The results obtained with NGL-IGS08 solution show a similar latitude dependence, but noticeable changes in amplitudes. To illustrate this, the latitude dependence for both solutions, are presented in Figure (10.6) for $\hat{\sigma}_w$, in Figure (10.7) for $\hat{\kappa}$, and in Figure (10.8) for $\hat{\sigma}'_{pl}$. The latitude dependence of the more commonly used Root Mean Square Error (RMSE) is presented in Figure (10.9). To ease the comparison between the NLG-IGS08 and NGL-IGS14 solution, we reported the running median latitude dependence of the NGL-IGS08 on that of NGL-IGS14 solution.

In Figure (10.6), the latitude dependence of the white noise amplitude $\hat{\sigma}_w$ reveals an equatorial bulge for all coordinates and solutions. The bulge is particularly marked for the Up coordinate where $\hat{\sigma}_w \approx 4$ mm inside the inter-tropical band, and $\hat{\sigma}_w \approx 0$ mm outside. It can also be noted that the bulge is systematically smaller in the latest solution. Also, an asymmetry between the southern and northern hemispheres present in NGL-IGS08 is systematically reduced in NGL-IGS14.

In Figure (10.7), there is a clear difference in the latitude dependence of the spectral index $\hat{\kappa}$ between Up and horizontal coordinates. For horizontal coordinates, the median $\hat{\kappa}$ show a mild latitude dependence with a $-0.1/-0.2$ decrease near the equator. The Up coordinates rather show a bi-modal distribution with an increase of about $+0.55$ at mid-latitudes. In words, for the Up component, time-correlated processes tend to show longer memory near the equator and near the poles than at mid-latitudes. The effect of the reprocessing is relatively negligible for the North and East coordinates. For Up coordinates, the median $\hat{\kappa}$ shifted down by about -0.1 in the NGL-IGS14.

In Figure (10.8), the latitude dependence of the (modified) power-law process amplitude $\hat{\sigma}'_{pl}$ shows a strong latitude dependence for the Up coordinates, where $\hat{\sigma}'_{pl}$ estimates double near the poles. The most dramatic effect is visible for the NGL-IGS14 solution, which also reveals a peak for latitudes near $+65^\circ$ corresponding to a large increase of $\hat{\sigma}'_{pl}$ in Alaska and around the Baltic sea, as visible in (Figure 10.5). For horizontal coordinates, $\hat{\sigma}'_{pl}$ showed a North-South asymmetry in NGL-IGS08, which is corrected in NGL-IGS14.

10.6.2 Spatial correlation

For each direction, and each solution, we estimated the spatial correlation $\hat{\rho}(t_k, d_l)$ every 15 days between 2008 and 2018, with a time averaging window of 30 days, and for 199 classes of distance d_l ranging from 50 km to 19950 km, with a spatial averaging window δ_d of 100 km. The correlation estimates $\hat{\rho}(t_k, d_l)$ along with the time average over the 2008-2018 period $\bar{\rho}(d_l)$ are presented in Figure (10.10) for the NGL-IGS08 solution, and in Figure (10.11) for the NGL-IGS14 solution.

The average spatial correlation $\bar{\rho}(d_l)$ in Figure (10.11) is close to $+0.45$ for short separation distances (that is, $d_l \approx 0$ km) and decreases close to zero for larger separation distances (that is, $d_l \geq 7500$ km). This behaviour confirms that the stochastic variability in these data-sets shows a specific spatial correlation structure. To model this average spatial correlation, we approximated the mean correlogram $\bar{\rho}(d_l)$ by a Gaussian correlation function defined by

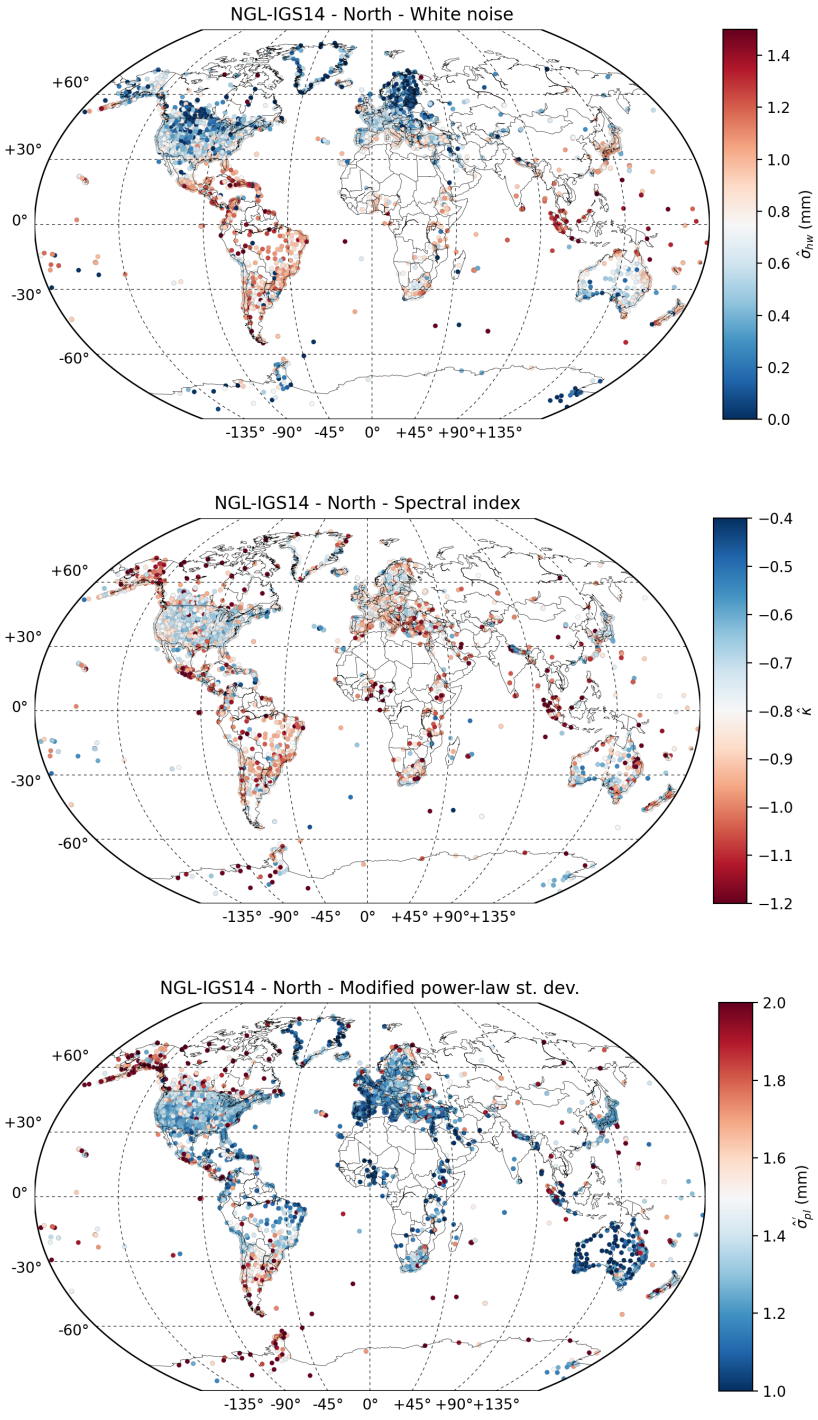

FIGURE 10.3: Global variability of white noise amplitudes (*top*), spectral indices (*middle*); and power-law amplitudes (*bottom*) for the North coordinates of the NGL-IGS14 solution.

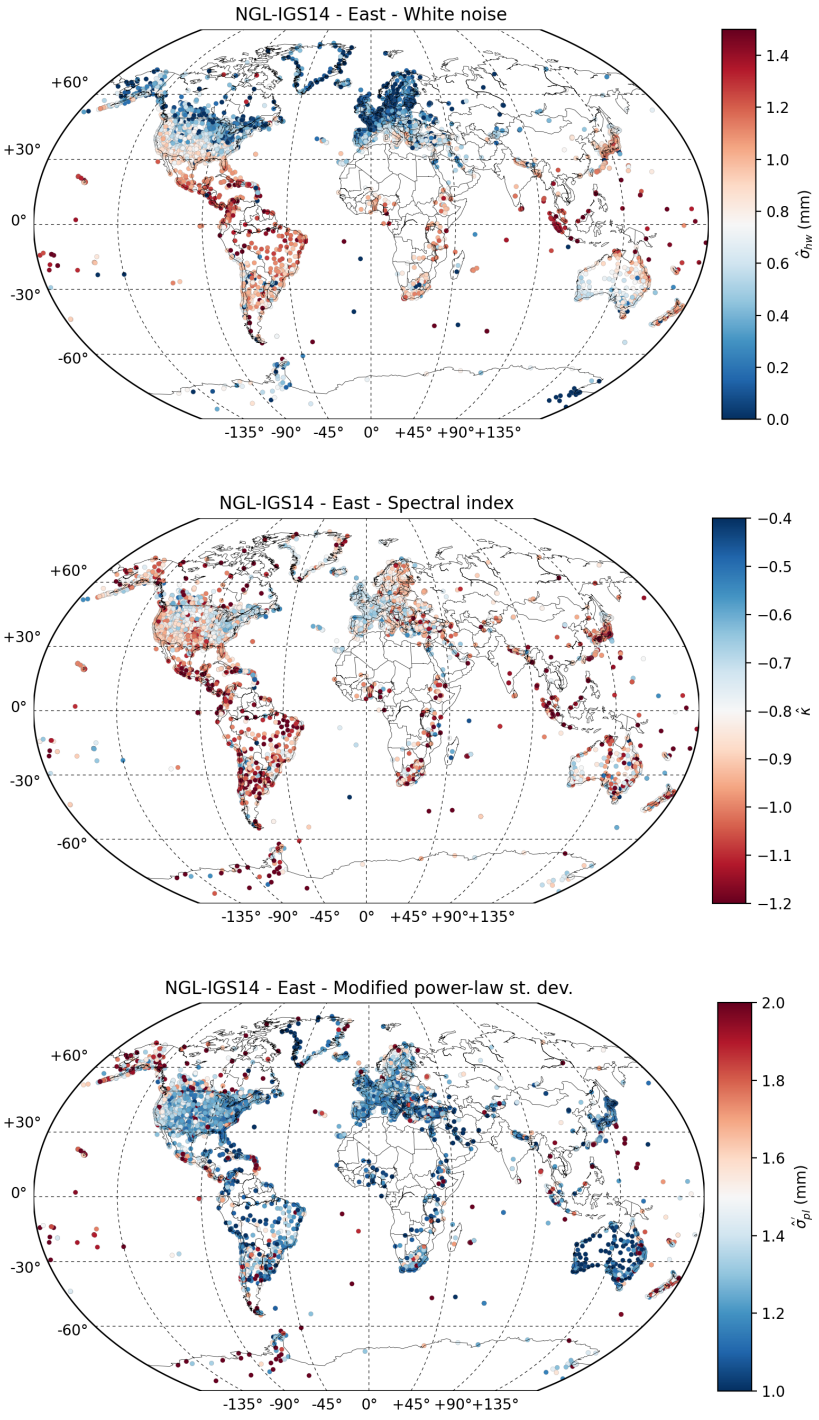

FIGURE 10.4: Global variability of white noise amplitudes (*top*), spectral indices (*middle*); and power-law amplitudes (*bottom*) for the East coordinates of the NGL-IGS14 solution.

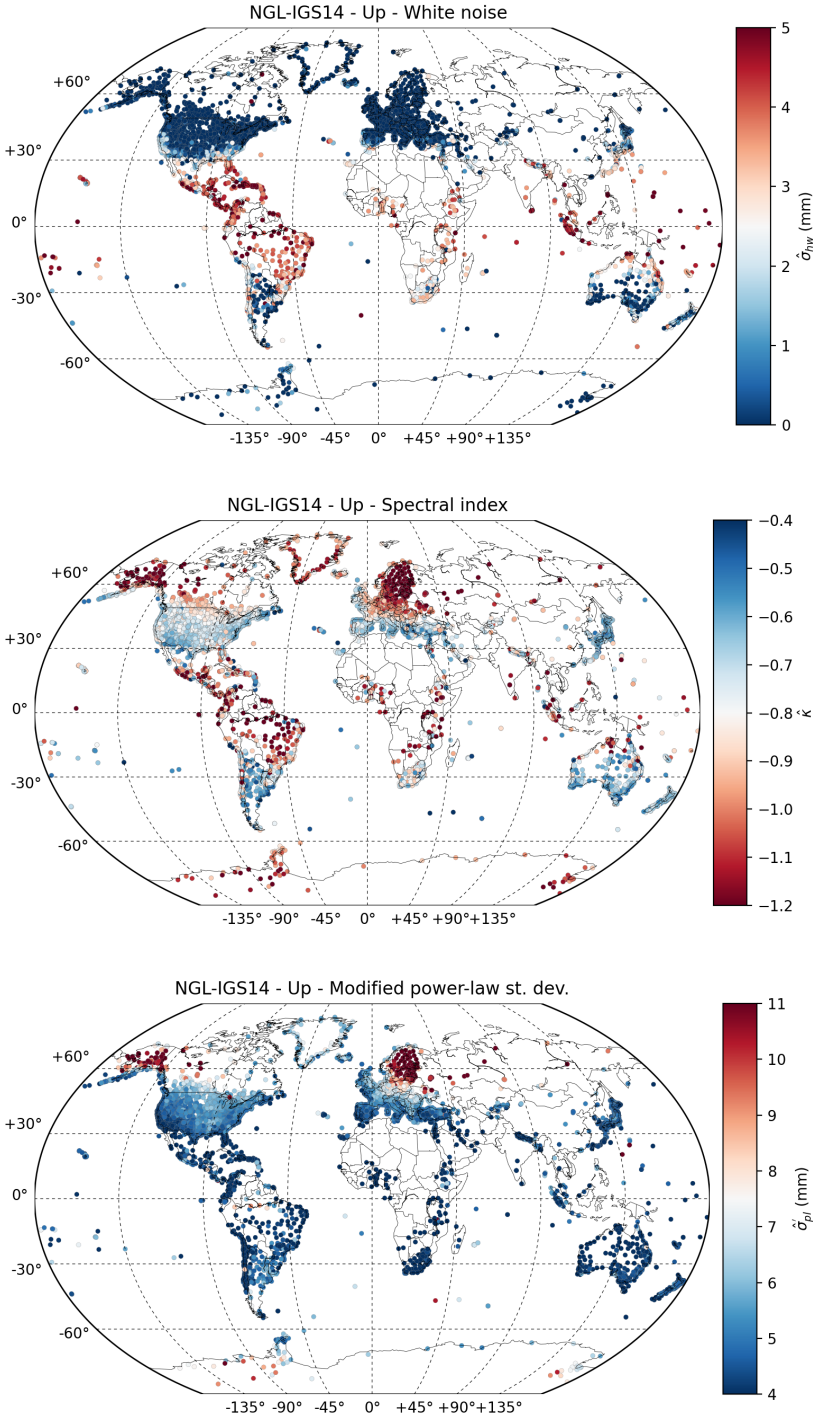

FIGURE 10.5: Global variability of white noise amplitudes (*top*), spectral indices (*middle*); and power-law amplitudes (*bottom*) for the Up coordinates of the NGL-IGS14 solution.

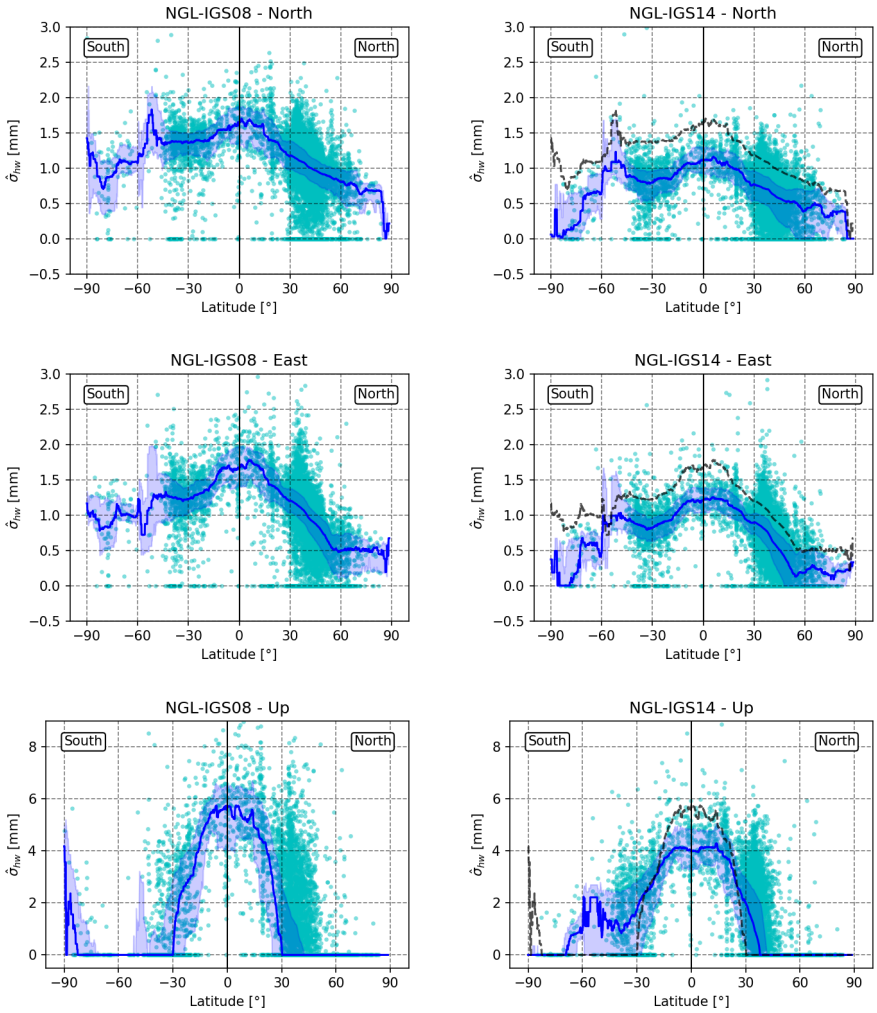

FIGURE 10.6: Latitude dependence of white noise amplitudes $\hat{\sigma}_w$ for the North coordinate (*top*), for the East coordinate (*middle*), for the Up coordinate (*bottom*), and for both the NGL-IGS08 solution (*left*) and NGL-IGS14 solution (*right*). The dots in cyan represent the raw estimates. The continuous blue lines represent running medians over centered 10° latitude windows. The shaded areas represent the corresponding interquartile ranges. The dashed black lines correspond to the running median for NGL-IGS08, reported on the NGL-IGS14 results to ease the comparison.

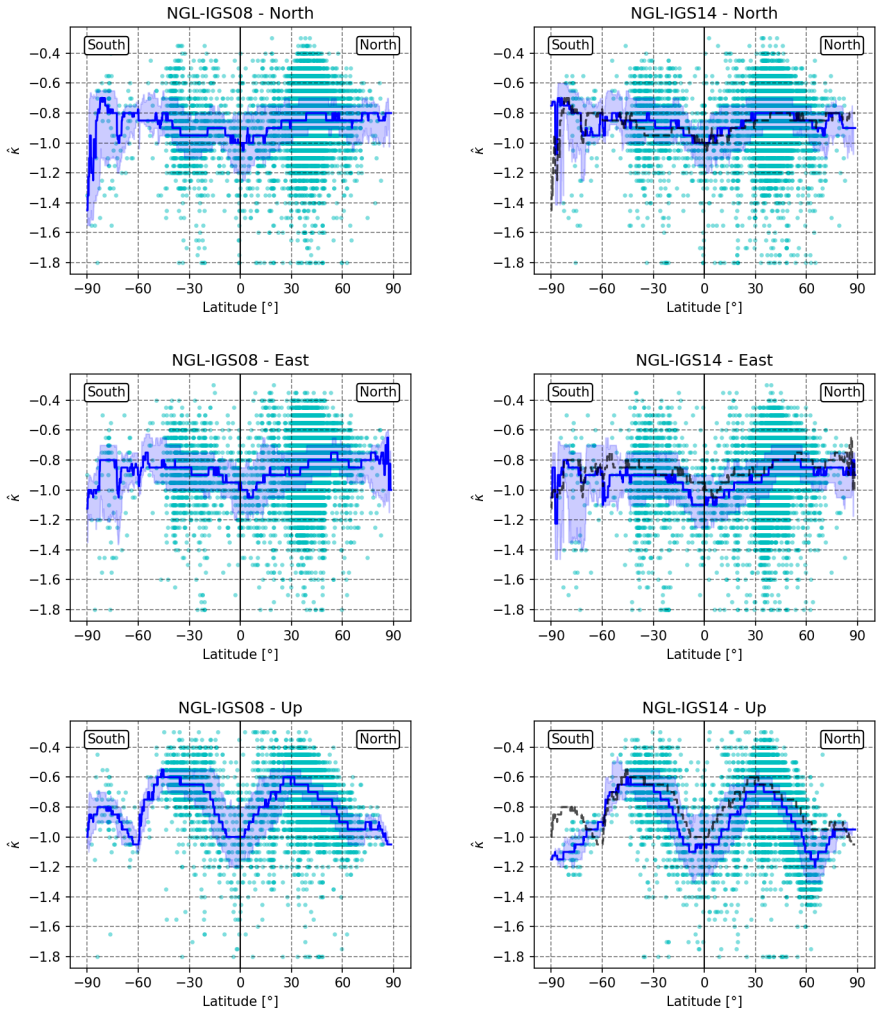

FIGURE 10.7: Latitude dependence of spectral indices $\hat{\kappa}$ for the North coordinate (*top*), for the East coordinate (*middle*), for the Up coordinate (*bottom*), and for both the NGL-IGS08 solution (*left*) and NGL-IGS14 solution (*right*). The dots in cyan represent the raw estimates. The continuous blue lines represent running medians over centered 10° latitude windows. The shaded areas represent the corresponding interquartile ranges. The dashed black lines correspond to the running median for NGL-IGS08, reported on the NGL-IGS14 results to ease the comparison.

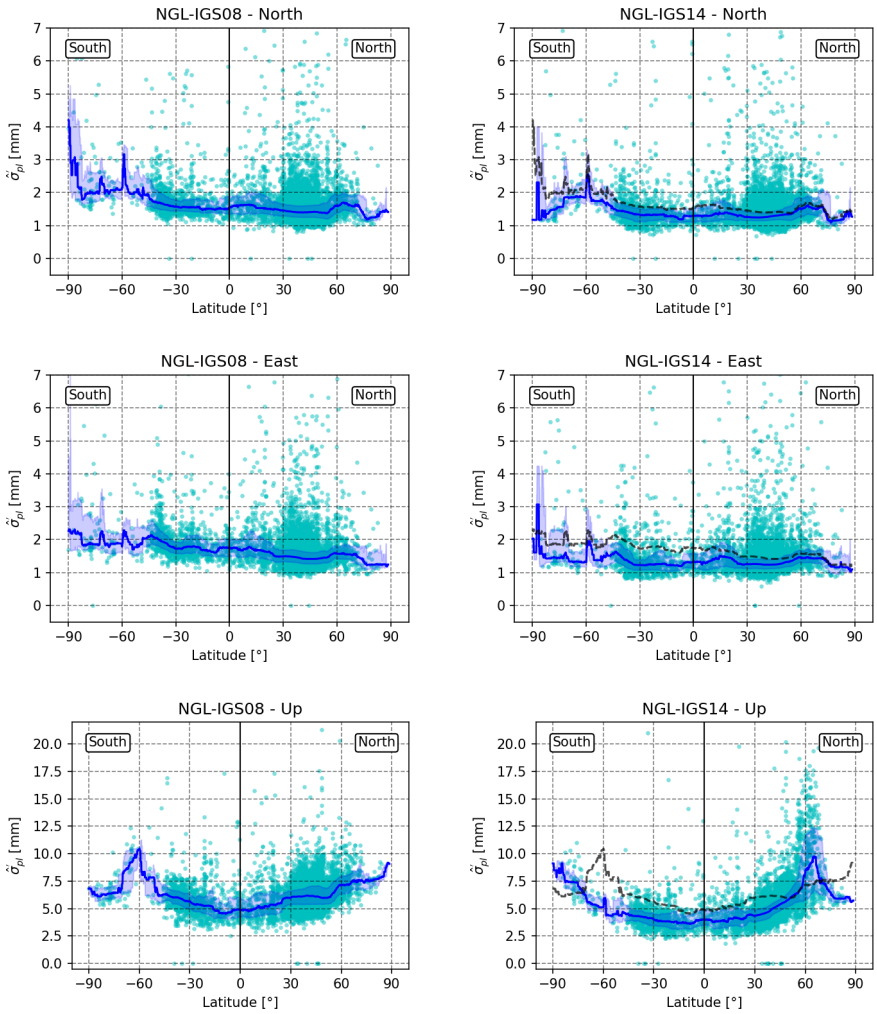

FIGURE 10.8: Latitude dependence of power-law process amplitudes $\hat{\sigma}'_{pl}$ for the North coordinate (*top*), for the East coordinate (*middle*), for the Up coordinate (*bottom*), and for both the NGL-IGS08 solution (*left*) and NGL-IGS14 solution (*right*). The dots in cyan represent the raw estimates. The continuous blue lines represent running medians over centered 10° latitude windows. The shaded areas represent the corresponding interquartile range. The dashed black lines correspond to the running median for NGL-IGS08, reported on the NGL-IGS14 results to ease the comparison.

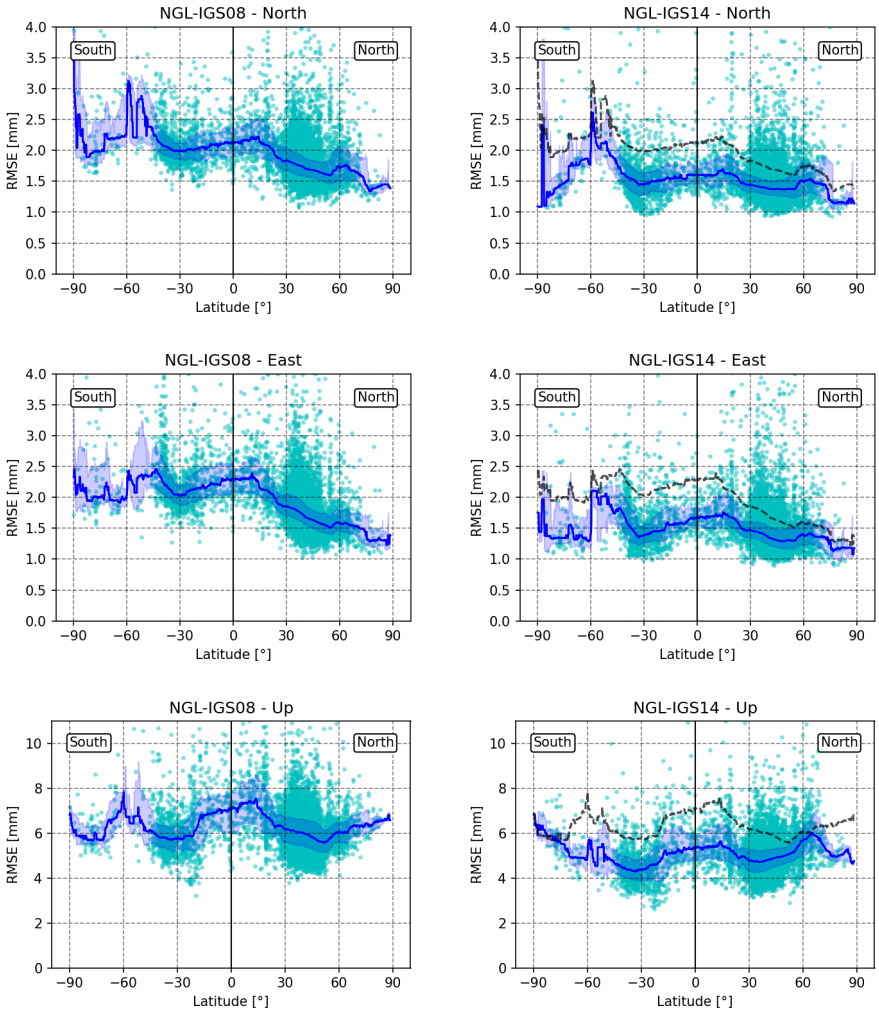

FIGURE 10.9: Latitude dependence of the Root Mean Square Error (RMSE), for the North coordinate (*top*), for the East coordinate (*middle*), for the Up coordinate (*bottom*), and for both the NGL-IGS08 solution (*left*) and NGL-IGS14 solution (*right*). The dots in cyan represent the raw estimates. The continuous blue lines represent running medians over centered 10° latitude windows. The shaded areas represent the corresponding interquartile ranges. For each time series, RMSE are based on residual estimates computed using the best stochastic model. The dashed black lines correspond to the running median for NGL-IGS08, reported on the NGL-IGS14 results to ease the comparison.

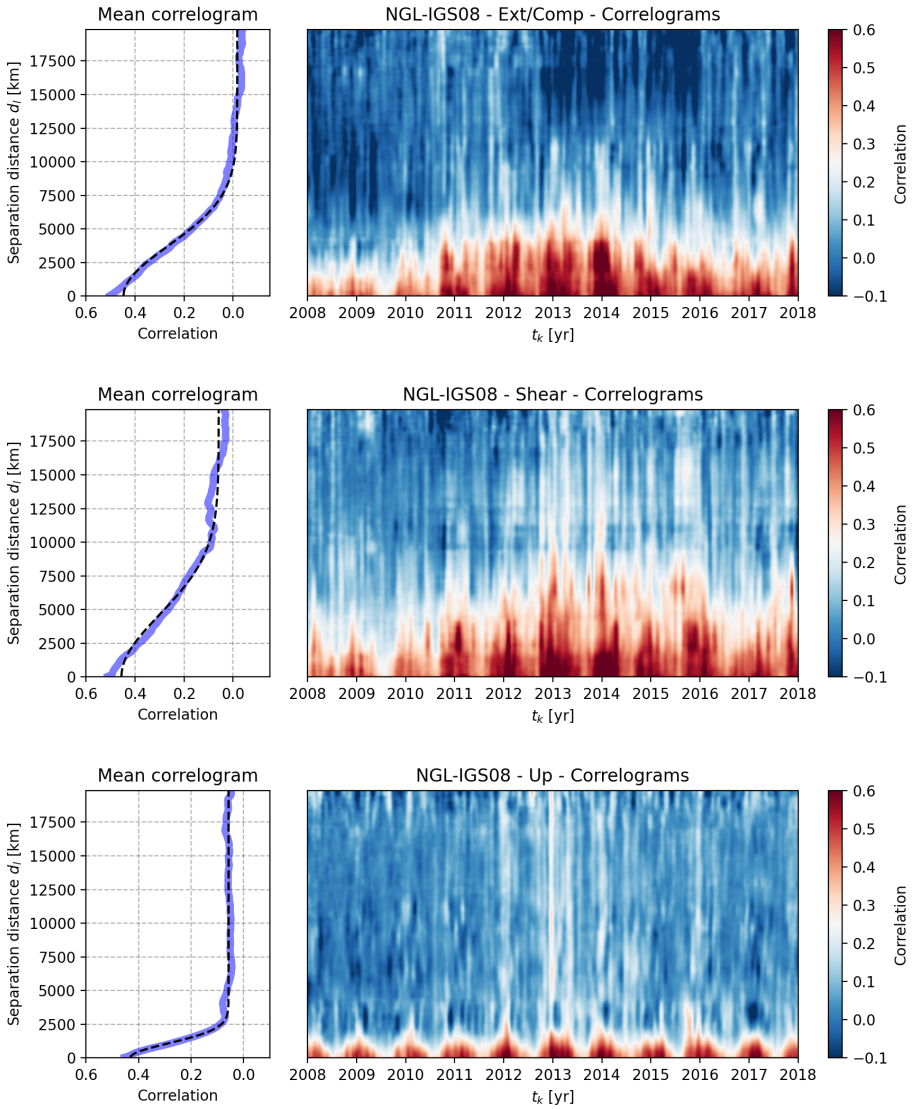

FIGURE 10.10: Spatial correlation estimates $\hat{\rho}(t_k, d_l)$ for the NGL-IGS08 solution. Extension/compression direction (*top*), Shear direction (*middle*) and Up direction *(bottom*). The blue curve represents the median correlogram. The dashed black line represents the Gaussian correlation function.

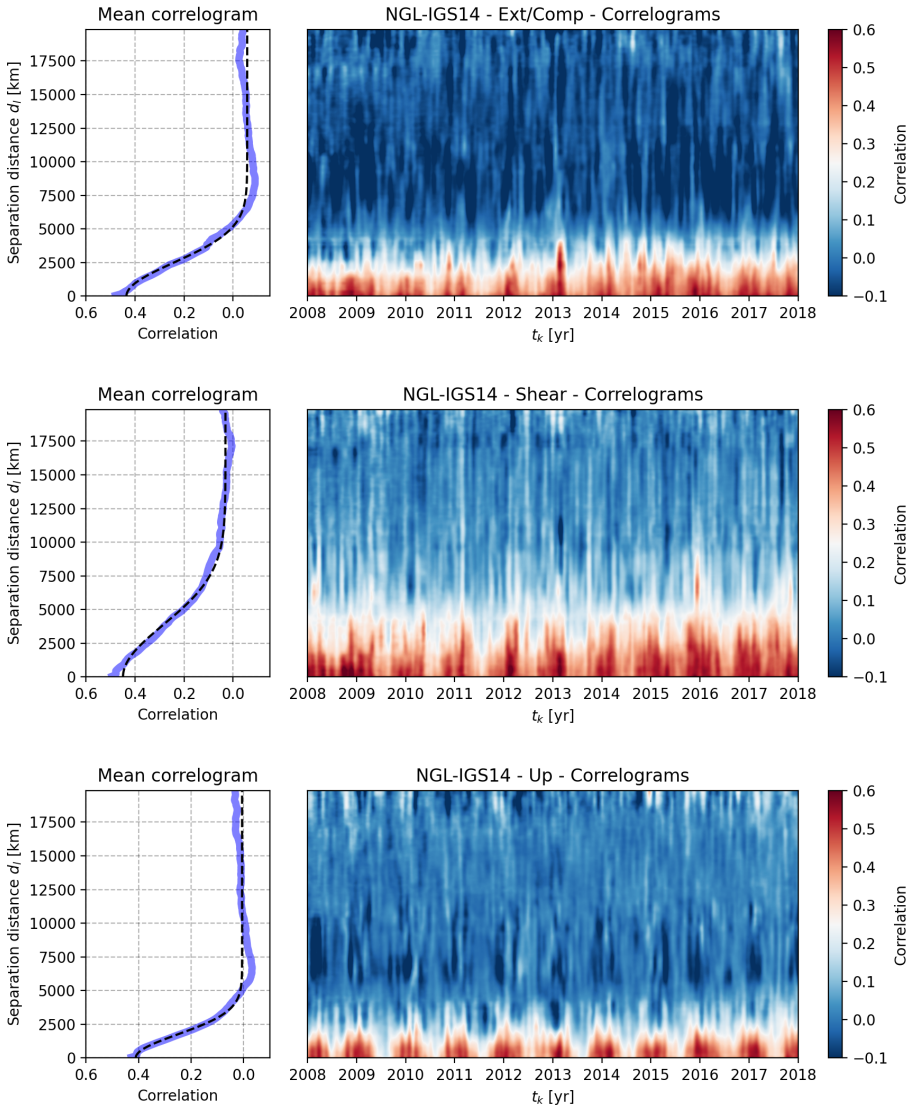

FIGURE 10.11: Spatial correlation estimates $\hat{\rho}(t_k, d_l)$ for the NGL-IGS14 solution. Extension/compression direction (*top*), Shear direction (*middle*) and Up direction *(bottom*). The blue curve represents the median correlogram. The dashed black line represents the Gaussian correlation function.

Solution	Direction	$\bar{\rho}_c$	$\bar{\rho}_0$	$\bar{\rho}_{\max}$	r [km]
NGL-IGS08	Ext/Comp	-0.02	0.46	0.44	5299
	Shear	0.06	0.40	0.46	6562
	Up	0.06	0.37	0.43	1551
NGL-IGS14	Ext/Comp	-0.06	0.49	0.43	3539
	Shear	0.03	0.42	0.45	5479
	Up	0.01	0.40	0.41	2470

TABLE 10.1: Parameter estimates for the average Gaussian correlation function (Equation 10.27).

$$\bar{\rho}(d_l) = \begin{cases} 1, & d_l = 0 \\ \bar{\rho}_c + \bar{\rho}_0 \cdot \exp\left(-\left(d_l/r\right)^2\right), & d_l > 0 \end{cases}, \tag{10.27}$$

where $\bar{\rho}_c$ represents the constant correlation, $\bar{\rho}_0$ the distance-variable correlation, and r the distance range parameter. From this Gaussian correlation function, a maximum of correlation can be defined as $\bar{\rho}_{\max} = \bar{\rho}_c + \bar{\rho}_0$.

Estimates for $\hat{\bar{\rho}}_c$, $\hat{\bar{\rho}}_0$, $\hat{\bar{\rho}}_{\max}$ and $\hat{r}$ resulting from the least-squares adjustment of the Gaussian correlation function on $\bar{\rho}(d_l)$ are presented, for all coordinates, in Table (10.1). The corresponding models are represented in Figures (10.10) and (10.11). In Table (10.1) $\hat{\bar{\rho}}_{\max}$ is similar for all coordinates. However, the range parameter $\hat{r}$ is systematically larger for the horizontal directions than for the vertical one. Note that the parameter r corresponds to the argument of the inflexion point of the Gaussian correlation function. To ease physical interpretations, one can define an effective range, $d_e = 2r$, which corresponds to the separation distance above which the distance-variable part of spatial correlation $\hat{\bar{\rho}}_0$ is decreased by 98%. For instance, in NGL-IGS14, the distance-variable part tend to be negligible for separation distances above $d_e \approx 4000$ km for the Up direction, and above $d_e \approx 7000$ km for horizontal directions (Figure 10.10 and 10.11).

With the reprocessing, the range parameter $\hat{r}$ systematically decreased for both compression/extension and shear directions and increased of the vertical direction. This change is particularly visible in Figure (10.10) for horizontal directions where a large scale correlation pattern, maximum in 2014, disappears in Figure (10.11).

In addition to the mean correlation, the time-variable correlograms in Figures (10.10) and (10.11) reveal an annual oscillation for short separation distances, that is $d_l < 4000$, and for all directions. An example of the estimated spatial correlation time series $\hat{\rho}(t_k, d_l)$ is presented in Figure (10.12) for separation distances of 50, 950, 1950, and 3950 km. Figure (10.12) shows that the amplitude of this annual oscillation is maximum for short distances ($d_l = 50$ km) and is synchronous for all distances classes. Besides, both the mean correlation and the annual amplitude gets closer to 0.00 for increasing separation distances. Hence, to describe the oscillation of $\hat{\rho}(t_k, d_l)$, we simply added an annual modulation to the distance-variable part of the mean Gaussian correlation function defined Equation (10.27), that is

Solution	Direction	$\hat{\rho}_c$	$\hat{\rho}_0$	$\hat{\rho}_A$	$\hat{\phi}_A$ [rad]	$\hat{\rho}_{\text{min}}$	$\hat{\rho}_{\text{max}}$	$\hat{r}$ [km]
NGL-IGS08	Comp/Ext	-0.02	0.46	0.06	1.709	0.38	0.50	5299
	Shear	0.06	0.40	0.07	1.509	0.39	0.53	6562
	Up	0.05	0.37	0.16	1.266	0.26	0.58	1551
NGL-IGS14	Comp/Ext	-0.06	0.49	0.05	1.273	0.38	0.48	3539
	Shear	0.03	0.42	0.06	1.311	0.39	0.51	5479
	Up	0.01	0.4	0.11	1.333	0.30	0.52	2469

TABLE 10.2: Parameter estimates for the time-variable Gaussian correlation function (Equation 10.28). The $\hat{r}$ are reported values from Table (10.1).

$$\rho(t_k, d_l) = \begin{cases} 1, & d_l = 0 \\ \rho_c + [\rho_0 + \rho_A \sin(\omega_A t_k + \phi_A)] \cdot \exp\left(-\left(d_l/r\right)^2\right), & d_l > 0 \end{cases} \tag{10.28}$$

where ρ_c is the constant correlation, ρ_0 is the average of the time-variable part of the correlation, ω_A is the annual angular frequency, ρ_A is the annual amplitude, ϕ_A is the corresponding phase, and $\hat{r}$ is the range parameter estimated from the mean correlogram. From these estimates, we can define a maximum and a minimum of correlation for $d_l \approx 0$ as $\rho_{\text{min}} = \rho_c + \rho_0 - \rho_A$ and $\rho_{\text{max}} = \rho_c + \rho_0 + \rho_A$.

Estimates of $\hat{\rho}_c$, $\hat{\rho}_0$, $\hat{\rho}_A$, $\hat{\phi}_A$, ρ_{min} and ρ_{max}, based on the least-squares adjustment of the model (10.28) on $\hat{\rho}(t_k, d_l)$ are presented in Table (10.2). The corresponding model is visible, for the Up coordinates of the NGL-IGS14 solution, in Figure (10.12). With an amplitude $\hat{\rho}_A$ of 0.16 for the NGL-IGS08 solution and of 0.11 for the NGL-IGS14 solution, Up coordinates show annual amplitudes about 2 times larger than horizontal coordinates for both solutions. According to the estimated phases $\hat{\phi}_A$, the correlation maximum lies around January and the minimum around July for all coordinates. In NGL-IGS14, the phase $\hat{\phi}_A$ are similar, but the annual amplitudes $\hat{\rho}_A$ are reduced compared to the NGL-IGS08 solution, especially for the Up coordinates, which shows a 31% reduction.

Given that most stations are either located in the USA and Europe, the observed phases and amplitudes of the short-distance time-variable correlation are probably representative of a phenomenon occurring in these areas rather than in other parts of the world. In particular, if the periodic nature of the correlation is linked to seasons through surface mass loading, a change in phase and amplitude with the location may occur.

To explore the possible spatial variability of this oscillation, we use the fact that an annual periodicity of the spatial correlation is likely to result in an annual periodicity of the variance of each station. Hence, from the estimated residual time series, we computed monthly standard deviations every 15 days for all stations. From the resulting standard deviation time series, we adjusted a trend and annual sine function. The annual amplitude and phase estimates obtained from time series presenting more than 4 years of cumulative observations are visible, for the NGL-IGS14 Up

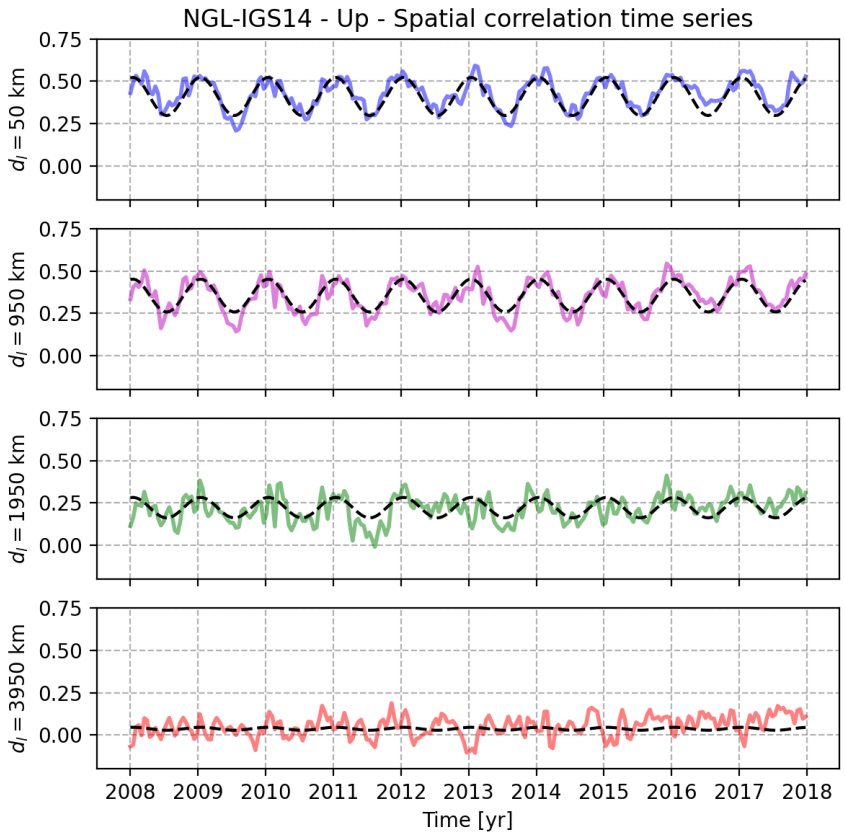

FIGURE 10.12: Time-variable spatial correlation estimates $\hat{\rho}(t_k, d_l)$ for 4 classes of distances of the Up coordinate of the NGL-IGS14 solution. The coloured lines represent the estimates at given class of distance d_l. The dashed black lines represent the corresponding time-variable Gaussian correlation function.

coordinates, in Figure (10.15). Figure (10.15) demonstrates that the annual amplitudes and phases of the Up coordinates have a spatial signature and a global-scale variability. In particular, between 0° and +30° of latitude, and in Antarctica, the standard deviation is not maximum in January but in June. Hence, other amplitudes and phases for the time-variable correlation should be expected for time-variable correlation analyses performed on different parts of the globe, or more evenly distributed networks of stations.

10.7 Discussion

10.7.1 Spatial dependence of noise parameters

Figures (10.3), (10.4) and (10.5) show that all stochastic parameters present a global-scale variability, and this, for all coordinates and solutions. Hence, considering spatially variable spectral indices and noise amplitudes is preferable when analyzing global solutions, and especially for vertical land motions. In particular, the assumption of a constant ratio between all variance components in a global network, as used in Amiri-Simkooei (2013) and Amiri-Simkooei, Mohammadloo, and Argus (2017) does not seem realistic in NGL solutions.

The latitude dependence of white noise amplitudes, clearly visible in Figures (10.6), was already detected and discussed by Williams et al. (2004). More recently, it was also observed in IGS station position time series by Klos et al. (2019). Williams et al. (2004) suggested that the equatorial bulge visible in Figure (10.6) could have an atmospheric origin. The NGL-IGS14 solution benefits from improvements in the modelling of both tropospheric and ionospheric effects. With the reprocessing, we observe a reduction of −30% of white noise amplitudes for all coordinates and latitudes. Such a systematic reduction is not observed for power-law amplitudes. Hence, our results tend to support the hypotheses that non-modelled atmospheric effects result in an increase of white noise in position time series (Munekane and Boehm, 2010). However, the improvements in the modelling of atmospheric delays do not seem to reduce the equatorial bulge. Hence, the origin of this bulge remains to be investigated.

Unlike white noise amplitudes, power-law noise amplitudes $\hat{\sigma}'_{pl}$ of the Up coordinates increase more at high latitudes than near the equator. This dependence is even more pronounced in the NGL-IGS14 solution. Moreover, NGL-IGS14 also shows a distinct peak for the Up coordinates in Figure (10.8) around latitude +65°, which corresponds to a localized increase of $\hat{\sigma}'_{pl}$ in Alaska, around the Baltic Sea and in Russia (Figure 10.5). This localized increase in the NGL-IGS14 suggests the influence of non-modelled geophysical phenomena, which became more apparent with modelling improvements. A possible candidate for this latitude pattern is the Non-Tidal Atmospheric and Oceanic Loading (NTAOL), whose predicted RMSE shows a strong latitude dependence in Mémin, Boy, and Santamaria-Gomez (2020). We must mention that a process does not have to show a power-law behaviour to affect power-law process estimates: in principle, any non-modelled processes can bias the noise estimation, even if the energy of the process is localized on a frequency band.

The latitude distribution of the spectral index $\hat{\kappa}$ in Figure (10.7) shows similarities with that of $\hat{\sigma}'_{pl}$. In particular, the Up coordinates shows a pattern that is very different from that of horizontal coordinates, for both solutions, and exhibits longer

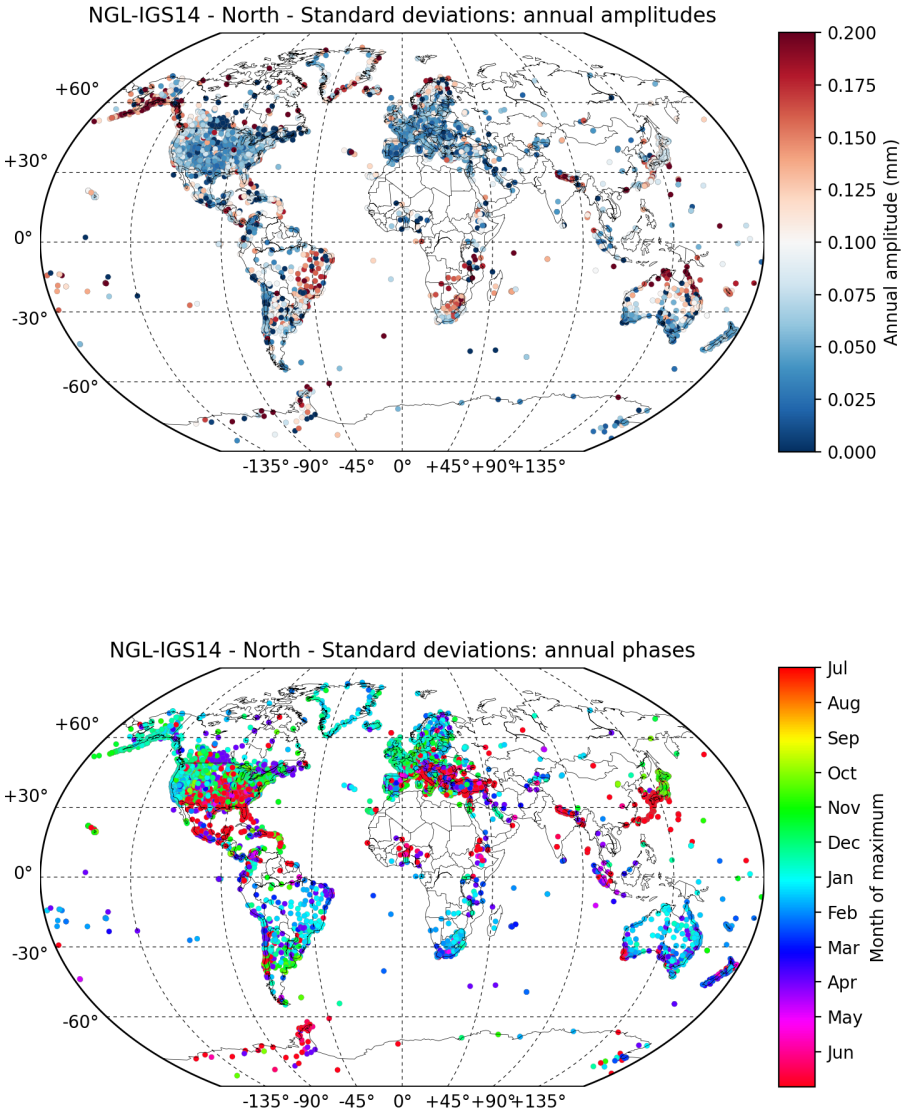

FIGURE 10.13: Spatial variability of the standard deviation oscillation for the North coordinates of NGL-IGS14. Annual amplitudes of monthly standard deviations (*top*). Annual phase of monthly standard deviations. The phase is expressed as the month corresponding to the maximum of dispersion (*bottom*). for time series presenting more than 4 years of cumulative observations

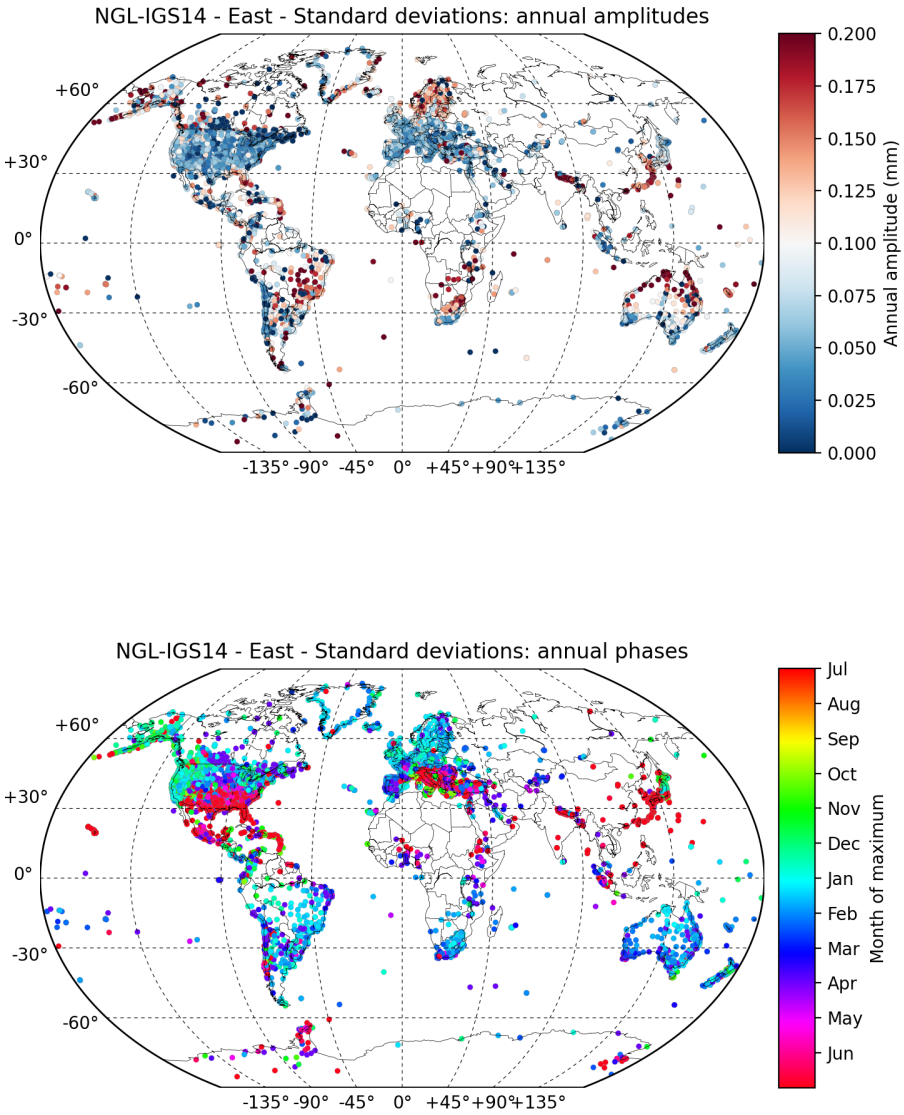

FIGURE 10.14: Spatial variability of the standard deviation oscillation for the East coordinates of NGL-IGS14. Annual amplitudes of monthly standard deviations (*top*). Annual phase of monthly standard deviations. The phase is expressed as the month corresponding to the maximum of dispersion (*bottom*).

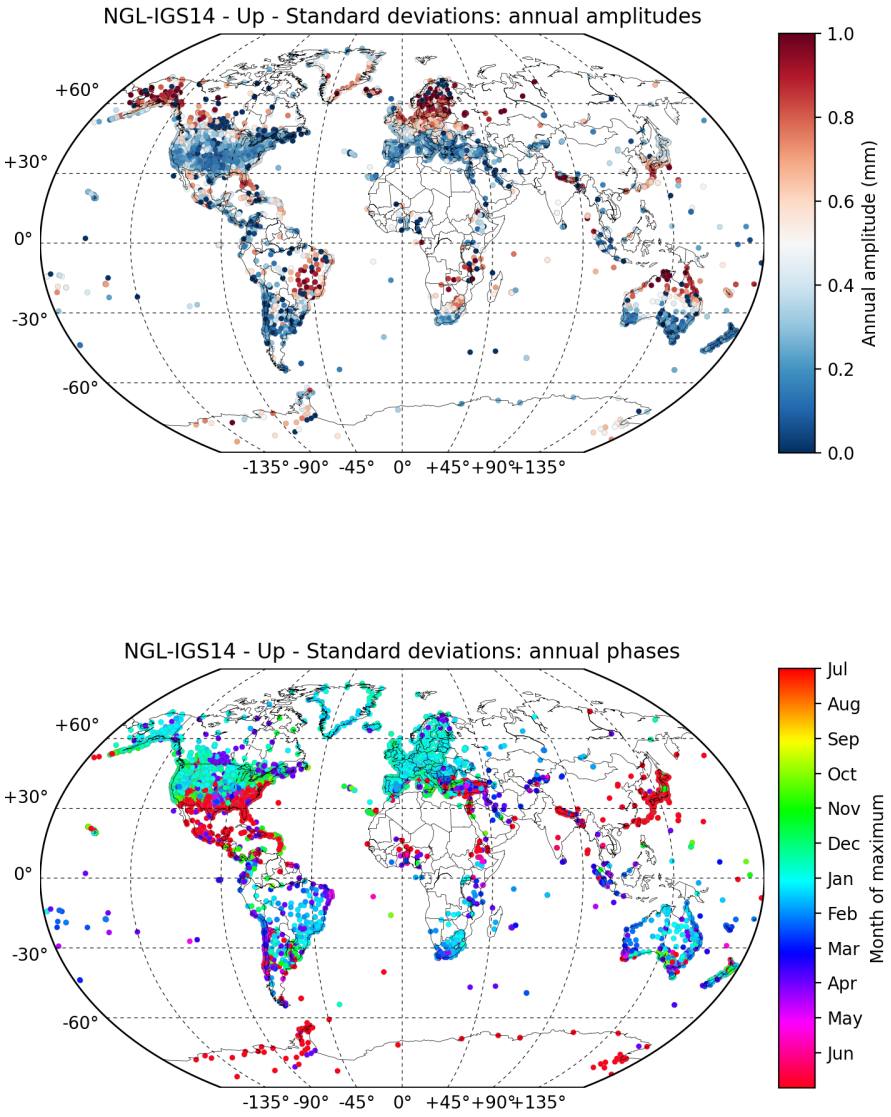

FIGURE 10.15: Spatial variability of the standard deviation oscillation for the Up coordinates of NGL-IGS14. Annual amplitudes of monthly standard deviations (*top*). Annual phase of monthly standard deviations. The phase is expressed as the month corresponding to the maximum of dispersion (*bottom*).

memory at high latitudes. Hence, it is possible that at high latitudes, NTAOL could also influence κ estimates. However, unlike $\hat{\sigma}'_{pl}$, $\hat{\kappa}$ also shows a reduction at the equator. Hence, it is possible that another type of loading effect, namely the hydrological loading, whose RMSE is expected to higher near the equator (Mémin, Boy, and Santamaria-Gomez, 2020) also affect $\hat{\kappa}$.

10.7.2 Time-variable spatial correlation

As for previous studies (Williams et al., 2004; Amiri-Simkooei, Mohammadloo, and Argus, 2017; Benoist et al., 2020), our results evidence a clear spatial correlation of the residual time series. The spatial scale of the spatial correlation in horizontal directions is systemically larger than in the vertical direction (Table 10.1, Figures 10.10 and 10.11). The spatial correlations estimated in this study are slightly smaller than in previous studies, even for NGL-IGS08, undoubtedly due to the very different number of stations analyzed.

Thanks to the considerable number of stations, our correlation analysis also revealed a new characteristic of spatial correlation, namely an annual periodicity of its amplitude (Figure 10.12). For instance, for the Up coordinates of NGL-IGS14, the spatial correlation increases by 73% in January (+0.52) compared to July (+0.30). Since most stations are located in the United States and Europe, this periodicity is probably specific to these areas. This annual variability is also observed in standard deviation residual time series. Still, it shows different phase and amplitudes depending on the position of the station on the globe (Figure 10.15). Such seasonality of the spatial correlation and monthly standard deviations, for the Up coordinates, seems consistent with the possible influence of surface mass loading.

For horizontal coordinates, the seasonality of the standard deviation is visible, but about 5 times smaller than for the Up coordinates in Figures (10.13), (10.14) and (10.15). For the seasonality of the spatial correlation, it is 2 times smaller than for the Up coordinates. In NGL-IGS14, the spatial correlation in the extension-compression direction is similar to that in the Up direction (Figure 10.11). This spatial autocorrelation also seems similar to that induced by atmospheric loading on the vertical direction (Fig. 4 in Petrov and Boy (2004)). This is not the case for the Shear direction. Hence, the spatial and temporal variability in horizontal directions may partly result from atmospheric loading. This, however, requires further investigations.

Although the causes of the spatial correlation and its temporal variability remain unknown, accounting for it in future studies may result in more realistic uncertainties for spatial analyses. In particular, the uncertainty on deformation parameters, such as displacement, strain or rotation parameters has shown to be sensitive to the spatial correlation of observations (Razeghi, Amiri-Simkooei, and Sharifi, 2016).

10.7.3 Influence of the reprocessing

Figures (10.9), (10.6) and (10.8) shows that the recent reprocessing resulted in a reduction of RMSE, white noise amplitudes, and power-law amplitudes. On average, the estimated reduction of RMSE is of 20% for horizontal components and 18% for vertical components. These results are in agreement with the 17% reduction in repeatability reported by Blewitt et al. (2019). Complementarily, our analysis shows that this improvement is not equally distributed between white noise and power-law processes, and neither between the south and north hemisphere.

With an average reduction of about 34%, the white noise component is the most affected by the reprocessing, which we attribute to the improvements in the modelling of atmospheric delays. This supports the hypothesis that tropospheric error appears as white noise in position time series (Munekane and Boehm, 2010).

In addition to this average white noise reduction, the reprocessing has removed a North/South asymmetry for both white noise and power-law process amplitudes of horizontal coordinates (Figures (10.6) and (10.8)). A possible cause of latitude dependence could be an improvement of nominal hydrostatic delays resulting from the change from GMF to VMF1. As the global pressure shows a large North/South asymmetry (Böhm, Heinkelmann, and Schuh, 2007), the shift from the Temperature (GPT) model to the more accurate ECMWF numerical weather model may induce latitude dependent improvements, especially in the southern hemisphere.

The improvements observed on Up coordinates are not similar to that on horizontal coordinates. The reduction of white noise amplitude is only visible near the equator in Figure (10.6) because white noise was undetected at high latitudes in NGL-IGS08. Besides, the reduction in power-law process amplitudes does not show a South/North asymmetry. Instead, power-law process amplitudes rather show global reduction with exceptions localized in Alaska, around the Baltic Sea and in Antarctica. A possible cause for these localized exceptions is atmospheric loading deformations, which could be more visible in NGL-IGS14 likely thanks to the use of the VMF1 (Steigenberger, Boehm, and Tesmer, 2009).

In addition to noise amplitude reductions, the reprocessing also had a visible impact on spatial correlations. For instance, although the noise level of each time series is reduced, the spatial correlation level remained similar for both solutions (Table 10.1). It implies that part of the atmospheric errors in NGL-IGS08 was spatially correlated; otherwise, the spatial correlation would have increased. Another evidence of the impact of reprocessing is the change of spatial correlation on horizontal direction, visible in Figure (10.10) and Figure (10.11), where large scale correlation patterns disappeared. This also results in reduction of the range parameter r in Table (10.1). The amplitude of this spatial pattern seems correlated with the solar cycle 24. It may result from second-order ionospheric effects, non-modelled in NGL-IGS08, which can cause large scale spatial pattern in the residuals of horizontal position time series (Kedar et al., 2003).

Unlike horizontal components, we observe that the Up coordinates show an increase in the range parameter. This could also be due to the influence of atmospheric loading deformations, expected to be more visible in NGL-IGS14 than in NGL-IGS08. However, Table (10.2) shows that the annual amplitude of the spatial correlation reduced by 31% in NGL-IGS14. Hence, part of the correlation periodicity in NGL-IGS08 may be due to atmospheric errors.

10.8 Conclusion

Assessing the spatial and temporal dependence of the stochastic variability in GNSS position time series is necessary for both the uncertainty assessment and for the understanding of the source of stochastic variability in GNSS position time series. By analyzing the time series of over 10,000 globally distributed stations, we demonstrated that all stochastic parameters, namely white noise amplitudes, power-law

amplitudes and power-law spectral indices show global scale variability for North, East and Up coordinates. This global variability, most visible for the Up coordinates, shows a strong latitude dependence. This latitude pattern could result from the influence of surface mass loading deformations or atmospheric errors. In all cases, our study evidence that it is not realistic to assume that position time series have similar stochastic properties in a global solution.

In complement to the time correlation analysis, our study evidence a significant and time-variable spatial correlation of positions time series for both the extension/compression, shear and up directions. Such a time variable correlation is compatible with the influence of surface mass loading. However, as a fraction of the time variability was removed in the NGL reprocessing, atmospheric errors may also be the cause of the time variable spatial correlation.

Finally, our comparative analysis demonstrated the NGL reprocessing resulted in an apparent reduction of the stochastic variability, especially in the Southern Hemisphere. The only observed exception is for the Up coordinates. Among the possible cause for this increase, there is the influence of non-tidal atmospheric loading deformations. We investigate this hypothesis in chapter 11.

Chapter 11

Influence of non-tidal atmospheric and oceanic loading on the stochastic properties of vertical land motion time series

In the previous chapter, our stochastic analyses revealed that Up coordinates show a spatial variability standing out from that of horizontal coordinates for both NGL solutions. This spatial variability is even more pronounced in the most recent solution, suggesting the possible influence of non-modelled geophysical processes. In this chapter, we investigate to which extent surface mass loading deformations, known to be more significant on the Up coordinates, can explain such a specific variability.

11.1 Introduction

Trends estimated from vertical land motion time series obtained using global navigation satellite systems (GNSS) are central to the understanding of geophysical processes such as glacial isostatic adjustment (GIA) (Nocquet, Calais, and Parsons, 2005; Schumacher et al., 2018) and coastal vertical land motions (Pfeffer and Allemand, 2016; Wöppelmann and Marcos, 2016). One of the main factor limiting the precision of these trend estimates is that VLMs present a stochastic variability showing not only white noise but also power-law properties.

In chapter 10, we evidenced a spatial dependence of both white noise amplitudes, power-law amplitudes, and spectral indices that is specific to the Up coordinates. This spatial dependence is especially visible in the latest Nevada Geodetic Laboratory reprocessing, suggesting the possible influence of non-modelled physical processes. At the global scale, this variability is mostly latitude dependent. White noise amplitudes are maximum near the equator and null outside the inter-tropical band. Power-law process amplitudes show a minimum at the equator and maximum near the poles, with hot-spots around the Baltic Sea, Alaska, Canada, Russia and Antarctica. Finally, spectral indices are systematically smaller near the equator and near the poles than at mid-latitudes, identifying processes with longer memory in these areas.

In this study, we investigate whether the spatial variability observed in chapter 10 is caused by non-modelled vertical deformation induced by surface mass loading. To do so, we apply the exact same temporal and spatial noise estimation strategy than in chapter 10, but to Up coordinates time series for which we subtract the

non-tidal loading deformation computed by the Earth System Modelling group of GeoForschungsZentrum Potsdam (ESMGFZ) (Dill and Dobslaw, 2013).

11.2 Computing loading deformations

The global surface mass distribution induced by atmospheric, oceanic and hydrologic phenomena results in spatial and temporal variations of pressure at the surface of the Earth, which we refer to as load and denote as $\mathcal{L}(t, \lambda, \phi)$ and express in kg/m^2. Information about the spatial and temporal variability of the load is usually obtained from space missions, such as Gravity Recovery And Climate Experiment (GRACE) or from physical numerical models.

In this study, we use ground deformation predictions based on numerical models and not satellite observations because it is presently the only way to obtain both daily and globally distributed estimates of the load $\mathcal{L}(t, \lambda, \phi)$.

When dealing with globally distributed load estimates, $\mathcal{L}(t, \lambda, \phi)$ can be expressed in the spherical harmonic domain as

$$\mathcal{L}(t, \lambda, \phi) = \sum_{l=0}^{\infty} \sum_{m=0}^{l} \mathcal{L}^C_{lm}(t) Y^C_{lm}(\lambda, \phi) + \mathcal{L}^S_{lm}(t) Y^S_{lm}(\lambda, \phi), \tag{11.1}$$

in which $\mathcal{L}^C_{lm}(t)$ and $\mathcal{L}^S_{lm}(t)$ are the Stokes coefficients of the load and Y^C_{lm} and $Y^S_{lm}(\lambda, \phi)$ are the spherical harmonic functions of degree l and order m, defined as

$$\begin{bmatrix} Y^C_{lm}(\lambda, \phi) \\ Y^S_{lm}(\lambda, \phi) \end{bmatrix} = P^m_l(\sin(\phi)) \begin{bmatrix} \cos(m\lambda) \\ \sin(m\lambda) \end{bmatrix}, \tag{11.2}$$

where $P^m_l(x)$ is the Legendre associated function defined by

$$P^m_l(x) = \frac{(-1)^l}{2^l l!} \sqrt{(1 - x^2)^m} \frac{d^{l+m}}{dx^{l+m}} (x^2 - 1)^l. \tag{11.3}$$

According to Farrell (1972), the vertical displacements $\delta_{\mathsf{Up}}(t, \lambda, \phi)$ resulting from the effect of surface mass loading can be expressed, under the approximation of a spherical, isotropic, non-rotating, and elastic Earth model, as a function of the spherical harmonic coefficients of the load, $\mathcal{L}^C_{lm}(t)$ and $\mathcal{L}^S_{lm}(t)$, such that

$$\delta_{\mathsf{Up}}(t, \lambda, \phi) = \frac{1}{\rho_E} \sum_{l=0}^{\infty} \sum_{m=0}^{l} \frac{3h_l}{2l + 1} \left[\mathcal{L}^C_{lm}(t) Y^C_{lm}(\lambda, \phi) + \mathcal{L}^S_{lm}(t) Y^S_{lm}(\lambda, \phi) \right] \tag{11.4}$$

in which ρ_E denotes the mean density of the earth model in kg/m^2, and h_l denotes the radial Love number of degree l.

The radial load Love numbers h_l depend on the rheological properties of the considered Earth model. For $l = 1$, the load Love number h_1 also depends on the considered frame.

11.2.1 Loading deformation predicted by the EMSGFZ-Postdam

In this study, we consider 3 types of loading effects, namely non-tidal atmospheric loading (NTAL), non-tidal oceanic loading (NTOL) and hydrological loading (HYDL).

Vertical deformations caused by NTAL are calculated by ESMGFZ using 3-hourly atmospheric surface pressure from the European Center for Mean Weather Forecast (ECMWF). Deformations caused by NTOL are calculated using 3-hourly ocean bottom pressure from the Max-Planck-Institute Meteorology Ocean Model (MPIOM). The deformations caused by HYDL deformation is calculated using 24-hourly terrestrial water storage from the Land Surface Discharge Model (LSDM). To obtain vertical displacements from each type of load, the load Love numbers h_l used by EMSGFZ are computed from the elastic Earth model ak135 (Kennett, Engdahl, and Buland, 1995; Wang et al., 2012) and expressed in the centre of Earth's figure frame (CF).

Given the sampling rate of ECMWF and MPIOM models, NTAL and NTOL deformation time series are estimated every 3 hours. To obtain daily time series, we linearly interpolated the NTAL and NTOL predictions at the epochs of the GNSS VLM time series. To illustrate the expected variability from each type of load, the standard deviation of the deformation predictions are presented, for all time series in Figure (11.1).

11.3 Preliminary change in repeatability analysis

Performing stochastic parameter estimation for over 10,000 daily position time series, is computationally expensive. Hence, before doing so, we preferred finding the right candidate for the latitude pattern observed in chapter 10 using a more straightforward method.

Before variance component estimation, we estimate the change in repeatability of the residuals after subtracting the NTAL, NTOL, and HYDL models from observations. As a measure of repeatability, we use the percentage of Root Mean Square Error (RMSE) reduction. For each type of load, we compute the RMSE assuming white noise, that is through an unweighted least-squares adjustment.

The RMSE reduction percentages obtained after removing either NTAL, NTOL, or HYDL deformations are presented in Figure (11.2). NTAL corrections result in an RMSE reduction for 97.5% of the stations, more effective near the poles than near the equator. NTOL corrections result in an RMSE reduction for 87.6% of the stations. Just like NTAL, the impact of NTOL is latitude dependent and most effective near the poles. HYDL corrections result in RMSE reduction for 67.5% of the stations. This reduction is not maximum near the pole, but near the equator and in the USA. Also, HYDL corrections result in a slight RMSE increase in Eastern Europe.

As RMSE reductions resulting from HYDL corrections are not as systematic as for NTAL and NTOL corrections and not maximum near the poles, we preferred not to include it for stochastic parameter estimation. Hence, in the following, we focus on the impact of the non-tidal atmospheric and oceanic loading (NTAOL) on the spatial and temporal properties of the stochastic variability in VLM time series.

11.4 Influence of NTAOL on noise properties

We estimated the stochastic properties of the VLM time series after removing the NTAOL time series computed by the ESMGFZ. The observations, the functional

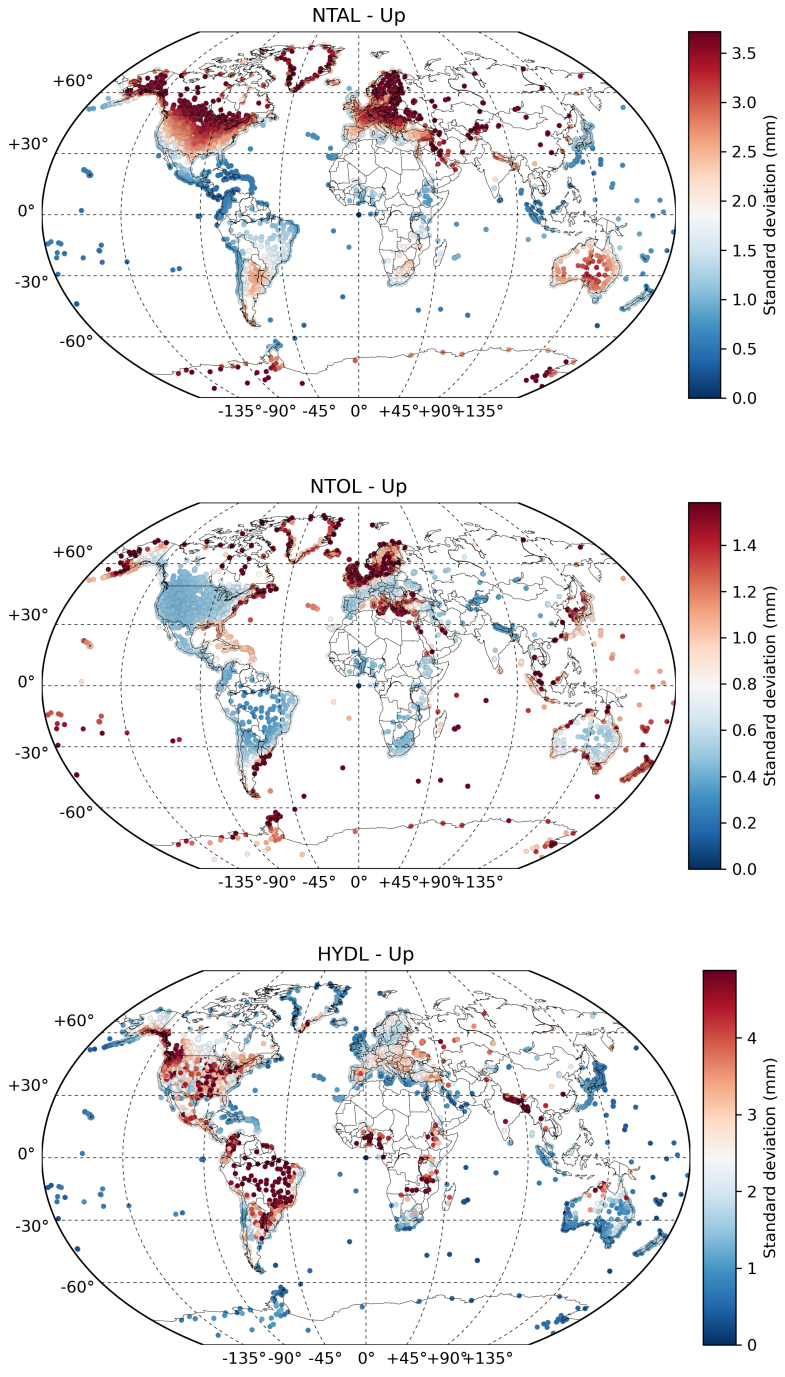

FIGURE 11.1: Standard deviation of the predicted vertical land motions resulting from surface mass loading. Non-Tidal Atmospheric Loading (NTAL) (*top*). Non-Tidal Oceanic Loading (NTOL) (*middle*). Hydrological loading (HYDL) (*bottom*).

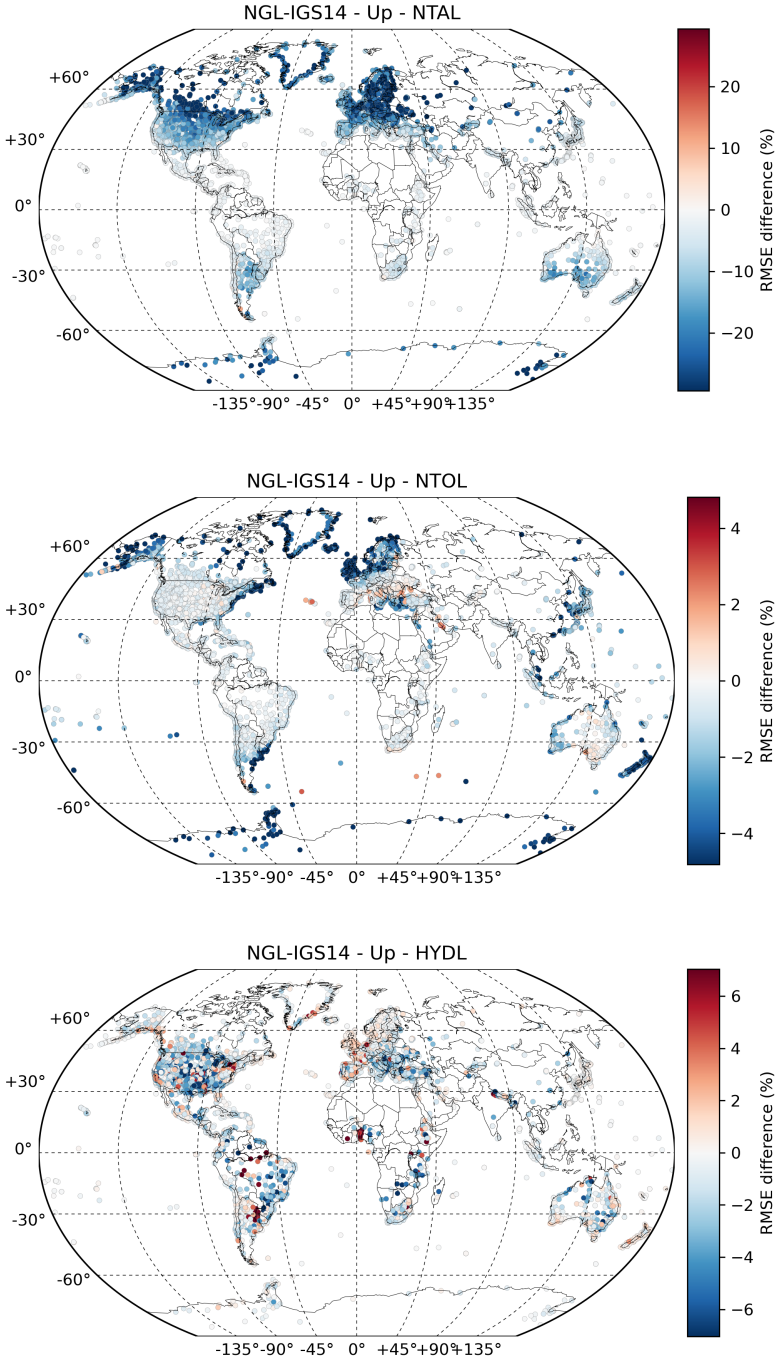

FIGURE 11.2: Global distribution of RMSE reduction percentages after removing deformation predictions. Non-tidal atmospheric loading (NTAL) (*top*). Non-tidal oceanic loading (NTOL) (*middle*). Hydrological loading (HYDL) (*bottom*).

models, the stochastic models, and the statistical methods used for temporal and spatial correlation estimation are identical to those used in chapter 10.

11.4.1 Impact on temporal correlation

Figure (11.3) presents the global distribution of variance components estimates for VLM time series with NTAOL deformations removed. Estimates presented in Figure (11.3) significantly differs from that observed in chapter 10 (Figure 10.5) for all stochastic parameters. Instead of being null everywhere but near the equator, the white noise amplitudes are now positive everywhere and reveal an equatorial bulge more similar to that of horizontal coordinates. The spectral index estimates also show a reduction in spatial variability, with, on average, longer memory than in Figure 10.5. The most significant changes are visible for the power-law amplitudes, whose latitude dependence have completely disappeared and show a more homogeneous distribution. In particular, the power-law amplitudes observed around the Baltic Sea, Alaska, Canada, Russia and Antarctica are now comparable to those of the rest of the world.

To better illustrate the changes as a function of the latitude, Figure (11.4) depicts the latitude dependence of RMSE, white noise amplitude, spectral index, and power-law amplitude estimates before and after NTAOL corrections. The running medians before NTAOL corrections are reported on that after corrections to ease the comparison. In Figure (11.4), the RMSE (here estimated using a weighted least-square adjustment) shows a median reduction reaching 30% at high latitudes, and especially near $+60°$. This overall RMSE reduction is linked to a clear change in the nature of the stochastic variability. Out of the inter-tropical band, white noise amplitudes now show a floor of about 2.0 mm. The equatorial bulge visible in white noise amplitudes now appears as a 2 mm increase at the equator. For the spectral index, the bi-modal latitude distribution visible before NTAOL corrections disappears and become mostly centred on -1 with a decrease to -1.2 near the equator. Spectral indices also show a clear increase in dispersion between $[-60°, -30°]$ and $[+30°, +60°]$. Based on the map in Figure (11.3), this increase suggests that the latitude representation for this parameter is not appropriate anymore. Finally, Figure (11.4) confirms the observation made with Figure (11.3), namely that removing NTAOL deformations completely removes the high latitude dependence of power-law amplitudes. More precisely, power-law amplitudes now show only slight deviations from a 3.5 mm average.

Note that, although we observe a decrease in RMSE at high latitudes, the change in stochastic parameters in Figure (11.4) does not indicate whether considering white noise and power-law noise model is still relevant when accounting for NTAOL deformations. This can however be tested by analysing the change in restricted log-likelihood. Figure (11.5) shows the global distribution of restricted log-likelihood changes after removing NTAOL deformations. The restricted log-likelihood increased for 97.9% of the stations. Hence, the observations are now more consistent with a combination of white noise and power-law noise than before removing NTAOL deformations.

Since the stochastic model directly influences deterministic parameter uncertainties, we evaluated how trend uncertainties are affected by the change in stochastic parameters. Because all stations do not have the same number of observations nor the same functional model, we simulated the influence of the estimated stochastic

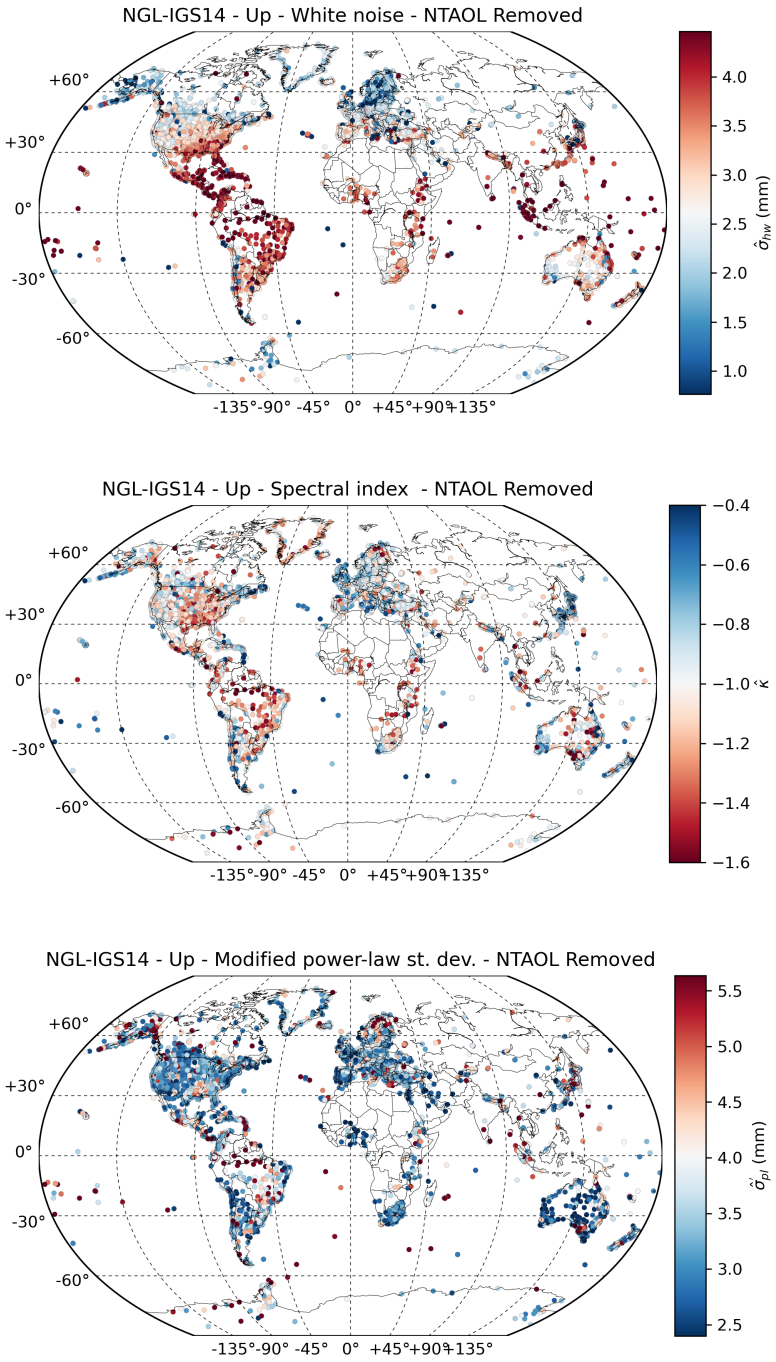

FIGURE 11.3: Global variability of white noise amplitudes (top), spectral indices (middle); and power-law amplitudes (bottom) for the Up coordinates with NTAOL deformations removed.

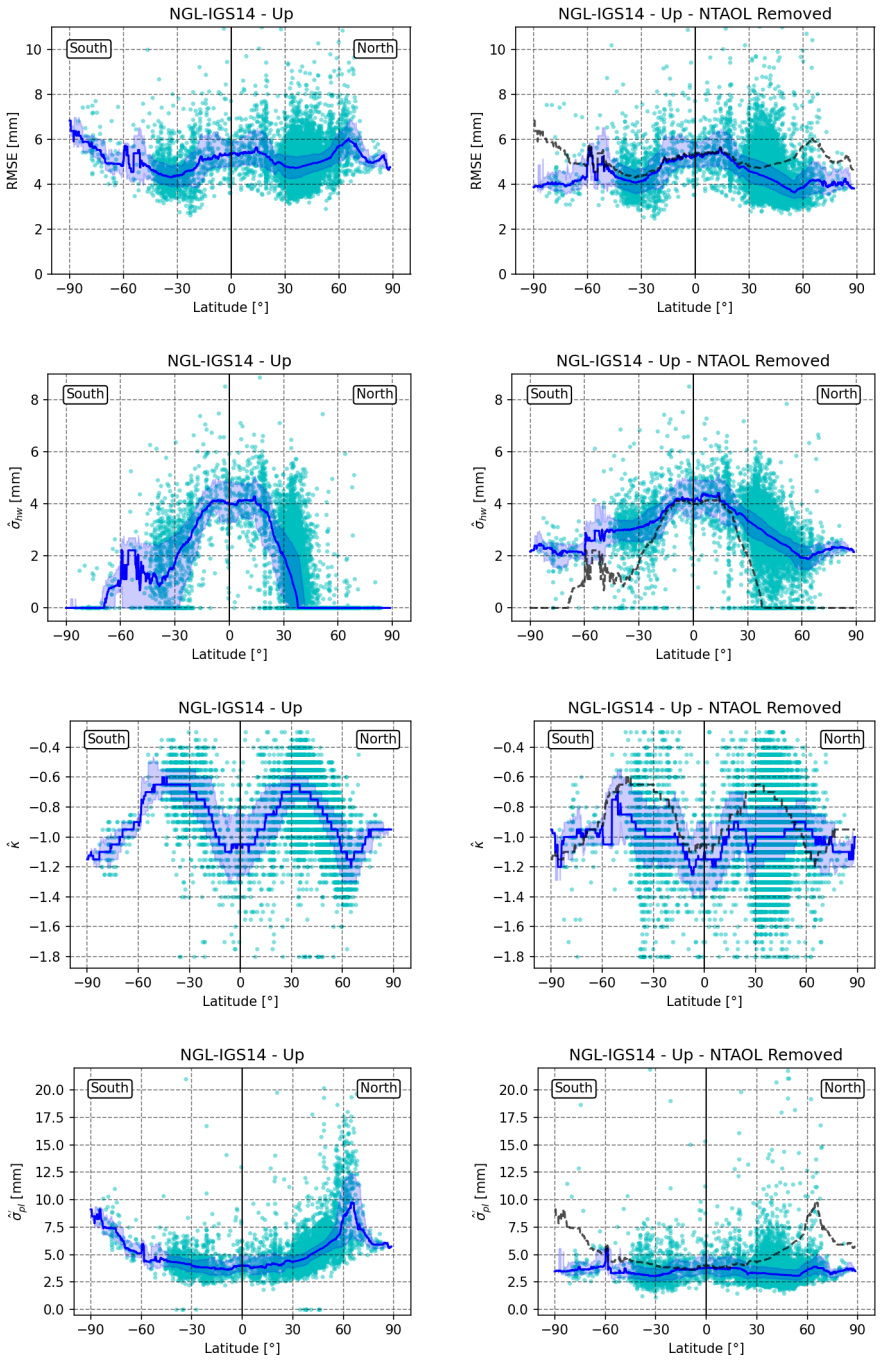

FIGURE 11.4: Latitude dependence of RMSE (*top*), white noise amplitudes (*upper middle*), spectral indices (*lower middle*), power-law amplitudes (*bottom*) in the presence of NTAOL deformations (*left*) and with NTAOL deformations removed (*right*). The dots in cyan represent the raw estimates. The continuous blue lines represent running medians over centered 10° latitude windows. The shaded areas represent the corresponding interquartile ranges. The dashes black lines are running medians before NTAOL corrections reported on that after corrections.

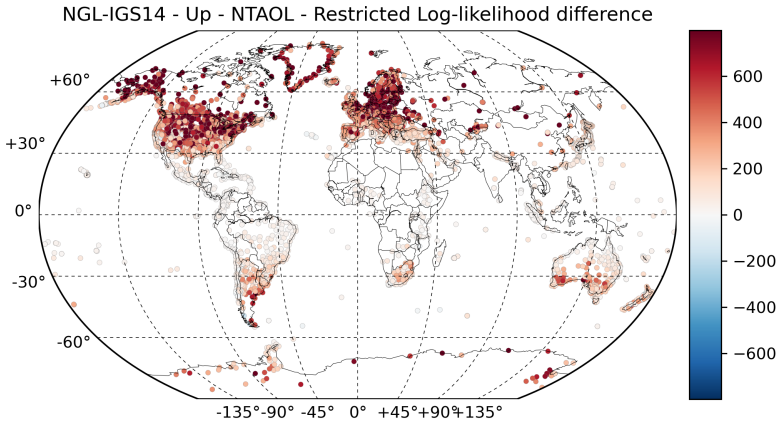

FIGURE 11.5: Global variability of the restricted log-likelihood differences.

Dataset	$\bar{\rho}_c$	$\bar{\rho}_0$	$\bar{\rho}_{\max}$	r [km]
With NTAOL	0.01	0.40	0.41	2470
Without NTAOL	0.02	0.27	0.29	2485

TABLE 11.1: Parameter estimates of the Gaussian model adjustment on the mean correlograms.

models on a fictional time series with continuous observations for 8 years, and no deterministic effect other than a linear trend. The estimated trend uncertainties with and without NTAOL corrections are presented in Figure (11.6).

Figure (11.6) demonstrates that NTAOL corrections also results in a dramatic reduction – that is, up to -75% – of the trend uncertainties in Alaska, in Canada, around the Baltic Sea, in Russia and Antarctica.

11.4.2 Impact on spatial correlation

Figure (11.7) presents estimates of both the time variable correlograms and the average correlation, with and without NTAOL corrections. In Figure (11.7), NTAOL corrections results in a reduction of both the average spatial correlation and the time variable correlation for separation distances $d_l < 4000$ km. The change of time correlation is also visible in Figure (11.8), which presents the spatial correlation time series for 4 separation distances, namely $d_l = 50$ km, $d_l = 950$ km, $d_l = 1950$ km and $d_l = 3950$ km.

As in chapter (10), we approximated the average spatial correlation by a Gaussian correlation function (Equation (10.27)). The adjusted parameters of the Gaussian correlation function with and without NTAOL corrections are presented in Table (11.1). Table (11.1) shows that the maximum of correlation for $d_l > 0$ km, noted $\rho_{\max}$, changed from 0.41 to 0.29, that is, a reduction of -29.3% caused by NTAOL corrections. The range parameter of the Gaussian correlation is unaffected by the NTAOL corrections.

Pursuing the application of the method of chapter (10), the time variable spatial correlation has been approximated by an annually modulated Gaussian correlation

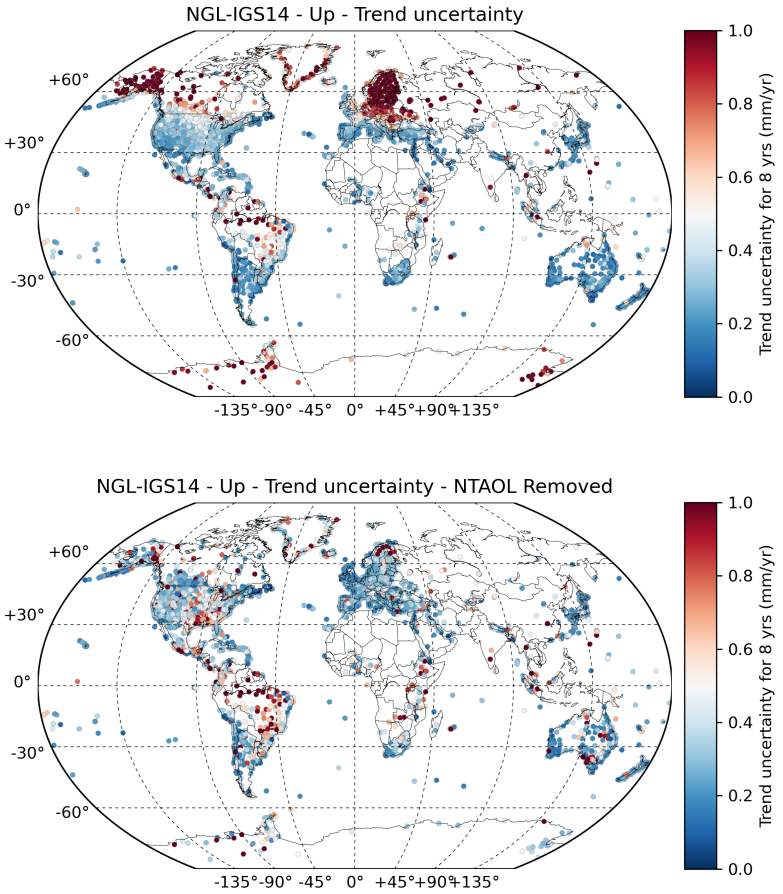

FIGURE 11.6: Global variability of the trend uncertainties for a 8 years long time series. Without NTAOL corrections (*top*). With NTAOL corrections (*bottom*).

Dataset	$\hat{\rho}_c$	$\hat{\rho}_0$	$\hat{\rho}_A$	$\hat{\phi}_A$ [rad]	$\hat{\rho}_{\text{min}}$	$\hat{\rho}_{\text{max}}$	$\hat{r}$ [km]
With NTAOL	0.01	0.40	0.11	1.333	0.30	0.52	2469
Without NTAOL	0.02	0.27	0.05	1.19	0.24	0.34	2485

TABLE 11.2: Parameter estimates of the time variable Gaussian model adjustment. The $\hat{r}$ are reported values from Table (11.1).

function (Equation (10.28)). The adjusted parameters of the time-variable correlation model, with and without NTAOL corrections, are presented in Table (11.2). Table (11.1) shows that in addition to reducing the average spatial correlation, NTAOL corrections also results in a reduction of annual amplitude, which changed from 0.11 to 0.05, that is, a -54.5% reduction. The estimated phases are similar in both data sets (the maximum of correlation is still in around January). In this study, annual correlation estimates are also likely biased by the oversampling of the USA and Europe. To get some insight into the possible spatial variability of the remaining spatial correlation in the rest of the world, we estimated the spatial variability of the annual oscillation of monthly standard deviations of the residuals. The global distribution of annual amplitude and phase estimates are presented in Figure (11.9). The effect of NTAOL corrections appears in Figure (11.9) as a clear reduction of the amplitude standard deviation periodicity in Alaska, Canada, around the Baltic Sea, Russia and Antarctica. Areas now showing the largest annual amplitudes seem to be that affected by hydrological loading (Figure (11.2)).

11.5 Discussion

NTAOL deformations strongly influence the estimated temporal correlation of VLM time series. Correcting for NTAOL deformations results in a significant improvement in the stochastic modelling for all time series, as shown by the change in restricted log-likelihood in Figure (11.5). This means that VLM observations corrected from NTAOL are more consistent with a stochastic model combining white noise and power-law process. In other words, NTAOL deformations introduce a stochastic temporal variability in VLM times that cannot be modelled by a combination of white noise and power-law stochastic model process. The annual oscillation of the variance observed in Figure (11.9) is an example of such variability. Hence, without NTAOL correction, the considered stochastic model is under-parametrized, which can ultimately lead to biased parameter estimates. This biased estimation issue explains why white noise was systematically absent at some latitudes and why power-law noise amplitudes were so large in chapter 10, but also in Klos et al. (2019). This dramatic change in stochastic parameters estimates has a direct impact on trend uncertainties. In particular, NTAOL corrections results in a -75% reduction in trend uncertainties in Alaska, Canada, around the Baltic Sea, Russia and Antarctica (Figure (11.6)). Hence, applying NTAOL corrections could lead to clear improvements in trend uncertainty estimates in these areas.

The spatial correlation analysis also shows that NTAOL strongly influences the spatial correlation between time series. In particular, NTAOL corrections reduced the average spatial correlation by -29.3% and its annual amplitude by -54.4%. Hence, applying NTAOL corrections could benefit not only to trend estimation but also to spatial deformation estimation.

Although smaller than before NTAOL corrections, a time variable spatial correlation is still clearly visible in Figure (11.8). Similarly, the stochastic parameters still exhibit spatial variability in Figure (11.3). Further investigations are needed to test whether hydrological loading can explain the remaining variability in VLM time series.

11.6 Conclusion

The stochastic modelling of vertical land motion time series is a crucial step to obtain both realistic trend estimates and uncertainties, but also for spatial analyses. In a previous study, the global distribution of stochastic parameters showed a latitude dependence specific to VLM time series.

This chapter investigated the impact of non-tidal atmospheric and oceanic loading (NTAOL) on stochastic parameters and their spatial variability. We showed that a large part of the latitude dependence previously observed is caused by NTAOL deformations, which introduces a non-modelled variability that biases the stochastic parameter estimation. Our results show that applying NTAOL corrections improves the agreement between the observations and the white noise + power-law stochastic model and reduces the uncertainty of trend estimates at high latitudes.

Besides the time correlation analysis, we also showed that NTAOL deformations also explains part of the spatial correlation and its time variability. Hence, applying NTAOL corrections could also be beneficial to spatial analyses.

After NTAOL corrections, both the spatial distribution of stochastic parameters and the spatial correlation still show some residual variability. Hence, it is possible that hydrological loading and atmospheric errors could explain part of the remaining variability, especially near the equator.

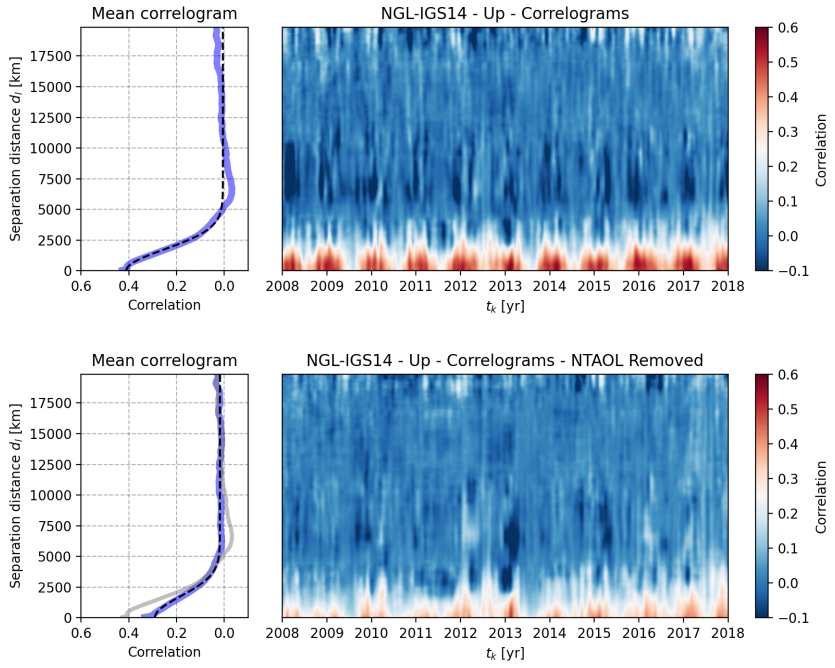

FIGURE 11.7: Spatial correlation estimates. Without NTAOL corrections (*top*). With NTAOL corrections (*bottom*). The blue curve represents the mean correlogram. The dashed black lines represent the Gaussian correlation model.

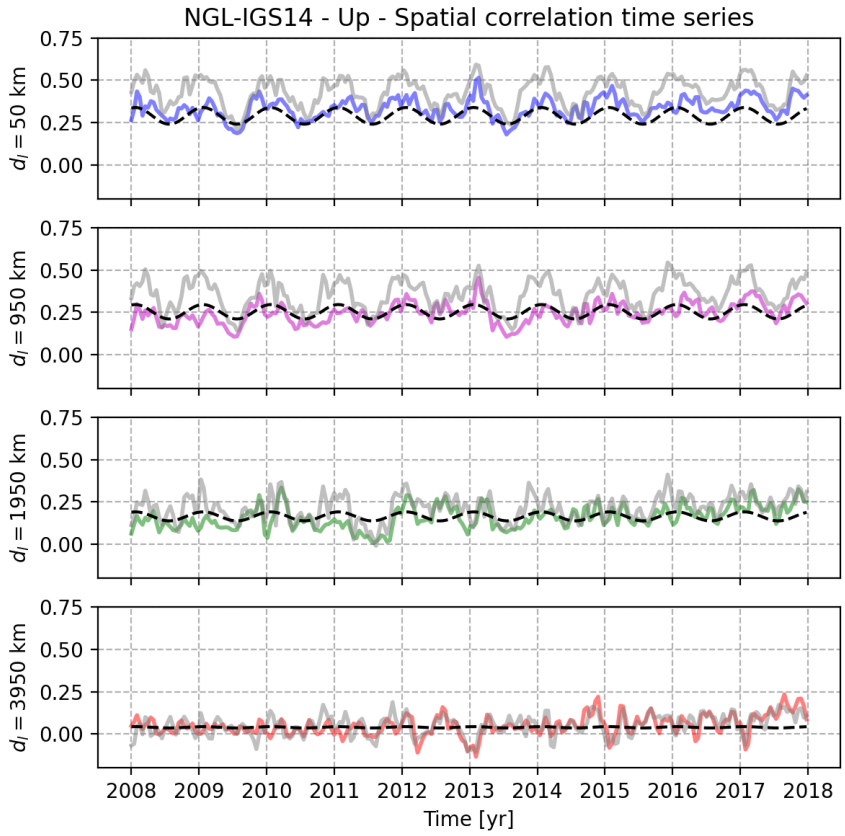

FIGURE 11.8: Time variable spatial correlation estimates for 4 classes of distances with NTAOL corrections. The colored lines represent the correlation estimates with NTAOL corrections. The light grey lines represent the correlation estimates without NTAOL corrections. The dashed black lines represent the time variable Gaussian correlation model.

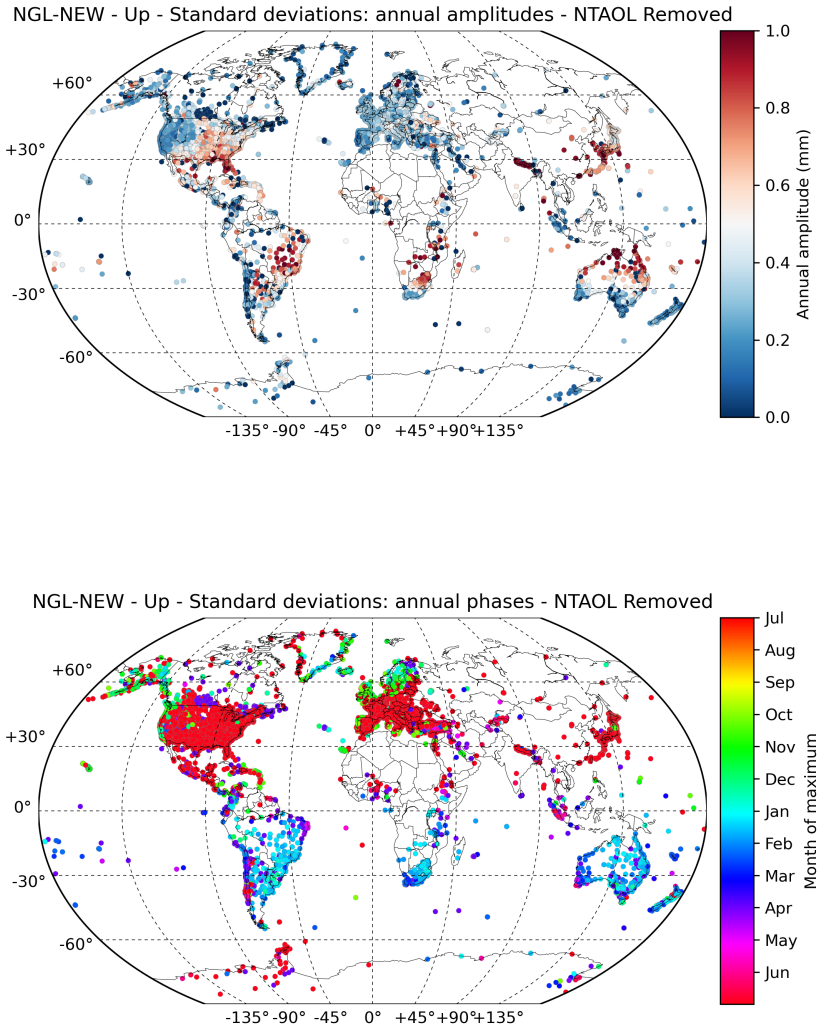

FIGURE 11.9: Spatial variability of the standard deviation oscillation with NTAOL corrections. Annual amplitudes of monthly standard deviations (*top*). Annual phase of monthly standard deviations. The phase is expressed as the month corresponding to the maximum of dispersion (*bottom*).

Chapter 12

Conclusions & perspectives

Throughout this thesis, we aimed at improving the knowledge about the uncertainty and accuracy of both sea-level observations provided by tide gauges and the vertical land motions estimated using Global Navigation Satellite Systems (GNSS). In this final chapter, we summarize how this work contributes to this objective, and we draw, for each considered quantity, new research perspectives.

12.1 Tide gauge observations

12.1.1 Contributions

Our work on tide gauge observations, presented in chapters 6 and 7, mostly focused on the calibration process, essential to provide sea-level time series stable in the long term.

In chapter 6, we started by providing a detailed description of the link between the sea-level differences between two collocated tide-gauges and the measurement biases, that is, the sensor offset and the scale error of the tested gauge. By providing an expression of sensor offset as a function of the sea-level bias parameters, we also showed that it is possible to determine the minimum calibration experiment duration necessary to reach a desired level of precision. Based on a simple simulation, we showed that to get a few millimetre precision on the sensor offset requires a few days of observation, especially in locations with small tidal ranges.

Then, in chapter 7, we demonstrated that, in the presence of at least three collocated tide-gauges, it is possible to obtain a more precise and a complete assessment of the performance of each gauge, by combining all available time series instead of comparing them by pairs. Using a Variance Component Estimation method, we could estimate the standard deviation of each gauge and demonstrate that, in the absence of biases, the uncertainty of each gauge is close to the 1.0 cm requirement suggested by the Intergovernmental Oceanographic Commission (IOC).

12.1.2 Perspectives

The results presented in chapter 7 are based on a calibration experiment carried out over a few hours and under calm conditions only. Performing calibration experiments over more extended periods would certainly allow deriving more realistic uncertainty estimates, but also enable investigating the possible dependence of tide-gauge uncertainty with environmental conditions.

One by-product of the combination method presented in chapter 7 is the estimation of a combined solution. This combined solution has several attractive features, namely

- it is corrected from biases,
- it automatically fills the observation gaps of some sensors,
- it is more precise than the individual observations,
- its variance accounts for the bias model used and on the number of sensors available.

Hence, in locations where two or more collocated permanent sensors are installed, it is possible to provide more precise and complete sea-level time series with realistic uncertainties. Thus, investigating the applicability of the combination method in an observatory with permanent collocated tide gauges is an interesting perspective to obtain more precise sea-level time series. Besides, given that the combination method in chapter 7 is expressed as a linear model of observation equations, one could, in addition, implement automatic discontinuity detectors using hypothesis testing method presented in chapter 5. In the presence of at least three gauges, these detectors would automatically identify the problematic gauge, which could be convenient for the operational maintenance of the tide gauges.

12.2 GNSS-based vertical land motion estimates

12.2.1 Contributions

Our work on GNSS position time series analysis, presented in chapters 8, 9, 10 and 11, first focused on the study of the statistical methods dedicated to the determination of the stochastic model of GNSS time series, and then on the description of the global variability of the stochastic properties of GNSS time series.

So far, studies devoted to the assessment of the time-correlated properties of GNSS position time series using the Least-Squares Variance Component Estimation (LS-VCE) method assumed linear stochastic models. In particular, when considering power-law processes, the spectral index was supposed to be known. To overcome this limitation, in chapter 8, we demonstrated that the nonlinear LS-VCE method could also estimate the spectral index of power-law processes along with the other parameters. Because the LS-VCE method provides unbiased and minimum variance estimates, the method could be an interesting alternative to the more widely used Maximum Likelihood Estimation method, which is only asymptotically unbiased. Besides, using the nonlinear LS-VCE method, one immediately obtains the uncertainty of all stochastic parameters, including that of the spectral index. We also tested the application of the nonlinear LS-VCE method to the determination of the variance component of generalized Gauss-Markov processes. In our study, the nonlinear LS-VCE method failed to converge in most of the cases. Hence, further work on these models is required to improve this result.

In the introduction, we mentioned that offsets, even when modelled, are a nuisance when estimating functional parameters, and especially for the estimation of trends. In chapter 9, we showed that the presence of offsets was also a nuisance for the determination of low-frequency stochastic parameters. In particular, the presence of offsets systematically biased the results expected from both the Lomb-Scargle

periodogram and the MLE method, which incorrectly promoted generalized Gauss-Markov models over power-law models. Although such biases can be avoided using unbiased methods, such as the Least-Squares Harmonic Estimation periodogram or the restricted MLE (RMLE), the presence of offsets systematically increased the uncertainty of low-frequency variance component estimates. Hence, we showed that, in the future, offsets may restrain the identification of the low-frequency properties of GNSS position time series.

In complement to the study of estimation methods, in chapter 10, we investigated the spatio-temporal properties of the stochastic variability in the position time series of over ten thousand globally distributed GNSS stations produced by the Nevada Geodetic Laboratory. By doing so, we were able to show that the stochastic properties of the north, east, and vertical coordinates all presented a spatial variability at the global-scale. In addition, by performing spatial correlation analyses over monthly windows, we demonstrated that the stochastic variability of all coordinates was spatially correlated, but also that this correlation presented a seasonal signature. The coordinate that showed the most pronounced spatio-temporal variability was the vertical one, which displayed a latitude dependence and hot-spots of time-correlation near polar regions. Through the analysis of two different NGL solutions, we evidenced that this latitude dependence was even more apparent in the solution using improved orbits and clocks products and GNSS modelling, which suggested the influence of non-modelled geophysical processes rather that positioning errors.

Finally, based on the loading deformation time series computed by the GeoForschungsZentrum (GFZ), in chapter 11, we investigated whether the spatio-temporal variability of the vertical coordinate could result from the influence of surface mass loading. To this end, we applied the same methodology as in chapter 10, but on VLM observations from which the non-tidal atmospheric and oceanic loading (NTAOL) deformations predicted by the GFZ were removed. Our results showed that the latitude dependence of the time-correlated properties of this coordinate was fully explained by NTAOL effects. Besides, NTAOL deformations also explained part of the average spatial correlation and its seasonality.

12.2.2 Perspectives

The chapter 8 was a proof of concept for the application of the nonlinear LS-VCE method to the variance component estimation of power-law stochastic models. A natural follow-up to this study would be to test the applicability of this method to real GNSS time series and to find ways to obtain reliable estimates for generalized Gauss-Markov processes. For the latter, change the formulation of the problem, or investigating fast and robust ways to get close approximates values could lead to a better convergence rate of the method.

In chapter 10, our results evidenced a spatial correlation with a seasonality. Investigating how this spatial correlation affects the covariance of trend estimates between stations could be helpful to future scientific applications using spatial vector fields as inputs. This could theoretically be done by simultaneously estimating all functional parameters from all observations of a given solution. However, in practice, this strategy would be far too computationally expensive with a realistic spatio-temporal covariance model. Therefore, other methods, for instance based on Monte-Carlo simulations of temporally correlated spatial-fields, could be considered.

Finally, the study in chapter 11 was restricted to the influence of NTOAL deformations. Now the influence of NTOAL deformations on stochastic parameters is understood, studying the influence of hydrological loading could be useful in understanding the remaining variability in the spectral indices and in white noise amplitudes.

Bibliography

Abbondanza, C et al. (2015). "Three-Corner Hat for the assessment of the uncertainty of non-linear residuals of space-geodetic time series in the context of terrestrial reference frame analysis". In: *Journal of Geodesy* 89.4, pp. 313–329.

Agnew, Duncan Carr (1992). "The time-domain behavior of power-law noises". In: *Geophysical research letters* 19.4, pp. 333–336.

Akaike, Hirotugu (1974). "A new look at the statistical model identification". In: *IEEE transactions on automatic control* 19.6, pp. 716–723.

Allen, Myles R and Leonard A Smith (1996). "Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise". In: *Journal of climate* 9.12, pp. 3373–3404.

Altamimi, Zuheir et al. (2016). "ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions". In: *Journal of Geophysical Research: Solid Earth* 121.8, pp. 6109–6131.

Amiri-Simkooei, AliReza (2007). "Least-squares variance component estimation: theory and GPS applications". PhD thesis. TU Delft, Delft University of Technology.

Amiri-Simkooei, Alireza (2020). "Least Squares Contribution to Geodetic Time Series Analysis". In: *Geodetic Time Series Analysis in Earth Sciences*. Springer, pp. 185–209.

Amiri-Simkooei, AR (2009). "Noise in multivariate GPS position time-series". In: *Journal of Geodesy* 83.2, pp. 175–187.

— (2013). "On the nature of GPS draconitic year periodic pattern in multivariate position time series". In: *Journal of Geophysical Research: Solid Earth* 118.5, pp. 2500–2511.

— (2016). "Non-negative least-squares variance component estimation with application to GPS time series". In: *Journal of Geodesy* 90.5, pp. 451–466.

Amiri-Simkooei, AR and J Asgari (2012). "Harmonic analysis of total electron contents time series: methodology and results". In: *GPS solutions* 16.1, pp. 77–88.

Amiri-Simkooei, AR, TH Mohammadloo, and DF Argus (2017). "Multivariate analysis of GPS position time series of JPL second reprocessing campaign". In: *Journal of Geodesy* 91.6, pp. 685–704.

Amiri-Simkooei, AR, PJG Teunissen, and CCJM Tiberius (2009). "Application of least-squares variance component estimation to GPS observables". In: *Journal of Surveying Engineering* 135.4, pp. 149–160.

Amiri-Simkooei, AR and CCJM Tiberius (2007). "Assessing receiver noise using GPS short baseline time series". In: *GPS solutions* 11.1, pp. 21–35.

Amiri-Simkooei, AR, CCJM Tiberius, and sand PJG Teunissen (2007). "Assessment of noise in GPS coordinate time series: methodology and results". In: *Journal of Geophysical Research: Solid Earth* 112.B7.

Amiri-Simkooei, AR, S Zaminpardaz, and MA Sharifi (2014). "Extracting tidal frequencies using multivariate harmonic analysis of sea level height time series". In: *Journal of Geodesy* 88.10, pp. 975–988.

Amiri-Simkooei, AR et al. (2019). "Offset detection in GPS position time series using multivariate analysis". In: *GPS Solutions* 23.1, p. 13.

André, Gaël et al. (2013). "Measuring sea level with GPS-Equipped Buoys: A multi-instruments experiment at Aix island". In: *The International Hydrographic Review* 10.

Argus, Donald F and Michael B Heflin (1995). "Plate motion and crustal deformation estimated with geodetic data from the Global Positioning System". In: *Geophysical Research Letters* 22.15, pp. 1973–1976.

Baarda, W (1968). "A testing procedure for use in geodetic networks". In: *Publication on Geodesy, New Series* 2.

Ballu, Valérie et al. (2011). "Comparing the role of absolute sea-level rise and vertical tectonic motions in coastal flooding, Torres Islands (Vanuatu)". In: *Proceedings of the National Academy of Sciences* 108.32, pp. 13019–13022.

Ballu, Valérie et al. (2019). "Vertical land motion in the Southwest and Central Pacific from available GNSS solutions and implications for relative sea levels". In: *Geophysical Journal International* 218.3, pp. 1537–1551.

Becker, Melanie et al. (2012). "Sea level variations at tropical Pacific islands since 1950". In: *Global and Planetary Change* 80, pp. 85–98.

Benoist, C et al. (2020). "Accounting for spatiotemporal correlations of GNSS coordinate time series to estimate station velocities". In: *Journal of Geodynamics*, p. 101693.

Bertiger, Willy et al. (2020). "GipsyX/RTGx, A New Tool Set for Space Geodetic Operations and Research". In: *Advances in Space Research*.

Bevis, Michael and Abel Brown (2014). "Trajectory models and reference frames for crustal motion geodesy". In: *Journal of Geodesy* 88.3, pp. 283–311.

BIPM, IEC et al. (2008). "International vocabulary of metrology–basic and general concepts and associated terms, 2008". In: *JcGM* 200, pp. 99–12.

Blewitt, Geoffrey, William C Hammond, and Corné Kreemer (2018). "Harnessing the GPS data explosion for interdisciplinary science". In: *Eos* 99, pp. 1–2.

Blewitt, Geoffrey et al. (2019). "Improved GPS position time series spanning up to 25 years for over 18,000 stations resulting from IGS Repro3 products, JPL's GipsyX software, and more advanced modeling techniques". In: *AGUFM* 2019, G12A–04.

Boehm, Johannes, Birgit Werl, and Harald Schuh (2006). "Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for

Medium-Range Weather Forecasts operational analysis data". In: *Journal of geophysical research: solid earth* 111.B2.

Böhm, Johannes, Robert Heinkelmann, and Harald Schuh (2007). "Short note: a global model of pressure and temperature for geodetic applications". In: *Journal of Geodesy* 81.10, pp. 679–683.

Böhm, Johannes et al. (2006). "Global Mapping Function (GMF): A new empirical mapping function based on numerical weather model data". In: *Geophysical Research Letters* 33.7.

Bonnefond, P, BJ Haines, and C Watson (2011). "In situ absolute calibration and validation: a link from coastal to open-ocean altimetry". In: *Coastal altimetry*. Springer, pp. 259–296.

Bos, Machiel S et al. (2020). "Introduction to Geodetic Time Series Analysis". In: *Geodetic Time Series Analysis in Earth Sciences*. Springer, pp. 29–52.

Bos, MS et al. (2008). "Fast error analysis of continuous GPS observations". In: *Journal of Geodesy* 82.3, pp. 157–166.

— (2013). "Fast error analysis of continuous GNSS observations with missing data". In: *Journal of Geodesy* 87.4, pp. 351–360.

Bos, MS et al. (2014). "The effect of temporal correlated noise on the sea level rate and acceleration uncertainty". In: *Geophysical Journal International* 196.3, pp. 1423–1430.

Calais, Eric (1999). "Continuous GPS measurements across the Western Alps, 1996–1998". In: *Geophysical Journal International* 138.1, pp. 221–230.

Camargo, Carolina ML et al. (2020). "Exploring Sources of Uncertainty in Steric Sea-Level Change Estimates". In: *Journal of Geophysical Research: Oceans*, e2020JC016551.

Cartwright, David Edgar (2000). *Tides: a scientific history*. Cambridge University Press.

Cazenave, Anny and Gonéri Le Cozannet (2014). "Sea level rise and its coastal impacts". In: *Earth's Future* 2.2, pp. 15–34.

Cazenave, Anny et al. (2018). "Global sea-level budget 1993-present". In: *Earth System Science Data* 10.3, pp. 1551–1590.

Chupin, Clémence et al. (2020). "Mapping Sea Surface Height Using New Concepts of Kinematic GNSS Instruments". In: *Remote Sensing* 12.16, p. 2656.

Church, John A et al. (2013). *Sea level change*. Tech. rep. PM Cambridge University Press.

Davis, James L, Brian P Wernicke, and Mark E Tamisiea (2012). "On seasonal signals in geodetic time series". In: *Journal of Geophysical Research: Solid Earth* 117.B1.

Dill, Robert and Henryk Dobslaw (2013). "Numerical simulations of global-scale high-resolution hydrological crustal deformations". In: *Journal of Geophysical Research: Solid Earth* 118.9, pp. 5008–5017.

Dmitrieva, K, P Segall, and AM Bradley (2017). "Effects of linear trends on estimation of noise in GNSS position time-series". In: *Geophysical Journal International* 208.1, pp. 281–288.

Farrell, WE (1972). "Deformation of the Earth by surface loads". In: *Reviews of Geophysics* 10.3, pp. 761–797.

Feissel-Vernier, M, O de Viron, and K Le Bail (2007). "Stability of VLBI, SLR, DORIS, and GPS positioning". In: *Earth, planets and space* 59.6, pp. 475–497.

Fotopoulos, Georgia (2003). "An analysis on the optimal combination of geoid, orthometric and ellipsoidal height data". PhD thesis.

Franc, Vojtěch, Václav Hlaváč, and Mirko Navara (2005). "Sequential coordinate-wise algorithm for the non-negative least squares problem". In: *International Conference on Computer Analysis of Images and Patterns*. Springer, pp. 407–414.

Galloway, Devin L and Thomas J Burbey (2011). "Regional land subsidence accompanying groundwater extraction". In: *Hydrogeology Journal* 19.8, pp. 1459–1486.

Gazeaux, Julien et al. (2013). "Detecting offsets in GPS time series: First results from the detection of offsets in GPS experiment". In: *Journal of Geophysical Research: Solid Earth* 118.5, pp. 2397–2407.

Gobron, Kevin et al. (2019). "Assessment of Tide Gauge Biases and Precision by the Combination of Multiple Collocated Time Series". In: *Journal of Atmospheric and Oceanic Technology* 36.10, pp. 1983–1996.

Gouriou, Thomas, Belén Martín Míguez, and Guy Wöppelmann (2013). "Reconstruction of a two-century long sea level record for the Pertuis d'Antioche (France)". In: *Continental Shelf Research* 61, pp. 31–40.

Granger, Clive WJ and Roselyne Joyeux (1980). "An introduction to long-memory time series models and fractional differencing". In: *Journal of time series analysis* 1.1, pp. 15–29.

Gray, James E and David W Allan (1974). "A method for estimating the frequency stability of an individual oscillator". In: *28th Annual Symposium on Frequency Control. 1974*. IEEE, pp. 243–246.

Griffiths, Jake and Jim Ray (2016). "Impacts of GNSS position offsets on global frame stability". In: *Geophysical Journal International* 204.1, pp. 480–487.

Harris, Charles R et al. (2020). "Array programming with NumPy". In: *Nature* 585.7825, pp. 357–362.

Harville, David A (1977). "Maximum likelihood approaches to variance component estimation and to related problems". In: *Journal of the American statistical association* 72.358, pp. 320–338.

He, X et al. (2019). "Investigation of the noise properties at low frequencies in long GNSS time series". In: *Journal of Geodesy* 93.9, pp. 1271–1282.

Hosking, J. R. M. (1981). "Fractional differencing". In: *Biometrika* 68.1, pp. 165–176.

Hughes, Chris W and Simon DP Williams (2010). "The color of sea level: Importance of spatial variations in spectral shape for assessing the significance of trends". In: *Journal of Geophysical Research: Oceans* 115.C10.

Hurst, Harold Edwin (1951). "Long-term storage capacity of reservoirs". In: *Trans. Amer. Soc. Civil Eng.* 116, pp. 770–799.

— (1957). "A suggested statistical model of some time series which occur in nature". In: *Nature* 180.4584, pp. 494–494.

IOC (1985). *Manual on sea level measurement and interpretation. Volume I-Basic procedures.* UNESCO.

— (1994). *Manual on sea level measurement and interpretation. Volume II-Emerging technologies.* UNESCO.

IOC et al. (2002). *Manual on sea level measurement and interpretation. Volume III-Reappraisals and Recommendations as of the year 2000.* UNESCO.

— (2006). *Manual on sea level measurement and interpretation. Volume IV-An update to 2006.* UNESCO.

— (2012). "Global Sea Level Observing System (GLOSS): Implementation plan 2012". In: *Intergov. Oceanogr. Comm. Tech. Ser* 100.

— (2016). *Manual on sea level measurement and interpretation. Volume V-Radar gauges*. UNESCO.

Kasdin, N Jeremy (1995). "Discrete simulation of colored noise and stochastic processes and 1/f/sup/spl alpha//power law noise generation". In: *Proceedings of the IEEE* 83.5, pp. 802–827.

Kedar, Sharon et al. (2003). "The effect of the second order GPS ionospheric correction on receiver positions". In: *Geophysical Research Letters* 30.16.

Kennett, Brian LN, ER Engdahl, and R Buland (1995). "Constraints on seismic velocities in the Earth from traveltimes". In: *Geophysical Journal International* 122.1, pp. 108–124.

Khazraei, SM and AR Amiri-Simkooei (2019). "On the application of Monte Carlo singular spectrum analysis to GPS position time series". In: *Journal of Geodesy* 93.9, pp. 1401–1418.

King, Matt A et al. (2010). "Improved constraints on models of glacial isostatic adjustment: a review of the contribution of ground-based geodetic observations". In: *Surveys in geophysics* 31.5, pp. 465–507.

Klos, Anna et al. (2019). "Noise-dependent adaption of the Wiener filter for the GPS position time series". In: *Mathematical Geosciences* 51.1, pp. 53–73.

Koch, KR (1986). "Maximum likelihood estimate of variance components". In: *Bulletin Gædésique* 60.4, pp. 329–338.

Kreemer, Corné, Geoffrey Blewitt, and Paul M Davis (2020). "Geodetic evidence for a buoyant mantle plume beneath the Eifel volcanic area, NW Europe". In: *Geophysical Journal International* 222.2, pp. 1316–1332.

Kreemer, Corné, Geoffrey Blewitt, and Elliot C Klein (2014). "A geodetic plate motion and Global Strain Rate Model". In: *Geochemistry, Geophysics, Geosystems* 15.10, pp. 3849–3889.

Langbein, John (2004). "Noise in two-color electronic distance meter measurements revisited". In: *Journal of Geophysical Research: Solid Earth* 109.B4.

— (2012). "Estimating rate uncertainty with maximum likelihood: differences between power-law and flicker–random-walk models". In: *Journal of Geodesy* 86.9, pp. 775–783.

— (2017). "Improved efficiency of maximum likelihood analysis of time series with temporally correlated errors". In: *Journal of Geodesy* 91.8, pp. 985–994.

Langbein, John and Hadley Johnson (1997). "Correlated errors in geodetic time series: Implications for time-dependent deformation". In: *Journal of Geophysical Research: Solid Earth* 102.B1, pp. 591–603.

Larson, Kristine M, Richard D Ray, and Simon DP Williams (2017). "A 10-Year Comparison of Water Levels Measured with a Geodetic GPS Receiver versus a Conventional Tide Gauge". In: *Journal of Atmospheric and Oceanic Technology* 34.2, pp. 295–307.

Lennon, GW (1968). "The evaluation of tide-gauge performance through the Van de Casteele test". In: *Cah. Oceanogr* 20, pp. 867–877.

Lentz, Steven J (1993). "The accuracy of tide-gauge measurements at subtidal frequencies". In: *Journal of atmospheric and oceanic technology* 10.2, pp. 238–245.

Lomb, Nicholas R (1976). "Least-squares frequency analysis of unequally spaced data". In: *Astrophysics and space science* 39.2, pp. 447–462.

MacAulay, Phillip, Charles O'Reilly, and Keith Thompson (2008). "Atlantic Canada's Real-Time Water Level System Observations, Predictions, Forecasts and Datums on the Web". In:

Mandelbrot, Benoit B and John W Van Ness (1968). "Fractional Brownian motions, fractional noises and applications". In: *SIAM review* 10.4, pp. 422–437.

Mao, Ailin, Christopher GA Harrison, and Timothy H Dixon (1999). "Noise in GPS coordinate time series". In: *Journal of Geophysical Research: Solid Earth* 104.B2, pp. 2797–2816.

Marone, Chris J, CH Scholtz, and Roger Bilham (1991). "On the mechanics of earthquake afterslip". In: *Journal of Geophysical Research: Solid Earth* 96.B5, pp. 8441–8452.

Mémin, Anthony, Jean-Paul Boy, and Alvaro Santamaria-Gomez (2020). "Correcting GPS measurements for non-tidal loading". In: *GPS Solutions* 24.2, pp. 1–13.

Míguez Martín, Belén, Ronan Le Roy, and Guy Wöppelmann (2008). "The use of radar tide gauges to measure variations in sea level along the French coast". In: *Journal of coastal research* 24.sp3, pp. 61–68.

Míguez Martín, Belén, Laurent Testut, and Guy Wöppelmann (2008). "The Van de Casteele test revisited: an efficient approach to tide gauge error characterization". In: *Journal of Atmospheric and Oceanic Technology* 25.7, pp. 1238–1244.

— (2012). "Performance of modern tide gauges: towards mm-level accuracy". In: *Scientia Marina* 76.S1, pp. 221–228.

Munekane, Hiroshi and Johannes Boehm (2010). "Numerical simulation of troposphere-induced errors in GPS-derived geodetic time series over Japan". In: *Journal of Geodesy* 84.7, pp. 405–417.

Nicholls, Robert J and Anny Cazenave (2010). "Sea-level rise and its impact on coastal zones". In: *science* 328.5985, pp. 1517–1520.

Nocquet, J-M, E Calais, and B Parsons (2005). "Geodetic constraints on glacial isostatic adjustment in Europe". In: *Geophysical Research Letters* 32.6.

Pálinkáš, Vojtech et al. (2017). "Regional comparison of absolute gravimeters, EUROMET. MG-K2 key comparison". In: *Metrologia* 54.1A, p. 07012.

Park, Joseph, Robert Heitsenrether, and William Sweet (2014). "Water level and wave height estimates at NOAA tide stations from acoustic and microwave sensors". In: *Journal of Atmospheric and Oceanic Technology* 31.10, pp. 2294–2308.

Patterson, H Desmond and Robin Thompson (1971). "Recovery of inter-block information when block sizes are unequal". In: *Biometrika* 58.3, pp. 545–554.

Peltier, William R and John T Andrews (1976). "Glacial-isostatic adjustment—I. The forward problem". In: *Geophysical Journal International* 46.3, pp. 605–646.

Pérez, B et al. (2014). "Overlapping sea level time series measured using different technologies: an example from the REDMAR Spanish network". In: *Natural Hazards and Earth System Sciences* 14.3, p. 589.

Petrov, Leonid and Jean-Paul Boy (2004). "Study of the atmospheric pressure loading signal in very long baseline interferometry observations". In: *Journal of Geophysical Research: Solid Earth* 109.B3.

Pfeffer, Julia and Pascal Allemand (2016). "The key role of vertical land motions in coastal sea level variations: a global synthesis of multisatellite altimetry, tide gauge data and GPS measurements". In: *Earth and Planetary Science Letters* 439, pp. 39–47.

Pfeffer, Julia et al. (2017). "Decoding the origins of vertical land motions observed today at coasts". In: *Geophysical Journal International* 210.1, pp. 148–165.

Press, William H (1978). "Flicker noises in astronomy and elsewhere". In: *Comments on Astrophysics* 7, pp. 103–119.

Pugh, David and Philip Woodworth (2014). "Sea-level measuring systems". In: *Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes*. Cambridge University Press, 17–35.

Pytharouli, Stella, Spyros Chaikalis, and Stathis C Stiros (2018). "Uncertainty and bias in electronic tide-gauge records: Evidence from collocated sensors". In: *Measurement* 125, pp. 496–508.

Rao, C Radhakrishna (1971). "Minimum variance quadratic unbiased estimation of variance components". In: *Journal of Multivariate Analysis* 1.4, pp. 445–456.

Ray, J et al. (2008). "Anomalous harmonics in the spectra of GPS position estimates". In: *GPS solutions* 12.1, pp. 55–64.

Razeghi, SM, AR Amiri-Simkooei, and MA Sharifi (2016). "Coloured noise effects on deformation parameters of permanent GPS networks". In: *Geophysical journal international* 204.3, pp. 1843–1857.

Rebischung, P et al. (2012). "IGS08: the IGS realization of ITRF2008". In: *GPS solutions* 16.4, pp. 483–494.

Rebischung, Paul et al. (2016). "The IGS contribution to ITRF2014". In: *Journal of Geodesy* 90.7, pp. 611–630.

Rebischung, Paul et al. (2017). "Flicker Noise in GNSS Station Position Time Series: How much is due to Crustal Loading Deformations?" In: *AGU Fall Meeting Abstracts*.

Santamaría-Gómez, Alvaro and Anthony Mémin (2015). "Geodetic secular velocity errors due to interannual surface loading deformation". In: *Geophysical Journal International* 202.2, pp. 763–767.

Santamaría-Gómez, Alvaro and Jim R Ray (2020). "Chameleonic noise in GPS position time series". In:

Santamaría-Gómez, Alvaro et al. (2011). "Correlated errors in GPS position time series: Implications for velocity estimates". In: *Journal of Geophysical Research: Solid Earth* 116.B1.

Santamaría-Gómez, Alvaro et al. (2012). "Mitigating the effects of vertical land motion in tide gauge records using a state-of-the-art GPS velocity field". In: *Global and Planetary Change* 98, pp. 6–17.

Savage, JC and WH Prescott (1978). "Asthenosphere readjustment and the earthquake cycle". In: *Journal of Geophysical Research: Solid Earth* 83.B7, pp. 3369–3376.

Scargle, Jeffrey D (1982). "Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data". In: *The Astrophysical Journal* 263, pp. 835–853.

Schumacher, Maike et al. (2018). "A new global GPS data set for testing and improving modelled GIA uplift rates". In: *Geophysical Journal International* 214.3, pp. 2164–2176.

Schwarz, Gideon et al. (1978). "Estimating the dimension of a model". In: *The annals of statistics* 6.2, pp. 461–464.

Steigenberger, Peter, Johannes Boehm, and Volker Tesmer (2009). "Comparison of GMF/GPT with VMF1/ECMWF and implications for atmospheric loading". In: *Journal of Geodesy* 83.10, p. 943.

Syvitski, James PM et al. (2009). "Sinking deltas due to human activities". In: *Nature Geoscience* 2.10, pp. 681–686.

Takasu, Tomoji (2013). "RTKLIB ver. 2.4. 2 Manual". In: *RTKLIB: An Open Source Program Package for GNSS Positioning*, pp. 29–49.

Teunissen, PJG (1988). "Towards a least-squares framework for adjusting and testing of both functional and stochastic models." In:

— (2000a). "Adjustment theory, Series on Mathematical geodesy and positioning". In: Delft University Press.

— (2000b). "Testing theory an introduction, Series on Mathematical geodesy and positioning". In: Delft University Press.

Teunissen, PJG and AR Amiri-Simkooei (2008). "Least-squares variance component estimation". In: *Journal of geodesy* 82.2, pp. 65–82.

Valty, Pierre et al. (2013). "Assessing the precision in loading estimates by geodetic techniques in Southern Europe". In: *Geophysical Journal International* 194.3, pp. 1441–1454.

Van Camp, Michel and Olivier Francis (2007). "Is the instrumental drift of superconducting gravimeters a linear or exponential function of time?" In: *Journal of Geodesy* 81.5, pp. 337–344.

Van Camp, Michel, Simon DP Williams, and Olivier Francis (2005). "Uncertainty of absolute gravity measurements". In: *Journal of Geophysical Research: Solid Earth* 110.B5.

Van Dam, Tonie, Geoffrey Blewitt, and M Heflin (1994). "Detection of atmospheric pressure loading using the Global Positioning System". In: *Journal of Geophysical Research* 99.B12, pp. 23939–23950.

Van Dam, Tonie et al. (2001). "Crustal displacements due to continental water loading". In: *Geophysical Research Letters* 28.4, pp. 651–654.

Vaníček, Petr (1969). "Approximate spectral analysis by least-squares fit". In: *Astrophysics and Space Science* 4.4, pp. 387–391.

Viron, O de et al. (2008). "Retrieving earthquake signature in grace gravity solutions". In: *Geophysical Journal International* 174.1, pp. 14–20.

Wackernagel, Hans (2013). *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media.

Wang, Hansheng et al. (2012). "Load Love numbers and Green's functions for elastic Earth models PREM, iasp91, ak135, and modified models with refined crustal structure from Crust 2.0". In: *Computers & Geosciences* 49, pp. 190–199.

Wang, Lei and Thomas Herring (2019). "Impact of estimating position offsets on the uncertainties of GNSS site velocity estimates". In: *Journal of Geophysical Research: Solid Earth* 124.12, pp. 13452–13467.

Watson, Christopher, Richard Coleman, and Roger Handsworth (2008). "Coastal tide gauge calibration: a case study at Macquarie Island using GPS buoy techniques". In: *Journal of Coastal Research*, pp. 1071–1079.

Watson, Christopher et al. (2011). "Absolute calibration in bass strait, Australia: TOPEX, Jason-1 and OSTM/Jason-2". In: *Marine Geodesy* 34.3-4, pp. 242–260.

Watts, D Randolph and Harilaos Kontoyiannis (1990). "Deep-ocean bottom pressure measurement: Drift removal and performance". In: *Journal of Atmospheric and Oceanic Technology* 7.2, pp. 296–306.

Wdowinski, Shimon et al. (1997). "Southern California permanent GPS geodetic array: Spatial filtering of daily positions for estimating coseismic and postseismic displacements induced by the 1992 Landers earthquake". In: *Journal of Geophysical Research: Solid Earth* 102.B8, pp. 18057–18070.

Williams, SDP (2003a). "The effect of coloured noise on the uncertainties of rates estimated from geodetic time series". In: *Journal of Geodesy* 76.9-10, pp. 483–494.

Williams, Simon DP (2003b). "Offsets in global positioning system time series". In: *Journal of Geophysical Research: Solid Earth* 108.B6.

— (2008). "CATS: GPS coordinate time series analysis software". In: *GPS solutions* 12.2, pp. 147–153.

Williams, Simon DP et al. (2004). "Error analysis of continuous GPS position time series". In: *Journal of Geophysical Research: Solid Earth* 109.B3.

Williams, Simon DP et al. (2020). "Demonstrating the Potential of Low-Cost GPS Units for the Remote Measurement of Tides and Water Levels Using Interferometric Reflectometry". In: *Journal of Atmospheric and Oceanic Technology* 37.10, pp. 1925–1935.

Wöppelmann, Guy and Marta Marcos (2016). "Vertical land motion as a key to understanding sea level change and variability". In: *Reviews of Geophysics* 54.1, pp. 64–92.

Wöppelmann, Guy et al. (2007). "Geocentric sea-level trend estimates from GPS analyses at relevant tide gauges world-wide". In: *Global and Planetary Change* 57.3-4, pp. 396–406.

Zhang, Jie et al. (1997). "Southern California Permanent GPS Geodetic Array: Error analysis of daily position estimates and site velocities". In: *Journal of geophysical research: solid earth* 102.B8, pp. 18035–18055.

Zumberge, JF et al. (1997). "Precise point positioning for the efficient and robust analysis of GPS data from large networks". In: *Journal of geophysical research: solid earth* 102.B3, pp. 5005–5017.

Analyse statistique des mesures des mouvements verticaux du sol et du niveau de la mer à la côte.

Assurer la stabilité à long-terme des mesures du niveau de la mer à la côte est un enjeu essentiel à une gestion durable des territoires littoraux. Pour ce faire, les scientifiques reposent essentiellement sur deux techniques de mesure complémentaires : la marégraphie, qui permet de mesurer l'évolution du niveau de la mer par rapport à la côte, et la géodésie spatiale, qui permet de mesurer les mouvements verticaux de la côte elle-même, par rapport au centre de la terre. Ces techniques sont complémentaires car la correction des mouvements verticaux estimés par géodésie permet d'obtenir des mesures géocentriques du niveau de la mer à la cote, et donc de mieux comprendre l'origine des changements du niveau de la mer.

Afin d'estimer et améliorer la stabilité et la qualité de ces deux sources d'informations, cette thèse propose des développements méthodologiques dédiés à l'estimation de la précision de chaque type d'observation, et en étudie les sources d'erreurs potentielles.

Mots clés : mouvements verticaux du sol, marégraphes, géodésie, séries temporelles, analyses statistiques.

Statistical analysis of vertical land motions and sea level measurements at the coast

Ensuring the long-term stability of sea-level measurements at the coast is essential for the sustainable management of coastal territories. To do this, scientists mainly rely on two complementary measurement techniques: tide-gauges, which measure changes in sea-level with respect to the coast, and space geodesy, which allows measuring the vertical movements of the coast itself, with respect to the centre of the Earth. These two techniques are complementary because the correction of the vertical land motions estimated using geodetic techniques allows estimating geocentric sea-level change at the coast, and thereby to understand the origin of sea-level change better.

To estimate and improve the stability and quality of these two sources of information, this thesis proposes methodological developments dedicated to the estimation of the precision of each type of observation and investigates the potential sources of errors.

Keywords: vertical land motions, tide gauges, geodesy, time series, statistical analysis.