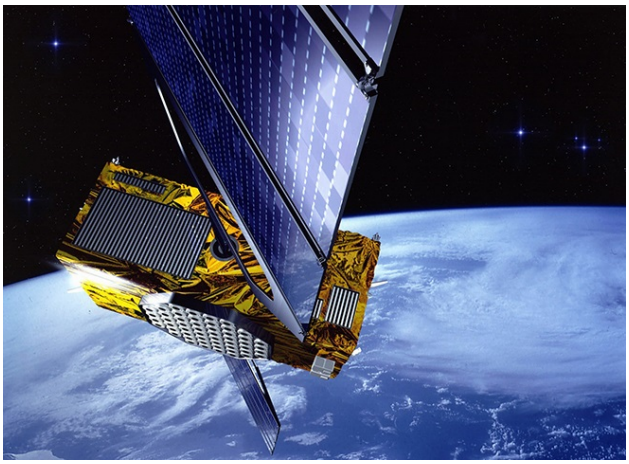
University of Stuttgart

Institute of Geodesy

Noise characterization of GPS time series from the second IGS reprocessing campaign



Bachelor Thesis

Ganyi Wang

Stuttgart, May 2015

Supervisor: Prof. Dr.-Ing. Nico Sneeuw
University of Stuttgart

M.Sc. Qiang Chen

Erklärung der Urheberschaft

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form in keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Ort, Datum Unterschrift

Abstract

By late 2013, the Analysis Centers (ACs) of the International Global Navigation Satellite System Service (IGS) begun the second reanalysis of the full history of Global Positioning System (GPS) data collected by the IGS global network since 1994 in a fully consistent way using the latest models and methodology. In this thesis, GPS time series of selected sites from different analysis centers were analysed with the software CATS. The purpose of this thesis is to discuss noise content using maximum likelihood estimation (MLE) and determine a noise model that can best describe the noise process in GPS time series of 20 years or more.

The noise was assumed to be a combination of white noise and power-law noise. In the first part, we used fixed noise models with integer spectral indices for the MLE. The three noise models are: white noise only, white noise plus flicker noise ($\kappa = -1$) and white noise plus random walk noise ($\kappa = -2$). The estimated MLE values of different noise models were compared to determine the preferred noise model. Another noise parameter, amplitude, was also estimated. Latitude dependence of the noise amplitudes and the influence of data cleaning on the noise analysis were discussed.

In the second part, the noise model was again assumed to be white noise plus power-law noise, but the spectral index of the power-law noise component was not fixed and was also estimated. The estimated spectral indices were compared with the results of the first part, in order to acquire a more precise description of the realistic noise components.

At last, a comparison between the noise features of GPS time series from different ACs was made. Besides the results of noise analysis, the importance and practical value of noise analysis was also discussed.

Contents

1 Introduction **1**
1.1 Importance and Motivation of Noise Analysis . . . 1
1.2 Previous Studies . . . 3
1.3 The Second IGS Data Reprocessing Campaign . . . 4
1.3.1 IGS Products . . . 4
1.3.2 Analysis Centers . . . 5
1.3.3 New Features of Data Analysis . . . 7
1.4 Outline of the Thesis . . . 11

2 Methodology **13**
2.1 Model Parameters . . . 14
2.2 Data Preprocessing . . . 16
2.3 Maximum Likelihood Estimation . . . 18
2.4 Power-law Noise Covariance Matrix . . . 19
2.4.1 Covariance Matrix for Noise with Integer spectral indices . . . 20

3 Data Processing and Software **23**
3.1 Data . . . 23
3.1.1 Source . . . 23
3.1.2 Download . . . 23
3.1.3 SINEX File . . . 24
3.1.4 Data Transformation . . . 26
3.2 Software CATS . . . 27
3.2.1 Nested Maximization . . . 28
3.2.2 The Time Series File . . . 29
3.2.3 Command Line Options . . . 31
3.2.4 Typical CATS Output . . . 32

4 Results **37**
4.1 Noise Models with Integer Spectral Index . . . 39
4.1.1 Preferred Noise Model . . . 39
4.1.2 Amplitudes estimation . . . 43
4.2 Noise Models with Non-integer Spectral Index . . . 46
4.3 Comparison between Different ACs . . . 49

5 Conclusion **53**

A Appendix **XVII**
A.1 MLE Value of Different Noise Model . . . XVII
A.2 Noise Amplitude Estimation . . . XXIII

List of Figures

1.1 Deep-Drilled Braced Monument (DDBM) (Courtesy: Borsa et al. (2008)) 2
1.2 Short-Drilled Braced Monument (SDBM) (Courtesy: Borsa et al. (2008)) 2
1.3 Comparison between white and power-law noise amplitudes in north and east of DDBM and SDBM (Courtesy: Borsa et al. (2008)) 3
1.4 WRMS of the station position residuals (Courtesy: Rebischung et al. (2014)) 5
1.5 WRMS of the station position residuals (combined) (Courtesy: Rebischung et al. (2014)) 6
1.6 3D root mean square of repro1 combination residuals (Courtesy: Rebischung et al. (2014)) 8
1.7 Scale factors estimated between the daily combined repro2 solution and IGb08 (Courtesy: Rebischung et al. (2014)) 8
1.8 GPS Coverage of different Analysis Centers (Courtesy: Rebischung et al. (2014)) . 9
1.9 IGS tracking network (Source: https://igscb.jpl.nasa.gov/network/complete.html) 10

2.1 Time series for Arequipa, Peru before and after earthquake 14
2.2 Data Cleaning after level 1 (3IQR) and 2 (2IQR) 17

3.1 Log-likelihood estimates for the north component of GPS site BRST. **(a)** Log-likelihood values greater than 2130 plotted in amplitude space, that is, as a function of flicker noise and white noise amplitude. Black triangles indicate the path of the uphill simplex from $\sigma_1 = 1$; $\sigma_2 = 1$. White dots are the predictions from the 1D brent maximizations. White line is the exact maximum at all possible angles. **(b)** as a except in ϕ and r space. **(c)** The results of the 1D maximization which calculates the maximum log-likelihood values as a function of ϕ. (Courtesy: Williams (2008)) 30

4.1 IGS Tracking Network Europe Region 38
4.2 Percentage as preferred noise model between flicker + white noise and random walk + white noise 41
4.3 Selected sites to demonstrate latitude dependence 45
4.4 Latitude dependence of noise amplitudes 46
4.5 Histogram of the estimated spectral index 48
4.6 East components of site MATE 49
4.7 Up components of site GRAZ 50

List of Tables

2.1 Number of data points before and after cleaning 18

3.1 SINEX File content 26
3.2 SOLUTION/ESTIMATION Block 26

4.1 Selected sites for noise analysis 38
4.2 MLE results of North components before data cleaning 40
4.5 Percentage of time series with preferred noise model 40
4.3 MLE results of North components after data cleaning level 1 41
4.4 MLE results of North components after data cleaning level 2 42
4.6 Estimated noise amplitudes before data cleaning 43
4.7 Selected sites to demonstrate latitude dependence 44
4.8 Estimated Spectral index of Power Law Noise 47
4.9 Percentage of time series from different ACs with preferred noise model 50
4.10 Average of estimated spectral indices from different ACs 51

A.1 MLE results of East components before data cleaning XVII
A.2 MLE results of East components after data cleaning level 1 XVIII
A.3 MLE results of East components after data cleaning level 2 XIX
A.4 MLE results of Up components before data cleaning XX
A.5 MLE results of Up components after data cleaning level 1 XXI
A.6 MLE results of Up components after data cleaning level 2 XXII
A.7 Estimated noise amplitudes after data cleaning level 1 XXIII
A.8 Estimated noise amplitudes after data cleaning level 2 XXIV

Chapter 1

Introduction

1.1 Importance and Motivation of Noise Analysis

Due to a mixture of effects, it is difficult to determine the source of the noise and the accuracy of the GPS time series. Understanding the content of the noise is very important so that realistic uncertainties can be assigned to parameters estimated from them. Previous studies showed that, if a correlated noise model is neglected and the white noise only model is used, the rate uncertainties will be greatly underestimated by as much as an order or magnitude. So it is obvious that the assumed noise model will significantly affect the estimated rate uncertainties.

As mentioned above, we use 3 noise type and their combination as fixed noise models for the analysis. First, the noise will be assumed to be white noise only, a combination of white noise plus flicker noise, or a combination of white noise plus random walk noise. A preferred noise model will be determined among these 3 possibilities. For the second part the spectral indices and amplitudes of the power-law noise will also be estimated simultaneously with the white noise. Maximum likelihood estimation (MLE) will be used to analyse the noise parameter and determine a preferred noise model, which will be described and explained later.

However the realistic purpose and practical value of noise analysis, is to possibly remove the noise or at least reduce it. After the analysis realistic parameter uncertainties are provided, a means for minimizing the noise still needs to be determined. That is why classifying and quantifying the noise components are so important, because understanding the noise components can provide clues as to the source of the noise and to help increase the accuracy and precision of the data. With noise information of long time series of over two decades, There is even possibility to prognose the trend and take precautions against possible future source of the noise.

In this thesis we would love to compare different noise models, estimate noise parameters such as amplitudes and spectral indices of the power-law noise and hopefully determine a generally preferred noise model for all GPS time series. As showed in next section, some early studies have already worked on this subject. The purpose is to compare our results with them and discover whether the new data reprocessing campaign and a extended length of GPS time series have influences on the noise analysis.

An example from Williams et al. (2004) illustrates the practical use of noise analysis: because of the varying conditions of the anchoring media the instability of geodetic monument is considered an significant source of random walk noise. To minimize the noise caused by it some GPS arrays have adopted deep drill braced monument. The new western North America Plate

Boundary Observatory (PBO) installed 850 new continuous GPS monuments using the deep-drilled braced monument. Figure 1.1 and Figure 1.2 compared the configuration of deep-drilled and short-drilled braced monuments (SDBM and DDBM).

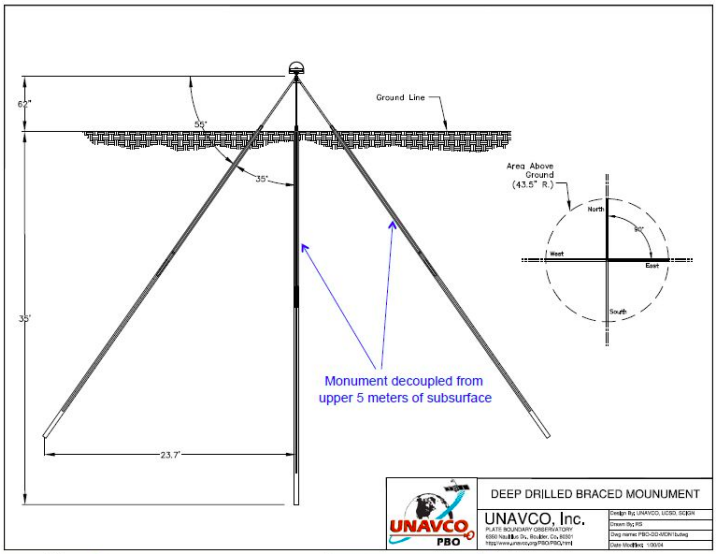


Figure 1.1: *Deep-Drilled Braced Monument (DDBM) (Courtesy: Borsa et al. (2008))*

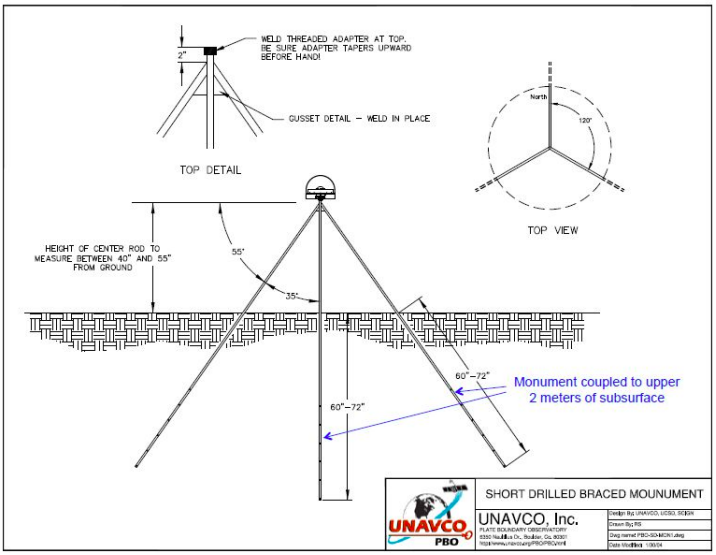


Figure 1.2: *Short-Drilled Braced Monument (SDBM) (Courtesy: Borsa et al. (2008))*

Later study of Borsa et al. (2008) concluded that although the white noise amplitude, power-law noise amplitude and spectral index, and annual/semi-annual sinusoid amplitudes of all show statistical equivalence between the performance of SDBM and DDBM stations, Annual and semi-annual sinusoid fitting shows a difference in amplitude between north and east components.

Figure 1.3 is an example of noise analysis and the practical use of the study. We can see that, noise analysis is not only a means to classify the noise source, but also a way to test the efficiency of any noise reducing methods.

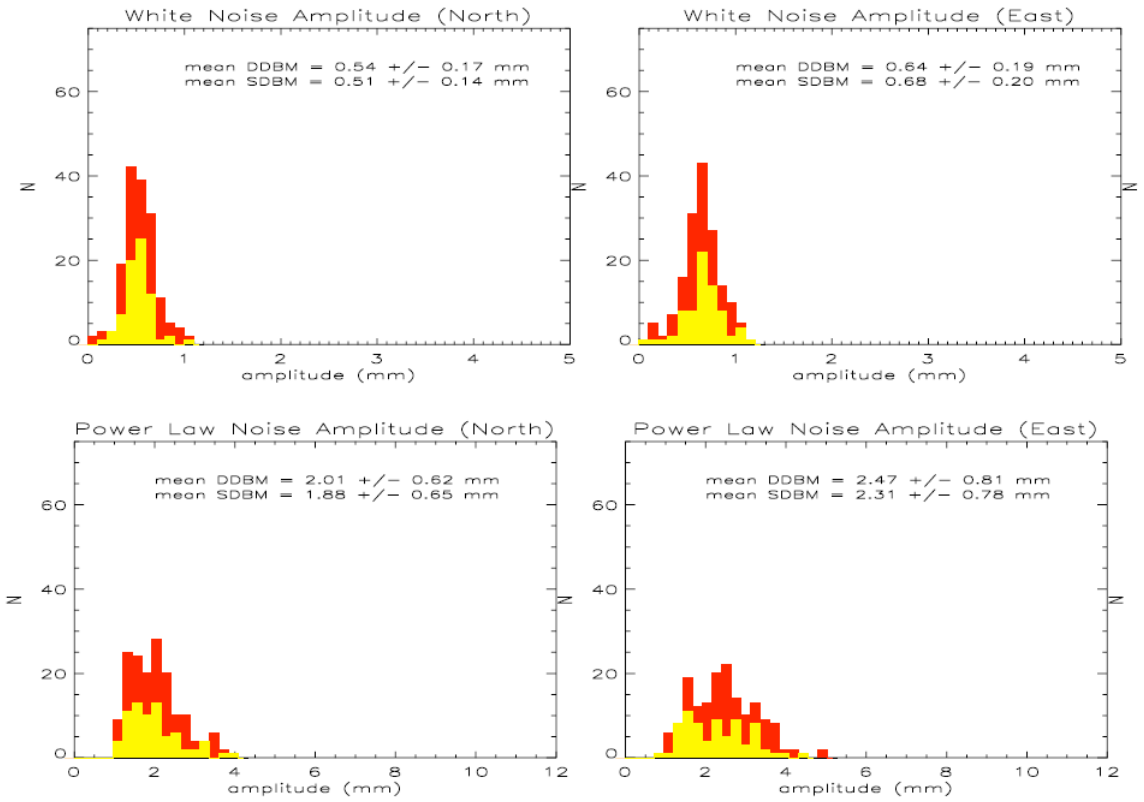


Figure 1.3: *Comparison between white and power-law noise amplitudes in north and east of* DDBM *and* SDBM *(Courtesy: Borsa et al. (2008))*

1.2 Previous Studies

Until recently, it was typically assumed that only white noise was present in geodetic time series. However, several data sets have now provided evidence for the presence of power-law noise. In addition, equations to analyse data sets and obtain a rate uncertainty when computational speed is improved are provided, which enables researchers to not only classify but also quantify the noise components of GPS time series.

Zhang et al. (1997) examined 19 months of continuous GPS data from 10 sites in southern California. Using maximum likelihood estimation (MLE) with integer spectral indices, Zhang et al. (1997) found that the noise in the data was best described as a combination of white noise and flicker noise. Using the power spectra, the noise was characterized by a fractal noise process with spectral index of -0.4.

Mao et al. (1999) examined 3 years of data from a global distribution of 23 CGPS (Continuous GPS) sites. Using both MLE at integer spectral indices and power spectra, they concluded that

white noise plus flicker noise best describe the noise content of the time series. A latitude dependence was also found for the white noise part of the vertical component.

Calais (1999) analysed three permanent GPS sites in Europe. A combination of white noise and flicker noise was the prferred model.

Several studies also acknowledge the importance of random walk noise in GPS data. Cumulative disturbances from the soils and weather displace geodetic monuments with respect to the deeper crust (Johnson and Agnew, 1995).

Whether or not the random walk noise is detectable depends on the length of the time series, the sampling frequency, and the relative amplitudes of the other noise components. It has been identified in continuous strainmeter data (Wyatt, 1982, 1989) and two-color electronic distance meter data (Langbein and Johnson, 1995). However, this type of disturbance can be minimized by carefully designed monuments, like GPS stations deployed in southern California in which the base is securely anchored at depth and decoupled from surface (Wyatt, 1989; Zhang et al., 1997).

However, regional GPS networks have much longer inter-station spacing so that other sources of error such as known random atmospheric propagation effects could dominate the error budget (Williams et al., 1998).

1.3 The Second IGS Data Reprocessing Campaign

By late 2013, the Analysis Centers (ACs) of the International GNSS Service began the second reanalysis of the full history of GPS data collected by the IGS global network since 1994 in a fully consistent way using the latest models and methodology. This effort follows the successful first full reprocessing by the IGS, which provided the IGS input for ITRF2008, among other things.

1.3.1 IGS Products

The expected Analysis Center and combined IGS products include:

- daily GPS & GLONASS orbits & GPS satellite clocks
- daily GPS satellite & tracking station clocks
- daily Earth rotation parameters
- terrestrial coordinate frames with ERPs with full variance-covariance matrix (SINEX format).

  And this is the product which will be used in the noise analysis.

1.3.2 Analysis Centers

The possible Analysis Centers contributing complete solutions to this effort are:

- CODE (Center for Orbit Determination in Europe)
- EMR (Natural Resources Canada)
- ESA (European Space Operations Centre)
- GFZ (GeoForschungsZentrum)
- GRGS (Groupe de Recherche de Geodesie Spatiale)
- JPL (Jet Propulsion Laboratory)
- MIT (Massachusetts Institute of Technology)
- NGS (National Geodetic Survey)
- SIO (Scripps Institution of Oceanography)

In addition, some centers might provide SINEX solutions to help densify the terrestrial reference frame, particularly for GPS stations located near tide gauges (associated with the TIGA Pilot Project):

- GTZ (GFZ TIGA, Germany)
- ULR (University of La Rochelle, France)

As time series used in this thesis are from different analysis centers. A comparison between the noise of time series from these ACs is necessary. At the Fall 2014 American Geophysical Union Meeting, Rebischung et al. (2014) compared the residuals of data from different ACs with the weighted root mean square of station position residuals.

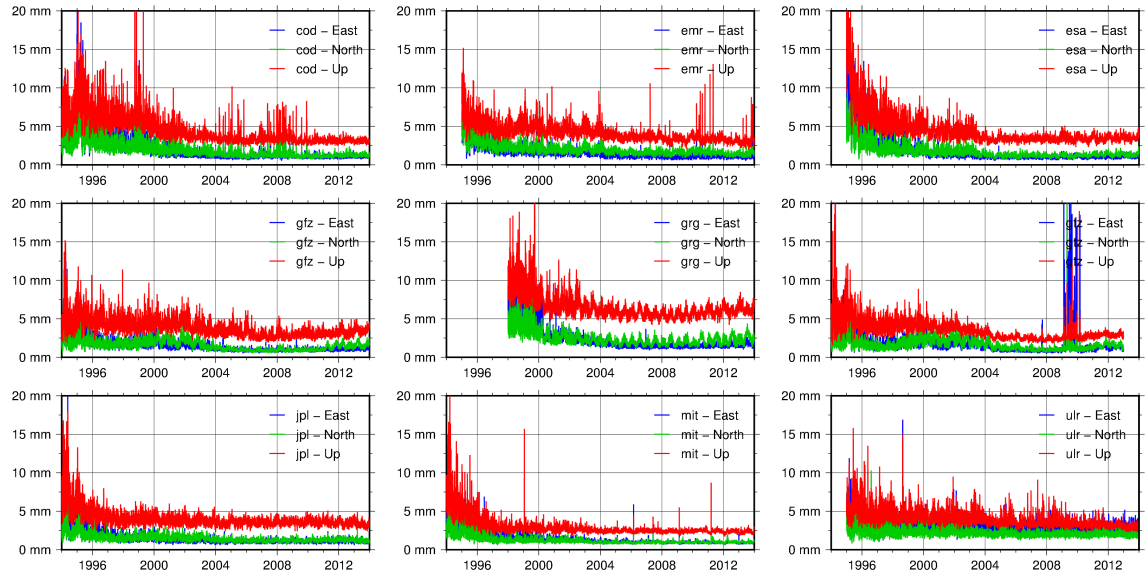


Figure 1.4: WRMS *of the station position residuals (Courtesy: Rebischung et al. (2014))*

Rebischung et al. (2014) combined smoothed WRMS of the station position residuals from 8 different analysis centers together. As shown in 1.5, a few features can be summarized:

- All ACs: the daily station position estimates provided by the different ACs are, in general, of homogeneous quality. The inter-AC agreement is at the level of 1 mm to 1.5 mm in horizontal and 3 mm to 4 mm in vertical after 2004.
- CODE: shows higher noise level in early years (before 2000)
- EMR: higher WRMS in North than the other components, possibly due to station biases
- ULR: high level of high-frequency white noise, especially in horizontal, possibly systematic errors
- GRGS: substantially higher WRMS in Up and in East before 2000, obvious semi-annual variations in North and Up
- GTZ: peak in East between 2009 and 2010 with unknown reason, , possibly due to station biases which later corrected
- MIT: almost dominant in all 3 components

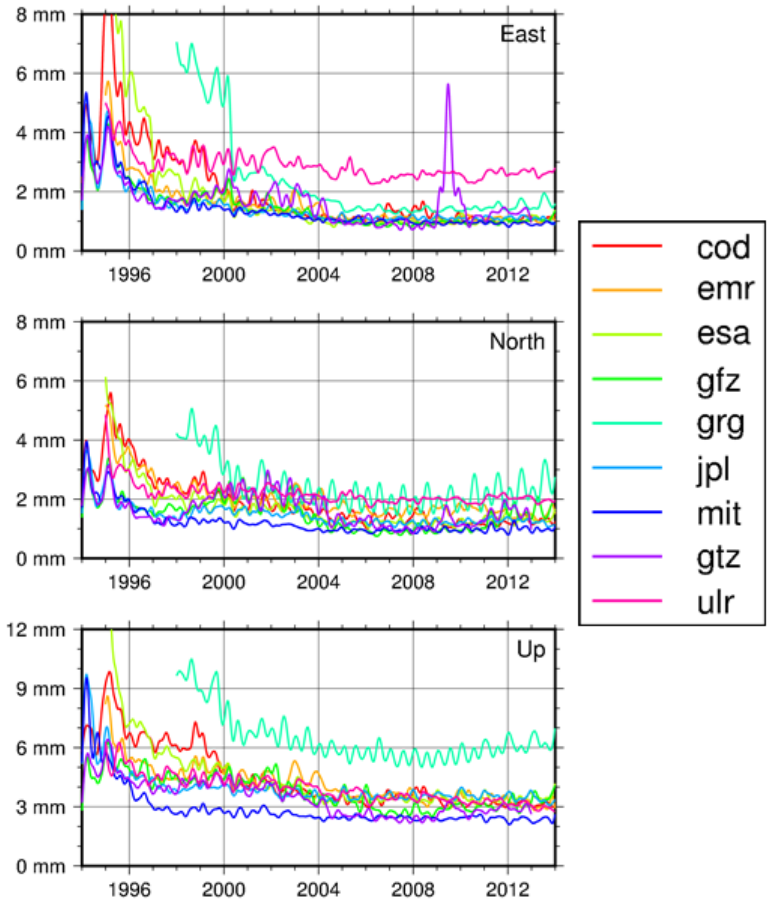


Figure 1.5: WRMS *of the station position residuals (combined) (Courtesy: Rebischung et al. (2014))*

1.3.3 New Features of Data Analysis

In the first data reprocessing campaign there are several elements of the analysis procedures used by all participating centers include:

- try to complete archives for missing or useful data sets
- use absolute antenna calibrations for satellite transmit and ground receive antennas
- IERS Conventions 2003 generally implemented
- updated model for station displacements due to ocean tidal loading
- no non-tidal loading displacements should be applied to station positions
- updated models for tropospheric propagation delays
- the IGS master SINEX template file, igs.snx, is updated daily based on the history of IGS site logs and should be used to validate AC SINEX file metadata
- P1-C1 satellite code biases

Compared with the first IGS data reprocessing campaign, there are new models and methodology used and few changes made in the second campaign. The purpose of the re-analysis of GNSS data collected by the IGS network since 1994 using the latest models and methodology is to reduce systematic errors in IGS products and provide IGS contribution to ITRF2013.

The main updates since repro 1 are:

- use IGB08 reference frame (aligned to ITRF2008)
- use updated igs08.atx "absolute" antenna calibrations
- yaw attitude variations are taken into consideration
- reflected (albedo) and retransmitted radiation from the Earth are taken into consideration
- EGM2008 geopotential field now recommended

  to update values for time-variations of low-degree coefficients and use new model for the mean pole trajectory, geopotential ocean tide and ocean pole tide.
- move from weekly to daily frame integrations in order to fully preserve loading signals in IGS position time series
- addition of UT1 libration effects for the tidal EOP variation
- higher-order ionospheric corrections
- for repro2, ACs are asked to avoid any solution over-constraints, applying pre-removed or unremovable constraints no tighter than noise levels

To compare with the result of the second data reprocessing campaign, figure 1.6 shows the combination of root mean square of the station position residuals in all three components of the first campaign.

- In the aspect of station positions:

– the post-2004 inter-AC agreement of the second campaign, despite all of the possible systematic errors, is comparable to the first campaign.

– The residuals before 1996 are substantially higher.

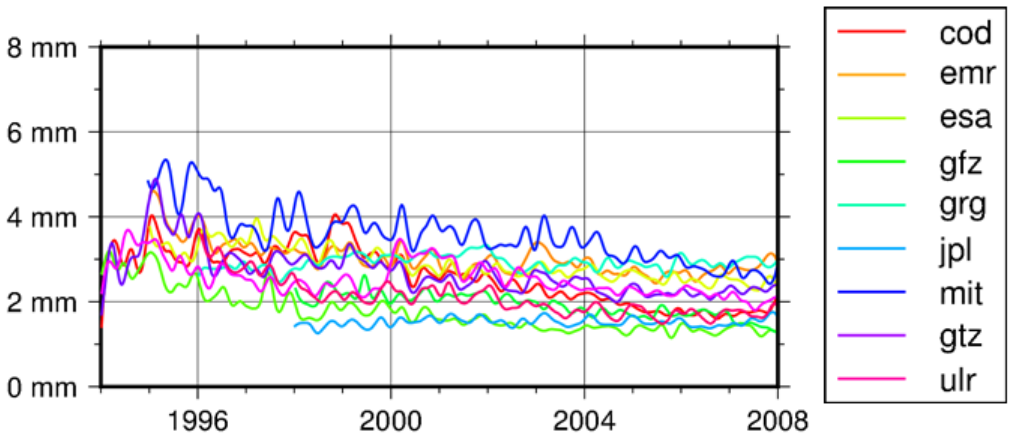


Figure 1.6: *3D root mean square of repro1 combination residuals (Courtesy: Rebischung et al. (2014))*

- Although it is not researched in this thesis, the second reprocessing campaign also shows excellent inter-AC agreement for the geocenter coordinate. (at the level of 2 mm to 5 mm for the X and Y geocenter coordinates and 5 mm to 10 mm for the Z geocenter coordinate)
- The temporal scale variation of the combined repro2 solutions seems geophysically reasonal as showed in figure 1.7. It demonstrates an identical variation pattern with IGb08.

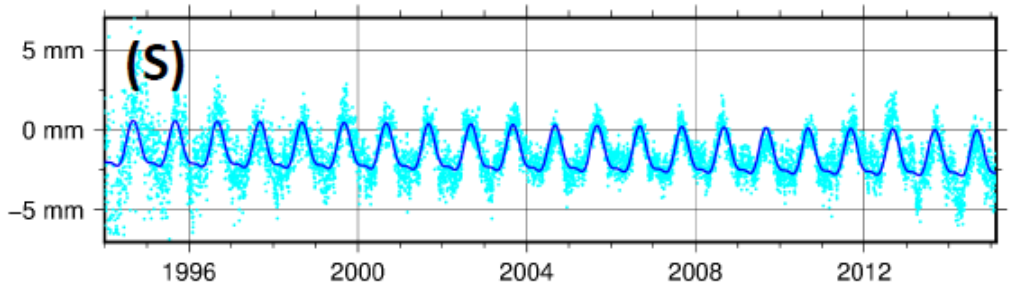


Figure 1.7: *Scale factors estimated between the daily combined repro2 solution and IGb08 (Courtesy: Rebischung et al. (2014))*

In the previous studies of GPS time series noise analysis, an important disadvantage is the lack of data points, which means the overall length of the time series is not long enough for some specific noise models to be detected over other noises. As a result, the estimated parameters can be biased, even a correlated noise model can be recognized as one noise only model.

Figure 1.8 below shows the time coverage of the Analysis Centers from the second Data Reprocessing Campaign. GPS time series of different institutes over two decades are now available.

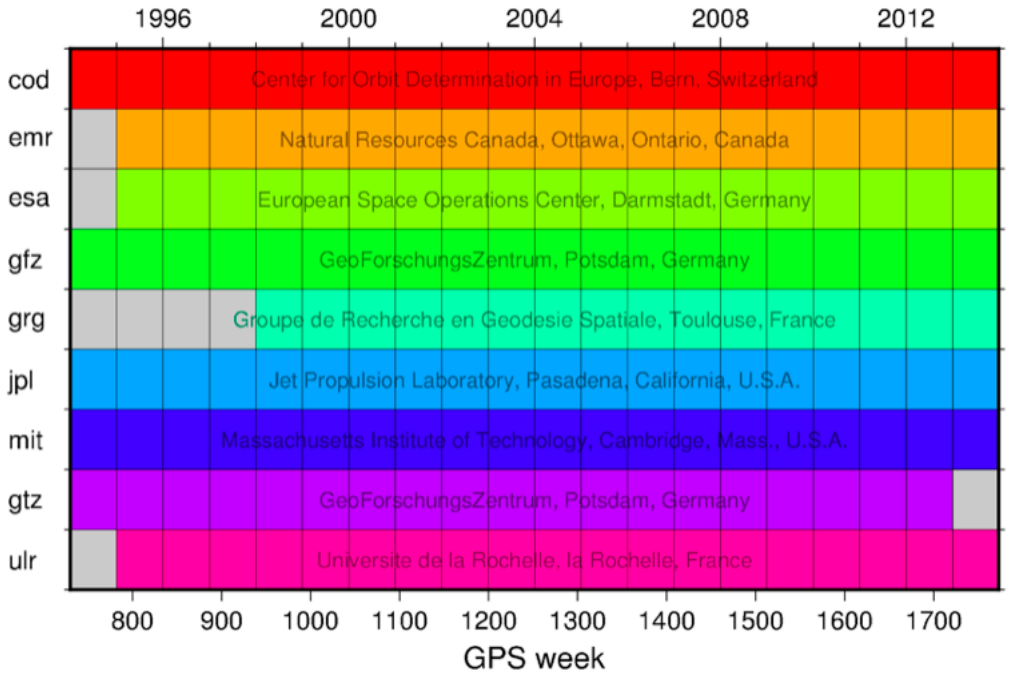


Figure 1.8: GPS *Coverage of different Analysis Centers (Courtesy: Rebischung et al. (2014))*

Among all the combined IGS products from the second data reprocessing campaign, sufficient SINEX data provide us with continuous GPS position time series of hundreds of stations over recent 20 years, which can be used for a variety of analysis and researches and are certainly a huge advantage for a precise determination of the parameter uncertainties and classification of the noise component.

For many previous studies, a drawback was the lack of long time series with sufficient data points. Some time series were simply not long enough for some noise components to be detected.

Figure 1.9 is a map of the IGS tracking network. GPS time series from different sites all over the world can be downloaded for research purpose.

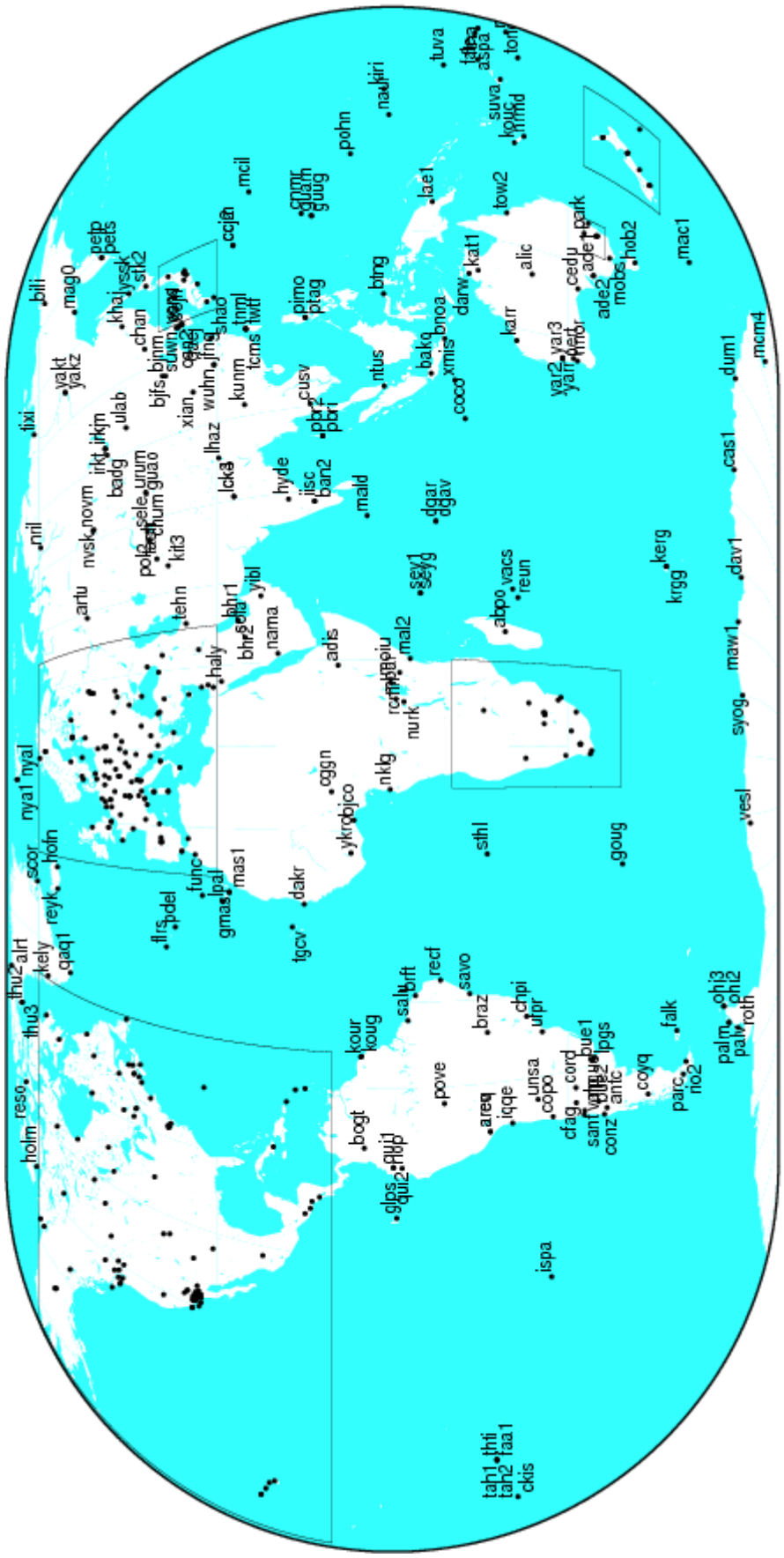


Figure 1.9: IGS *tracking network (Source: https://igscb.jpl.nasa.gov/network/complete.html)*

1.4 Outline of the Thesis

In the first chapter, the background and purpose of this thesis were introduced. We discussed the importance of noise analysis and the motivation of the research. Previous studies on this subject were presented and explained to help the readers understand the development and progress in noise analysis of GPS time series. A detailed introduction of the second IGS data reprocessing campaign was made including the products, analysis centers and the upgraded features compared with the first reprocessing campaign.

The second chapter of the thesis explained the methods used in the noise analysis. First of all, the concept of power-law noise and the three noise models we used as fixed models in the analysis was introduced. Second, the various steps involved in modeling the coordinate time series were described. To eliminate outliers, a detecting algorithm was introduced and discussed. Then we explained the theory of maximum likelihood estimation to estimate the noise component and determine a preferred model to best describe a time series. At last, the derivation of covariance matrices for different noise models was explained.

The third chapter is mainly about data processing and the related software. For data processing, we explained in detail the source of the data, the method to download, the content of the data and the transformation of data to required format. To introduce the software (CATS), we explained the method used to reduce computational burden, the operational principle of the software and typical input/output files.

In the fourth chapter, the estimation results were presented and discussed. In the first part, the results of estimation with noise models with integer spectral indices were compared to determine a generally preferred noise model among the three fixed models. The amplitudes of noise components were also estimated and their latitude dependence was tested. In the second part, noise models with any spectral indices were used for the estimation. The spectral index of power-law noise was also estimated, the results were compare with the results from the first part. At last, the GPS time series from different ACs and their noise characteristics were compared and discussed.

The last chapter concluded the results of our estimation and explained the results with possible reasons. Thoughts inspired by this research were also expressed in this chapter.

Chapter 2

Methodology

This chapter will explain the methods used in the noise analysis including model parameters and maximum likelihood estimation.

Since the beginning of the present century, geodetic measurements have been used to monitor crustal deformations. Without significant episodic deformations, such as large earthquakes, it is typical to fit a linear trend to repeated measurement and obtain deformation rate. The traditional means to determine the rate is linear regression of coordinate component and geodetic network adjustment. Usually the measurement errors are assumed to be statistically uncorrelated from one another (white noise). However studies have showed that there are evidences for large correlations in the geodetic data. The assumption that the noise is purely white noise leads, for example, to grossly underestimated site rate uncertainties.

Previous studies suggest that noise in GPS position time series can be described as a power-law process or one with time domain behavior that has power spectrum of the form:

$$P_x(f) = P_0 \left(\frac{f}{f_0} \right)^{\kappa}, \tag{2.1}$$

where f is the temporal frequency, P_0 and f_0 are constant, and κ is the spectral index (Agnew, 1992).

Typically the spectral index, κ, lies within the range $[-3, 1]$. Noise with different spectral index can be divided into stationary and nonstationary process.

- fractional Brownian motions (nonstationary)

  Naturally occurring process often have negative indices ranging from $-3 < \kappa < -1$ including classical Brownian motion "random walk" with $\kappa = -2$.

- fractional Gaussian process (stationary)

  Stationary processes with $-1 < \kappa < 1$ are called "fractional Gaussian" process including the special case of uncorrelated white noise $(\kappa = 0)$.

As the most typical and common noise components, the following 3 specific noise models will be used to form fixed noise combinations for the later analysis.

- **White noise**: a random signal with a constant power spectral density.

- **Random walk noise**: a classical Brownian motion with $\kappa = -2$.

  It is often caused by monument instabilities and has been identified in continuous strainmeter data and two-color electronic distance meter data.

- **Flicker noise**: pink power density spectrum with $\kappa = -1$.

  It is commonly observed in a wide variety of dynamical processes, including sunspot variability, the wobble of the Earth about its axis, undersea currents and uncertainties in time measured by atomic clocks.

We will use the term coloured noise to refer to power-law process other than classical white noise.

2.1 Model Parameters

For the analysis the daily position for each site are first combined to form time series of geodetic positions. The cartesian XYZ coordinate time series are then rotated to a topocentric north, east, up (NEU) coordinate system with:

$$\begin{bmatrix} N \\ E \\ U \end{bmatrix} = \begin{bmatrix} -\sin(\lambda)\cos(\phi) & -\sin(\lambda) - \sin(\phi) & \cos(\phi) \\ -\sin(\phi) & \cos(\phi) & 0 \\ \cos(\lambda)\cos(\phi) & \cos(\lambda)\sin(\phi) & \sin(\lambda) \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} . \tag{2.2}$$

Figure 2.1 shows an example of modeled coordinate time series for site AREQ (Arequipa, Peru) in the north direction. The model for this time series contains 1 offset in year 2001 and a linear trend.

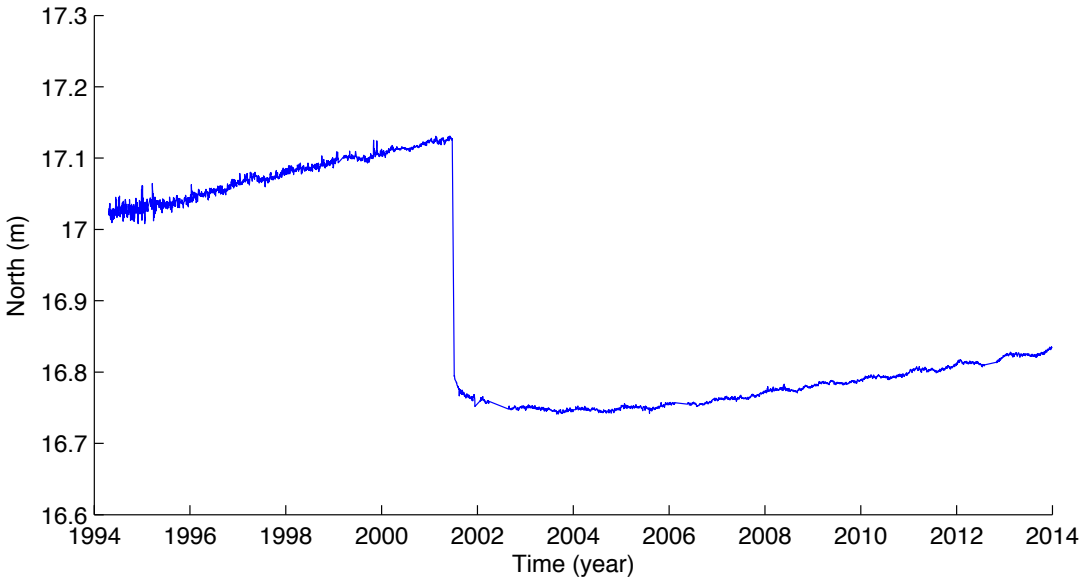


Figure 2.1: *Time series for Arequipa, Peru before and after earthquake*

The various steps involved in modeling the coordinate time series are described below.

The observed motion $y(t)$ of each site in each direction can be written as:

$$y(t_i) = a + bt_i + c\sin(2\pi t_i + d\cos(2\pi t_i) + e\sin(4\pi t_i) + f\cos(4\pi t_i) + \sum_{j=1}^{n} g_j H(t_i - E_j) + v_i \,, \quad (2.3)$$

where t_i for $i = 1...N$ are the daily solution epochs in units of years, and H is the heaviside step function. The first two terms are the site position a, and linear rate b, respectively. c and d are the coefficients that describe the annual periodic motion, e and f describe semi-annual motion. The next term corrects for n offsets, with magnitudes g_j and epochs E_j. v is a vector of independent measurement errors.

Assuming that the offset epochs are known, the model is linear with respect to the coefficients

$$x = \begin{bmatrix} a & b & c & d & e & f & g_1 ... g_n \end{bmatrix} \,, \quad (2.4)$$

so that

$$y = Ax + v \,, \quad (2.5)$$

where A is the design matrix of partial derivatives.

For the case of uncorrelated white noise, the observation covariance matrix is defined by the individual variances of the daily position solutions:

$$T = \Delta T^{-\kappa/4} \begin{bmatrix} \phi_0 & 0 & 0 & ... & 0 \\ \phi_1 & \phi_0 & 0 & ... & 0 \\ \phi_2 & \phi_1 & \phi_0 & ... & 0 \\ ... & ... & ... & ... & ... \\ \phi_N & \phi_{N-1} & \phi_{N-2} & ... & \phi_0 \end{bmatrix} \,. \quad (2.6)$$

The weighted least squares solution for the best linear unbiased estimates of the unknown parameters is then

$$\hat{x} = (A^T C^{-1} A)^{-1} A^T C^{-1} y \,, \quad (2.7)$$

with parameter covariance

$$\hat{C_x} = \chi^2 (A^T P^{-1} A)^{-1} \,, \quad (2.8)$$

The postfit residuals are

$$\hat{v} = y - Ax \,, \quad (2.9)$$

and the reduced square statistic

$$\chi_v^2 = \frac{\hat{v}^T C^{-1} \hat{v}}{(N-\nu)^2}\,, \tag{2.10}$$

where N is the number of daily position solutions and ν is the number of model parameters.

It is impossible to automatically locate the offsets. To identify the offsets the data residual need to be carefully inspected and a knowledge of the site history is also necessary. The identification of subtle offsets often required several iterations through the model. Generally offsets were placed in the epochs when either hardware changes at the site or nearby seismic events. However exceptions can also be found with large offsets and no known explanation.

Figure 2.1 is an obvious example of post-seismic deformation. It is caused by the earthquake near the coast Peru, 175 km west of Arequipa, on 23 July 2001. Not all the terms in the motion equation are necessarily relevant. For example, estimating very small offsets can increase the uncertainty of more important parameters, such as the linear rate.

2.2 Data Preprocessing

To eliminate the outliers, each position time series was "cleaned" using a robust outlier detection algorithm applied to the postfit residual. The cleansing algorithm is based on the median and interquartile range (IQR) statistics to describe the central value and spread of the data. The IQR of a data sample is the difference between its 75th and 25th percentile, and is a robust measure of scale. The median and IQR are computed within a sliding window centered on each measurement. Outliers are defined as having:

$$|\hat{v}_i - \text{median}(\hat{v}_{i-w/2,i+w/2})| > 3 \cdot \text{IQR}(\hat{v}_{i-w/2,i+w/2})\,, \tag{2.11}$$

where median and IQR operate on the residuals within the window of size w, chosen to be 1 year. The window center for all the measurement within 6 months of the start or end of observation remains fixed at 6 months from the time series endpoints. For each site, outlier epochs are identified separately in each coordinate direction and then applied to all three coordinate directions.

As a result the cleaning algorithm removed 0.1–0.4% of the data from each global site, which did not reach the average 1–4% level of data cleaning from Nikolaidis (2002). To adjust the intensity of the cleansing algorithm, the constant in the criterion was changed to eliminate more data points and compare the results:

$$|\hat{v}_i - \text{median}(\hat{v}_{i-w/2,i+w/2})| > 2 \cdot \text{IQR}(\hat{v}_{i-w/2,i+w/2})\,. \tag{2.12}$$

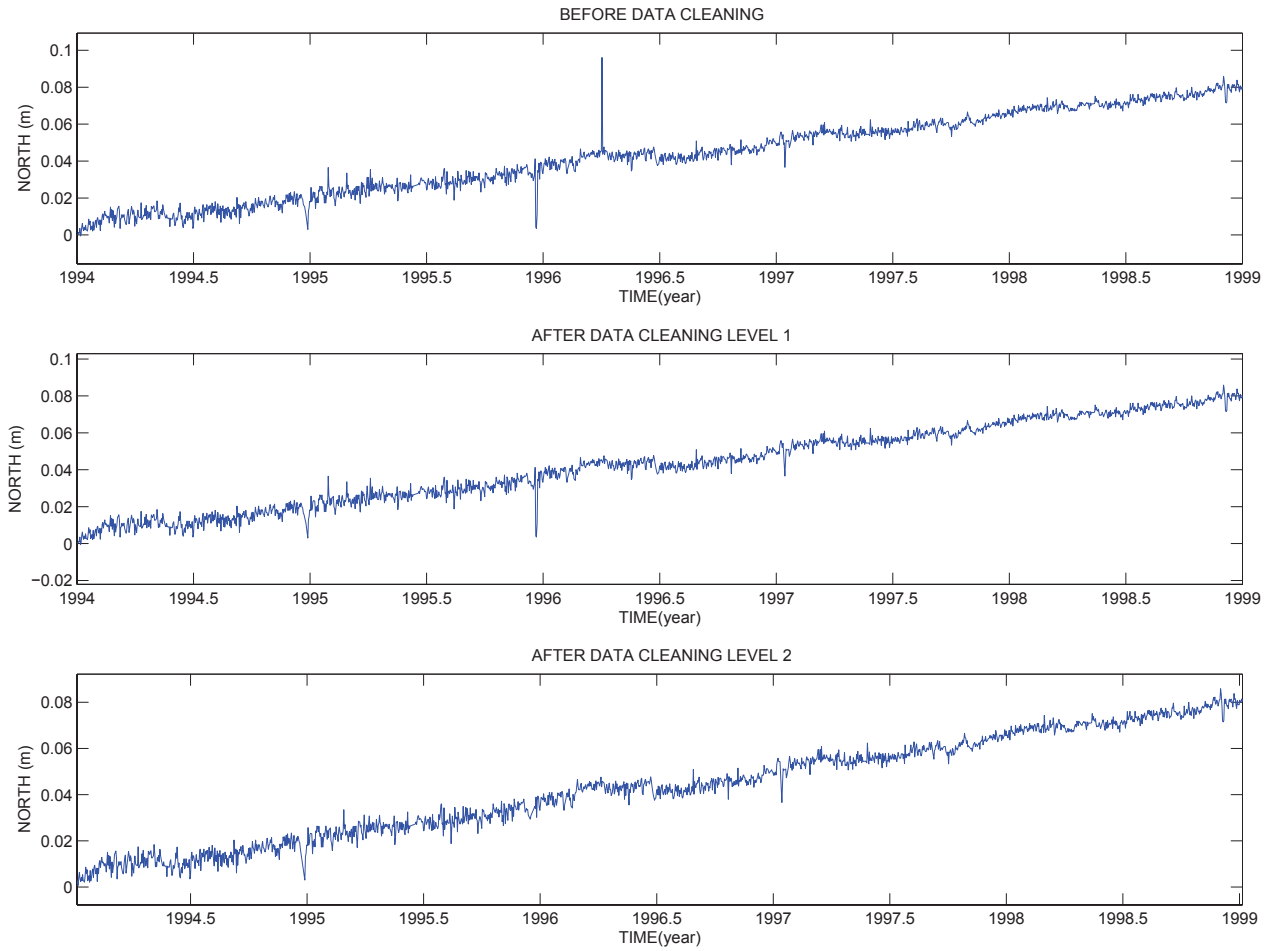


Figure 2.2: *Data Cleaning after level 1 (3IQR) and 2 (2IQR)*

Figure 2.2 is a comparison of the time series for site GRAZ in north direction before and after the data cleaning with different level of intensity.

As shown in the plot, data points that have significant differences with the median of the residuals within the one year window are considered outliers and are eliminated from the time series. These data points possibly indicate a temporary change such as a sudden movement of the geodetic momuments, which would soon be recovered. The amplitudes of the noise after the cleaning are, as expected, reduced.

However whether the eliminated data points are outlier or not is still not certain. To determine the influence of the data cleaning on the noise analysis, we can compare the results of maximum likelihood estimation on fixed noise models before and after the cleaning and treated the north, east, and vertical components separately.

With this edit criterion, the cleaning algorithm typically removed 0.1–0.4% of the data in level 1 and 1–3% of the data in level 2 from each global site. Table 2.1 shows the comparison between the number of remaining data points from different ACs and sites.

Site	AC	Data Points		
		Before	After Level 1	After Level 2
GLSV	CODE	5504	5502	5423
	ESA	4200	4185	4117
	GFZ	6532	6529	6499
	GRGS	6733	6728	6691
	MIT	6574	6568	6510
GRAZ	CODE	7146	7136	7074
	ESA	5896	5876	5820
	GFZ	6470	6454	6384
	GRGS	6312	6305	6254
	MIT	6998	6983	6926
MATE	CODE	7168	7117	7073
	ESA	6329	6277	6213
	GFZ	6713	6690	6579
	GRGS	6410	6386	6303
	MIT	7036	7001	6917
WTZZ	CODE	5307	5301	5261
	ESA	6787	6764	6711
	GFZ	6002	5983	5923
	GRGS	6332	6317	6290
	MIT	5176	5166	5130
ZIMM	CODE	7193	7181	7125
	ESA	2636	2631	2578
	GFZ	6958	6938	6865
	GRGS	6573	6565	6487
	MIT	7073	7054	6989

Table 2.1: *Number of data points before and after cleaning*

2.3 Maximum Likelihood Estimation

To estimate the noise component and measure the amount of white noise and power law noise in the time series, we used maximum likelihood estimation (MLE). The best-fit noise model to the data is that which maximizes the probability function

$$\mathrm{lik}(\hat{v}, C) = \frac{1}{(2\pi)^{N/2}(\mathrm{det}C)^{1/2}} \exp(-0.5 v^T C^{-1} v) \,, \tag{2.13}$$

where lik is likelihood and det is the determinant of a matrix, C is the covariance matrix representing the assumed noise in the data, N is the number of epochs and v are the postfit residuals to the linear function using weighted least squares with the same covariance matrix C.

For greater numerical stability, we maximize the logarithm of the likelihood:

$$\ln[\mathrm{lik}(\hat{v}, C)] = -0.5[\ln(\mathrm{det}C) + v^T C^{-1} v + N \ln(2\pi)] \,, \tag{2.14}$$

where ln is the natural logarithm, N is the number of epochs, C is the data covariance matrix, and v are the postfit residual from some linear model applied to the same covariance matrix C.

The covariance matrix C can represent many form of Gaussian stochastic noise such as white, power law, first-order Gauss Markov, band pass together with a multitude of combinations of the above. In this thesis we assume that C is the combination of two noise sources, a white noise component and a power law noise component.

Then the data covariance matrix can be written as:

$$C = a_w^2 I + b_\kappa^2 J_\kappa \,, \tag{2.15}$$

is composed of white noise and power-law noise with amplitudes a and b respectively, where I is the identity matrix and J_κ is the covariance matrix for the appropriate power-law noise.

2.4 Power-law Noise Covariance Matrix

For power-law noise with any spectral index $-3 < \kappa < 1$, the covariance matrix can be derived using a method described by Johnson and Wyatt (1994). Consider a transformation matrix T such that a sequence of random variables x with covariance matrix J_κ is created from a vector d of IID (independent and identically distributed random variables) unit-variance random variables by $x = Td$. The desired covarince matrix J_κ is, by propagation of errors, therefore:

$$J_\kappa = TC_dT^T \,. \tag{2.16}$$

Since the covariance matrix $C_d = I$ then

$$J_\kappa = TT^T \,, \tag{2.17}$$

where the transformation matrix T is described by Hosking (1981) using the method of fractional differencing/integrating. The transformation matrix is:

$$T = \Delta T^{-\kappa/4} \begin{bmatrix} \phi_0 & 0 & 0 & ... & 0 \\ \phi_1 & \phi_0 & 0 & ... & 0 \\ \phi_2 & \phi_1 & \phi_0 & ... & 0 \\ ... & ... & ... & ... & ... \\ \phi_N & \phi_{N-1} & \phi_{N-2} & ... & \phi_0 \end{bmatrix} \,, \tag{2.18}$$

with the sampling interval ΔT and with

$$\phi_N = \frac{-\frac{\kappa}{2}(1-\frac{\kappa}{2})...(N-1-\frac{\kappa}{2})}{N!} = \frac{\Gamma(N-\frac{\kappa}{2})}{N!\Gamma(-\frac{\kappa}{2})} \,. \tag{2.19}$$

The transformation matrix T is scaled by $\Delta T^{-\kappa/4}$, where ΔT is the sampling interval, before forming the covariance matrix by TT^T. This ensures that the power spectra for any noise source with spectral index κ will cross at the same frequency given the same sampling interval ΔT and equal noise amplitude b_κ.

The equation for the power spectrum is then:

$$P = \frac{D_\kappa b_\kappa^2}{f_s^{\frac{\kappa}{2}+1}} f^\kappa \,, \tag{2.20}$$

where

$$D_\kappa = 2(2\pi)^\kappa (24 \cdot 60 \cdot 60 \cdot 365.25)^{\frac{\kappa}{2}} \,, \tag{2.21}$$

and f_s is the sampling frequency in Hz. The crossover frequency is therefore

$$f_0 = \frac{f_s^{1/2}}{2\pi\sqrt{24 \cdot 60 \cdot 60 \cdot 365.25}} \,. \tag{2.22}$$

If the data are not equally spaced then each of column of the transformation matrix T is scaled by the individual ΔT, i.e., $\Delta T = \left|t_j - t_0\right|$, prior to forming the covariance matrix.

2.4.1 Covariance Matrix for Noise with Integer spectral indices

With the general equation for power-noise covariance matrix above, we can now discuss the noise with integer spectral indices separately.

2.4.1.1 White Noise

From the equation of the transformation matrix we can see that if $\kappa = 0$, then $\phi_0 = 1$ and $\phi_N = 0$ for $N > 0$.

Therefore $T = I$ and $J_0 = I$, which means the covariance matrix is scalar and independent of time.

2.4.1.2 Random-Walk Noise

If $\kappa = -2$, then $\phi_N = 1$ for any N

$$T = \begin{bmatrix} 1 & 0 & 0 & ... & 0 \\ 1 & 1 & 0 & ... & 0 \\ 1 & 1 & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 1 & 1 & 1 & ... & 1 \end{bmatrix} \,. \tag{2.23}$$

The covariance matrix is therefore equal to

$$J_{-2} = \begin{bmatrix} \Delta T_1 & \Delta T_1 & \Delta T_1 & ... & \Delta T_1 \\ \Delta T_1 & \Delta T_2 & \Delta T_2 & ... & \Delta T_2 \\ \Delta T_1 & \Delta T_2 & \Delta T_3 & ... & \Delta T_3 \\ ... & ... & ... & ... & ... \\ \Delta T_1 & \Delta T_2 & \Delta T_3 & ... & \Delta T_N \end{bmatrix} , \tag{2.24}$$

and is exactly the same as the matrices decribed in Zhang et al. (1997); Mao et al. (1999).

2.4.1.3 Flicker Noise

For flicker noise, the covariance matrix was approximated by Zhang et al. (1997). The constant in this matrix was selected so that the power spectrum of flicker noise and random-walk noise cross at a period of 1 year with a sampling interval of 1 day and equal amplitude ($b_{-1} = b_{-2}$). Now the amplitude of this earlier matrix still needs scaling to be exactly the same as that derived from the transformation matrix. The scaling can be estimated from their respective power-spectrum equations. That is

$$P_{\text{old}} = \frac{b_{\text{old}}^2 f^{-1}}{2\pi^2} , \tag{2.25}$$

$$P_{\text{new}} = \frac{b_{\text{new}}^2 f^{-1}}{\pi\sqrt{24 \cdot 60 \cdot 60 \cdot 365.25}} , \tag{2.26}$$

therefore

$$b_{\text{new}} = \frac{(f_s \cdot 24 \cdot 60 \cdot 60 \cdot 365.25)^{1/4}}{\sqrt{2\pi}} b_{\text{old}} . \tag{2.27}$$

When the sampling frequency is once per day (in Hz) then

$$b_{\text{new}} = 1.744 \cdot b_{\text{old}} . \tag{2.28}$$

Therefore values quoted for the amplitude of flicker noise covariance matrix must be scaled before being used.

There are many previous error analysis studies of GPS time series. Some of them only chose specific spectral indices because there were no general form for the covariance matrix of power-law noise at that time. The exact covariance matrix for random walk noise was given by Langbein and Johnson (1997), while an approximation for the flicker noise covariance matrix was given by Zhang et al. (1997) and derived from an algorithm described by Gardner (1978) for generating such noise. Another reason is that there are some advantages to assume a specific type of noise model for the time series. First of all, it is easy to analyse the noise if one parameter is fixed. Second, assuming all the sites are affected by the same noise source such as atmospheric loading, monument motion and reference frame makes the determination of noise component

easier. Therefore they should have a similar noise spectrum. Third, with long time series to analyse, the computational burden is usually huge. Limited numbers of fixed noise models would reduce the computational burden and make the MLE more efficient. If the purpose is to provide more robust uncertainties on the linear model parameters, then a model such as white noise plus flicker noise might not be exact, but it may be close enough.

However, there are reasons why the spectral index should also be estimated. While many studies have concentrated on the integer spectral indices as potential models there is no reason why a non-integer spectral index may not be appropriate.

First, now we can compute the power-law covariance matrix by means of fractional differencing and integration (described by Williams (2003)). Second, stochastic models can provide clues in the search for the physical phenomena that affects the time series. Third, with the development of computer capacity, the computational burden is now not such a problem. However, estimating the noise parameters or spectral index of a long time series with thousands of data points, for example, in this thesis, more than 7000, still takes several days.

To improve the efficiency of the analysis, the software CATS splits the maximization up into two parts, which will be explained in detail in chapter 3.

Chapter 3

Data Processing and Software

3.1 Data

3.1.1 Source

The data used for this thesis are downloaded from the data base CDDIS [1] (The Crustal Dynamics Data Information System). The CDDIS was established in 1982 as a dedicated data base to archive and distribute space geodesy related data sets and has served as a global data center for the IGS since 1992. Today, the CDDIS archives and distributes mainly Global Navigation Satellite Systems (GNSS), laser ranging (both to artificial satellites, SLR, and lunar, LLR), Very Long Baseline Interferometry (VLBI), and Doppler Orbitography and Radio-positioning Integrated by Satellite (DORIS) data for an ever increasing user community of geophysists.

For our purpose of analysing noise content of GPS time series we select SINEX file provided by ACs after the second data reprocessing campaign in the time range from GPS week 730 to 1773 (the year 1994 to 2014).

3.1.2 Download

For our research required SINEX files from each analysis center are listed under a specific catalogue ("repro2") and yet are separated from each week. To meet my requirement and increase the efficiency of the data acquisition, computer program Wget is used for the Download.

Wget is a computer program that retrieves content from web servers. It supports downloading via HTTP, HTTPS, and FTP protocols. Hier is the command line written for the download:

```
#!/bin/bash
DESTINATION=/Users/visagod/Desktop/DATA
SOURCE=ftp://cddis.nasa.gov/gps/products/

for i in {730..999}
do
        wget -P $DESTINATION -r --no-parent --no-directories -A
        "*snx*"  $SOURCE"0"$i"/repro2"
done
```

[1] ftp://cddis.nasa.gov/gps/products/

```
for i in {1000..1773}
do
        wget -P $\$$DESTINATION -r --no-parent --no-directories -A
        "*snx*"  $SOURC$i"/repro2"
done
```

3.1.3 SINEX File

To unify the form of initial GPS data and enable associated analysis of different data, IGS present a standard data format SINEX.

Each SINEX line has at most 80 ASCII characters. The SINEX file is subdivided in groups of data called blocks. Each block is enclosed by a header and trailer line. Each block has a fixed format. The blocks contain information on the file, its input, the sites and the solution. All elements within a line are defined.

The first character of each line identifies the type of information that the line contains. Five characters are reserved. They have the following meaning when they are at the beginning of a line, they identify:

```
"%" Header and trailer line,
"*" Comment line within the header and trailer line,
"+" Title at the start of a block
"-" Title at the end of a block
" " Data line within a block
```

The following tables 3.1 list the contents of each main block included in one SINEX File.

FILE	FILE/REFERENCE	This block provides information on the Organization, point of contact, the software and hardware involved in the creation of the file.
	FILE/COMMENT	This block can be used to provide general comments about the SINEX data file.
INPUT	INPUT/HISTORY	This block provides information about the source of the information used to create the current SINEX file.
	INPUT/FILES	This block identify the input files (and the current SINEX file) and allow for a short comment to be added to describe those files.
	INPUT/ACKNOWLEDGEMENTS	This block defines the agency codes contributing to the SINEX file.
SITE	SITE/ID	This block provides general information for each site containing estimated parameters.

	SITE/DATA	This block gives the relationship between the estimated station parameters in the SINEX file and in the input files.
	SITE/RECIEVER	List the receiver used at each site during the observation period of interest.
	SITE/ANTENNA	List of antennas used at each site used in the SINEX file.
	SITE/GPS_PHASE_CENTER (Mandatory for GPS)	List of GPS phase center offsets for the antennas described in the Antenna block. The offset is given from the Antenna Reference Point (ARP) to the L1 and L2 phase centers respectively.
	SITE/ECCENTRICITY (Mandatory)	List of antenna eccentricities from the Marker to the Antenna Reference Point (ARP) or to the intersection of axis.
SATELLITE	SATELLITE/ID (Recommended for GNSS, if available)	List of GNSS satellites used in the SINEX file
	SATELLITE/PHASE_CENTER (Mandatory for GNSS, if satellite antenna offsets are not estimated)	List of GNSS satellite antenna phase center corrections. The antenna offsets are given from the center of mass (CM). More than one line per satellite is necessary, if the satellite transmits on more than two frequencies.
SOLUTION	SOLUTION/EPOCHS (Mandatory)	List of solution epoch for each Site Code/Point Code/Solution Number/Observation Code (SPNO) combination.
	SOLUTION/STATISTICS	Statistical information about the solution contained in the SINEX file.
	SOLUTION/ESTIMATE (Mandatory)	Estimated parameters
	SOLUTION/APRIORI (Mandatory)	Apriori information for estimated parameters, either the used apriori values for the adjustment or the parameters of a Helmert Transformation for applied inner constraints with the constraint given in the field 'Standard Deviation'.
	SOLUTION/MATRIX_ESTIMATE (Mandatory)	The matrix contents can be: CORR - Correlation Matrix, COVA - Covariance Matrix, INFO - Information Matrix (of Normals)
	SOLUTION/MATRIX_APRIORI	The matrix contents can be: CORR - Correlation Matrix, COVA - Covariance Matrix, INFO - Information Matrix (of Normals)

	SOLUTION/NORMAL_EQUATI -ON_VECTOR	If the SINEX file shall provide the normal equation directly this block is mandatory and contains the vector of the right hand side of the unconstrained (reduced) normal equation.

Table 3.1: SINEX *File content*

3.1.4 Data Transformation

As we use software CATS to analyse the time series, all the necessary information need to be collected from the SINEX file and rearranged into a new format (`.neu`) that can be processed by the software.

For the transformation into NEU file, the most important information we get from a SINEX File is the SOLUTION/ESTIMATION block, including data showed in table 3.2.

Data	Information	Unit
STAX	station X coordinate	m
STAY	station Y coordinate	m
STAZ	station Z coordinate	m
VELX	station X velocity	m/year
VELY	station Y velocity	m/year
VELZ	station Z velocity	m/year
XGC	geocenter X coordinate	m
YGC	geocenter Y coordinate	m
ZGC	geocenter Z coordinate	m
RS_RA	radio source right ascension	rad
RS_DE	radio source decline	rad
RS_RAR	radio source right ascension rate	rad/year
RS_DER	radio source declination rate	rad/year

Table 3.2: SOLUTION/ESTIMATION *Block*

For each interested site one CATS file will be created containing the information:

- site name
- reference position:

  initial cartesian coordinates (X, Y, Z)

  initial latitudes, longitudes and heights
- information of the site in each available epoch including:

  Time of the epoch (years+decimal)

  coordinate differences in local geodetic system in north direction

  coordinate differences in local geodetic system in east direction

coordinate differences in local geodetic system in height

standard deviation in north direction

standard deviation in east direction

standard deviation in height

The positions are assumed to be in meters. The individual formal errors are only used in the MLE estimation to define a variable white noise component.

To enable the described transformation a package of program is required with the features of:

- reading SINEX files and computes NEU coordinates and standard deviation of a site with regard to initial position and writes to CATS time series file.(`snx2cats`)
- reading XYZ coordinates and covariance matrix from SINEX file(`rdsnx`)
- Computes reference ellipsoid parameters(`refell`)
- form rotation matrix to convert from cartesian system to local geodetic system (`rotct2lg`)
- transformation between cartesian system and local geodetic system (`cct2clg, ct2lg`)
- transformation between cartesian system and ellipsoidal system (`xyz2ell`)
- transformation between julian date and calendar date (`cal2jd, doy2jd, jd2cal`)
- transformation between julian date and year with decimal of year (`jd2yr`)

The functions described above are written by Mike Craymer in Matlab code and can be downloaded as geodetic toolbox from his homepage.[2]

3.2 Software CATS

CATS (Create and Analyse Time Series) is a standalone C program developed by Williams (2008) to study and compare stochastic noise processes in continuous GPS coordinate time series and assign realistic uncertainties to parameters derived from them. It uses maximum likelihood estimation to fit a multi-parameter model to a time series while analysing the residuals to determine the type and magnitude of the noise.

The software uses the method of least squares to fit a multi-parameter model to a time series while concurrently analysing the residuals to assess the type and magnitude of the stochastic noise. The program solves for all parameters in a two-part procedure:

- the linear function which includes an intercept, a slope, the possibility of abrupt steps (earthquake or equipment changes) and any known periodicities (annual and semi-annual terms)
- the non-linear part which allows the parameters and amplitudes of several specific noise models to be estimated.

[2] `http://www.craymer.com`

There are currently three methods implemented in CATS for assessing the noise content: maximaun likelihood estimation, sepctral estimation and an empirical method based on the work described in Williams (2003). The principle method, and the one we used in this thesis is MLE.

3.2.1 Nested Maximization

To estimate the noise components and the parameters from the linear function the likelihood that these values have occured for a given set of observation is maximized. As described in equation 2.13.

Since the maximum is unaffected by monotonic transformations. The algorithm of choice for solving the maximum likelihood problem is the Nelder-Mead uphill simplex (Williams (2008)). It is a straightforward method to program, but the drawback is its lack of efficiency in the number of function evaluations used. The algorithms used in CATS have been optimized in two broad categories to improve the computational speed: a single dimension reduction in the estimation of amplitudes and nested maximization algorithms.

Very often one may wish to estimate parameters of the stochastic model other than just the noise amplitude: for example, estimating the spectral index κ in power-law noise models. These parameters could simply be added as extra dimensions in a general uphill simplex together with the set of noise amplitudes. However at most iterations the uphill simplex may choose a new value for one of these parameters and therefore a new unit covariance matrix has to be created. It appears better to split the maximization up into two parts:

- inner maximization: simply estimate the noise amplitudes for fixed noise model parameters
- outer maximization: estimate the other noise model parameters

This results in far fewer calls to the routines dealing with the covariance matrices (creation, inversion and determinant estimation) and is likely to be more stable numerically.

In reality the desired covariance matrix can be any combination of different stochastic models, that is:

$$C = \sum_{i=1}^{m} \sigma_i^2 J_i \,, \tag{3.1}$$

where there are m different unit covariance matrices. For a fixed set of J_i the MLE is often thought to be an m dimensional problem. However, it can be shown to be $m-1$ dimensional. If the covariance matrix consists of one noise source then we have:

$$C = \sigma^2 J \,, \tag{3.2}$$

and the MLE can be re-written as:

$$\ln[\mathrm{lik}(\hat{v},\sigma)] = -0.5[2N\ln(\sigma) + \ln(\det J) + \frac{v^T J^{-1} v}{\sigma^2} + N\ln(2\pi)] \,. \tag{3.3}$$

Now, with the method of least squares we find that the residuals v and the estimated parameters are invariant to a scale change in the covariance matrix. Therefore we can differentiate the log likelihood equation with respect to σ and find the value of σ that gives us the maximum likelihood is:

$$\sigma = \sqrt{\frac{v^T J^{-1} v}{N}} \,. \tag{3.4}$$

So in this case the maximum log-likelihood can be found explicitly. By association then a two-dimensional covariance matrix should only need a one-dimensional numerical maximization. Instead of two noise amplitudes σ_1 and σ_2 we can transform these to two alternative variable namely an angle, ϕ and a scalar r such that:

$$\sigma_1 = r\cos\phi \,, \sigma_2 = r\sin\phi \,. \tag{3.5}$$

For a given angle we can compute a unit covariance matrix containing the correct ratios of the two, that is:

$$C_{\text{unit}} = r^2(\cos^2(\phi)J_1 + \sin^2(\phi)J_2) \,, \tag{3.6}$$

which is similar in form to the one-dimensional covariance matrix. This reconfiguration can be extended to any number of dimensions. For a given angle, ϕ we can explicitly solve for r and therefore only need to maximize the log-likelihood numerically as a function of the angle, thereby reducing dimensionality by one. Typically, this results in far fewer steps than the equivalent m-dimensional maximization.

Figure 3.1 is an example from Williams (2008) to illustrate the nested maximization.

3.2.2 The Time Series File

The time series file is fairly free format consisting of two parts: header information and the time series. The header information is basically a list of parameters relevant to the GPS site the time series comes from. The header information (plus any line you want to eliminate from the processing) starts with a # symbol. The only header information relevant to CATS is the list of offsets. The data set consists of seven columns (no particular format) corresponding to time (in decimal years), north, east, up, north error, east error and up error. The positions are, by default, assumed to be in meters, however this can be altered with the scale factor option.

If there are any known discontinuities or offsets in the time series then these are specified in the header of the file in the following format:

```
# offset  decimal_date  component_code
```

This is an example of a typical CATS time series file (`.neu`) of the site GLSV:

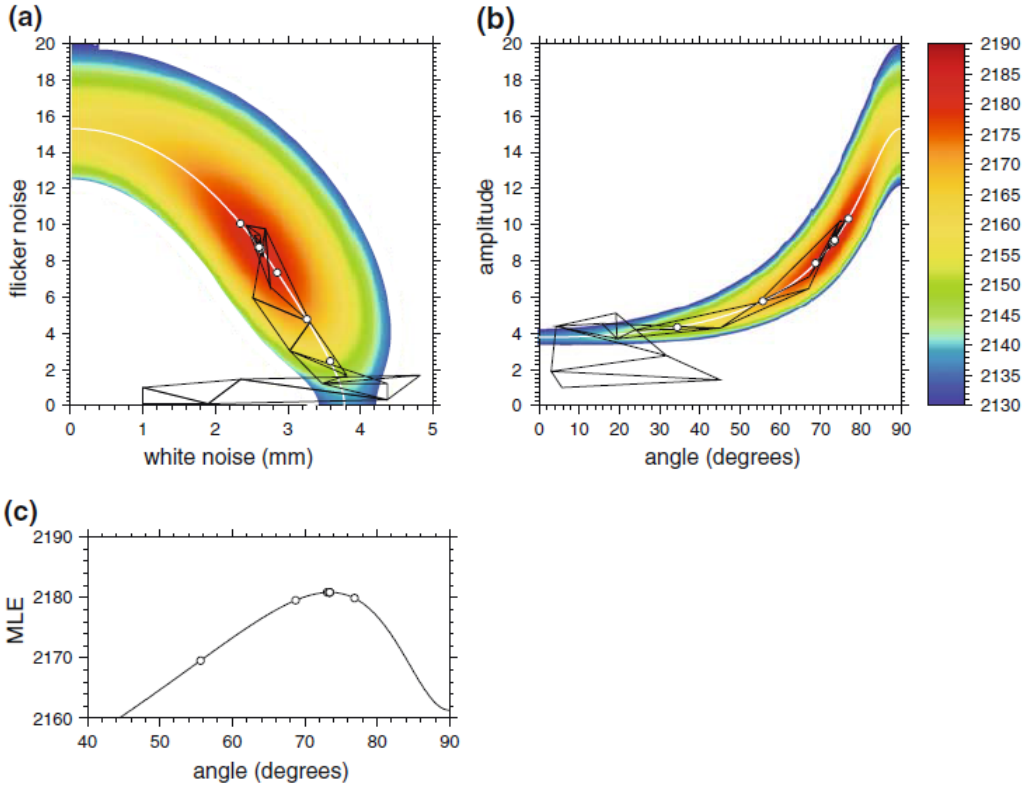


Figure 3.1: *Log-likelihood estimates for the north component of* GPS *site* BRST. ***(a)*** *Log-likelihood values greater than 2130 plotted in amplitude space, that is, as a function of flicker noise and white noise amplitude. Black triangles indicate the path of the uphill simplex from $\sigma_1 = 1$; $\sigma_2 = 1$. White dots are the predictions from the 1D brent maximizations. White line is the exact maximum at all possible angles.* ***(b)*** *as a except in ϕ and r space.* ***(c)*** *The results of the 1D maximization which calculates the maximum log-likelihood values as a function of ϕ. (Courtesy: Williams (2008))*

```
# Site: GLSV
# Reference Position:
# X (m): 3512889.1006
# Y (m): 2068979.7810
# Z (m): 4888903.1601
# Latitude  (deg):  50.3641820
# Longitude (deg):  30.4967301
# Height      (m):    226.3294
# Reference Frame:
# offset
1998.1575    0.0000    0.0000     0.0000   0.0016   0.0011   0.0016
1998.2178    0.0026    0.0037    -0.0147   0.0008   0.0006   0.0008
1998.2205    0.0046    0.0032    -0.0198   0.0008   0.0006   0.0008
1998.2233    0.0049    0.0044    -0.0212   0.0008   0.0006   0.0008
1998.2260    0.0052    0.0010    -0.0177   0.0008   0.0006   0.0008
```

```
1998.2288    0.0027    0.0028    -0.0180   0.0008   0.0006   0.0009
1998.2315    0.0039    0.0010    -0.0201   0.0008   0.0006   0.0008
1998.2342    0.0007    0.0047    -0.0197   0.0008   0.0007   0.0009
1998.2370    0.0041    0.0047    -0.0153   0.0008   0.0006   0.0008
1998.2425    0.0033    0.0043    -0.0144   0.0008   0.0007   0.0009
1998.2863    0.0042    0.0064    -0.0087   0.0008   0.0006   0.0009
1998.2890    0.0033    0.0072    -0.0059   0.0008   0.0006   0.0009
1998.2918    0.0056    0.0062    -0.0061   0.0009   0.0007   0.0009
1998.2945    0.0055    0.0070    -0.0008   0.0008   0.0006   0.0009
1998.2973    0.0063    0.0075    -0.0027   0.0008   0.0006   0.0008
1998.3000    0.0054    0.0065    -0.0053   0.0008   0.0006   0.0008
1998.3027    0.0077    0.0096    -0.0092   0.0008   0.0006   0.0008
1998.3055    0.0064    0.0037    -0.0139   0.0008   0.0006   0.0008
1998.3082    0.0049    0.0040    -0.0127   0.0008   0.0006   0.0009
1998.3110    0.0062    0.0054    -0.0156   0.0008   0.0007   0.0009
1998.3137    0.0062    0.0068    -0.0152   0.0009   0.0007   0.0009
1998.3164    0.0077    0.0063    -0.0078   0.0009   0.0007   0.0009
1998.3192    0.0072    0.0074    -0.0088   0.0008   0.0006   0.0008
1998.3219    0.0073    0.0073    -0.0170   0.0008   0.0007   0.0009
1998.3247    0.0053    0.0067    -0.0067   0.0011   0.0008   0.0012
1998.3274    0.0072    0.0082    -0.0119   0.0008   0.0006   0.0008
```

3.2.3 Command Line Options

The program is command line oriented, so it does not require any other files except for the input time series file in order to run. The options can be divided into three categories:

- Stochastic model:

  The stochastic model is defined using the model option and can be used as many times as required on the command line.

  For example:

  ```
  --model pl:k-1 --model wh:
  ```

  is the option for a flicker noise plus white noise model, where `pl:` stands for power law noise and `wh:` stands for white noise. The second part of the power law noise option `k-1` sets the spectral index to -1 (flicker noise). If the spectral index is not specified, then the program will solve for this parameter. Currently CATS can produce a covariance matrix for the following noise models:

  - White noise <`wh`>
  - Power-law noise <`pl`>
  - First-order Gauss Markov noise <`gm`>

  - Band-pass noise <bp>
  - Generalized Gauss Markov <gg>
  - Variable white noise <vw>
  - Step-variable white noise <sw>

- Functional model:

  The functional model is already predefined as having a slope, an intercept and any offsets that are found in the time series file. The sinusoid switch can be used if you wish to solve for the amplitude and phase of any periodic signals. If you do not require the program to estimate a trend in the data, then this option can be switched off using notrend.

- Other options:

  `Method` can define which estimation method to use.

  `Verbose` can output extra information.

  `Output` forces the results to be printed to a file instead of the screen.

To estimate the amount of flicker noise and white noise in the file `vyas.neu` assuming there is an annual signal in the series we can use the command:

```
cats vyas.neu --sinusoid 1y --model pl:k-1 --verbose
--output vyas.fn_mle
```

To estimate the spectral index plus the amplitudes of white and power-law noise in the east and up component of file `vyas.neu` assuming an annual and semi-annual signal and using the fast option we can use the command:

```
cats vyas.neu --sinusoid 1y1 --model pl: --column 3 --verbose
--speed 2 --output vyas.pl_mle
```

3.2.4 Typical CATS Output

Output from the software is either to the screen or to a file named using the output option. The first block of lines provide a brief summary of the command issued, the software version and some information on the user and the machine the program was excited on.

If the `verbose` option is included, the second block provides a summary of the data set used. The probable sampling frequency of the time series is usually 1 day. This is required because the Power-Law and First-Order Gauss Markov covariance matrices are scaled by the sampling interval.

The lines containing actual results are preceded with a "+" symbol, followed by `NORT, EAST, VERT` or `PSMSL` and are generally self-evident. `MLE` is the calculated, maximum log likelihood value, `INTER` is the intercept at t_0 which defined as the start of the first year of data. Quoted uncertainties

This is a example of a typical CATS output of the site ZIMM with the data from Analysis Center GFZ:

```
 Cats Version : 3.1.2
 Cats command : cats --model pl:k-1 --model wh: --sinusoid 1y
 --verbose --output ZIMM_es.wh_fn_mle ZIMM_es.neu
 Data from file : ZIMM_es.neu
 cats : running on GanyideMacBook-Air.local
 Darwin release 13.4.0 (version Darwin Kernel Version 13.4.0: Sun
 Aug 17 19:50:11 PDT 2014; root:xnu-2422.115.4~1/RELEASE_X86_64)
 on x86_64
 userid : visagod

 Sampling frequency 1.17366e-05 (Hz), 0.99 days
 Number of samples 1 period apart = 2556 of 2636
 Number of points in full series  = 2922
 Number of series to process : 3
 Start Time : Thu Jan 29 13:16:25 2015

+NORT  POWER LAW NOISE
+NORT  INDEX  :       -1.0000
+NORT  PL     :       IS FREE

+NORT  WHITE NOISE
+NORT  WH     :       IS FREE

 Time taken to create covariance matrix : 2.027801 seconds
 Time taken to compute eigen value and vectors : 37.4681 seconds
 wh_only = 4.39569426 (-7646.1503),
 cn_only = 14.48962722 (-6907.3018)
 Starting a one-dimensional minimisation : initial angle 45.00
 Angle = 45.000000 mle = -7101.93384974 radius =     4.836850
 wh =     3.420169 cn =     3.420169
 Next choice of angle = 27.811530
 angle = 27.811530 mle = -6938.21084593 radius =     6.375862
 wh =     2.974752 cn =     5.639368
 Next choice of angle = 17.188471
 angle = 17.188471 mle = -6864.78197440 radius =     8.500878
 wh =     2.512144 cn =     8.121211
 Next choice of angle = 10.623060
 angle = 10.623060 mle = -6853.97157537 radius =    10.763168
 wh =     1.984157 cn =    10.578700
 Next choice of angle = 11.218347
 angle = 11.218347 mle = -6853.06847674 radius =    10.519784
 wh =     2.046608 cn =    10.318782
```

```
 Next choice of angle = 12.352145
 angle = 12.352145 mle = -6852.52719981 radius =    10.077179
 wh =     2.155705 cn =     9.843905
 Next choice of angle = 12.182229
 angle = 12.182229 mle = -6852.51300093 radius =    10.141716
 wh =     2.140121 cn =     9.913339
 Next choice of angle = 12.195395
 angle = 12.195395 mle = -6852.51292788 radius =    10.136693
 wh =     2.141338 cn =     9.907937
 Next choice of angle = 12.193692
 angle = 12.193692 mle = -6852.51292619 radius =    10.137343
 wh =     2.141181 cn =     9.908636
 Next choice of angle = 12.193570
 angle = 12.193570 mle = -6852.51292619 radius =    10.137389
 wh =     2.141170 cn =     9.908686
 Next choice of angle = 12.193814
 angle = 12.193814 mle = -6852.51292620 radius =    10.137296
 wh =     2.141192 cn =     9.908585

 Finished :

 Angle = 45.000000 mle = -7101.93384974 wh = 3.420169 cn = 3.420169
 Angle = 27.811530 mle = -6938.21084593 wh = 2.974752 cn = 5.639368
 Angle = 17.188471 mle = -6864.78197440 wh = 2.512144 cn = 8.121211
 Angle = 10.623060 mle = -6853.97157537 wh = 1.984157 cn = 10.57870
 Angle = 11.218347 mle = -6853.06847674 wh = 2.046608 cn = 10.31878
 Angle = 12.352145 mle = -6852.52719981 wh = 2.155705 cn = 9.843905
 Angle = 12.182229 mle = -6852.51300093 wh = 2.140121 cn = 9.913339
 Angle = 12.195395 mle = -6852.51292788 wh = 2.141338 cn = 9.907937
 Angle = 12.193692 mle = -6852.51292619 wh = 2.141181 cn = 9.908636
 Angle = 12.193570 mle = -6852.51292619 wh = 2.141170 cn = 9.908686
 Angle = 12.193814 mle = -6852.51292620 wh = 2.141192 cn = 9.908585
+NORT  MLE    :    -6852.512926
+NORT  INTER  :        4.4748 +-      1.4017
+NORT  SLOPE  :       15.4407 +-      0.3986
+NORT  SIN    :        2.3377 +-      0.4632
+NORT  COS    :        0.3258 +-      0.4516
+NORT  POWER LAW NOISE
+NORT  INDEX  :       -1.0000
+NORT  PL     :        9.9086 +-      0.4075

+NORT  WHITE NOISE
+NORT  WH     :        2.1412 +-      0.0914

End Time : Thu Jan 29 13:20:13 2015
```

```
Total Time : 228
```

For the example shown above, the amplitudes of a white plus flicker noise model are estimated. Since this is a one-dimensional maximization problem and a special class of model where one of the components is white, the software uses the Brent maximization instead of the uphill simplex. The steps of this algorithm are seen above the results as a series of angles and next choice of angles. The software also calculates the two end members of the noise spectrum: white noise only and colored noise only to ensure the correct answer should the angle get close to 0 or 90.

Chapter 4

Results

In this chapter we will discuss the estimated noise parameters under all circumstances.

During the research the most significant problem is the processing speed. The time taken to process a time series depends on the length of time series and the model you are running. For example, the white noise only model is almost instantaneous whereas a stochastic model such as power-law noise plus white noise where the spectral index is also being estimated takes the longest. Although the computational speed of modern computer has been greatly improved, the great number of data points (20 years) we used in this thesis still resulted in a prolonged processing time. For example, the estimation of a noise model with non-integer spectral index for a single site can, in some cases, take $4-5$ days. That is also the reason why It took much longer time than expected to finish this thesis. However, the number of sites and ACs we researched in this period is still very limited.

In order to have a relatively good understanding of the noise components of a GPS time series and, to some extent, have a comparison between different analysis centers and with previous studies, the Europe region with a high GPS station density was chosen as our research field. Different sites were selected all around Europe for different researches. For specific purposes, GPS time series from some stations outside Europe were also used. Time series of these sites from different analysis centers were analysed with the software CATS. Table 4.1 lists the selected sites for our noise analysis and Figure 4.1 is a map of IGS tracking network of the Europe region. Some of the sites observed in this thesis are marked.

Site	Longitude	Latitude	Location
METS	24°23'43.2"	60°13'02.9"	Kirkkonummi, Finland
MOBJ	36°34'11.0"	55°06'53.6"	Obninsk, Russian Feder
POTS	13°03'57.9"	52°22'45.5"	Potsdam, Germany
GLSV	30°29'48.2"	50°21'51.1"	Kiev, Ukraine
POLV	34°32'34.6"	49°36'09.4"	Poltava, Ukraine
BRST	355°30'12.3"	48°22'49.8"	Brest, France
WTZZ	12°52'44.1"	49°08'39.2"	Bad Koetzting, Germany
GANP	20°19'22.6"	49°02'05.0"	Ganovce, Slovakia
ZIMM	7°27'55.0"	46°52'37.6"	Zimmerwald, Switzerland
GRAZ	15°29'36.5"	47°04'01.7"	Graz, Austria
BUCU	26°07'32.7"	44°27'50.2"	Bucuresti, Roma
ZECK	41°33'54.2"	43°47'18.2"	Zelenchukskaya, Russia
MADR	355°45'01.2"	40°25'45.0"	Madrid, Spain
MATE	16°42'16.1"	40°38'56.9"	Matera, Italy
NICO	33°23'47.2"	35°08'27.6"	Nicosia, Cyprus

NYAL	11°51'54.3"	78°55'46.5"	Ny Alesund, Norway
TIXI	128°51'59.1"	71°38'04.1"	Tixi, Russian
YELL	245°31'09.5"	62°28'51.2"	Yellowknife, Canada
IRKT	104°18'58.5"	52°13'08.5"	Irkutsk, Russia
WUHN	114°21'26.1"	30°31'53.9"	Wuhan, P.R. China
IISC	77°34'13.4"	13°01'16.2"	Bangalore, India
BOGT	285°55'08.6"	4°38'24.3"	Bogota, Colombia
NTUS	103°40'47.8"	1°20'44.9"	Singapore, Republic of Singapore
KARR	117°05'49.9"	-20°58'53.1"	Karratha, Australia
PERT	115°53'06.9"	-31°48'07.1"	Perth, Australia
KERG	70°15'19.9"	-49°21'05.3"	Port aux Francais, Kerguelen Island
MAC1	158°56'09.0"	-54°29'58.3"	MacQuarie Island, Sub-Antarctic
MAW1	62°52'14.6"	-67°36'17.2"	Mawson, Antarctica
MCM4	166°40'09.6"	-77°50'18.1"	Ross Island, Antarctic

Table 4.1: *Selected sites for noise analysis*

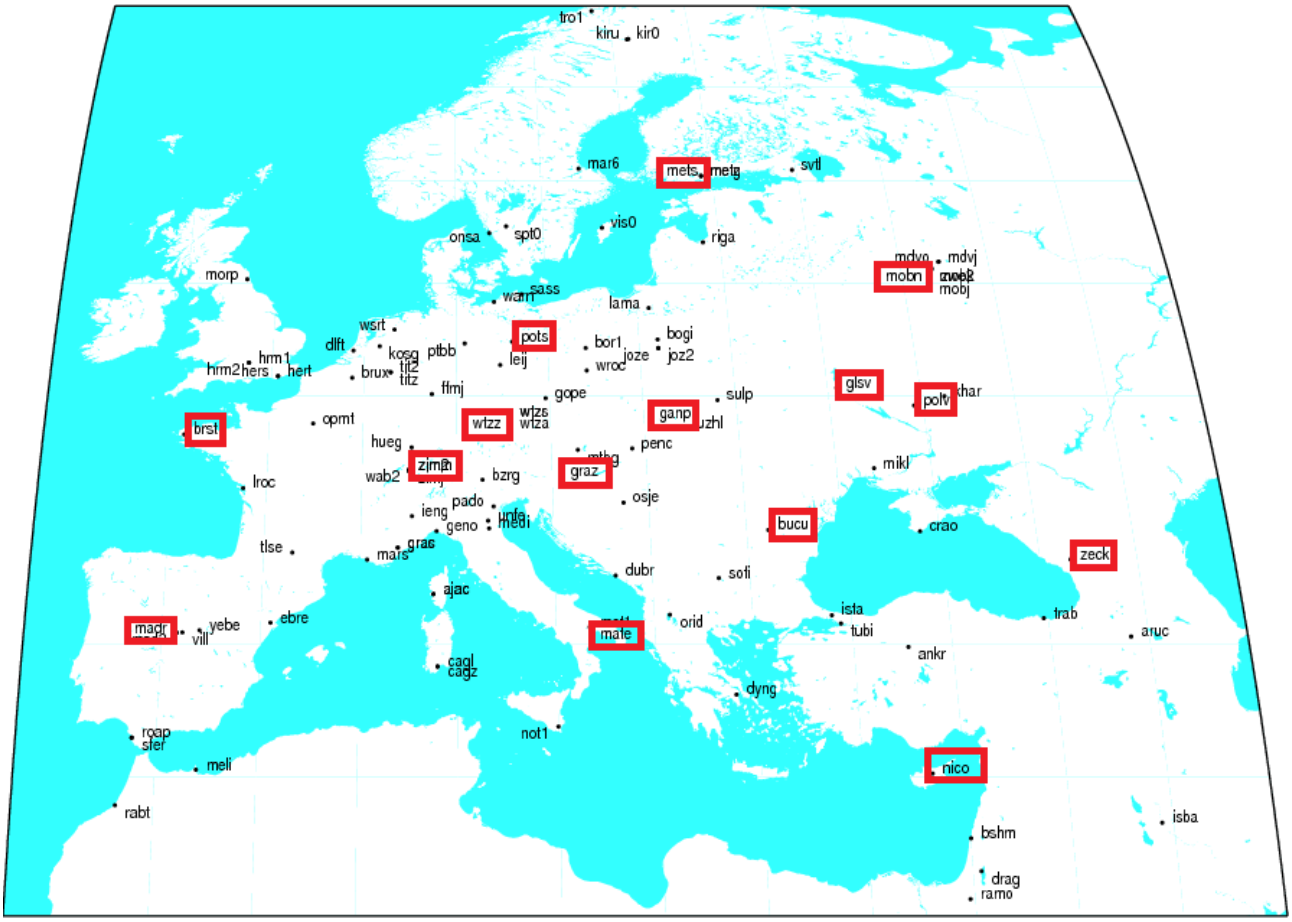


Figure 4.1: IGS *Tracking Network Europe Region*

4.1 Noise Models with Integer Spectral Index

Many previous error analysis studies of GPS time series only chose specific spectral indices because there were no general form for the covariance matrix of power law noise at that time. Fixed spectral index for power-law noise are also used to simplify the analysis and determination of noise components. Another reason is that with long time series to analyze, The computational burden is usually larger. limited numbers of fixed noise models would reduce the computational burden and make the MLE more efficient.

In this section we chose specific spectral indices for our noise models. Three noise models (white noise only, white noise plus flicker noise and white noise plus random walk noise) are tested on each time series to determine which one is preferred. Noise amplitudes of each component are also estimated with regard to different level of data cleaning.

4.1.1 Preferred Noise Model

As already described in chapter 2, the covariance matrix of a combination of white noise and power-law noise with amplitudes a and b respectively can be written as:

$$C = a_w^2 I + b_\kappa^2 J_\kappa \tag{4.1}$$

Three different types of noise models were evaluated with the MLE algorithm: white noise + flicker noise ($\kappa = -1$), white noise + random walk noise ($\kappa = -2$), and white noise only ($b_\kappa = 0$).

we treated North, East and Up separately as three different time series. There are all together over 200 selected time series from 4 analysis centers estimated. Every time series is tested 3 times to determine which noise model is preferred. The value of maximum likelihood estimation for different noise models are compared. The higher the log likelihood, the more significant and likely the model is as a candidate for the noise in the time series. The results are showed in table 4.2, 4.3, 4.4, A.1, A.2, A.3, A.4, A.5 and A.6 to demonstrate the change before and after the data cleaning.

As an example, we extracted the results of the north components of all the GPS time series, and compared them between different noise models in table 4.2, 4.3 and 4.4. The noise models with MLE values in bold are considered to be the preferred model.

We counted the time series with different preferred noise models before and after the data cleaning. The columns of table 4.5 represent the percentage of time series for which the log likelihood value for a specific model is the highest. The result shows that all of the time series prefer a white noise plus power-law noise model than a white noise only model. The following bar diagram compares the result of MLE between flicker plus white noise and random walk plus white noise model.

Site	AC	MLE Value (North)		
		White only	White + Flicker	White + Random Walk
GLSV	MIT	-13954.67	**-10027.97**	-10165.46
	ESA	-11563.53	**-8644.69**	-8688.13
GRAZ	MIT	-16160.77	**-14539.83**	-14706.88
	CODE	-14619.74	-9824.05	**-9621.29**
	ESA	-15883.14	**-13464.14**	-13523.58
	GFZ	-14888.38	**-13500.19**	-13650.08
MADR	MIT	-12057.43	**-9845.36**	-10023.76
	GFZ	-13012.54	**-8976.23**	-9945.74
MATE	MIT	-15799.23	**-13686.42**	-13781.36
	CODE	-14984.85	**-9894.9**	-9911.41
	ESA	-17001.92	-18291.78	**-14673.55**
	GFZ	-15606.28	**-13935.1**	-13977.07
POLV	MIT	-15234.12	**-13699.54**	-13878.23
	CODE	-15641.66	**-10054.32**	-10149.79
	GFZ	-16874.12	**-12315.38**	-13214.54
WTZZ	MIT	-9408.8	**-7937.01**	-8029.23
	CODE	-8577.51	-4819.57	**-4455.13**
ZECK	MIT	-16397.21	**-13487.48**	-13654.21
	ESA	-15621.35	**-9654.23**	-9987.65
	GFZ	-15421.02	**-10022.16**	-11238.32
ZIMM	MIT	-15089.67	**-14440.3**	-14593.44
	CODE	-13265.6	-10122.26	**-10038.79**
	ESA	-7646.15	**-6852.51**	-6882.44
	GFZ	-15002.81	-18822.80	**-14188.32**

Table 4.2: MLE *results of North components before data cleaning*

Best-fit noise model	Before	After Level 1	After Level 2
white noise only	0%	0%	0%
white + flicker noise	77.78%	59.72%	56.94%
white + random walk	22.22%	40.28%	43.06%

Table 4.5: *Percentage of time series with preferred noise model*

In figure 4.2 we can see that, in general, white noise plus flicker noise or white noise plus random walk noise is preferred to a white noise only model. Before the cleaning white noise plus flicker noise is clearly the dominant preferred noise model (77.78%). However with more and more "outliers" eliminated, the number of time series that are closer to a white noise plus random walk noise model increased. After level 1, the percentage of white noise plus flicker noise and white noise plus random walk noise are 59.72% and 40.28%. After level 2, numbers of time series that prefer these two noise models are very close.

Site	AC	MLE Value (North)		
		White only	White + Flicker	White + Random Walk
GLSV	MIT	-1385.05	**-12897.82**	-16286.18
	ESA	-11492.38	**-8493.26**	-8531.13
GRAZ	MIT	-15693.94	**-13890.29**	-14062.15
	CODE	-14562.21	**-9580.93**	-9792.12
	ESA	-15773.54	**-13316.92**	-13357.51
	GFZ	-14443.53	-17438.14	**-13062.22**
MADR	MIT	-13241.54	**-11246.17**	-12341.49
	GFZ	-15210.64	-9857.65	**-9612.74**
MATE	MIT	-15632.73	**-13578.72**	-19816.01
	CODE	-14878.34	**-9827.79**	-19647.45
	ESA	-16840.25	-18140.78	**-14481.02**
	GFZ	-15453.98	-19188.84	**-13819.65**
POLV	MIT	-16151.72	**-12414.39**	-12841.78
	CODE	-14512.14	-11234.57	**-11149.81**
	GFZ	-15471.01	-11235.85	**-10147.32**
WTZZ	MIT	-11746.79	**-10768.77**	-10898.28
	CODE	-10723.34	-7787.19	**-7666.12**
ZECK	MIT	-13214.81	**-9865.34**	-10041.18
	ESA	-12154.12	-11214.94	**-10175.54**
	GFZ	-11474.15	-9898.21	**-9547.24**
ZIMM	MIT	-14617.22	**-13848.77**	-14008.95
	CODE	-13173.96	-9973.45	**-9849.13**
	ESA	-7627.11	**-6831.7**	-6861.15
	GFZ	-14558.3	-18609.63	**-13545.17**

Table 4.3: MLE *results of North components after data cleaning level 1*

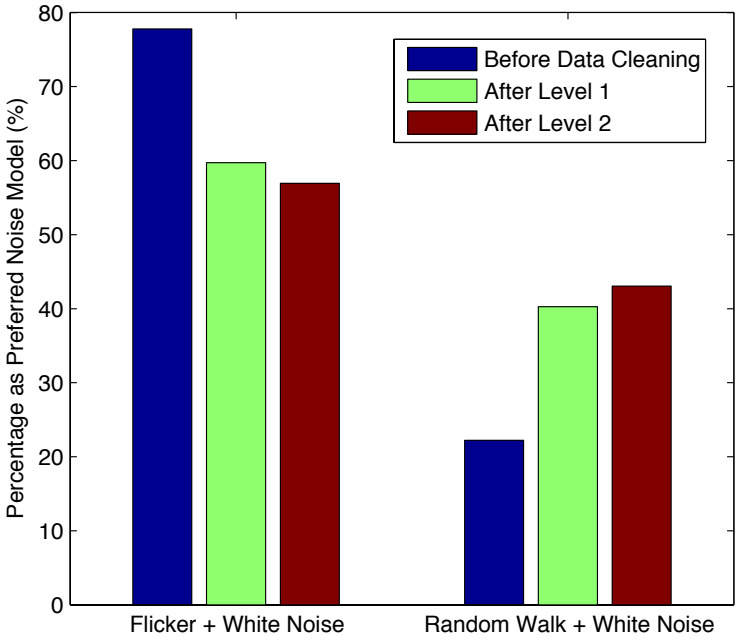


Figure 4.2: *Percentage as preferred noise model between flicker + white noise and random walk + white noise*

Site	AC	MLE Value (North)		
		White only	White + Flicker	White + Random Walk
GLSV	MIT	-13654.59	**-9450.27**	-9585.83
	ESA	-11305.55	**-8322.74**	-8361.34
GRAZ	MIT	-15130.72	**-13168.21**	-13285.61
	CODE	-14431.97	-9675.18	**-9442.4**
	ESA	-15559.05	**-13093.54**	-16083.91
	GFZ	-14077.26	-17226.99	**-12601.12**
MADR	MIT	-12545	**-9754.15**	-9841.81
	GFZ	-13547.45	-13154.12	**-12347.24**
MATE	MIT	-15444.79	**-13396.98**	-19614.72
	CODE	-14755.39	-19547.22	**-9740.76**
	ESA	-16640.79	**-14256.47**	-17957.8
	GFZ	-15019.62	**-13286.13**	-18846.39
POLV	MIT	-15412.214	**-19874.23**	-13154.72
	CODE	-13487.31	-13214.73	**-11544.54**
	GFZ	-19884.07	-154621.94	**-13457.87**
WTZZ	MIT	-9274.63	**-7909.88**	-12539.43
	CODE	-8488.4	-12520.92	**-4437.46**
ZECK	MIT	-12872.21	**-9824.16**	-11247.21
	ESA	-9845.21	-15412.65	**-9576.21**
	GFZ	-13564.46	**-12545.41**	-15634.29
ZIMM	MIT	-14154.8	-18795.35	**-13541.79**
	CODE	-12956.03	**-9714.8**	-18526.42
	ESA	-7448.66	**-6667.96**	-7114.53
	GFZ	-14200.37	-18395.37	**-13162.77**

Table 4.4: MLE *results of North components after data cleaning level 2*

To determine which one of the two model is generally preferred, one problem is to determine the impact of data cleaning on the realistic noise nature of GPS time series. There are no specific standard or limit for an outlier. The data cleaning algorithm in a certain level can possibly eliminate some realistic data points with characteristics of an outlier. And the preferred noise model can therefore be altered. For example, in one case, the first data cleaning only eliminated 6 data points from the time series, but the preferred noise model changed from white noise plus flicker noise to white noise plus random walk noise. In general, flicker noise plus white noise is always more likely to be the preferred noise model of any GPS time series.

Previous studies of Zhang et al. (1997) , Mao et al. (1999) and Williams et al. (2004) all showed that of the three models, white noise plus flicker noise is the preferred model. In Williams et al. (2004) study, a total number of 2962 log likelihood of 954 sites were compared and 83.46% of them suggest that white plus flicker noise is the preferred model.

Another reason that support the white noise plus flicker noise model is the prevalence of flicker noise in so many naturally occuring phenomenon, such as sunspot variability, the wobble ot the Earth about its axis, undersea currents, and uncertainties in time measured by atomic clocks (Gardner, 1978). On the other hand, several studies also acknowledge the importance of random walk noise in GPS data. Cumulative disturbances from the soils and weather displace

geodetic monuments with respect to the deeper crust (Johnson and Agnew, 1995). Whether or not the random walk noise is detectable depends on the length of the time series, the sampling frequency, and the relative amplitudes of the other noise components.

4.1.2 Amplitudes estimation

Another important parameter of the noise analysis is amplitudes. Again noise amplitudes for each site and for each component are summarized.

4.1.2.1 Amplitude of White Noise and Flicker Noise

As the flicker noise plus white noise is the generally preferred noise model. The noise amplitudes of these components are estimated in each time series. The results of the amplitude estimation of GPS time series before data cleaning are showed in table 4.6. The rest are shown in the appendix in table A.7 and A.8.

Site	AC	White Noise [mm]			Flicker Noise [mm]		
		North	East	Vertical	North	East	Vertical
GLSV	MIT	0.35± 0.10	0.13± 0.26	0	6.32± 0.14	6.39± 0.14	19.65±0.19
	ESA	0.49± 0.12	0.64±0.10	0	7.87± 0.19	7.88± 0.21	20.60± 0.22
GRAZ	MIT	1.26± 0.04	1.81± 0.03	1.30± 0.24	5.93± 0.17	5.64± 0.19	19.17± 0.42
	CODE	0	0	0	4.19±0.04	4.85± 0.04	12.05± 0.10
	ESA	1.44± 0.05	1.75± 0.06	1.94± 0.20	7.73± 0.21	8.21± 0.25	20.13± 0.49
	GFZ	1.31± 0.04	1.74± 0.05	3.20± 0.10	5.84± 0.19	6.89± 0.23	15.40± 0.45
MADR	MIT	1.33± 0.14	2.12± 0.05	2.59± 0.13	6.28± 0.12	7.32± 0.21	15.61± 0.22
	GFZ	1.45± 0.03	1.82± 0.06	3.06± 0.13	5.66± 0.23	5.76± 0.24	17.32± 0.37
MATE	MIT	1.19± 0.03	2.13± 0.03	2.92± 0.10	4.83± 0.14	4.60± 0.17	17.08± 0.40
	CODE	0	0	0	4.22± 28.39	4.99± 0.04	11.66± 10.27
	ESA	0	0	1.65	18.95	24.01	22.82
	GFZ	1.49± 0.03	2.14± 0.03	4.07± 0.08	4.75± 0.14	5.31± 0.20	14.14± 0.41
POLV	MIT	1.24± 0.06	1.76± 0.03	3.45± 0.12	5.62± 0.15	6.30± 0.15	14.33± 0.35
	CODE	0	0	0	6.72± 8.25	5.22± 0.06	12.25± 7.23
	GFZ	1.35± 0.04	3.14± 0.04	5.25± 0.06	5.13± 0.12	4.81± 0.19	15.25± 0.36
WTZZ	MIT	0.73± 0.02	0.65± 0.02	0	3.47± 0.10	3.32± 0.10	17.48± 5.82
	CODE	0	0	0	2.63± 39.17	2.52± 40.81	10.77± 9.57
ZECK	MIT	1.23± 0.04	2.14± 0.05	2.05± 0.13	5.45± 0.13	6.13± 0.18	18.10± 0.34
	ESA	1.54± 0.06	2.43± 0.08	2.89± 0.16	6.89± 0.32	8.42± 0.41	17.45± 0.83
	GFZ	0	0	3.15± 0.24	5.79± 0.25	6.43± 0.24	15.68± 0.43
ZIMM	MIT	1.35± 0.03	1.89± 0.04	1.88± 0.15	5.12± 0.17	5.37± 0.21	17.10± 0.40
	CODE	0	0	0	4.33± 0.04	4.79± 0.04	11.07± 10.84
	ESA	2.14± 0.09	2.73± 0.10	3.79± 0.19	9.91± 0.41	10.29± 0.48	17.90± 0.84
	GFZ	0	0	2.9	15.82± 0.13	20.4	14.37

Table 4.6: *Estimated noise amplitudes before data cleaning*

As expected, the data cleaning had some what reduced the noise amplitudes of all three components. The horizontal component are less noisy for white and flicker noise magnitudes than the vertical components by a factor of 2-3. And the east components are some what noisier than the north components.

In a few of the cases before the data cleaning and many of the cases after, the algorithm would often prefer a power-law noise only model, when the noise ratio is high. The more data were eliminated , the closer the noise was to a power-law noise only model. This is a great proof that data cleaning can strongly reduce the amplitude of white noise, and therefore with a high noise ratio, with MLE, a cleaned time series often prefers a power-law noise only model.

4.1.2.2 Latitude Dependence

Williams et al. (2004) found a latitude dependence for white noise and flicker noise. First of all, we should determine whether the noise magnitude in all three components was equal at all latitudes using the standard F test, and the null hypothesis was rejected. A function was arbitrarily derived to highlight any mid-latitude dependence and a hemisphere bias by Williams et al. (2004). The function take the form:

$$\sigma = \begin{cases} a + be^{-c\lambda^2}, & \text{if } \lambda = 0 \\ a + d(e^{-c\lambda^2} - 1) + b, & \text{otherwise} \end{cases} . \tag{4.2}$$

Where σ is the noise amplitude, λ is the site latitude, and a, b, c, d are the estimated parameters.

Site	Latitide [°]	Noise Amplitudes [mm]					
		Flicker Noise			White Noise		
		North	East	Vertical	North	East	Vertical
NYAL	78.93	7.08	5.46	12.7	2.56	2.99	5.92
TIXI	71.63	5.76	7.21	14.84	2.31	2.85	5.06
YELL	62.48	5.9	7.95	14.15	2.38	3.12	5.22
IRKT	52.22	5.5	10.81	14.54	2.26	3.27	5.10
WTZZ	49.14	2.63	2.52	10.77	1.95	3.36	5.17
GRAZ	47.07	4.19	4.85	12.05	2.10	2.94	5.92
ZIMM	46.88	4.33	4.79	11.07	2.17	3.14	6.03
MATE	40.65	4.22	5	11.66	3.11	2.51	5.13
WUHN	30.53	5.45	12.8	12.16	4.09	3.78	6.99
IISC	13.02	13	16.85	12.51	3.89	4.22	8.12
BOGT	4.64	6.73	4.59	18.83	4.01	4.64	9.42
NTUS	1.35	4.95	11.6	12.04	3.48	4.79	8.74
KARR	-20.98	21.77	14.1	9.02	2.55	5.12	8.15
PERT	-31.8	23.5	14.4	8.59	1.75	7.20	4.29
KERG	-49.35	5.42	4.79	13.92	2.21	2.39	5.96
MAC1	-54.5	11.76	5.98	7.99	1.88	2.49	4.00
MAW1	-67.6	3.72	3.23	8.6	1.86	1.62	4.80
MCM4	-77.84	4.99	4.98	10.69	1.50	2.49	3.85

Table 4.7: *Selected sites to demonstrate latitude dependence*

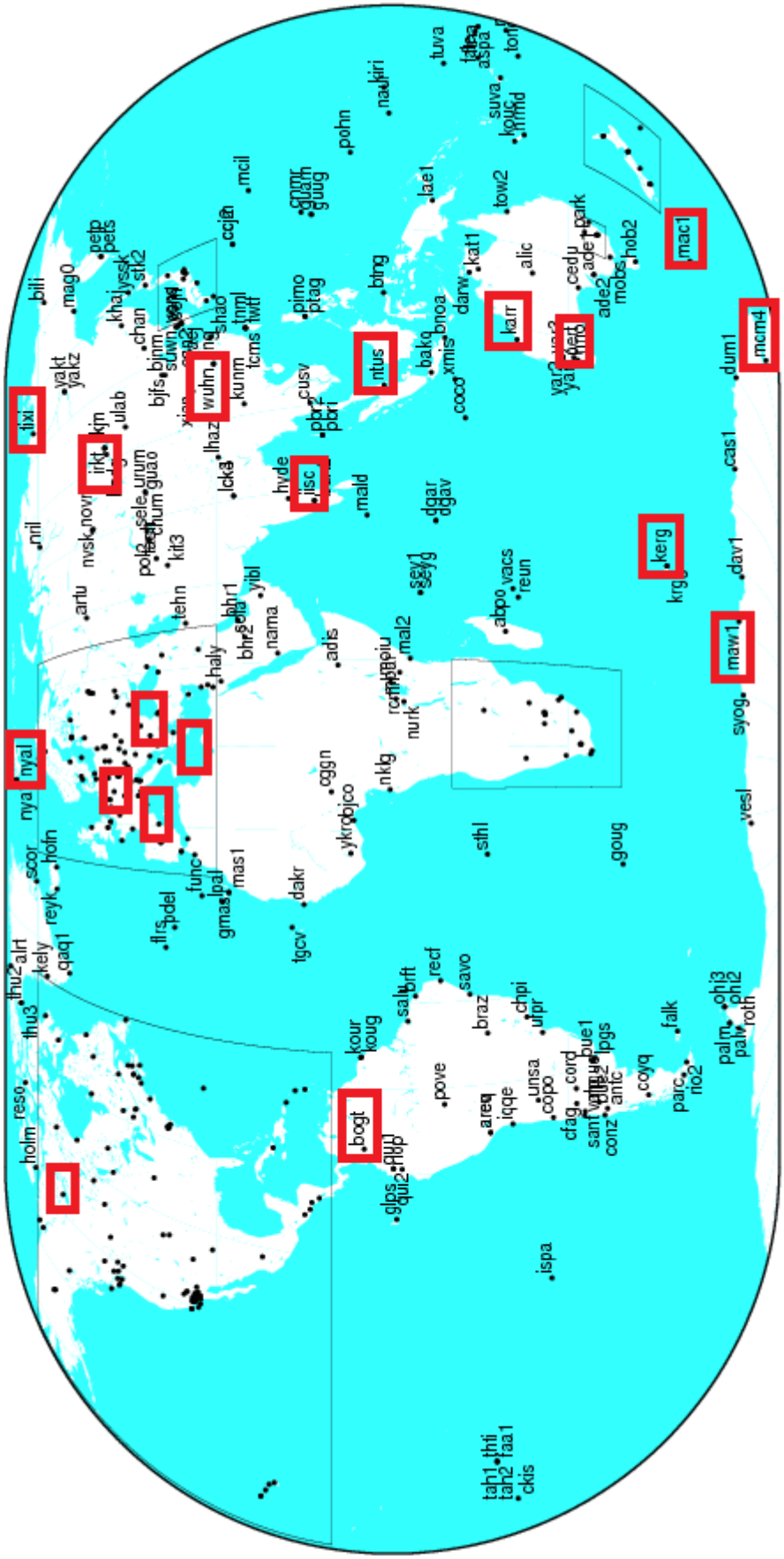


Figure 4.3: Selected sites to demonstrate latitude dependence

To demonstrate the latitude dependence of noise components, 18 sites with different latitudes were selected as showed in table 4.7 and figure 4.3. The noise amplitudes of flicker noise and white noise for each of the three components were compared and a polyfit was calculated for each noise group and plotted, so that we can better understand the variation tendency of the amplitudes with the change of the latitude. The time series was from analysis center CODE.

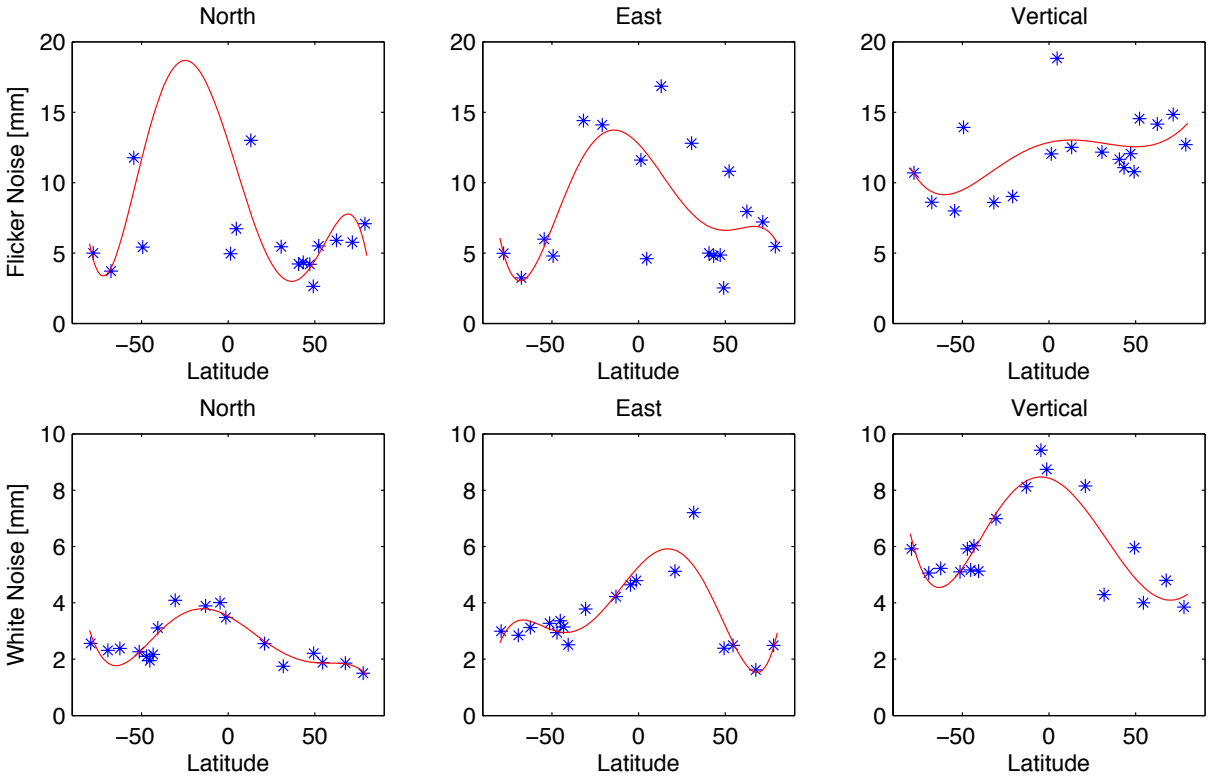


Figure 4.4: *Latitude dependence of noise amplitudes*

As showed in figure 4.4, in all cases the fit to data does suggest a latitude dependence. The fit to equation showed a significant positive bias for Southern Hemisphere sites. They are noisier than sites in the Northern Hemisphere And sites near the equator usually have the noise with greatest amplitudes.

4.2 Noise Models with Non-integer Spectral Index

As we have discovered, a noise model with integer spectral index may not be the optimal description for GPS time series. With limited knowledge of the noise source, site history and criteria for outliers, it is hard to decide between the white plus flicker noise model and white plus random walk noise model. With estimated spectral index we can have a more realistic understanding of the noise components. More accuracy and efficiency are therefore guaranteed in the search for the physical phenomena that affects the time series.

On the assumption that all the noise are composed of white noise and power-law noise, 96 time series of 10 selected stations from different ACs were analyzed again to estimate the spectral

index of their power-law noise components. With time series of this size (around 7000 data points), a MLE for one station takes several days. North, east and up components are treated separately. Table 4.8 shows the estimation results. Spectral indices estimated from different analysis centers for a same station show, to some extent, agreement with each other.

Site	AC	Spectral Index κ		
		North	East	Vertical
GLSV	CODE	-1.06	-1.01	-1.07
	ESA	-1.10	-1.00	-0.99
	GFZ	-0.97	-1.05	-0.95
GRAZ	CODE	-1.57	-1.39	-0.93
	ESA	-0.90	-1.01	-0.89
	GFZ	-1.03	-0.84	-0.89
	MIT	-1.36	-0.95	-0.89
MADR	GFZ	-0.99	-0.93	-0.89
	MIT	-0.92	-1.20	-0.93
MATE	CODE	-1.41	-1.14	-1.62
	ESA	-0.97	-1.02	-1.30
	GFZ	-0.71	-1.00	-1.01
	MIT	-0.94	-0.82	-0.83
NICO	CODE	-1.45	-1.32	-1.02
	GFZ	-1.23	-1.00	-1.04
POLV	MIT	-0.92	-0.98	-1.21
	CODE	-1.35	-1.46	-1.10
	GFZ	-0.87	-0.97	-1.03
POTS	CODE	-1.48	-1.33	-1.54
	ESA	-0.94	-0.78	-1.02
	GFZ	-0.70	-0.98	-1.00
	MIT	-1.02	-1.01	-0.97
WTZZ	CODE	-1.68	-1.58	-1.43
	GFZ	-0.92	-1.25	-1.08
	MIT	-0.98	-0.89	-0.99
ZECK	MIT	-0.99	-1.23	-0.86
	ESA	-0.76	-1.01	-0.90
	GFZ	-1.00	-0.60	-1.33
ZIMM	CODE	-1.51	-1.33	-0.84
	ESA	-1.05	-1.03	-0.72
	GFZ	-0.80	-0.86	-0.76
	MIT	-0.87	-0.52	-0.82

Table 4.8: *Estimated Spectral index of Power Law Noise*

The spectral indices range from –0.5 to –1.6 with a mean index of –1.05 and standard deviation of 0.26. The results show a mean and standard deviation that is consistent with the values seen for the simulated time series containing white noise plus flicker noise, which also confirms the previous conclusion of Williams et al. (2004) that the noise processes in GPS time series can be adequately described by a white noise plus flicker noise model.

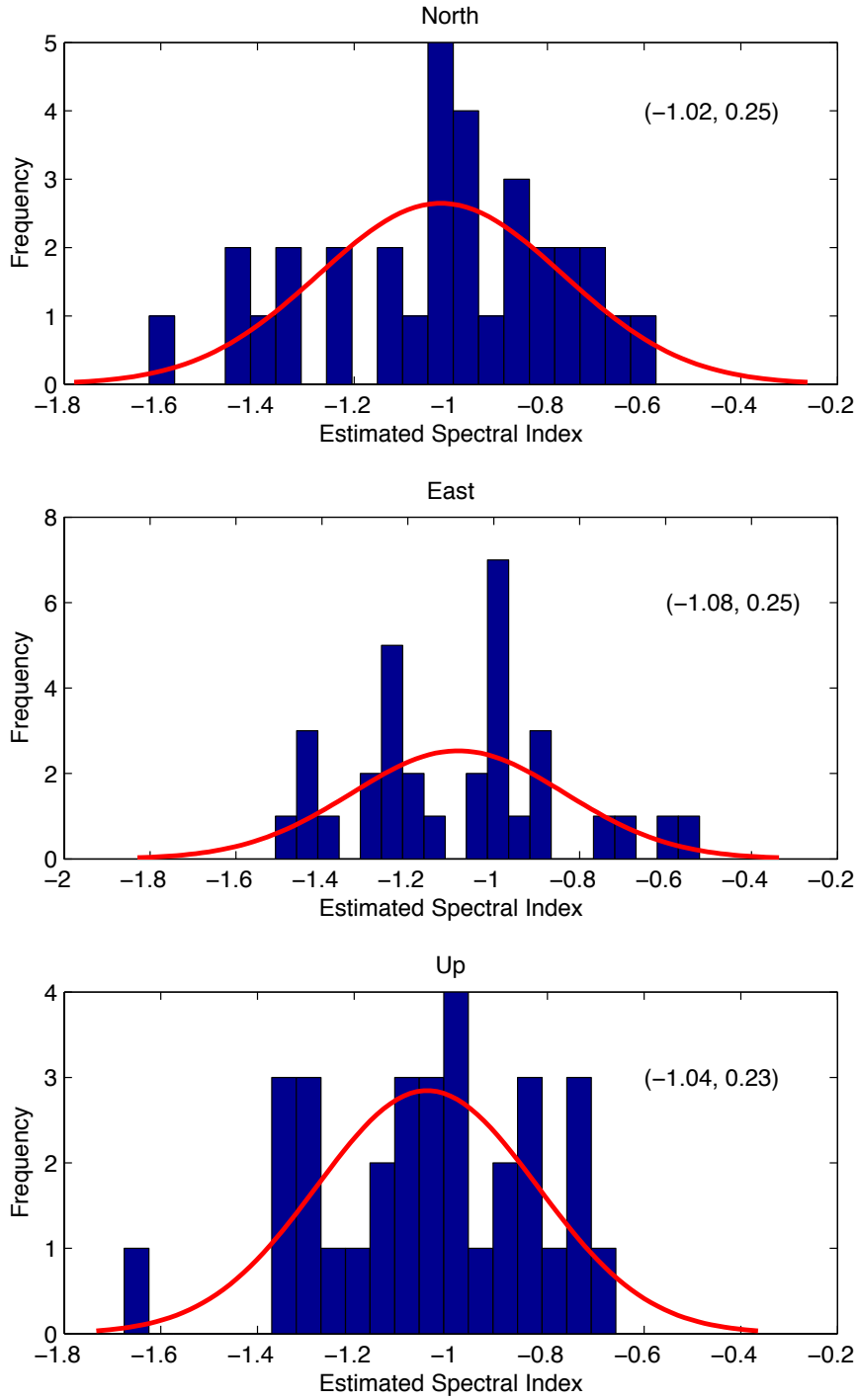


Figure 4.5: *Histogram of the estimated spectral index*

Figure 4.5 is a histogram of the estimated spectral index for the north, east, and vertical components at different sites, the red lines are calculated and plotted with matlab function $polyfit$. Because of the limited number of sites estimated, all of the results of a same direction were

combined together. Values inside the parentheses indicate the mean and standard deviation of the spectral estimates.

4.3 Comparison between Different ACs

For our analysis we mainly used the GPS data from four ACs (CODE, ESA, GFZ and MIT). A comparison between these analysis centers was made, in order to find out the differences between GPS time series from different analysis centers.

We plotted the cleaned time series of the same site from different ACs with Matlab. Figure 4.6 shows us the east components of the GPS time series of site MATE. In order to make the demonstration of the differences more significant, we extracted two years period ($2000 - 2002$) out of the time series.

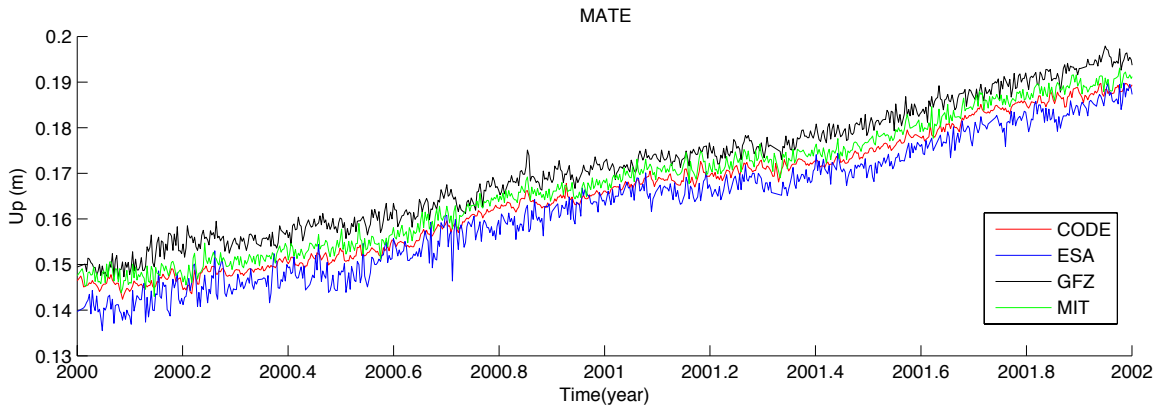


Figure 4.6: *East components of site* MATE

As we can see, time series from different ACs showed a relatively good inter-AC agreement. They exhibited similar patterns with time and had approximately equal noise amplitudes. This is also proved by the noise analysis using maximum likelihood estimation.

However, during our research, we found that time series from analysis center ESA, in some of the cases, showed relatively greater differences compared with the other three ACs. Figure 4.7 is an example of the case, the vertical components of the GPS time series of site GRAZ were showed in the plot.

As demonstrated below in figure 4.7, while time series from three analysis centers (CODE, GFZ and MIT) showed great inter-AC agreement with almost identical forms, the time series from ESA had significantly lower value of the vertical components compared with the others by around $4cm$.

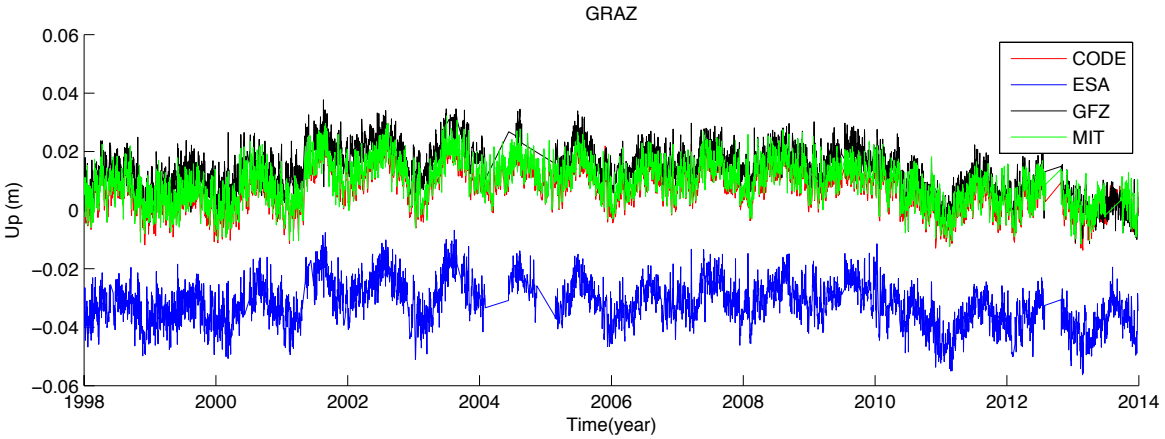


Figure 4.7: *Up components of site* GRAZ

It is obvious from the plot that the time series from ESA shares an identical pattern with the other time series. The possible reason for the difference is that at the beginning of some time series from ESA there are offsets, which may affect the value, but does not change the noise characteristic of a time series. However, the reason why there is significant offsets only in some of the time series from analysis center ESA is unknown.

The MLE results of different ACs were also compared. We calculated the percentage of time series from different ACs for which the likelihood value for a specific model is the highest. In another word, this noise model is determined to be the preferred model to describe the time series. Since no time series prefer a white noise only model, we only compare the two noise models of white noise plus power-law noise, the influence of data cleaning was also taken into consideration.

Data Cleaning	AC	Percentage	
		White + Flicker	White + Random Walk
Before	CODE	**40%**	**60%**
	ESA	86.7%	13.3%
	GFZ	72.2%	27.8%
	MIT	**100%**	**0%**
Level 1	CODE	**33.3%**	**66.7%**
	ESA	73.3%	26.7%
	GFZ	38.9%	61.1%
	MIT	**83.3%**	**16.7%**
Level 2	CODE	**26.7%**	**73.3%**
	ESA	73.3%	26.7%
	GFZ	38.9%	61.1%
	MIT	**79.2%**	**20.8%**

Table 4.9: *Percentage of time series from different* ACs *with preferred noise model*

From table 4.9 we know that time series from analysis center CODE are more likely to prefer a noise model which combines white noise and random walk noise, while time series from ESA and MIT are more likely to prefer a white noise plus flicker noise model.

AC	Spectral Index κ (mean)		
	North	East	Vertical
CODE	-1.44	-1.32	-1.19
ESA	-1.01	-1.04	-1.03
GFZ	-0.99	-1.02	-1.02
MIT	-0.99	-0.99	-0.99
Combined	**-1.02**	**-1.08**	**-1.04**

Table 4.10: *Average of estimated spectral indices from different* ACs

In the second part of our thesis, the spectral indices of the power-law noise components of GPS time series were estimated. To make a comparison between different ACs, the averages of the estimated spectral indices from each AC were calculated as showed in table 4.10. While the estimated spectral indices from most of the ACs are very close to the combined average and, at the same time, close to the spectral index of flicker noise ($\kappa = -1$), the estimated spectral indices from analysis center CODE are significantly lower, especially in horizontal components. This is probably also the reason why time series from CODE are more likely to prefer the white noise plus random walk noise model.

Chapter 5

Conclusion

For this thesis time series from different sites have been analysed in order to assess their noise characteristics. All of the data are provided by different analysis centers of the International GNSS Service after the second data reprocessing campaign. The average length of time series was around 6500 points (over 18 years).

This study as well as previous work indicates that significant coloured noise is ubiquitous in continuous GPS time series. This should be taken into account for any derived parameters and their uncertainties. Compared with white noise only model and white noise plus random walk noise model, the MLE analysis indicates that a combination of white noise and flicker noise is an appropriate stochastic model for all three coordinate components, which is also proved by the spectral index estimation of power-law noise. The estimated spectral index has a mean of –1.05 with a standard deviation of 0.26, consistent with the value of flicker noise ($\kappa = -1$).

For the white noise and flicker noise amplitudes, the horizontal components are less noisy than the vertical components, and the east components are some what noisier than the north components. The elimination of outliers shows a significant influence on the amplitudes of white noise, which may greatly increase the ratio of power-law noise and white noise and therefore make power-law noise only as a preferred noise model for many cleaned time series.

Both white noise and flicker noise amplitudes also show a latitude dependence with a maximum at the equator. For both noise components a southern hemisphere bias is apparent. southern hemisphere sites are noisier.

Compared with recent years, the noise amplitudes of GPS time series before 2000 are substantially higher. The possible reason is that the spatially correlated noise is partly related to the reference frame, which means as the number of sites available for the global reference frame stabilization increases, the noise related to it will be reduced. This may also be the reason for the southern hemisphere bias mentioned above.

A comparison between time series from different ACs were made. In general the four analysis centers we used in this thesis (CODE, ESA, GFZ and MIT) showed a good inter-AC agreement. In some of the cases, time series from analysis center ESA have significantly lower values while they share almost identical patterns with time series from the other three ACs. The possible reason is that at the beginning of these time series there are offsets, which may affect the value, but does not change the noise characteristic of a time series. While white noise plus flicker noise is generally regarded as preferred noise model, time series from analysis center CODE are more likely to prefer a white noise plus random walk noise model. The average of estimated spectral indices of the power-law noise components from CODE are significantly lower than those from the other ACs, and the estimated spectral indices from the other three ACs (ESA, GFZ and MIT)

are very close to the spectral index of flicker noise ($\kappa = -1$). This is probably the reason why time series from CODE are more likely to prefer the white noise plus random walk noise model ($\kappa = -2$).

From this thesis and many previous work we know that although white noise plus flicker noise is dominant as preferred noise model for GPS time series, there are a variety of other noise sources that may dominate including, for example, monument instabilities (random walk noise) and localized deformation due to changes in groundwater (unknown power-law noise plus annually repeating signals). That is the reason why classification and quantification of the noise components are so important in providing means for reducing the noise and increasing the accuracy and precision.

Bibliography

Agnew, DC (1992). The time domain behavior of power law noises. *Geophys Res Lett*, **19**(4):333–336. doi:10.1029/91GL02832.

Borsa, AA, Meertens, C and Jackson, M (2008). Noise characteristics of short drilled and deep drilled braced monuments in the PBO continuous GPS network. *AGU Fall Meeting Abstracts*, p. B626.

Calais, E (1999). Continuous GPS measurements across the Western Alps, 1996-1998. *Geophysical Journal International*, **138**(1):221–230. doi:10.1046/j.1365-246x.1999.00862.x.

Gardner, M (1978). White and brown music, fractal curves and one-over-fluctuation. *Mathematical Games, Scientific American, New York*, pp. 16–21.

Hosking, JRM (1981). Fractional differencing. *Biometrika*, **68**(1):165–176. doi:10.1093/biomet/68.1.165.

Johnson, H and Wyatt, F (1994). Geodetic network design for fault-mechanics studies. *Manuscripta geodaetica*, **19**(5):309.

Johnson, HO and Agnew, DC (1995). Monument motion and measurements of crustal velocities. *Geophys Res Lett*, **22**(21):2905–2908. doi:10.1029/95GL02661.

Langbein, J and Johnson, H (1995). Noise level of geodetic monuments. *Eos Trans AGU*, **76**:46.

——— (1997). Correlated errors in geodetic time series: Implications for time-dependent deformation. *J Geophys Res*, **102**(B1):591–603. doi:10.1029/96JB02945.

Mao, A, Harrison, CGA and Dixon, TH (1999). Noise in GPS coordinate time series. *J Geophys Res*, **104**(B2):2797–2816. doi:10.1029/1998JB900033.

Nikolaidis, R (2002). *Observation of Geodetic and Seismic Deformation with the Global Positioning System*. Ph.D. thesis, University of California.

Rebischung, P, Garayt, B, Collilieux, X and Altamimi, Z (2014). The IGS contribution to ITRF 2013 - preliminary results from the IGS repro2 SINEX conbinations (2014). In: *2014 IGS Workshop,Pasadena, California,USA*.

Williams, S (2008). CATS: GPS coordinate time series analysis software. *GPS Solutions*, **12**(2):147–153. doi:10.1007/s10291-007-0086-4.

Williams, S, Bock, Y and Fang, P (1998). Integrated satellite interferometry: Tropospheric noise, GPS estimates and implications for interferometric synthetic aperture radar products. *Journal of Geophysical Research: Solid Earth*, **103**(B11):27,051–27,067. doi:10.1029/98JB02794.

Williams, SDP (2003). The effect of coloured noise on the uncertainties of rates estimated from geodetic time series. *Journal of Geodesy*, **76**(9):483–494. doi:10.1007/s00190-002-0283-4.

Williams, SDP, Bock, Y, Fang, P, Jamason, P et al. (2004). Error analysis of continuous GPS position time series. *J Geophys Res*, **109**(B3):B03,412. doi:10.1029/2003JB002741.

Wyatt, F (1982). Displacement of surface monuments: Horizontal motion. *Journal of Geophysical Research: Solid Earth*, **87**(B2):979–989. doi:10.1029/JB087iB02p00979.

Wyatt, FK (1989). Displacement of surface monuments: Vertical motion. *J Geophys Res*, **94**(B2):1655–1664. doi:10.1029/JB094iB02p01655.

Zhang, J, Bock, Y, Johnson, H, Fang, P et al. (1997). Southern california permanent gps geodetic array: Error analysis of daily position estimates and site velocities. *J Geophys Res*, **102**(B8):18,035–18,055. doi:10.1029/97JB01380.

Appendix A

Estimation Results

A.1 MLE Value of Different Noise Model

Site	AC	MLE Value (East)		
		White only	White + Flicker	White + Random Walk
GLSV	MIT	-13496.35	**-9916.05**	-10072.89
	ESA	-10823.06	**-8772.12**	-8904.11
GRAZ	MIT	-16857.89	**-15823.51**	-15940.01
	CODE	-14839.69	**-10870.45**	-10907.84
	ESA	-15774.67	**-14178.98**	-14265.95
	GFZ	-15965.6	**-14989.57**	-15137.07
MADR	MIT	-15321.12	**-12354.21**	-12547.35
	GFZ	-15412.34	**-9845.98**	-9987.35
MATE	MIT	-18046.94	**-16349.95**	-16381.23
	CODE	-16418.13	**-11103.35**	-11363.21
	ESA	-17394.35	**-15758.84**	-19791.13
	GFZ	-17472.97	**-15851.51**	-15916.72
POLV	MIT	-15241.21	**-12471.34**	-13564.49
	CODE	-12143.21	**-9845.32**	-10021.24
	GFZ	-15412.68	**-9874.65**	-9987.74
WTZZ	MIT	-9345.14	**-7529.93**	-7634.78
	CODE	-8837.64	-4601.72	**-4288.69**
ZECK	MIT	-12145.65	**-10047.54**	-10121.78
	ESA	-16547.54	**-12087.61**	-12236.36
	GFZ	-15657.31	**-12001.32**	-12354.83
ZIMM	MIT	-16603.83	**-16100.2**	-16215.71
	CODE	-13674.71	**-10848.69**	-10976.92
	ESA	-7676.44	**-7243.19**	-7277.5
	GFZ	-17192.74	**-16712.82**	-20590.65

Table A.1: MLE *results of East components before data cleaning*

Site	AC	MLE Value (East)		
		White only	White + Flicker	White + Random Walk
GLSV	MIT	-13431.27	-15634.79	**-9911.31**
	ESA	-10736.25	**-8646.8**	-8776.84
GRAZ	MIT	-15771.11	**-14175.34**	-14330.2
	CODE	-14730.65	**-10679.32**	-10721.65
	ESA	-15668.71	**-14054.53**	-14130.43
	GFZ	-15401.83	-19450.9	**-14411.83**
MADR	MIT	-13241.21	**-11241.32**	-11345.21
	GFZ	-13541.54	**-10032.53**	-10159.83
MATE	MIT	-16976.27	-21064.97	**-14595.41**
	CODE	-16278.87	-20900.32	**-11282.56**
	ESA	-17199.03	-19615.64	**-15525.51**
	GFZ	-17015.14	-20582.92	**-15319.36**
POLV	MIT	-14514.41	**-12147.53**	-112357.94
	CODE	-16519.55	-15476.95	**-13654.54**
	GFZ	-15648.32	**-10254.61**	-10578.94
WTZZ	MIT	-12107.8	**-11087.34**	-11200.74
	CODE	-11354.55	**-8567.08**	-8615.67
ZECK	MIT	-11459.71	-8569.36	**-8474.38**
	ESA	-15674.95	-10254.98	**-10023.43**
	GFZ	-14378.53	**-12001.24**	-12044.79
ZIMM	MIT	-15523.67	**-14896.41**	-15008.97
	CODE	-13603.73	**-10769.85**	-10900.56
	ESA	-7645.83	**-7209.3**	-7243.45
	GFZ	-15424.96	-19933.85	**-14492.59**

Table A.2: MLE *results of East components after data cleaning level 1*

Site	AC	MLE Value (East)		
		White only	White + Flicker	White + Random Walk
GLSV	MIT	-13215.37	**-9555.14**	-9709.01
	ESA	-10503.88	**-8464.19**	-8579.16
GRAZ	MIT	-15184.7	**-13550.1**	-13651.39
	CODE	-14404.59	**-10275.33**	-10315.12
	ESA	-15342.93	-17687.14	**-13843.5**
	GFZ	-14911.71	-19198.85	**-13848.28**
MADR	MIT	-145873.47	-19476.33	**-12478.64**
	GFZ	-15947.26	**-13613.28**	-13862.57
MATE	MIT	-16725.57	**-14345.2**	-20837.48
	CODE	-16169.98	-20793.25	**-11200.54**
	ESA	-16952.41	-19418.99	**-15254.72**
	GFZ	-16530.43	-20202.83	**-14673.79**
POLV	MIT	-16781.27	**-13468.18**	-13784.59
	CODE	-15329.26	-19945.12	**-14678.46**
	GFZ	-14653.26	-18472.59	**-12534.79**
WTZZ	MIT	-9250.72	**-7655.57**	-13847.71
	CODE	-8749.47	**-4274.59**	-13912.06
ZECK	MIT	-13378.54	-15487.65	**-11547.96**
	ESA	-16541.57	-120487.32	**-14213.73**
	GFZ	-15542.67	**-11354.92**	-15412.35
ZIMM	MIT	-14902.42	-19977.35	**-14407.6**
	CODE	-13370.76	-19701.3	**-10724.05**
	ESA	-7397.55	**-7031.94**	-7522.49
	GFZ	-14948.28	-19683.16	**-13976.99**

Table A.3: MLE *results of East components after data cleaning level 2*

Site	AC	MLE Value (Up)		
		White only	White + Flicker	White + Random Walk
GLSV	MIT	-17974.25	**-16073.2**	-16305.39
	ESA	-13903.21	**-12505.04**	-12701.06
GRAZ	MIT	-23396.84	**-20633.86**	-20913.06
	CODE	-23145.19	-17366.47	**-16610.36**
	ESA	-20404.94	**-17979.81**	-18175.44
	GFZ	-22377.99	**-19540.86**	-19708.67
MADR	MIT	-20457.35	**-16587.64**	-19846.37
	GFZ	-22478.65	**-16541.54**	-18474.65
MATE	MIT	-24017.44	**-21379.41**	-21445.38
	CODE	-23140.29	-17182.46	**-16800.22**
	ESA	-23139.69	**-19842.46**	-19881.87
	GFZ	-22653.88	**-20898.37**	-20940.8
POLV	MIT	-18465.74	**-16403.52**	-16742.38
	CODE	-21310.24	-18656.53	**-18452.37**
	GFZ	-22485.31	-18569.66	**-18214.75**
WTZZ	MIT	-15939.9	**-14509.5**	-14737.31
	CODE	-15469.38	-12300.43	**-11284.68**
ZECK	MIT	-18741.25	**-15431.65**	-15687.32
	ESA	-21966.81	**-18432.68**	-19457.86
	GFZ	-15542.67	**-11354.92**	-15412.35
ZIMM	MIT	-22247.99	**-20562.48**	-20817.02
	CODE	-21277.51	-16868.08	**-16257.73**
	ESA	-8869.8	**-8384.52**	-8454.07
	GFZ	-21934.73	**-20437.91**	-20537.45

Table A.4: MLE *results of Up components before data cleaning*

Site	AC	MLE Value (Up)		
		White only	White + Flicker	White + Random Walk
GLSV	MIT	-17960.13	**-16107.71**	-16286.18
	ESA	-13795.51	**-12403.64**	-12602.11
GRAZ	MIT	-23181.87	**-20322.97**	-20586.68
	CODE	-23033.76	-17224.7	**-16511.01**
	ESA	-20170.7	**-17744.2**	-17912.07
	GFZ	-22137.85	**-19246.41**	-19339.39
MADR	MIT	-19578.36	**-17428.62**	-17658.59
	GFZ	-20164.74	-18465.91	**-18367.64**
MATE	MIT	-22708.5	**-20827.51**	-20836.23
	CODE	-21053.92	-16922.59	**-16616.09**
	ESA	-20672.63	**-19124.89**	-19138.56
	GFZ	-22046.52	-20526.73	**-20503.97**
POLV	MIT	-22647.32	-18635.24	**-18432.65**
	CODE	-19346.54	-16472.55	**-16325.41**
	GFZ	-17462.94	-16474.46	**-16354.87**
WTZZ	MIT	-16859.63	**-15203.76**	-15375.65
	CODE	-16662.77	-12904.62	**-12515.25**
ZECK	MIT	-17645.28	**-14354.33**	-14487.78
	ESA	-18465.84	**-16352.12**	-16472.51
	GFZ	-18848.29	**-15211.46**	-15343.62
ZIMM	MIT	-21791.84	**-19994.16**	-20245.99
	CODE	-21102.17	-16671.87	**-16065.54**
	ESA	-8736.6	**-8225.05**	-8291.77
	GFZ	-21572.89	**-19953.83**	-20042.39

Table A.5: MLE *results of Up components after data cleaning level 1*

Site	AC	MLE Value (Up)		
		White only	White + Flicker	White + Random Walk
GLSV	MIT	-17431.03	**-15606.22**	-15850.22
	ESA	-13376.61	**-12013.07**	-12205.06
GRAZ	MIT	-22831.61	**-19943.41**	-20198.37
	CODE	-22695.51	-16851.79	**-16156.48**
	ESA	-19733.96	**-17375.82**	-17484.2
	GFZ	-21722.37	**-18619.49**	-18710.57
MADR	MIT	-22145.32	-18465.36	**-183241.68**
	GFZ	-19642.38	-16433.52	**-16358.72**
MATE	MIT	-21540.55	**-20316.82**	-20369.67
	CODE	-20595.25	-16666.74	**-16437.6**
	ESA	-19944.57	**-18682.24**	-18739.38
	GFZ	-20928.82	-19698.7	**-19688.81**
POLV	MIT	-21632.57	**-18135.65**	-18343.62
	CODE	-20418.67	-18422.65	**-17896.58**
	GFZ	-19632.44	-17475.23	**-16989.67**
WTZZ	MIT	-15603.23	**-14224.06**	-14426.63
	CODE	-15197.83	**-11161.94**	-12172.77
ZECK	MIT	-15421.37	**-13004.56**	-13298.73
	ESA	-16287.59	**-14259.68**	-14322.71
	GFZ	-20014.39	**-18239.61**	-18456.37
ZIMM	MIT	-21301.73	**-19587.47**	-19748.32
	CODE	-20721.85	-16458.17	**-15773.26**
	ESA	-8397.11	**-7906.93**	-7957.58
	GFZ	-21101.54	**-19428.33**	-19514.48

Table A.6: MLE *results of Up components after data cleaning level 2*

A.2 Noise Amplitude Estimation

Site	AC	White Noise [mm]			Flicker Noise [mm]		
		North	East	Vertical	North	East	Vertical
GLSV	MIT	0	0	0	11.04± 0.11	18.16± 0.17	19.79± 0.19
	ESA	0.44± 0.12	0.62± 0.10	0	7.69± 0.18	7.71± 0.21	20.31± 4.50
GRAZ	MIT	1.09± 0.04	1.20± 0.04	1.39± 0.20	5.71± 0.16	5.68± 0.16	18.21± 0.40
	CODE	0	0	0	4.18± 0.04	4.74± 25.23	11.85± 10.08
	ESA	1.45± 0.05	1.75± 0.06	2.25± 0.16	7.48± 0.20	8.01± 0.25	18.65± 0.16
	GFZ	0	0	2.93	15.77± 0.14	21.54± 0.19	15.47
MADR	MIT	0	0	2.83± 0.12	7.41± 0.10	15.12± 0.13	16.23± 0.21
	GFZ 0	0	2.48± 0.13	12.32± 0.15	11.36± 0.21	15.22± 0.28	
MATE	MIT	0	0	2.86	17.96	21.46± 0.18	15.58
	CODE	0	0	0	16.77± 0.14	19.99± 0.17	11.43± 0.10
	ESA	0	0	2.81	18.94	23.95	17.6
	GFZ	0	0	3.72	18.54± 0.16	22.84± 0.20	14.42
POLV	MIT	0	0	2.45± 0.26	7.63± 0.17	11.23± 0.21	15.62± 0.28
	CODE	0	0	0	5.86± 0.25	4.78± 0.13	12.03± 0.34
	GFZ	0	0	2.23± 0.06	5.17± 0.15	7.34± 0.17	14.25± 0.36
WTZZ	MIT	1.22± 0.05	1.36± 0.05	2.08± 0.16	6.22± 0.20	6.34± 0.20	17.35± 0.47
	CODE	0	0	0	4.61± 0.04	5.34± 0.05	12.09± 0.12
ZECK	MIT	1.25± 0.04	1.62± 0.06	3.12± 0.12	4.73± 0.15	4.84± 0.17	12.37± 0.28
	ESA	0	0	1.89± 0.13	10.27± 0.14	6.32± 0.15	16.21± 0.35
	GFZ	0	0	2.79± 0.18	7.79± 0.23	7.25± 0.04	11.23± 0.22
ZIMM	MIT	1.20± 0.03	1.59± 0.03	1.62± 0.15	4.98± 0.15	4.65± 0.17	16.19± 0.37
	CODE	0	0	0	4.25± 0.03	4.75± 0.04	10.81± 0.09
	ESA	2.14± 0.09	2.70± 0.10	3.59± 0.18	9.85± 0.41	10.25± 0.48	16.94± 0.79
	GFZ	0	0	2.67	15.46	18.71	13.77

Table A.7: *Estimated noise amplitudes after data cleaning level 1*

Site	AC	White Noise [mm]			Flicker Noise [mm]		
		North	East	Vertical	North	East	Vertical
GLSV	MIT	0	0	0	6.05± 0.06	6.16± 0.06	18.81± 0.18
	ESA	0.36± 0.15	0.73± 0.20	0	7.71± 0.18	7.37± 0.20	19.38± 0.21
GRAZ	MIT	1.10± 0.03	1.26± 0.03	1.56± 0.17	4.79± 0.14	4.57± 0.14	17.24± 0.39
	CODE	0	0	0	4.16± 0.03	4.53± 26.27	11.47± 10.37
	ESA	0	0	1.49	16.69± 0.15	21.98	19.53
	GFZ	0	0	2.63	15.67	21.34	14.88
MADR	MIT	0	0	2.76	16.79	15.68	15.13
	GFZ	0	0	3.12	17.42± 0.10	18.35	15.67
MATE	MIT	0	0	2.98	18.04	21.53	14.05
	CODE	0	0	0	16.81	20.05	11.19± 0.09
	ESA	0	0	2.98	18.93	23.95± 0.21	15.87
	GFZ	0	0	3.51	18.45± 0.16	22.67	13.08
POLV	MIT	0	0	2.98	16.23	19.42	16.15
	CODE	0	0	0	17.91	22.05± 0.09	14.26
	GFZ	0	0	2.51	16.37± 0.12	18.24	14.95
WTZZ	MIT	0	0	0	12.2	15.74	16.94
	CODE	0	0	0	11.45± 0.11	14.91± 0.15	10.72± 0.10
ZECK	MIT	0	0	0	16.32± 0.11	21.32	15.61
	ESA	0	0	2.94	17.14± 0.15	17.28± 0.24	14.82
	GFZ	0	0	3.17	16.43± 0.11	22.18	17.02
ZIMM	MIT	0	0	1.35	15.59± 0.13	18.46	16.16
	CODE	0	0	0	14.27	16.83	10.68
	ESA	0	0	3.4	16.64± 0.23	19.49± 0.27	15.9
	GFZ	0	0	2.5	15.40± 0.13	18.58	13.33

Table A.8: *Estimated noise amplitudes after data cleaning level 2*