Norwegian University of Life Sciences
Faculty of Science and Technology (REALTEK)

Philosophiae Doctor (PhD)
Thesis 2018:14

Implementation of GNSS data integrity for precise real-time positioning

Implementering av GNSS dataintegritet for presis sanntidsposisjonering

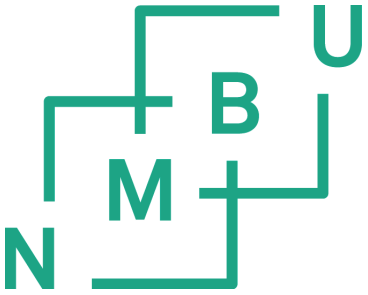
Mohammed Ouassou

Implementation of GNSS data integrity for precise real-time positioning

Implementering av GNSS dataintegritet
for presis sanntidsposisjonering

Philosophiae Doctor (PhD) Thesis
Mohammed Ouassou

Norwegian University of Life Sciences
Faculty of Science and Technology
Ås (2017)



Thesis 2018:14
ISSN 1894-6402
ISBN 978-82-575-1488-4

Abstract

In recent years, the area of global navigation satellite systems (GNSS) has received a lot of attention, with a large range of fruitful results. This includes a modernization of the existing American (GPS) and Russian (GLONASS) systems; the advent of European (Galileo) and Chinese (Beidou) systems; the adoption of multiple civilian frequencies for the broadcasts; and novel algorithms needed to process the ever-growing data sets. In total, this has resulted in more reliable positioning services for the GNSS end-users.

Network real-time kinematic (NRTK) positioning services have especially benefited from these recent developments. However, one weakness that the current generation of NRTK software still suffers from, is that the end-users do not know to what degree they can trust the results. To perform such a quality-check, post-processing of the data is required – and a decision has to be made regarding whether to acquire new data. The current situation will not satisfy the needs of future NRTK users, who are increasingly demanding real-time information about the data quality, i.e. that integrity information is made available simultaneously with the positioning data itself.

The purpose of this thesis is to design and implement such an NRTK data integrity solution. The guiding concept behind the algorithms presented will be to exert a continuous control over the GNSS information across three different levels: the network corrections, the baseline, and the rover data.

The integrity monitoring algorithms developed and described in this thesis are NRTK-agnostic, and work regardless of the method used for generation of the NRTK corrections. However, for derivation and development of the integrity and quality control algorithms we have based our work on the conceptual approach as given by the *NetAdjust* method [89, 90]. Most of the NRTK techniques are developed commercially and details about these are not readily available. But the NetAdjust method is well described in literature, it is therefore suitable as a starting point for our work, and we provide a review of the method in coming chapters.

The algorithms were tested using data from the Norwegian RTK network known as CPOS operated by the Norwegian Mapping Authority (NMA). The test area is from the Rogaland region in the south west of Norway, covering an area of 150 km × 150 km and composed of six reference receivers. Reference receivers are equipped with *Trimble NetR9* receivers, tracking GPS and GLONASS satellite signals. Baselines vary between 35–112 km and the height difference between the sites is about 225 m.

Acknowledgements

This dissertation presents the work of one individual, and could not have been realized without the help and support of the people around me. It is therefore a pleasure to express my gratitude to everyone who helped me make this thesis possible.

First, I wish to express my sincere gratitude to Prof. Jon G. O. Gjevestad for his continuous supervision, guidance, support during my time as a PhD candidate.

I also wish to thank the head of the Geodetic Institute at the Norwegian Mapping Authority (NMA), Per Erik Opseth, for believing in and supporting the project from the beginning to the end. Furthermore, I wish to thank the Faculty of Science and Technology at the Norwegian University of Life Sciences (NMBU) for offering this opportunity to me.

Note that during the time that the material in this thesis was developed, I was simultaneously employed as Research GNSS Analyst at the Geodetic Institute, NMA, whom I also thank for providing financial support for this project. The NMA also provided the data sets used for this investigation, which were obtained from the CPOS reference network. I wish to thank Tor-Ole Dahlø for his assistance with preparing this data for analysis.

Prof. Anna B. O. Jensen deserves my sincere gratitude for her inspiring guidance, support, advice, discussions, and encouragement during my research. She was a true friend to me during my research and was always available when I needed any kind of assistance. Thank you, Anna!

Thanks to Knut S. Jacobsen and Åsmund Skjæveland for the excellent work editing my thesis.

For their deep scientific discussions and encouraging support, I also wish to thank Prof. John F. Raquet (AFIT), Prof. Börje Forssell (NTNU), Prof. Hossein Nahavandchi (NTNU), Prof. Bent Natvig (UiO), Prof. Fred Godtliebsen (UiT), Jabir A. Ouassou (NTNU), and Oskar Johnsrud (NMA).

I would like to express my deepest gratitude and sincere appreciation to my whole family: my wife Gunn for her special help and understanding; and to my son Ali; and my daughters Nina, Nadia and Hedda. Finally, I dedicate this thesis to the most beautiful women involved in my life: my mother Hadhoume who gave me birth and an enjoyable life, and my wife Gunn R. Ouassou who committed to sharing life responsibilities with me.

Abbreviations

Abbreviation	Description
ACF	Auto-Correlation Function
ACVF	Auto-CoVariance Function
AIC	Akaike Information Criterion
AICC	Corrected AIC
AR	Auto-Regressive
ARMA	AR Moving Average
ARIMA	AR Integrated Moving Average
AS	Anti-Spoofing
BNR	Bias to Noise Ratio
CGPS	Continuous GPS
CME	Common-Mode Error
CPOS	Centimeter POSition based on NRTK
DD	Double-Difference
DFA	Dynamic Factor Analysis
DIA	Detection, Identification, and Adaptation
DID	Detection, Isolation, and Decision
DOP	Dilution Of Precision
ECEF	Earth-Centered, Earth-Fixed
EGNOS	European Geostationary Navigation Overlay System
EVD	Extreme Value Distribution
EVT	Extreme Value Theory
FA	Factor Analysis
GAGAN	GPS-Aided GEO-Augmented Navigation
GEO	Geostationary Earth Orbit
GDOP	Geometrical DOP
GARCH	Generalized AR Conditional Heteroskedasticity
GIVE	Grid Ionospheric Vertical Errors
GMRF	Gauss-Markov Random Field
GNSS	Global Navigation Satellite System
GPD	Generalized Pareto Distribution
GPS	Global Positioning System
HPL	Horizontal Protection Level
IGP	Ionospheric Grid Point
INLA	Integrated Nested Laplace Approximation
IPP	Ionospheric Piercing Point

Abbreviation	Description
KDE	Kernel Density Estimation
LOS	Line Of Sight
LLI	Loss of Lock Indicator
MA	Moving Average
MLE	Maximum Likelihood Estimation
MSL	Mean Sea Level
NMA	Norwegian Mapping Authority
NRISN	Norwegian Regional Ionosphere Scintillation Network
NRTK	Network RTK
PE	Position Error
PC	Principal Component
PFA	Principal Factor Analysis
PVT	Position, Velocity and Time.
QoS	Quality of Service
RMS	Root-Mean-Square
ROTI	Rate Of TEC Index
RTK	Real-Time Kinematic
RV	Random Variable
SARIMA	Seasonal ARIMA
SA	Selective Availability
SATREF	SATellite-based REFerence system
SD	Single-Difference
SIS	SATREF Ionospheric System
SLM	Single-Layer Model
SNR	Signal to Noise Ratio
SV	Stochastic Volatility
TEC	Total Electron Content
TS	Time Series
TTA	Time To Alert
UDRE	User Differential Range Errors
UTM	Universal Transverse Mercator
VLBI	Very Long Baseline Interferometry
VPL	Vertical Protection Level
WAAS	Wide Area Augmentation System
WGS84	World Geodetic System 1984

Frequently used symbols

Name	Definition
$\boldsymbol{x}_k$	State vector $(n \times 1)$
$\boldsymbol{z}_k$	Measurement vector $(m \times 1)$
$\hat{\boldsymbol{x}}_{k,k-1}$	Predicted state vector $(n \times 1)$
$\boldsymbol{w}_k$	System noise $(n \times 1)$
$\boldsymbol{\epsilon}_k$	Measurement noise $(m \times 1)$
$\boldsymbol{R}$	Measurement noise covariance matrix $(m \times m)$
$\boldsymbol{Q}$	Process noise covariance matrix $(n \times n)$
$\boldsymbol{\Phi}$	State transition matrix $(n \times n)$
$\boldsymbol{K}_G$	Kalman gain matrix $(n \times m)$
$\boldsymbol{P}_{k,k-1}$	Predicted error covariance matrix $(n \times n)$
$\boldsymbol{P}_{k,k}$	Corrected error covariance matrix $(n \times n)$
$\boldsymbol{H}_k$	Design matrix $(n \times m)$
v_k	Velocity (m/s)
E_l	Satellite elevation angle (rad or deg)
A_z	Satellite azimuth (rad or deg)
α	Integrity risk value
z	Zenith angle (rad or deg)
T_s	Test statistics
T_h	Threshold values
D_h	Decision rules
dbHz	Decibel-Hertz

Definitions related to GNSS

Name	Definition
Accuracy	A measure of how close an estimate of a GPS position is to the true one.
Ambiguity	During the signal acquisition by the receiver on the ground, the initial oscillator phase offset, the so-called ambiguity shall be determined and corresponds to the number of whole carrier cycles in the propagation path.
Clock Bias	The difference between the indicated clock time in the GPS receiver and true universal time (or GPS satellite time).
Clock offset	A constant difference in the time reading between two clocks, normally used to indicate a difference between two time zones.
DOP	Dilution of precision, a measure of the satellite-receiver geometry, describing factors which affects the geometric satellite distributions.
Error ellipse	Uncertainty of the measurements are described as an error ellipse, which has its center at the correct position and the size and the orientation are given by the coefficients of the ellipse equation.
Fixed effect	Unknown constant that we try to estimate from data.
GPS	A global system based on 24 satellite orbiting the earth at an altitude of 12000 statute miles and providing precise worldwide positioning and navigation information 24 hours a day.
Kalman Filter	Recursive data processing algorithm.
Magnetic North	Direction of the magnetic north pole relative to the observers position. In other words, the direction that a compass points.

Name	Definition
Multipath	Interference caused by reflected GPS signals arriving at the receiver.
Navigation Message	The message transmitted by each GPS satellite containing system time, clock correction parameters, ionospheric delay model parameters, and the satellite's ephemeris data and health. The information is used to process GPS signals to give the user time, position, and velocity. Also known as the data message.
Outlier	Outliers are usually not just biased observations, but rather artifacts caused by filters, instruments, and other chaotic phenomena. They significantly deviate from the distribution of regular observations, and this makes them straight-forward to eliminate.
Particle Filter	Recursive data processing algorithm for non-stationary and non-Gaussian processes based on simulation.
Position	An exact unique location based on geographic coordinate system.
Precision	A measure of how close an estimate of a GPS position is to the sample mean.
Random effect	A random variable and we try to estimate the parameters that describe the distribution of this effect.
Residual	Difference between observation and prediction.
Smoothing	Backward processing algorithm.
True North	The direction of the north pole from your current position. Magnetic compasses indicate north differently due to the variation between true north and magnetic north. A GPS receiver can display headings referenced to true north or magnetic north.
Troposphere	The lowest region of the atmosphere between the surface of the earth and the tropopause, characterized by decreasing temperature with increasing altitude.
WGS-84	World Geodetic System 1984, the primary map datum used by GPS. Secondary datums are computed as differences from the WGS-84 standard.

Publications

As a part of my doctoral research, I have also contributed to a number of self-contained research papers. The relevant papers are listed below, and the full manuscripts have also been enclosed at the end of this thesis.

Paper #1:
H.P. Kierulf, M. Ouassou, M.J.R. Simpson & O. Vestøl
A continuous velocity field for Norway
Journal of Geodesy **87(4)**, 337–349 (2013)

Paper #2:
M. Ouassou, A.B.O. Jensen, J.G.O. Gjevestad & O. Kristiansen
Next generation network real-time kinematic interpolation segment to improve the user accuracy
International Journal of Navigation and Observation, 346498 (2015)

Paper #3:
M. Ouassou, O. Kristiansen, J.G.O. Gjevestad, K.S. Jacobsen & Y.L. Andalsvik
Estimation of scintillation indices: a novel approach based on local kernel regression methods
International Journal of Navigation and Observation, 3582176 (2016)

Paper #4:
M. Ouassou, A.B.O. Jensen & J.G.O. Gjevestad
Network real-time kinematic data integrity by means of multivariate statistical analysis
Submitted to International Journal of Navigation and Observation: 06/11/2017

Paper #5:
M. Ouassou, B. Natvig, A.B.O. Jensen & J.I. Gåsemyr
Reliability Analysis of Network Real-Time Kinematic
Submitted to International Journal of Navigation and Observation: 30/11/2017

.

Contents

1 Introduction

"All in all, doing data analysis is simply about calculating eigenvectors; all the science and art in it is about finding the right matrix to diagonalize."

— Jean-Paul Benzecri (1973)

This chapter provides an introduction to the thesis, starting off with the motivation behind the study, and discussing the specific problems that we wish to address. The subsequent sections outline what an appropriate solution should look like, and how this thesis contributes to designing and implementing such a solution. The last section of the chapter gives a rough overview of the organization and contents of the rest of the thesis.

1.1 Background materials

One inherent weakness of the GNSS Network Real-Time Kinematic (NRTK) software supplies is that the system does not provide the users with a quality indicator for the transmitted signal. This type of information is missing today and routinely requested by the professional users of the GNSS data. This piece of information is referred to as NRTK data integrity.

Our aim is to extend the NRTK service with a new layer that is able to inform the user in the field when a faulty satellite or individual corrections should not be involved in the solution computation. In addition, an integrity solution is a set of control procedures that notifies the user in the field when the positioning service is unreliable.

1.1.1 Related research

In order to ease the discussions, the integrity-related definitions are given below and are originally taken from the European Geostationary Navigation Overlay System (EGNOS) [135].

Definition 1. ***Integrity****: Is a measure of the trust that can be placed in the correctness of the information supplied by the total system. Integrity includes the ability of a system to provide valid and timely warnings to the user, known as alerts, when the system must not be used for the intended operation.*

The integrity requirements are expressed quantitatively with three parameters:

Definition 2. ***Integrity Risk****: Is the probability that an error caused by the system causes a computed position error (PE) to exceeds the alarm limit (AL) without providing the information to the user within the specified time to alert.*

Definition 3. ***Time to alert****: The time is from the occurrence of the system error, causing the computed position error to exceed the alarm limit to the display of the alarm at the user's.*

Definition 4. ***Alarm limit****: Is the maximum allowable error in the user's position solution before an alarm must be given within the specified TTA.*

Definition 5. ***Protection level****: The protection level (PL) is the worst-case predicted position error based on geometry and user range error (URE).*

The integrity of the user position comes with two values, namely the horizontal protection level (HPL) and vertical protection level (VPL). The VPL tells us by how much the position in worst-case could be off vertically and HPL says the same about horizontal position.

From definitions 1–5, we can conclude that the computation of the integrity can be parametrized by three parameters: test statistic T_s, decision threshold D_h and protection level α, respectively.

1.1.2 GNSS augmentation systems integrity solution

GNSS augmentation systems were developed specifically for applications where safety of life is an absolute priority, for instance the aviation, maritime and land users. These systems include WAAS, GAGAN, EGNOS and others. Failures and system malfunctions must be detected, repaired if possible and reported to the users. The information provided by such systems can be grouped into three main components:

1. extra ranging sources using GEO communication satellites;
2. a vector of corrections to the GPS signal-in-space, including the ionospheric path delay, clock and ephemeris corrections;
3. integrity monitoring function to alert the users of out of tolerance operations.

Since the ionosphere path delay is a major threat to navigation in nominal operation, the augmentation system processing facilities calculate the error bounds for ionospheric corrections, called Grid Ionospheric Vertical Error (GIVE), and a combined error bound for the clock and the ephemeris corrections called the User Differential Range Errors (UDRE). The correction vector is broadcast to the users via geostationary satellites and used

to compute the confidence bounds of the position errors VPL and HPL, respectively.

Anyway, the user needs to know the average error level and the corresponding uncertainty of parameters that have direct influence on position solution degradation.

Interesting literature on how to estimate the ionospheric error bounds for WAAS [22, 27], guaranteeing the integrity to all users in the coverage area [91], general information on WAAS including message types [97] and detailed algorithms on WAAS performance standards [57].

Test statistics for exceeding events determining the VPL and HPL are based on the Markov or Chebyshev inequalities [94, p. 77]. These inequalities enable us to compute the error bounds on probabilities when only the average error level (the mean error function) and the associated uncertainty (the variance function) measured in terms of standard deviation are known. The probability distribution is assumed to follow the Gaussian distribution with a finite variance σ^2.

Let us use the notation $\boldsymbol{p}_e$ for the computed position error, by α the integrity risk value and by κ the protection levels VPL or HPL, respectively. Then the integrity equation is given by Eq. (1.1).

$$\mathrm{P}\Big[\|\boldsymbol{p}_e\| \geq \kappa\Big] \leq \alpha \tag{1.1}$$

A more powerful version of the above is given by the Borell-Tsirelsen-Ibraginov-Sudakov inequality, which gives a universal bound for the excursion probability, Eq. (1.4). For interesting discussions on the computation of the excursion probabilities of Gaussian processes, the reader is referred to [2, Chap. 2].

1.1.3 GNSS receiver-based integrity solution

The next type of integrity solution is the receiver-based measurement integrity. The algorithms developed for this purpose are receiver autonomous integrity monitoring (RAIM) family and fault detection and exclusion (FDE) algorithms. For more details about these methods, we refer the reader to e.g. Kaplan and Hegarty [54, pp. 346-360], Prasad and Ruggieri [88, pp. 62-65], and Grewal et al. [39, pp. 106-201]. These algorithms are developed by and used to support primarily aviation applications that are safety-critical in nature.

These approaches simply detect and exclude unhealthy satellites from the computations; in other words, the receiver ensures good measurements at the cost of some information. However, when multiple unhealthy satellites are observed, they are statistically likely to cancel each others effects. This means that the receivers throw away information that can be quite valuable

when processed using the right statistical techniques. Due to this fact, the integrity drawn from augmentation systems is typically considered more available than receiver-based integrity [54, Section 7.5].

The inputs to the RAIM algorithm are the standard deviation of the measurement noise, the measurement geometry, and the maximum allowable probabilities for a false alert and a missed detection. The output of the algorithm is the horizontal protection level (HPL), which is the radius of a circle centered at the true aircraft position that is assured to contain the indicated horizontal position with the given probability of false alert and missed detection.

During the last decade, a lot of efforts and research have been conducted to improve the RAIM and FDE algorithms. Tsai [134] proposed a moving average filter to perform satellite failure detection and exclusion by operating on residuals, while Wang [137] proposed an improved particle filter (Section 2.7.4) using a genetic algorithm to compute the weights.

Combining different navigation systems, for instance the GPS, GLONASS, Galileo and Beidou, the user can obtain a reliable position solution by applying RAIM/FDE robust algorithms. These types of data integrity are very economical compared to augmentation systems and I believe they will dominate in the future to provide the GNSS integrity solution. That is, with a lot of satellites in view, the user can discard the ones with bad data and still preserve a good geometry in solution computation.

Three RAIM methods have been proposed to implement GPS data integrity. These are:

- Range comparison method proposed by Lee [66].
- Least-square residual method proposed by Parkinson and Axelrad [85].
- Parity method proposed by Sturza [119].

For simplicity, we introduce the functionality of the RAIM algorithm based on least-square residuals for illustration. Assume that a number n of satellites in view and the error sources are uncorrelated between the different line-of-sight. One starts out by linearizing the measurements, and the model is given by Eq. (2.35). The linearizion process take place around the estimate minus actual position deviation vector $\boldsymbol{\beta}$, which is 4-dimensional (one temporal and three spatial coordinates).

The error along each satellite line-of-sight is taken to be zero-mean Gaussian noise with finite variance σ_i^2. The resulting position estimator can be expressed as:

$$\hat{\boldsymbol{x}} = (\boldsymbol{H}^T \boldsymbol{W} \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{W} \boldsymbol{z} = \boldsymbol{R} \boldsymbol{z} \tag{1.2}$$

where $\boldsymbol{R}$ is the weighted pseudo-inverse of the design matrix $\boldsymbol{H}$, and $\boldsymbol{W}$ is the precision matrix, defined as the inverse of the measurement noise covariance matrix Σ. Often $\Sigma^{-1} = \boldsymbol{W}$ is referred to as the weighting matrix and is a diagonal $n \times n$ matrix.

LPV-200 is the level of service based on the most demanding flight operation delivered by the WAAS system. The Federal Aviation Administration (FAA) navigation requirement for LVP-200 approaches (150 sec) set the probability α in Eq. (1.1) to be 10^{-7}. Since the vertical component x_v is critical for aviation, the computation of VPL confidence interval reads:

$$\text{VPL} = x_v \pm k(\alpha) \cdot \sigma_v \qquad \sigma_v^2 = \left[(\boldsymbol{H}^T \boldsymbol{W} \boldsymbol{H})^{-1} \right]_{3,3} \tag{1.3}$$

$k(\alpha) \sim 5.33$ for integrity risk value $\alpha = 10^{-7}$

1.1.4 GNSS reliability monitoring

The quality checks to assess the correctness of the model is achieved through reliability and precision analysis. Here, reliability refers to the trustworthiness of the observations. This is assessed using redundant observations to detect and identify any model errors (internal reliability), while also expressing the influence of undetected model errors (external reliability). Precision, on the other hand, is characterized by the variance in the observations. In any case, the procedure is straight-forward: we construct test statistics from the model residuals, and use rigorous hypothesis testing to detect and estimate problems with our model.

The elementary building blocks of the GNSS reliability monitoring are: minimal detectable bias (MDB), minimal detectable effect (MDE), and bias to noise ratio (BNR). For deeper insight on this subject, the reader is referred to Leick [68, Chp. 4.10], Kuusniemi et al. [63], Baarda [3] and de Jong et al. [53, Section 3.2].

Anyway, the key in reliability analysis are the improved stochastic model by imposing the variance function based on elevation angle (E_l) and check for blunders based on Detection Identification and Adaptation (DIA) principle [123].

1.1.5 Network RTK data integrity solution

The integrity solutions presented in sections 1.1.2 and 1.1.3 are not directly applicable for carrier-phase based positioning. These methods use the code pseudo-range observables and the user receiver has to compute the traveling time of the signal transmitted by the satellites. In contrast to users that require centimeter-level accuracy, the use of the carrier-phase observables is

necessary, and the method is based on the determination of the difference in phase of the carrier transmitted by the satellite and the signal replica generated by the receiver itself. The main obstacle of this method is the presence of ambiguity.

However, RAIM and FDE were developed as pseudo-range residual data analysis algorithms for GNSS safety-critical applications, such as e.g. the approach phase of flight. For high-accuracy applications, an extension of pseudo-range RAIM (PRAIM) known as carrier-phase based RAIM (CRAIM) was proposed by Feng et al. [31].

NRTK services provide GNSS users with a centimeter-level accuracy and is becoming indispensable tool for many users in various fields. The aim of the NRTK method is to minimize the influence of the correlated biases on the rover position computation within the bound of the average error level provided by the network. Often, these biases are referred to as distance-dependent errors. More precisely, these can be divided into the atmospheric (ionosphere and troposphere path delays) and the orbital errors.

Due to the fact that the atmospheric error behaves in a complex way and cause the position parameters to vary drastically from epoch to epoch, new methods have recently been developed to build better spatio-temporal models. This topic is treated in Section 2.8.

Li et al [70] introduced an elegant method to decorrelate the troposphere path delay from the height component, assuming that ambiguities are correctly resolved. This method improves the altitude estimate.

Several NRTK techniques exist and the most commonly used at present are for instance the Master Auxiliary Concept (MAC) [30, 122], the Virtual Reference Station (VRS) concept [64], the FKP techniques [141] and the Network Adjustment (NetAdjust) concept developed by John F. Raquet [89, 90].

The NRTK data integrity solution developed in this thesis will be independent of the method used for generation of the NRTK corrections. However, for derivation and development of the integrity and quality control algorithms we have based our work on the conceptual approach as given by the *NetAdjust* method. Most of the NRTK techniques mentioned above are developed commercially and details about these are not readily available. But the *NetAdjust* method is well described in literature, it is therefore suitable as a starting point for our work.

A new concept to check the correctness of the information provided by systems employing the carrier-phase measurements is needed. In recent years, the data integrity has become a big issue. For PPP applications, the integrity solution using the GPS and GLONASS data was introduced by Jokinen [52]. For railway applications, for instance train control systems, integrity monitoring was introduced by Capua [15].

Anyway, the right place to evaluate the quality of the rover receiver's position solution is at the receiver itself. This includes a check of correctness of the information provided by NRTK system and the local effects. The local effects refers to any disturbances affecting the rover receiver, for instance multipath, scintillation and others. Therefore a realistic stochastic model also has to account for such local effects.

The main topic of this thesis is to design and implement such algorithms at the system and the user levels in order to improve the rover position. The users of high accuracy GNSS NRTK positioning systems have requested the development of data integrity for a long time. In this dissertation, we consider how such a service can be designed and implemented, which can be of interest to both the NRTK service providers and their users.

1.2 Motivation

GNSS are primarily used for estimating the position, velocity and time for a user, also known as their PVT and it is very popular in navigation and is fundamental in surveying. During the past decade, these navigation systems have become indispensable research tools, and have been put to widespread use in fields such as surveying, machine guiding, and geophysics. New applications for the technology are discovered all the time, so this is still a field in rapid growth.

The NMA, who are funding this study, maintains and operates a real-time positioning service called CPOS. This service generates NRTK corrections for users of both GPS and GLONASS, and allows centimeter-level positioning in most of the mainland of Norway. The users are primarily the companies and organizations that build and maintain the Norwegian infrastructure, e.g. those involved in road building, rail-way maintenance, construction, surveying, and machine guiding.

The requirements regarding accuracy, coverage, reliability, and operational security are ever-increasing, and the positioning service is expected to be available in real-time. The number of users is also increasing drastically. For instance, if the current growth rate continues, a tenfold increase in the number of CPOS users is expected over the next decade. These two developments implicate that the positioning services of the future will have to, in real time, provide higher accuracy positioning to more users.

Figure 1.1 shows the number of CPOS users connected at the same time and the historical data of the users using the CPOS services in Norway. Both curves increase exponentially in a period of one decade, and if the development continues to follow the same pattern, the existing tools will not be sufficient to process the large data sets.

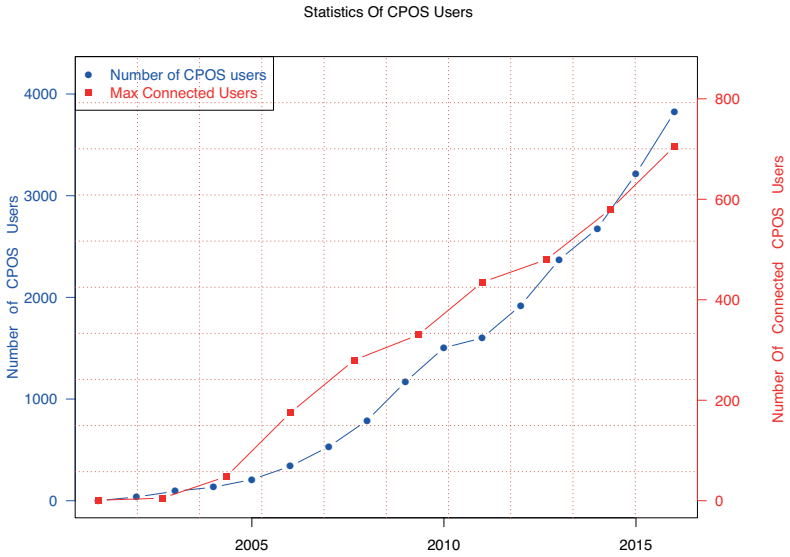


Figure 1.1: During the past couple of decades, the total number of CPOS NRTK users (blue) has increased from none to roughly 4000, while the number of simultaneously connected users (red) increased to about 800.

The users are also becoming more conscious of the quality of their positioning information. Although they may have quite different needs when it comes to accuracy and continuity, it is essential to know whether or not the data is good enough for their needs. The navigation systems of today are insufficient in this regard; e.g. many field workers have to wait until post analysis before they discover whether or not they have to make new measurements. The navigation systems of the future should allow the clients to generate a real-time estimate of the error ellipsoid around the user; in other words, an *integrity solution* for NRTK users is required.

1.3 Problem statement

The last type of integrity solutions are those based on an NRTK (Section 1.1.5), and this is the topic of this thesis. In other words, we will not attempt to redesign the GNSS receivers, and neither will we enhance the available NRTK correction algorithms; our sole concern in this thesis is to perform quality-control on the NRTK corrections generated by existing software solutions. When it comes to high-precision navigation and posi-

tioning, real-time services can offer substantial improvements in efficiency (for instance, observers in the field can immediately be provided with an estimate of the signal quality and resulting error ellipsoid). Extra information that says something about the quality of signal, ensure the quality of work in the fields with a major time saving.

Figure 1.2 shows both radical and incremental product development. The concept starts with the market needs. In our situation, the users of CPOS services require e.g. accuracy, availability, integrity, stability, and reliability of the system. These performance parameters define the functionality of the product to be developed. Assumptions are taken under the development time-span. You can't get what you want, there are limitations for instance the existing technology is not good enough to develop the product. The red lines indicate the interactions between these two categories and are limited by existing resources and technology in order to deliver the final product. In our case, we would like to add layer on top of the NRTK service, *NRTK data integrity solution*.

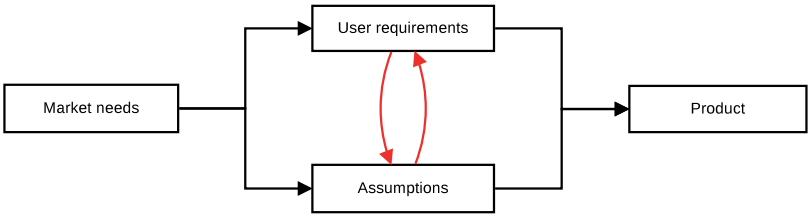


Figure 1.2: This figure depicts the relationship between the factors that affect product development.

1.4 Thesis objective

This thesis introduces a novel and effective optimal alarm system for GNSS NRTK carrier-phase users, capable of detecting, identifying and adapting extremal events that can cause the degradation of the user accuracy in the field. It is a new layer, performing the data integrity on different levels, the system and the user level. This layer can be regarded as an extension of GNSS NRTK conventional processing with responsibility of checking the correctness of information. The specific objective of this work can be formulated as follow:

- *Network correction data integrity*: On satellite basis, the typical average error level and the associated variance are determined. Mahalanobis distance takes into consideration the correlation and is used as a global

test for detection of extremal events, then the proposed local tests (normalized residuals, t-internally/externally standardized) are used interchangeably for detection. For adaptation, an anomaly satellite list is obtained.

The variance-covariance matrices are inspected for Heywood cases, i.e. situations with negative variance. This is achieved by monitoring the Kalman filter gain and the trace of variance-covariance matrices.

- *Rover raw data integrity*: A filter based on an empirical signal-to noise-ratio (SNR) weighting model is implemented to exclude all satellites that degrades the rover accuracy. In this case the variance of SNR is monitored to capture the multipath, scintillation and low elevation.

- *Baseline data integrity*: The double-difference residuals (DD) and corresponding variance-covariance are checked. The procedure is similar to the network data integrity with a minor modification. The length used to determine average error level is the correlation time of DD. The trace of the DD error covariance matrix is used for prediction of carrier phase error statistics.

- *Rover position error data integrity*: The correlation and accuracy are provided by the position error variance-covariance matrix. This information is used to construct the error ellipsoid based on significance error level α.

1.5 Contribution of thesis

Data quality checks and integrity monitoring techniques has been a research topic for many years in geodesy, surveying and navigation. For instance Baarda [3] developed a test procedure for use in geodetic networks, which has been used to check data against outlying observations in many different applications, for instance the analysis of the deformation problem in geodesy [61]. The DIA procedure [123] can be applied to any set of GNSS observation equations, such as the GPS quality control [58], geodetic networks [126] or integrated navigation system [126]. Another approach to error modeling is to perform a reliability and quality control procedure [63], using good statistical methods for the analysis [69].

The right place to help the CPOS NRTK users is in the field by providing them with any needed information to achieve high precision. This is an

ideal situation. Unfortunately, this help is impossible to provide to the CPOS NRTK users. Instead, we carry out quality control at the system and the user levels and generate alarms when the corrections from a specific satellite can not be trusted.

According to Kaplan [54, p. 302], the GNSS solution error can be factorized into a geometric factor (DOP, Dilution Of Precision) and a pseudo-range factor (UERE, User Equivalent Range Error). For more information on the computation of the UERE factor, the reader is referred to Kaplan [54, Ch. 7].

1.5.1 Optimal alarm system

This thesis introduces a novel and effective optimal alarm system of type detection, isolation and decision (DID) for kinematic carrier-phase users. It is able to carry out the quality control at different level of data processing. The NRTK corrections field is a pure spatio-temporal process and the determination of extremal events causing the degradation of the rover position accuracy is based on excursion probability.

$$\mathbb{P}\left\{\sup_{\substack{s\in S\\ t\in T}} f(s,t) \geq T_h\right\} \tag{1.4}$$

where (s,t) is a 4-dimensional vector in space s and time t and T_h is threshold value determining the level crossing events.

In addition, the covariance structure of such processes is too complex to capture and model. Often we end by imposing stationarity and modeling the processes by a separable model. The interested reader on this topic is referred to Gelb [35, p. 84].

1.5.2 Mathematical statistics as an NRTK tool

Regardless of the GNSS system used, the question of whether the information can be trusted or not is a question of parameter estimation and error analysis. Such questions lie at the heart of mathematical statistics, which can therefore be a very useful tool for analyzing NRTK systems.

GNSS observations are affected by observation errors (or uncertainties) and what we do in integrity monitoring is to estimate the reliability of the (NRTK correction or position) solutions given the observations affected by errors. This means that the parameters are stochastic variables and as such are described by statistics.

Advanced statistical methods are needed because GNSS NRTK positioning is rather complex and so are the characteristics of the error sources (e.g. the

ionosphere) and we need robust and reliable solutions. The work described in this thesis applies a number of modern statistical methods for integrity monitoring.

During the last decades there has been a large progress in the field of statistics and many of the new statistical methods have not yet been applied or tested in GNSS data processing. This is for instance the case for spatial-temporal modeling by Integrated Nested Laplace Approximation (INLA), hierarchical modeling for spatial processes, non-stationarity and non-Gaussian processes, Stochastic Partial Differential Equation (SPDE) and Gauss-Markov Random Fields (GMRF).

The current methods presented in sections 1.1.2 - 1.1.4 are using advanced statistics to identify and isolate data which can lead to wrong estimates. The work in this thesis builds upon the existing methods and applies advanced statistical methods in subsequent steps of the NRTK data processing chain to obtain more robust integrity and in the end a better positioning performance at the user level.

1.6 Published articles

There are multiple similarities between the dynamics of the financial market and the GNSS data processing. Both fields experience drastic changes from one second to next, e.g. due to natural disasters or major political events, causing for instance the stock indices to act unstable and create unpleasant situations. GNSS data processing experiences the same problems, rapid variations of the TEC known as scintillation, multipath interference, bad satellite geometry change, or any other phenomena that can cause large variations in the data. Therefore the GNSS data processing may gain a lot by considering the methods and algorithms that have been devised for financial applications.

The main goal of this dissertation is continuous check of the correctness of information provided by the NRTK with respect to the dynamic imposed by the GNSS data processing algorithms. Therefore, I tried to integrate some methods and algorithms from the financial market to GNSS data processing and apply the advanced statistical methods in GNSS data modeling. More precisely, the commonalities include how to control the variations of the variance, handling discontinuities, missing observations, spatial-temporal process modeling, and prediction of stochastic processes.

Some considerations around the NRTK data integrity are introduced below as a background for the design process and to ease the discussions in the following sections.

- Ambiguity resolution: The key to precise positioning is the correct

ambiguity resolution and validation. With ambiguities resolved to wrong integer numbers, there will offsets in the position solution, and with float ambiguities (ambiguities that are not fixed to integer values) the position solution is inaccurate and also very unstable and sensitive to changes in satellite geometry.

- Geometry: Good satellite-receiver geometry, as for instance expressed by the so-called DOP factor (dilution of precision) is important to perform successful ambiguity resolution and achieve centimeter level accuracy in real time.
- Covariance structure definition: Spatio-temporal models that describe well the variations of the spatially correlated errors in the corrections field is also an important key for reliable NRTK positioning.
- Large data sets: Robust estimation algorithms to handle large data sets are also a key factor becoming more important in the future as observations from several GNSS systems to a larger degree will be combined in one processing loop. So, we are experiencing a computational paradigm shift. Today, most NRTK systems operate with data from the American GPS and the Russian GLONASS system. Including data from the European Galileo as well as the Chinese Beidou systems in NRTK operations will soon be the norm for most NRTK services. With satellites from more GNSS systems being available the satellite-receiver geometry on the rover side is improved. This is especially important when the user is operating in restricted environments such as narrow street canyons or forest areas.
- Non-stationarity and non-Gaussian processes: The main challenge here is how to construct better spatio-temporal models to monitor spatial-temporal processes that are not stationary and non-Gaussian. The combination of stochastic partial differential equation (SPDE) and Gauss-Markov Random Fields (GMRF) offer flexibility to generate a good result with minimal processing time. That is, robust against the big-n problem .

In order to meet the need of the GNSS NRTK data processing today and in the future, the key is to use the most advanced methods of mathematical statistics that are available to date. For instance, construction of better models for spatio-temporal processes (Paper #2 of this thesis), imputation algorithms for spatio-temporal processes (Paper #1 of this thesis), constructing reliable test statistics based on some well-defined metrics (Paper #4 of this thesis), computing the QoS parameters of the NRTK corrections field (Paper #2 of this thesis), reliable scintillation index and more realistic

stochastic models (Paper #3 of this thesis), and identification and amelioration of the weaknesses and the strengths of the NRTK system (Paper #5 of this thesis).

We end our discussions by a simple comparative example. In the thirties, scientists in the fields of physics and chemistry had achieved a lot by discovering new chemical elements. After that, we entered a new period of technological progress, which changed our way of life by combining the newly discovered elements and constructing something new, for instance electronics communication devices, nuclear weapons, and other interesting devices.

In our case, we already have enough mathematical tools to do the job in NRTK data integrity solution, so we do not need to develop new ones. These tools can be considered as elementary building blocks in the hands of the data analyst or modeler. The main challenge is to know what tools exist, what they can do for us, what their strengths and weaknesses are, and how to combine them in appropriate ways to describe the observed variations as well as possible.

All published articles are based on this principle, and use the existing advanced mathematical statistics to construct something new.

1.7 Thesis outline

The six chapters of this dissertation are organized in this order.

1. Chapter 1: Introduces the background, objective and main contribution of this work, as well as the outline of dissertation.

2. Chapter 2: An introduction to the mathematical statistics used in this dissertation. This includes model selection, spatial-temporal processes, hypothesis testing, and multivariate statistical analysis. The reader can either read this chapter first as an introduction to the field, or skip it initially and then use it as a reference when needed.

3. Chapter 3: An overview of GNSS error sources related to the use of network of reference receivers. Understanding the nature of the GNSS errors present in measurement are absolutely necessary. Included are definitions of GNSS observables, differencing techniques, linear combinations of observables, and GNSS error modeling and mitigation.

4. Chapter 4: Double-difference of linear combination of carrier phase measurements are formed, the ambiguities are resolved, the covariance structure based on kriging are computed, the noises are empirically determined and the user corrections are computed.

5. Chapter 5: Elementary main building blocks to design the GNSS NRTK data integrity solution will be introduced and is the core of the articles 2 and 4. This chapter can be read together with article 4 as a supplement. Some of the material is new, and some is just summarized.

6. Chapter 6: Summary of the results of the GNSS NRTK data integrity solution and the recommendations for further research are presented. How to standardize the final product is described. This chapter can be considered as interface to the author's work in this dissertation.

The core of this dissertation are the chapters 1, 5, and 6. Chapters 2–4 are optional and can be used as reference.

2 Mathematical background

"The best time to plan an experiment is after you have done it."
— R.A. Fisher (1962)

This chapter introduces the GNSS data processing tools and the basic mathematical background applied in this dissertation. This includes the parameter estimation techniques, modeling in space and time, forecasting, prediction, and hypothesis testing.

Section 2.1 introduces in general the problems that need to be addressed seriously, for instance, the covariance structure, stationarity, and massive data processing. In addition, the basic statistical methods are introduced. Section 2.2 introduces the test statistics including the normality tests, serial correlations and test for randomness. Model selection and hypothesis testing are the topics of the Sections 2.3 and 2.4.

Least square adjustment techniques have been used in GNSS data processing for a long time. Section 2.5 outlines the general regression model, Gauss-Markov model (GMM) and parameter estimations in general, while Section 2.6 introduces a class of estimations known as penalized methods.

Section 2.7 handles the Bayesian statistics, including the most used GNSS data processing tool, the Kalman filter. Handling the non-linear and non-Gaussian processes are subject of Section 2.7.4.

Section 2.8 handles spatial-temporal processes. This class of processes play a key role in NRTK corrections generation, prediction and covariance structure definition.

Finally, Section 2.9 outlines the concept of the multivariate statistical analysis (MVA). MVA is chosen as the implementation tool of the NRTK data integrity in this dissertation.

2.1 Introduction to the field of statistics

A key problem in NRTK data processing is the estimation of the distance-dependent errors, composed of a portion of ionosphere and troposphere path delays and the satellite position errors. Due to the limitations imposed by these biases, the user needs to work within a short distance from the reference receiver to obtain centimeter-level position accuracy. These biases can be modeled as a pure spatio-temporal process. In addition, NRTK data processing faces new challenges: (1) realistic models for massive data sets; (2) parameter estimations of non-Gaussian and non-linear processes; (3)

determination of valid covariance functions; (4) handling non-stationarity of spatial-temporal processes. (5) quality check in form of accuracy and reliability.

- **Massive data sets**: Since new navigation satellite systems (e.g. Galileo, Beidou) in addition to the old GPS and GLONASS become operational, a massive data set needs to be processed in real-time, so we are experiencing a computational paradigm shift. Building realistic spatial-temporal models for huge data sets require integrations over hundreds or even thousands of unknown parameters, and is computationally expensive. Banerjee [4] introduced the hierarchical Bayesian methods, and where the dependency between parameters are organized in hierarchical manner, where the level describes the dependency.
  H. Rue [98] introduced the Integrated Nested Laplace Approximation (INLA) computation machinery to approximate the Gaussian field by Gauss-Markov random field (GMRF).

- **Non-stationarity for spatial processes**: Capturing the covariance structure of the process under study in NRTK is a very complex and challenging task. Often one chooses a short-cut by imposing stationarity even if the process is non-stationary. Section 2.8 covers this topic.

- **Covariance structure**: Valid covariance functions must satisfy the Bochner's theorem [102, p. 487]. We can construct a complex valid covariance function by linear combinations of existing ones. Section 2.8.2 gives a list of valid covariance functions commonly used.

- **Quality checks and reliability**: Checking the correctness of the information provided by the NRTK is the core of this thesis. Details are given by articles on NRTK data integrity (Paper #4 of this thesis) and NRTK data reliability (Paper #5 of this thesis), respectively.

In conclusion, our aim is always directed to improve the rover position accuracy in the field and to provide the user with quality indicators. This will be accomplished by building better models describing the variations in the data, mode identifications, covariance structure definitions, better large data processing algorithms and monitoring the extremal events. We are always in the position of deciding how to solve the conflict between cost and accuracy in the best way. GNSS data processing requires two main questions to be answered on epoch-by-epoch basis.

- How large is large ?
- How much where ?

2.1.1 Elementary statistics

GNSS data processing requires at least a minimum knowledge in the field of statistics. This means that mastering how to build a test statistics and carry out the hypothesis testing, normality and randomness testing, time-series and spatial-temporal processes modeling and prediction are necessary to carry out the data quality and integrity assessment of any system.

Our aim is to measure the variability and the examine the centrality of the data. Let our observations be presented by $\{\boldsymbol{z}\}_{i=1}^{N}$, then the sample mean and variance are given by $\bar{\boldsymbol{z}} = n^{-1}\sum z_i$ and $\boldsymbol{S}^2 = (n-1)^{-1}\sum_{i=1}^{n}(z_i - \bar{z})^2$, respectively.

The variance is a measurement tool to measure the variability with respect to centrality indicators such as the mean or the median. The most used tool for variability measurement is the standard deviation, defined as the square root of the variance and is expressed in the same unit as the mean and the median. Contrarily, the variance is expressed in square units.

2.1.2 Central Limit Theorem (CLT)

One the most important applied theorem in statistics is the central limit theorem [67]. The concern of CLT is directed to a sum of a large number of random variables, which has a normal limit distribution.

Theorem 1. *Classical CLT: Let* $\{Z_i\}_{i=1}^{n}$ *be identical, independent, distributed with* $E(Z_i) = \mu$ *and* $Var(Z_i) = \sigma^2 < \infty$. *Then*

$$\sqrt{n}(\bar{Z} - \mu)/\sigma \quad \xrightarrow{L} \quad \mathcal{N}(0,1) \tag{2.1}$$

or equivalently

$$\sqrt{n}(\bar{Z} - \mu) \quad \xrightarrow{L} \quad \mathcal{N}(0,\sigma^2) \tag{2.2}$$

When n increases, the corresponding density becomes more and more symmetric and closer to the normal distribution. Working with Gaussian distribution it is easy to carry out the inference about the mean or the variance of a population and construct the confidence regions of the test statistics.

Thanks to de-Moivre (1733) and Laplace (1810), the CLT allows us to approximate the probability as the area under the standard normal curve between a and $-a$.

$$P\left[\sqrt{n}|\bar{z} - \mu|/\sigma \leq a\right] \tag{2.3}$$

Reader interested in the CLT with application to different distributions are referred to Lehman [67, pp. 73–75].

2.2 Test statistics

In general, under any model building, the noise component after decomposition of time-series is expected to follow the Gaussian distribution, with a finite mean μ and variance σ^2. If not, work has to be done to identify the hidden signal. This operation is referred to as a test for serial correlation and is the subject of Section 2.2.4.

Assumptions about normality of observations are commonly used in data analysis by the modeler and help to carry out the inference and hypothesis testing. Most test statistics, for instance the Student t-test and Fisher F-test, rely on normality assumptions. Our aim is to assess the deviation from the Gaussian distribution.

We start by formulating the null and the alternative hypothesis, and most commonly normality tests based on different principles are presented.

Let $\{\boldsymbol{z}_i\}_{i=1}^n$ denote the observations, and let μ_z and σ_z^2 denote the mean and the variance, respectively. Let $\bar{z}$ denote the sample arithmetic mean.

Define $\mu_j = \mathbb{E}(z_i - \mu_z)^j$, so $\sigma_z^2 = \mu_2$. Note that the symmetry of the probability density function (PDF) is often measured by the coefficient of skewness, γ_1, where $\gamma_1 = \mu_3/\mu_2^{3/2}$. The shape parameter is denoted by the coefficient of kurtosis, $\gamma_2 = (\mu_4/\mu_2^2) - 3$.

2.2.1 Jarque-Bera normality test

The Jarque-Bera normality test [49] uses the third and the fourth moments about the mean known as the skewness and kurtosis, respectively, to assess the deviation from normality. The test statistics $T_{\mathbf{JB}}$ reads:

$$T_{\mathbf{JB}} = \frac{n}{6}\left\{g_1^2 + g_2^2/4\right\} \tag{2.4}$$

where g_1 and g_2 are sample skewness and kurtosis, and are given by the two expressions, respectively.

$$g_1 = \frac{n^{-1}\sum_{i=1}^n (z-\bar{z})^3}{[n^{-1}\sum_{i=1}^n (z-\bar{z})^2]^{3/2}} \tag{2.5}$$

$$g_2 = \frac{n^{-1}\sum_{i=1}^n (z-\bar{z})^4}{[n^{-1}\sum_{i=1}^n (z-\bar{z})^2]^2} \tag{2.6}$$

Note that g_1 and g_2 are both asymptotically normal distributed. So, we can conclude that the Gaussian distribution will have skewness equal to zero and kurtosis equal to three. According to the large sample theory, the test

statistic $T_{\mathbf{JB}}$ has an asymptotic χ^2-distribution with two degrees of freedom (i.e, $T_{\mathbf{JB}} \mapsto \chi^2_2$). Since the distribution of the test statistic $T_{\mathbf{JB}}$ is known, we carry out the hypothesis testing of normality. At significance level α, we reject the null hypothesis H_0 if the computed value of test $T_{\mathbf{JB}} > \chi^2_{2,1-\alpha}$.

In order that the test $T_{\mathbf{JB}}$ shall work properly, various influential factors must be investigated, for instance outliers. It is well known that the sample moments are influenced by outliers, likewise the $T_{\mathbf{JB}}$. Missing observations, changes in the mean, in variance, sample size, and discontinuities can cause the failure of the test. In addition, the test is poorly suited to small sample sizes.

Improvement of the test can be achieved by Monte Carlo simulation [75], using different definition of the sample skewness and kurtosis. In the presence of outliers, a robust modification of the rescaled moments test for normality is given by Gel and Gastwirth [34], utilizing a robust measure of variance.

2.2.2 Kolmogorv-Smirnov test

The empirical test statistic $D = T_{\mathbf{KS}}$ is based on the empirical distribution function (EDF). The interested reader of the original work of the Kolmogorov's 1933 paper is referred to [117]. Kolmogorv-Smirnov test statistics are used to determine if the observations follow a certain distribution.

For independent and identically distributed variable $\{z_i\}_{i=1}^n \sim i.i.d\ \mathbb{N}(0, \sigma^2)$, parameter vector $\boldsymbol{\theta}$ will typically be unchanged if we permute indices. The order z has no significance and what is matters is how the data is distributed. Let $\{z_i\}_{i=1}^n$ be a sequence sampled from a stochastic process. The empirical CDF is a discrete distribution defined by:

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{z_i \leq x\}} \tag{2.7}$$

where $F(x) = P\{z_i \leq x\}$ is the CDF. We have the following results:

- The distribution of $n\hat{F}(x) \sim \text{Bin(n}, F(x))$ follows the binomial distribution, due to the following facts:
  - we have n independent observations.
  - only two possible outcomes ($<$ or $>$).
  - the probability $F(x) = P\{z_i \leq x\}$ remains constant.
- The variance computation. For an arbitrary $x \in \mathbb{R}^d$.

$$\mathbb{E}(\hat{F}(x) - F(x))^2 = \frac{F(x)(1 - F(x))}{n} \tag{2.8}$$

This leads to that the convergence in probability is in order of $-\frac{1}{2}$, that is $\hat{F}(x) - F(x)) = \mathbf{O}_p(n^{-1/2})$

2.2.3 Quantile-Quantile plot

The Q-Q plot allows comparison of two sets, better than the median or the mean comparison. It is often used to test for normality. It is a visualization tool between the theoretical and empirical distributions.

2.2.4 Serial correlation

After fitting the model to the data and estimating the parameters, we expect that the residuals follow the Gaussian distribution with zero mean (uncorrelated) and finite variance. Note that the residuals are not explained by the selected model and are the most diagnostic method used to construct the goodness-of-fit test. Plotting the residuals is absolutely a useful technique for detecting any unexpected patterns in the data, and can be regarded as a validation process of the selected model. Clive [73, p.25] gives a summary for plotting the residuals. Visual inspection of the residuals by plotting w_i versus w_{i-1} is a good start. If serial correlation is detected, then the hypothesis test shall be constructed, for instance the Durbin-Watson test.

$$w_t = \phi\ w_{t-1} + v_t \qquad v \sim \mathcal{N}(0, \sigma_v^2) \tag{2.9}$$

Note that $(-1 < \phi < 1)$. If $\phi = 1$, the Eq. (2.9) becomes a random-walk process. The hypothesis testing for serial correlation reads:

$$\begin{aligned} H_0 \quad &: \quad \phi = 0, \\ H_1 \quad &: \quad \phi \neq 0, \end{aligned} \tag{2.10}$$

Under H_0, the Eq. (2.9) becomes $w_t = v_t$, while the alternative hypothesis H_1, corresponding to $\phi \neq 0$ means that there is a serial correlations between the error terms w. The Durbin-Watson test statistic $T_{\mathbf{wd}}$ on a data set of size N reads:

$$T_{\mathbf{wd}} = \frac{\sum_{t=2}^{N}(w_t - w_{t-1})^2}{\sum_{i=1}^{N} w_t^2} \tag{2.11}$$

The one-sided alternatives $H_1 : \phi < 0$ or $H_1 : \phi > 0$ can be used to test for negative or positive correlations.

Any test statistic requires the determination of the upper and lower bound known as critical values of the test. For Durbin-Watson test, tables are used to make a decision.

Let d_l and d_u denote the lower and upper bound. We reject H_0 if the value of $T_{\mathbf{wd}} < d_l$ when testing for positive correctness, we accept H_0 if $T_{\mathbf{wd}} > d_l$ and last possibility is to test the case if $d_l < T_{\mathbf{wd}} < d_u$.

2.2.5 Test for randomness

Our aim is to assess if the model residuals are random and independent. There are several tests for randomness.

One way is to inspect the auto-correlations of the residuals for any patterns and large values. According to large sample theory, the auto-correlation of a white noise sequence is independent and approximately normal distributed with zero mean and variance $1/N$, where N is the size of data set. That is $\hat{\rho}(h) \sim \mathcal{N}(0, 1/N)$.

Checking the correlation structure of the residuals are achieved by plotting the $\hat{\rho}(h)$ versus h along with error bounds of $\pm 2/\sqrt{N}$. This procedure can be regarded as a visual inspection of $\hat{\rho}(h)$ to detect the correlation.

The more general test is the Ljung-Box-Pierce statistic, known as the Portmanteau test and is given by the expression:

$$Q = N(N+2)\sum_{h=1}^{H}\frac{\hat{\rho}^2(h)}{N-H} \tag{2.12}$$

where $\hat{\rho}(h)$ denotes the sample auto-covariance function an lag j and H is the number of lags considered in the test statistic Q. The distribution of Q is approximately χ^2 with H degrees of freedom. That is $Q \sim \chi^2_H$. The value of H is chosen arbitrary, typically, $H = 20$.

2.3 Model selection

So far we have assumed that the model is given. The statistician starts typically by visualizing the data to determine whether the selected model is applicable or not. If the model appears to explain variations in the data in a good way, the analysis can provide estimates of parameters, and carry out the prediction.

An alternative is to include $k > 1$ possible models $\{M_1, M_2, \cdots, M_k\}$ in the analysis, hopefully at least one of these k models describe the variation in the data. Selecting a model among a list of candidates is a trade-off between explanation of the variations in the data as much as possible, often referred to as goodness-of-fit and the simplicity of the model.

The statisticians main tasks are:

1. select a model $M \in \{M_1, M_2, \cdots, M_k\}$.
2. estimate the model parameters in M.
3. validation, investigate how good the chosen model is. The data modeler shall include the possibility that none of the models are

good enough to describe the variations in the data in a satisfactory way.

Anyway, a mechanism is needed to use the data and assign a score to each candidate model. This tool functions in the same manner as the one defined in Section 2.6 in estimation.

2.3.1 Akaike's information criterion (AIC)

Comparing the log-likelihood function of the candidate models is not good enough to conclude which model is appropriate to the data. Including more parameters in the model will automatically increase the value of the log-likelihood and we end up selecting the model with the most parameters. Therefore, some punishment strategy is needed, which punishes model complexity in the sense of containing many parameters. Akaike's method aims is to find such models that balance the complexity and presenting the variation in data indicated by high value of log-likelihood. The general formula of AIC reads:

$$\text{AIC(M)} = 2 \quad \text{log-likelihood}_{\max}(M) - \quad 2 \ \dim(M) \tag{2.13}$$

for each candidate model M, where dim(M) is the length of the parameter vector θ.

2.3.2 Bayesian information criterion (BIC)

Inside the Bayesian framework, our main concern in model selection is to pick the candidate model with the highest probability given the data. The BIC of Schwarz (1978) and Akaike (1977, 1978) is similar to AIC, is takes the form of a penalised log-likelihood function where the penalty is equal to the logarithm of the sample size n times the number of estimated parameters in the model. The general formula of BIC reads:

$$\text{BIC(M)} = 2 \quad \text{log-likelihood}_{\max}(M) - \quad 2\log(n) \ \dim(M) \tag{2.14}$$

for each candidate model M. dim(M) and n are the number of parameters estimated in the model and the sample size of the data, respectively. The best model is the one with largest BIC value.

2.3.3 Others information criterion

The AIC and BIC methods allow the data modeler to select a single model from a list of candidates, and this is considered to be the best model. The

selected model shall explain all patterns underlying the data and possibly predict the future data points.

Other methods exist as competitors to AIC and BIC, for instance focussed information criterion (FIC) introduced by Gerda and Nils [19, Chap. 6]. The FIC focuses on parameters of interest, and the best model should depend on them, such as the mean, median, variance, or covariate values. In addition, the FIC allows and encourages different models to be selected for different purposes. For comparison of different information criteria, the interested reader is referred to Gerda and Nils [19, Chap. 4].

2.4 Statistical hypothesis testing

2.4.1 Error types

Statistical hypothesis testing is a process of conducting and reporting the outcome of an experiment with diagnostic evaluation and decision making. Two types of errors can occur. The first is a *Type I error*, that we reject H_0 when it is true. The second is a *Type II error*, we fail to detect that H_0 is false. Figure 2.1 summarizes these errors.

	H_0 true	H_0 false
Accept H_0	None	Type II
Reject H_0	Type I	None

Table 2.1: Error types in statistical hypothesis testing. The diagonal entries, i.e. accepting a true hypothesis or rejecting a false one, are not errors at all. The off-diagonal elements, i.e. rejecting a true hypothesis or accepting a false one, are referred to as Type I and II errors, respectively.

Results of any test statistics depend strongly on two pieces of information:

1. whether the null-hypothesis H_0 was rejected or accepted.

2. the level of significance α at which the test is carried out.

$$\alpha = P(\text{Type I error}) \qquad \beta = P(\text{Type II error})$$
$$= P(\text{Reject } H_0 | H_0 \text{ true}) \qquad = P(\text{Accept } H_0 | H_0 \text{ false AND } H_1 \text{ is true})$$

2.4.2 p-value

The p-value is defined as the observed significance level. A small value indicates there is a strong evidence for the alternative hypothesis H_1 and will lead to rejection of H_0. Mathematically, the p-value reads:

$$p = P(T \leq c_0 | H_0) \tag{2.15}$$

where c_0 is the observed value of the test statistics $T(\cdot)$, and $P(\cdot|H_0)$ denotes the probability under H_0. One major advantage of using p-value is that the hypothesis testing becomes very easy, because the probability tables are no longer necessary. The interested reader can find a procedure on how to compute the p-value for a specific test statistics $T(\cdot)$ in [77, p. 200].

2.4.3 Power of test

The power of the test statistics is defined as the determination of the probability of not making a Type II error, and measures the performance of the test. That is, the ability to detect a false null hypothesis and is defined as one minus the probability of II error.

$$\text{Power} = 1 - \beta \tag{2.16}$$

2.4.4 Generalized likelihood ratio test (GLRT)

GLRT is a statistical tool used in hypothesis testing and to define a rejection region $\mathbb{R}$. The procedure is defined as follow:

If we let our data be presented by a $\{z_i\}_{i=1}^N$ and Θ is a parameter vector, we will consider two hypotheses:

$$\begin{aligned} H_0 &: \quad \Theta \in \Omega_0 \\ H_1 &: \quad \Theta \in \Omega_1 \end{aligned} \tag{2.17}$$

The ratio of the two likelihoods measures the closeness of the two hypotheses H_0 and H_1. H_0 is the reduced model while H_1 is considered as the complete model.

Definition 6. *Generalized likelihood ratio*
Let $\{Y_i\}_{i=1}^N$ be a random sample from $f_Y(y;\boldsymbol{\theta})$. The generalized likelihood ratio, Λ, is defined to be

$$\Lambda = \frac{max_{\Omega_0} L(\boldsymbol{\theta})}{max_{\Omega_1} L(\boldsymbol{\theta})} = \frac{L(\hat{\Omega_0})}{L(\hat{\Omega_1})} \tag{2.18}$$

Λ is a random variable because it is defined as a function of random sample.

Definition 7. *Generalized likelihood ratio test*
A Generalized likelihood ratio test (GLRT) is the one that reject, H_0, whenever $0 < \lambda < \lambda^\star$, where $\lambda^\star$ is chosen so that

$$P\{0 < \Lambda < \lambda^\star | H_0 \text{ is true}\} = \alpha \tag{2.19}$$

The critical value $\lambda^\star$ can be determined by solving the equation:

$$\alpha = \int_0^{\lambda^\star} f_\Lambda(\lambda|H_0)d\lambda \tag{2.20}$$

Taking the logarithms of likelihood ratio Λ and applying Wilks' theorem [51, p. 217], two times the difference of log likelihoods converges to a χ^2 distribution. Large values of this test statistic indicate a deviance from the null H_0 which in turn leads us to reject the null hypothesis H_0. For more details, the interested reader is referred to [51, Chaps. 5-6]

2.4.5 Inference about the mean and the variance

A common problems encountered in GNSS data processing can be summarized as follow: (1) changes in the location (mean); (2) changes in the shape (variance); (3) changes in both, the location and the shape. Therefore establishing the procedures of hypothesis testing is absolutely the right remedy and a necessity. Start by isolating the parameter of interest, form a time series, and carry out the inference about the mean and the variance. We start by defining the test for univariate distributions. The generalization of $(p \times 1)$ parameter vector is the subject of the Section 2.9. Starting by reviewing the correspondence between the acceptance regions and the confidence interval for the mean value define by the hypothesis testing:

$$\begin{aligned} H_0 &: \quad \mu = \mu_0 \\ H_1 &: \quad \mu \neq \mu_0 \end{aligned} \tag{2.21}$$

this means that:

$$\left\{ \text{Do not reject } H_0 : \mu = \mu_0 \text{ at level } \alpha \right\} \tag{2.22}$$

or

$$\left|\frac{\bar{Z}-\mu_0}{S/\sqrt{(N)}}\right| \leq t_{N-1}(\alpha/2) \tag{2.23}$$

is equivalent to:

$$\left\{\mu_0 \text{ lies in the } 100(1-\alpha) \text{ confidence interval } \bar{Z} \pm t_{N-1}(\alpha/2)\ S/\sqrt{N}\right\} \tag{2.24}$$

where t_{N-1} denotes the upper $100(\alpha/2)$'th percentile of the t-distribution with $(N-1)$ degree of freedom. $\bar{Z}$ and S denote the sample mean and variance, respectively.

Moving from one to two sample test is straightforward. The assumptions about the two populations are:

- Independence: Both data sets are independent
- Common variance: Both data sets experience the same variations. If not the problem is known as "*Behrens-Fisher problem*" and is unsolved problem in statistics.

Theorem 2. *Two sample test*

Let $\{X_i\}_{i=1}^{N_1} \sim \mathcal{N}(\mu_x, \sigma^2)$ *and* $\{Y_i\}_{i=1}^{N_2} \sim \mathcal{N}(\mu_y, \sigma^2)$ *and let X's and Y's be independent. Let* S_X^2 *and* S_Y^2 *be the two sample variances, and* S_p^2 *the pooled variance, where*

$$S_p^2 = \frac{\sum_{i=1}^{N_1}(X_i-\bar{X}) + \sum_{i=1}^{N_2}(Y_i-\bar{Y})}{N_1+N_2-2} \tag{2.25}$$

Then

$$T = \frac{\bar{X}-\bar{Y}-(\mu_x-\mu_y)}{S_p\sqrt{\frac{1}{N_1}+\frac{1}{N_2}}} \sim T_{N_1+N_2-2} \tag{2.26}$$

At the α level of significance, the GLRT for testing

$$H_0 \quad : \quad \mu_x = \mu_y, \tag{2.27}$$

$$H_1 \quad : \quad \mu_x \neq \mu_y, \tag{2.28}$$

We reject H_0 if

$$t = \frac{\bar{X}-\bar{Y}}{S_p\sqrt{\frac{1}{N_1}+\frac{1}{N_2}}} \sim T_{N_1+N_2-2} \tag{2.29}$$

is either if $t \leq -t_{\alpha/2,N_1+N_2-2}$ or $t \geq -t_{\alpha/2,N_1+N_2-2}$.

The expression of the squared $t-$distribution is given by the expression:

$$t^2 = \sqrt{N}(\bar{X} - \mu)\ (S^2)^{-1}\ \sqrt{N}(\bar{X} - \mu) \tag{2.30}$$

This is equivalent to:

$$\left\{\text{Normal RV}\right\} \quad \left\{\frac{\text{Scaled chi-square RV}}{\text{degree of freedom}}\right\}^{-1} \quad \left\{\text{Normal RV}\right\}$$

The two-sample problems are set up to detect the possible shifts in the locations. Comparing the variability parameters is more important. The next theorem introduces the F-test to carry out the variance comparison.

Theorem 3. *F-test*
Let $\{X_i\}_{i=1}^{N_1} \sim \mathcal{N}(\mu_x, \sigma_x^2)$ *and* $\{Y_i\}_{i=1}^{N_2} \sim \mathcal{N}(\mu_y, \sigma_y^2)$ *and let X's and Y's be independent. An an approximate GLRT for*

$$\begin{aligned} H_0 &: \quad \sigma_x^2 = \sigma_y^2, \\ H_1 &: \quad \sigma_x^2 \neq \sigma_y^2, \end{aligned} \tag{2.31}$$

At the α *level of significance, the test calls for* H_0 *to be rejected if*

$$\frac{S_Y^2}{S_X^2} \text{ is either } \begin{cases} \leq F_{\alpha/2,N_1-1,N_2-1} \\ \geq F_{1-\alpha/2,N_1-1,N_2-1} \end{cases} \tag{2.32}$$

Recall that testing for equal means depend on whether or not the two population variances were equal. This means that testing for equal variance shall precede the test for the means equality.

2.5 Estimation theory

Systematic errors and uncertainties are common in GNSS observations data processing. Finding optimal estimates for the unknown parameters is absolutely required. Estimators with lower bias and high precision (small variance) are preferable over other estimators. Anyway, one always tries to balance between these two terms in best possible way. Since the bias presents systematic error, it may be reasonable to control the bias before considering the variance reduction of an estimate.

The statistical terms of estimate like consistency, sufficiency, convergence in probability and in distribution are interesting characteristics of any estimates. These estimators are key for building the test statistics in general.

We start by giving some important definitions and basic notations to ease discussions and methods used in this thesis. Let (Ω, Σ, ψ) be a probability space, where Ω is a sample space, Σ is a $\sigma-$algebra of subset Ω and ψ is probability measure defined over all members of Σ.

Denote by $Z = (Z_1, \cdots, Z_n)$ the observations, $q(\theta) = q(\theta_1, \cdots, \theta_k)$ parameters vector, $T(Z)$ the test statistics, $\mathbb{E}(T(.)) = q(\theta)$ is the expectation, $B(\theta, T(.))$ is the bias and $V = (\theta, T)$ is the variance.

Definition 8. *Unbiased Estimator : An estimator $T(Z)$ is said to be an unbiased estimator for $q(\theta)$ if $\mathbb{E}(T) = q(\theta)$, where $\mathbb{E}$ is the expectation operator.*

Definition 9. *Minimum variance estimator: An estimator with minimum variance is often preferred over all other estimators and guarantee the stability of the system. Since our estimator satisfy the unbiasedness (Def. 8), the mean square error(MSE) is equivalent to the variance.*

$$\mathbb{E}\{[(q(\hat{\theta}) - E(q(\hat{\theta}))]^T [q(\hat{\theta}) - E(q(\hat{\theta}))]\} = min \tag{2.33}$$

Definition 10. *Consistency: If $T_n(Z_1, \cdots, Z_n)$ is reasonable consistent estimator for $q(\theta)$, if its converges in probability. That is, for all $\epsilon \geq 0$.*

$$\lim_{N \to \infty} \mathbb{P}\Big\{|T_N(Z_1, \cdots, Z_N) - q(\theta)| \geq \epsilon\Big\} \xrightarrow{\mathbb{P}} 0 \tag{2.34}$$

Often we prove the consistency by applying the Chebyshev inequality, Eq. 4.17.

Definition 11. *Sufficiency : We say that $T_n(Z_1, \cdots, Z_n))$ is reasonable sufficient statistic for $q(\theta)$, if there exists a determination of the conditional distribution $\mathbb{P}(Z|T(.))$ that is independent of $q(\theta)$.*

The information about $q(\theta)$ contained in the sample Z_n is concentrated in the statistics $T(.)$.

2.5.1 Linear models

A simple linear model is given by the Eq. (2.35), relating the response vector $\boldsymbol{z}$ to the variates $\boldsymbol{H}$. $\boldsymbol{x} \in \mathbb{R}^{(n \times 1)}$ is the model parameter. The measurement noise $\boldsymbol{\epsilon} \in \mathbb{R}^{(m \times 1)}$ has zero-mean and variance-covariance $\boldsymbol{\Sigma} \in \mathbb{R}^{(m \times m)}$, and $\boldsymbol{H} \in \mathbb{R}^{(m \times n)}$

$$\boldsymbol{z} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{\epsilon} \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}) \tag{2.35}$$

To obtain an valid solution, the number of measurements must be greater than unknowns ($m > n$). According to the principle of least-square, the

optimal estimator $\hat{\boldsymbol{x}}$ is obtained by minimizing the cost function $C(\boldsymbol{x})$ of the sum of the squares of the residuals $\boldsymbol{\epsilon}$.

$$C(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{m} \boldsymbol{\epsilon}_i^2 = \frac{1}{2}[\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}]^T[\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}] \tag{2.36}$$

The residual vector $\boldsymbol{\epsilon}$ is defined as the difference between the true value of $\boldsymbol{z}$ and the estimated $\hat{\boldsymbol{z}}$.

Minimization of the cost function, Eq. (2.36), with respect to the vector parameter $\boldsymbol{x}$ is obtained by setting its gradient with respect to $\hat{\boldsymbol{x}}$ to zero. The necessary condition:

$$\nabla_{\hat{\boldsymbol{x}}} C = \frac{\partial C}{\partial \hat{\boldsymbol{x}}} = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\hat{\boldsymbol{x}} - \boldsymbol{H}^T\boldsymbol{z} = 0 \tag{2.37}$$

The sufficient condition, see computational assumptions Section 2.5.2.

$$\nabla^2_{\hat{\boldsymbol{x}}} C = \frac{\partial^2 C}{\partial \hat{\boldsymbol{x}} \partial \hat{\boldsymbol{x}}^T} = (\boldsymbol{H}^T\boldsymbol{H})^{-1} \geq 0 \tag{2.38}$$

where $\nabla_{\hat{\boldsymbol{x}}} C$ and $\nabla^2_{\hat{\boldsymbol{x}}} C$ are respectively the Jacobian and the Hessian. From Eq. (2.37), the normal equation (NEQ) is obtained.

$$(\boldsymbol{H}^T\boldsymbol{H})^{-1}\hat{\boldsymbol{x}} = \boldsymbol{H}^T\boldsymbol{z} \tag{2.39}$$

2.5.2 Assumptions and distributions

Imposing assumptions in statistical analysis is necessary in order to build test statistics, estimations, computing confidence region and decision making through hypothesis testing. The drawback of assumptions is that they limits the generalization. Assumptions are always linked to linearity, distribution, independence and variance stability. The interested reader can find more on this topic in [17].

- **Linearity assumptions**: Each observation response $\boldsymbol{z}_i$ can be written as linear function of the i'th row $\boldsymbol{H}_i^T$. That is, $\boldsymbol{z}_i = \boldsymbol{H}_i^T\boldsymbol{x}$
- **Computational assumptions**: The term $(\boldsymbol{H}^T\boldsymbol{H})^{-1}$ exists.
- **Distribution assumptions**: Test statistics mostly used in hypothesis testing for instance the Fisher F-test, student t-test and others, assume that:
  - $\boldsymbol{H}$ is measured without error.

- $\boldsymbol{\epsilon}_i$ is independent of $\boldsymbol{z}_i$.
- $\boldsymbol{\epsilon}_i$ is identical with a common defined variance σ^2.
- we can write that $\boldsymbol{\epsilon}_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2 I)$.

- **Implicit assumptions**: All observations are equally reliable and should have an equal role in determining the least square results and influencing conclusions.

- **Additional comments**

  - Vector of stochastic error term $\boldsymbol{\epsilon}$, is not directly observed and often is assumed to follow Gaussian distribution.
  - Vector $\boldsymbol{e}$ reflects the variance function of the true model error.
  - Taking consideration of GNSS temporal correlation, the LS residuals have different variabilities from one another. This phenomena is called heteroscedastic and is due to dependency of satellite(s) elevation angle E_l. A trade-off between observation quality and the satellite geometry is critical in a GNSS application. Therefore transformations are needed to maintain the temporal correlation properties with homogeneous variance (homoscedastic).

The least square solution of Eq. (2.35) reads:

$$\hat{\boldsymbol{x}} = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{z} \tag{2.40}$$

$$\hat{\boldsymbol{z}} = \boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{z} = \boldsymbol{P}\boldsymbol{z} \tag{2.41}$$

$$\boldsymbol{e} = (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\epsilon} \tag{2.42}$$

where $\boldsymbol{P}$ is called the prediction (hat) matrix. The ordinary residuals $\boldsymbol{e}$ depend only on $\boldsymbol{P}$ and the relationship between $\boldsymbol{e}$ and $\boldsymbol{\epsilon}$ is established by Eq. (2.42).

Assuming that $\boldsymbol{\epsilon}$ are independent and have a common variance, then Eq. (2.42) indicates that $\boldsymbol{e}$ are not independent (unless $\boldsymbol{P}$ is diagonal) and they do not have the same variance (unless diagonal elements of $\boldsymbol{P}$ are equal).

Residuals $\boldsymbol{e}$ can be regarded as a reasonable substitute of $\boldsymbol{\epsilon}_i$ if the rows of $\boldsymbol{H}$ is homogeneous. This means that the diagonal element of prediction matrix P are approximately equal and off-diagonal elements of $\boldsymbol{P}$ are sufficiently small. For these reasons, it is preferable to use the transformed version of the ordinary residuals for diagnostic purposes. That it is, instead of using e_i one may use the transformed expression given the function $\boldsymbol{G}$:

$$\boldsymbol{G}(\boldsymbol{e}_i, \sigma_i) = \boldsymbol{e}_i/\sigma_i \tag{2.43}$$

2.5.3 Fitting Parametric Model

Various methods can be considered to fit the parametric model tothe sample population.

1. Least square parameter estimations (LSPE)
   - Ordinary Least Square (OLS): The cost function $C_1(\theta)$ is given by the Eq. (2.36).
   - Weighted Least Square (WLS): We choose $\boldsymbol{x}$ to minimize the cost function $C_2(\boldsymbol{x})$

$$C_2(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{m} \boldsymbol{\epsilon}_i^2 = \frac{1}{2}[\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}]^T \boldsymbol{W}[\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}] \tag{2.44}$$

     The estimate becomes: $\hat{\boldsymbol{x}}_W = (\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{z}$ and $\boldsymbol{W}$ is a diagonal matrix. The elements of $\boldsymbol{W}(\boldsymbol{x})$ present the variances of the measurements.
   - Generalized Least Square (GLS): We choose $\boldsymbol{x}$ to minimize the cost function $C_3(\boldsymbol{x})$

$$C_3(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{m} \boldsymbol{\epsilon}_i^2 = \frac{1}{2}[\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}]^T \boldsymbol{G}[\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}] \tag{2.45}$$

     The estimate becomes: $\hat{\boldsymbol{x}}_G = (\boldsymbol{H}^T\boldsymbol{G}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{G}\boldsymbol{z}$. The elements of $\boldsymbol{G}$ present the variances and the covariances of the measurements.

2. Best Linear Unbiasedness Estimator (BLUE): Our aim is to determine an estimator which is unbiased and has the smallest error variance. The generator of such an estimator is the GMM, see Section (2.5.4). The cost function is the mean square error:

$$C_4(\boldsymbol{x}) = \mathbb{E}\{(\hat{\boldsymbol{x}} - \boldsymbol{x})^T(\hat{\boldsymbol{x}} - \boldsymbol{x})\} \tag{2.46}$$

   We are seeking an estimator of the form $\hat{\boldsymbol{x}} = \boldsymbol{B}\boldsymbol{z}$ that is linear function of the observations. The solution derivation is given [80, p. 226] where $\boldsymbol{B} = (\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H})\boldsymbol{H}^T\boldsymbol{R}^{-1}$, and $\boldsymbol{R} = \mathbb{E}(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon})$. $\mathbb{E}(.)$ is the expectation operator.

   Replacing $\boldsymbol{R}^{-1} = \boldsymbol{G}$ in Eq. (2.45) as the weighting matrix, we get the BLUE estimator. The inverse of the weighting matrix $\boldsymbol{G}$ is referred to as the precision matrix $\boldsymbol{P}$.

3. Maximum Likelihood (ML): If our observations come from a Gaussian process, then the procedure is simple by computing the exact likelihood and maximize it with respect to unknown parameter vector $\boldsymbol{x}$. If the number of observation is larger, this method suffers to compute the inverse of matrix and determinant of the model covariance matrix.

   If the observations follow the Gaussian distribution, the MLE and LSE provide the same estimate.

4. Maximum a posteriori Estimator (MAP): Maximization of the posterior distribution, multiplication of the likelihood and the prior distribution.

$$\hat{\boldsymbol{x}}_{\text{map}} = \text{argmax}\{p(\boldsymbol{x}_k|\boldsymbol{z}_k)\} \tag{2.47}$$

5. Minimum Means Square Error (MMSE): Defined as the expectation of the posterior distribution.

$$\hat{\boldsymbol{x}}_{\text{mmse}} = \mathbb{E}\{p(\boldsymbol{x}_k|\boldsymbol{z}_k)\} \tag{2.48}$$

2.5.4 Gauss-Markov Model

The Gauss-Markov model (GMM) is given by the Eq. (2.35) with additional assumptions. The expectation $\mathbb{E}(\boldsymbol{z}) = \boldsymbol{H}\boldsymbol{x}$ and the variance/deviance is defined $\boldsymbol{V}(\boldsymbol{z}) = \sigma^2\boldsymbol{P}$, where $\boldsymbol{P} = \boldsymbol{W}$ is the precision/weight matrix and is defined as the inverse of Σ.

The matrix $\boldsymbol{P}$ is assumed to symmetric and positive defined (SPD), $\boldsymbol{H}$ has a full column rank and the number n of observation is always kept greater that the number m of parameters. This model states that the unbiased (Def. 8) least squares estimator $q_1(\boldsymbol{\theta})$ has the smallest mean squared error (Def. 9) of all linear estimators with no bias and is preferable over all other estimator candidates. However, there may exists a biased estimator $q_2(\boldsymbol{\theta})$ with minimum mean square error. Such an estimator sacrifices/trades little bias for larger reduction in variance. Next section handles this class of estimation.

The GMM is very popular and widely used in GNSS data processing. Interested reader is referred to [60, Chap. 3.2]. In addition, the model bears the name of two respected scientists Gauss and Markov. Both have used different approaches and reached the same conclusion. Gauss used the least-square method while Markov used the best unbiased estimator to estimate the parameter vector $q(\boldsymbol{\theta})$. The interested reader is referred to the original papers of Gauss [33] and Markov [76].

2.6 Penalized methods

The power of the penalized methods is to improve the variance of an estimator. The reduction of the variance is accomplished by either imposing the constraint on regression coefficients or by regularization of the rate of change of residual errors and local variations. That is, find a best compromise between the goodness-of-fit and the smoothness. The results are the prediction accuracy (low variance) and easy interpretation (lower dimensions). Such biased estimators are commonly used and any method that shrink or sets some coefficients of least squares to zero may results in a biased estimator. The two best-known techniques for shrinking the regression coefficients towards zero are Ridge and Lasso regressors. For a detailed mathematical formulation and discussions, please refer to references [145, Chap. 7.3] and [43, Chap.3].

2.6.1 Penalized least squares

In case of the penalized least square (PLS), the cost function defined by Eq. (2.36) becomes [38, p. 5]:

$$C(g) = \sum_{i=1}^{m} [z_i - g(t_i)]^2 + \alpha \int_a^b \{g''(x)\}^2 \tag{2.49}$$

where $\alpha > 0$ is the smoothing parameter and defines the rate of change between the residuals and local variations. Anyway, minimizing $C(g)$ gives the best compromise between smoothness and goodness-of-fit. A large value of α will make the penalty term more in action while a small value of α, the first term will be the main contribution.

2.6.2 Penalized maximum likelihood

MLE is the most attractive method used for parameters estimation. Like any other methods, MLE sometimes fail to deliver what is expected from the method. An amelioration potential is the penalized ML and profile ML and will not be treated here. The interested reader is referred to Cole et al. [20].

2.6.3 Shrinkage methods

The Gauss-Markov model [60, p. 153] and [42, p. 49] states that for any unbiased estimator, the LS estimates for $\boldsymbol{x}$ have the smallest variance among all other existing unbiased estimators. Anyway, we can find a biased estimator with a smaller mean square error. That is, sacrificing little bias for a larger reduction of variance.

Bias estimators are used in estimation theory. Shrinkage methods are such techniques that impose constraints on regression coefficients and can be considered as a real competitors to the Gauss-Markov model. The most used are Ridge and Lasso regressors.

Ridge regression

Ridge regression shrinks the coefficients by imposing a penalty on their size in L_2 and minimizing a penalized residual sum of squares.

$$\hat{\boldsymbol{x}}_{\text{ridge}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \left\{ \sum_{i=1}^{m} \left\{ z_i - x_0 - \sum_{j=1}^{p} h_{ij} x_j \right\}^2 + \lambda \sum_{j=1}^{p} x_j^2 \right\} \tag{2.50}$$

Equivalently, we can write the Eq. (2.50) differently:

$$\hat{\boldsymbol{x}}_{\text{ridge}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \sum_{i=1}^{m} \left\{ z_i - x_0 - \sum_{j=1}^{p} h_{ij} x_j \right\}^2, \quad \text{subject to } \sum_{j=1}^{p} x_j^2 \leq s \tag{2.51}$$

Lasso regression

The Lasso estimate replace the L_2 Ridge penalty $\sum_{j=1}^{p} x_j^2$ by the L_1 Lasso penalty $\sum_{j=1}^{p} |x_j|$. The Lasso estimate reads:

$$\hat{\boldsymbol{x}}_{\text{lasso}} = \underset{x}{\operatorname{argmin}} \sum_{i=1}^{m} \left\{ z_i - x_0 - \sum_{j=1}^{p} h_{ij} x_j \right\}^2, \quad \text{subject to } \sum_{j=1}^{p} |x_j| \leq k \tag{2.52}$$

2.7 Bayesian statistics

The popularity of the Bayesian method is at the peak for the moment. The observations, $\boldsymbol{z}$ and the unknowns, $\boldsymbol{\theta}$ are treated as stochastic. The Kalman filter belongs to this class. For historical development of this amazing data processing methodology, we refer the interested reader to Blangiardo and Cameletti [9, Chap. 3].

In many applications, it is advantageous to consider parameter vector $\boldsymbol{\theta}$ as observed by a stochastic variable, with apriori distribution $\pi(\boldsymbol{\theta})$. In order to compute the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{z})$, which represents the

uncertainty about $\boldsymbol{\theta}$. Two components are needed. The likelihood $l(\boldsymbol{z}|\boldsymbol{\theta})$ and $\pi(\boldsymbol{\theta})$. Applying the Bayesian formula on distributions, we get:

$$p(\boldsymbol{\theta}|\boldsymbol{z}) = \frac{l(\boldsymbol{z}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{p(\boldsymbol{z})} \tag{2.53}$$

where $p(\boldsymbol{z})$ is the marginal distribution of the data and is considered as a normalization constant, so the Equation (2.53) can be written as $p(\boldsymbol{\theta}|\boldsymbol{z}) \propto l(\boldsymbol{z}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$.

2.7.1 Prior selection

Based on the nature of parameter of interest $\boldsymbol{\theta}$, two choices are considered to derive the prior distribution $\pi(\boldsymbol{\theta})$, informative and non-informative priors. With informative, the prior information is available from previous experiments or from expert's opinions. Non-informative is to use for instance the uniform distribution, but the most famous is the one proposed by Jeffery [50] and is given by $\pi(\boldsymbol{\theta}) \propto |I(\boldsymbol{\theta})|^{1/2}$ where

$$I(\theta) = -\mathbb{E}\left[\left.\frac{\partial^2 \log f(\boldsymbol{z}|\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}}\right] \tag{2.54}$$

is the expected Fisher information matrix about the parameter vector $\boldsymbol{\theta}$.

Berger [7] presented a prior based on expected discrepancy measures of information. Another class of generating the prior is the conjugate distributions. The posterior and the prior distributions belong to the same distribution family, often the exponential family.

Anyway, it's impossible to cover all the material for the Bayesian modeling. The reader is referred to the books by Berger [6, Chaps. 4–7], and Rowe [96].

2.7.2 Kalman-Filter

Kalman Filter (KF) is a weighted recursive least square algorithm, satisfying the following three general requirements of optimality conditions;

- Unbiasedness (Definition 8).
- Consistency (Definition 10).
- Minimum variance (Definition 9).

KF focuses attention on the state vector of the system $\boldsymbol{x}$, because it contains all necessary and relevant informations needed to investigate the system. This class of estimation is referred to as Bayesian state estimation. One try

to estimate the posterior distribution based on all available information provided by the system, including current measurements and the historical data known as prior info.

KF equation definitions

Popularity of KF is due to its recursive characteristic, the user does not need to save previous observations; instead, all filter observations are carried forward in the filter.

In order to make inference about the dynamic system, at least two main equations are needed; the measurement and system models. In addition, their corresponding stochastic models as illustrated in equations. One of the characteristics of the noises, namely, w_k and v_k, of the two models are mutually independent and zero-mean normal distributed with variance-covariance $\boldsymbol{Q}$ and $\boldsymbol{R}$.

Following the notation given by Brown and Hwang [13, Chap. 5], the system and the measurement equations are given respectively by Eqs.(2.55) and (2.56).

State-space system model:

$$\boldsymbol{x}_k = \boldsymbol{\Phi}_k \, \boldsymbol{x}_{k-1} + \boldsymbol{w}_k \quad \boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k) \tag{2.55}$$

State-space measurement model:

$$\boldsymbol{z}_k = \boldsymbol{H}_k \, \boldsymbol{x}_k + \boldsymbol{v}_k \quad \boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_k) \tag{2.56}$$

where

$\boldsymbol{x}_k$	$(n \times 1)$ state vector at epoch k,
$\boldsymbol{\Phi}_k$	transition matrix from epoch $k-1$ to k,
$\boldsymbol{x}_{k-1}$	state vector at epoch $k-1$
$\boldsymbol{w}_k$	system noise at epoch k
$\boldsymbol{Q}_k$	covariance matrix of system noise $\boldsymbol{w}_k$
$\boldsymbol{z}_k$	$(m \times 1)$ measurements vector at epoch k,
$\boldsymbol{H}_k$	$(m \times n)$ design matrix that relate the state vector to measurements at epoch k
$\boldsymbol{v}_k$	$(m \times 1)$ measurement noise at epoch k and assumed to white noise
$\boldsymbol{R}_k$	$(m \times m)$ covariance matrix of measurement noise $\boldsymbol{v}_k$

Eq. 2.55 is known as vector auto-regressive process of order 1 (VAR(1)) and describes the evolution of the state $\boldsymbol{x}_k$ with time. Generally speaking, one way to introduce the dependency between $\boldsymbol{x}_k$ and $\boldsymbol{x}_{k-1}$ is through an

auto-regressive process. For detailed information of VAR model, the reader is referred to [108, p.302]. The uncertainties are expressed by the error covariance matrix $\boldsymbol{Q}$. The second equation (2.56) relating the state vector $\boldsymbol{x}_k$ and the measurement $\boldsymbol{z}_k$ is modeled as GMM Section 2.5.4. However, the meaning of the Kalman filter equation pairs is given. We can replace any model of these two equations by a competitive one that can improve the results and provide filter stability. The procedures of how to alternate between estimations, corrections and innovations computation are given by algorithm 1.

The computation steps needed for the filter are:

Algorithm 1: Conventional Kalman filter main processing loop

1 Step 1: Enter prior estimate of the state vector $\hat{\boldsymbol{x}}_0^-$ and its corresponding error covariance $\hat{\boldsymbol{P}}_0^-$.

2 **for** $k = 1, 2, \cdots$ **do**

3 Step 2: Compute the Kalman gain $\boldsymbol{K}_k = \hat{\boldsymbol{P}}_k^- \boldsymbol{H}_k (\boldsymbol{H}_k \hat{\boldsymbol{P}}_k^- \boldsymbol{H}_k^T + \boldsymbol{R}_k)^{-1}$

4 Step 3: Update the estimate $\hat{\boldsymbol{x}}_k^+$ with new information $\boldsymbol{z}_k$. $\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k (\boldsymbol{z}_k - \boldsymbol{H}_k \hat{\boldsymbol{x}}_k^-)$

5 Step 4: Compute the error covariance $\hat{\boldsymbol{P}}_k^+ = (\boldsymbol{I} - \boldsymbol{K_k}\, \boldsymbol{H_k}) \hat{\boldsymbol{P}}_k^-$

6 Step 5: Project a head: $\hat{\boldsymbol{x}}_{k+1}^- = \boldsymbol{\Phi} \hat{\boldsymbol{x}}_k^+$ and $\hat{\boldsymbol{P}}_{k+1}^- = \boldsymbol{\Phi}_k \hat{\boldsymbol{x}}_k^+ \boldsymbol{\Phi}_k^T + \boldsymbol{Q}_k$

7 **end**

Optimal choice of the Kalman filter gain

The key of the KF is the optimal choice of the Kalman filter gain $\boldsymbol{K}_k$. $\boldsymbol{K}_k$ is chosen to minimize the length of the estimation error vector. This is equivalent to minimize the trace of the error covariance matrix (diagonal elements), corresponding to the total variance of our estimator. It is clear that the relationship between these two quantities is:

$$\mathrm{E}\{(\hat{\theta} - E(\hat{\theta}))^T (\hat{\theta} - E(\hat{\theta}))\} = \mathrm{trace}\ \ \mathrm{E}\{(\hat{\theta} - E(\hat{\theta}))(\hat{\theta} - E(\hat{\theta}))^T\} \tag{2.57}$$

The Kalman filter K_k gain is a very important parameter to monitor in order to guarantee the stability of the filter. The $\boldsymbol{K}_k$ is directly involved in the updating process of the state vector $\hat{\boldsymbol{x}}_k^+$ and the error covariance matrix $\hat{\boldsymbol{P}}_k^+$ as shown in steps 3–4 in algorithm 1. This is easily done by inspection of diagonal elements of $\boldsymbol{K}_k$. Total and specific variances are other options.

2.7.3 Heywood Case

The output from Kalman filter(s) need to be investigated based on epoch basis. Iteration between estimation and updating steps causes some element

of the main diagonal of the error covariance matrix to become negative. The main diagonal elements correspond to the variance. For statisticians, this phenomena is referred to as the Heywood case [44].

In social science studies where the correlation analysis is the main issues, the most common methods used are for instance the factor analysis (FA) and the structural equation modeling (SEM). The result from different analysis shows the existence of improper solutions and the Heywood case. That is, the solution can not be obtained due to the correlation and the variance estimates are respectively greater that one and less than zero. A statistical effort devoted to understand when these events are most likely to occur and the identification of the main source of the problem. Among them are; outliers [10, Bollen 1987], non-convergence and under-identification[Van Driel 1978] [26], empirical under-identification [93, Rindskopf 1984], and structurally misspecified models [18, Chen 2001].

In the field of navigation and surveillance, similar problems have occurred and been examined while the employment of Kalman filer and the most appropriate reasons are summarized as follow; process noise uncertainties, transition matrix is ill-conditioned, this means that a small change in the constant coefficient matrix results in a large change in the solution, effect of incorrect a priori statistics (the need for initialization of ill-conditioned caused by large a priori, we often use scaling), scaling some variables and leaving others cause the filter to produce the negative variance, it's better to be avoided. Scaling is a dangerous task and can produce greater errors and unreliable Kalman gain.

Improvement suggested by experienced modelers are; Gelb [35, p. 286], Bierman [8, pp. 68-112], Maybeck [79, p. 368], Grewal et al [40] and others can be summarized as follow:

- State-noise compensation.
- Joseph's stabilized Kalman filter.
- Conventional Kalman filter with lower bound.
- Potter-Schmidt square root filter.
- U-D factorization, the modified weighted Gram Schmidt (MWGS) time propagation.
- Fading memory filter.
- Sequential root square filter (SRIF).

New suggestions to avoid the Heywood case

Another possible way to avoid the Heywood case of the conventional Kalman filter is by model substitutions. That is, replacing the Eqs. (2.56) and (2.55) by a competitive models.

- Measurement model, Eq. (2.56): The GMM can be replaced with any penalized methods defined in Section 2.6.
- System model, Eq. (2.55): The VAR(1) can be replaced with any higher order VAR(.) process, for instance VAR(2), an $ARMA(p,q)$ process or Holt-Winter algorithm [12, pp. 322-326].

Anyway, the key in GNSS data processing is to provide a continuous, stable and correct error covariance matrix (EVCM) on epoch by epoch basis.

2.7.4 The particle filter

Recently, the particle filter becomes a very popular tool used in GNSS state estimation and found a various range applications in scientific and engineering. For instance, multipath monitoring [37], ambiguity resolution [131] and receiver coordinate estimation [121].

In recent years, non-linear, non-stationary and non-Gaussian filtering has been the main focus in statistics and engineering. GNSS data processing falls in this category, and linear Gaussian estimation problem carried out by the conventional Kalman filter fails to deliver an optimal solution. In addition, the computed PDF remains Gaussian at each iteration of the filter. In case of non-linear and non-Gaussian, there is no general expression for the required PDF, instead one is forced to use the approximations or sub-optimal solutions. Among these approximations are; (1) analytical approximations; (2) numerical approximations; (3) simulation approaches; and others. Anyway, the particle filter falls in the third category, namely the simulation approaches. The main ingredients of the PF are listed as follow:

1. The system and measurement equations are given by nonlinear functions $f(.)$ and $h(.)$:

$$\boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k, \boldsymbol{w}_k); \quad \text{System equation} \tag{2.58}$$

$$\boldsymbol{z}_k = h(\boldsymbol{x}_k, \boldsymbol{v}_k); \quad \text{Measurement equation} \tag{2.59}$$

   where $\{\boldsymbol{w}_k\}$ and $\{\boldsymbol{v}_k\}$ are mutually independent white Gaussian processes with known PDF's. k is time index, $\boldsymbol{x}_k$ is the state and $\boldsymbol{z}_k$ is the observations.

2. Assuming that the PDF of the initial state $p(x_0)$ is known. Randomly generate N initial particles on the basis of the pdf $p(x_0)$. These particles are denoted $\boldsymbol{x}_{0,i}^{+}$, for $i = \{1, \cdots, N\}$. The parameter N is chosen carefully to balance between the computational complexity and the estimation accuracy.

3. For $k = 1, 2, \cdots, N$, do the following tasks:

   a) Perform the time propagation step to obtain a priori particles $\boldsymbol{x}_{k,i}^{-}$ using the known process equation and the known PDF of the process noise:

   $$\boldsymbol{x}_{k,i}^{-} = f_{k-1}(\boldsymbol{x}_{k-1,i}^{+}, \boldsymbol{w}_{k-1}^{i}) \quad i = 1, \cdots, N \tag{2.60}$$

   where each $\boldsymbol{w}_{(k-1)}^{i}$ noise vector is randomly generated on the basis of the known pdf of $\boldsymbol{w}_{(k-1)}$

   b) Compute the relative likelihood q_i of each particle $\boldsymbol{x}_{k,i}^{-}$ conditioned on the measurement $\boldsymbol{z}_k$. This is done by evaluating the PDF $p(\boldsymbol{z}_k|\boldsymbol{x}_{k,i}^{-})$ on the basis of the nonlinear measurement equation and the PDF of the measurement noise.

   c) Scale the relative likelihoods obtained in the previous step as follows:

   $$q_i = \frac{q_i}{\sum_{j=1}^{N} q_j} \tag{2.61}$$

   Normalization: the sum of all the likelihoods q_i is equal to one.

   d) Generate a set of a posteriori particles $\boldsymbol{x}_{k,i}^{+}$ on the basis of the relative likelihoods q_i. This is called the re-sampling step.

   e) Now that we have a set of particles $\boldsymbol{x}_{k,i}^{+}$ that are distributed according to the PDF $p(\boldsymbol{x}_k|\boldsymbol{z}_k)$, we can compute any desired statistical measure of this PDF, the mean vector and the covariance matrix.

The PDF of the initial state $p(\boldsymbol{x}_0)$ can be computed from the historical data or from the filter's warm-up stage. In case of no strong opinion about the prior distribution, it is desirable to start with a prior that is non-informative PDF, the uniform distribution with Jeffery rules Section 2.7.1.

2.7.5 Regularized Particle Filter (RPF)

The divergence of the PF; The drawbacks of the PF are the computation demanding, filter instability and sample impoverishment. In order to overcome these difficulties, the use of regularized PF (RPF) is recommended.

The PF is not a perfect filter and suffers from the phenomena called the sample impoverishment. This phenomena occurs when there is no overlapping of the posterior $p(\boldsymbol{z}_k|\boldsymbol{x}_k)$ and the prior distribution $p(\boldsymbol{x}_k|\boldsymbol{z}_{k-1})$ over the region of state space. This means that only a few particles are re-sampled to become a posteriori particles. In the worst case particles can take the same value and result in a black hole often called *"particles collapse"*.

Suggestions from the experts in the field is simply to increase the sample size N, the number of particles, but this can quickly lead to intensive computational demands, and often causes unwanted delays. In order to improve the performance of the PF, some amelioration has to be taken, for instance considering that we are sampling from a continuously function instead of discrete and applying the technique of the kernel smoothing in the re-sampling stage. The re-sampling strategy is to replaces step (3 (iv) in the particle filter algorithm 2.7.4 by this procedure:

Assume that we have an n-state system, the N a priori particles $\boldsymbol{x}_{k,i}^-$ and the N corresponding (normalized) a priori probabilities q_i. Generate a posteriori particles $\boldsymbol{x}_{k,i}^+$ as follow:

1. Compute the ensemble mean $\boldsymbol{\mu}$ and sample covariance $\boldsymbol{S}$ of the a priori particles as follows.

$$\boldsymbol{\mu} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_{k,i}^- \tag{2.62}$$

$$\boldsymbol{S} = \frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{x}_{k,i}^- - \boldsymbol{\mu})(\boldsymbol{x}_{k,i}^- - \boldsymbol{\mu})^T \tag{2.63}$$

2. Perform a square root factorization of $\boldsymbol{S}$ (e.g., a Cholesky factorization) to compute the $n \times n$ matrix $\boldsymbol{A}$ such that $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{S}$.

3. Compute the volume of the n-dimensional unit sphere as $v_n = 2\pi v_{n-2}/n$. The starting values for this recursion are $v_1 = 2, v_2 = \pi$ and $v_3 = 4\pi/3$

4. Compute the optimal kernel bandwidth h as follows:

$$h = \frac{1}{2}\left\{8v_n^{-1}(n+4)(2\sqrt{\pi})^n\right\}^{1/(n+4)}N^{1/(n+4)} \tag{2.64}$$

   The bandwidth h is the smoothing parameter of the particle filter.

5. Approximate the pdf $p(\boldsymbol{x}_k|\boldsymbol{z}_k)$ as follows:

$$\hat{p}(\boldsymbol{x}_k|\boldsymbol{z}_k) = \sum_{i=1}^{N} q_i K_h(x_k - x_{k,i}) \tag{2.65}$$

where the kernel $K_h(x)$ is given as:

$$K_h(x) = (\det A)^{-1} h^{-n} K(A^{-1}x/h) \tag{2.66}$$

and the Epanechnikov kernel $K(x)$ is given as

$$k(x) = \begin{cases} \frac{n+2}{2v_n}(1 - \|x\|_2^2) & ; \quad \text{if } \|x\|_2 < 1 \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{2.67}$$

Other kernels can also be used in the PDF approximation.

6. Now that we have $\hat{p}(\boldsymbol{x}_k|\boldsymbol{z}_k)$ from the previous step, we generate the a posteriori particles $\boldsymbol{x}_{k,i}^+$ from the approximate $\hat{p}(\boldsymbol{x}_k|\boldsymbol{z}_k)$.

Other approaches exist to prevent the sample impoverishment, for instance Markov Chain Monte Carlo (MCMC), auxiliary PF and others. The interested reader is referred to [109, pp. 476-480].

2.8 Spatial statistics

According to the Tooler's first law in geology:

"Near things are more related than distant things".

This phenomena was studied intensively by Kolmogorov in more details in 1948 by calculating the variance between two locations in fluid studies. The South African mining engineer D.G. Krige used the same analogy and developed the linear spatial interpolation algorithm carrying his name the kriging algorithm. The Danish mathematician Torben Krarup [3] was the first to build the theoretical foundation for this new concept, namely collocation. Since then, the method has been considered by geodesists as an algorithm for performing geodetic computations.

2.8.1 Introduction to spatial processes

In order to carry out the estimation, inference, predictions of spatial processes, it's worth to start by some basic definition. This class of processes are widely used in GNSS data processing.

Definition 12. *Spatial processes: A spatial process or random field is a collection of random variables, indexed by some set $D \subset \mathbb{R}^d$, containing spatial coordinates $s_1, s_2, \ldots, s_N$.*

Definition 13. *Gaussian processes: The process $z(s)$ is said to be Gaussian, if for any $p \geq 1$ and locations $\{s_1, s_2, ..., s_N\}$ the vector $(z(s_1), z(s_2), \cdots, z(s_p))$ has multivariate normal distribution.*

$$f(\mathbf{z}) = |2\pi|^{-p/2}|\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{z}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{z}-\boldsymbol{\mu})\right\} \tag{2.68}$$

Definition 14. *Strictly stationary: A random field is called strict stationary field if the spatial distribution under translation of coordinates is invariant, i.e.*

$$\begin{aligned} &\mathbf{Pr}\{Z(s_1) < z_1, Z(s_2 < z_2, \cdots, Z(s_k) < z_k\} = \\ &\mathbf{Pr}\{Z(s_1 + h) < z_1, Z(s_2 + h) < z_2, ..., Z(s_k + h) < z_k\} \end{aligned} \tag{2.69}$$

For all k and h

Definition 15. *Second order staionary: A process $\{Z(s_j)\}_{j=1}^N$ is a second-order stationary if the process has a constant mean $E\{Z(s_j)\} = \mu$ for all j and a valid covariance function that depend only on spatial separation lag h between locations s_j.*

Definition 16. *Variogram: Is a tool of spatial statistics that measures how quickly the spatial autocorrelation $\gamma(h)$ falls off with increasing distance h.*

The variogram is defined as follow:

$$\begin{aligned} \mathbf{var}\left\{Z(s_i) - Z(s_j)\right\} &= \text{var}\{Z(s_i)\} + \text{var}\left\{Z(s_j)\right\} - 2\ \text{cov}\left\{Z(s_i), Z(s_j)\right\} \\ &= 2\ \boldsymbol{C}(0) - 2\ \boldsymbol{C}(s_i - s_j) \\ &= 2\{\boldsymbol{C}(0) - \boldsymbol{C}(h)\} \\ &= 2\boldsymbol{\gamma}(h) \end{aligned} \tag{2.70}$$

Note that $\mathbf{C}(0)$ is the variance, $\mathbf{C}(h)$ is the covariance function and $h = |s_i - s_j|$ is the lag separation.

Definition 17. *Isotropy: If the process is intrinsically stationary with a semi-variogram for some function. The semi-variogram depends on its vector argument only through its length, and then the process is isotropic.*

2.8.2 Valid covariance functions

The isotropic processes are easy to deal with, due to the existence of widely parametric forms for semi-variogram. Here we list some of the most used:

1. Linear

$$\gamma_0(h) = \begin{cases} 0 & \text{if } h = 0; \\ c_0 + c_1 h & \text{if } h > 0; \end{cases} \tag{2.71}$$

2. Gaussian

$$\gamma_0(h) = \begin{cases} 0 & \text{if } h = 0; \\ c_0 + c_1(1 - \exp -\frac{h^2}{R^2}) & \text{if } h > 0; \end{cases} \tag{2.72}$$

3. Wave

$$\gamma_0(h) = \begin{cases} 0 & \text{if } h = 0; \\ c_0 + c_1(1 - \frac{R}{h}\sin(\frac{h}{R})) & \text{if } h > R; \end{cases} \tag{2.73}$$

4. Exponential

$$\gamma_0(h) = \begin{cases} 0 & \text{if } h = 0; \\ c_0 + c_1(1 - \exp -\frac{h^2}{R^2}) & \text{if } h > 0; \end{cases} \tag{2.74}$$

5. Power Law

$$\gamma_0(h) = \begin{cases} 0 & \text{if } h = 0; \\ c_0 + c_1 h^{\lambda} & \text{if } h > 0; \end{cases} \tag{2.75}$$

Positive definiteness

It is often desirable to construct a new covariance model as a linear combination of basic covariance models. To do so, the covariance function of the second-order stationary spatial process must satisfy the positive-defined condition for any locations and arbitrary real numbers w_i for all i.

$$\sum_{i=1}^{k}\sum_{j=1}^{k} w_i w_j C(s_i - s_j) \geq 0 \tag{2.76}$$

The left side is just the variance of $\sum_{i=1}^{N} w_i Z(s_i)$

2.8.3 Estimation Process

After establishing the main concepts of spatial covariance and variogram, now it is time to think about their estimation. The simplest estimator is the method of moments (MoM) estimator and is given by:

$$\hat{\gamma}(h) = \frac{1}{|N(h)|}\sum_{(s_i,s_j)\in N(h)} \left[Z(s_i) - Z(s_j)\right]^2 \tag{2.77}$$

Where the $N(h)$ denotes all pairs (s_i, s_j) for which the difference is $s_i - s_j = h$ and $|N(h)|$ denotes the cardinality. This estimator is not robust against outliers.

Cressie [23] suggested an estimator as an approximately unbiased estimator for variogram.

$$2\hat{\gamma}(h) = \frac{1}{0.457 + \frac{0.494}{|N(h)|}} \times \left(\frac{1}{|N(h)|} \sum_{(s_i,s_j)\in N(h)} \left[Z(s_i) - Z(s_j) \right]^{\frac{1}{2}} \right)^4 \tag{2.78}$$

2.8.4 Fitting Parametric Model

Two methods can be considered to fit the parametric models to the sample variogram

1. Least Square Estimations
   - Ordinary Least Square(OLS): We choose θ to minimize

     $$\left\{ (\hat{\gamma} - \gamma(\theta))^T (\hat{\gamma} - \gamma(\theta)) \right\} \tag{2.79}$$

   - Generalized Least Square(GLS): We choose θ to minimize

     $$\left\{ (\hat{\gamma} - \gamma(\theta))^T V(\theta)^{-1} (\hat{\gamma} - \gamma(\theta)) \right\} \tag{2.80}$$

   - Weighted Least Square(WLS): We choose θ to minimize

     $$\left\{ (\hat{\gamma} - \gamma(\theta))^T W(\theta)^{-1} (\hat{\gamma} - \gamma(\theta)) \right\} \tag{2.81}$$

   $W(\theta)$ is a diagonal matrix, the diagonal elements present the variances of the entries of $\hat{\gamma}$

2. Maximum Likelihood (ML)
   If our observations come from a Gaussian process, then the procedure is simple by computing the exact likelihood and maximize it with respect to unknown parameters. If the number of observation is larger, this method suffers to compute the inverse of matrix and determinant of the model covariance matrix.

2.8.5 Prediction and Interpolation

Given the observation $\{Z(s_i)\}_{i=1}^N$, we like to predict the value of $Z(s_0)$ where we have not observed anything. Our goal is to find an estimator $\hat{Z}_0 = \hat{Z(s_0)} = \sum_{i=1}^N w_i Z(s_i)$ so that the following requirements are met:

1. **Unbiasedness:** $E[Z(s_0)] = E[Z_0]$: This will be accomplished if $\sum_{i=1}^{N} w_i = 1$ and the mean is stationary (constant).

2. **Minimum Prediction Variance:** We make some assumptions about the mean value of random filed Z(s). If the mean is constant across the entire region of interest and is unknown, then we have ordinary kriging, otherwise the method is known as the simple kriging.

   Any estimator meeting the conditions of unbiasedness and minimum prediction variance is said to be a best linear unbiased predictor (BLUP).

   Let examine the components of the Mean Square Prediction Error (MSPE).

$$\begin{aligned}
\sigma_\epsilon^2 &= \text{var}(Z_0 - \hat{Z}_0) \\
&= \text{var}(Z_0) + \text{var}(\hat{Z}_0) - 2\text{cov}(Z_0, \hat{Z}_0) \\
&= \sigma^2 + \text{var}(\sum_{i=1}^{N} w_i Z_i) - 2\ \text{cov}(Z_0, \sum_{i=1}^{N} w_i Z_i) \\
&= \sigma^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \text{cov}(Z_i, Z_j) - 2\ \sum_{i=1}^{N} w_i \text{cov}(Z_i, Z_j) \\
&= \sigma^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j C_{ij} - 2\ \sum_{i=1}^{N} w_i C_{i0} \qquad (2.82)
\end{aligned}$$

where w_i and w_j are weights, C_{i0} and C_{ij} are the variance and the covariance between two locations.

2.8.6 Spatial-temporal covariance definition

The aim of analysis of GNSS data is describing the leading modes of variability in time-space processes. The observation of the process Z at locations $s_i = (x_i, y_i)$ and at time t is given by the expression $\left\{Z(s,t) : s \in D \in \mathbb{R}^2, t \in T\right\}$.

That is, the spatial domain changes over time and we cannot treat spatio-temporal data as spatial data with additional dimension. The main reason is that time and space cannot be compared due to the fact that they possess different unity and in addition the space has no past, present and future [102].

Textbooks covering spatial-temporal processes with different emphasis are listed bellow:

1. Cressie [23]: Is the classical and will always be used as main reference to any spatial process study.

2. Matérn [78]: Major advantages over other covariance functions. Matérn covariance function is governed by three parameters (location, scale and shape).

3. Guttorp and Sampson proposed a two-step approach to handle non-stationarity. A nonparametric algorithm to estimate the spatial covariance structure for the entire random field without assuming stationarity. The interested reader is referred to [65, pp. 93–95].

4. Smith L. R. [112] addresses the environmental issues, since much of the environmental work involves spatial sampling, and heavy emphasis of spatial statistics.

5. Stein [116, p. 12] recommended the use of the Matérn model due to its flexibility, the ability to model the smoothness of physical processes and possibility to handle non-stationarity. The interested reader on the original work of Matérn is referred to [78].

6. Good discussions of the impact of the directional effect on covariance structure is given by Sherman [107].

If the covariance structure of such process is known, then the distribution is known. A reasonable and acceptable representation of the covariance structure of such process is a separable covariance.

$$\begin{aligned} F\colon \mathbb{R}^2 \times \mathbb{R}^{(+)} &\longmapsto \mathbb{R} \\ (s,t) \quad &\longmapsto Z(s,t) \end{aligned} \tag{2.83}$$

A reasonable covariance structure from exponential model read

$$\operatorname{cov}(Z(s_i,t_i),Z(s_j,t_t)) = \exp\left\{-\theta_s \|h_{i,j}\|\right\} \times \exp\left\{-\theta_t \|t_i - t_j\|\right\} \tag{2.84}$$

This type of presentation plays a central role when expressing the correlation of geophysical phenomena in space and time, for instance the ionosphere or the troposphere variations. These types are often modeled by the first Gauss-Markov processes or random walk process. The interested reader is referred to Gelb [35], and Skone [110].

2.9 Multivariate statistical analysis

Multivariate statistical analysis is the most attractive tool to administrate and analyze the GNSS data, because the GNSS data is actually a multivariate in nature.

Another important characteristic of the multivariate approach is to monitor the spatio-temporal process. Taking the mean vector comparison for instance, we can compare the GNSS atmospheric variations before and after some occurred events.

Let $Y = Y_{i,j,k}$, where $i = 1, \cdots, n_{rec}$ are the reference receivers, $j = 1, \cdots, n_{sat}$ are the satellites observed at each site i, and k is size of the moving window controlling the dynamic of well defined sub-network. The size of the moving window is often chosen to be equal to the correlation length of the observations used.

$$\begin{aligned} \boldsymbol{Y}_1 &= Y_{1,1}\ Y_{1,2}\ Y_{1,3} \cdots Y_{1,k} \\ \boldsymbol{Y}_2 &= Y_{2,1}\ Y_{2,2}\ Y_{2,3} \cdots Y_{2,k} \\ \vdots & \qquad\qquad \vdots \\ \boldsymbol{Y}_i &= Y_{i,1}\ Y_{i,2}\ Y_{i,3} \cdots Y_{i,k} \end{aligned}$$

Our aim is to build a test statistics in order to carry out the inference about the mean and the variance of the population $\boldsymbol{Y}$. That is, seeking for methods that are more flexible and simple to carry out the mean and the variance comparison of two spatial-temporal processes. Such methods have to take into account the following

- Missing observations. Difference in data length shall is not affecting the comparison test.
- Large value of k. Requires enough data to do the comparison.
- Independence: Satellites observations are independent from one satellite to another.

A p-dimensional probability density function $f(y)$, which is parametrized by a mean vector μ and covariance matrix Σ reads:

$$f(\mathbf{z}) = |2\pi|^{-p/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{z}-\boldsymbol{\mu})^{\mathrm{T}} \Sigma^{-1} (\mathbf{z}-\boldsymbol{\mu})\right\} \tag{2.85}$$

2.9.1 Generalized likelihood ratio test (GLRT)

We like to construct a rejection region R which controls the size of error type I based on GLRT. Error Type I is defined as the probability of rejecting H_0 when H_0 is true.

To make the discussions process move smoothly, the univariate case will be treated first. Let our data be presented by a $\{X_i\}_{i=1}^{N}$ and θ is a parameter vector. We will consider two hypotheses:

$$\begin{aligned} H_0 &: \quad \Theta \in \Omega_0 \\ H_1 &: \quad \Theta \in \Omega_1 \end{aligned} \tag{2.86}$$

GLRT measures the closeness of the two hypotheses H_0 and H_1, often the H_0 is regarded as the reduced model and H_1 as complete model. What the GLRT does for hypothesis testing is the same thing as the maximum likelihood does for the estimations process.

Taking the logarithms of likelihood ratio Λ and applying the Wilks' theorem, two times the difference of log likelihoods converges to a χ^2 distribution. Large values of this test statistic indicate a deviance from the null H_0 that leads to reject the null hypothesis H_0. For more details, the reader is referred to [51, Chaps. 5-6].

Multivariate normal random variable (MVNRV)

Moving from univariate to multivariate random variables of dimension p is straightforward. The procedures defined for univariate yield for multivariate distribution. Our start point is the generalization of the squared $t-$distribution given by the Eq. (2.30).

$$T^2 = \sqrt{N}(\bar{\boldsymbol{X}} - \boldsymbol{\mu})\ (\boldsymbol{S})^{-1}\ \sqrt{N}(\bar{\boldsymbol{X}} - \boldsymbol{\mu}) \tag{2.87}$$

This is equivalent to:

$$\begin{aligned} T^2_{p,N-1} &= \left\{\text{MVNRV}\right\}^T \quad \left\{\frac{\text{Wishart random matrix}}{\text{degree of freedom}}\right\}^{-1} \quad \left\{\text{MVNRV}\right\} \\ &= \mathcal{N}_p(\mathbf{0},\boldsymbol{\Sigma})^T \quad \left\{\frac{1}{N-1} W_{p,N-1}(\boldsymbol{\Sigma})\right\} \quad \mathcal{N}_p(\mathbf{0},\boldsymbol{\Sigma}) \end{aligned} \tag{2.88}$$

Note that the Wishart distribution is equivalent to chi-square in univariate random variables.

Hotelling's T^2 can be considered as the basic of several multivariate control charts. The sampling distribution of the Hotelling's T^2 has the same shape as an F-distribution. To be more specific:

$$T^2 = \frac{(N-1)\,p}{N-p} F_{p,N-p}(\alpha) \tag{2.89}$$

where $F_{p,N-p}$ denote the random variable with Fisher $F-$distribution and with p and $N-p$ degree of freedom.

Application of multivariate analysis in GNSS is huge, for instance comparing the ionospheric and tropospheric models. The procedure is simple, one start by constructing a grid model where we have registered the values generated by both models for ionosphere or for troposphere. Then we generate p time-series of length N, where p is also the number of grid points. At α level of significance, we can investigate if both models have the same level or vary differently.

Inference about Mean Vectors

Inference about the mean vectors at some level α rely on the assumption that the population variance-covariance matrices of both populations are equal. Test for equality of mean vectors is defined as follow:

$$\begin{aligned} H_0 &: \quad \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \\ H_1 &: \quad \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2 \end{aligned} \tag{2.90}$$

In order to carry out the inference about the mean vectors, a test statistic T^2 shall be derived. Let $\{X_{1i}\}_{i=1}^{N_1} \sim \mathcal{N}_p(\boldsymbol{\mu}_1, \Sigma)$ and $\{X_{2i}\}_{i=1}^{N_2} \sim \mathcal{N}_p(\boldsymbol{\mu}_2, \Sigma)$.

$$T^2 = \left[\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\right]^T \left[\left(\frac{1}{N_1} + \frac{1}{N_2}\right) \boldsymbol{S}_{\text{pooled}}\right]^{-1} \left[\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\right] \tag{2.91}$$

is distributed as

$$\frac{(N_1 + N_2 - 2)\, p}{(N_1 + N_2 - p - 1)} F_{p, N_1 + N_2 - p - 1} \tag{2.92}$$

where the pooled variance, S_{pooled} is defined by the expression:

$$\boldsymbol{S}_{\text{pooled}} = \frac{\sum_{i=1}^{N_1} (X_{1i} - \bar{X}_1)(X_{1i} - \bar{X}_1)^T + \sum_{i=1}^{N_2} (X_{2i} - \bar{X}_2)(X_{2i} - \bar{X}_2)^T}{N_1 + N_2 - 2} \tag{2.93}$$

Consequently

$$P\left\{ T^2 \leq \frac{(N_1 + N_2 - 2)\, p}{(N_1 + N_2 - p - 1)} F_{p, N_1 + N_2 - p - 1}(\alpha) \right\} = 1 - \alpha \tag{2.94}$$

Why the multivariate methods is absolutely the right instrument to carry out the NRTK data integrity solution

Statistical method refers to the process of analyzing, interpreting, displaying, and decision making based on the observed data. Therefore the multivariate statistical analysis are collection of methods and procedures.

GNSS data is a pure multivariate in nature. All techniques and methods of multivariate will be used to carry out the quality control of the information generated the GNSS systems.

Concerning the NRTK data integrity solution, a strong criterion is the trade-off between the quick detection of anomalies and fewer generated alarms.

Any way, the multivariate statistical analysis is a powerful method and is chosen to carry out the quality control. Determination of the sample distribution of the statistics is considered as a good start.

Small sample hypothesis testing of T^2

Let $\boldsymbol{X} \sim \mathcal{N}_p(\boldsymbol{\mu}, \Sigma)$ and a random sample N of observations is collected, where $N - p < 40$. Then testing the population mean vector $\boldsymbol{\mu}$ is equal to target $\boldsymbol{\mu}_0$ goes as follow:

$$\begin{aligned} H_0 &: \quad \boldsymbol{\mu} = \boldsymbol{\mu}_0 \\ H_1 &: \quad \boldsymbol{\mu} \neq \boldsymbol{\mu}_0 \end{aligned}$$

The test statistics T_0^2 is given by Eq. (2.87) and is distributed as $\frac{(N-1)p}{N-p} F_{p,N-p}(\alpha)$, where $F_{p,N-p}$ denotes the random variable with Fisher F distribution and with p and $N - p$ degree of freedom. We reject the null hypothesis is if:

$$T_0^2 \geq T_c^2 = \frac{(N-1)\, p}{N-p} F_{p,N-p}$$

This means that T_0^2 is no longer distributed as $\dfrac{(N-1)\, p}{N-p} F_{p,N-p}$ and its value will be significantly larger.

For large sample ($N - p > 40$), the distribution of $T_0^2 \sim \chi_p^2$ and we reject the null hypothesis if the value of $T_0^2 > \chi_p^2$.

Variance-covariance comparison

The main objective of such a test is to find out if two different models behave in the same way. Models can be any spatio-temporal model administrated by the multivariate methods, for instance the ionospheric models or tropospheric.

Let our observations be generated by two Gaussian processes with different means and variance-covariance matrices

$$\boldsymbol{X}_{ih} \sim \mathcal{N}_p(\boldsymbol{\mu}_h, \Sigma_h) \quad i = 1, 2, \cdots, N_h \quad \text{and} \quad h = 1, 2 \tag{2.95}$$

Testing for equality of Variance-Covariance can be written as:

$$\begin{aligned} H_0 &: \quad \Sigma_1 = \Sigma_2 \\ H_1 &: \quad \Sigma_1 \neq \Sigma_2 \end{aligned} \tag{2.96}$$

$\boldsymbol{S}_h$ is a sample estimation of Σ_h for $h = 1, 2$ with Wishart distribution.

$$N_h S_h \sim W_p(\Sigma_h, N_h - 1) \tag{2.97}$$

The common variance-covariance matrix (weighted overage of both sample variance-covariance matrices), $\boldsymbol{S} = \frac{N_1}{N_1+N_2}\ \boldsymbol{S}_1 + \frac{N_2}{N_1+N_2}\ \boldsymbol{S}_2$ is also Wishart distributed.

$$\sum_{h=1}^{2} N_h\ \boldsymbol{S}_h \sim W_p(\Sigma, N-2) \tag{2.98}$$

The Likelihood Ratio Test (LRT) leads to the test statistic:

$$T_s = -2\log(\Lambda) \tag{2.99}$$

$$= N\log(|\ \boldsymbol{S}\ |) - \sum_{h=1}^{2} N_h \log(|\ \boldsymbol{S}_h\ |) \tag{2.100}$$

where $N = N_1 + N_2$ and $\boldsymbol{\Lambda}$ is obtained from Wilks theorem. Under H_0, the test statistics T_s is approximately distributed as χ^2_m where the degree of freedom $m = (1/2)(N-1)\ p\ (p-1)$.

The procedure is simple, from data:

- Compute $\boldsymbol{S}, \boldsymbol{S}_1, \boldsymbol{S}_2, |\boldsymbol{S}_1|, |\boldsymbol{S}_2|$, and $|\boldsymbol{S}|$
- compute LRT $= -2\log(\boldsymbol{\Lambda})$

At α significance level, we can test whatever if we reject the null hypothesis and conclude that the variation generated by both models are different.

Some important remarks to consider when testing for equality of two matrices in general:

- Comparing two matrices of the same size to assure the homogeneity variance-covariance can easily be accomplished in numerical linear algebra by calculating the difference of both matrices and examine the entries of differenced matrix. In case all entries are close to zero, then we achieve the homogeneity.
- The cosine of the angle between two vectors is directly related to the correlation between the vectors. We can apply this idea to measure the similarity of two matrices.

Introduction to GNSS

3

"The first data set is always garbage."

— Ron Snee (2014)

This chapter provides an introduction to the principles behind Global Navigation Satellite Systems (GNSS). In particular, we give a quick overview of how the GPS system works; present the most common error sources and how to mitigate them; and the different approaches to GNSS positioning.

3.1 Introduction

The GNSS systems – including GPS, GLONASS, Galileo, Beidou, and so on – are all based on the same principle of timed signals. The ranging signal, the navigation data messages and the radio frequency carrier in L-band are sent from satellite(s) and received by the user receiver, which generate observables such as code and carrier-phase pseudo-ranges. The determination of the position at the observed site is achieved at any place at the Earth surface or in space.

For multi-GNSS receivers, signals can be received from many constellations and the observations can be processed together, which come to benefit to the users. This will improve the satellite geometry, position determination and the navigation performance parameters (accuracy, continuity, availability, and integrity) and system reliability.

It is impossible to cover all GNSS systems in a single chapter. We limit our discussions to the GPS system, because it is stable and provides a global coverage. Various text books on the GPS theory and applications are given with different emphasis. Starting by the classical, Strang and Borre [118, Chaps. 14-15] offer a gentle introduction to GPS with matlab code for experimentation, Seeber [105, Chap. 7] introduces the GPS system in general and introduction to the theory and practice of GPS is given by Hofmann-Wendolf [45]. A detailed signal structure description is given by Pratap and Enge [80, Part 3], and Kaplan and Hegarty [54, Chaps. 4–5]. Leick [68, Chap. 4] provides an elegant statistical method for parameter estimation and quality control of the GNSS error sources. Xu [144] gives a detailed description on data processing using robust Kalman filter and least square method. Prasad and Ruggieri [88] gives a detailed description of how the integration of GNSS can be done and meet the future challenges.

Modern and next-generation books on GNSS and their applications, recommended by the expert in the field and aimed at the industries are Leick [69], and P.J.G. Teunissen and O. Montenbruck [124].

Since the signal decoding techniques has shifted from binary phase shift keying (BPSK) to binary offset carrier (BOC) modulation, the technology of the GNSS receivers has changed to adapt to this new flavor of signals. Interesting textbooks covering the signal acquisition with different emphasis are given respectively by Forssell [32, Chap. 11] spread spectrum, Borre [11] for GPS, and by Jaizki and el. [101] for GPS and Galileo radio frequency (RF). Finally, development of efficient GNSS receiver algorithms to work with weak signals are treated by Zeidan [146].

The rest of this chapter is organized as follow; Section 3.2 introduces the GPS constellation and signals. Next, Section 3.3 gives a brief introduction to the GPS observables and error sources mitigation. Sections 3.4 – 3.6 give a short introduction to GNSS data processing, including the observations combinations, differencing and GNSS functional and stochastic models. Section 3.7 presents positioning with GPS. Section 3.8 presents a multipath mitigation algorithm based on Kalman filter. Finally, Section 3.9 presents the TEC monitoring algorithm.

3.2 GPS constellation, segments and signals

NAVigation Satellite Timing And Ranging(NAVSTAR) GPS is a satellite navigation system capable of providing accurate, continuous global positioning and navigation services. The system consists of 31 operational satellites in space, approximately uniformly dispersed around six circular orbits with four or more satellites each. The orbits are inclined at an angle of 55° relative to the equator and separated from each other by multiple of 60°. The nominal orbital period is 11 hours, 58 minutes or one-half sidereal day.

The definition of the satellite constellation for navigation system is an optimization problem. Maximizing the cost function $C(.)$ with respect to key performance parameters Θ. Our aim is to obtain a constellation with a global coverage worldwide with a minimum number of satellites. Intensive efforts and studies has been conducted in this field in order to determine the fewest satellites for a given constellation and it turns out that the generalization of Walker constellation [136] is the most appropriate method of solving the optimization problem. In addition to inclination angle i and the period T, three other parameters are needed to describe the constellation and are given by the triplet called the Walker notation $N/P/F$, where N is the number of satellite in the constellation, P is the number of orbital planes, and F is the phasing factor ($F = 0, 1, \cdots, P-1$).

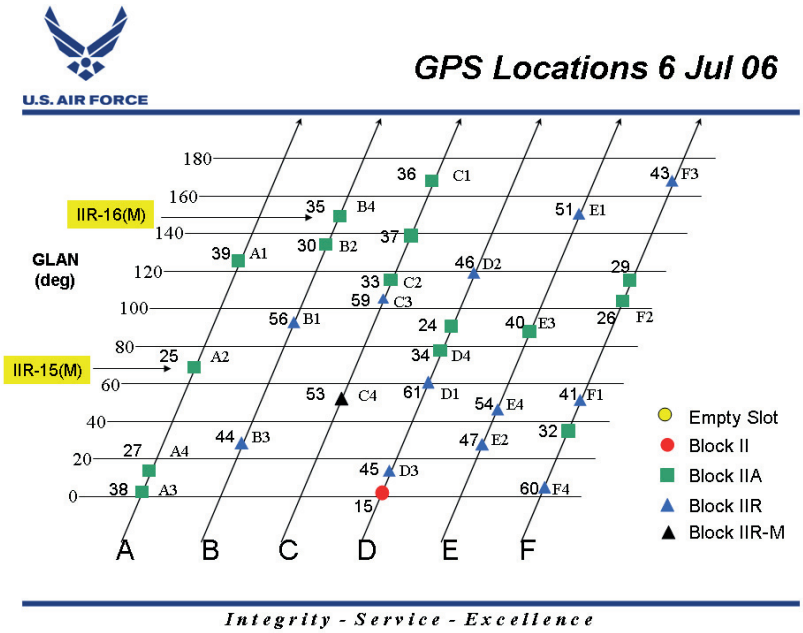


Figure 3.1: GPS satellites constellation.
The right of the figure shows the different generation of GPS satellites.
Image credit::https://en.wikipedia.org/wiki/File:GPS_constellation_6_jul_06.png

Other constellation techniques exists, for instance Rider [92] studied the class of orbit that are circular with equal latitude and inclination.

For more discussions on constellation design for satellite navigation with focus on key parameters performance, for instance the effects of layers of a constellation, the number of satellites, orbit types, and the orbital altitude are given in [54, p. 49]. Theoretically, three or more GPS satellites will always be visible anywhere on the surface of the Earth. Figure 3.1 gives the GPS satellite constellation.

The heart of the GPS satellite is the precise atomic clocks, two rubidium and two cesium, that are used to generate two coherent carriers L1 and L2 in the L-band. The GPS satellites transmit on two L-band frequencies with one at 1575.42 MHz (L1) and the other on 1227.60 MHz (L2). These frequencies are integral multiples $f_1 = 154 f_0$ and $f_2 = 120 f_0$, where the fundamental frequency $f_0 = 10.23$ MHz.

3.2.1 GPS segments

The GPS system is composed of four segments:

- Space Segment (SS): Includes the satellite constellation of GPS satellites and is run by the US Air Force, responsible for operation and maintenance of the system. The main Control Centre is at Falcon Air Force Base, Colorado Springs, USA.

- The Control Segment (CS): Responsible for monitoring the space segment by using signals from channels L1 and L2 to estimate and predict the satellite orbits and clock errors. This information is uploaded to the satellite, which broadcasts the navigation message part by part in frames or sub-frames to the Earth (users).

- The User Segment (US): Different types of GPS receivers and users of various applications are considered the user segment of the GPS. The receiver uses the signals and navigation messages to compute the position, velocity and precise time.

- The Ground Segment (GS): Includes civilian tracking networks that provide the user segment with reference control, precise ephemeris, and real-time services (DGPS).

3.2.2 GPS signals

From the user point of view, the signals are the most interesting part of the GPS system and can be regarded as the main interface component to the user.

- Coarse/Acquisition(C/A) code: Modulated on the carrier L1. Each satellite has a different C/A code, so that they can be uniquely identified.

- Precision(P) code: Modulated on the carrier L1 and L2 and is better for more precise positioning.

- Navigation Message: Can be found on L1 and L2 and includes information on the broadcast ephemeris, and used to determine the satellite orbital parameters, satellite clock corrections, almanac data, ionosphere information, and satellite health information status.

GPS new civil signals

Due to the advance in satellite and receiver technology, the modernization of GPS becomes possible. The improvement of the space and ground segments allow the performance and accuracy of the GPS to improve radically. The additional signals (L1C, L2C and L5) transmitted by modernized satellites improve the accuracy and the performance. An excellent review of the subject is provided by Leick [69, p. 243] which includes an extensive list of references.

- L1C (1575.42 MHz): L1C signal contains pilot and data channels. Both signals use Binary Offset Carrier (BOC) modulation techniques. The data and pilot channels are multiplexed in a such a way that the pilot channel has 75% of the power and the data channel has 25%. That is, L1C include advanced design for enhanced performance and will be broadcasted at a higher power level.
- L2C (1227.6 MHz): L2C is the second civil frequency and enables the development of lower-cost, dual-frequency civil GPS receivers that allow for correction of ionospheric path delay.
- L5 (1176.45 MHz): L5 is the new third civil frequency and will transmit at a higher power than current civil GPS signals, and have a wider bandwidth.

3.3 GNSS error sources

Degradation of GNSS signal can be grouped into three main categories, satellite dependent errors, signal propagation and user receiver dependent errors. The atmosphere affect the GNSS signals in three different ways; bending, absorption and ionization. From Eqs. (3.20) and (3.21), the GNSS measurements are subject to many error sources, which degrades the obtainable accuracy of the user positions. In order to process the GNSS data, it is absolutely desirable to understand the nature of the error sources and how to model, mitigate and estimate them.

3.3.1 Mitigation of the ionospheric path delay

The ionosphere is part of the Earth's upper atmosphere and it's layer extends from approximately 50 to 1000 km above the surface of the Earth. It consists of gases that have been ionized by solar radiation. The ionization process produces a clouds of free electrons that act as a dispersive medium for GNSS signals (velocity is a function of frequency). It is a pure spatio-temporal

process. The spatial and temporal variability has a very significant impact on GNSS signals traveling from satellites and reaching the user receivers. The dynamics of the ionosphere has remarkable characteristics, the structure and the peak electron density vary strongly with time, geographic location (equator, North and South hemisphere). Moreover, it's variability can be classified as follow:

- Diurnal variation: vary with time of the day.
- Seasonal variation: season of the year.
- Periodical variation: 11 years sunspot cycle.
- Geographical location: Latitude variation.

In addition, the solar radiation strikes the atmosphere with an average power density of $1.37 kw/m^2$. This value is known as the solar constant. When solar radiation strikes the atmospheric molecules, these will absorb part of this radiation and a free electron and a positive ion are produced. The ionosphere path delay in GPS pseudo-range observations is the largest error source and can retard radio waves from their velocity in free space by more than 300 ns, on a worst case basis, corresponding to a range error of 100 m. A GNSS signal exhibits code delay and carrier-phase advance when traversing the ionosphere. The important parameter for ionosphere path delay is the total number of electrons (TEC) encountered by the radio waves on its path from satellite to the GNSS receiver. The slant total electron content (TEC) is defined as the electron density integrated along the LOS path. The TEC is expressed in TEC units (TECU), where 1 TECU is equal to 10^{16} electrons in a cylindrical volume with a 1 m^2 cross section aligned along the LOS.

$$\mathrm{TEC} = \int_s^r N_e(s)ds \tag{3.1}$$

The slant TEC (STEC) is calculated along the LOS from a satellite to a receiver and therefore is a function of the satellite and the receiver antenna position. Any way, we are interested in TEC along the local vertical, namely the VTEC. The mapping function that maps the STEC to the VTEC is needed. All users have to correct for this error, and the technique used differ based on the user's GNSS receiver.

A. Single Layer Model

The ionosphere is often modeled as a SLM shown in Figure 3.2. The model well known as the thin-shell model. SLM assumes that all free electrons

are concentrated in a shell of infinitesimal thickness. The height of this idealized layer is usually correspond to the expected height of the maximum electron content. Usually, the value of 350 – 450 km are used.

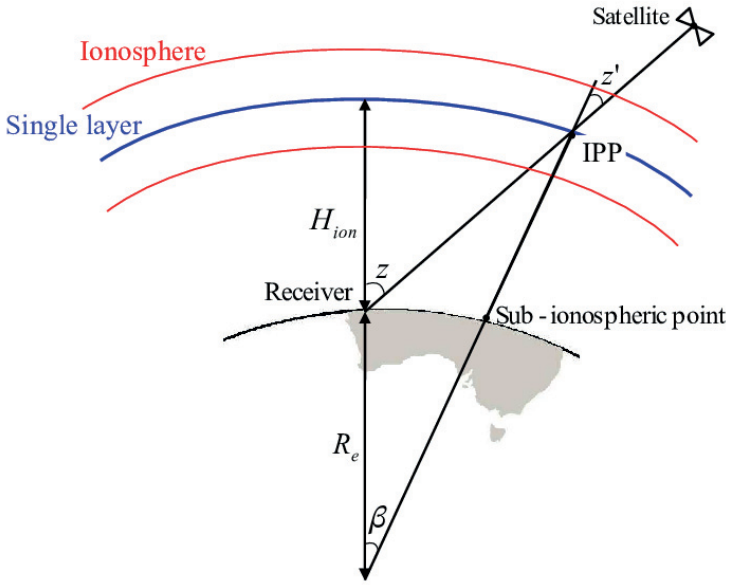


Figure 3.2: Single Layer Ionosphere Model

The derivation of the latitude ϕ_{ipp} and longitude λ_{ipp} of the IPP is done by the following expression.

$$z = \frac{\pi}{2} - E_l \tag{3.2}$$

$$z^{'} = \arcsin\left\{\frac{R}{R+H}\sin(z)\right\} \tag{3.3}$$

$$\beta = z - z^{'} \tag{3.4}$$

$$\phi_{ipp} = \arcsin\left\{\cos(\beta)\sin(\phi_r) + \sin(\beta)\cos(\phi_r)\cos(A_r^s)\right\} \tag{3.5}$$

$$\lambda_{ipp} = \lambda_r + \arcsin\left\{\frac{\sin(\beta)\sin(A_r^s)}{\cos(\phi_{ipp})}\right\} \tag{3.6}$$

where R is the radius of the Earth. E_l is the elevation angle. H is the height of the ionosphere shell, usually 350 km. The (ϕ_r, λ_r) are longitude and latitude of the receiver. A_r^s is the azimuth angle. The model approximate the shell as a sphere.

B. Linear combinations of observables

Due to the dispersive nature of the ionosphere, forming a geometry-free linear combination Eq. (3.33) of L1 and L2 pseudo-range or carrier-phase measurements, the ionospheric delay can be corrected for dual frequency receivers. The ionosphere effects in the GNSS signal measurements can be completely eliminated by forming the ionosphere-free linear combination, Eq. (3.32).

C. Broadcast model

For single frequency receivers, the ionosphere broadcast algorithms developed by J. A. Klobuchar [59] can be used to reduce the effect of the ionosphere. It can correct up to 50%. The algorithm assumes that all electrons are concentrated in a single layer at 350 km above the Earth surface. The algorithm is simple and take into consideration the number of coefficients used and the computation time. All angles are in units of semi-circle, and the time is in seconds.

1. Compute Earth- centered angle Ψ

$$\Psi = \frac{0.0137}{E_l + 0.11} - 0.022 \quad (semicircles)$$

2. Compute the sub-ionosphere latitude

$$\Phi_I = \phi_U + \Psi \cos(A_z) \quad (semicircles)$$

3. Compute the sub-ionosphere longitude

$$\lambda_I = \lambda_U + \frac{\Psi \sin(A_z)}{\cos(\Phi_I)} \quad (semicircles)$$

4. Find the geomagnetic latitude

$$\Phi_m = \Phi_U + 0.064 \cos(\lambda_I - 1.617) \quad (semicircles)$$

5. Find the local time:

$$t = 43200 \frac{\lambda_I}{\pi} + t_{\text{GPS}}$$
$$\text{if} \quad t > 86400 \quad \Rightarrow \quad t = t - 86400$$
$$\text{if} \quad t < 0 \quad \Rightarrow \quad t = t + 86400$$

6. Compute the slant factor:

$$F = 1.0 + 1.6(0.53 - E_l)^3$$

7. Compute the ionospheric time delay :

$$\begin{aligned} T_{iono} &= F.\left[5.10^{-9} + \sum_{n=0}^{3} \alpha_n \left(1 - \frac{x^2}{2} + \frac{x^4}{24}\right)\right] \\ \text{where } x &= \frac{2\pi(t - 50400)}{\sum_{n=0}^{3} \beta_n \Phi_m^n} \end{aligned}$$

The α_n and β_n coefficients are uploaded by the master station to the satellites and broadcasted from the satellites to the user in navigation message. In addition, the algorithm uses the elevation angle E_l of satellite and azimuth angle A_z. These parameters must be calculated before calling the algorithm to estimate the ionospheric time-delay.

The Klobuchar broadcast algorithm is very popular and is used in more sophisticated ionospheric monitoring algorithms, for instance in interpolation algorithm Delay Inverse Distance Weighting (IDW) implemented in augmentation system like WAAS, EGNOS, GAGAN systems. IDW uses Klobuchar as apriori information about IPP and IGP points. Chinese satellite navigation system, Beidou, uses the Klobuchar broadcast model while Galileo uses the NeQuick model. No ionospheric model is broadcast by GLONASS satellites.

D. Global Ionospheric Map

The Centre for Orbit Determination in Europe (CODE), located at the University of Berne in Switzerland, generates daily global ionospheric maps (GIMs) based on data collected from up to 300 globally distributed GNSS stations and using spherical harmonics for interpolation, Schaer [103].

Maps are grid based ($2.5° \times 5°$) and updated every two hours.

Figure 3.3 presents a sample example of a global VTEC map at 12:00 UT on 21th March, 2009 form the CODE Analysis Centre Global Ionospheric Map (GIM), available at ftp://ftp.unibe.ch/aiub/CODE/)

3.3.2 Troposphere error

The lower part of the atmosphere, called the troposphere, is electrically neutral and non-dispersive for frequencies as high as about 15 GHz. The raw tropospheric delay in pseudo-range units depends on the vertical delay related to the dry and wet components. About 90% of the tropospheric

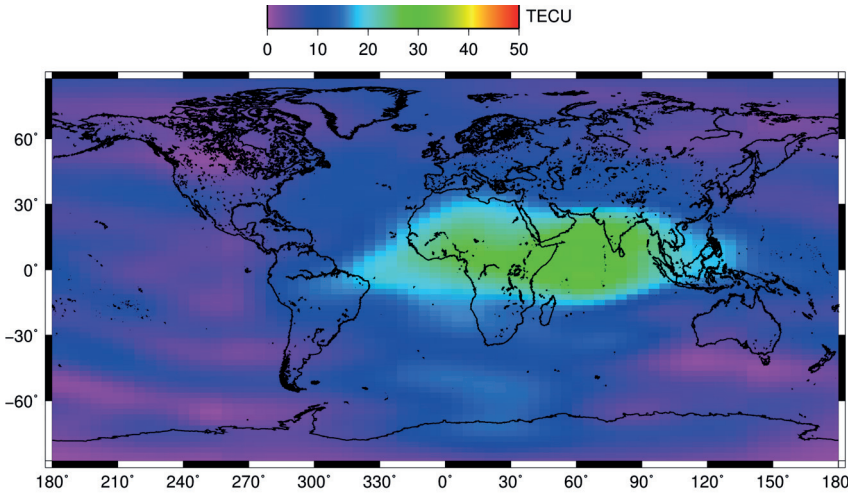


Figure 3.3: Image credit: Royal Observatory of Belgium (ROB) GNSS Research Group

refraction arise from the dry component and about 10% from the wet component. Due to the temperature, pressure and humidity variations of the atmosphere, the wet component is difficult to model.

Besides the path delay caused by the ionosphere, the troposphere path delay is considered as a threat to precise GNSS applications. The aim is to correct the tropospheric delays as accurately as possible by identification of the correct empirical tropospheric model with a mapping function. Wang and Li [138] showed that there is no practical difference between mapping functions with or without meteorological data for wet component derivation. Since the troposphere is height dependent, Feng [138] applied a decorrelation algorithm to achieve improvement of z–component.

Several models exist for correction of troposphere path delay. The most popular used in this thesis is the empirical Saastamoinen model [100].

$$\Delta_{trop} = \frac{0.002277}{\cos(z)} \left\{ p + \left(\frac{1255}{T} + 0.05 \right) e - \tan^2(z) \right\} \tag{3.7}$$

where p, T, and e are the total pressure in hPa, absolute temperature of the air in Kalvin, and partial pressure of water vapor in hPa, respectively, derived from the standard atmosphere. h_{rel} is the relative humidity and h is the geodetic height above the MSL.

$$P = 1013.25(1 - 2.2557 \times 10^{-5} \times h)^{5.2568} \tag{3.8}$$

$$T = 15.0 - 6.5 \times 10^{-5} + 2.73.15 \tag{3.9}$$

$$e = 6.108 \times \exp\left\{\frac{17.15T - 4684.0}{T - 38.45}\right\} \times \frac{h_{rel}}{100} \tag{3.10}$$

Tropospheric delay modeling

The tropospheric path delay is shortest in the zenith direction. The GNSS signal takes a long time to traverse the air mass. Due to the spatial and temporal variations of the troposphere, it is difficult to define the exact functional relationship caused by this dependency. Often we define the mapping functions to describe the dependence of the slant hydrostatic (SHD) and the slant wet delays (SWD) into their corresponding zenith delays, ZHD and ZWD, respectively. Note that the mapping functions are always a function of the elevation angle E_l of the slant signal.

Denote by z the zenith angle, then the computation of SHD and SWD reads:

$$\text{SHD} = \text{ZHD} \times M_h(z) \tag{3.11}$$

$$\text{SWD} = \text{ZWD} \times M_w(z) \tag{3.12}$$

The slant total delay (STD) caused by the troposphere is then equal to sum of two quantities, SHD, and SWD, respectively.

$$\text{STD} = \text{SHD} + \text{SWD} \tag{3.13}$$

Zenith dry delay

Several mapping functions have been developed in the past. The most used one is the Niell [82] mapping function (NMF). The NMF is accurate, independent of the meteorological data and take into consideration the time of the year.

$$M_N(z) = \frac{1 + \dfrac{a}{1 + \dfrac{b}{1 + c}}}{\cos(z) + \dfrac{a}{\cos(z) + \dfrac{b}{\cos(z) + c}}} + h_{[km]} \left\{ \frac{1 + \dfrac{a_h}{1 + \dfrac{b_h}{1 + c_h}}}{\cos(z) + \dfrac{a_h}{\cos(z) + \dfrac{b_h}{\cos(z) + c_h}}} \right\} \tag{3.14}$$

Table 3.1: Coefficients for Niell's DMF.

ϕ	$\tilde{a} \times 10^3$	$\tilde{b} \times 10^3$	$\tilde{c} \times 10^3$
≤ 15	1.2769934	2.9153695	62.610505
30	1.2683230	2.9152299	62.837393
45	1.2465397	2.9288445	63.721774
60	1.2196049	2.9022565	63.824265
≥ 75	1.2045996	2.9024912	64.258455
ϕ	$a_p \times 10^5$	$b_p \times 10^5$	$c_p \times 10^5$
≤ 15	0	0	0
30	1.2709626	2.1414979	9.0128400
45	2.6523662	3.0160779	4.3497037
60	3.4000452	7.2562722	84.795348
≥ 75	4.1202191	11.723375	170.37206
	$a_h \times 10^3$	$b_h \times 10^3$	$c_h \times 10^3$
	0.0000253	5.49	1.14

The coefficients of Eq. (3.14) are provided by tables 3.1 and 3.2, respectively, as a function of station latitude ϕ. If $\phi \in]15, 75[$, the interpolation algorithm must be applied. In addition, the coefficients a, b and c must be corrected for periodic terms by application of the Eq. (3.15).

$$a(z, DOY) = \tilde{a} - a_p \cos\left\{2\pi \frac{DOY - \mathrm{DOY}_0}{365.25}\right\} \tag{3.15}$$

where DOY is the day of year and DOY_0 is equal to 28 or 211 depending on the station location in the Southern or Northern hemisphere, respectively.

Table 3.2: Coefficients for Niell's WMF.

ϕ	$a \times 10^4$	$a \times 10^3$	$c \times 10^2$
15	5.8021897	1.4275268	4.3472961
30	5.6794847	1.5138625	4.6729510
45	5.8118019	1.4572752	4.3908931
60	5.9727542	1.5007428	4.4626982
75	6.1641693	1.7599082	5.4736038

Zenith wet delay

One precise way to measure the accuracy of wet delay is by water vapor radiometer. Methods used in GNSS are using empirical wet delay correction and the performance depends on the accuracy of the zenith wet delay model and the wet mapping function. Due to the non-uniform distribution of water vapor horizontally and vertically, the determination of the covariance structure of the spatio-temporal process in question is very complicated. Often we end by imposing stationarity. Anyway, the process noise of the troposphere wet component is modeled as a random walk or first-order Gauss-Markov processes, where the covariance is an exponential function. A Kalman filter can be used for this purpose.

3.3.3 Orbital error

Satellites positions are calculated using the broadcast navigation message [54, p. 42], which includes Keplerian elements and time derivative of some parameters, known as the time of epoch or time of ephemeris. The computation of the navigation message is carried out by averaging the measurements generated at different monitor stations distributed around the globe, then the navigation messages are uploaded to satellites every two hours approximately. In order to minimize the orbital errors, one remedy is to increase the number of monitor stations and decrease the uploading time, updating the GPS ephemeris more frequently. For post-processing applications, one can use the precise orbits rather than navigation message. The computation of precise orbits are carried out by using huge data sets collected at many reference stations before and after the time period. The interpolation algorithms for instance Lagrange interpolating polynomial [14, p. 100] is commonly used to compute the satellite position for any given time t. The precise orbits are delivered in SP3-x format from IGS. In order to determine the satellite position errors one can compute the difference between the satellite position computed using the broadcast navigation message and the precise orbits.

3.3.4 Satellite clock error

The satellite clock offset can be eliminated by forming a single difference of two carrier-phase (or code) observations.

The GPS time of transmission of the GPS message is $t = t_{sv} - \Delta t_{sv}$, where the t_{sv} is the SV PRN (Space Vehicle Pseudo Random Noise) code phase time at the time of transmission. The satellite clock correction term Δt_{sv} is

approximated by a second order polynomial:

$$\Delta t_{sv} = a_{f0} + a_{f0}(t - t_{oc}) + a_{f2}(t - t_{oc})^2 + \Delta t_R$$

Here a_{f0}, a_{f1},a_{f2} are the polynomial correction coefficients corresponding to phase error, frequency error and rate of change of frequency error (in second, second per second and second per square second), and t_{oc} is the clock data reference time in seconds. The interested reader is referred to ICD [48, pp. 88–90] for more details, including the explanation and derivation of the group delay.

3.3.5 Receiver clock error

The receiver clock offset can be eliminated by forming a double-difference, using the technique described in Section 3.5.

3.3.6 Multipath

GNSS signal reaching the receiver from different paths in addition to direct signal is called multipath and is caused by reflections from surrounding objects (near the receiver). As a result, it is highly dependent upon the conditions surrounding the receiver antenna, the type of antenna, and the internal tracking loop algorithms of the receiver. This makes multipath a difficult error to remove, because there is no model that can be used for the general case. The received signal at the receiver is composed of direct and indirect signals, S_d and S_i, respectively.

$$S(t) = S_d(t) + \sum_{i=1}^{N} S_i(t) \tag{3.16}$$

Indirect signal differs from the direct one in amplitude, phase and in delay (delayed/advanced) depending on surrounding objects reflection characteristics. These differences destroy the shape of the auto-covariance function and make the process to trace the peak very difficult. As a result, reduction of the effect of the multipath can be carried out by any algorithm that is capable to construct all important parts of the auto-covariance function in the tracking loop. Some important characteristics of multipath are:

1. Multipath signal will always arrive after the direct path signal because it must travel a longer propagation path.

2. The multipath signal will normally be weaker than the direct path signal since some signal power will be lost from the reflection. It can be stronger if the direct path signal is hindered in some way.

3. If the delay of the multipath is less than two pseudo-random noise (PRN) code chip lengths, the internally generated receiver signal will partially correlate with it. If the delay is greater than two chips, the correlation power will be negligible

4. Multipath delays are periodic signal with zero average with time correlation period sufficiently large, up to a couple of minutes.

Since the multipath error is not modeled in the GNSS functional model, several methods are proposed to reduce the effect of the multipath, for instance Souza et al. [114] proposed the use of wavelet method to correct for multipath effect. The interested reader on how to reduce the code multipath errors is referred to [80, Chap. 10].

Improvement of multipath

The improvement of multipath can be achieved by acting at different areas:

1. GNSS receiver: Adding more correlators in tracking loop algorithms in order to construct the damaged auto-covariance function and detect the peak.

2. Receiver antenna: Choke ring that damp signals from a low elevation angle (E_l).

3. Surrounding area: Choosing place for placement for the receiver where less refractions are possible.

The measurement bias caused by signal multipath acts differently. Unlike the other error sources, multipath is normally uncorrelated between antenna locations. Hence, the base and remote receivers experience different multipath interference, and differencing between them will not cancel the errors. Also, modeling of multipath for each antenna location is difficult and impractical. In the presence of multipath, most GPS positioning methods suffer a degradation in accuracy and an increase in processing time. Pseudorange multipath at a real-time differential GPS (DGPS) monitor station will result in errors creeping into the differential corrections, causing large position biases for DGPS users.

Multipath is serious for static receiver rather than for the mobile receiver. Since the investigation of multipath remains an active area of research, we propose a mitigation algorithm in Section 3.8 based on Kalman filter.

3.3.7 Receiver measurement noise

This type of noise refers to any noise generated by the receiver itself in the process of tracking the code or carrier-phase measurements. It is considered to be Gaussian distributed with zero mean and finite variance σ^2. That is, if ϵ denotes the receiver measurement noise, then $\epsilon \sim \mathcal{N}(0, \sigma^2)$.

Testing for the white noise is carried out by zero baseline test where two GNSS receivers are connected to the same antenna. We give some important facts:

- The measurement noise increases as the quality of the received signal decreases, and it's given by the SNR value.
- The variance increases in case of any linear combinations of observations or under application of the difference techniques. Let $\Delta\nabla\Phi_{ab}^{xy}$ be the double-difference between two reference receivers a and b observing the same satellites x and y. The noise component reads:

$$\Delta\nabla\epsilon_{ab}^{xy} = \epsilon_b^x - \epsilon_b^y - \epsilon_a^x + \epsilon_a^y \tag{3.17}$$

  Assuming that the variance of each line of sight are uncorrelated and constant, the variance is computed by the expression:

$$\begin{aligned} \mathrm{var}(\Delta\nabla\epsilon_{ab}^{xy}) &= \mathrm{var}(\epsilon_b^x) + \mathrm{var}(\epsilon_b^y) + \mathrm{var}(\epsilon_a^x) + \mathrm{var}(\epsilon_a^y) \\ &= 4\sigma^2 \end{aligned} \tag{3.18}$$

  The standard deviation under double-difference operation increased by a factor of 2. In case of single difference operation, the increasing factor is $\sqrt{2}$.

3.4 GNSS data processing

The aim of this section is to introduce the GNSS data processing methods. In order to carry out the precise positioning, the carrier-phase observables are used. GNSS observation equations, linear combinations of the observables, and difference techniques are introduced. The mathematical functional and stochastic models are given including the process noise.

The parameters estimation by means of least-square, weighted and general least-squares, filtering, smoothing and forecasting will be introduced.

The user accuracy depends strongly on the various interactive factors. In order to have a real-feeling on range determination, the reader is referred to Kaplan [54, Figure 7.1]. The figure describes the different range measurement timing relationship.

3.4.1 GNSS observations equations

The observation types provided by the GPS receivers depend strongly on the technique used. A single frequency receiver can only provide the user with observation on L_1 frequency, while code-less receivers can provide observations on dual frequencies (L_1 and L_2). There are three basic types of GNSS observables, the code pseudo-range, carrier-phase and Doppler measurements. Suppose that the C/A, P, or Y-code (Encrypted P code) are transmitted by satellite s at emission time t_e and registered by receiver r at time t_r. The fundamental observation equation is defined by:

$$\rho_r^s(t_r, t_e) = c\ (t_r - t_e) = c\ \tau \tag{3.19}$$

where

$\rho_r^s(t_r, t_e)$	:	pseudo-range, expressed in unit of length
c	:	speed of light.
t_e	:	transmission or emission time of the signal.
t_r	:	observation/receiving time of the signal.
τ	:	signal traveling time.

Omitting the instrumental biases and taking into account the satellite and the receiver clock errors, the ionospheric effects, the tropospheric path delay, Earth tide and loading effects, multipath and the effects of the relativity, then the complete model for code pseudo-range observables reads:

$$P \quad = \quad \rho + c\ (\delta t_r - \delta t^s) + I + T + M + \delta_{\text{rel}} + \delta_{\text{tide}} + \epsilon_p \tag{3.20}$$

where $\rho = \sqrt{(x_s - x_r)^2 + (y_s - y_r)^2 + (z_s - z_r)^2}$ is the geometric range. (x_s, y_s, z_s) and (x_r, y_r, z_r) denote the satellite and receiver coordinates in ECEF frame, respectively. The remainder errors denoted by ϵ_p which are considered to be zero-mean Gaussian distributed ($\epsilon_p \sim \mathcal{N}(0, \sigma^2)$).

P	:	pseudo-range measurement [m].
c	:	speed of light [m/s].
δt_r	:	receiver clock bias [s].
δt^s	:	satellite clock bias [s].
I	:	measurement delay due to ionosphere [m].
T	:	measurement delay due to troposphere [m].
M	:	delay caused by multipath [m].
δ_{rel}	:	relativistic effects [m].
δ_{tide}	:	Earth tide and ocean loading effects [m].

The carrier-phase observation is measured as phase difference between the signal received from satellite and the one generated by the receiver itself at the reception time t_r. Carrier phase observables have a similar equation as code pseudo-range Eq. (3.20) and expressed in unit of length.

$$\begin{aligned} \Phi &= \lambda\phi \\ &= \rho + c\,(\delta t_r - \delta t^s) - I + T + m + \delta_{\text{rel}} + \delta_{\text{tide}} + \lambda\, N + \epsilon_\phi \end{aligned} \tag{3.21}$$

Note that the parameters ρ, δt_r, δt^s, T, δ_{rel}, and δ_{tide} are defined in Eq. (3.20). ϵ_ϕ consists of noise and unmodeled effects, and mainly accounts for the remaining biases (e.g., carrier-phase offset, variation and wind-up, receiver noise, etc.). The multipath m refers to the carrier-phase measurement. The ionosphere path delay I has a minus sign compared to the code. That is, the ionosphere delays the code and advance the carrier-phase with the same quantity of time. λ is the wavelength of L_1, L_2 or L_5 in unit of meter. One additional term in Eq. (3.21) is N, which is the carrier-phase integer ambiguity. Determination of N is the key for precise positioning and is assumed to be constant and unknown. In addition, the values of N are different and independent between satellite and receivers, but constant between successive measurement epochs. All terms in the observation Eqs (3.20) and (3.21) must be mitigated. The mitigation process can be carried out by modeling, estimation, or combinations of observables.

The last GPS observation type is the Doppler. The Doppler effect is a phenomenon of frequency shift of the electromagnetic signal caused by the relative motion of the transmitter and the receiver. Let the signal emitted by the satellite have the nominal frequency f. The radial velocity of the satellite related to the receiver are given respectively by Xu [144, p. 41] and Kaplan [54, p. 58] and will not be repeated here.

3.5 Forming GNSS observations differences

In GNSS data analysis it has become very common to form linear combinations and/or to generate difference between observations to reduce or eliminate some of the measurement delays or to reduce the number of unknown parameters. Differences between the observations are usually formed to eliminate parameters, especially the receiver and satellite clocks and also the unknown phase ambiguities. For this purpose one can form single, double, and triple-differences. Differences that can be made are:

- **Between stations**: The difference between two (almost) simultaneous observations by two stations observing the same GNSS satellite. In this difference the satellite specific terms are (almost) perfectly eliminated,

in particular the satellite clock and the relativistic delay. Satellites specific biases are eliminated as well.

- **Between satellites**: The difference between two (almost) simultaneous observations by one station observing two different GNSS satellites. In this difference, the station specific terms are (almost) perfectly eliminated, in particular the station clock (δ_k). But also other station specific biases are eliminated.

- **Between epochs**: The difference between two observations from one station observing one GNSS satellites on two different epochs. In this difference, the initial phase ambiguity is eliminated. But also the constant receiver-transmitter biases are eliminated.

Undifferenced method has an important aspect in that the absolute information of the system described by the observations equations is retained. The estimation process often referred to as filtering can be easily implemented.

Single difference

Under single-difference operation, three types of difference among the observations are obtained. These are; across-receivers, across-satellites and across-time, respectively. Details on the derivation of the observation equations of single-difference, the reader is referred to Leik (2015) [69, pp. 271-273]. If the pair of the receivers are not identical, produced by different manufactures, and not running the same software, the hardware biases under single-difference can become critical. Therefore the double-difference techniques is preferable because the receiver and satellite biases are canceled and the nature of the phase ambiguities have the integer properties. The variance is increased and can create unpleasant situations. Single-difference between two reference receivers a and b observing the same satellite i for the same carrier frequency f reads:

$$P^i_{ab} = \rho^i_{ab} - c\ \delta t_{ab} + I^i_{ab} + T^i_{ab} + v^i_{ab} + \epsilon^i_{ab} \tag{3.22}$$

$$\lambda^i\ \phi^i_{ab} = \rho^i_{ab} - c\ \delta t_{ab} - I^i_{ab} + T^i_{ab} + V^i_{ab} + \lambda^i N^i_{ab} + \zeta^i_{ab} \tag{3.23}$$

where v^i_{ab} and V^i_{ab} present phase center offsets and variations at the single-difference level. λ^i and N^i_{ab} are wavelength and ambiguities at the single-difference level.

Double difference

The reason for the popularity of the double-difference techniques is the elimination of the satellite and the receiver biases, and in addition the

ionosphere and the troposphere are largely reduced in the case of short-baseline. The drawback is that the variance of noise level is increased (Eq. (3.67)) and the complexity will increase as well. The double-difference between two reference receivers a and b observing the same satellites i and j for the same carrier frequency f reads:

$$P_{ab}^{ij} = \rho_{ab}^{ij} + I_{ab}^{ij} + T_{ab}^{ij} + v_{ab}^{ij} + \epsilon_{ab}^{ij} \tag{3.24}$$

$$\lambda^j \phi_{ab}^j - \lambda^i \phi_{ab}^i = \rho_{ab}^{ij} - I_{ab}^{ij} + T_{ab}^{ij} + V_{ab}^{ij} + \lambda^{ij} N_{ab}^{ij} + \zeta_{ab}^{ij} \tag{3.25}$$

where v_{ab}^{ij} and V_{ab}^{ij} present phase center offsets and variations at the double-difference level. λ^{ij} and N_{ab}^{ij} are wavelength and ambiguities at the double-difference level.

Triple difference

The triple-difference is defined as the difference between two double-differences (Eqs. (3.24) and (3.25)) over two epochs t_1 and t_2.

$$P_{ab}^{ij}(t_1,t_2) = P_{ab}^{ij}(t_2) - P_{ab}^{ij}(t_1) \tag{3.26}$$

$$\lambda^{ij} \phi_{ab}^{ij}(t_1,t_2) = \Phi_{ab}^{ij}(t_2) - \Phi_{ab}^{ij}(t_1) \tag{3.27}$$

One important characteristic of triple difference is that the initial ambiguity cancels and can be used to get a good initial value of positions for subsequent double-difference solution. Triple-difference technique can be used as cycle-slip detection algorithm Section 4.1.3.

3.5.1 GPS stochastic model

The functional model (FM) of the GPS observations, given by Eqs. (3.20) and (3.21) defines the mathematical relationship between the measurements and the unknowns, while the stochastic model (SM) describes the statistical property of the observations characterized by the variance-covariance matrices (VCM). VCM provides general information of the observations, including the precision of the observations expressed by the matrix $\boldsymbol{P} = \Sigma^{-1}$, the variance is given by the main diagonal elements while the different physical correlations are given by off-diagonal elements of the VCM. The covariance of the off-diagonal elements of the VCM presents different types of physical correlations between the GPS observables, such as spatial correlation between different channels, the cross-correlations between the observables L_i, for $i = 1, 2,$ and 5, and the temporal correlation between epochs. That is, observational dependency over time, space and between frequencies, respectively. In the estimation process, the precision matrix $\boldsymbol{P}$ is chosen to

be the weight in order to obtain the best linear unbiased estimator (BLUE) in least-square.

Carrier-phase variance models

In order to define a realistic stochastic model for GPS observables, the processing algorithms have to assign weights to observations based on the scintillation distortion of each satellite link, elevation angles, multipath and the signal strength measured by the parameter SNR.

Various variance model for GPS carrier-phase observables are available and vary from simple weighting schemes where the weights are equal to complex ones. Today, the stochastic model includes the following: correlation among observations [128]; satellite elevation dependency modeled by exponential function [29]; temporal and cross-correlations [132]; and multipath detection and monitoring [139]. That is, a suitable robust weighting algorithm that reduces the influence of the satellite exposed by scintillation, multipath, and will enhance the ability to resolve the carrier-phase ambiguity and to improve the stochastic model for GNSS processes. The variance stabilization is monitored without causing the Heywood case.

Carrier-phase variance model complexity

Due to the fact that signal with a low elevation angle experiences a large delay when traversing the atmosphere to reach the user receiver. The signal becomes noisy and suffers from multipath and SNR degradation. The intention is to improve and stabilize the carrier-phase variance without causing any damage. For high precision GPS applications one can risk to introduce unwanted signals, for instance outliers that can bias the estimation process. Anyway, stabilization of the variance is indispensable for precise GPS applications.

3.6 GNSS linear combinations of observations

The key of forming a particular linear combinations (LC) of the basic observables (code and carrier-phase), is that the LC possesses a characteristics that is used for different purposes. For instance the ability to detect the cycle-slip, elimination of ionosphere, ambiguity resolution, noise level determination, and other biases. The drawback is that the noise level is increased under LC compared to L1 or L2. Anyway, for any linear combination formed by using the Eq. (3.28), the following parameters must be computed and investigated for the derived signal, namely the frequency f_i, the wavelength λ_i, carrier-phase ambiguities N_i, noise level σ_i and the ionosphere amplification factor

I_i^1 with respect to Φ_1. This task can be carried out on any difference level (un-differenced, single-differenced or double-differenced) of the observations. Also linear combinations can be made out of a combination of code and phase observations. The linear combinations currently most used in GNSS data processing are:

- Ionosphere-free Linear Combination (IFLC).
- Melbourne-Wubbena linear combination (MWLC).
- Wide-lane linear combinations (WLLC).
- Geometry-free linear combination (GFLC).

Linear combinations is application specific. That is, a suitable combination of code and carrier-phase can offer a new possibility to understand and solve many of GNSS problems, for instance elimination or modeling the ionosphere, cycle slip detection or carrier-phase ambiguity resolution. Taking the phase observations Φ_1, and Φ_2 in the unit of meter, the resulting linear combination (LC) defined as

$$\Phi_i = k_{1,i}\Phi_1 + k_{2,i}\Phi_2 \tag{3.28}$$

where $k_{1,i}$ and $k_{2,i}$ are real-valued coefficients defined in Table 3.3. The corresponding frequency f_i and the wavelength λ_i of the combined signal Φ_i are given respectively by Eqs. (3.29) and (3.30):

$$f_i = k_{1,i}f_1 + k_{2,i}f_2 \tag{3.29}$$

$$\lambda_i = \frac{c}{f_i} = \frac{c}{k_{1,i}f_1 + k_{2,i}f_2} \tag{3.30}$$

Applying the error propagation law and assuming that the observations on L_1, and L_2 are uncorrelated and have the same defined noise level σ_1, then the noise of the new derived signal is given by the expression:

$$\sigma_i = \sigma_1\sqrt{k_{1,i}^2 + k_{2,i}^2} \tag{3.31}$$

3.6.1 Ionosphere-Free linear combination Φ_3

IFLC is the one most widely used in GNSS data processing. The advantage of the new derived signal is that the effect of the ionosphere is removed, at least to the first order. The second and third order effect of the ionosphere is at a few millimeter level and still uncorrected. Nevertheless, in recent years models have been developed to account for the higher order ionosphere effects (Bassiri & Hajj, 1993) [5], (Kedar & Hajj, 2003) [55] and (Hoque & Jakowski, 2007) [47]. For the Ionosphere-free linear combination we use the Eq. (3.28) with the following values for $k_{1,i} = f_1^2/(f_1^2 - f_2^2)$ and $k_{2,i} = f_2^2/(f_1^2 - f_2^2)$. The new derived observation equations for code and carrier-phase respectively are:

$$\begin{aligned} \Phi_{[\mathrm{IF}]} &= \frac{1}{(f_1^2 - f_2^2)}(f_1^2\Phi_1 - f_2^2\Phi_2) \\ P_{[\mathrm{IF}]} &= \frac{1}{(f_1^2 - f_2^2)}(f_1^2 P_1 - f_2^2 P_2) \end{aligned} \tag{3.32}$$

The drawback of the new signal Φ_3, is that the phase ambiguity N_3 is no longer an integer, which make the process of fixing the ambiguity very difficult and almost impossible to fix. In addition, the noise level is much higher. It should be pointed out that this linear combination does not have true wavelength. However, when in ambiguity fixing using the wide-lane to derive the difference between the number of cycles on L_1 and L_2 in a second step the ionosphere-free linear combination is used to resolve the true number of cycles on L_1 and L_2. In this step the ionosphere-free linear combination is often referred to as narrow-lane combination. The term "narrow" in this case comes from the fact that the (artificial) wavelength of this linear combination for GPS is only 109 mm.

3.6.2 Geometry-Free linear combination Φ_4

GFLC is independent of receiver clocks, satellite clocks and geometry (orbits, station coordinates). The linear combination cancels out all the geometry information leaving only, or mainly, the ionosphere effects. Thus this observation is ideally suited, and most commonly, used for the estimation of the state of the ionosphere. It is also well suited for the cleaning of un-differenced data.

$$\Phi_{[GF]} = \Phi_1 - \Phi_2 \tag{3.33}$$

3.6.3 Wide-lane linear combination (WLLC)

WLLC is interesting for ambiguity resolution because just like the Melbourne-Wubbena combination it has a very long wavelength of 860 mm. This is also the reason why the wide-lane and the Melbourne-Wubbena linear combinations cause for some confusion. However, contrary to Melbourne-Wubbena the wide-lane combination uses phase only and it still contains the full geometry information and thus also the clock, troposphere, and ionospheric effects. Thus its use for ambiguity resolution is mainly on short(er) baselines where the ionosphere, and to a lesser extend the geometry, effects cancel out to a certain extent in the differencing. The wide-lane combination of the code observations does not really serve any purpose. For the wide-lane linear combination we use the Eq. (3.28) with the following values for $k_{1,i} = f_1/(f_1 - f_2)$ and $k_{2,i} = f_2/(f_1 - f_2)$. Leading to the following observation equations for carrier-phase:

$$\Phi_{[WL]} = 1/(f_1 - f_2)\{f_1\Phi_1 - f_2\Phi_2\} \tag{3.34}$$

3.6.4 Melbourne-Wubbena linear combination (MWLC)

MWLC is widely used for observation screening and for integer ambiguity resolution. It is a combination of code and phase measurements which results in a combined measurement with a relatively long wavelength (e.g. 86 cm for GPS) and a noise that is lower than that of the individual code measurements. In fact, the combination is such that all information is lost and only an ambiguity and noise remains. This combination eliminates the effect of the geometry, clocks, ionosphere and the troposphere. So if good code measurements are available this linear combination makes screening the data for cycle slips (jumps in the carrier-phase measurements by an integer number of cycles due to the temporal interruption of the GNSS signal) and outliers very easy as it easily allows the detection of small cycle slips and outliers, e.g. at the 1 to 2 cycle level (1 to 2 m level). Because of this functionality it is also perfectly suited to estimate the integer value of the wide-lane ambiguities in the integer ambiguity resolution procedures.

$$\Phi_{[MW]} = \frac{1}{(f_1 - f_2)}(f_1\Phi_1 - f_2\Phi_2) - \frac{1}{(f_1 + f_2)}(f_1P_1 + f_2P_2) \tag{3.35}$$

3.7 Positioning with GPS

Position determination using GPS depends strongly on the method used. For point positioning, known as standalone, one receiver is used by employing

Table 3.3: L1–L2 carrier-phase linear combinations.

LC	$k_{1,i}$	$k_{2,i}$	$\alpha_{1,i}$	$\alpha_{2,i}$	λ_i [m]	σ_i/σ_1	I_i
Φ_1	1	0	1	0	0.190	1	1
Φ_2	0	1	0	1	0.244	1	1.65
Φ_{IF}	2.55	−1.55	77	−60	0.006	2.98	0
Φ_{GF}	1	−1	−60	77	−1	1.41	0.65
Φ_{WL}	4.53	−3.53	1	−1	0.862	5.74	1.28

the code pseudorange from at least four GPS satellites. This techniques is used when the required accuracy is very low. Other positioning methods exists, for instance, the relative positioning (RP), precise point positioning (PPP), real-time kinematic (RTK) and PPP-RTK. These methods will be addressed in coming sections. For each these positioning methods, the functional, the stochastic models and the limiting factors will be provided. Anyway, our aim is to obtain a clean observations equations involving only measurements and position coordinates. Any parameter limiting the accuracy of the positioning method requires monitoring and will ease the data integrity solution.

3.7.1 GPS single point positioning

The single point positioning requires at least four satellites in view and uses a code-pseudorange measurements and navigation message to determine the unknown position within a positioning accuracy at the meter level. This method is known as a code-based single point positioning (SPP) method.

The functional model of SPP is defined by Eq. (3.20), where the troposphere and the ionosphere path delays are determined by models given by Eqs. 3.7 and 3.3.1, respectively.

Let ρ_u^j represent the Euclidean distance between the satellite position in ECEF coordinate system (x^j, y^j, z^j) at emission time t_e and the receiver position (x_u, y_u, z_u) at the reception time t_r, the geometric distance reads:

$$\rho_u^j = \sqrt{(x_u - x^j)^2 + (y_u - y^j)^2 + (z_u - z^J)^2} \tag{3.36}$$

In case of the measurements are linear of the receiver position and clock synchronization error $(x_u, y_u, z_u, c\delta t_u)$, the weighted least square techniques given by Eq. (2.44) is applicable. That is, $\boldsymbol{z} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{\epsilon}$, Unfortunately, the measurements are not linear and the measurement equations can be written by a general non-linear vector function $\boldsymbol{z} = H(\boldsymbol{x}) + \boldsymbol{\epsilon}$. Applying the Taylor

expansion up to the first order around the initial vector $\boldsymbol{x}_0$, we get :

$$H(\boldsymbol{x}) = H(\boldsymbol{x}_0) + \frac{\partial H(\boldsymbol{x})}{\partial \boldsymbol{x}}(\boldsymbol{x} - \boldsymbol{x}_0) + \cdots \tag{3.37}$$

assuming that $\boldsymbol{x}_0$ is closer to the true value and neglecting the second order of the Eq. (3.37), the approximation reads:

$$H(\boldsymbol{x}) \approx H(\boldsymbol{x}_0) + \frac{\partial H(\boldsymbol{x})}{\partial \boldsymbol{x}}(\boldsymbol{x} - \boldsymbol{x}_0) + \boldsymbol{\epsilon} \tag{3.38}$$

We can obtain the linear equation $\boldsymbol{z} - H(\boldsymbol{x}_0) = \boldsymbol{H}(\boldsymbol{x} - \boldsymbol{x}_0) + \boldsymbol{\epsilon}$. Applying the linear weighted least techniques, the normal equation reads:

$$H^T W H(\hat{\boldsymbol{x}} - \boldsymbol{x}_0) = H^T W(\boldsymbol{z} - H(\boldsymbol{x}_0)) \tag{3.39}$$

The estimated parameter vector $\hat{\boldsymbol{x}}$ reads:

$$\hat{\boldsymbol{x}} = \boldsymbol{x}_0 + (H^T W H)^{-1} H^T W(\boldsymbol{z} - H(\boldsymbol{x}_0)) \tag{3.40}$$

The improvement of the final solution can be achieved by iteration:

$$\begin{aligned} \hat{\boldsymbol{x}_0} &= \boldsymbol{x}_0 \\ \hat{\boldsymbol{x}}_{i+1} &= \hat{\boldsymbol{x}}_i + (H^T W H)^{-1} H^T W(\boldsymbol{z} - H(\hat{\boldsymbol{x}}_i)) \end{aligned}$$

The final estimated parameter $\hat{\boldsymbol{x}} = \lim_{i \to \infty} \hat{\boldsymbol{x}}_i$ is obtained under convergence. Anyway, the convergence will not guarantee the correct user position.

The iterated least square method is often referred to as one-step Gauss-Newton method [108, p. 133] and suffers from ill-conditioned effect where a small variations in the input results a large variations in the output. The Levenberg-Marquardt ([145, pp. 261-266], [86, Chap. 15.5]) method is preferable for non-linear least square estimation and guarantee the convergence of the non-linear system.

3.7.2 GPS relative positioning

GNSS relative positioning known as differential positioning employs at least two GNSS receivers simultaneously tracking at least four common satellites. Since two receivers are employed, one is selected as a reference with known coordinates, the other receiver known as the rover or remote where the coordinate is unknown and shall be determined relative to the reference or base site. Relative positioning uses carrier-phase pseudo-range and the differencing technique to compute the rover position with a high accuracy at centimeter level. The method is precise compared to code pseudo-range treated in Section 3.7.1. The interested reader is referred to Leick (2015) [69, Chap. 7], and Wellenhof (2008) [46, Chap. 6].

Functional model of relative positioning

The kinematic relative positioning can be divided into two categories, the post-processing (PPK) and real-time kinematic (RTK).

Luo [74, p. 90] and El-Rabbany [28, Chap. 5] describe different techniques between these two methods. The material will not be repeated here. Our attention is directed to RTK because it is considered as one of the main components in NRTK data processing. Key to obtain a centimeter level accuracy is the rapid and reliable ambiguity resolution. The initial ambiguities are resolved instantaneously using the *"on-the-fly"* ambiguity resolution techniques, Hoffmann-Wellenhof et al. (2008) [46, p. 217]. Better processing techniques that reduces the impact of the nuisance parameters to centimeter level is also a key. In addition, the baseline length should not exceed 20 km. This limitation is imposed by the distance dependent biases such as the signal refraction in the atmosphere (ionosphere and troposphere) and the orbit errors.

The functional model of the relative positions describes the mathematical relationship between the double-difference observations and the unknown parameters. The model shall be given in both cases, the short and the long baselines. In case of the GPS, applying the code division multiple access (CDMA) techniques. The simplified carrier-phase double-difference (DD) observation equations between two receivers a and b and two satellites i and j reads:

$$\Phi_{a,b}^{i,j} = \rho_{a,b}^{i,j} + \lambda_f N_{a,b}^{i,j} + \epsilon_{a,b}^{i,j} \tag{3.41}$$

where

$\rho_{a,b}^{i,j}$: geometrical range between receivers a and b and satellites
: i and j [m].
λ_f : wavelength of the carrier combination [m]
$N_{a,b}^{i,j}$: integer DD phase ambiguity [cycles]
$\epsilon_{a,b}^{i,j}$: random noise of DD [m]

Since the coordinates of the base station are known, the resulting state vector reads:

$$\boldsymbol{x} = (x_b, y_b, z_b, N_{a,b}^{i,j}), \quad \text{for } i = 1, 2, \cdots, n_{ns} \tag{3.42}$$

where n_{ns} is the number of observed satellites pair, (x_b, y_b, z_b) is the rover coordinates and $N_{a,b}^{i,j}$ are ambiguities. Note that the ionosphere and troposphere residuals are ignored for short baselines.

Stochastic model of relative positioning

The stochastic model of the relative positioning is very complex compared to the Precise Point Positioning (PPP) techniques. Due to the fact that the variance of the noise level under the DD increased by a factor of 4, the complexity has increased as well. Due to the fact that the DD observation is composed of four undifferenced observations, one has to study the mathematical and physical correlations. That is, the covariance structure related to the satellite, station, frequency and epoch. The spatial, temporal and cross correlations are always important to investigate for the process under study. That is, fully populated covariance matrix is attractive to investigate. For short baselines, the spatial-correlation of the observations is much stronger than those from long baselines. For PPP, the spatial correlation exists only between one receiver and two satellites while in relative positioning, the correlation exists between observation from one satellite observed in two stations. El Rabbany (1994) [28, Chap. 3] investigated the effects of the correlation on the baseline and its accuracy. The conclusion is that the physical correlation is typically inversely proportional to both observation sampling rate and the baseline length.

Relative positioning limiting factors

Both models (functional and stochastic) of the relative positioning are presented. Now it is time to present the error sources and their effects. These errors impose some limitation that have direct effect on the position solution. The classification of these errors is divided into the spatial separation of the rover and the base station, namely the baseline length and distance independent effects known as the site specific error limitations.

Distance dependent errors

Distance-dependent errors refer to satellite orbit, ionosphere and troposphere errors. Under the double-difference operation, the ionospheric effect is reduced.

Satellite orbit error

The impact of the effect of un-modeled orbit errors on the estimated station coordinates estimates is given by Dash et al. (2015) [24, p. 35], giving the error δx in a component of a baseline of length l as a function of an orbit error of size ΔX.

$$\Delta x[m] \approx \frac{l}{d} \Delta X[m] \tag{3.43}$$

where $d \approx 25.000$ km is the approximate distance between the satellite system and the survey area, and l is the baseline length.

Ionospheric effects

The ionosphere is considered as the major threat to positioning and navigation. In general, one must keep in mind that the ionosphere impacts on GNSS signals depend on the TEC which varies with time and space. That is, a pure spatio-temporal process. Details of high order ionospheric effects on GPS signals can already be found in Bassiri and Hajj (1993) [5]. An excellent review of the subject is provided by Petrie et al. (2011) which includes an extensive list of references. The later reference reviews several approaches to quantify geometry bending errors. Our aim is to compute the TEC value and study the variations in ionosphere model. The TEC variations can be modeled as random-walk [110].

Troposphere effects

Some guidelines to improve the impact of the troposphere path delay on relative positioning are given by Luo [74, p. 95].

- Differencing the observations as in case of relative positioning, the troposphere solution may be biased by a constant offset. External troposphere calibration may be an appropriate solution.
- Strong correlation between the zenith tropospheric delay and the height of the station [70].
- Advanced mapping functions (e.g. GMF, VMF1) [115] to handle the low elevation angle E_l and to improve the stochastic model.

Distance independent errors

Distance independent error refers to the site specific errors, for instance the multipath, and station hardware devices.

Relative positioning improvement

Up to now, we have presented the functional and the stochastic models of the relative positioning as well as the limitations imposed by the method. The next step is to move little further by proposing some amelioration potential to the relative positioning.

Regarding the stochastic model, the weighting variance functions is absolutely a recommendation. Taking into considerations the spatial, temporal

and cross correlation to build a realistic stochastic model. The most variance models used today are weighting schemes that take into consideration the elevation angle E_l, the signal-to-noise ratio (SNR), ARMA process [74], and other methods. Another well established method is the model residuals analysis. Analyzing the time series of the observations residuals by means of an ARIMA process and estimate the parameters in question. Furthermore, mathematical determination of the temporal correlation, statistical verification of results by determination of the distribution, change in the means and variance and physical interpretation will enhance the possibility to carry out the check of the correctness of the information provided by the relative positioning.

3.7.3 GPS precise point positioning

PPP techniques use undifferenced dual-frequency pseudo-range and carrier-phase measurements to determine the receiver's position in a three dimensional space and the receiver clock error. The position accuracy is at cm level. The main obstacle of this method is the initialization process, the long convergence time of the carrier-phase ambiguities. The initialization process requires at least 20 minutes. The PPP is based on the use of the precise GPS orbit and clock data provided by the analysis center for example by the International GNSS Service (IGS) and Jet Propulsion Laboratory (JPL). Depending on the delay of the availability of the data from real-time, different accuracy of the products are provided. For example *Ultra-rapid* products are available in real time with around 10 cm for the orbits and 5 ns for the clocks. The PPP method is economical compared to relative positioning and requires one single receiver and no need for additional correction methods to mitigate errors that are not canceled under differencing. To overcome the limitation imposed by the ambiguities, Wübbena et al [142] introduced a PPP-RTK method that reduces the convergence time up to 50 seconds. In addition, a position accuracy can be improved for a single frequency receiver by correcting for the ionospheric path delay provided by the network.

Various textbooks are available and the reader is referred to Zumberge et al. [147], Witchayangkoon [140], Mohamed [1], and Geng [36] for deeper insight on this subject.

GPS PPP functional model

The basic concept of the PPP method is described. The first order ionosphere path delay is eliminated by means of IFLC Eq. (3.32). Other nuisance parameters limiting the accuracy of the PPP are estimated or modeled.

The functional model of the GPS observables relates the parameters to be estimated to the observations. For static PPP, the observation equations of the IFLC from receiver r to satellite s reads:

$$\Phi^s_{r,3} = \rho^s_r + c\,(\delta t_r - \delta t_s) + T + \lambda_3\, N^s_{r,3} + \epsilon^s_{r,3} \tag{3.44}$$

where

ρ^s_r : geometrical range between the receiver r and satellite s [m]
c : speed of light [m/s]
δt_r : receiver clock bias [s]
δt_s : satellite clock bias [s]
T : measurement delay due to troposphere [m]
λ_3 : wavelength of LC3 measurement [m]
$N^s_{r,3}$: non-integer phase ambiguity of the LC3 measurement [cycles]
$\epsilon^s_{r,3}$: random noise of LC3 measurement [m]

where the range $\rho^s_r = \sqrt{(x_r - x^s)^2 + (y_r - y^s)^2 + (z_r - z^s)^2}$ is the Euclidean distance between the satellite position in ECEF coordinate system (x^s, y^s, z^s) at emission time t_e and the receiver position (x_r, y_r, z_r) at the reception time t_r. From Eq. (3.44), the troposphere path delay can be approximated by applying the Eq. (3.13). The orbits (x^s, y^s, z^s) and clock (δt_s) products from IGS, then the simplified version of the Eq. (3.44) reads:

$$\Phi^s_{r,3} = \rho^s_r + c\dot{\delta} t_r + T_w \cdot m_{T_w} + \lambda_3\, N^s_{r,3} + \epsilon^s_{r,3}, \quad s = 1, \cdots, n_s \tag{3.45}$$

Resulting the state vector $\boldsymbol{x} = (x_r, y_r, z_r, c\ \delta t_r, T_w, N^s_{r,3})^T$, where T_w is the troposphere zenith wet delay, and n_s is the number of observed satellites.

Taking the expectation of the Eq. (3.45) and noting that $\mathbf{E}(\epsilon^s_{r,3}) = 0$, then the functional model of PPP reads:

$$\mathbf{E}(\Phi^s_{r,3}) = \rho^s_r + c\dot{\delta} t_r + T_w + \lambda_3 \dot{N}^s_{r,3} = F(\boldsymbol{x}) \tag{3.46}$$

The function $F(\boldsymbol{x})$ is non-linear, applying the Taylor series expansion Eq. (3.40) around the initial value $\boldsymbol{x}_0$. The approximated observation linear model in matrix form reads:

$$\Delta\Phi_3 = \boldsymbol{H}\, \Delta\hat{\boldsymbol{x}} + \boldsymbol{\epsilon} \tag{3.47}$$

where $\Delta\hat{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{x}_0$ is the estimated reduced parameter, $\boldsymbol{\epsilon}$ is the model uncertainty (residuals), $\Delta\Phi_3 = \Phi_3 - F(\boldsymbol{x}_0)$ is the vector of reduced observations, and finally $\boldsymbol{H}$ is the design matrix defined by Eq. (3.48)

$$\boldsymbol{H} = \left[\frac{\partial F(\boldsymbol{x})}{\partial x_r}, \frac{\partial F(\boldsymbol{x})}{\partial y_r}, \frac{\partial F(\boldsymbol{x})}{\partial z_r}, \frac{\partial F(\boldsymbol{x})}{\partial \delta t_r}, \frac{\partial F(\boldsymbol{x})}{\partial T_w}, \frac{\partial F(\boldsymbol{x})}{\partial N^s_{r,3}}\right]^{x=x_0}_{s=1,\cdots,n_s} \tag{3.48}$$

$$= \left[\frac{\boldsymbol{x}_r - \boldsymbol{x}^s}{\rho^s_r}, \quad \frac{\boldsymbol{y}_r - \boldsymbol{y}^s}{\rho^s_r}, \quad \frac{\boldsymbol{z}_r - \boldsymbol{z}^s}{\rho^s_r}, \quad c, \quad m_{T_w}, \quad 0\right]^{x=x_0}_{s=1,\cdots,n_s}$$

where m_{T_w} is the zenith wet delay mapping function and noting that $\boldsymbol{P} = \boldsymbol{W}$ is the precision/weight matrix and is defined as the inverse of the cofactor of $\Sigma_{\Delta\Phi_3}$. The solution is given by the expression:

$$\hat{\boldsymbol{x}} = \boldsymbol{x}_0 + (\boldsymbol{H}^T\boldsymbol{P}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{P}\,(\Delta\Phi_3) \tag{3.49}$$

Statistical investigation of the PPP solution

In order to construct the test statistics, our starting point is to compute the residuals, expectation and variance of the solution vector $\hat{\boldsymbol{x}}$.

The residual vector $\boldsymbol{\epsilon}$ is computed from the GMM Eq. (3.47).

$$\boldsymbol{\epsilon} = \boldsymbol{H}\cdot\boldsymbol{\Delta}\hat{x} - \Delta\Phi_3 \tag{3.50}$$

Applying the variance-covariance propagation law to Eq. (3.49), we get:

$$\mathrm{cov}(\hat{\boldsymbol{x}}) = \sigma_0^2\cdot(\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H})^{-1} \tag{3.51}$$

where the variance σ_0^2 is estimated from the weighted sum of squared residuals and is given by the expression

$$\hat{\sigma}_0^2 = \frac{(\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H})^{-1}}{n-m} \tag{3.52}$$

where n is the number of observations and m is the number of unknown. Clearly from the definition of F-test, the $\hat{\sigma}_0^2$ follows the F-distribution with $(n-m)$ degree of freedom.

GPS PPP nuisance parameters

Unlike the differencing methods where some nuisance parameters are canceled such at the receiver and the satellite biases, the PPP method introduces some unusual complications that requires additional correction terms and models to be considered. For instance, to account for satellite-specific, atmosphere and site-specific errors. These errors must be considered, otherwise the degradation of the accuracy of the PPP is obtained. Among these corrections are the satellite and receiver antenna phase center offsets, carrier-phase wind-up effect, and Earth deformation effects (solid tides, ocean loading, and pole tide). Excellent treatment of these issues can be found in [120, pp. 129-138] and will not be repeated here.

GPS PPP stochastic model

The stochastic model describes the statistical properties of observations precision and correlation and is expressed by the variance-covariance matrix (VCM). The VCM of the PPP solution, Eq. (3.51) contains valuable information. The main diagonal elements present the variance of the estimated state vector, $\hat{x}$ while the off-diagonal elements present the correlation between parameters.

In order to construct a more realistic VCM of estimate, a more appropriate observations weighting algorithm, the variance model and taking into consideration the physical correlations are the right remedy to improve the stochastic model.

3.7.4 PPP-RTK functional model

PPP-RTK can be regarded as the PPP augmented high-precision positioning service. The network is used to provide the user with satellite clocks information as well with the interpolated ionospheric path delays. The precise IGS orbits are used to determine the satellite positions.

Reducing the GNSS errors to minimum will enhance the possibility to resolve the ambiguities at the user side within 50 seconds.

A few techniques have been proposed for PPP-RTK concept. The interested reader is referred to Teunissen et al. [125] and Teunissen and Khodabandeh [129].

3.8 GPS multipath mitigation algorithm

The most common methods for reducing multipath are improved antenna design (e.g., choke ring ground planes) and careful site selection. Unfortunately, it is often not possible to change accuracy or processing time. Therefore, the method of reducing multipath that will be most transparent to the user is to remove it at the signal level within the GPS receiver itself. In order to mitigate the multipath, we propose a concept based on iterative weighted least square method. Our inspiration is taken from PPP method by following the same procedure. Start by the observations equations, estimate/eliminate the nuisance parameters, and finally isolate the parameter of interest, namely the multipath.

Derivation of multipath equation

A Kalman filter is used to monitor the multipath at observation domain. The parameters that will be investigated are the signal to noise ratio on L_1 and L_2,

and double difference phase residuals. Starting with basic GPS observables (P_1, P_2, Φ_1 and Φ_2), equations of the code and carrier-phase observations from satellite k to receiver i in unit of length reads:

$$P_i^k = \rho_i^k + c(\delta t_i - \delta t^k) + c(b_i + b^k) + T_i^k + I_i^k + M_i^k + r_i^k + V_i^k + \epsilon_i^k \quad (3.53)$$

$$\Phi_i^k = \rho_i^k + c(\delta t_i - \delta t^k) + c(a_i + a^k) + T_i^k - I_i^k + m_i^k + r_i^k + v_i^k + \lambda N_i^k + \varepsilon_i^k \quad (3.54)$$

Assuming that hardware biases b_i, b^k, a_i and a^k are estimated and removed and the term ρ_i^k also includes relativistic effects r_i^k, phase center offsets and variations (V_i^k and v_i^k) and orbits errors O_i^k. In addition, assume that no cycle slip is present and ambiguities λN_i^k are resolved.

Let $R_i^k = \rho_i^k + c(\delta t_i - \delta t^k)$ and omitting the indices i and k. The simplified observations model reads:

$$P_1 = R + T + I_1 + M_1 + \epsilon_1 \quad (3.55)$$

$$P_2 = R + T + I_2 + M_2 + \epsilon_2 \quad (3.56)$$

$$\Phi_1 = R + T - I_1 + m_1 + \varepsilon_1 \quad (3.57)$$

$$\Phi_2 = R + T - I_2 + m_2 + \varepsilon_2 \quad (3.58)$$

The ionospheric delay on L_1 and L_2 are related by the expression:

$$I_2 = \gamma\, I_1 \quad \text{where } \gamma = \left(\frac{f_1}{f_2}\right)^2 \quad (3.59)$$

Performing ionosphere free linear combination on pseudo-range and carrier-phase we obtain

$$P_4 = R + T + \frac{\gamma}{\gamma - 1}(M_1 + \epsilon_1) - \frac{1}{\gamma - 1}(M_2 + \epsilon_2) \quad (3.60)$$

$$\Phi_4 = R + T + N_4 + \frac{\gamma}{\gamma - 1}(m_1 + \varepsilon_1) - \frac{1}{\gamma - 1}(m_2 + \varepsilon_2) \quad (3.61)$$

N_4 is the phase ambiguity of the iono-free linear combination of carrier-phase. We have ignored the multipath (m_i) and noise (ε) delays in carrier-phase compared to the code pseudo-range. Taking the difference between the two Eqs. ((3.60) and (3.61)) we get the approximation:

$$\Phi_4 - P_4 \approx N_4 + \frac{\gamma}{\gamma - 1}(M_1 + \epsilon_1) - \frac{1}{\gamma - 1}(M_2 + \epsilon_2) \quad (3.62)$$

Smoothed measurements

Changes in geometric range ρ to the satellite are independent of the technique used code or carrier-phase. Note that the carrier-phase measurement has better noise properties than the code. Therefore smoothing the pseudo-range by carrier-phase reduces the noise level and improve the overall accuracy. Before smoothing the code pseudo-range with carrier-phase, the measurements shall be cleaned by removing the ionosphere path delay and the hardware biases. The smoothing process is carried out according to the formula:

$$P_{\text{sm}} = \Phi_t - \frac{\left\{\sum_{i=1}^{n-1} \Phi_{t-i} - P_{t-i}\right\}}{n} \tag{3.63}$$

where P_t and Φ_t are respectively the code and carrier-phase pseudo-ranges at epoch t and are given by Eqs. (3.53) and (3.54). n is the sliding window size and is used as the smoothing buffer.

For practical considerations, the smoothing process is initialized in case of carrier-phase discontinuity caused by a cycle-slip. The smoothing parameter n is a configurable and can take values no less than 200 seconds and no more than 1200 seconds. The multipath will not be removed during the initialization step (less that 200 seconds). Combining the Eqs. (3.63) and (3.62) and arranging, we get:

$$P_{\text{sm}} = \Phi_t - N_4 + \frac{\gamma}{\gamma-1} \underbrace{\frac{\left\{\sum_{i=1}^{n-1}(M_1+\epsilon_1)\right\}}{n}}_{(M_1+\epsilon_1)\textbf{ mean value}} - \frac{1}{\gamma-1} \underbrace{\frac{\left\{\sum_{i=1}^{n-1}(M_2+\epsilon_2)\right\}}{n}}_{(M_2+\epsilon_2)\textbf{ mean value}} \tag{3.64}$$

From Eq. (3.64) we observe that the we are actually computing the mean value of the noise and the multipath on $L_1$and L_2 frequencies. Typically, the multipath delays are zero mean on time periods sufficiently large ($N/2$), and at least can be chosen to correspond to the correlation time of the multipath.

Multipath monitoring algorithm

Monitoring the multipath behavior in real time for each observed satellite from a reference receiver is a challenge. The aim is to detect the satellite influenced by multipath and exclude it from the computation or assign a weight to reduce the influence on position estimation. That is, avoid biased solution caused by multipath. For this purpose, the Kalman filter is chosen due to it's recursive characteristic. The user does not need to save previous observations; instead, all filter observations are carried forward in the filter.

Kalman filter includes measurement and system models as well as their corresponding stochastic models as illustrated in equations. The noises w_k and v_k of the models are assumed to be independent from each other and distributed according to the normal distribution. Following the notation proposed by Brown and Hwang [13], the system and the observations equations are given by Eqs. (2.55) and (2.55). The multipath Eq. (3.62) relating the pseudo-range and carrier-phases is supposed to be constant in time and the variation is dominated by the P code multipath and the measurement accuracy.

Multipath state vector definition x

The state vector is defined as $\boldsymbol{x} = [M^1 \cdots M^x]$, where M^k refers to the multipath of satellite k observed at the reference receiver y, omitted in this case. From the Eq. (2.55), the state vector $\boldsymbol{x}_k$ depends only on the previous solution $\boldsymbol{x}_{k-1}$. This is an auto-regressive process of order 1. Since the noise $\boldsymbol{w}_k$ is normal distributed, the process is referred to as the first-order Gauss-Markov process.

Multipath transition matrix definition Φ

From the Eq. (3.62), theoretically, the multipath variation is considered to be constants in time. Since the process under study is a first-order Gauss-Markov process and its auto-covariance function is an exponential. Therefore the choice of $\boldsymbol{\Phi} = \exp\{-T/\tau\}$ for multipath is an appropriate one and is decreasing exponentially. τ and T denote respectively the correlation time and the time difference between epochs. In many applications, the exponential correlation function is often used, because it fits a large number of physical processes with remarkable accuracy. For interested reader is referred to Huang and Brown [13, p. 95], Gelb [35, p. 81] and Leick [68, p. 169].

$$\boldsymbol{\Phi} = \begin{bmatrix} \exp\{-T/\tau\} & 0 & \cdots & 0 \\ 0 & \exp\{-T/\tau\} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \exp\{-T/\tau\} \end{bmatrix} \tag{3.65}$$

or

$$\boldsymbol{\Phi} = \left[\exp\{-T/\tau\}\right] \tag{3.66}$$

If $\frac{T}{\tau}$ is constant, then the multipath is modeled as a constant.

Multipath process noise definition Q

The variance of the process noise for correlation time τ is given by Gelb [35, p. 82].

$$\boldsymbol{Q}_{w_k} = E\{\boldsymbol{w}\boldsymbol{w}^T\} = \frac{\tau}{2}(1 - e^{-\frac{2T}{\tau}})\, q_k \tag{3.67}$$

where q_k is the variance of the process noise. The initial values of the quantities τ and q_k can be determined empirically by computing the sample mean and the auto-covariance functions.

When τ approaches 0, the process noise is a white noise and no correlation exist between epochs, when τ approaching infinity, the process is a pure random walk.

Updating the state vector

The updating step at epoch $k+1$ is carried out by multiplying previous state vector with transition matrix $\boldsymbol{\Phi}$.

$$\{M_i^j\}_{k+1} = \boldsymbol{\Phi}\ \{M_i^j\}_k \tag{3.68}$$

The covariance matrix $\boldsymbol{P}_{k+1} = [\mathrm{var}(M_i^j)] = \boldsymbol{\Phi}\boldsymbol{P}_k\boldsymbol{\Phi}^T + \boldsymbol{Q}$, where the variance is given by the expression:

$$\mathrm{var}\{M_i^j\}_{k+1} = \mathrm{var}\{M_i^j\}_k \quad e^{\frac{-2T}{\tau}} + (1 - e^{-\frac{2T}{\tau}})\, q_k \tag{3.69}$$

Innovation computation

In order to control the variation of the process under study, the standardized innovations must be computed and tested for variance upper bound. From Eq. (3.62), we see that the computation of innovations requires that the ambiguities are known. Therefore, the right thing to do is to resolve ambiguities with multipath.

3.9 TEC observables from code and carrier-phase

Two main ionospheric parameters needed to be monitored are the TEC and scintillation indices (S_4 and σ_ϕ). S_4 refers to the amplitude scintillation while σ_ϕ refers to the carrier-phase scintillation.

We start by forming the basic TEC observable and the algorithm will be applied only to GPS satellite. In addition, we assume that valid observations are present for a given epoch, this means that all four types of measurements (L_1, L_2, P_1 and P_2) are present with valid signal-to-noise ratio (SNR) values and loss-of-lock (LLI) indicators.

GPS signals transmitted by satellites are influenced by the ionosphere and depend strongly on frequency f and on the slant TEC between the GPS satellite and the receiver. The group delay is given in unit of distance or unit of second by the following expressions:

$$I_{\phi,g} = \{40.3/(f^2)\}\text{TEC} \quad [m] \tag{3.70}$$

$$\delta t = \{40.3/(cf^2)\}\text{TEC} \quad [s] \tag{3.71}$$

The sensitivity of the ionospheric range delay to slant TEC for the L1 GPS signal is 0.162 m per TECU and 0.267 m for L2.

With a dual-frequency GPS receiver we can measure the difference in ionospheric delays between the L_1((1575.42MHz) and L_2(1227.42MHz) of the GPS frequencies up to first order, with the assumption that the satellite signal travels along the same path through the ionosphere.

We start by computing the basic slant TEC observables (see Eqs. ((3.20) and (3.21)) and their corresponding variances. We assume the P_1, P_2, Φ_1 and Φ_2 are uncorrelated.

Applying the geometry linear combination Section 3.6.2 to the code and carrier-phase pseudo-ranges to derive the TEC value.

$$\begin{aligned} P_1 - P_2 &= 40.3\ \text{TEC}\{(1/f_2^2) - (1/f_1^2)\} \\ &= 40.3\ \text{TEC}\{(f_1^2 - f_2^2)/f_1^2 f_2^2\} \end{aligned} \tag{3.72}$$

TEC computation from code pseudo-range is straightforward:

$$\begin{aligned} P_{\text{TEC}} &= \frac{1}{40.3}\left\{\frac{f_1^2 f_2^2}{f_1^2 - f_2^2}\right\}(P_1 - P_2) \\ &= k\ (P_1 - P_2) \end{aligned} \tag{3.73}$$

The variance computation of ρ_{TEC} reads:

$$\begin{aligned} \text{var}(P_{\text{TEC}}) &= k^2\Big\{\text{var}(P_1) + \text{var}(P_2) - 2\ \text{cov}(P_1, P_2)\Big\} \\ &= 2\ k^2\ \text{var}(P_2) \\ &= 2\ k^2\ \sigma^2_{P,\text{gps}} \end{aligned} \tag{3.74}$$

For carrier phase observations, we obtain,

$$\phi_{\text{TEC}} = -k\ (\Phi_1 - \Phi_2) \tag{3.75}$$

The corresponding variance reads:

$$\begin{aligned} \text{var}(\phi_{\text{TEC}}) &= k^2\Big\{\text{var}(\Phi_1) + \text{var}(\Phi_2) - 2\ \text{cov}(\Phi_1, \Phi_2)\Big\} \\ &= 2\ k^2\ \text{var}(\Phi_2) \\ &= 2\ k^2\ \sigma^2_{\phi,\text{gps}} \end{aligned} \tag{3.76}$$

Assuming that $\mathrm{var}(\Phi_1) = \mathrm{var}(\Phi_2) = \sigma^2_{\phi,\mathrm{gps}}$ and $k = 9.519643E16$. The constant 9.519643 is the conversion factor from unit of distance (meter) to TECU.

4 NRTK corrections generation

"Every time I think what's going on, suddenly there's another layer of complications. I just want to get this damn thing solved"

— John Scalzi

The aim of this chapter is to demonstrate how the NRTK corrections are generated in general. I will start by presenting the methods used to resolve the carrier-phase ambiguity (determination of the initial number of cycles at the first observation epoch), the network covariance structure determination of the correlated errors and finally how the corrections are generated.

Section 4.1 presents how double-difference data processing is carried out, including how to generate a maximum of differenced observables, processing modes, outliers and cycle-slip detection and repair, and computation difficulties. Section 4.2 handles the ambiguity resolution in details. Section 4.3 gives a detailed procedure on how to generate the NRTK corrections. Finally, Section 4.4 handles the NRTK covariance structure definitions.

4.1 Processing double-differenced observables

Our aim is to generate a maximum set of independent double-differenced observables from raw carrier-phase measurement for each satellite observed at each configured site of the region of interest. Two common techniques are used for this purpose. The first method determines all common visible satellites in the network (sub-network), then a *base station* and a *base satellite* are selected. This method refers to *reference double-differencing*. Assume five satellites are viewed from two sites A and B, we have:

$$\left\{\Delta\nabla\Phi_{AB}^{ij} = \Delta\nabla\Phi_{1B}^{1j} : j = \{2,3,4,5\}\right\} \tag{4.1}$$

where the site 1 and satellite 1 are chosen as reference site and satellite, respectively (A=1, i=1).

The process gets more complex when the sub-network becomes larger and the selection of common viewed satellites will be violated. The second method is sequential double-differencing, instead of selecting a base satellite, the satellites are sorted in ascending order and the double differenced observable is formed between pairs of satellites.

$$\left\{\Delta\nabla\Phi_{AB}^{ij} : (i,j) = \{(1,2),(2,3),(3,4),(4,5)\}\right\} \tag{4.2}$$

We are facing an optimization problem. That is, many solutions exist but we are interested only in the one that can provide a more precise and reliable solution.

Saalfield algorithm [99] uses a boolean matrix of zeros and ones to convert 4-tuples carrier-phase observables $\{\phi_a^x, \phi_a^y, \phi_b^x, \phi_b^y\}$ into a single double-difference $\{\phi_a^x - \phi_a^y - \phi_b^x + \phi_b^y\}$ for all independent sets $\{\phi_v^w\}$, where $v = 1, 2, \cdots, n_r$, $w = 1, 2, \cdots, n_s$, n_s is the number of observed satellites and n_r number of reference receivers.

This task will generate the connection matrix $\boldsymbol{R}$ that transforms the vector ϕ_a^x into $\Delta\nabla\boldsymbol{\phi}_{ab}^{xy}$.

$$\Delta\nabla\boldsymbol{\Phi} = \boldsymbol{R}\,\boldsymbol{\Phi} \tag{4.3}$$

Three types of receiver-satellite connection matrix $\boldsymbol{R}$ can be computed, conventional (common satellite), optimal and total double difference observables.

4.1.1 Double difference processing mode

Data collected by network GNSS receivers can be processed in two ways: taking into account the between baseline correlations, this approach is referred to as multi-baseline processing (MBP). GNSS software post-processing package, GAMIT uses this approach. Experts in the field of GNSS data processing suggest that this approach is the correct one to chose. This method is one step approach. Another approach is neglecting the correlation between baselines and is referred to as baseline-by-baseline processing (BBP) and is a two step approach. This approach is the most used by the most commercial GNSS software, for instance, post-processing package Bernese. Different path selection to process data will generate different solutions. Usually the shortest baselines are selected in order to reduce the impact of distance-dependent biases. Assume our sub-network is composed of four sites and is given by $N = \{A, B, C, D\}$, then the sets $\left\{\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}\right\}$ and $\{\overrightarrow{AB}, \overrightarrow{BC}, \overrightarrow{CD}\}$ will generate different solutions in BBP.

Han and Rizos [41] algorithm can be used to generate a more reliable set of independent double differenced observable by processing all possible baselines in BBP mode. The computation of the differencing operator matrix $\boldsymbol{R}$ is easily carried out by applying the Kronecker tensor product to sites and satellites order processing matrices. Three different configurations are given to demonstrate the computation of $\boldsymbol{R}$. Suppose that the un-differenced carrier-phase vector $\boldsymbol{\Phi}$ is given by the expression:

$$\boldsymbol{\Phi} = \left\{\Phi_A^1, \Phi_A^2, \Phi_A^3, \Phi_A^4, \Phi_A^5, \Phi_B^1, \Phi_B^2, \Phi_B^3, \Phi_B^4, \Phi_B^5, \Phi_C^1, \Phi_C^2, \Phi_C^3, \Phi_C^4, \Phi_C^5\right\} \tag{4.4}$$

Considering the alternative when the sites are processed in order to generate the following baselines $\left\{\overrightarrow{AB}, \overrightarrow{AC}\right\}$ and satellites in order $\Big\{(1,2),(1,3),(1,4),(1,5)\Big\}$. The corresponding information matrices are given by $\boldsymbol{V}$ and $\boldsymbol{W}$. The site A and satellite 1 are chosen as references.

$$\mathbf{V} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \qquad \mathbf{W} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix} \tag{4.5}$$

The differencing operator matrix $\boldsymbol{R}$ is given by the Kronecker tensor product $\mathbf{R} = \mathbf{V} \otimes \mathbf{W}$.

$$\mathbf{R} = \left[\begin{array}{ccccc|ccccc|ccccc} 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \end{array}\right] \tag{4.6}$$

The corresponding covariance matrix is obtained by applying the law of propagation of variance we have

$$\begin{aligned} \mathbf{cov}(\Delta\nabla\Phi) &= \mathbf{cov}(\boldsymbol{R}\,\boldsymbol{\Phi}) \\ &= \boldsymbol{R}\,\mathbf{cov}(\Phi)\,\boldsymbol{R}^T \\ &= \sigma^2\,\boldsymbol{R}\boldsymbol{R}^{\boldsymbol{T}} \end{aligned} \tag{4.7}$$

The errors ε_i of un-differenced carrier phases Φ_a^x assumed to be independent, identical standard Gaussian distributed with a common variance $\sigma^2 = 1$, that is $\varepsilon_i \sim \mathcal{N}(0,1)$ and the variance-covariance matrix reads:

$$\mathbf{C}_{\Delta\nabla\Phi} = \boldsymbol{R}\,\boldsymbol{R}^{\boldsymbol{T}} = \left[\begin{array}{cccc|cccc} 4 & 2 & 2 & 2 & 2 & 1 & 1 & 1 \\ 2 & 4 & 2 & 2 & 1 & 2 & 1 & 1 \\ 2 & 4 & 2 & 1 & 1 & 2 & 1 & 1 \\ 2 & 2 & 2 & 4 & 1 & 1 & 1 & 2 \\ \hline 2 & 1 & 1 & 1 & 4 & 2 & 2 & 2 \\ 1 & 2 & 1 & 1 & 2 & 4 & 2 & 2 \\ 1 & 1 & 2 & 1 & 2 & 2 & 4 & 2 \\ 1 & 1 & 1 & 2 & 2 & 2 & 2 & 4 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \hline \mathbf{C}_{21} & \mathbf{C}_{22} \end{array}\right] \tag{4.8}$$

Keeping the baselines unchanged: $\left\{\overrightarrow{AB}, \overrightarrow{AC}\right\}$ and consider the satellites processing order:$\left\{(1,2),(2,3),(3,4),(4,5)\right\}$, then the baselines and satellites information matrices $\boldsymbol{V}$ and $\boldsymbol{W}$ are given by the following expressions:

$$\mathbf{V} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \qquad \mathbf{W} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \tag{4.9}$$

The corresponding double difference transformation matrix $\boldsymbol{R}$ reads:

$$\mathbf{R} = \left[\begin{array}{ccccc|ccccc|ccccc} 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{array}\right] \tag{4.10}$$

The corresponding covariance matrix $\boldsymbol{RR^T}$ is given by:

$$\mathbf{C}_{\Delta\nabla\boldsymbol{\Phi}} = \boldsymbol{R}\,\boldsymbol{R}^T \left[\begin{array}{cccc|cccc} 4 & -2 & 0 & 0 & 2 & -1 & 0 & 0 \\ -2 & 4 & -2 & 0 & -1 & 2 & -1 & 0 \\ 0 & -2 & 4 & -2 & 0 & -1 & 2 & 1 \\ 0 & 0 & -2 & 4 & 0 & 0 & -1 & 2 \\ \hline 2 & -1 & 0 & 0 & 4 & -2 & 0 & 0 \\ -1 & 2 & -1 & 0 & -2 & 4 & -2 & 0 \\ 0 & -1 & 2 & -1 & 0 & -2 & 4 & -2 \\ 0 & 0 & -1 & 2 & 0 & 0 & 2 & 4 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \hline \mathbf{C}_{21} & \mathbf{C}_{22} \end{array}\right] \tag{4.11}$$

Consider the last option by processing the baselines in order $\left\{\overrightarrow{AB}, \overrightarrow{BC}\right\}$ and satellites in order $\left\{(1,2),(1,3),(1,4),(1,5)\right\}$, the information matrices $\boldsymbol{V}$ and $\boldsymbol{W}$ is given the expressions:

$$\mathbf{V} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \qquad \mathbf{W} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \tag{4.12}$$

The differencing operator matrix $\boldsymbol{R}$ is given by the Kronecker tensor product $\mathbf{R} = \mathbf{V} \otimes \mathbf{W}$ as before.

$$\mathbf{R} = \left[\begin{array}{ccccc|ccccc|ccccc} 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 1 \end{array}\right] \tag{4.13}$$

The corresponding covariance for our configuration is given by

$$\mathbf{C}_{\Delta\nabla\mathbf{\Phi}} = \left[\begin{array}{cccc|cccc} 4 & -2 & 0 & 0 & -2 & 1 & 0 & 0 \\ -2 & 4 & -2 & 0 & 1 & -2 & 1 & 0 \\ 0 & 2 & 4 & -2 & 0 & 1 & -2 & 1 \\ 0 & 0 & -2 & 4 & 0 & 0 & 1 & -2 \\ \hline -2 & 1 & 0 & 0 & 4 & -2 & 0 & 0 \\ 1 & 2 & 1 & 0 & -2 & 4 & -2 & 0 \\ 0 & 1 & -2 & 1 & 0 & -2 & 4 & -2 \\ 0 & 0 & 1 & -2 & 0 & 0 & -2 & 4 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \hline \mathbf{C}_{21} & \mathbf{C}_{22} \end{array}\right] \tag{4.14}$$

The variance-covariance of sub-matrices $\boldsymbol{C}_{11}, \boldsymbol{C}_{22}, \boldsymbol{C}_{12}, \boldsymbol{C}_{21}$ generated in three different configurations correspond to the processed baselines. We see that the off-diagonal matrices are different from 0, this explains that the correlation exists between baselines.

4.1.2 Computation difficulties

Denote by $\boldsymbol{C}_{\phi} = \sigma^2$ the covariance function of raw carrier-phase observable and applying the variance-covariance propagation law to the double-difference carrier-phase observable, we obtain $\boldsymbol{C}(\Delta\nabla\mathbf{\Phi}) = \sigma^2 \boldsymbol{R}\boldsymbol{R}^{\boldsymbol{T}}$. Computation of the precision matrix $\boldsymbol{P} = (\boldsymbol{R}\boldsymbol{R}^{\boldsymbol{T}})^{-1}$ can be a challenging task if the size of $\boldsymbol{R}$ becomes larger.

Realistic approach of generating a maximum set of independent double differenced observables is to allow a different reference satellite to be selected for different baseline and allowing the correlation between the baselines. This approach is referred to as *base satellite per baseline*. The

precision matrix with correlation $\boldsymbol{M}$ between baselines reads:

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{M}_{S_1} & \boldsymbol{M}_{\overrightarrow{AB}} & \boldsymbol{M}_{\overrightarrow{AC}} \\ \boldsymbol{M}_{\overrightarrow{BA}} & \boldsymbol{M}_{S_2} & \boldsymbol{M}_{\overrightarrow{BC}} \\ \boldsymbol{M}_{\overrightarrow{CA}} & \boldsymbol{M}_{\overrightarrow{CB}} & \boldsymbol{M}_{S_3} \end{bmatrix} \tag{4.15}$$

where $\boldsymbol{M}_{S_1}$, $\boldsymbol{M}_{S_2}$, and $\boldsymbol{M}_{S_3}$ correspond to the variance of the configured stations while off-diagonal elements of $\boldsymbol{M}$ correspond to the correlation between baselines. Now, the new challenge that we are facing is that the block matrices are not identical and the use of Kronecker matrix product to reduce the computation time is violated. Introduction of *missing observations* allow us to construct the identical block matrices and perform the inversion of block matrix. Another option is to invert the full matrix $(\boldsymbol{R}\boldsymbol{R}^{\boldsymbol{T}})^{-1}$. As a result, different patterns of $\boldsymbol{R}$ are obtained based on strategy and processing mode utilized. The next step will focus on "resolving ambiguity in double-difference level" and is the subject of next section.

4.1.3 GPS data screening

GPS data pre-processing refers to the process of data cleaning, and includes the detection and removable of the outliers, carrier-phase discontinuities and dealing with missing observations. The data pre-processing is an important issue in terms of NRTK data integrity.

A. Outliers detection and removable

In order to produce high quality network corrections, the double-difference phase observations are formed and historical data are used to compute statistical values for instance the mean, the median, the variance and the root-mean square (RMS) errors. The typical average error level is computed from historical data and the observations that lie outside the acceptance region are ignored. The historical data is determined by a sliding window. The size of the window is a user defined parameter and corresponds to the temporal correlation of the observation. The procedure applied to handle low quality is given by the expression:

$$\left|\Delta\nabla\Phi_i - \mathbb{E}(\Delta\nabla\boldsymbol{\Phi})\right| \geq k\ \sqrt{\mathrm{var}(\Delta\nabla\boldsymbol{\Phi})} \tag{4.16}$$

where the $\mathbb{E}(.)$ is the mathematical expectation operator. Typical value of $k \in [3, 6]$ must be determined from the data. Lower values of k risk to reject good observations. Trade-off between a good geometry preservation and satellites rejection is important to maintain. $\Delta\nabla\boldsymbol{\Phi}_i$ is the current observation between two reference receivers and two satellites, $\mathbb{E}(\Delta\nabla\Phi)$ is the expectation of historical data (empirical mean).

Since the median is a robust estimator than the mean value, we can replace the mean by the median value in Eq. (4.16). If we take a closer look at Eq. (4.16), we see that the similarity exists to Chebyshev inequality (4.17).

Chebyshev inequality (4.17) defines an upper bound for the probability that $\Delta\nabla\mathbf{\Phi}_i$ lies outside a k-neighborhood of $\mathbb{E}(\Delta\nabla\mathbf{\Phi})$. This equation is often used to prove the consistency of an estimator Def. (10).

$$\mathbb{P}\{|\Delta\nabla\mathbf{\Phi}_i - \mathbb{E}(\Delta\nabla\mathbf{\Phi})| < k\} \geq 1 - \frac{\sigma}{k^2} \tag{4.17}$$

or equivalently

$$\mathbb{P}\{|\Delta\nabla\mathbf{\Phi}_i - \mathbb{E}(\Delta\nabla\mathbf{\Phi})| \geq k\} < \frac{\sigma}{k^2} \tag{4.18}$$

B. Cycle slip detection and repair algorithm

The cycle-slip is caused by a blockage of the satellite signal due to the man-made objects (buildings), trees, mountains, high ionospheric activities, multipath, low satellite elevation angle (E_l), signal processing tracking loops and other unknown phenomena. Anyway, the jumps must be detected and repaired in order to carry out precise positioning. To accomplish the goal, we have to construct a time series from the carrier-phase measurements, by forming a geometry-free linear combination. To ease the notation, let $\kappa = \dfrac{f_2^2 - f_1^2}{f_1^2\, f_2^2}$, then we have:

$$\begin{aligned} \Phi^k_{i,\mathrm{GF}} &= \Phi^k_{i,1} - \Phi^k_{i,2} \\ &= -\kappa\frac{I^k_i}{f_1^2} - \lambda_1\, N^k_{i,1} + \lambda_2 N^k_{i,2} + \epsilon^k_{i,\mathrm{GF}} \end{aligned} \tag{4.19}$$

The error term $\epsilon^k_{i,\mathrm{GF}}$ for carrier pseudo-range is too small and can be neglected. Taking the expectations of both sides of Eq. (4.19) we get:

$$\mathbb{E}\{\Phi^k_{i,\mathrm{GF}}\} = -\kappa\frac{I^k_i}{f_1^2} - \lambda_1\, N^k_{i,1} + \lambda_2 N^k_{i,2} \tag{4.20}$$

The next step is to compute the difference between two successive epochs, this will give the variation of the ionosphere in time stamp $\Delta t = t_2 - t_1$ when the cycle slip is absent. The integer ambiguities are constant between epoch consecutive epochs and will be canceled out.

$$\begin{aligned} \mathbb{E}\{\Delta\Phi^k_{i,\mathrm{GF}}\} &= \Phi^k_{i,\mathrm{GF},t2} - \Phi^k_{i,\mathrm{GF},t1} \\ &= \kappa\Big(\frac{I^k_{i,t1}}{f_1^2} - \frac{I^k_{i,t2}}{f_1^2}\Big) \\ &\quad + \lambda_1 N^k_{i,1,t2} - \lambda_2 N^k_{i,2,t2} - \lambda_1 N^k_{i,1,t1} + \lambda_2 N^k_{i,2,t1} \end{aligned} \tag{4.21}$$

Differencing operator is just another way to achieve stationarity. Stationarity allows us to compute the statistics of the time series and use them to construct the test to detect any anomalies, for instance discontinuities. This approach uses the difference between two successive observations in double-difference level, this is the triple difference.

As usual, the procedure is simple. if the computed absolute value of the difference is greater than a predefined constant T_h, the detection of cycle-slip is accomplished. The repair process requires always to determine the size of the jumps.

The cycle-slip is characterized by three parameters, the time of occurrence t_i, the threshold value (T_h) and the magnitude of the jump (δ). So the cycle-slip parameter vector $\boldsymbol{\theta} = (t_i, T_h, \delta)$. The values used in this study are $\delta = .2$ cycles and $T_h = .35$ cycles. The last step is to choose the threshold T_h value empirically based on some test statistics. We expect the output from the differencing operator to be stable. This task is accomplished by inspection of the variance components.

C. Prediction by auto-regressive process

For time-series $\Delta\nabla\Phi_{ab}^{xy}$, a prediction of the current $\Delta\nabla\Phi_{ab,p}^{xy}(t)$ is done by an auto-regressive process of order p (previous epochs).

$$\Delta\nabla\Phi_{ab,p}^{xy}(t) = \sum_{i=1}^{p} \alpha_i \Delta\nabla\Phi_{ab}^{xy}(t-i) \tag{4.22}$$

where the constant α_i must be determined from the data. The detection criteria is based on computing the residual ϵ between the predicted and the computed values:

$$\epsilon = |\Delta\nabla\Phi_{ab,p}^{xy}(t) - \Delta\nabla\Phi_{ab,c}^{xy}(t))| \tag{4.23}$$

Detection process is based on choosing a threshold value T_h, if the value of $\epsilon > T_h$. Typical values are $T_h = 0.2$ cycles for L_1, and value of 0.3 cycles for wide-lane linear combinations.

D. Practical considerations

The correct cycle-slip detection and repair algorithm shall count for additional parameters that can degrade the performance of the algorithm, for instance the missing observations, receiver clock jumps and how many cycle-slip reparations are allowed in the working sliding window. The size

of the sliding window corresponds to the correlation length of the observations. The following rules are applied in the detection of the cycle-slip algorithm.

- The receiver clock jumps produce the same effects as cycle-slip, so the algorithm shall include procedures to handle the receiver clock jumps.

- The processing buffer must be re-initialized in case of high number of missing values. Allowable minimum is set to 2 or 3.

- Detection of cycle-slip is successfully accomplished but the repair is not possible. In this case the working buffer must be initialized.

4.2 Ambiguity resolution

This section is devoted to the ambiguity resolution. The double-difference carrier-phase observation is chosen as a case study. The procedure is applicable to un-differenced and single-differenced observations as well.

In order to generate a reliable user corrections, it is absolute necessary to determine the double-differenced phase ambiguities $\Delta\nabla N$ between reference receivers. The challenge in real time, is that we like to get the correct answer within a short period of time. Fixing all double differenced ambiguities $\Delta\nabla N$ relative to a common satellite S_1 in the whole network, is a process referred to as; *network operating at a common ambiguity level*.

A high expectation of the estimation process, requires a good satellite geometry with significant changes over time, redundant measurements, absence of the multipath at the reference receivers used under network processing and sometimes the noise referred to as un-modeled errors can pose problems.

Before going any further, it is nice to distinguish between a functional and a stochastic model of the process under study. The first describes the mathematical relationship between measurements and parameters to be estimated while the second focuses on statistic characteristic of the measurement noise.

Starting by parameterization of GPS positioning parameters and forming

the linearized double-difference observation equation:

$$\mathbb{E}\underbrace{\begin{bmatrix} \Phi^{kl}_{ij,1} - \Phi^{kl}_{ij,1,c} \\ \Phi^{kl}_{ij,2} - \Phi^{kl}_{ij,2,c} \\ P^{kl}_{ij,1} - P^{kl}_{ij,1,c} \\ P^{kl}_{ij,2} - P^{kl}_{ij,2,c} \end{bmatrix}}_{\textbf{response vector } z} = \underbrace{\begin{bmatrix} D_i & D_j & D_T & D_N & D_I \end{bmatrix}}_{\textbf{design matrix D}} \underbrace{\begin{bmatrix} \Delta x_i \\ \Delta y_i \\ \Delta z_i \\ \Delta x_j \\ \Delta y_j \\ \Delta z_j \\ \Delta TZD_i \\ \Delta TZD_j \\ N^{kl}_{ij,1} \\ N^{kl}_{ij,1} - N^{kl}_{ij,2} \\ \frac{I^{ij}_{ij}}{f_1^2} \end{bmatrix}}_{\textbf{parameter vector } \theta} \tag{4.24}$$

Before going any further, the computation of the corrected $\{\Phi^{kl}_{ij,x,c}\}$ for $x = 1, 2$, for all satellites (k and l) and sites (i and j) combinations are given by the following procedure. Starting by computing the receiver position and clock offset using the point positioning techniques(code pseudo-range) as first guess, then the final approximation of geometric ranges between satellites and the receivers are obtained.

The design matrices are:

$$D_i = \begin{bmatrix} \frac{x^k - x_i}{\rho^k_{i,c}} + \frac{x^l - x_i}{\rho^l_{i,c}} & \frac{y^k - y_i}{\rho^k_{i,c}} + \frac{y^l - y_i}{\rho^l_{i,c}} & \frac{z^k - z_i}{\rho^k_{i,c}} + \frac{z^l - z_i}{\rho^l_{i,c}} \\ \frac{x^k - x_i}{\rho^k_{i,c}} + \frac{x^l - x_i}{\rho^l_{i,c}} & \frac{y^k - y_i}{\rho^k_{i,c}} + \frac{y^l - y_i}{\rho^l_{i,c}} & \frac{z^k - z_i}{\rho^k_{i,c}} + \frac{z^l - z_i}{\rho^l_{i,c}} \\ \frac{x^k - x_i}{\rho^k_{i,c}} + \frac{x^l - x_i}{\rho^l_{i,c}} & \frac{y^k - y_i}{\rho^k_{i,c}} + \frac{y^l - y_i}{\rho^l_{i,c}} & \frac{z^k - z_i}{\rho^k_{i,c}} + \frac{z^l - z_i}{\rho^l_{i,c}} \\ \frac{x^k - x_i}{\rho^k_{i,c}} + \frac{x^l - x_i}{\rho^l_{i,c}} & \frac{y^k - y_i}{\rho^k_{i,c}} + \frac{y^l - y_i}{\rho^l_{i,c}} & \frac{z^k - z_i}{\rho^k_{i,c}} + \frac{z^l - z_i}{\rho^l_{i,c}} \end{bmatrix} \tag{4.25}$$

and

$$D_j = \begin{bmatrix} \frac{x^k - x_j}{\rho^k_{j,c}} + \frac{x^l - x_j}{\rho^l_{j,c}} & \frac{y^k - y_j}{\rho^k_{j,c}} + \frac{y^l - y_j}{\rho^l_{j,c}} & \frac{z^k - z_j}{\rho^k_{j,c}} + \frac{z^l - z_j}{\rho^l_{j,c}} \\ \frac{x^k - x_j}{\rho^k_{j,c}} + \frac{x^l - x_j}{\rho^l_{j,c}} & \frac{y^k - y_j}{\rho^k_{j,c}} + \frac{y^l - y_j}{\rho^l_{j,c}} & \frac{z^k - z_j}{\rho^k_{j,c}} + \frac{z^l - z_j}{\rho^l_{j,c}} \\ \frac{x^k - x_j}{\rho^k_{j,c}} + \frac{x^l - x_j}{\rho^l_{j,c}} & \frac{y^k - y_j}{\rho^k_{j,c}} + \frac{y^l - y_j}{\rho^l_{j,c}} & \frac{z^k - z_j}{\rho^k_{j,c}} + \frac{z^l - z_j}{\rho^l_{j,c}} \\ \frac{x^k - x_j}{\rho^k_{j,c}} + \frac{x^l - x_j}{\rho^l_{j,c}} & \frac{y^k - y_j}{\rho^k_{j,c}} + \frac{y^l - y_j}{\rho^l_{j,c}} & \frac{z^k - z_j}{\rho^k_{j,c}} + \frac{z^l - z_j}{\rho^l_{j,c}} \end{bmatrix} \tag{4.26}$$

The troposphere path delay is estimated by Eq. (3.13) at each station. The design matrix D_T reads

$$D_T = \begin{bmatrix} M_w(Z^k_i) - M_w(Z^l_i) & M_w(Z^k_j) - M_w(Z^l_j) \\ M_w(Z^k_i) - M_w(Z^l_i) & M_w(Z^k_j) - M_w(Z^l_j) \\ M_w(Z^k_i) - M_w(Z^l_i) & M_w(Z^k_j) - M_w(Z^l_j) \\ M_w(Z^k_i) - M_w(Z^l_i) & M_w(Z^k_j) - M_w(Z^l_j) \end{bmatrix} \tag{4.27}$$

where Z is the zenith angle and M_w is the mapping function for the wet delay.

The design matrix D_N depends strongly on the selected type of linear combination of the observations. Useful linear combination for the GPS positioning is the wide-lane ($\lambda_{\text{wl}} = 86.19$) cm. The design matrix D_N for L_1 and wide-lane read:

$$D_N = \begin{bmatrix} \lambda_1 & 0 \\ \lambda_2 & -\lambda_2 \\ 0 & 0 \\ 0 & \end{bmatrix} \tag{4.28}$$

The design matrix D_I corresponding to the ionospheric parameter on L_1 reads:

$$D_I = \begin{bmatrix} -1 \\ -\frac{f_1^2}{f_2^2} \\ 1 \\ \frac{f_1^2}{f_2^2} \end{bmatrix} \tag{4.29}$$

The linearized double-difference observations equation can be modeled as a mixed model with the smooth terms as random effects.

$$\boldsymbol{z} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{A}\,\boldsymbol{y} + \epsilon \quad , \quad \boldsymbol{y} \sim \mathcal{N}(0, \boldsymbol{\psi}), \quad \epsilon \sim \mathcal{N}(0, \boldsymbol{\Lambda}\sigma^2) \tag{4.30}$$

where random vector $\boldsymbol{y}$ contains random effects, with zero expected value and covariance matrix $\boldsymbol{\psi}$, and $\boldsymbol{A}$ is a model matrix for the random effects. $\boldsymbol{\psi}$ is a positive definite matrix of simple structure, which is typically used to model residuals auto-correlation: its elements are usually determined by a simple model, with a few (or no) unknown parameters.
Often $\boldsymbol{\Lambda}$ is chosen to be an identity matrix.

The process of determining the ambiguities known as ambiguity resolution is actually a four step processing:

- **Float solution**: Estimating the real-valued ambiguities known as float solution. Least-squares techniques Eq. (4.30) or Kalman filter Eqs. (2.55) and (2.56) are typically used for this purpose.

- **Fix solution**: Determination of ambiguity search space by mapping or variable transformation to decorrelate the ambiguities. The aim is to obtain a fix solution. That is, the initial number of cycles, N, at the first observation epoch is correct and is an integer. LAMBDA method of Teunissen [127] can be used for this purpose.

- **Ambiguity validation**: Validation known as discrimination test based on some well defined metric. It is used to determine the correct integer set and is based on difference or ratio tests. This task has gained less attention than the other mentioned tasks.

- **Ambiguity administration**: This task is totally ignored and refers to the administration of the ambiguities. The ambiguity administration process refers to the following activities:
  - How to handle the rising and setting satellites ?
  - Base satellite changes cause the jump in the carrier-phase. How the jump is fixed and monitored ?
  - Monitoring the variance of the ambiguities. If the variance of a specific satellite exceeds the upper limit, everything shall be re-initialized.
  - Definitions of the threshold value (T_h) for outliers and cycle-slips (T_{cs}), respectively.

Monitoring the Kalman gain

One remedy to deal with the Heywood effect [44] is to carry out continuously inspection of the diagonal elements of the error covariance matrix, monitoring the Kalman filter gain and computation of condition number to measure the degree of ill-conditioned of a matrix (error covariance matrix of the state vector).

4.2.1 Ambiguity resolution float solution

The ambiguity resolution float solution is estimated by least squares technique Section 2.5.4 or Kalman filter Section 2.7.2. Assumed that the process and measurement noises are Gaussian distributed, otherwise the regularized particle filter or Extended Kalman Filter (EKF) are an appropriate choice. The state vector $\boldsymbol{x}$ for unknown model parameters and the corresponding variance-covariance matrix $\boldsymbol{P}$ can be estimated. Consider at epoch t_k we observe the measurement vector y_k.

First the state vector $\boldsymbol{x}$ reads:

$$\boldsymbol{x} = \begin{bmatrix} \delta x & \delta y & \delta z & N_{1_{WL}} & \cdots & N_{n_{WL}} \end{bmatrix}^T \tag{4.31}$$

where

$\left(\delta \boldsymbol{x}\ \delta y\ \delta z\right)$ errors in the nominal position of remote receiver in ECEF frame.

$\left(N_{i_{WL}}\right)_{i=1}^{n}$ double difference wide-lane ambiguities to be estimated.

As we mentioned before, the Kalman filter contains two main models; the first one is an VAR(1) process describing the state dynamics evolution from one epoch to the next, and how the error covariance of the state vector develops over time. The measurement model relates the observations to the state vector. The components of the Kalman machinery are described in more details in Section 2.7.2.

4.2.2 Kalman filter dynamic model definition

In order to determine the dynamics of the system under study, it's absolute necessary to select a suitable stochastic model for noise sources that affects the system error state covariance. These models are chosen with attention that are physically realizable. Realizability is achieved if the derivative model has a finite variance (controllable). In most navigation system, the models used are random-walk or the Gauss-Markov process.

The double-difference wide-lane state vector $\boldsymbol{x}$ contains states that correspond to the position coordinates, troposphere wet delay, ionospheric slant delay and ambiguities. In case of short baseline, the ionospheric and troposphere path delays are ignored. Anyway, the position, multipath, and atmospheric delay are modeled as random-walk (RW) or first order Gauss-Markov process. The ambiguity states are assumed to be constant if no cycle-slip is present and modeled as random constant. The state vector reads:

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_p & \boldsymbol{x}_{T'} & \boldsymbol{x}_{I'} & \boldsymbol{x}_M & \boldsymbol{x}_N \end{bmatrix}^T \tag{4.32}$$

where $\boldsymbol{x}_p$, $\boldsymbol{x}_{T'}$, $\boldsymbol{x}_{I'}$, $\boldsymbol{x}_M$ and $\boldsymbol{x}_N$ are respectively state vector of position, troposphere and ionospheric residuals, multipath and ambiguities. The transition and the noise matrices ($\boldsymbol{\Phi}$ and $\boldsymbol{Q}$) will be presented as a sub-blocks.

Gauss-Markov process fits a large number of physical processes with reasonable accuracy. In addition, the mathematical representation of such process is relatively simple. The differential equation of GMP reads:

$$\dot{\boldsymbol{x}} = -\beta\, \boldsymbol{x} + \boldsymbol{w} \tag{4.33}$$

and the corresponding auto-correlation functions reads:

$$\phi_{xx} = \sigma^2\, \exp\{-\beta|\tau|\} \tag{4.34}$$

Eq. (4.34) shows that the auto-correlation of GMP decreases exponentially with time.

Starting by defining the state and transition matrices of state vector $\boldsymbol{x}$.

State vector position estimation

The GMP is used to estimate $\boldsymbol{\delta p} = (\delta x, \delta y, \delta z)$, the transition and the process noise matrices of positioning blocks are given by Eq. (4.35). Assuming a spectral density for the driving vector of $(Sp_{\delta x}, Sp_{\delta y}, Sp_{\delta z})$, with units of m^2/s^3.

$$\boldsymbol{\Phi_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{Q_1} = \begin{pmatrix} \frac{Sp_{\delta x}}{3}\Delta t^3 & 0 & 0 \\ 0 & \frac{Sp_{\delta y}}{3}\Delta t & 0 \\ 0 & 0 & \frac{Sp_{\delta z}}{3}\Delta t \end{pmatrix} \tag{4.35}$$

Δt is the transition time interval in seconds. Derivation of process noises for position, velocity and acceleration can be found respectively in [62] and [13, chap. 11].

State vector for ambiguities

Ambiguity is expected to be constant as long as the reference receiver tracks the satellite signal. The change occurs when there is a loss of carrier-phase lock and creates carrier-phase discontinuities. Therefore the ambiguities state is modeled as random constant. The corresponding transition and process noise matrices are given by Eq. (4.36).

$$\mathbf{\Phi}_2 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \qquad \mathbf{Q}_2 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \tag{4.36}$$

State vector for correlated errors

Correlated errors refer to the residuals of the ionosphere and the troposphere path delays. In case of short baseline (< 20) km, the correlated double difference measurement error terms are almost canceled, and we expect that the floating solution should be closer to the integer with minor difference. The complete transition and noise sub-block matrices read:

$$\mathbf{\Phi} = \begin{pmatrix} \mathbf{\Phi_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}_2 \end{pmatrix} \qquad \mathbf{Q} = \begin{pmatrix} \mathbf{Q_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{pmatrix} \tag{4.37}$$

In case of large distances, the atmospheric correlated errors, the ionosphere ($\Delta\nabla I^{'}$) and troposphere residuals ($\Delta\nabla T^{'}$) increase with separation distance between reference receivers and are not canceled. Assuming a first order GMP with driving noises w_I and w_T of spectral densities Sp_I and Sp_T respectively in units of m^2/s and time constant T_c in seconds. The complete state vector $\boldsymbol{x}_k$ at time k reads:

$$\boldsymbol{x}_k = [\delta x\ \delta y\ \delta z\ \ \Delta\nabla N_{1_{WL}} \cdots \Delta\nabla N_{n_{WL}}\ \ \Delta\nabla I_1 \cdots \Delta\nabla I_n\ \ \Delta\nabla T_1^{'} \cdots \Delta\nabla T_n^{'}]^T \tag{4.38}$$

The temporal correlation in the TEC is consistent with the first order GMP Skone[111, p.40], which implies that the variable $\Delta\nabla I$ is exponentially correlated. The continuous and discrete state space equations for $\Delta\nabla I$ are:

$$d\{\Delta\nabla I\} = -\alpha\Delta\nabla I + w_I \quad ; \quad \Delta\nabla I_{k+1} = e^{\{-\alpha\Delta t\}}\Delta\nabla I_k + w_{I_k} \tag{4.39}$$

where $T_c = 1/\alpha$ is the correlation time and $\Delta t = t_{k+1} - t_k$ is the prediction interval. The corresponding process noise and the transition sub-block

matrices are respectively given by the Eqs. (4.40) and (4.41).

$$\mathbf{\Phi_3} = \begin{pmatrix} e^{-\alpha_1 \Delta t} & 0 & \cdots & 0 \\ 0 & e^{-\alpha_2 \Delta t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{-\alpha_n \Delta t} \end{pmatrix} \tag{4.40}$$

$$\mathbf{Q_3} = \begin{pmatrix} \sigma_I^2(1 - e^{-2\alpha_1 \Delta t}) & 0 & \cdots & 0 \\ 0 & \sigma_I^2(1 - e^{-2\alpha_2 \Delta t}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_I^2(1 - e^{-2\alpha_n \Delta t}) \end{pmatrix} \tag{4.41}$$

Parameters α_i and σ_I^2 are empirical values for the correlation time and spectral density of the ionosphere parameter.

The troposphere residuals will be handled similarly.

$$\mathbf{\Phi_4} = \begin{pmatrix} e^{-\gamma_1 \Delta t} & 0 & \cdots & 0 \\ 0 & e^{-\gamma_2 \Delta t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{-\gamma_n \Delta t} \end{pmatrix} \tag{4.42}$$

$$\mathbf{Q_4} = \begin{pmatrix} \sigma_T^2(1 - e^{-2\gamma_1 \Delta t}) & 0 & \cdots & 0 \\ 0 & \sigma_T^2(1 - e^{-2\gamma_2 \Delta t}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_T^2(1 - e^{-2\gamma_n \Delta t}) \end{pmatrix} \tag{4.43}$$

Parameters γ_i and σ_T^2 are empirical values for the correlation time and the spectral density of troposphere parameter.

State vector for multipath

The state transition and the process noise of the multipath read:

$$\mathbf{\Phi_5} = \begin{pmatrix} e^{-\zeta_1 \Delta t} & 0 & \cdots & 0 \\ 0 & e^{-\zeta_2 \Delta t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{-\zeta_n \Delta t} \end{pmatrix} \tag{4.44}$$

$$\mathbf{Q_5} = \begin{pmatrix} \sigma_M^2(1-e^{-2\zeta_1\Delta t}) & 0 & \cdots & 0 \\ 0 & \sigma_M^2(1-e^{-2\zeta_2\Delta t}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_M^2(1-e^{-2\zeta_n\Delta t}) \end{pmatrix} \tag{4.45}$$

Parameters ζ_i and σ_M^2 are empirical values for the correlation time and spectral density of multipath parameter.

Finally, the complete transition and the process noise matrices for state vector $\boldsymbol{x}$ are given by the Eq. (4.46).

$$\mathbf{\Phi} = \begin{pmatrix} \mathbf{\Phi_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi_2} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Phi_3} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{\Phi_4} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{\Phi_5} \end{pmatrix} \qquad \mathbf{Q} = \begin{pmatrix} \mathbf{Q_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q_2} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q_3} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q_4} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q_5} \end{pmatrix} \tag{4.46}$$

4.2.3 Kalman Filter measurement model definition

The Kalman filter dynamic model equation describes the state evolution by monitoring the temporal variations of the unknowns while the measurement model describes the functional relationship between the observations and the unknowns trough the design matrix $\boldsymbol{H}$.

In order to carry out the filtering algorithm, the parameter vector $\mathbf{\Theta} = (\boldsymbol{H}, \mathbf{\Phi}, \boldsymbol{Q}, \boldsymbol{R})$ shall be defined, four matrices must be specified. In the previous section, the transition matrix $\mathbf{\Phi}$ and the process noise Q are defined. In this section, the design matrix $\boldsymbol{H}$ and the measurement covariance matrix $\boldsymbol{R}$ will be given.

The design matrix $\boldsymbol{H} = D_i - D_j$, where D_i and D_j are given respectively in Section 4.2. The standard deviations of the errors [80, p.250] in the code and carrier-phase differences are given by the following expression:

carrier-phase

$$\sigma\left(\epsilon_{(ij)}^{(kl)}\right) \approx 0.05 \;\; cycles \; (\approx 1 \; cm)$$

code

$$\sigma\left(e_{(ij)}^{(kl)}\right) \approx 1 \; m$$

We have to take into account the variance amplification under any linear combinations of carrier-phase observables. To evaluate the variance of the sum of two dependent random variables, we have to include in the computation the covariance.

$$\mathrm{Var}(aX_1 + bX_2) = a^2\mathrm{Var}(X_1) + b^2\mathrm{Var}(X_2) + 2ab\mathrm{Cov}(X_1, X_2) \tag{4.47}$$

In case of independence of random variables, the covariance term is ignored in Eq. (4.47). In case of wide-lane, the standard deviation is amplified by $\sqrt{2}$ by assuming that L_1 and L_2 are independent and with equal variance.

All parameters needed to estimate the float ambiguities by using the Kalman filter are determined. The filter's warming period is defined as the time when the filter starts to give the measurement the weight and reduce the estimation variance. In this case, the filter starts to depend more on measurements and far less on the prior distribution.

The next step is the use of the estimated state and covariance obtained from the filter update operation, the float solution can be resolved into integer by a well known algorithm, for instance ambiguity de-correlation algorithm discussed earlier.

4.2.4 Ambiguity validation test

Once the ambiguity is solved, the popular ratio test is performed for validation of correctness of ambiguity. Three parameters are chosen for evaluation of the performance of the algorithm, namely the percentage of correct fixes. A good performance results in a high percentage of correct fixes. The second performance measure is the percentage of incorrect fixes and we expect a low value for good performance The third and final parameter is the mean time to fix, which is defined as the average time required to determine the ambiguities which were calculated correctly. Any incorrect ambiguities are not included in the mean time to fix.

4.3 NRTK Corrections generation

This section is devoted to the generation of user-level corrections for positioning services. Note that centimeter-level accuracy also requires the observation and analysis of the carrier-phase. In addition, the nuisance parameters must be be determined to the same level of accuracy.

To better understand NRTK positioning, a good place to start is the principles behind the classical RTK method. This method uses only one reference station, and the rover has to work in close proximity to this station. This is mostly due to ionospheric, tropospheric, and orbital errors, which

scale with the distance from the station. Such an approach is sufficient for many users, who can usually carry out their work without issues. However, a major drawback is the lack of redundancy: if the base station experiences any kind of malfunction, the service is immediately interrupted.

NRTK methods attempt to remedy this issue, and has therefore been gaining traction. There are many ways to extend the RTK concept to a network service, including virtual reference stations (VRS), the master auxiliary method (MAC), the area parameter correction (FKP), and the pseudo-reference method. Herein, we focus on the VRS concept, because it is already widely employed and comparatively simple.

4.3.1 NRTK biases

We will start by writing out the basic observation equation, which summarizes the relationship between the carrier-phase observations and error sources,

$$\Phi = \rho + c(\delta t_r - \delta t^s) - I + T + m + \lambda N + \epsilon_\phi, \tag{4.48}$$

where Φ has units of length, and ϵ_ϕ contains the remaining errors from e.g. tidal effects, relativistic effects, and so on. Based on (4.48), and the linear combination techniques introduced in sections 3.6.1 and 3.6.2, we can implement two independent filters to estimate the so-called *dispersive* and *non-dispersive* biases. Dispersive biases exhibit a frequency-dependence, and are introduced when the signal passes through the ionosphere. Non-dispersive or geometric biases are frequency-independent, and are largely due to the tropospheric and clock errors. Implementing two filters has the advantage that we can handle these qualitatively different errors independently. This method is sometimes referred to as solving the network positioning in *un-differenced mode*.

4.3.2 Virtual reference station

The NRTK method requires at least three stations, with pairwise separations of less than 100 km. As illustrated in figure 5.2, it can be divided into three distinct segments:

- Data collection: Multiple reference stations collect GNSS satellite data and send them to the central processing facilities.

- Data manipulation: The raw observations are checked for outliers and their ambiguities are resolved.

- Bias interpolation: Once the network biases are computed at the reference receiver, the distance-dependent errors need to be interpolated to the location of the user. Several methods can be used for this purpose, and the most advanced ones can be found in Ouassou et al. [84].

The basic idea is then to simulate an unphysical "virtual" base station in closer proximity to the end-user than any of the physical "real" base stations, and use this virtual base station to generate synthetic RTK corrections.

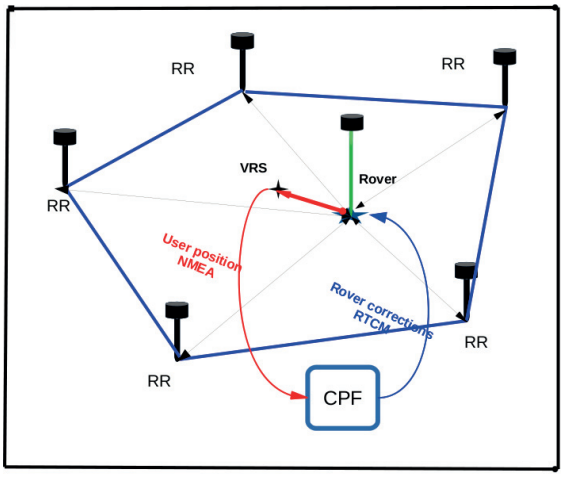


Figure 4.1: Illustration of NRTK VRS concept. CPF stands for *central processing facility*, RR for *reference receiver*, and VRS for *virtual reference receiver*.

The procedure of the NRTK VRS method is defined as follows:

- Approximate the rover position: A user in the field sends their estimated position to the CPF via NMEA protocol. The initial position can be computed by employing a code pseudo-range observables. This can be done using the algorithm in section 3.7.1, or alternatively using the Bancroft algorithm [118, Section 15.7].

- The CPF chooses a sub-network around the rover and estimates the biases, followed by an interpolation to the rover location.

- The rover uses the corrections to compute the baseline between the true position and the approximated position using the carrier-phase double-difference method. This step is known as the relative positioning. The algorithm described in section 3.7.2 is applicable. The methods

described in section 4.1.1 can then be used to generate a maximum set of independent double-differenced carrier-phase measurement.

In this sub-section, we tried to provide a short summary of how the network corrections can be generated. All mathematical processes defined in section 4.2 can be used to generate reliable corrections. For additional information about the VRS concept, the reader is referred to [46, pp. 188-191].

4.4 NRTK covariance structure definition

After determination of the correlated biases, which are composed of ionospheric, tropospheric, and satellite position biases, the next step is to determine the covariance structure of the remaining biases in the configured sub-network. The methods needed to construct the covariance structure are described in detail in section 2.8.

Ouassou et al. [84] showed that ordinary kriging with the Matérn covariance function is the most appropriate choice under normal circumstances, and produces a smooth solution with acceptable accuracy. The Matérn covariance function is well-behaved even for non-stationary fields, and is governed by only three parameters: location, scale, and shape. Stein [116] also recommended the use of the Matérn model due to its flexibility, i.e. its ability to model the smoothness and non-stationarity of physical processes.

.

5 NRTK data integrity design

"Every problem becomes very childish when once it is explained to you."
— Arthur Conan Doyle

The aim of this chapter is to demonstrate how the design process of the NRTK data integrity solution is carried out in this dissertation. Section 5.1 presents the results of marketing research conducted by NMA on NRTK services. For derivation of the NRTK data integrity algorithms, the network adjustment method *NetAdjust* is used as case study and is the topic of the Section 5.2. Section 5.3 presents the multivariate statistical analysis as a vital implementation tool. Section 5.4 presents in more details the overarching philosophy behind the NRTK design process. Finally, Section 5.5 presents some results of testing this procedure on real data sets.

5.1 NMA marketing research on NRTK

A marketing research was conducted by the NMA in 2010, where the CPOS NRTK users were questioned about the NRTK services. The main focus of the research is directed to the navigation performance parameters (availability, continuity, integrity, accuracy), reliability, and support. The objective is to help the NRTK users to carry out their job in satisfying conditions by developing the right product. The conclusion of the report is as follows:

1. Support: Highly requested. All NRTK users answered that they like to have a help in the field to carry out our job. Many of these users are not expert in GNSS, for instance machine guiding, road building, and others. Their message is clear, we like to avoid going next day to do the same job.
2. Integrity: Quality control of the work in the field is strongly desired.
3. Accuracy: Different levels of accuracy are required.
4. Continuity: For some NRTK users, the continuity of the service is critical. For these users, they are in the middle of something.

The follow-up actions considered were:

- Since NRTK users have different accuracy, this item must be configurable in the rover.

- All NRTK users required an error ellipse showing the true position and the rover position displayed in the rover display. In addition, height component.
- Stanford diagram was proposed but rejected. It is too complex for these users.

Based on this report, the NMA has defined a strategy to help the NRTK users regarding the design and the implementation of the NRTK data integrity. Two separate studies were carried out: the investigation of the network and the local data integrity. The best way to help the CPOS NRTK users is in the field, but this is impossible to provide. Therefore, our aim is to try to develop something that can be used by the NRTK users, and our conclusion is based on the following points:

- The NRTK data integrity service should be implemented as a new layer on top of the existing NRTK service.
- It should be independent of the NRTK software supply.
- It should be able to help the NRTK users in the field to carry their job in satisfying conditions.

Grid model proposed by augmentation systems is out of the question and the integrity concept of such system is not a good choice for NRTK users.

In addition, providing the user with integrity indicators such as ionosphere indicator I95 proposed by Trimble will not help to improve the user position, but it provides general information to identify the main reason of the problem.

5.2 Network adjustment method

Several NRTK techniques exist, and the most common used ones are the Master Auxiliary Concept (MAC) [30, 122], the Virtual Reference Station (VRS) concept [64], and the FKP techniques [141]. However, the Network Adjustment (NetAdjust) method developed by [89, 90] is chosen as case study for derivation and development of the integrity and quality control algorithms.

Most of the NRTK techniques mentioned above are developed commercially and details about these are not readily available. But the NetAdjust method is well-described in literature, it is therefore suitable as a starting point for our work, and we provide a review of the method in the following.

Our aim is to identify the exact locations in the NRTK data processing chain where data quality ought to be inspected and diagnosed. The result of

this analysis should be a list of suspicious satellites that generate anomalous data. The integrity monitoring algorithms developed and described in the remainder of this paper work independent of the method used for generation of the NRTK corrections.

5.2.1 Mathematical background of NetAdjust

The NetAdjust method is based on carrier-phase double-difference techniques. Taking the difference between the original observation signals allows us to eliminate or reduce the effect of uncanceled differential biases. In addition, the technique has the advantage of a reduction in both the measurement and parameter count. One need not to include the entire set of double-difference measurements because it contains redundant information. In the case of double-difference observations, receiver and satellite clock errors are eliminated, i.e. the results become independent of the receiver and satellite clock biases. In this work, the effect of residual double-differenced code and phase hardware delays is assumed to be negligible.

The overarching philosophy behind the NetAdjust design can be summarized as follows [90]:

1. *Main equation*:

$$\Delta\nabla\ell = \Delta\nabla\delta\ell + \Delta\nabla N \tag{5.1}$$

   Note that $\Delta\nabla$ is the double-difference operator and $\Delta\nabla\ell$ is the double-difference carrier-phase measurements, after subtracting range observables and the troposphere delay. This equation states that after correcting for double-difference ambiguity $\Delta\nabla N$, this is equivalent to the double-difference error $\Delta\nabla\delta\ell$, which is composed of residual atmospheric effects (ionosphere and troposphere), residual effects of the satellite position errors, as well as residual effects of multipath, receiver noise, antenna phase center variation, etc.

2. *NetAdjust signature*: Regardless of what ambiguity resolution algorithm one uses, the resolution is improved when the GNSS errors are minimized. This can be accomplished by reducing the uncertainties in the first term of Eq. (5.1), which facilitates the estimation of the second term, yielding improved ambiguity resolution.

3. *Error characteristics*: The NetAdjust method describes the error as a function of the position.

4. *Optimization*: Given the network measurements minus range observables and troposphere delay, one can estimate the differential measurement error $\delta\ell$ that minimizes the total variance. The optimal

estimator is determined using a Bayesian method, i.e. selecting a suitable loss function $L(\cdot)$ and thus an appropriate Bayes risk function $B(\cdot) = \mathbb{E}[L(\cdot)]$, where $\mathbb{E}$ is the expectation operator. For more details, e.g. [6] offers an elegant explanation of decision theory and Bayesian analysis.

5. *Prediction*: Least-squares collocation is a statistical estimation method that combines least-squares adjustment and prediction methods. The NetAdjust method uses the least-squares covariance analysis for accuracy prediction, i.e. to predict the carrier-phase error statistics for a given network configuration. For more details of this technique, the reader is referred to e.g. [87].

We will now provide a brief discussion of the mathematical details of the method. We assume that the relationship between the parameter vector $\mathbf{x}$ and observation vector $\mathbf{Y}$ is a simple linear model $\mathbf{Y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, where $\mathbf{e}$ is an error vector. The Bayesian optimal estimator $\hat{\mathbf{x}}_{\text{opt}}$ with quadratic loss function is then obtained by minimizing the Bayes risk

$$\hat{\mathbf{x}}_{\text{opt}} = \mathbf{C}_{\mathbf{xY}}\mathbf{C}_{\mathbf{Y}}^{-1}\mathbf{Y}, \tag{5.2}$$

where $\mathbf{C}_{\mathbf{Y}}$ is the covariance matrix between sample locations, and $\mathbf{C}_{\mathbf{xY}}$ the covariance matrix between sample and prediction locations. This is also known as the *kriging equation*, and is used to compute the weights $\mathbf{W} = \mathbf{C}_{\mathbf{xY}}\mathbf{C}_{\mathbf{Y}}^{-1}$. To be more specific:

1. The elements of the covariance matrix $\mathbf{C}_{\mathbf{Y}}$ for the locations $\mathbf{Y}$ in the sample are defined as:

$$\forall i, j: \quad \left[\mathbf{C}_{\mathbf{Y}}\right]_{ij} = \text{Cov}(Y_i, Y_j). \tag{5.3}$$

2. The elements of the covariance matrix $\mathbf{C}_{\mathbf{xY}}$ between the prediction points $\mathbf{x}$ and the sample locations $\mathbf{Y}$ are:

$$\forall i, j: \quad \left[\mathbf{C}_{\mathbf{xY}}\right]_{ij} = \text{Cov}(x_i, Y_j). \tag{5.4}$$

3. The NetAdjust estimator $\hat{\mathbf{x}}_{\text{opt}}$ is the optimal minimum variance error estimator. Note that Eq. (5.2) can also be written in the simple form $\hat{\mathbf{x}}_{\text{opt}} = \mathbf{W}\mathbf{Y}$, which is a linear function of the observation vector $\mathbf{Y}$, and takes into consideration the covariance structure of the problem when estimating the weight matrix $\mathbf{W}$.

   Computationally, the bottleneck when calculating the weight matrix $\mathbf{W}$ is the matrix inversion $\mathbf{C}_{\mathbf{Y}}^{-1}$. If the covariance matrix is large, the matrix inversion can become very time consuming. Moreover, if the matrix is ill-conditioned, there is also a risk of negative variance generation [44].

NetAdjust uses the kriging equation [Eq. (5.2)] to compute the network corrections. The corrections are then transmitted to the user, and involves them in the position computation process. For more details, the reader is referred to [90].

5.2.2 Prediction with NetAdjust method

NetAdjust method uses the covariance analysis to predict the code and carrier-phase error statistics. The key equation relating the total variance of an optimal estimator and the error covariance matrix is given.

For any estimator $e(Y)$ of type Linear Unbiased Minimum Variance (LUMV), the following expression holds [90, p. 54] .

$$\begin{aligned}\mathbf{E}(\mathbf{x}-e(\boldsymbol{Y}))^2 &= \mathbf{E}[(\mathbf{x}-e(\boldsymbol{Y}))^T(\mathbf{x}-e(\boldsymbol{Y}))] \\ &= \operatorname{trace}\left\{\mathbf{E}[(\mathbf{x}-e(\boldsymbol{Y}))(\mathbf{x}-e(\boldsymbol{Y}))^T]\right\}\end{aligned} \tag{5.5}$$

The left expression of the Eq. (5.5) is the Bayesian risk with quadratic loss function, while the right side is the total variance given by covariance of the estimator $e(Y)$.

Our aim is to construct the prediction function from our estimator. The judging parameter to be considered is the total variance of error variance-covariance matrix C_{err}.

$$g(C_{err}) = \mathrm{rms}_{\mathrm{pred}} = \operatorname{trace}(C_{err}) = \left\{\frac{1}{k}\sum_{i=1}^{k} C_{ii}\right\}^2 \tag{5.6}$$

where k is the number of satellites, C_{ii} are diagonal elements of the covariance matrix C_{err} and $g(C_{err})$ function is the Root Mean Square (RMS) of the diagonal elements of C_{err}.

5.2.3 Rover time-series analysis

All activities carried out so far are observation domain related. This section provides the analysis of rover position error. Using the techniques of Section 2.3, we are able to define the exact model describing the position error of the rover. In ideal situations, the rover position error shall follow the Gaussian distribution with a mean $\mu \approx 0$ and a finite small variance σ^2.

This type of analysis can be regarded as a feedback correction component to the applied algorithms in the observation domain and can help to clarify the main reason behind the detected anomalies. This way, we can understand the system better and determine the right procedures to minimize complications.

5.3 Multivariate statistical analysis as an implementation tool

Multivariate statistical analysis is a powerful and a vital tool to model and carry out quality control of GNSS data. The main tasks of GNSS data processing are the implementation of algorithms and procedures to estimate the nuisance parameters, analyze the data, interpretation of the data and decision making based on observed data. On the other hand, with combination of the GNSS constellations, for instance the GPS, GLONASS, Galileo and Beidou, massive data sets are generated. The data are multivariate in nature.

Anyway, the multivariate data analysis is not just a function of many variables but maybe these variables are correlated. In addition, the application of the multivariate analysis is huge and used in classification, cluster analysis, data reduction, processes control, path analysis, and pattern recognition and diagnosis.

In this study, we will use multivariate analysis to achieve two main goals. The first one is the administration of the GNSS data, it provides an elegant way to arrange and control all dynamic activities of GNSS constellations under data processing. The second one is to carry out the quality control. Note that all elementary materials needed are well described in Section 2.9, including hypothesis testing, distributions and test statistics.

5.3.1 MVA as GNSS administration tool

Administration of the dynamic of the GNSS data is not an easy task. The data analyst have to take into account the spatial and temporal correlations, handling the missing values problem, removing old satellites from computations, handling outliers, including new raising satellites and much more. The right and vital tool to achieve and control the dynamic of GNSS constellations is to apply the multivariate statistical analysis (MVA).

MVA offers all functions needed to treat the GNSS data with respect and independent of the differencing methods and linear combinations of the observations used to generate the network corrections.

Let $Y = \{Y_{ijk}\}$ be observations, where $i = 1,\ldots,n_{\text{rec}}$ are the reference receivers, $j = 1,\ldots,n_{\text{sat}}$ are the satellites observed at each site i, and k is the size of the moving window. The size of the moving window is equal to the correlation length of the observations used. According to [104], this correlation length is in the range of 300–600 seconds in the widelane case. Odolinski [83] presented two methods to estimate the correlation length, and found ∼17 min for the horizontal component, and ∼37 min for the vertical one. In any case, the larger the moving window, the lower the

correlation separation time.

The correlation time can also vary depending on the baseline length. For example, for short baselines of only a few kilometers, we expect only multipath errors and internal receiver effects to be relevant, and that these two factors will determine the correlation time. However, for longer baselines, larger correlation times can be expected if any residual atmospheric delays still remain.

We can describe Y as a matrix-valued sequence of length k, describing the dynamics of the network correction field $G(\boldsymbol{s}, t)$, where $\boldsymbol{s}$ and t are the spatial and the temporal spaces, while $G(\boldsymbol{s}, t)$ is an empirical Green function that is constructed from the data $Y = \{Y_{ijk}\}$.

The average error level and the corresponding uncertainties are computed and the quality control and data integrity are carried out. This is the subject of the next Section 5.3.2.

5.3.2 MVA as quality control and data integrity tool

According to the large sample theory, the correction field $G(\boldsymbol{s}, t)$ should be well described by a multivariate normal distribution known as a Gaussian field. This means that the distribution should converge to this regardless of the parent population we sample from. A p-dimensional multivariate normal distribution is given respectively by the Eq. (2.85).

This expression contains all necessary information needed to construct the test statistics and variance monitoring. Therefore the Eq. (2.85) can be rewritten as follows:

$$f(\mathbf{z}) = T_A(\mathbf{z}) \exp\left\{-\frac{1}{2} T_B(\mathbf{z})\right\} \tag{5.7}$$

where the notation $|\cdot|$ refers to the matrix determinant, and the functions T_A and T_B are defined respectively by the expressions $|2\pi|^{-p/2}|\Sigma|^{-1/2}$ and $(\mathbf{z}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\mathbf{z}-\boldsymbol{\mu})$. T_A and T_B are elementary building blocks of the test statistics used to this dissertation. Probabilities in multivariate case are presented by volumes under the surface over regions defined by intervals of the observations x_i. This constant is T_A. T_B is known as the squared Mahalanobis distance and T_A contains some interesting information that need to be monitored, namely the generalized variance given by the expression $|\Sigma_x|^{-1}$. The generalized variance provides a way of writing the information contained in Σ as a single judging number. In addition, the generalized variance appears in the expression of maximum likelihood estimation machinery [51, p. 172]:

$$L(\hat{\boldsymbol{\mu}}, \hat{\Sigma}) = (\text{const}) \times |\Sigma|^{-n/2} \tag{5.8}$$

where n is the number of observations. These are main reasons why this stochastic variable must be monitored by controlling the variations.

The advantages of using the MVA is huge. We benefit from the available mathematics needed to construct the hypothesis testing, studying the correlation among satellites, handling the missing observations, dimension reduction and more. Therefore the MVA is chosen as a design and implementation tool for NRTK data integrity.

5.4 Design of NRTK data integrity solution

The NRTK data integrity solution is a set of control procedures that check the correctness of the information provided by the NRTK system. Figure 5.1 presents four different levels of quality check that are required to carry out the check of the correctness of information provided by the NRTK data processing techniques.

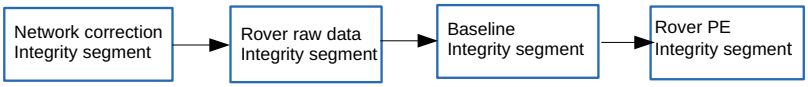


Figure 5.1: Work-flows NRTK data integrity segments for precise positioning. PE stands for position error.

5.4.1 Introduction to the design process

Our aim is to combine and use the information from the previous Sections 5.1–5.3 properly to design a product that complies with the user requirements and can be used by the NRTK users. This includes the output result from the NRTK users marketing research campaign, the expert opinions and information from the case study. The conclusion is that the design process must take into consideration the following parameters.

- **Data pedigree**: This is an expression used to study the performance of an animal. Horse racing study is carried out to find out if the animal has the chance to become a champion. The result from the study shows that if the mother or the father is a champion, there is a big chance of a descendant to become a champion.

  That is, assessing the pedigree of the data can help the data modeler to avoid accepting poor quality data and performing the wrong analysis. Therefore make sure that the process generating the NRTK corrections is well understood.

- **Nuisance parameters**: Models and algorithms used to estimate, mitigate the nuisance parameters must be understood in more details. A list of weakness and strengths of statistical methods used must be determined.
- **Error levels**:
  - Make sure that the error level and the corresponding uncertainty are well defined.
  - Define the distribution of the critical parameters.
  - Avoid shortcuts by imposing assumptions. This task will help to generalize the integrity solution and will work independently of the method used to generate the corrections.
- **Advanced tools**: Use modern statistical methods to make the right decision.
- **System improvements**: There is always a room for improvements. Identification of the weakness and the strength of the system will enhance the ability to improve the service and the integrity solution. This means one has to use modern statistics to build better spatio-temporal models, handle non-stationarity and non-Gaussian processes, and determination of correct covariance structure. Finally, one can avoid assumptions to generalize the method or the procedure.

Our aim is to design and implement a NRTK data integrity that is independent of the network processing techniques. The best way to achieve this goal is to extend the NRTK approach with a new segment known as the NRTK data integrity segment. Figure 5.2 shows all segments of the NRTK. The red box presents the new segment, namely the NRTK data integrity segment. The main function of this new segment is to carry out quality control, checking the correctness of corrections and the user in the field will be informed if the corrections cannot be trusted.

The new segment known as NRTK data integrity segments is composed into three components namely the network, baseline and the rover quality checks. Check barriers are key for implementations. Figure 5.3 shows the check barriers used as check points to ensure that the correctness of information is performed before going further to the next step. Each component will be treated separately in the next sections.

5.4.2 Network correction integrity segment

This is the system level check in order to produce the high quality synthetic observation data known as virtual reference station. This process requires

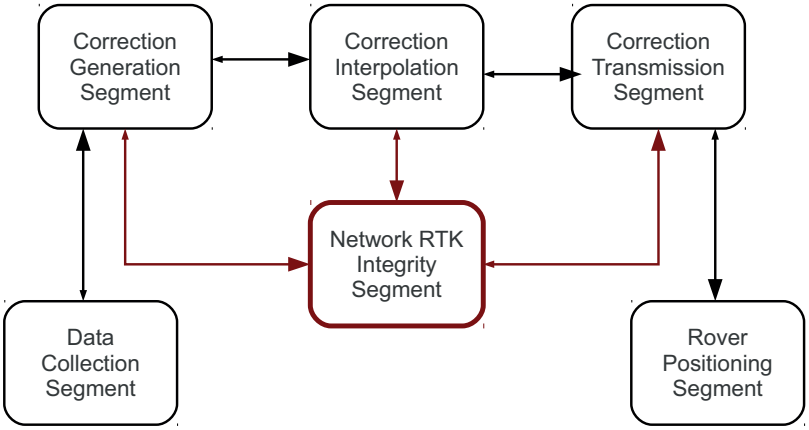


Figure 5.2: NRTK data processing segments. The red box presents the new segment, namely the NRTK data integrity segment.

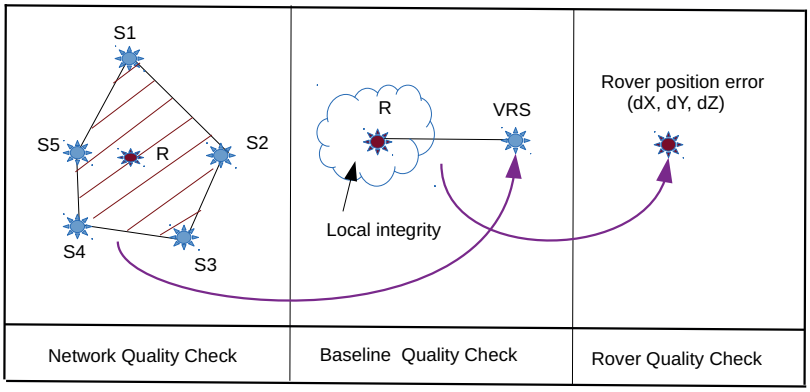


Figure 5.3: Check barriers of the network RTK data integrity. The left panel shows NRTK data integrity and produces a high-quality computation point, known as VRS. In this level, we have produced a high quality network average error level. The curved line indicates the output for this computation point. Middle panel shows the local and the baseline data integrity. Right panel shows the rover position solution.

a three steps processing, network reference receiver data integrity, corrections integrity and interpolation to generate the user synthetic corrections. These will be handled separately in the coming subsections. This task requires checking the input to the algorithms that are raw observations and the output that is the corrections and corresponding variance-covariance structure.

5.4.3 Network reference station data integrity

The main objective of this check barrier is to allow only high quality raw observations to compute the user network corrections. The quality check is at the observations level and often called station data integrity.

With station data integrity we mean performing the quality check on pseudo-range and carrier-phase measurements collected at the stations. This process ensures that the following checks are done.

- Satellite data integrity: Discarding measurements from unhealthy satellites or from satellites for which we do not have the orbital data.
- Controlled/uncontrolled cycle slip: Examination of loss-of-lock indicator (LLI), and signal-to-noise ratio (SNR) flags. Linear observation combinations defined in Section 3.6 are used to detect carrier-phase measurement discontinuities. Computation of Doppler and phase difference are recommended in this check barrier.
- Receiver clock offset: The receiver clock offset refers to the jump of the magnitude ±30 ms. The jump must be detected and corrected.
- Outlier detection: Robust algorithm to detect and remove outliers.
- Ground station coordinate: Procedure to detect the station coordinate is out of range.
- Prune satellites based on their elevation angle.
- Check and prune satellites based on their measured SNR and RMS error.

We expect at least the following valid parameters are present.

$$\Theta = (L_1, L_2, P_1, P_2, \mathrm{SNR}_1, \mathrm{SNR}_2, \mathrm{LLI}_1, \mathrm{LLI}_2, \mathrm{GDOP}_r, \mathrm{E}_l, \mathrm{A}_z).$$

5.4.4 Heywood case

Heywood cases are treated in Section 2.7.3 in more detail. In this section, we supplement additional information that is relevant for monitoring and events detection.

Key for precise positioning is correct determination and validation of the carrier phase ambiguity resolution. Often, this task is carried out by a Kalman filter [13, Figure 5.8]. Kalman gain $\boldsymbol{K}_k$ is involved in the computation of state vector update $\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k(\boldsymbol{z}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_k^-)$ and the corresponding error covariance matrix $\hat{\boldsymbol{P}}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\hat{\boldsymbol{P}}_k^-$.

The main diagonal elements of the variance-covariance matrix ($\hat{P}_k^+$) may become negative. These values correspond to the variance and by definition shall be positive. This phenomena is known as Heywood case [44]. Continuous inspection of the diagonal elements of the error variance-covariance ($\hat{P}_k^+$) is carried out on epoch-by-epoch basis to avoid the filter instability.

5.4.5 User correction data integrity

The output from NRTK processing filters are the ambiguities and the user corrections. These corrections address e.g. the unmodelled parts of the ionosphere,troposphere and satellite positions errors. This type of information usually are the residuals and corresponding variance-covariance matrices.

In order to construct the test statistics, it is necessary to determine the typical average level of biases on satellite basis and the corresponding uncertainties.

In addition, the sliding window is used to control the dynamic of the NRTK corrections. The size of the sliding window is set to the correlation length, i.e. the time span for which the observations can be considered completely decorrelated.

The procedure to construct the list of anomalous satellites is based on detection, isolation and decision. Two candidates are considered to carry out the global test and 4 possibilities to carry out the local test. For decision, the corrections from high residual values and variances are flagged for exclusion (NOT USE).

5.4.6 Selection of global test statistics

The correction contribution from each configured site in the network is assumed to follow multivariate normal distribution, with the mean vector $\boldsymbol{\mu}$ and variance-covariance matrix Σ.

As a result of the dynamic motion of satellites, new satellites periodically appear on the horizon, while others disappear. This event creates gaps in data set and make it difficult to construct the average error level for each site. The method chosen to handle this problem is the empirical orthogonal function known as principal component in the field of statistics. This task is repeated on epoch-by-epoch basis without destruction of the structure of the data and not introducing outliers. The empirical stochastic correction field $G(\mathbf{z})$ can be regarded as a function of $\mathbf{Y}_i$, where $\mathbf{z} = (\mathbf{s}, t)$ is a 4-dimensional vector in space $\mathbf{s}$ and time t. We will assume that it is a Gaussian field with a probability density function $f(\mathbf{z})$, which is parametrized by a mean vector $\boldsymbol{\mu}$ and variance-covariance matrix Σ and is given by the Eq. (2.85). Our aim is to detect the influential residuals that contribute to the

degradation of the rover position based on some distance metrics. Consider for instance the Euclidean metric given by the Eq. (5.9) that measures how far the observation x_i is from the mean μ_i. The method performs well mathematically, but is sensitive to the specific units of measurements.

$$E_i = (\mathbf{x}_i - \boldsymbol{\mu}_i)^{\mathrm{T}}(\mathbf{x}_i - \boldsymbol{\mu}_i) \tag{5.9}$$

Therefore one may wonder if there is a more informative way, particularly in a statistical sense, to measure if the distance $\mathbf{x}_i$ is far from the mean $\boldsymbol{\mu}_i$. One such metric is given by the *squared Mahalanobis distance* (SMD) defined in Eq. (5.10), which accounts for the correlations between the observations and measures the distance in units of standard deviations.

$$M_i = (\mathbf{x}_i - \boldsymbol{\mu}_i)^{\mathrm{T}}\, \Sigma^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_i) \tag{5.10}$$

The statistical characteristics of Eq. (5.10) need to be clarified. Consider now the set A:

$$A = \{\boldsymbol{x} \in \mathbb{R}^p : (\boldsymbol{x} - \bar{\boldsymbol{\mu}})\, \boldsymbol{S}^{-1}\, (\boldsymbol{x} - \bar{\boldsymbol{\mu}}) = c^2\}, \tag{5.11}$$

The set A is an ellipsoid centered at the mean $\bar{\mu}$, that is the data points which have equal Mahalanobis distance from the mean, that are located in the same ellipsoid centered at the mean. $\boldsymbol{S}$ is the sample variance-covariance matrix. An alternative metric is the *Mahalanobis depth* (MD):

$$m_i = \frac{1}{1 + M_i} \tag{5.12}$$

This time, we measure how far the observations $\mathbf{x}_i$ are from the median, and we note that large values of m_i correspond to values of x_i that are deep inside the distribution. Two test statistics are chosen to carry out the global test on network corrections. The Mahalanobis distance and depth. This task is often referred to as detection process. Other distance metrics exist to assess the influence of the observations, for instance the Cook's distance [95, p. 227], Andrews-Pregibon [71] and others that are not mentioned in this dissertation. Since our NRTK data integrity design and implementation is based on multivariate data analysis, and if we take a closer look at the formula of the multivariate distribution, Eq. (2.85), we see that the squared Mahalanobis distance appear in the exponent. We have explained that SMD accounts for the correlations between the observations and measures the distance in units of standard deviations and independent of the specific units of measurements. SMD is used to carry out the classification and used to detect the outliers, that is the observations that are far away from the mean field.

In addition, the most used test statistics in geodesy, surveying and navigation are the standard normal distribution, F-test, Student t-test and χ^2-test. So the SMD test actually follows the χ^2 distribution. For more information about the properties and benefits of SMD-based approaches, please consult Rousseeuw and Leroy [95]; S. Dasgupta [113]; Timm [133]. For Mahalanobis depth based approaches are described by Djauhari and Umbara [25]; Liu et al [72]; Mosler [81].

5.4.7 Selection of local test statistics

The main objective of local test is to isolate the suspicious satellite causing large deviance from the empirical average error level. This process is referred to as the isolation. Test statistics have to be constructed in order to carry out the isolation process. The test statistics is dependent on available information, starting with typical linear model such as Eq. (2.35), relating the response vector $\boldsymbol{z}$ to the variates $\boldsymbol{H}$. Assumptions about linearity, computational, and distribution are treated in Section 2.5.2 and will not be repeated. The stochastic error $\boldsymbol{\epsilon}$ is not directly observed and often is assumed to follow Gaussian distribution. The vector $\boldsymbol{e}$ reflects the variance function of the true model error. Considering GNSS temporal correlation, the least-squares residuals ($\boldsymbol{\epsilon}$) have different variabilities from others. This phenomena is called heteroscedastic and is due to dependency of satellite(s) elevation angle (trade-off between observation quality and the satellite geometry). Therefore, transformations are needed to maintain the temporal correlation properties with homogeneous variance (homoscedastic).

5.4.8 Rover raw data integrity

The best place to help the user is in the field. The raw carrier-phase observations must be investigated for multipath, scintillation and signal distortion before carrying out any further processing. That is, ensuring the quality of the carrier-phase observables can be viewed as a new check barrier in the NRTK data processing chain. In addition, all inspections carried out in Section 5.4.3 are applicable.

Low satellite elevation data may be included in processing and will increase the noise and systematic errors due to the large path of signal traveling through the ionosphere and the troposphere mediums. Several weighting schemes based on measured carrier-to-noise-power density ratio (C/N_0) are used to model the random error and the distortion of the signal. This shows that the C/N_0 values are highly correlated with satellite elevation angle. The need of the variance stabilization functions for carrier-phase observables to reduce the impact of biased observations are highly desirable.

Several variance functions are proposed to reduce the effects of biased observations on the final solution. These methods will not be repeated here, but we refer the interested reader to the work of Schon and Brunner [104] and A. Wieser [139]. A summary of the most currently applied re-weighting variance functions can be found X. Luo [74].

5.4.9 Baseline data integrity

Almost all high precision GNSS data processing strategies are based on double-difference carrier-phase observations (Section 4.1).

This is the last check barrier of the NRTK in the observation domain. The middle panel of Figure 5.3 shows the baseline data integrity. From the network an average biased non physical reference station is produced, known as a virtual reference station. The rover raw observation data are inspected (Section 5.4.8). Computation of the rover position in the relative positioning mode requires the analysis of the double-difference residuals and the inspection of the corresponding variance-covariance matrix for Heywood cases. Test statistics defined in sections 5.4.6 and 5.4.7 are applicable here.

5.5 Results of NRTK data integrity solution

Output from the processing filters (Section 5.4.9) are the rover position in topocentric coordinates and the corresponding variance covariance matrix $\boldsymbol{C}_p$. $\boldsymbol{C}_p$ describes how the measurement errors have been mapped into the calculated model parameters. It can be used to assess the uncertainty in the position components $(\delta e, \delta n, \delta u)$.

The covariance matrix can be used to define an error ellipsoid. Our aim is to present the rover position error ellipsoid at the rover display. Figure 5.4 shows the concept. It is easily seen that our concept can be seen as the NRTK data quality augmentation and the data integrity.

.

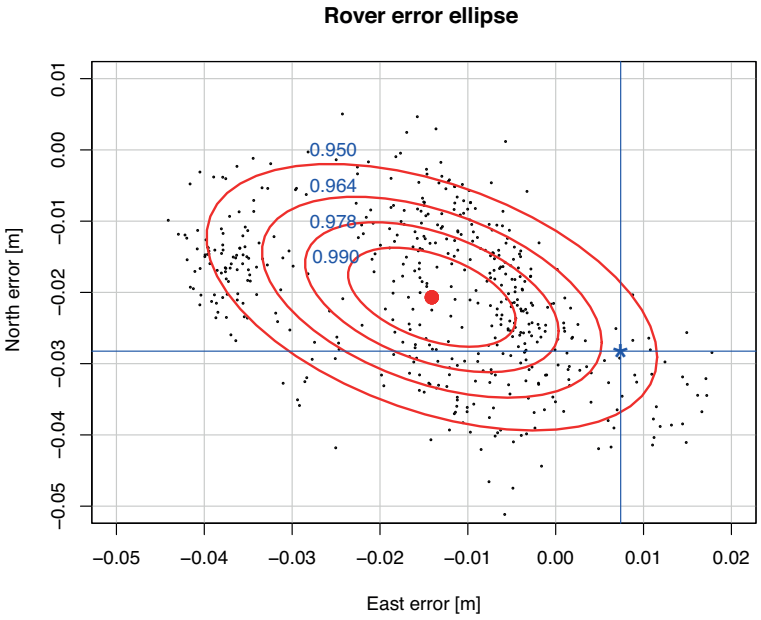


Figure 5.4: Error ellipse displaying the rover position error in the horizontal plane. The number of observations correspond to the correlation length. The center of the ellipse is displayed by the red point, while the user location is given by the intersection between the horizontal and vertical blue lines. Each ellipse correspond to the probability of acceptance of the null hypothesis

6 Summary and conclusion

"If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is."

— John von Neumann

The aim of this chapter is to summarize the design and the implementation process of the NRTK data integrity solution provided in this dissertation. Section 6.1 provides abstracts of published articles. Summaries of the work is the topic of the Section 6.2. Section 6.3 presents the discussions and conclusion. Finally, Section 6.6 gives the future work to be considered.

6.1 Paper abstracts

6.1.1 A continuous velocity field for Norway

Full citation: Halfdan Pascal Kierulf, Mohammed Ouassou, Matthew James Ross Simpson and Olav Vestøl. A continuous velocity field for Norway. Journal Of Geodesy April 2013, Volume 87, Issue 4, pp 337–349. DOI:10.1007/s0-0190-012-0603-2

Abstract: In Norway, as in the rest of Fennoscandia, the process of Glacial Isostatic Adjustment causes ongoing crustal deformation. The vertical and horizontal movements of the Earth can be measured to a high degree of precision using GNSS. The Norwegian GNSS network has gradually been established since the early 1990s and today contains approximately 140 stations. The stations are established both for navigation purposes and for studies of geophysical processes. Only a few of these stations have been analyzed previously. We present new velocity estimates for the Norwegian GNSS network using the processing package GAMIT. We examine the relation between time-series length and precision. With approximately 3.5 years of data, we are able to reproduce the secular vertical rate with a precision of 0.5 mm/year. To establish a continuous crustal velocity field in areas where we have no GNSS receivers or the observation period is too short to obtain reliable results, either interpolation or modeling is required. We experiment with both approaches in this analysis by using (i) a statistical interpolation method called Kriging and (ii) a GIA forward model. In addition, we examine how our vertical velocity field solution is affected by the inclusion of data from repeated leveling.

Results from our geophysical model give better estimates on the edge of the network, but inside the network the statistical interpolation method performs better. In general, we find that if we have less than 3.5 years of data for a GNSS station, the interpolated value is better than the velocity estimate based on a single time-series.

6.1.2 Next generation network real-time kinematic interpolation segment to improve the user accuracy

Full citation: Mohammed Ouassou, Anna B. O. Jensen, Jon G. O. Gjevestad, and Oddgeir Kristiansen. Next generation network real-time kinematic interpolation segment to improve the user accuracy. International Journal of Navigation and Observation, Volume 2015 (2015). DOI:10.1007/s00190-012-0603-2

Abstract: This paper demonstrates that automatic selection of the right interpolation/smoothing method in the Network Real-Time Kinematic (NRTK) interpolation segment can improve the accuracy of rover position estimates and the processing time in the NRTK processing center considerably.

The methods investigated for this purpose include Inverse Distance Weighting (IDW) methods; bi-linear and bi-cubic spline interpolation; kriging interpolation; thin-plate splines; and numerical approximation methods for spatial processes. Three days of selected Norwegian Centimeter POSition (CPOS) RTK network data sets with an average baseline between reference stations of order 60–70 km were used in this analysis. The data was selected from days with ionospheric activity that was characterized as high (up to 15 ppm), moderate (up to 4 ppm), and normal.

12 prediction locations were used to analyze the performance of the interpolation methods, by computing and comparing different measures of the goodness of fit: Root Mean Square Error (RMSE), Absolute Mean Error (AME), Goodness Of Fit (GOF) and Computation Time (CT). The tests were performed based on minimum values of variance and computation time, and the results show that ordinary kriging with the Matern covariance function and the thin plate spline method were usually the most appropriate choices. For the data sets with small variations, the Akima methods were the most effective due to processing time. However, when the covariance structure preserves sparsity, the numerical approximation methods require far less memory and computation time than all other methods. Finally, the IDW methods are simple and stable; but they do not conserve the spatial covariance structure, and were less accurate than the other algorithms.

6.1.3 Estimation of scintillation indices: A novel approach based on local kernel regression methods with bias corrected Akaike Information Criteria (AICC)

Full citation: Mohammed Ouassou, Oddgeir Kristiansen, Jon G. O. Gjevestad, Knut Stanley Jacobsen, and Yngvild L. Andalsvik. Estimation of Scintillation Indices: A Novel Approach Based on Local Kernel Regression Methods with bias corrected Akaike Information Criteria (AICC). International Journal of Navigation and Observation Volume 2016 (2016). DOI:10.1155/2016/3582176

Abstract: We present a comparative study of computational methods for estimation of ionospheric scintillation indices. First, we review the conventional approaches based on Fourier transformation and low-pass/high-pass frequency filtration. Next, we introduce a novel method based on nonparametric local regression with bias corrected Akaike Information Criteria (AICC). All methods are then applied to data from the Norwegian Regional Ionospheric Scintillation Network (NRISN), which is shown to be dominated by phase scintillation and not amplitude scintillation. We find that all methods provide highly correlated results, which demonstrates the validity of the new approach to this problem. All methods are shown to be very sensitive to filter characteristics, and the averaging interval. Finally, we find that the new method is more robust to discontinuous phase observations than conventional methods.

6.1.4 Network real-time kinematic data integrity by means of multivariate statistical analysis

Full citation: Mohammed Ouassou, Anna B. O. Jensen, and Jon G. O. Gjevestad. Network real-time kinematic data integrity by means of multivariate statistical analysis.
Submitted to International Journal of Navigation and Observation: 06/11/2017

Abstract: We introduce a novel approach to the computation of network real-time kinematic (NRTK) data integrity, which can be used to improve the position accuracy for rover receiver in the field. Our approach is based on multivariate statistical analysis and Stochastic Generalized Linear Model (SGLM). The network average error corrections and the corresponding variance fields are computed from data, while the squared Mahalanobis distance (SMD) and Mahalanobis depth (MD) are used as test statistics to

detect and remove the satellites that supply inaccurate data. The variance-covariance matrices are also inspected and monitored to avoid the Heywood effect, i.e. negative variance generated by the processing filters. The quality checks were carried out at both the system and user levels in order to reduce the impact of extreme events on the rover position estimates. The SGLM is used to predict the user carrier-phase and code error statistics. Finally, we present analyses of real-world data sets to establish the practical viability of the proposed methods.

6.1.5 Reliability Analysis of Network Real-Time Kinematic

Full citation: Mohammed Ouassou, Bent Natvig, Anna B. O. Jensen, and Jørund I. Gåsemyr. Reliability analysis of the Network real-time kinematic. Submitted to International Journal of Navigation and Observation: xx/yy/2017

Abstract: In this paper, the multi-state reliability theory was applied to the network real-time kinematic (NRTK) data processing chain, where the quality of the network corrections baseline residuals and the associated variance-covariance matrices are considered as the system state vectors. The state vectors have direct influence on the rover receiver position accuracy. The penalized honored stochastic averaged standard deviation (PHSASD) is used to map the NRTK sensitive data, represented by the states vectors to different levels of performance. The study shows that the improvement is possible by identification of critical components in the NRTK system and implementation some parallelism that makes the system more robust.

6.2 Summaries

This section summarizes the author's work.

6.2.1 Paper 1: A continuous velocity field for Norway

Author contribution: The author's main contribution was *chapter 4*, in addition to contributing to the general discussion of the ideas and results.

A novel spatio-temporal imputation algorithm based on combination of spatial linear interpolation known as kriging and resampling algorithms was developed and tested. The aim of the new method was the construction of the Matérn field for any spatio-temporal process in general. For validation, we have applied the algorithm to construct the velocity field for the entire mainland of Norway using the GNSS data for the first time.

The algorithm starts by predicting and generating a time series at locations inside the network where we have no observations, then the predicted values are treated as observations in next iteration. The covariance structure of the spatial process is estimated and the prediction at new locations are carried out. This process is repeated until the entire region of interest is treated.

To evaluate the performance of the algorithm, the validation process is carried out by excluding a location with real observations from computation and compare the predicted and the observed time series. The accuracies of the implemented algorithm are measured in terms of the velocity field [56, Fig. 5] of the topocentric components east, north and up. The final product is the construction of the velocity field for Norway based on GNSS data.

The article gained a lot of attention, especially from a Chinese research group. China mainland is a huge area and their interest is to use the same methods to predict and generate time series for some locations inside the network where no observations are available. For more details, see [106].

The relation of this method to the overarching topic of this thesis, is that the main objective of the spatio-temporal imputation algorithm is to study the impact of the missing observations on the NRTK data integrity solution.

Since the atmospheric effect on GNSS satellite signals is one of the largest error sources in real-time GNSS systems, such an algorithm is highly desirable by the NRTK data processing facilities. The secret of NRTK is to measure the correlated error that is composed of atmospheric path delays and orbit errors. If we can construct a complete regional correction field, it will enhance the ability to ameliorate the rover position. The integrity in this case is the analysis of the residual effects on rover position accuracy. In case of computing statistics in NRTK data integrity, missing observations occur often. The benefit of such method will avoid biased results of the estimation process.

6.2.2 Paper 2: Next generation network real-time kinematic interpolation segment to improve the user accuracy

Article goals: Processing huge data sets and be able to study the impact of the interpolation/smoothing algorithms on the NRTK data integrity solution.

Processing large data sets is a challenge of the future, and our suggestion for how to handle this is formulated as follows. First of all, we already have enough mathematical tools to do the job, so we do not need to develop new ones. These tools can be considered as elementary building blocks

in the hands of the data analyst/modeler. The main challenge is to know that the right tools exist, what they can do for us, what their strengths and weaknesses are, and how to combine them in appropriate ways to describe the observed variations as well as possible.

As massive data sets need to be processed in real-time, we are experiencing a computational paradigm shift. The main challenging issues in the NRTK data processing are:

- How do we generate high quality NRTK corrections ?

  Based on raw observations generated by a network of reference receivers, five sophisticated algorithms run in parallel and compete about the quality of service based on epoch-by-epoch basis. The transmitted user corrections are from the algorithm with a highest score. The main challenge here is how to construct better spatio-temporal models to monitor spatial-temporal processes that are not stationary and non-Gaussian.

- How do we handle the huge data sets ?

  Figure 1.1 shows how the number of NRTK users increased with time. By today we have around $\sim$ 4000 CPOS-users, and the maximum connected users at the same time is around 800. If the development continues the way it is today, we expect in one decade nearly 10 times the number of users. So we will have approximately 30000 users using the service and around 8000 users connecting simultaneously to the network.

  Therefore the solution to meet future needs to be able to process data from large geographical areas, as well as computing the necessary corrections and quality indicators ready for use, so that any RTK user that connects will be served immediately.

- What is the implementation tools ?

  The combination of stochastic partial differential equation (SPDE) and Gauss-Markov Random Fields (GMRF) offer flexibility to generate a good result with minimal processing time, and is robust against the big-n problem.

6.2.3 Paper 3: Estimation of scintillation indices: A novel approach based on local kernel regression methods with bias corrected Akaike Information Criteria (AICC)

Article goals: To study the impact of local disturbances on NRTK data integrity caused by the rapid variations of the TEC known as scintillation.

A good understanding of the ionosphere morphology can aid developing a suitable mitigation algorithm that assign weights to observations based on the scintillation distortion of each satellite link.

To be more specific, a more realistic stochastic model for GNSS observables would have to take into account the variance caused by scintillation. This would avoid biased solutions. Today, the stochastic model includes: the correlation among observations [128]; satellite elevation dependency modeled by exponential function [29]; temporal and cross correlations [16] and multipath detection and monitoring [139]. That is, a suitable robust weighting algorithm that reduces the influence of the satellite exposed to scintillation. The algorithm will enhance the ability to resolve the carrier-phase ambiguity and improve the stochastic model for the GNSS processes.

A reliable scintillation indices and definition of more realistic stochastic models for NRTK data processing. This task can be considered as *local data integrity* and is intended to protect the user(s) from errors caused by the scintillation.

Some of our main results were:

1. The performance of the digital filters increase by implementing the windowing techniques. The study shows that the Kaiser window [130, Chap. 7.2] is the most recommended window. The windowing techniques have the same function as the kernel in KDE.

2. The study shows that there is a direct connection between the filter cut-off frequency f_c and the optimal smoothing parameter h_{opt} in kernel density estimation (KDE).

3. The study shows also that the local regression with bias corrected Akaike Information Criteria (AICC) is a generalization of digital filters.

4. The study shows also that our concept is robust against carrier-phase discontinuities compared to all other methods.

6.2.4 Paper 4: Network real-time kinematic data integrity by means of multivariate statistical analysis

Article goals: This paper forms the backbone of this dissertation. Quality check of the correctness of the information provided by NRTK systems to the user are monitored. The average corrections and variances fields are determined from data. Global and local tests statistics are constructed, and variance-covariance matrices are checked for Heywood case. The carrier-phase and code statistics are predicted by employing the stochastic generalized linear model (SGLM).

As we mentioned in the introduction (Chap. 1) is that one inherent weakness of the GNSS NRTK software supplies not providing the user with a quality indicator that says something about the signal transmitted, i.e. the user corrections. The main objective of this paper is to augment the quality control of the NRTK data processing chain and to provide the user in the field with NRTK data integrity indicators. Indicators shall be displayed on the rover display. We conclude that our concept can be regarded as NRTK data quality augmentation and data integrity provider. The concept will protect the user in the field from the errors that are not detected by the NRTK processing algorithms. The integrity indicators help the user in the field to identify the main reason of position degradation.

6.2.5 Paper 5: Reliability Analysis of Network Real-Time Kinematic

Article goals: Application of the multi-state reliability theory to the NRTK data processing for the first time. The concept is able to predict the upper and lower performance levels of the rover position accuracy, in a fixed time.

The concept contributes to the improvement of the NRTK services to achieve the robustness, for instance the navigation performance parameters (availability, continuity, integrity and the accuracy).

This paper was written in direct cooperation between the NMA and the University of Oslo, Department of Mathematics.

The main objective of this paper is to study the performance of NRTK processing algorithms. It is a tool based on multi-state reliability theory and is able to detect the weakness of the system and try to improve it. A novel prediction algorithm known as penalized honored average standard deviation is used to retrieve the integrity indicators of the network, baseline data processing and the raw rover carrier-phase. The algorithm honors the

NRTK algorithms when they are performing well and punishes them in cases of destruction or degradation of the HDOP values. Generated indicators are key to study the impacts of the ambiguity algorithms, network corrections filters, and instruments (receivers, antenna, etc) on NRTK data integrity.

6.3 Conclusion

In this dissertation we have designed and implemented a GNSS NRTK data integrity solution that can facilitate the activities of the NRTK users in the field. Our approach can be viewed as a data quality control augmentation and data integrity provider. It is built on the top of the existing NRTK modules generating the network corrections. In addition, the concept is independent of the observations combinations and differencing methods used to generate the user corrections.

Multivariate statistic analysis is chosen as an implementation tool, where the network average error corrections and the corresponding variance fields are computed from the raw data and monitored. The Stochastic Generalized Linear Model (SGLM) is proposed as a prediction function of the rover carrier-phase and code error statistics.

An improvement of the rover position estimation can be achieved by applying procedures for data integrity monitoring at the system and user levels in NRTK processing chain. The methods tested makes it possible to identify satellites with bad data so these can be eliminated or down-weighted in the positioning process leading to an improvement in the rover position from epoch-to-epoch. Tests carried out as described in the dissertation show that there is indeed an improvement in the rover position after applying the new method.

It is expected that the suggested approach will reduce the number of wrong or inaccurate rover positions encountered by NRTK users in field, which subsequently will lead to a more efficient work flow for NRTK users. All test results shown in this paper are based on GPS data only, but the algorithms will work just as well with data from e.g. GLONASS or Galileo satellites. We addressed the problem of massive data processing. Our suggestion is formulated as follows: First of all, we already have enough mathematical tools to do the job, so we do not need to develop new ones. These tools can be considered as elementary building blocks in the hands of the data analyst or modeler. The main challenge is to know that the right tools exist, what they can do for us, what their strengths and weaknesses are, and how to combine them in appropriate ways to describe the observed variations as well as possible, for instance mapping the covariance structure of the network corrections.

6.4 Discussions

1. **Benefit from NRTK data integrity**:

   NRTK data integrity helps the user in the field. To benefit from the NRTK data integrity, use of the new RTCM 3.x [21] message types is recommended. From network data integrity, the anomaly list is produced and suspicious satellites are sent to the rover. The rover software must also be upgraded to be able to decode and use the data properly. This task requires a new software module to be implemented in the rover.

2. **Ambiguity resolution**: The key to precise positioning is correct determination and validation of the carrier phase ambiguity resolution. Often, this task is carried out by a Kalman filter [13, Figure 5.8]. Kalman gain K_k is involved in the computation of state vector update $\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k(\boldsymbol{z}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_k^-)$ and the corresponding error covariance matrix $\hat{\boldsymbol{P}}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\hat{\boldsymbol{P}}_k^-$. Therefore, $\hat{\boldsymbol{P}}_k^+$ must be inspected for Heywood case and $\boldsymbol{K}_k$ must be monitored correctly to avoid filter instability.

3. **Future GNSS data processing**: In order to meet the challenges of the future GNSS data processing, it is required that experts from different fields work closely together. These fields are computer science, mathematical statistics, and geophysics. The hope is that such a cooperation can be used to solve the GNSS big data problem.

The procedures of NRTK data integrity solution design and implementation defined in this dissertation are applicable to any GNSS positioning techniques, for example PPP, RTK, and RTK-PPP. The key is to define the functional and the stochastic models of the method, isolate and monitor the parameters that have direct influence on position degradation. Then compute the statistics and provide some level of trust to the user in form numbers or plots.

6.5 Future work

"The best time to plan an experiment is after you have done it."
R. A. Fisher

The work presented in this dissertation uses a few days of data and considering only the ionosphere activities as a parameter, varying from quiet to severe conditions. To improve the quality control procedures of NRTK data integrity it is necessary to include other constellations, for

instance GLONASS, Galileo, and Beidou and process data over a large period of time and for all seasons. Data rejection rate will be increased up to 10% while preserving the geometry.

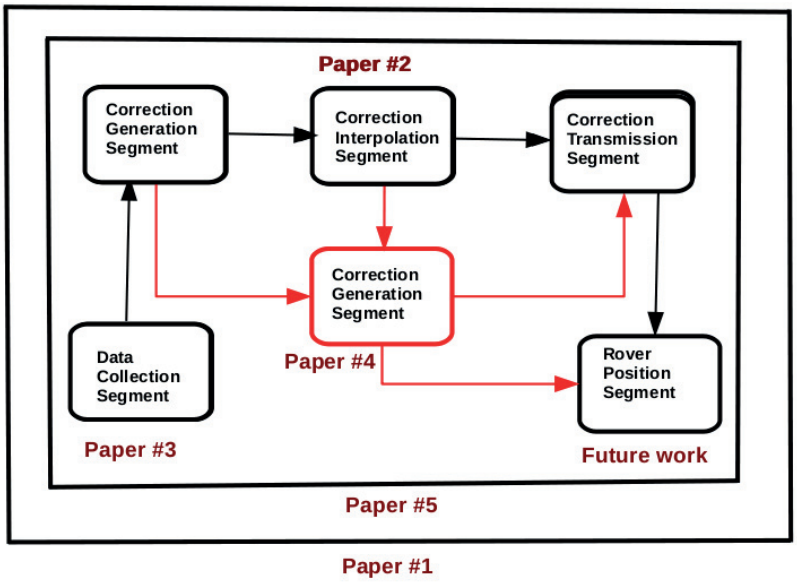


Figure 6.1: Figure links each Paper to the NRTK segment and specify the future work. Black boxes present the NRTK data processing segments while the red box presents the NRTK data integrity segment.

Figure 6.1 shows the mapping between the work done so far in form of papers and the NRTK segments. That is, shows the art of NRTK data integrity. The future work will be focused on rover position error (position domain) by enabling the algorithms described in the Papers. This includes:

- Impact of missing observations on NRTK data integrity solution. Topic of the Paper # 1 of this thesis.
- Impact of simultaneously parallel interpolations algorithms on NRTK data integrity. Topic of the Paper # 2 of this thesis.
- Impact of the GNSS stochastic variance models on NRTK integrity solution. Topic of the Paper # 3 of this thesis.
- Impact of various GNSS constellations on NRTK data integrity solution.

- Impact of ambiguity resolution based on particle filter on NRTK data integrity solution.

6.6 Future considerations

Although, the future NRTK service is moving from local corrections provider to global, the procedures developed in this dissertation to check for the correctness of the information are still applicable. That is, securing the quality of the user corrections. An area that is growing rapidly and changing completely our life is the mobile applications. The GNSS positioning solutions with a quality control for handset-based users is already in demand.

In 2011, Trimble NRTK software provider is responding with the same remedy by introducing a new product called Trimble CenterPoint RTX [143]. This new service utilizes real-time data from a global reference station infrastructure in order to compute satellite orbits and clocks and other system adjustments to the receiver, resulting in centimeter level positions that deliver repeatable high accuracy positions worldwide. That is, global corrections information with integrity solution to mobile devices. Of course, the NRTK SW providers competitors will follow immediately.

Standardization will facilitate the way the information is provided to the users, how to be processed and amelioration potential that come to the user benefits. In addition, amelioration of the services in need such as police, ambulance and fire department in order to reach the people in need quickly with well defined position.

Bibliography

[1] Abdel-salam, M.A.: Precise Point Positioning Using Un-Differenced Code and Carrier Phase Observations. Ph.D. thesis, Department of Geomatics Engineering, University of Calgary, Canada, UCGE Reports No. 20229 (2005)

[2] Adler, R.J., Taylor, J.E.: Random Fields and Geometry. Springer Monographs in Mathematics. Springer New York (2009)

[3] Baarda, W.: A Testing Procedure for Use in Geodetic Networks. v. 2, no. 5. Netherlands Geodetic Commission (1968)

[4] Banerjee, S., Carlin, B.P., Gelfand, A.E.: Hierarchical Modeling and Analysis for Spatial Data. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis (2004). URL https://books.google.no/books?id=A{_}R4AgAAQBAJ

[5] Bassiri, S and Hajj, G.: Higher-Order Ionospheric Effects on the GPS Observable and Means of Modeling them. AASI/AIAA Space flight Mechanics Meetings, Pasadena, California pp. 93 – 169 (1993)

[6] Berger, J.O.: Statistical Decision Theory and Bayesian Analysis. Springer Series in Statistics. Springer (1985)

[7] Berger, J.O., Bernardo, J.M., Mendoza, M.: On Priors that Maximize Expected Information. Purdue University. Department of Statistics (1988). URL https://books.google.no/books?id=YWUrrgEACAAJ

[8] Bierman, G.J.: Factorization Methods for Discrete Sequential Estimation. Mathematics in science and engineering. Academic Press (1977). URL https://books.google.no/books?id=96VpawEACAAJ

[9] Blangiardo, M., Cameletti, M.: Spatial and Spatio-temporal Bayesian Models with R - INLA. Wiley (2015). URL https://books.google.no/books?id=QQ{_}QBwAAQBAJ

[10] Bollen, K.A.: Outliers and improper solutions: A confirmatory factor analysis example. Sociological Methods & Research **15**(4), pp. 375 – 384 (1987)

[11] Borre, K., Akos, D.M., Bertelsen, N., Rinder, P., Jensen, S.H.: A Software-Defined GPS and Galileo Receiver: A Single-Frequency Approach. Applied and Numerical Harmonic Analysis. Birkh{ä}user Boston (2007). URL https://books.google.no/books?id=x2g6XTEkb8oC

[12] Brockwell, P.J., Davis, R.A.: Introduction to Time Series and Forecasting. v. 1. Springer (2002). URL https://books.google.no/books?id=VHB4OSAmwcUC

[13] Brown, R.G., Hwang, P.Y.C.: Introduction to random signals and applied Kalman filtering: with MATLAB exercises and solutions. v. 1. Wiley (1997)

[14] Burden, R.L., Faires, J.D.: Numerical Analysis. Prindle, Weber & Schmidt series in mathematics. PWS-Kent Publishing Company (1993). URL https://books.google.no/books?id=75PBQgAACAAJ

[15] Capua, R.: High Integrity Augmentation for the European Train Control System. ION conferance - Tampa 2015 (2015)

[16] C.C.J.M. Tiberius: Recursive data processingfor kinematic GPS surveying. Netherlands Geodetic Commission (1998)

[17] Chatterjee, S., Hadi, A.S.: Sensitivity Analysis in Linear Regression. Wiley Series in Probability and Statistics. Wiley (1988). URL https://books.google.no/books?id=D-IfpBnR0IEC

[18] Chen, F., Bollen, K.A., Paxton, P., Curran, P.J., B., K.J.: Improper Solutions in Structural Equation Models: Causes, Consequences, and Strategies. Sociological Methods & Research **29**(4), 468–508 (2001). DOI 10.1177/0049124101029004003. URL http://smr.sagepub.com/cgi/doi/10.1177/0049124101029004003

[19] Claeskens, G., Hjort, N.L.: Model Selection and Model Averaging. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press (2008). URL https://books.google.no/books?id=sjQFwB2XEFAC

[20] Cole, S.R., Chu, H., Greenland, S.: Maximum Likelihood, Profile Likelihood, and Penalized Likelihood: A Primer. American Journal of Epidemiology **179**(2), 252–260 (2014). DOI 10.1093/aje/kwt245. URL https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwt245

[21] Commission-RTCM: RTCM Standard 10403.1 for Differential GNSS (Global Navigation Satellite Systems) Services: Version 3. Radio Technical Commission for Maritime Services (2006)

[22] Conker, R.S., El-arini, M.B., Aalbetson, T.W., Klobuchar, J.A., Doherty, P.H.: Description and Assessment of Real-Time Algorithms to Estimate the Ionospheric Error Bounds for WAAS*. Navigation **44**(1), pp. 77–88 (1997). DOI 10.1002/j.2161-4296.1997.tb01941.x. URL http://doi.wiley.com/10.1002/j.2161-4296.1997.tb01941.x

[23] Cressie, N.A.C.: Statistics for Spatial Data, revised edn. Wiley Series in Probability and Mathematical Statistics. Wiley (1993)

[24] Dach, R., Lutz, S., Walser, P., Fridez, P.: Bernese GNSS Software, 5.2 edn. Astronomical institute, University of Bern, Bern (2015)

[25] Djauhari, M., Umbara, R.: A redefinition of Mahalanobis depth function. Malaysian Journal of Fundamental and Applied Sciences **3**, pp. 150–157 (2007)

[26] van Driel, O.P.: On various causes of improper solutions in maximum likelihood factor analysis. Psychometrika **43**(2), pp. 225–243 (1978). DOI 10.1007/BF02293865. URL http://link.springer.com/10.1007/BF02293865

[27] El-arini, M.B., Conker, R.S., Albertson, T.W., Reagan, J.K., Klobuchar, J.A., Doherty, P.H.: Comparison of Real-Time Ionospheric Algorithms for a GPS Wide-Area Augmentation System (WAAS)*. Navigation **41**(4), pp. 393–414 (1994). DOI 10.1002/j.2161-4296.1994.tb01887.x. URL http://doi.wiley.com/10.1002/j.2161-4296.1994.tb01887.x

[28] El-Rabbany, A.E.S.: The Effect of Physical Correlations on the Ambiguity Resolution and Accuracy Estimation in GPS Differential Positioning. Technical report. Department of Geodesy and Geomatics Engineering, University of New Brunswick (1994). URL https://books.google.no/books?id=h4C{_}MwEACAAJ

[29] Eueler, H.J., Goad, C.C.: On optimal filtering of GPS dual frequency observations without using orbit information. Bulletin Géodésique **65**(2), pp. 130–143 (1991). DOI 10.1007/BF00806368. URL http://link.springer.com/10.1007/BF00806368

[30] Euler, H.J., Keenan, C.R., Zebhauser, B.E., Wübbena, G.: Study of a Simplified Approach in Utilizing Information from Permanent Reference Station Arrays. In: Proceedings of the National Technical

Meeting of the Satellite Division of the Institute of Navigation, (ION GPS 2001), vol. 104, pp. 371–391 (2001)

[31] Feng, S., Ochieng, W., Moore, T., Hill, C., Hide, C.: Carrier phase-based integrity monitoring for high-accuracy positioning. GPS Solutions **13**(1), pp. 13–22 (2009). DOI 10.1007/s10291-008-0093-0

[32] Forssell, B.: Radionavigation Systems. GNSS technology and applications series. Artech House (2008). URL https://books.google.no/books?id=xGdTPgAACAAJ

[33] Gauss, C.F., Bertrand, J., Trotter, H.F.: Gauss's Work (1803-1826) on the Theory of Least Squares. Statistical Techniques Research Group, Section of Mathematical Statistics, Department of Mathematical [sic], Princeton University (1957). URL https://books.google.no/books?id=-3ccGQAACAAJ

[34] Gel, Y.R., Gastwirth, J.L.: A robust modification of the Jarque–Bera test of normality. Economics Letters **99**(1), pp. 30–32 (2008). DOI 10.1016/j.econlet.2007.05.022. URL http://linkinghub.elsevier.com/retrieve/pii/S0165176507001838

[35] Gelb, A.: Applied Optimal Estimation. MIT Press (1974). URL https://books.google.no/books?id=KlFrn8lpPP0C

[36] Geng, J.: Rapid Integer Ambiguity Resolution in GPS Precise Point Positioning. University of Nottingham (2011). URL https://books.google.no/books?id=6TExnQAACAAJ

[37] Giremus, A., Tourneret, J.Y., Calmettes, V.: A Particle Filtering Approach for Joint Detection/Estimation of Multipath Effects on GPS Measurements. IEEE Transactions on Signal Processing **55**(4), 1275–1285 (2007). DOI 10.1109/TSP.2006.888895. URL http://ieeexplore.ieee.org/document/4133032/

[38] Green, P.J., Silverman, B.W.: Nonparametric Regression and Generalized Linear Models: A roughness penalty approach. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis (1993). URL https://books.google.no/books?id=-AIVXozvpLUC

[39] Grewal, M.S., Weill, L.R., Andrews, A.P.: Global Positioning Systems, Inertial Navigation, and Integration. Wiley (2007). URL https://books.google.no/books?id=799oi-elP0sC

[40] Grewal, M.S., Weill, L.R., Andrews, A.P.: Global Positioning Systems, Inertial Navigation, and Integration, vol. 14, 2 edn. John Wiley & Sons, Inc. (2007)

[41] Han, S., Rizos, C.: Selection and scaling of simultaneous baseline for GPS network adjustment, or correct procedures for processing trivial baselines. Geomatics Research Australasia **63**, pp. 51–66 (1995)

[42] Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. Springer Series in Statistics. Springer New York (2009). URL https://books.google.no/books?id=tVIjmNS3Ob8C

[43] Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics. Springer New York (2013). URL https://books.google.no/books?id=yPfZBwAAQBAJ

[44] Heywood, H.B.: On Finite Sequences of Real Numbers. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character pp. 486–501 (1931)

[45] Hofmann-Wellenhof, B., Lichtenegger, H., Collins, J.: Global Positioning System: theory and practice. Springer-Verlag (1994). URL https://books.google.no/books?id=1jBPAQAAIAAJ

[46] Hofmann-Wellenhof, B., Lichtenegger, H., Wasle, E.: GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and more. Springer Vienna (2008). URL https://books.google.no/books?id=Np7y43HU{_}m8C

[47] Hoque, M.M., Jakowski, N.: Higher order ionospheric effects in precise GNSS positioning. Journal of Geodesy **81**(4), 259–268 (2007). DOI 10.1007/s00190-006-0106-0. URL http://link.springer.com/10.1007/s00190-006-0106-0

[48] ICD-GPS-200C(1993): Interface Control Document

[49] Jarque, C.M., Bera, A.K.: A Test for Normality of Observations and Regression Residuals. International Statistical Review **55**(2), 163–172 (1987). DOI 10.2307/1403192

[50] Jeffreys, H.: An Invariant Form for the Prior Probability in Estimation Problems. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **186**(1007), 453–461 (1946). DOI 10.1098/

rspa.1946.0056. URL http://rspa.royalsocietypublishing.org/cgi/doi/10.1098/rspa.1946.0056

[51] Johnson, R.A., Wichern, D.W.: Applied Multivariate Statistical Analysis. Prentice Hall (2002)

[52] Jokinen, A., Feng, S., Schuster, C., Washington, W.O.: Precise Point Positioning and Integrity Monitoring with GPS and GLONASS. URL http://www.rin.org.uk/uploadedpdfs/conferenceproceedings/jokinenpaper2a-web.pdf

[53] de Jong, C.D.: Hydrography. Series on mathematical geodesy and positioning. DUP Blue Print (2002). URL https://books.google.no/books?id=GpgZAQAAIAAJ

[54] Kaplan, E., Hegarty, C.: Understanding GPS: Principles and Applications, Second Edition. Artech House mobile communications series. Artech House (2005)

[55] Kedar, S., Hajj, G.A., Wilson, B.D., Heflin, M.B.: The effect of the second order GPS ionospheric correction on receiver positions. Geophysical Research Letters **30**(16) (2003). DOI 10.1029/2003GL017639. URL http://doi.wiley.com/10.1029/2003GL017639

[56] Kierulf, H.P., Ouassou, M., Simpson, M.J.R., Vestøl, O.: A continuous velocity field for Norway. Journal of Geodesy **87**(4), 337–349 (2013). DOI 10.1007/s00190-012-0603-2. URL http://link.springer.com/10.1007/s00190-012-0603-2

[57] Kim, E., Walter, T., Powell, J.D.: Wide Area Augmentation System-Based Flight Inspection System. Journal of Aircraft **45**(2), 614–621 (2008). DOI 10.2514/1.32531. URL http://www.gps.gov/technical/ps/2008-WAAS-performance-standard.pdfhttp://arc.aiaa.org/doi/10.2514/1.32531

[58] Kleusberg, A., Teunissen, P.J.G.: GPS for geodesy. Environmental science. Springer (1998)

[59] Klobuchar, J.A.: Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users. Defense Technical Information Center (1987). URL https://books.google.no/books?id=r9j7jwEACAAJ

[60] Koch, K.R.: Parameter Estimation and Hypothesis Testing in Linear Models. Springer Berlin Heidelberg (2013). URL https://books.google.no/books?id=n3bvCAAAQBAJ

[61] Kok, J.J.: Statistical analysis of deformation problems using Baarda's testing procedures. In: Forty years of Thought, Delft **2**, 470–488 (1982)

[62] K.P. Schwarz, M.E.C., Wong, R.V.C.: A comparison of GPS kinematic models for the determination of position and velocity along a trajectory. manuscripta geodaetica (1989)

[63] Kuusniemi, H., Wieser, A., Lachapelle, G., Takala, J.: User-level reliability monitoring in urban personal satellite-navigation. IEEE Transactions on Aerospace and Electronic Systems **43**(4), 1305–1318 (2007). DOI 10.1109/TAES.2007.4441741

[64] Landau, H., Vollath, U., Chen, X.: Virtual Reference Station Systems. Journal of Global Positioning Systems **1**(2), 137–143 (2002). DOI 10.5081/jgps.1.2.137

[65] Le, N.D., Zidek, J.V.: Statistical Analysis of Environmental Space-Time Processes. Springer Series in Statistics. Springer New York (2006). URL https://books.google.no/books?id=I9IpfL{_}BrbsC

[66] Lee, Y.C.: Analysis of Range and Position Comparison Methods as a Means to Provide GPS Integrity in the User Receiver. Proceedings of the Annual Meeting of the Institute of Navigation, (Seattle, WA) pp. pp. 1 – 4 (1986)

[67] Lehmann, E.L.: Elements of Large-Sample Theory. Springer Texts in Statistics. Springer New York (2006). URL https://books.google.no/books?id=cNIRBwAAQBAJ

[68] Leick, A.: GPS Satellite Surveying. Wiley (2004). URL https://books.google.no/books?id=4qE6xYjYSHgC

[69] Leick, A.: GPS Satellite Surveying. John Wiley and Sons, Incorporated (2015)

[70] Li, B., Feng, Y., Shen, Y., Wang, C.: Geometry-specified troposphere decorrelation for subcentimeter real-time kinematic solutions over long baselines. Journal of Geophysical Research **115**(B11), B11,404 (2010). DOI 10.1029/2010JB007549. URL http://doi.wiley.com/10.1029/2010JB007549

[71] Little, J.K.: Influence and a Quadratic Form in the Andrews-Pregibon Statistic. Technometrics **27**(1), 13 (1985). DOI 10.2307/1270464

[72] Liu, R.Y., Serfling, R.J., Souvaine, D.L.: Data Depth: Robust Multivariate Analysis, Computational Geometry, and Applications. DIMACS series in discrete mathematics and theoretical computer science. American Mathematical Soc. (2003)

[73] Loader, C.: Local Regression and Likelihood. Statistics and Computing. Springer New York (2006). URL https://books.google.no/books?id=NpjeBwAAQBAJ

[74] Luo, X.: GPS Stochastic Modelling Signal Quality Measures and ARMA Processes. Ph.D. thesis, Karlsruhe Institute of Technology (KIT) (2013). DOI 10.1007/978-3-642-34836-5

[75] Mantalos, P.: Three different measures of sample skewness and kurtosis and their effects on the Jarque Bera test for normality. International Journal of Computational Economics and Econometrics 2(1), 47 (2011). DOI 10.1504/IJCEE.2011.040576. URL http://www.inderscience.com/link.php?id=40576

[76] Markov, A.A., Liebmann, H.: Wahrscheinlichkeitsrechnung: mit 7 Figuren. Teubner (1912). URL https://books.google.no/books?id=v{_}YovgAACAAJ

[77] Martinez, W.L., Martinez, A.R.: Computational Statistics Handbook with MATLAB. Chapman & Hall/CRC Computer Science & Data Analysis. CRC Press (2001). URL https://books.google.no/books?id=my-i9VNzCfAC

[78] Matérn, B.: Spatial variation: stochastic models and their application to some problems in forest surveys and other sampling investigations. Meddelanden fr{å}n Statens skogsforskningsinstitut. Stockholm. (1960). URL https://books.google.no/books?id=81wTAQAAIAAJ

[79] Maybeck, P.S.: Stochastic Models, Estimation, and Control. Mathematics in Science and Engineering. Elsevier Science (1982). URL https://books.google.no/books?id=L{_}YVMUJKNQUC

[80] Misra, P., Enge, P.: Global positioning system: signals, measurements, and performance. Ganga-Jamuna Press (2006)

[81] Mosler, K.: Depth Statistics. In: Robustness and Complex Data Structures, pp. 17–34. Springer Berlin Heidelberg, Berlin, Heidelberg (2013). DOI 10.1007/978-3-642-35494-6{_}2

[82] Niell, A.E.: Global mapping functions for the atmosphere delay at radio wavelengths. Journal of Geophysical Research: Solid Earth **101**(B2), 3227–3246 (1996). DOI 10.1029/95JB03048. URL http://doi.wiley.com/10.1029/95JB03048

[83] Odolinski, R.: Temporal correlation for network RTK positioning. GPS Solution **16**, 147–155 (2012). DOI 10.1007/s10291-011-0213-0

[84] Ouassou, M., Jensen, A.B.O., Gjevestad, J.G.O., Oddgeir, K.: Next Generation Network Real-Time Kinematic Interpolation Segment to Improve the User Accuracy. International Journal of Navigation and Observation **2015**, 1–15 (2015). DOI 10.1155/2015/346498

[85] Parkinson, B.W., Axelrad, P.: Autonomous GPS Integrity Monitoring Using the Pseudorange Residual. Navigation **35**(2), 255–274 (1988). DOI 10.1002/j.2161-4296.1988.tb00955.x. URL http://doi.wiley.com/10.1002/j.2161-4296.1988.tb00955.x

[86] Press, W.H.: Numerical Recipes 3rd Edition: The Art of Scientific Computing. Cambridge University Press (2007). URL https://books.google.no/books?id=1aAOdzK3FegC

[87] Pullen, S., Enge, P., Parkinson, B.: A New Method for Coverage Prediction for the Wide Area Augmentation System (WAAS) (1995)

[88] Ramjee, P., Ruggieri, M.: Applied satellite navigation using GPS, GALILEO, and augmentation systems. Boston ; London : Artech House, 2005 (2005)

[89] Raquet, J., Lachapelle, G.: Development and Testing of a Kinematic Carrier-Phase Ambiguity Resolution Method Using a Reference Receiver Network 1. Navigation **46**(4), 283–295 (1999). DOI 10.1002/j.2161-4296.1999.tb02415.x

[90] Raquet, J.F.: UCGE Reports Number 20116 Development of a Method for Kinematic GPS Carrier-Phase Ambiguity Resolution Using Multiple Reference Receivers By (20116) (1998)

[91] Rhonda Slattery, Steve Peck, John Anagnost, Moon, M.: Selected papaers on satellite based augmentation systems (SBASs). The institute of navigation (ION)

[92] Rider, L.: Analytic design of satellite constellations for zonal earth coverage using inclined circular orbits. Journal of the Astronautical Sciences **Vol. 34**, pp. 31–64 (1986). URL http://adsabs.harvard.edu/abs/1986JAnSc..34...31R

[93] Rindskopf, D.: Structural Equation Models. Sociological Methods & Research **13**(1), 109–119 (1984). DOI 10.1177/0049124184013001004. URL http://journals.sagepub.com/doi/10.1177/0049124184013001004

[94] Ross S.M.: Introduction to probability models, 8 edn. Academic Press (2003). URL https://books.google.no/books?id=wlI9PgAACAAJ

[95] Rousseeuw, P.J., Leroy, A.M.: Robust Regression and Outlier Detection. Wiley Series in Probability an. Wiley (2003)

[96] Rowe, D.B.: Multivariate Bayesian Statistics: Models for Source Separation and Signal Unmixing. CRC Press (2002). URL https://books.google.no/books?id=Z3XLBQAAQBAJ

[97] RTCA, Inc., S..: Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment (2001)

[98] Rue, H., Held, L.: Gaussian Markov Random Fields: Theory and Applications. CRC Monographs on Statistics & Applied Probability. CRC Press, London (2005)

[99] Saalfeld, A.: Generating basis sets of double differences. Journal of Geodesy **73**(6), 291–297 (1999). DOI 10.1007/s001900050246. URL http://link.springer.com/10.1007/s001900050246

[100] Saastamoinen, J.: Contributions to the theory of atmospheric refraction. Bulletin géodésique **107**(1), 13–34 (1973). DOI 10.1007/BF02522083. URL http://link.springer.com/10.1007/BF02522083

[101] Samper, J.M., Lagunilla, J.M., Perez, R.B.: GPS and Galileo: Dual RF Front-end receiver and Design, Fabrication, & Test. Mcgraw-hill (2008). URL https://books.google.no/books?id=RBoNswEACAAJ

[102] Schabenberger, O., Gotway, C.A.: Statistical Methods for Spatial Data Analysis. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis (2004)

[103] Schaer, S.: Mapping and Predicting the Earth's Ionosphere Using the Global Positioning System. French Studies of the Eighteenth and Nineteenth Centuries. Institut f{ü}r Geod{ä}sie und Photogrammetrie, Eidg. Technische Hochschule Z{ü}rich (1999). URL https://books.google.no/books?id=zinKAAAACAAJ

[104] Schön, S., Brunner, F.K.: A proposal for modelling physical correlations of GPS phase observations. Journal of Geodesy **82**(10), 601–612 (2008). DOI 10.1007/s00190-008-0211-3

[105] Seeber, G.: Satellite Geodesy, 2 edn. Walter de Gruyter GmbH & Co (2003)

[106] Shen, Y., Li, W., Xu, G., Li, B.: Spatiotemporal filtering of regional GNSS network's position time series with missing data using principle component analysis. Journal of Geodesy **88**(1), 1–12 (2014). DOI 10.1007/s00190-013-0663-y. URL http://link.springer.com/10.1007/s00190-013-0663-y

[107] Sherman, M.: Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties. Wiley Series in Probability and Statistics. Wiley (2011). URL https://books.google.no/books?id=ebvy63s{_}f6YC

[108] Shumway, R.H., Stoffer, D.S.: Time Series Analysis and Its Applications: With R Examples. Springer Texts in Statistics. Springer New York (2006). URL https://books.google.no/books?id=Z7hAm5kj1sMC

[109] Simon, D.: Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches. Wiley (2006). URL https://books.google.no/books?id=UiMVoP{_}7TZkC

[110] Skone, S.: Wide area ionosphere grid modelling in the auroral region. Ph.D. thesis, UCGE Report No. 20123, University of Calgary, Calgary, Canada

[111] Skone, S., Feng, M., Tiwari, R., Coster, A.: "Characterizing Ionospheric Irregularities for Auroral Scintillations". In: Proceedings of the 22nd International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2009), pp. 2551–2558. Savannah, GA (2009)

[112] Smith, R.L.: Environmental Statistics. University of North Carolina (2001). URL https://books.google.no/books?id=HE9RngEACAAJ

[113] Somesh Dasgupta: The evolution of the D^2 statistic of the Mahalanobis. Indian journal of pure applied Math. **26**(6), 485–501 (1995)

[114] de Souza, E.M., Monico, J.F.G., Pagamisse, A.: GPS Satellite Kinematic Relative Positioning: Analyzing and Improving the Functional Mathematical Model Using Wavelets. Mathematical Problems in

Engineering **2009**, 1–18 (2009). DOI 10.1155/2009/934524. URL http://www.hindawi.com/journals/mpe/2009/934524/

[115] Steigenberger, P., Boehm, J., Tesmer, V.: Comparison of GMF/GPT with VMF1/ECMWF and implications for atmospheric loading. Journal of Geodesy **83**(10), 943–951 (2009). DOI 10.1007/s00190-009-0311-8. URL http://link.springer.com/10.1007/s00190-009-0311-8

[116] Stein, M.L.: Interpolation of Spatial Data. Springer Series in Statistics. Springer New York, New York, NY (1999). DOI 10.1007/978-1-4612-1494-6

[117] Stephens, M.A.: An Appreciation of Kolmogorov's 1933 Paper (1992). URL https://books.google.no/books?id=DceLPwAACAAJ

[118] Strang, G., Borre, K.: Linear Algebra, Geodesy, and GPS. Wellesley-Cambridge Press (1997). URL https://books.google.no/books?id=MjNwWUY8jx4C

[119] Sturza, M.A.: Navigation System Integrity Monitoring Using Redundant Measurements. Navigation **35**(4), 483–501 (1988). DOI 10.1002/j.2161-4296.1988.tb00975.x. URL http://doi.wiley.com/10.1002/j.2161-4296.1988.tb00975.x

[120] Subirana, J., Zonrnoza, J., Hermandez-Pajares, M.: GNSS data processing Volume I: Fundamentals and algorithms. Eurtopean Space Agency (ESA) (2013)

[121] Sun, X., Zhang, S., Hu, Q.: Particle Filter for Positioning Accuracy Improvement in GNSS Receiver. In: Proceedings of the 3rd International Conference on Electric and Electronics. Atlantis Press, Paris, France (2013). DOI 10.2991/eeic-13.2013.108. URL http://www.atlantis-press.com/php/paper-details.php?id=10345

[122] Takac, F., Zelzer, O.: The Relationship Between Network RTK Solutions MAC, VRS, PRS, FKP and i-MAX. Proceedings of the 21st International Technical Meeting of the Satellite Division of the Institute Of Navigation (ION GNSS 2008) pp. 348–355 (2008)

[123] Teunissen, P.: An integrity and quality control procedure for use in multi sensor integration. Proceedings of the 3rd international technical meeting of the satellite division of the institute of navigation (ION GPS 1990), Colorado Spring pp. 513–522 (1990)

[124] Teunissen, P., Montenbruck, O.: Springer Handbook of Global Navigation Satellite Systems. Springer Handbooks. Springer International Publishing (2017). URL https://books.google.no/books?id=93goDwAAQBAJ

[125] Teunissen, P., Odijk, D., Zhang, B.: PPP-RTK: Results of CORS Network-Based PPP with Integer Ambiguity Resolution. Journal of Aeronautics Astronautics and Aviation **42**(4), pp. 223 – 230 (2010)

[126] Teunissen, P.J.G.: Quality Control in Geodetic Networks. In: Optimization and Design of Geodetic Networks, pp. 526–547. Springer Berlin Heidelberg, Berlin, Heidelberg (1985). DOI 10.1007/978-3-642-70659-2{_}18

[127] Teunissen, P.J.G., De Jong, P.J., Tiberius, C.C.J.M.: Performance of the LAMBDA Method for Fast GPS Ambiguity Resolution. Navigation **44**(3), 373–383 (1997). DOI 10.1002/j.2161-4296.1997.tb02355.x. URL http://doi.wiley.com/10.1002/j.2161-4296.1997.tb02355.x

[128] Teunissen, P.J.G., Jonkman, N.F., Tiberius, C.C.J.M.: Weighting GPS Dual Frequency Observations: Bearing the Cross of Cross-Correlation. GPS Solutions **2**(2), 28–37 (1998)

[129] Teunissen, P.J.G., Khodabandeh, A.: Review and principles of PPP-RTK methods. Journal of Geodesy **89**(3), 217–240 (2015). DOI 10.1007/s00190-014-0771-3. URL http://link.springer.com/10.1007/s00190-014-0771-3

[130] Thede, L.D.: Practical Analog and Digital Filter Design. Artech House microwave library. Artech House (2005)

[131] Tian, Y., Ge, M., Neitzel, F., Zhu, J.: Particle filter-based estimation of inter-system phase bias for real-time integer ambiguity resolution. GPS Solutions **21**(3), 949–961 (2017). DOI 10.1007/s10291-016-0584-3. URL http://link.springer.com/10.1007/s10291-016-0584-3

[132] Tiberius, C.: Recursive data processing for kinematic GPS surveying. Publications on Geodesy / Netherlands Geodetic Commission. Nederlandse Commissie Voor Geodesie (1998). URL https://books.google.no/books?id=qPpMAQAAIAAJ

[133] Timm, N.H.: Applied Multivariate Analysis. Springer Texts in Statistics. Springer New York (2007)

[134] Tsai, Y.H., Chang, F.R., Yang, W.C.: GPS fault detection and exclusion using moving average filters. IEE Proceedings - Radar, Sonar and Navigation **151**(4), 240 (2004). DOI 10.1049/ip-rsn:20040728. URL http://digital-library.theiet.org/content/journals/10.1049/ip-rsn{_}20040728

[135] Ventura-Traveset, J., Agency, E.S.: EGNOS: The European Geostationary Navigation Overlay System : a Cornerstone of Galileo. ESA SP. ESA Publications Division (2006). URL https://books.google.no/books?id=jpnyGAAACAAJ

[136] Walker, J.G.: Satellite constellations. Journal of the British Interplanetary Society **Vol. 37**, pp. 559–572 (1984)

[137] Wang, E., Pang, T., Zhang, Z., Qu, P.: GPS Receiver Autonomous Integrity Monitoring Algorithm Based on Improved Particle Filter. Journal of Computers **9**(9) (2014). DOI 10.4304/jcp.9.9.2066-2074. URL http://ojs.academypublisher.com/index.php/jcp/article/view/13181

[138] Wang, M., Li, B.: Evaluation of Empirical Tropospheric Models Using Satellite-Tracking Tropospheric Wet Delays with Water Vapor Radiometer at Tongji, China. Sensors **16**(2), 186 (2016). DOI 10.3390/s16020186. URL http://www.mdpi.com/1424-8220/16/2/186

[139] Wieser, A., Brunner, F.K.: An extended weight model for GPS phase observations. Earth, Planets and Space **52**(10), 777–782 (2000). DOI 10.1186/BF03352281

[140] Witchayangkoon, B.: Elements of GPS precise point positioning. Ph.D. thesis (2000)

[141] Wübbena, G., Bagge, A., Seeber, G., Boeder, V., Hankemeier, P.: Reducing distance dependent errors for real-time precise DGPS applications by establishing reference station networks. In: Proceedings of Ion Gps, vol. 9, pp. 1845–1852 (1996)

[142] Wübbena, G., Schmitz, M., Bagge, A.: PPP-RTK: Precise point positioning using state-space representation in RTK network. Proceedings of ION GNSS 2005, Long Beach, CA, 13-16 Sep pp. 25,585–2594 (2005)

[143] Xiaoming Chen, Timo Allison, Wei Cao, K.F., son, Simon Grünig, Victor Gomez, Adrian Kipka, J., Köhler, Herbert Landau, Rodrigo Leandro, Gang Lu, R.S., Z, N.T.: Trimble RTX, an Innovative New Approach for Network RTK. Trimble TerraSat GmbH, Germany

[144] Xu, G.: GPS: Theory, Algorithms and Applications. Springer Berlin Heidelberg (2007). URL https://books.google.no/books?id=peYFZ69HqEsC

[145] Yan, X.: Linear Regression Analysis: Theory and Computing. World Scientific Publishing Company (2009). URL https://books.google.no/books?id=MjNv6rGv8NIC

[146] Ziedan, N.I.: GNSS Receivers for Weak Signals. Artech House Space Technology and Applications. Artech House (2006). URL https://books.google.no/books?id=bYNTAAAAMAAJ

[147] Zumberge, J.F., Heflin, M.B., Jefferson, D.C., Watkins, M.M., Webb, F.H.: Precise point positioning for the efficient and robust analysis of GPS data from large networks. Journal of Geophysical Research: Solid Earth **102**(B3), 5005–5017 (1997). DOI 10.1029/96JB03860. URL http://doi.wiley.com/10.1029/96JB03860

.

Paper #1

Reference

H.P. Kierulf, M. Ouassou, M.J.R. Simpson & O. Vestøl
A continuous velocity field for Norway
Journal of Geodesy **87(4)**, 337–349 (2013)

Contribution

M. Ouassou: The author's main contribution was writing part of chapter 4. Derivation of spatial-temporal imputation algorithm based on kriging and re-sampling algorithms. This includes conceiving the idea, performing the numerical calculations, and validation. As a result, the velocity for the mainland of Norway is derived for the first time by using spatial-temporal imputation algorithm. In addition to contributing to the general discussion of the ideas and results.

H.P. Kierulf: First author, coordinator, responsible for the quality of the manuscript and putting all pieces together.

M.J.R. Simpson: The author's main contribution was writing part of chapter 4 (GIA modeling). This includes performing the numerical calculations, and validation. In addition to contributing to the general discussion of the ideas and results.

O. Vestøl: The author's main contribution is the examination of how the vertical velocity field solution is affected by the inclusion of data repeated leveling. In addition to contributing to the general discussion of the ideas and results.

J Geod (2013) 87:337–349
DOI 10.1007/s00190-012-0603-2

ORIGINAL ARTICLE

A continuous velocity field for Norway

Halfdan Pascal Kierulf · Mohammed Ouassou · Matthew James Ross Simpson · Olav Vestøl

Received: 3 March 2012 / Accepted: 18 October 2012 / Published online: 9 November 2012

Abstract In Norway, as in the rest of Fennoscandia, the process of Glacial Isostatic Adjustment causes ongoing crustal deformation. The vertical and horizontal movements of the Earth can be measured to a high degree of precision using GNSS. The Norwegian GNSS network has gradually been established since the early 1990s and today contains approximately 140 stations. The stations are established both for navigation purposes and for studies of geophysical processes. Only a few of these stations have been analyzed previously. We present new velocity estimates for the Norwegian GNSS network using the processing package GAMIT. We examine the relation between time-series length and precision. With approximately 3.5 years of data, we are able to reproduce the secular vertical rate with a precision of 0.5 mm/year. To establish a continuous crustal velocity field in areas where we have no GNSS receivers or the observation period is too short to obtain reliable results, either interpolation or modeling is required. We experiment with both approaches in this analysis by using (i) a statistical interpolation method called Kriging and (ii) a GIA forward model. In addition, we examine how our vertical velocity field solution is affected by the inclusion of data from repeated leveling. Results from our geophysical model give better estimates on the edge of the network, but inside the network the statistical interpolation method performs better. In general, we find that if we have less than 3.5 years of data for a GNSS station, the interpolated value is better than the velocity estimate based on a single time-series.

Keywords Velocity field · GNSS · Kriging · GIA

H. P. Kierulf ( ) · M. Ouassou · M. J. R. Simpson · O. Vestøl
Geodetic Institute, Norwegian Mapping Authority,
3511 Hønefoss, Norway
e-mail: halfdan.kierulf@kartverket.no

H. P. Kierulf
Department of Geosciences, University of Oslo,
0316Oslo, Norway

1 Introduction

The establishment of a permanent Global navigation satellite system (GNSS) network in Norway began in the early 1990s. By the end of 2000, the network contained 15 continuously operating GNSS receivers, with around a third of these contributing data to the International GNSS Service (IGS) and/or the European Permanent Network (EPN). Following 2003, the network has undergone a densification (largely for navigation purposes) and now contains approximately 140 stations on the Norway mainland with an average spacing of 60 km (Fig. 1). This network provides a means to establish a well-constrained velocity field for Norway and a powerful tool for the study of geophysical phenomena. In this study, we examine data from the Norway mainland GNSS network as of the start of 2011. For stations where we have sufficiently long time-series, the majority of which have not been analyzed before, we present new GNSS derived velocities.

Observations of Glacial Isostatic Adjustment (GIA) from across Fennoscandia (the geographic areas of Finland, Norway and Sweden) show the ongoing relaxation of the Earth in response to past ice mass loss (e.g., Fjeldskaar 1994; Lambeck et al. 1998a,b; Milne et al. 2001; Johansson et al. 2002; Kierulf et al. 2003; Lidberg et al. 2007, 2010). Data from permanent GNSS stations provide a measure of movements in both the vertical and horizontal components of motion (Nocquet et al. 2005). The GNSS observations show two main features. Firstly, a pattern of Earth uplift with highest rates (∼10 mm/year in Gulf of Bothnia) corresponding to

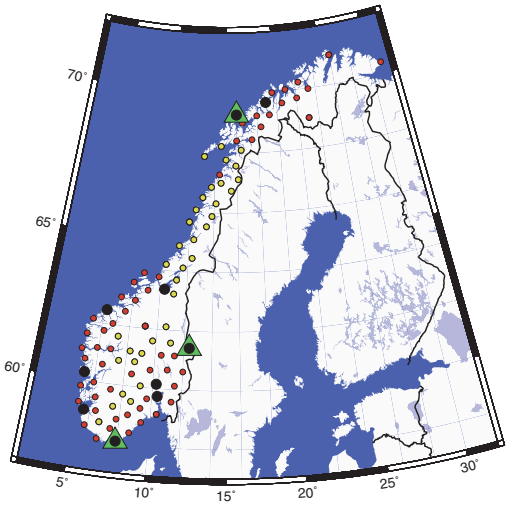


Fig. 1 Permanent GNSS stations on the Norway mainland. The *black circles* mark high quality sites established in 2000 or earlier for geodetic purposes (these stations are ANDO, ALES, BRGS HFSS, KRSS, OSLS, STAS, TROM, TRO1 and TRYS). *Red* and *yellow circles* mark stations established after 2001 and mainly built to serve the Norwegian positioning service. The *yellow circles* mark stations established after 2008 and are not used for the velocity estimates. *Green triangles* mark stations that are discussed in Sect. 4

areas of thickest ice during the last glacial period (∼21,000 years ago). Rates with lower, but still positive values, are shown for most of Norway (e.g., Vestøl 2006). Secondly, horizontal movements indicate a regional deformation characterized by an outward spreading from the centre of past maximum ice thickness.

Much of the previous work has been completed under the landmark Baseline Inferences for Fennoscandian Rebound, Sea-level, and Tectonics (BIFROST) project. Results from the BIFROST network have been published regularly (e.g., Milne et al. 2001; Johansson et al. 2002; Lidberg et al. 2007, 2010) and include some Norwegian and North European stations. In a separate investigation, Vestøl (2006) presented a model of land uplift based on a collocation method. He used observations from leveling, tide gauges and GNSS data from Fennoscandia and the nearby areas of continental Europe [the GNSS velocities used in his analysis are the same as in Lidberg et al. (2007)].

In this study, we use a scientific GNSS analysis software to derive daily results for the permanent Global Navigation Satellite System (GNSS) stations on the Norway mainland (Sect. 2). In the early years the Norwegian GNSS network went through several upgrades and equipment changes. After May 2000 the situation has been more stable. We have, therefore, opted to only include data from 2000-05-01 and later.

We have tested the precision of the time-series related to the total observation time (Sect. 3).To derive velocities in areas without GNSS or where the GNSS stations have observation periods that are to short to make reliable velocity estimates, we have developed two different interpolation routines based on the statistical concept of Kriging (Cressie 1993). We will also interpolate our results using geophysical model of Glacial Isostatic Adjustment (GIA) and leveling (Sect. 4).

GNSS results from the Arctic Norwegian islands Svalbard, Hopen, Bjørn Øya and Jan Mayen are not included in this paper. Results from Svalbard can be found in Kierulf et al. (2009a,b).

2 GNSS analysis-strategies

We use the GNSS analysis software GPS Analysis Software of MIT (GAMIT) (Herring et al. 2011). This software makes use of the so-called Double Difference (DD) approach, where a network of GNSS stations is analyzed in a single adjustment. A least square adjustment is used for parameter estimation. This implies that parameters which vary with time, for example, the troposphere, have to be estimated as piecewise linear parameters. The atmospheric zenith delay was estimated with a 2 hourly piecewise linear model together with a daily troposphere gradient. Ocean-loading coefficients (Scherneck 1991) from the FES2004-model are used. To model the tropospheric delay Vienna Mapping Functions (VMF) (Boehm et al. 2006) was used. The (igs05_*.atx) was used to model the phase centre variations. We have used a cut-off elevation of 7°.

To reduce the computational time, our network was divided into sub-networks analyzed individually and later on combined to daily results using GLOBK (Herring et al. 2011) to daily results. The daily result files were then transformed into ITRF2008 in a two step procedure. In the first step, a network of northern European IGS stations from the areas around Norway (GRAS, HERT, KIR0, MAR6, MDVO, METS, MORP, NYA1, NYAL, ONSA, POTS, RIGA, TRO1, TROM, WSRT and WTZR) were used to transform the minimally constrained daily solutions into ITRF2008. Preliminary results for all stations were extracted and velocities were computed. In the second step the procedure was repeated using the output from the first step, but this time including the vast majority of the Norwegian stations. This two step procedure using the dense-network stabilization will be more robust since we have a stronger realization of the frame on each day. This approach will remove most of the so-called common mode error, but since our connection to the ITRF is through a regional set of IGS stations, our results are partly de-coupled from the global reference frame. See Legrand et al. (2010) for a detailed analysis of the limitations

of regional reference frame realizations. Appendix 7.1 also includes a more general discussion on reference frame realizations and comparisons to previous studies.

3 Examination of time-series

It is widely recognized that the assumption of only white noise content is unrealistic for GNSS time-series (Johnson and Agnew 1995; Zhang et al. 1997; Mao et al. 1999; Williams et al. 2004). Williams et al. (2004) recommend a noise model combining both white noise and flicker noise for most GNSS sites. In this study the time-series analysis were performed using the CATS software (Williams 2008), using both a white noise model and a combination of white noise and flicker noise. We opt to include annual and semi-annual signals as additional parameters in our time-series analysis. In addition, parameters for offsets for all antenna and radome changes where included in the time-series analysis as well as parameters for offsets where breaks in the time-series were obvious after a visual inspection.

To examine the stability of velocity estimates we have performed a convergence analysis using solutions from stations established in year 2000 or earlier (black circles in Fig. 1). Velocities have been computed for each time-series, first using only the last 2.5 years of data, then the last 3.0 years of data and then extending the time period by 0.5 years until the last 10 years of data have been included. The RMS of the differences between the velocities for the shorter period and the velocities for the complete time-series (back to 2000-05-01, prior to this date equipment changes were performed more frequently and, therefore, results were less reliable) are calculated for each time span and component (Eq. 1).

$$RMS(\Delta t) = \sqrt{\sum_{i=1}^{n} \frac{(r_i(\Delta t) - \overline{r_i})^2}{n}}, \quad (1)$$

where $r_i(\Delta t)$ is the rate for station i with time-series of length Δt. $\overline{r_i}$ is the rate for station i using the complete time-series back to 2000-05-01 and n is the number of stations. This test gives a measure of the stability of the estimated secular rates as function of time-series length. Convergence studies for single stations time-series have earlier been performed for instance in Scherneck et al. (2002) and Vespe et al. (2002). The RMS differences are plotted in Fig. 2. We find that a precision of 0.5 mm/year is achieved after 3.0, 2.5 and 3.5 years in the north, east and vertical components, respectively. A 0.2 mm/year precision is achieved after 4.0, 5.5 and 5.5 years in the north, east and vertical component, respectively.

In Fig. 3, the mean uncertainties of the velocity estimates are plotted against time-series length. The uncertainty estimations are based on a noise model that includes white and flicker noise (Williams 2008). Stations with 3.5 years of data have a mean uncertainty of 1.0 mm/year in the vertical and 0.3 mm/year in the horizontal components. To achieve a mean uncertainty of 0.5 mm/year, 3.0, 2.5 and 7.5 years of data are needed for the north, east and vertical components, respectively.

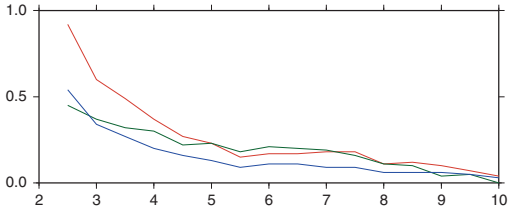


Fig. 2 Stability of the velocity estimates. RMS differences between velocity estimates for a time-series of a given length and the velocity for the complete time-series. The north component is blue, east is green and height is red. The *y* axis is in mm/year and the *x* axis is length of the time-series given as the start time of the time-series in year before 2011

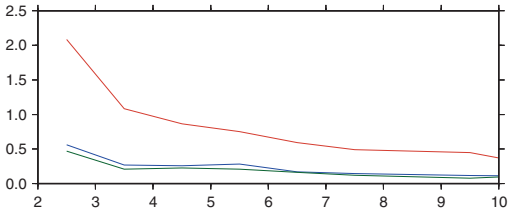


Fig. 3 Uncertainties of the velocity estimates. Mean uncertainties of the velocity estimates are plotted against time-series length. North component is *blue*, east *green* and height *red*. The *y* axis is in mm/year and the *x* axis is length of the time-series given as the start time of the time-series in year before 2011

3.1 Discussion on time-series analysis and time-series length

The accuracy and precision of velocity estimates from geodetic time-series strongly depend on the length of the time-series. According to Blewitt and Lavallée (2002) 2.5 years of data is sufficient to get precise velocity estimates if you account for periodic variations. If you only include secular rate and offset in the regression model then 4.5 years of data are required. Note that Blewitt and Lavallée (2002) implicitly assumed only white noise in the time-series. Bos et al. (2010) performed a similar analysis including power-law and white noise. They showed that the effect of including or not including an annual signal in the time-series analysis, is much larger when a power-law plus white noise model is used instead of a pure white noise model. Teferle et al. (2009) argue that annual and semi-annual signals can bias the velocity

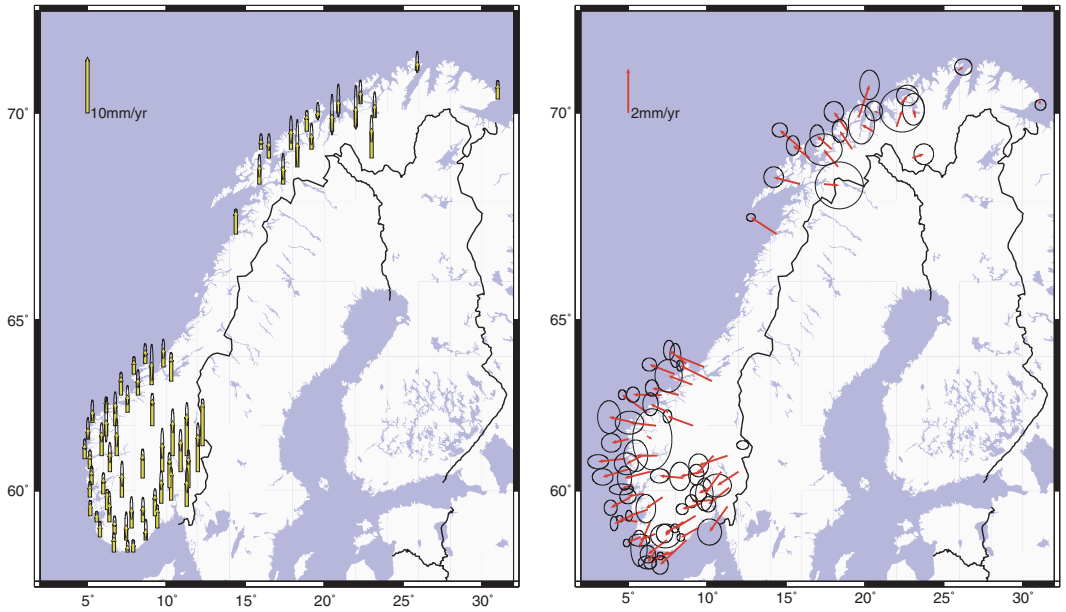


Fig. 4 Land uplift (*left*) and horizontal velocities (*right*) in Norway. The land uplift is in ITRF2008. The horizontal velocities are relative to the stable Eurasian plate (Boucher and Altamimi 2011). The probability ellipses (2σ) are based on a noise model including white and flicker noise

estimates also for time-series exceeding 4.5 years if the signal are large.

In King et al. (2010) theoretical vertical site velocity uncertainties for stations in the Northern Hemisphere of 1.41, 1.58 and 0.54 mm/year after 5 years of observations and 0.71, 0.79, 0.27 mm/year after 10 years of observations are derived using DD, Precise Point Positioning (PPP) and regional stacking, respectively. The uncertainties in the horizontal components are between a third and a fourth in the Double Difference (DD) solution and between half and a third in the Precise Point Positioning (PPP) solution. These theoretical uncertainties agree with the empirical values derived in the time-series analysis and summarized in Fig. 3. See also Santamaria-Gomez et al. (2011) for a study of the time evolution in the noise characteristics of the time-series.

We have opted to include stations with more than three years of data only, and find them sufficient for the precision needed in the analysis described further in this paper. However, we recognize that increasing the time span to five years of data improves the precision considerably, especially in the vertical component.

The velocities off all stations with at least three years of data are plotted in Fig. 4. The horizontal velocities are given relative to the stable Eurasian plate as described in Boucher and Altamimi (2011).

The tests conducted here provide information on the precision and stability of the velocity estimates, but not their absolute accuracies. That is, systematical errors in the reference frame realization, errors in the GNSS analysis strategy or local secular motion of the antenna monument may also affect our velocity solutions (see e.g. King et al. 2010; Wu et al. 2011) and Appendix 7.1 for a more detailed discussion).

4 Establishing a continuous velocity field for Norway

To establish a continuous crustal velocity field in areas where we have (1) no GNSS receivers or (2) the observation period is too short to obtain reliable results, either interpolation or modeling is required. In the first part of this Section we show results from a statistical interpolation method called Kriging. In the second part we present results from a GIA forward model constrained by the GNSS data. In Sect. 4.3 repeated leveling is used as as an additional constraint on the vertical velocities.

The different methods are tested in Sect. 4.4. The observed velocities used here are based on the GNSS time-series of three years or longer, which leaves us with a total of 66 stations.

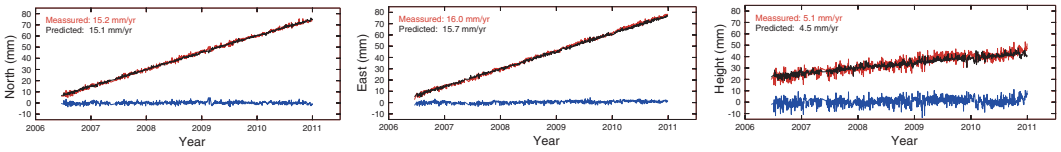


Fig. 5 Observed and predicted time-series for HFSS (Hønefoss). *Red* is observed, *black* is predicted and *blue* is the difference. The estimated rates for the measurements (resp. predictions) are given in *red* (resp. *black*)

4.1 Statistical interpolation

We have used the linear spatial interpolation algorithm, ordinary Kriging (Cressie 1993). The necessary theory and formulas for Kriging are described in Cressie (1993). One of the critical conditions for a successful global interpolation routines are accurate knowledge about the covariance structure. One tool, often used in Kriging, is the variogram or semivariogram. Semivariogram is defined in Cressie (1993).

We will use the interpolation routines in two different ways. First using the daily coordinates for stations in the network to predict daily coordinates for the new point and then using these daily values to estimate the velocities. Second, we will use the velocities already established (Sect. 3) to predict velocities in new points.

4.1.1 Predicting time-series

Daily 3D coordinates from the complete Norwegian GNSS network are used to predict new daily coordinates for a location inside the network, where we have no observations. To find the covariance structure we have used the semi-variogram. Time-series for the predicted location are extracted from these daily predictions. The RMS of the differences between the daily predictions and observations vary from 0.8 to 1.2, 0.8 to 1.4, 2.6 to 3.3 mm in the north, east and vertical component respectively. For HFSS, the corresponding numbers are 0.9, 0.9 and 3.0 mm. Figure 5 shows the predicted time-series for the location Hønefoss (HFSS) in southern Norway as well as the observed time-series for the same location (Note: results from HFSS was excluded before performing the daily predictions). We see a very good agreement between observations and predictions both for the long-term evaluation of the coordinates, but also for short-term fluctuations. This indicates that the method provides a good reconstruction of real observations and hence might be used both for predictions in areas where we have no observations or to extend short time-series into periods without observations. Statistical predictions based on this method are hereafter called SP-TS.

4.1.2 Predicting velocity field

As shown in Fig. 4, there is clearly a case of missing observations: no velocity observations are available for latitudes between 65° and 68°, as well as the locations north-east of Norway.

The missing value problem is handled by a set of advanced procedures with a common purpose: producing plausible values for the missing observations. One of the stronger approaches is the Bayesian one, which simply treats the missing data as extra parameters (Sorensen and Gianola 2002).

Our aim is to construct a continuous velocity field for the entire country by using the available observations. The imputation procedure used to accomplish this goal is an iterative and linear spatial interpolation algorithm, also known as local Kriging. This algorithm is capable of completing the data set by replacing missing observations with predictions. For an example where Kriging have been applied in a similar application see e.g. Teza et al. (2012) where a strain rate map was developed for Northern Victoria Land.

At first, the empirical covariance function is estimated using data solely from southern Norway. The result is then used to generate predictions for some locations above 65° latitude. These are in turn used, alongside the original data-set, to re-compute the empirical covariance function and subsequently make more predictions in the border region. The observation area is expanded, and the process is repeated until the velocity field covers the entire Norwegian mainland, and the missing observations have been reconstructed. The velocity field is shown in Fig. 6, and the accompanying variability is visualized by the variogram in Fig. 7.

The estimation process involves alternating between (1) a step that computes the empirical covariance function using local Kriging, and (2) a step that merges the predictions with the original data, as a complete observation. Hereafter we call this statistical interpolation method for SP-VF.

4.2 Modeling velocities

The GIA model employed is composed of three components: a model of grounded past ice evolution (for Fennoscandia and other ice-covered areas), a sea level model to compute the redistribution of ocean mass for a given ice and Earth model,

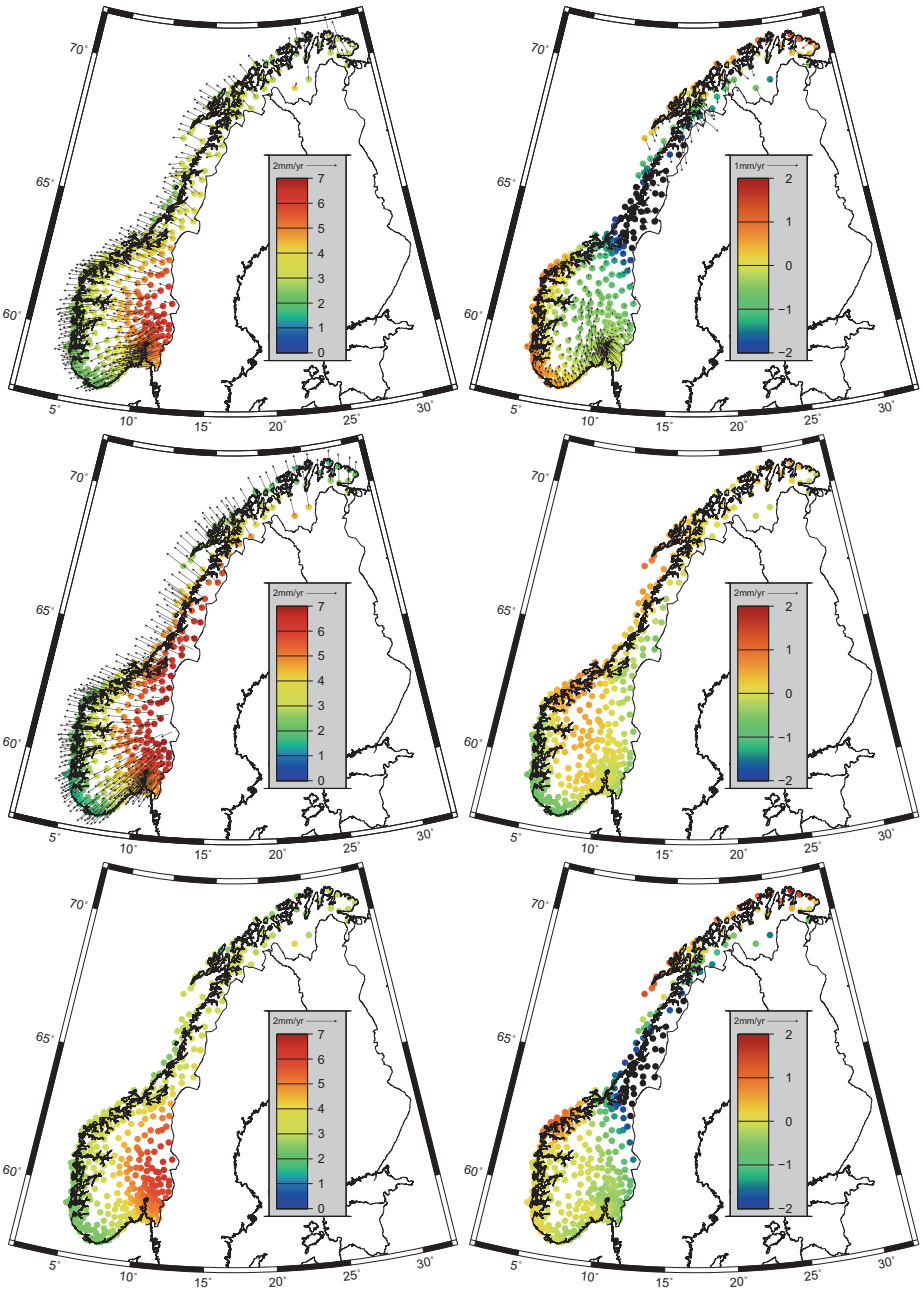


Fig. 6 Velocity fields for Norway. Vertical velocities are given in ITRF2008 while horizontal velocities are transformed to the stable Eurasian plates (ETRF2000). The *left panels* are from *top* to *bottom*; the statistical prediction (SP-VF), the GIA model (GIA-3D) and Leveling (LP). The *right panels* are from top to bottom the differences between; statistical prediction and the GIA model, the GIA model and leveling, and statistical prediction and leveling

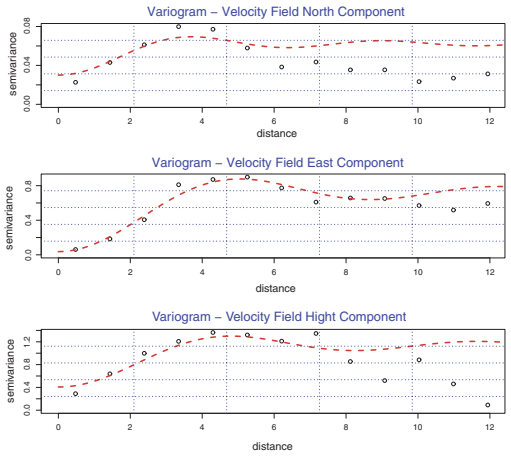


Fig. 7 Variogram for the three components of the velocity field. *Black circles* indicate the empirical semivariogram, while the *dashed red* lines indicate the theoretical models fitted by restricted maximum likelihood. The horizontal axis represents the lag distance in degrees, while the vertical one represents the variability of the velocity field

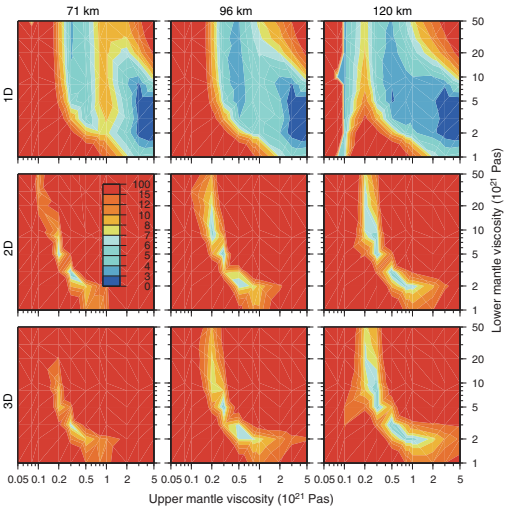


Fig. 8 Goodness of fit for GIA-model, using the reduced χ^2 criterion. The key gives the χ^2 values (see Eq. 2). Top panels show results for the vertical component (1D), middle panels for the horizontal components (2D) and bottom panels for all 3 components (3D). As noted previously, we correct for horizontal plate motions following Boucher and Altamimi (2011)

and an Earth model to compute the solid Earth deformation associated with the ice-ocean loading history. The GIA model used here, and the method used to calculate present-day land motion, is the same as applied by Milne et al. (2001) except that the sea level component of the model was improved as discussed in Mitrovica and Milne (2003) and Kendall et al. (2005).

Past GIA modeling studies have used both paleo sea level data (e.g. Lambeck et al. 1998b) and/or GNSS observations (e.g. Milne et al. 2001, 2004) to help constrain Earth model parameters. These investigations have shown that it is not yet possible to uniquely constrain the Earth viscosity structure for the Fennoscandia region but provide us with a range of Earth parameter values that satisfy the various GIA observables. The main aim of the GIA modeling work performed here, however, is to test how well the model performs in areas where we have no observations (in comparison to the statistical interpolation method), rather than as an investigation of the Earth viscosity structure. This assumes that, after correcting for horizontal plate motion, crustal deformation is solely attributable to the GIA signal.

To perform the test, we constrain the model using a subset of 56 of the total 66 GNSS observations available (the stations marked as red circles in Fig. 1). The 10 GNSS locations not used as a constraint (black circles in Fig. 1) are control stations. We use these to see how well the model reproduces the observed velocities (see following Section).

Given our limited knowledge of the Earth viscosity structure for Norway, we generate predictions of present-day vertical land motion using a suite of 297 Earth viscosity models. The range of values we explore is similar to those as in Milne et al. (2001, 2004), namely; lithospheric thickness is varied from 71 to 120 km, upper mantle viscosity from 0.05×10^{21} to 5×10^{21} Pas and lower mantle viscosity from 10^{21} to $50 \times 10^{21}21$ Pas. To determine an optimal Earth model (i.e. the model which gives best-fit to the GNSS data) we conduct a simple statistical test. We compute vertical and horizontal velocities at the 56 GNSS stations considered for each of the 297 Earth models introduced above and quantify the goodness of fit for each Earth model using the reduced χ^2 criterion:

$$\chi^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i^{\text{pred}} - y_i^{\text{obs}}}{\sigma_i}\right)^2. \quad (2)$$

The χ^2 value indicates the difference between the predictions from the GIA-model y_i^{pred} and the observed vertical velocity y_i^{obs} for a specified observational error σ_i and given GNSS station i (σ_i is the uncertainty of the velocity estimates found assuming a combination of white noise and flicker noise). A value of 1 or less indicates a good fit to the data.

Figure 8 shows how goodness of fit to the GNSS observations varies with Earth model parameters. We find broadly similar results to Milne et al. (2001, 2004), namely, that the vertical velocities favor an Earth model with a relatively stiff upper mantle whereas horizontal rates suggest a weaker one. Differences between χ^2 values for the various lithospheric thicknesses are small. Results from a more comprehensive

investigation, however, suggest a preference for a lithosphere of 100 km or thicker for Fennoscandia (Milne et al. 2004). For the models with a 120 km lithospheric thickness, an upper mantle viscosity of 5×10^{21} Pas and lower mantle viscosity of 3×10^{21} Pas gives best-fit to the GNSS data in the vertical component (hereafter called GIA-1D model). We note that other studies have inferred Earth viscosity values differing to ours and indicate significant lateral variations of Earth structure across Fennoscandia (Steffen and Wu 2011).

In comparison to the vertical component, the GIA model generally shows a poorer fit to the observed horizontal velocities. One reason for this is that the observational errors on the horizontal components are smaller (typically between 0.1 and 0.2 mm/year) which leads to higher χ^2 values. In addition, as the horizontal motions are dominated by a rigid rotation largely driven by plate tectonic processes, isolating the GIA signal is difficult. We correct for the rigid rotation following Boucher and Altamimi (2011). Note that relatively small errors in this correction will affect the determination of the GIA signal and, in turn, the χ^2 values. If we consider the vertical and both horizontal components, the best χ^2-fit is for a model with a 120 km thick lithosphere has an upper mantle viscosity of 5×10^{20} Pas and lower mantle viscosity of 3×10^{21} Pas (hereafter called GIA-3D model).

Isolating the different geophysical signals is unfortunately not a straightforward problem. As mentioned, the rigid rotation we apply here is from Boucher and Altamimi (2011), which has been determined from European GNSS stations. We are aware, however, that this rigid rotation may also contain (or be contaminated by) signals attributable to GIA and already taken into account in our GIA model. Past modeling studies have shown that ongoing GIA in North America (see Mitrovica et al. 1994) and rotational effects associated with GIA (see e.g. Milne et al. 2004) produce a relatively uniform and not insignificant signal of solid Earth motion over Europe.

As it is difficult to distinguish these uniform GIA signals from the horizontal motion dominated by plate tectonics, we conduct a sensitivity test to see how a 1 mm/year error in the plate velocity affects our χ^2 values and RMS of our velocity predictions. The results are plotted in Fig. 9. The grey curve shows the increase in χ^2 relative to the GIA-3D model. We see very little increase in the χ^2 for the best fit model when we add 1 mm/year to the south component of the plate velocity, but a large increase if we add 1 mm/year to the north and east velocities. The red (blue and green resp.) curve is the increase in RMS for the height (north and east resp.) component for a 1 mm/year change in the plate correction in different directions. The RMS values show similar results to the χ^2 values.

Small errors in the plate motion model will influence our χ^2 values and velocity predictions for the control points.

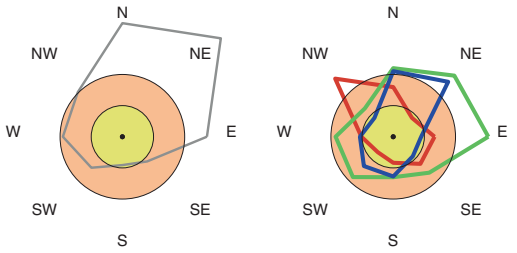


Fig. 9 Sensitivity test.Sensitivity of velocity predictions and χ^2 values when adding 1 mm/year to the plate correction in E, NE, N, NW, W, SW, S, SE and SE. The results are normalized with respect to the GIA-3D model, which is shown as the yellow circle. The orange circle marks the doubled normalized value of the GIA-3D model. The left panel shows the χ^2 values. The right panel shows the RMS of the velocity predictions for the control points. The vertical component is shown in red, north component in blue and east component in green. Results inside the yellow area show an improvement of the χ^2 values or RMS, while results outside the orange area show a doubling of the values

However, results from the sensitivity test (Fig. 9) demonstrate that, as long as the χ^2 stay low also the velocity predictions stay good. The GIA predicted velocities after adding 1 mm/year to the south component of the plate velocity are still good, while 1 mm/year to the north degenerate the GIA model (large χ^2) and make our predictions on the control points much worse.

Our technique for isolating the GIA signal is only one of several methods that can be used. For example, Kierulf et al. (2003) subtracted the GIA signal from the GNSS observations before estimating plate motion. Lidberg et al. (2007) solved for an additional rigid rotation before comparing GNSS results with GIA models. Whereas, in Hill et al. (2010) transformation parameters were included in the observation equation between observations and GIA models to account for possible reference frame problems.

To which extent another approach for isolating the rigid plate rotation and GIA would improve our velocity predictions, are not examined. However, the sensitivity test indicate that it will not improve our ability to predict velocities significantly.

4.3 Repeated leveling as an additional constraint on the vertical velocities

As stated in Sect. 4.1.2, there is clearly a case of missing observations in the mid- and north-eastern part of Norway. Vestøl (2006) shows that repeated leveling lines can be combined with GNSS observations in a common computation of the vertical velocity field. The leveling lines are plotted in Fig. 10.

Following the same procedure as in Vestøl (2006), repeating leveling lines have filled up the open gaps between the

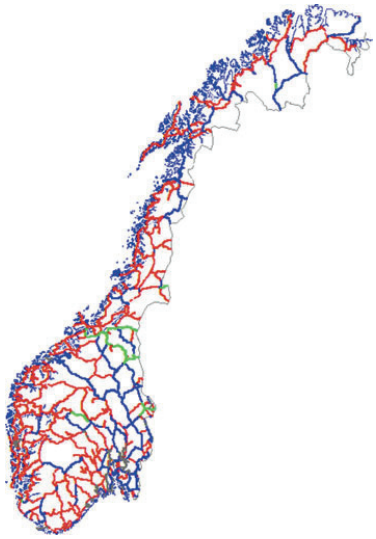


Fig. 10 Repeated leveling *lines* in Norway. The first order leveling network consists of lines from 1916 to 2011. The *green lines* are measured three times, the *blue two* and the *red* lines are measured once only. However, also the *red lines* from different years contain land uplift information when forming a loop. *Lines* from Sweden and Finland are also used in the computation, but are not shown here

Global Navigation Satellite System (GNSS) stations (the 56 stations marked as red circles in Fig. 1) and brought higher redundancy into the system. The predictions based on this method are hereafter called LP.

4.4 Evaluation of predicted and modeled velocities

To evaluate the precision of the statistical predictions we have compared our predictions with observed velocities for a subset of the stations (the 10 stations marked with black circles in Fig. 1). The statistical predictions were performed using the remaining 56 stations as data base. RMS values of the differences between observations and predictions are included in Table 1. The two statistical methods give almost identical results; SP-TS method is slightly better in the north component, while the SP-VF method is preferable in the east and vertical components. The vertical results are hampered by a few outliers. All the outliers are at the perimeter of the network (stations that geometrically can not be surrounded by any triangle of nearby stations). The stations are Trysil (TRYS) in the east, Andøya (ANDO) in north-west and Kristiansand (KRSS) in south (green triangles in Fig. 1). All three stations have a residual (difference between observed and predicted values) of above 1 mm/year in the height component for at least one of the methods used. The RMS after removing these stations is included in the parentheses in Table 1. We note large improvements for both methods, but the SP-VF method gives better agreement in the height component. Not surprisingly, we can conclude that such interpolation methods are more uncertain at the perimeter of the network.

Table 1 Uncertainties of predictions

	OBS.-SP-TS (mm/year)	OBS.-SP-VF (mm/year)	OBS.-GIA-1D (mm/year)	OBS.-GIA-3D (mm/year)	OBS.-LP (mm/year)
North	0.25 (0.24)	0.34 (0.32)	–	0.29 (0.29)	–
East	0.44 (0.30)	0.20 (0.18)	–	0.20 (0.21)	–
Height	0.78 (0.53)	0.62 (0.27)	0.55 (0.54)	0.86 (0.97)	0.72 (0.70)

The values are the RMS between observations and predictions for all sites established in 2000 or earlier (*black circles* in Fig. 1). For the values in parentheses the outliers at the perimeter of the network (*green triangles* in Fig. 1) are removed

It is also two other factors worth to remember. To perform these comparisons the stations with longest time-series are used as control sites. This implies that the stations with the presumable highest precision are not included in the base of the prediction. Furthermore are almost all the oldest stations in Norway located at the coast and therefore at the edge of the GNSS network, making the prediction of these particular sites more uncertain. The RMS values given in Table 1 may therefore be regarded as upper bound for what could be expected for such types of interpolation. Using also the control stations as base for the interpolation would have improved the results from the statistical predictions. Predicted velocities in center of the network would presumable have better precision.

This comparison indicates that Kriging based on already established velocities (SP-VF) is slightly better than Kriging based on daily coordinates (SP-TS). The method SP-TS is also by far much more time consuming; the Kriging procedure has to be repeated for the whole network for each single day. On the other hand the method SP-TS could be regarded as more robust in the way that, if a few days of interpolation fail it would easily be detected and removed in the final time-series analysis. Method SP-TS is also able to predict non-linear features for instance expected yearly fluctuations for a site.

The modeled velocities (GIA-3D) are compared with real observations in the same way as we did for the statistical interpolated velocities. The RMS values are included in Table 1. In the horizontal components GIA-3D and SP-VF perform equally, but in the vertical component the statistical interpolated values fit better to the observations. Horizontal motions are dominated by plate tectonic processes, which make isolating the GIA signal difficult. With GIA-1D we have constrained our GIA model only using the vertical velocities. The GIA-1D model has better fit to vertical observations than the

SP-VF. If we exclude the stations on the perimeter of the network (green triangles in Fig. 1), however, the SP-VF are closer to observations. To summarize, interpolation based on geophysical models gives better estimates on the edge of the network, but inside the network the statistical interpolation method is preferable.

To evaluate the interpolation in the vertical component using leveling (LP), we use the same 10 control stations (black circles in Fig. 1). The LP results perform at a level between the two results using Kriging (SP-VF and SP-TS). However, the LP results are hampered by the discrepancy in Trondheim (TRDS) (misfit of 1.3 mm/year). Removing Trondheim from the solution yields a RMS of 0.55 mm/year, similar to the GIA-1D. If we remove the three perimeter stations in addition the RMS is 0.45 mm/year. This is better than achieved with GIA-1D, but not at the level of SP-VF. In the Trondheim area the leveling shows a clear increase in the land uplift from west to east, a trend not so clear if we compare the GNSS time series in Trondheim with the neighbouring stations.

The GIA model indicates that the largest gradient in the uplift values is in mid-Norway. Unfortunately, this is also the area where we lack velocity estimates from the GNSS observations (the records here are currently too short to obtain reliable velocities). Figure 6 shows broad agreement between the leveling and GIA solutions in mid-Norway. Whereas, the Kriging solution shows large differences between both the leveling and GIA solutions in this area. For areas where we not yet have reliable velocity estimates from the GNSS observations, therefore, measurements from leveling can provide a useful additional constraint on velocity field solutions based on statistical methods.

5 Discussion

In Sect. 3 the agreement between part of the time-series and the complete time-series was evaluated for 10 different stations (black circles in Fig. 1). After 3.5 years of data the RMS was 0.3, 0.3 and 0.5 mm/year for the north, east and vertical components, respectively. These RMS values are at the same level as we achieved with SP-VF and GIA-3D in the horizontal components and slightly better in the vertical component. If we remove the outliers at the edge of the network (green triangles in Fig. 1), the precision of SP-VF in the vertical component is similar to what we achieved with 4.5 years time-series length compared to the complete time-series. To summarize, depending on the geometry of the network, statistical predictions and geophysical models give better precision than time-series with less than approximately 3.5 years of observations. For stations with longer time-series the observed velocities are preferable.

We have made two underlying assumption of the interpolation; (1) the deformations in Norway have a spatial wavelength longer than the distance between stations and (2) GIA and rigid plate tectonics are the dominant source of crustal movement in Norway.

Other geophysical processes that might introduce secular crustal deformation in Norway are neotectonics and loading from glaciers and large water reservoirs. All large water reservoirs in Norway were established before the Norwegian GPS network and will not introduce significant secular elastic deformations for the GNSS sites.

We do not expect that present-day glacier changes in Norway will have a large affect on the GNSS velocities as their mass changes are relatively small and such elastic deformations are confined to areas close to the mass changes (e.g. Khan et al. 2007; Kierulf et al. 2009b).

Although Norway is situated on the stable Eurasian plate and far from the plate margin, the seismological and neotectonical activities are relatively large. Geological evidence, seismological measurements as well as InSAR and GPS data indicate that the Ranafjord area (66.2° north) is the most tectonically active area in Norway (see Olesen et al. 2012). A GPS campaign network in the area has been occupied twice with 9 years separation. Results indicate a relative deformation of the network of 1 mm/year horizontally (Olesen et al. 2012). This is above the precision level found in this paper. Local deformation exceeding the 0.5 mm/year can not be excluded in other neotectonically active areas of Norway.

Comparing the Kriging and GIA-model solutions reveals an interesting pattern of differences in the horizontal components (Fig. 6 upper right panel). This likely reflects errors in the GIA model solution and/or the presence of non-GIA related signals. Indeed, examining deviations from the GIA model solution may help identify non-GIA effects, this is something that could be explored in future investigations.

6 Conclusion

In this paper we have analyzed data from the permanent GNSS network in Norway using the GNSS analysis package GAMIT. The results are presented as time-series and velocity estimates are calculated. The precision and accuracy of these velocity estimates are examined with respect to time-series length.

All tests show a decline of the results if you have less than three years of data, especially in the vertical component. Results improve gradually when you extend your time-series length.

In the second part of this paper we have looked at several methods to predict velocities in areas where we do not have permanent GNSS receivers or the observation period is too short (less than three years) to calculate good

Table 2 Uplift for some Norwegian stations

	K2012 ITRF2008 mm/year	L2010 ITRF2005 mm/year	L2007 ITRF2000 mm/year
OSLS	5.11 ± 0.33	6.51 ± 0.47 (-0.15 ± 0.28)	5.78 ± 0.42 (0.99 ± 0.53)
STAS	1.46 ± 0.17	2.90 ± 0.42 (-0.15 ± 0.28)	1.18 ± 0.51 (0.98 ± 0.53)
TRDS	4.31 ± 0.35	6.19 ± 0.51 (-0.13 ± 0.28)	3.80 ± 0.58 (1.05 ± 0.53)
TRO1	2.90 ± 0.33	4.61 ± 0.83 (-0.10 ± 0.28)	2.30 ± 0.49 (1.13 ± 0.53)
VARS	2.67 ± 0.28	5.74 ± 0.86 (-0.07 ± 0.28)	1.89 ± 1.13 (1.14 ± 0.53)

K2012 is from this study, L2010 is from Lidberg et al. (2010) and L2007 is from Lidberg et al. (2007). Numbers in parenthesis is the vertical component of the transformation parameters between the reference frames and should be added to the uplift value to transform them to ITRF2008

velocity estimates. Two different statistical interpolation methods based on Kriging theory are performed as well as an interpolation method using a geophysical GIA-model and a method using repeated leveling. The models are comparable, but the velocities based on geophysical models are more robust on the perimeter or outside the GNSS network, while the statistical method give better results inside the network. Our results indicate that velocity estimates based on the different interpolation methods are better than that estimated from a single GNSS station which has less than 3.5 years of data.

Acknowledgments We would like to acknowledge the SATREF group at the Norwegian Mapping Authority for there effort to establish the permanent GNSS network in Norway necessary for this study and especially Oddgeir Kristiansen for his work to collect and maintain an archive of Norwegian GNSS stations. We would like to thank IGS for providing necessary geodetic infrastructure and geodetic products. The authors also thank the reviewers for constructive feedback that helped to improved the original manuscript considerably.

7 Appendix

7.1 Earlier results and reference frame

In Table 2 we present vertical velocities for stations which are included in this study and also included in previous studies (Lidberg et al. 2007, 2010). We find differences between the different studies. Taking the uncertainties and transformation into account the differences between this study and Lidberg et al. (2007) are relatively small, but the uplift in Lidberg et al. (2010) seems a bit too large.

Differences between the different ITRFs have been discussed extensively in several papers (e.g., Argus 2007; Teferle et al. 2009; Lidberg et al. 2010; Altamimi et al. 2011). In Norway the differences between ITRF2000 and ITRF2008, based on the transformation parameters (Altamimi et al. 2007, 2011), are approximately 1, 0 and 1 mm/year, in the north-, east- and height-component, respectively. The formal uncertainties in the transformation parameters between different reference frames do not necessarily reflect the uncertainty of the reference frame relative to geophysical processes. Wu et al. (2011) find that ITRF2008 is consistent with the earth mean center of mass at the 0.2 mm/year level. In Collilieux and Wöppelmann (2011) the ITRFs are extensively discussed in the context of global sea-level.

Differences between results of geodetic studies represent an issue for the correct understanding of geophysical processes (see e.g. King et al. 2010). To obtain better constraint on the GNSS results we can use independent observations such as other geometric techniques like Very Long Baseline Interferometry (VLBI) and Satellite Laser Ranging (SLR) (e.g., Altamimi et al. 2011) non-geometric techniques like gravimetry (e.g., Teferle et al. 2009; Omang and Kierulf 2011) or geophysical evidence (e.g., Argus 2007; Kierulf et al. 2009b). Individual components of the Global Geodetic Observing System (GGOS) (e.g. Rummel et al. 2005) have to be maintained and improved.

References

Altamimi Z, Collilieux X, Legrand J, Garayt B, Boucher C (2007) ITRF2005: a new release of the international terrestrial reference frame based on time series of station positions and earth orientation parameters. J Geophys Res 112:b09401 OSE doi:10.1029/2007JB004949

Altamimi Z, Collilieux X, Métivier L (2011) ITRF2008: an improved solution of the international terrestrial reference frame. J Geodesy 85:457–473. doi:10.1007/s00190-011-0444-4

Argus FA (2007) Defining the translation velocity of the reference frame of earth. Geophys J Int 169(3):830–838. doi:10.1111/j.1365-246X.2007.03344.x

Blewitt G, Lavallée D (2002) Effect of annual signals on geodetic velocities. J Geophys Res 107(2145):11. doi:10.1029/2001JB000570

Boehm J, Werl B, Schuh H (2006) Troposphere mapping functions for GPS and very long baseline interferometry from European

Centre for medium-range weather forecasts operational analysis data. J Geophys Res 111:B02–406 doi:10.1029/2005JB003629

Bos M, Bastos L, Fernandes R (2010) The influence of seasonal signals on the estimation of the tectonic motion in short continuous GPS time-series. J Geodynamics 49:205–209. doi:10.1016/j.jog.2009.10.005

Boucher C, Altamimi Z (2011) Memo: specifications for reference frame fixing in the analysis of a EUREF GPS campaign, v8. http://etrs89.ensg.ign.fr/

Collilieux X, Wöppelmann G (2011) Global sea-level rise and its relation to the terrestrial reference frame. J Geodesy 85:9–22. doi:10.1007/s00190-010-0412-4

Cressie NAC (1993) Statistics For Spatial Data. Wiley Series in Probability and Mathematical Statistics, isbn: 978-0-471-00255-0.

Fjeldskaar W (1994) Viscosity and thickness of the asthenosphere detected from the Fennoscandian uplift. Earth Planet Sci Lett 126:399–410

Herring T, King R, McClusky S (2011) Introduction to GAMIT/GLOBK release 10.4. Tech. rep. Massachusetts Institute of Technology, Cambridge.

Hill EM, Davis JL, Tamisiea ME, Lidberg M (2010) Combination of geodetic observations and models for glacial isostatic adjustment fields in Fennoscandia. J Geophys Res 115:B07403. doi:10.1029/2009JB006967

Johansson J, Davis J, Scherneck HG, Milne G, Vermeer M, Mitrovica J, Bennett R, Jonsson B, Elgered G,Elosegui P, Koivula H, Poutanen M, Ronnang B, Shapiro I (2002) Continous GPS measurements of postglacial adjustment in Fennoscandia 1. Geodetic result. J Geophys Res 107(B8):2157. doi:10.1029/2001JB000400

Johnson HO, Agnew DC (1995) Monument motion and measurements of crustal velocities. Geophys Res Lett 22(21):2905–2908. doi:10.1029/95Gl02661

Kendall RA, Mitrovica JX, Milne GA (2005) On postglacial sea level: II. numerical formulation and comparative results on spherically symmetric models. Geophys J Int 161(3):679–706. doi:10.1111/j.1365-246X.2005.02553.x

Khan SA, Wahr J, Stearns LA, Hamilton GS, van Dam T, Larson KM, Francis O (2007) Elastic uplift in southeast Greenland due to rapid ice mass loss. Geophys Res Lett 34:L21701. doi:10.1029/2007GL031468

Kierulf HP, Plag HP, Kristiansen O (2003) Towards the true rotation of a rigid Eurasia. In: Torres JA, Hornik H (eds) EUREF Publication No. 12, Verlag des Bundesamtes für Kartographie und Geodäsie, Frankfurt am, Main, pp 118–124.

Kierulf HP, Pettersen B, McMillan D, Willis P (2009) The kinematics of Ny-Ålesund from space geodetic data. J Geodynamics 48:37–46. doi:10.1016/j.jog.2009.05.002

Kierulf HP, Plag HP, Kohler J (2009) Measuring Surface deformation induced by present-day ice melting in Svalbard. Geophys J Int 179(1):1–13. doi:10.1111/j.1365-246X.2009.04322.x

King M, Altamimi Z, Boehm J, Bos M, Dach R, Elosegui P, Fund F, Hernndez-Pajares M, Lavallèe D, Mendes Cerveira P, Penna N, Riva R, Steigenberger P, van Dam T, Vittuari L, Williams S, Willis P. (2010)Improved constraints on models of glacial isostatic adjustment: a review of the contribution of ground-based geodetic observations. Surv Geophys 31:465–507. doi:10.1007/s10712-010-9100-4

Lambeck K, Smither C, Ekman M (1998a) Tests of glacial rebound models for Fennoscandinavia based on instrumented sea- and lake-leve records. Geophys J Int 135:375–387

Lambeck K, Smither C, Johnston P (1998b) Sea-level change, glacial rebound and mantle viscosity for northern Europe. Geophys J Int 134:102–144

Legrand J, Bergeot N, Bruyninx C, Wöppelmann G, Bouin MN, Altamimi Z (2010) Impact of regional reference frame definition on geodynamic interpretations. J Geodynamics 49 (3–4):116–122.doi:10.1016/j.jog.2009.10.002, WEGENER 2008 - Proceedings of the 14th General Assembly of Wegener

Lidberg M, Johansson J, Scherneck HG, Davis J (2007) An improved and extended GPS derived velocity field for the glacial isostatic adjustment in Fennoscandia. J Geodesy 81(3):213–230. doi:10.1007/s00190-006-0102-4

Lidberg M Johansson JM, Scherneck HG, Milne GA (2010) Recent results based on continuous GPS observations of the GIA process in Fennoscandia from BIFROST. J Geodynamics 50:1

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

Milne G, Davis J, Mitrovica J, Scherneck HG, Johansson J, Vermeer M, Koivula H (2001) Constraints on glacial isostatic adjustments in Fennoscandia. Science 291:2381–2385

Milne GA, Mitrovica JX, Scherneck HG, Davis JL, Johansson JM (2004) Continuous GPS measurements of postglacial adjustment in Fennoscandia: 2. modeling results. J Geophys Res 109:B02412. doi:10.1029/2003JB002619

Mitrovica JX, Milne GA (2003) On postglacial sea level: I. general theory. Geophys J Int 154(2):253–267. doi:10.1046/j.1365-246X

Mitrovica JX, Davis JL (1994) A spectral formalism for computing three-dimensional deformations due to surface load, 2. present-day glacial isostatic adjustment. J Geophys Res 99:7075–7101

Nocquet JM, Calais E, Parsons B (2005) Geodetic constraints on glacial isostatic adjustment In Europe. Geophys Res Lett 32:L06308. doi:10.1029/2004Gl022174.l022174

Olesen O, Brönner M, Dehls J, Lindholm C, Bungum H, Kierulf H, Bockmann L (2012) Neotectonics and strandflat formation in Nordland, northern Norway. Geophys Res Abstr 14:EGU2012–13080

Omang OCD, Kierulf HP (2011) Past and present-day ice mass variation on Svalbard revealed by superconducting gravimeter and GPS measurements. Geophys Res Lett 38:L22304. doi:10.1029/2011GL049266

Rummel R, Rothacher M, Beutler G (2005) Integrated global geodetic observing system (IGGOS) - science rationale. J Geodynamics 40 (4–5):357–362

Santamaria-Gomez A, Bouin MN, Collilieux X, Wöppelmann G (2011) Correlated errors in GPS position time series: implications for velocity estimates. J Geophys Res 116:B01405. doi:10.1029/2010JB007701

Scherneck HG (1991) A parametrized solid earth tide model and ocean tide loading effects for global geodetic baseline measurements. Geophys J Int 106(3):677–694. doi:10.1111/j.1365-246X.1991.tb06339.x

Scherneck HG, Johansson J, Elgered G, Davis J, Jonsson B, Hedling G, Koivula H, Ollikainen M, Poutanen M, Vermeer M, Mitrovica J, Milne GA (2002) BIFROST: Observing the three-dimensional deformation of Fennoscandia. In: Mitrovica X, Vermeersen BLA (eds) Ice Sheets Sea Level and the Dynamic Earth. American Geophysical Union, Washington, pp 69–93

Sorensen D, Gianola G (2002) Likelihood, Bayesian, and MCMC methods in quantitative genetics. Springer, Berlin

Steffen H, Wu P (2011) Glacial isostatic adjustment in Fennoscandia: a review of data and modeling. J Geodyn 52(3–4):160–2004. doi:10.1016/j.jog.2011.03.002

Teferle FN, Bingley RM, Orliac EJ, Williams SDP, Woodworth PL, McLaughlin D, Baker TF, Shennan I, Milne GA, Bradley SL, Hansen DN (2009) Crustal motions in Great Britain: evidence from continuous GPS, absolute gravity and Holocene sea level data. Geophys J Int 178:23–46. doi:10.1111/j.1365-246X.2009.04185.x

Teza G, Pesci A, Casula G (2012) Strain rate computation in northern Victoria land (Antarctica) from episodic GPS surveys. Geophys J Int 189:851–862. doi:10.1111/j.1365-246X.2012.05403.x

Vespe F, Rutigliano P, Ferraro C, Nardi A (2002) Vertical Reference Systems, IAG Symposia, vol 124. In: Drewis H, Dodson AH,

Fortes LPS, Sanchez L, Sandoval P (eds) Vertical motions from geodetic and geological data: a critical discussion of the results. Springer, Berlin Heidelberg, pp 66–71

Vestøl O (2006) Determination of postglacial land uplift in Fennoscandia from leveling, tide-gauges and continuous GPS stations using least squares collocation. J Geodesy 80:248–258. doi:10.1007/s00190-006-0063-7

Williams SDP (2008) CATS: GPS coordinate time series analysis software. GPS Solut 12(2):147–153. doi:110.1007/s10291-10007-10086-10294

Williams SDP, Bock Y, Fang P, Jamason P, Nikolaidis RM, Prawirodirdjo L, Miller M, Johnson DJ (2004) Error analysis of continuous GPS position time series. J Geophys Res 109:1–19

Wu X, Collilieux X, Altamimi Z, Vermeersen B, Gross R, Fukumori I (2011) Accuracy of the international terrestrial reference frame origin and earth expansion. Geophys Res Lett 38:L13304. doi:10.1029/2011GL047450

Zhang J, Bock Y, Johnson H, Fang P, Williams S, Genrich J, Wdowinski S, Behr J (1997) Southern California permanent GPS geodetic array: error analysis of daily position estimates and site velocities. J Geophys Res 102(B8):18035–18055. doi:10.1029/97JB01380

Paper #2

Reference

M. Ouassou, A.B.O. Jensen, J.G.O. Gjevestad & O. Kristiansen
Next generation network real-time kinematic interpolation segment to improve the user accuracy
International Journal of Navigation and Observation, 346498 (2015)

Contribution

M. Ouassou conceived the idea, performed the numerical calculations, and wrote the majority of the manuscript. All authors contributed to the discussion of the results and revisions of the manuscript.

A.B.O. Jensen involved in validation process and writing the majority of section 1.

J.G.O. Gjevestad and O. Kristiansen involved in discussion, support and conclusion.

J.A. Ouassou also acknowledged for useful discussions, feedback, and proofreading.

Hindawi Publishing Corporation
International Journal of Navigation and Observation
Volume 2015, Article ID 346498, 15 pages
http://dx.doi.org/10.1155/2015/346498

Research Article

Next Generation Network Real-Time Kinematic Interpolation Segment to Improve the User Accuracy

Mohammed Ouassou,[1,2] Anna B. O. Jensen,[3] Jon G. O. Gjevestad,[2,4] and Oddgeir Kristiansen[1,2]

[1] *Norwegian Mapping Authority, Geodetic Institute, 3511 Hønefoss, Norway*
[2] *Department of Mathematical Sciences and Technology, NMBU, 1432 Akershus, Norway*
[3] *KTH Royal Institute of Technology, 100 44 Stockholm, Sweden*
[4] *Department of Electronics and Telecommunications, NTNU, 7491 Trondheim, Norway*

Correspondence should be addressed to Mohammed Ouassou; mohammed.ouassou@statkart.no

Received 19 September 2014; Revised 7 December 2014; Accepted 25 December 2014

Academic Editor: Aleksandar Dogandzic

This paper demonstrates that automatic selection of the right interpolation/smoothing method in a GNSS-based network real-time kinematic (NRTK) interpolation segment can improve the accuracy of the rover position estimates and also the processing time in the NRTK processing center. The methods discussed and investigated are inverse distance weighting (IDW); bilinear and bicubic spline interpolation; kriging interpolation; thin-plate splines; and numerical approximation methods for spatial processes. The methods are implemented and tested using GNSS data from reference stations in the Norwegian network RTK service called CPOS. Data sets with an average baseline between reference stations of 60–70 km were selected. 12 prediction locations were used to analyze the performance of the interpolation methods by computing and comparing different measures of the goodness of fit such as the root mean square error (RMSE), mean square error, and mean absolute error, and also the computation time was compared. Results of the tests show that ordinary kriging with the Matérn covariance function clearly provides the best results. The thin-plate spline provides the second best results of the methods selected and with the test data used.

1. Introduction

The use of GNSS and network real-time kinematic positioning to achieve GNSS positions with accuracy at the cm-level is increasing rapidly these years. This is partly due to the development and modernization of the GNSS systems themselves (GPS, GLONASS, Galileo, and Beidou), but it is also caused by a general quest for better position accuracy in many user communities.

High-accuracy GNSS positioning is based on the carrier phase being observable. Using the notation from [1], the basic observation equation that summarizes the relation between observations and error sources is given as follows:

$$\Phi = \lambda^{-1}(r - I + T) + f(\delta t_u - \delta t^s) + N + \varepsilon_\phi, \quad (1)$$

where Φ is the phase observation in cycles, λ is the wavelength in meters/cycle, r is the geometric distance between the receiver and satellite in meters, I is the ionospheric signal delay in meters, T is the tropospheric signal delay in meters, f is the frequency in Hertz, δt_u and δt^s are the clock errors of, respectively, the receiver and the satellite, N is the initial number of cycles at the first observation epoch (the *ambiguity*), and ε_ϕ is a noise term given in cycles that mainly accounts for multipath (reflected signals) and receiver noise.

When using the NRTK technique, a network of reference stations is used to estimate the errors in the positioning process, that is, the effects of the ionosphere and troposphere as well as inaccuracies in the satellite position as provided with the broadcast ephemerids from the satellites.

The accuracy of NRTK positioning systems depends on the ability to identify and mitigate the error sources in

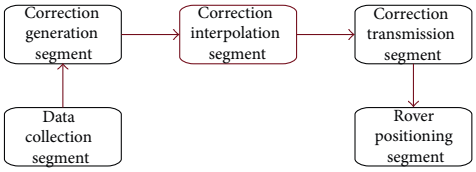


Figure 1: Network real-time kinematic segments.

the system as well as the residual biases. The biases include residual effects from the space segment, signal propagation, environment effects, and receiver noise in the reference network. The mitigation process can be carried out by modeling, estimation, or combinations of observables.

The NRTK processing chain can be summarized as follows: the first step is to collect raw measurements from the network of reference stations, solve for the ambiguities within the reference network, and generate error estimates. The next step is to apply the interpolation/smoothing scheme to generate the RTK corrections for the user location. The RTK corrections are then transmitted to users who can then perform real-time positioning with accuracy at the cm-level.

Figure 1 shows all segments involved in the NRTK processing chain. The figure illustrates the so-called virtual reference station (VRS) concept, which was developed by Landau et al. [2]. Other NRTK standards such as for instance the master auxiliary concept (MAC) also exist [3], but we limit the discussion in this paper to the VRS concept.

As the GNSS systems and users become more numerous, the amount of data that needs processing increases as well, which poses some interesting challenges for the NRTK system developers and service providers. This paper focuses on processing large data sets and high quality interpolators/smoothers that can be used to aid the data processing. Let us consider how the RTK processing is carried out. First the user sends his/her position to the control center, and then the network engine chooses a suitable subnetwork which is used to generate corrections, and these corrections are then transmitted back to the user. The first challenge to this model is the number of users, since each user has to be processed independently, and the number of users has increased dramatically in recent years. The solution to this is to construct new models and algorithms. These should be able to process data from large geographical areas, as well as computing the necessary corrections and quality indicators ready for use, so that any RTK user that connects will be served immediately.

In other branches of science and engineering, new analysis tools that satisfy these requirements have already been developed: neural networks, machine learning, classification and regression trees, hierarchical models, and so forth. In this paper, some existing interpolation/smoothing methods are applied to real datasets, and the strengths and weaknesses of each method are identified. The results are then used to combine these methods and construct models that describe the observed variations in the data as well as possible.

Interpolation methods can be divided into two categories: local methods and global methods. The local methods only use a subset of the data for interpolation, which implies that the required processing time is reduced. Conversely, the global techniques use all the data available to generate predictions. In this paper, both these approaches are considered. Referring to Figure 1, the main focus of this paper is directed at the correction interpolation segment and more specifically at the automatic selection of the right interpolation algorithm based on appropriate tests, such that the rover position estimation will be improved.

The rest of the paper is organized as follows: Section 2 gives a full description of the test data using the Norwegian GNSS network data, known as CPOS, and introduces the variational problem in general. Section 3 covers local interpolation algorithms, specifically the inverse-distance weighted and bilinear/bicubic methods by Akima. Section 4 and the following sections deal with global interpolation methods. First, thin-plate splines and the Bayesian model behind the smoothing are reviewed in this section. Section 5 introduces numerical approximation schemes for Gaussian random fields. Section 6 covers spatial interpolation algorithms and specifically the ordinary kriging method. Section 7: the performance parameters are defined in this section. Section 8: the results from Sections 3–6 are generalized in this section. Section 9 is the conclusion and discussion and covers applications of the results developed in Sections 3–6.

2. Test Data and Variational Reconstruction

2.1. Test Data. The main success of network real-time kinematic positioning has been the reduction of correlated errors in the network (e.g., ionospheric, tropospheric, and satellite position errors). This type of errors is collectively referred to as *distance-dependent errors* and can be subdivided into the *dispersive errors* which depend on frequency and the *nondispersive errors* which do not.

The size of the network varies with time, as the individual reference stations and satellites may not deliver data for a while, and observations are typically correlated to each other. Modeling the spatial and temporal variations of such a process is too complex to capture the covariance structure of the data, so often we end up imposing stationarity. In this paper, we apply techniques for handling spatial processes in order to capture the covariance structure in the data, such that high quality synthetic data can be provided. The main clue is to employ the right tool from epoch to epoch, based on some appropriate criteria.

We prefer to work with real data, and since the real network error estimates were not made available, we decided to analyze the ionospheric path delays for CPOS RTK network, given by absolute TEC values. If the ionosphere data is replaced with the full network corrections, the same algorithms should still function very well. Ionospheric path delay is considered the single largest source of inaccuracy for positioning and navigation, so the quality of the NRTK corrections is strongly affected in the case of moderate to high ionosphere activity. To test the algorithms against each other, a large ionospheric data set from the Norwegian CPOS network is investigated. The data is generated by a first-order geometry-free approach (Section 8.1).

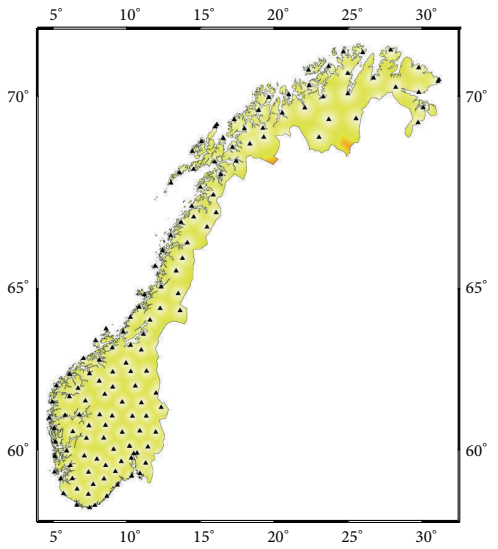


Figure 2: Norwegian CPOS RTK network. Filled triangles mark station locations. Height difference between sites up to 1000 meters.

At the time of writing, the CPOS RTK network contains approximately 180 stations on the Norwegian mainland (for a listing of stations of the CPOS RTK network, see the appendix.). The algorithms were tested with different station configurations (50, 75, and 110 stations), equipped with geodetic dual frequency GNSS receivers, which track both GPS and GLONASS satellites. In this investigation, however, only the GPS portion of the data was used. The distribution of the CPOS RTK network reference stations is given in Figure 2.

2.2. Variational Reconstruction

2.2.1. Problem Formulation. Let us assume that the observed noisy measurements $\{z_i\}$ at the locations $\{s_i : s_i \in \mathbb{R}^2\}$ is a random function, with the mean $m(s)$ and variance σ^2. Our goal is then to predict the value at other locations $\{s_k\}$ where we have no observations, under the assumption that the predicted values should resemble its neighbors. To achieve this, we can either interpolate or construct a smooth function $g(s)$ that represents the variation in the data and is robust against outliers.

The data that will be modeled is a pure spatiotemporal process, namely, the absolute total electron count (TEC). Assuming weak stationarity of the process under study, the mean and variance are not functions of the spatial location s. The model used to describe the variation in data in this paper, however, is assumed to have the form

$$z(s) = m(s) + \varepsilon(s). \tag{2}$$

The mean function $m(s)$, often referred to as the *trend* or *deterministic part*, determines the large-scale variation in the data. The function $\varepsilon(s)$ is called the *random part* and determines the small-scale variation. This process model will be assumed in the subsequent discussion of all the different interpolation/smoothing techniques presented in this paper.

Some data analysts prefer the *Cressie decomposition* [4, ch. 3] of the observed random field $Z(s)$, which takes the form

$$Z(s) = m(s) + T(s) + \varepsilon(s), \tag{3}$$

where $s \in \mathbb{R}^2$ is the spatial location; $Z(s)$ is the observation; $m(s) = X\beta$ is the trend (the mean component of the model); $T(s)$ is a stationary Gaussian process with variance σ^2 (partial sill), and a correlation function parameterized in its simplest form by ϕ (the range parameter); and finally ε is an error term, with a variance parameter τ^2 (nugget variance).

2.2.2. Model Parameter Estimations. Once the model is defined, the next step is to estimate the model parameters. In general, this is done numerically by minimizing the negative log-likelihood function. The most used optimization methods are, respectively, the conjugate gradient method, the quasi-Newtonian method, and the Nedler-Mead method. The details of these methods will not be treated in this paper, but the interested reader is referred to references [5, 6].

The algorithm may not converge to correct parameter values when called with the default options. The user should therefore try different initial values, and if the parameters have different orders of magnitude, a scaling of the parameters may be necessary. If such problems arise, some possible workarounds include

(i) rescaling data values by dividing by a constant,

(ii) rescaling coordinates by subtracting values and/or dividing by constants,

(iii) bootstrapping to accelerate the convergence. This method is used in our implementation of the kriging algorithm in Section 6.

2.2.3. Model Validation. In the field of statistical analysis, an appropriate way of analyzing data is to divide it into three distinct subsets. The training dataset is used to construct the model, the validation data is used to check the model, and the last data set is used to challenge the model. The main purpose is to determine whether or not our model is an accurate representation of the real world data. This process is called the *model validation assessment*. The most famous methods are the family of cross-validation, generalized maximum likelihood (GML) methods, Akaike information criterion (AIC), Bayesian information criterion (BIC), and so forth.

In our implementation, the generalized cross-validation is used to determine the optimal smoothing parameter λ (see Section 4). The computation AIC and BIC are computed in Section 6, when maximum likelihood estimation is used instead of weighted least squares in the kriging algorithm. The GML methods will be used in future work.

3. IDW and Akima Interpolation

3.1. IDW Interpolation. The inverse-distance weighted (IDW) scheme is an exact interpolator. It honors the data by assigning weights to all neighboring points according to their distance from the prediction location s_0. Locations that are closer to s_0 receive higher weights, and locations that are far from s_0 are given lower weights; this mechanism is administered by the parameter r in the IDW predictor formula. The user can freely choose the number of observations used to perform the interpolation. This is done by defining a radius R around the prediction location s_0.

The IDW predictor scheme is defined as

$$\widehat{z}(s_0) = \frac{\sum_{i=1}^{N} z(s_i)\, d_{0i}^{-r}}{\sum_{i=1}^{N} d_{0i}^{-r}}. \tag{4}$$

Here, s_0 is the prediction location, N is the number of observations, $z(s_i)$ are the neighboring observations, r is the weight decision parameter, and d is the distance (either spherical or Euclidean).

The IDW method is originally due to Shepard [7], which described a global method. All derived IDW methods are either generalizations or variations of this method. The basic Shepard's method can be expressed as

$$F(x, y) = \frac{\sum_{k=1}^{N} w_k(x, y)\, f_k}{\sum_{k=1}^{N} w_k(x, y)}, \tag{5}$$

where typically the weight w_k is the inverse Euclidean distance $w_k = 1/d_k^2 = \{(x - x_k)^2 + (y - y_k)^2\}^{-1/2}$. We will however define a disk with center (x_k, y_k) and a radius R and set the weight w_k to zero outside of this disk. A natural scheme suggested by many authors, like, for example, Renka and Brown [8], is given by the expression

$$w_k(x, y) = \left\{ \frac{(R_w - d_k)_+}{R_w d_k} \right\}^2, \tag{6}$$

where

$$(R_w - d_k)_+ = \begin{cases} R_w - d_k & \text{if } d_k < R_w \\ 0 & \text{if } d_k \geq R_w. \end{cases} \tag{7}$$

Impose the constraints such that

(i) the sum of all weights w_k inside the disk R_w should be normalized to unity, that is, $\sum w_k = 1$,

(ii) the predictor is a linear combination of the observations.

If the variance of the predictor is then controlled such that it is at a minimum, the IDW behaves almost like the local kriging interpolator (Section 6); however the covariance structure is not preserved.

For the implementation, the package *gstat* from Edzer Pebesma is used to carry out IDW (see Table 3 for more information).

3.2. Akima Algorithms. Bilinear or bicubic spline interpolation is applied using different versions of algorithms by Akima [9, 10]. Given a set of data points in a plane, our aim is to fit a smooth curve that passes through the given points considered as reference points. The method is local and is based on a piecewise function composed of a set of polynomials and applicable up to the third degree on each interval. The method produces remarkable results with minimum processing time. For a detailed mathematical formulation, please refer to references [9, 10].

3.2.1. Basics of Spline Interpolation/Smoothing. An understanding of the basic elementary building blocks of a 1D spline facilitates the understanding of 2D and 3D splines, for instance, the TPS (Section 4).

Given real numbers $\{t_1, t_2, \ldots, t_n\} \in [a, b]$, a function g defined on interval $[a, b]$ is a cubic spline if and only if the following conditions are satisfied:

(i) the function g is a cubic polynomial on each interval $(a, t_1), (t_1, t_2), \ldots, (t_n, b)$;

(ii) the function g and its first and second derivatives are continuous at each of the points t_i.

Condition (ii) implies that the cubic polynomials from condition (i) fit together on each t_i, where the t_i are called *knots*. Together these two conditions imply that $g(x)$ is a function with continuous first and second derivatives on the whole interval $[a, b]$.

For some given real constants a_i, b_i, c_i, d_i, the cubic spline function g can be expressed as

$$g(t) = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i(t - t_i)^3, \tag{8}$$

where the index $i = 0, 1, 2, \ldots, n$. The end-point knots correspond to the boundaries of the function domain; that is, $t_0 = a$ and $t_{n+1} = b$.

Finding a smoothing spline is not an easy task. Reinsch (1967) proposed an algorithm and showed that the solution of the minimum principle is actually cubic splines. The basic idea is to construct a nonsingular system of linear equations of the second derivative λ_i of $\widehat{g}$. The resulting equations are computationally efficient because of their banded structure. For an excellent exposition of the material, see also [11].

3.2.2. Output Result from Akima. Figure 3 shows the output from Akima with bilinear interpolation.

4. Thin Plate Spline Method

4.1. Mathematical Preliminaries. In this section, our main interest is not to construct a function $m(s_i)$ that exactly interpolates the data z_i at n distinct points but to find an attractive way to smooth noisy data. The method of thin-plate splines (TPS) will be used for this purpose.

Duchon [12] was the first to build the theoretical foundation for the TPS method. The name TPS comes from the physical situation of bending a thin surface, where the method minimizes the bending energy of a thin plate fixed

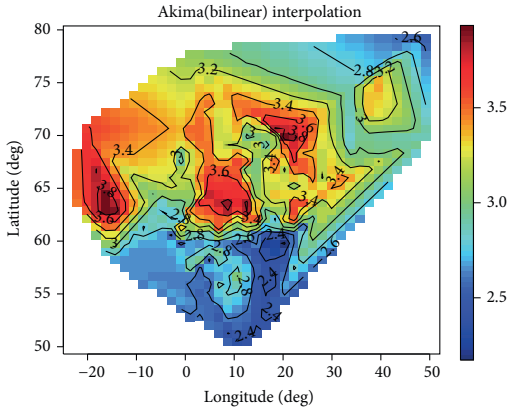


Figure 3: L1-VTEC variation surface generated by AKIMA (bilinear) interpolation. Color scale is expressed in meters. High ionospheric activity (year: 2013, doy: 152), 110 reference stations.

at the data sites. For our application, the TPS method is used to minimize the cost function:

$$J_m(f) = \int_{\mathbb{R}^d} \left\| D^m f(s) \right\|^2 \mathrm{d}s, \tag{9}$$

where D^m is a vector of partial differentiation operators of order m. In the two-dimensional case, that is, when $d = 2$, $m = 2$, and $f = f(x, y)$, the TPS penalty function can be written as

$$J_2(f) = \iint_{\mathbb{R}^2} \left\{ \left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right\} \mathrm{d}x\, \mathrm{d}y. \tag{10}$$

Let $\Delta^4(x, y)$ denote the differential operator in the integrand of (10). The thin-plate spline $f(x, y)$, which is the solution to the variational problem of minimizing the penalty $J_2(f)$, can then be found by solving the biharmonic equation $\Delta^4 f(x, y) = 0$.

The goal is to find the function f_λ in Sobolev space [13, p. 250] that minimizes the following expression:

$$S(f) = n^{-1} \sum_{i=1}^{n} \left(z(s_i) - f(s_i) \right)^2 + \lambda J_2(f), \tag{11}$$

where n is the total number of observations, J_2 is a smoothness penalty (the cost function), and λ is the smoothing parameter. The smoothing parameter is a trade-off factor between the rate of change of the residual error and local variation. Optimal minimization of $S(f)$ results in a good compromise between smoothness and goodness of fit.

Once the curve approximation of the data has been constructed, generating values at any location, where no observations are available, is accomplished by simply indexing the variables x and y and fetching the corresponding value. This is a major advantage of smoothing methods over interpolation methods; no extra interpolations are required after the curve has been constructed for a given epoch.

Validation is carried out by the GCV (Section 2.2.3). Let the cross-validation function $\hat{g}_{n,\lambda}^{[k]}$, with $k = 1, 2, \ldots, n$, be defined as

$$\hat{g}_{n,\lambda}^{[k]} = \arg\min_{\tilde{g} \in W_2^{(m)}} \sum_{i=1}^{n} \left(z_i - \tilde{g}(x_i) \right)^2 + \lambda \int_0^1 \left\{ \tilde{g}^{(m)}(x_i) \right\}^2 \mathrm{d}x, \tag{12}$$

where $\lambda > 0$. The OCV (ordinary cross-validation) and OCV MSE (ordinary cross-validation mean square error) $V_0(\lambda)$, respectively, are defined as

$$\begin{aligned} \mathrm{OCV}(\lambda) &= \arg\min_{\lambda \in \mathbb{R}^+} V_0(\lambda), \\ V_0(\lambda) &= \frac{1}{n} \sum_{i=1}^{n} \left\{ \hat{g}_{n,\lambda}^{[k]}(x_i) - z_i \right\}^2. \end{aligned} \tag{13}$$

The determination of the GCV (general cross-validation) goes as follows. First, the expression for V_0 has to be rewritten. There exists an $n \times n$ matrix $A(\lambda)$, the smoothing/influence/sensitivity matrix with the property. Consider

$$\begin{bmatrix} \hat{g}_{n,\lambda}(x_1) \\ \vdots \\ \hat{g}_{n,\lambda}(x_n) \end{bmatrix} = A(\lambda) Z, \tag{14}$$

such that $V_0(\lambda)$ can be written as

$$V_0(\lambda) = n^{-1} \sum_{j=1}^{n} \frac{\left(a_{kj} z_j - z_k \right)^2}{\left(1 - a_{kk} \right)^2}, \tag{15}$$

where $k, j \in \{1, 2, \ldots, n\}$ and a_{kj} is element $\{k, j\}$ of $A(\lambda)$.

Definition 1 (generalized cross-validation (GCV)). Let $A(\lambda)$ be the smoothing matrix defined in (14); then the GCV function is given by the expression

$$\mathrm{GCV}(\lambda) = \frac{n^{-1} \left\| (I - A(\lambda)) z \right\|^2}{\left\{ n^{-1} \operatorname{tr} (I - A(\lambda)) \right\}^2}. \tag{16}$$

4.2. Estimation of the Smoothing Parameter λ. The smoothing parameter λ plays a central role in the TPS method. By adjusting the value of λ, one can get the desired level of smoothness at the cost of accuracy at the data sites. When we set this parameter to zero, the problem is reduced to an interpolation with no smoothing. On the other hand, when the smoothing parameter tends to infinity, the method yields a plane which is least-square fitted to the data. The smoothness penalty method can be chosen by any criteria, but the most popular criterion is GCV (generalized cross-validation), also known as the "left-out one" method. The GCV criterion selects the smoothing parameter λ that minimizes the GCV function, equation (16), that is $\hat{\lambda} = \arg\min_{\lambda \in \mathbb{R}^+} \mathrm{GCV}(\lambda)$.

The GCV function $V(\lambda)$ is the predicted mean square error and can be viewed as a weighted version of the $\mathrm{OCV}(\lambda) = V_0(\lambda)$:

$$V(\lambda) = \frac{1}{n}\sum_{i=1}^{n}\left\{\widehat{g}_{n,\lambda}^{[i]}(x_i) - z_i\right\}^2 W_k(\lambda),$$
$$W_k(\lambda) = \left\{\frac{I - a_{kk}(\lambda)}{n^{-1}\operatorname{tr}(I - A(\lambda))}\right\}^2. \tag{17}$$

In geodesy, it is often interesting to estimate the accuracy f_λ. Two loss functions are considered: the mean square prediction error $T(\lambda)$, and the stricter Sobolev error is defined as $W(\lambda) = \|f_\lambda - f\|_W^2$,

$$W(\lambda) = \int_a^b \{f_\lambda - f\}^2 \,\mathrm{d}G + \int_a^b \left\{f_\lambda^{(m)} - f^{(m)}\right\}^2 \mathrm{d}G. \tag{18}$$

The performance of an estimator is often well characterized by the risk function, defined as the expectation value of the loss function:

$$R_T(\lambda) = \mathbb{E}(T(\lambda)), \qquad R_W(\lambda) = \mathbb{E}(W(\lambda)). \tag{19}$$

In this analysis, the GCV is used to estimate the smoothing parameter λ. Figure 12 shows the smoothed surface generated by the TPS with GCV.

For implementation, the CRAN package *rgcvpack* is used to implement the TPS algorithm (see Table 3 for more information).

5. Numerical Approximation Methods

Numerical approximation techniques will assist us in processing huge data sets with convergence. The main idea is based on the pioneering work of Besag [14].

Let us assume that our observations at different locations follow a multivariate Gaussian distribution with mean μ_z and variance-covariance Σ_z. Then the continuously Gaussian fields have the distribution

$$f(z) = \{2\pi\}^{-n/2}\left|\Sigma_z\right|^{-1}\exp\left\{-\frac{1}{2}(z-\mu_z)^T\Sigma_z^{-1}(z-\mu_z)\right\}. \tag{20}$$

Approximating the continuous Gaussian random field by the discrete Gauss-Markov random field is accomplished by introducing the Markov property. This is done as follows: we say that two locations s_i and s_j are conditionally independent if and only if

$$x_i \perp x_j \mid x_{-(i,j)}. \tag{21}$$

This property is very important when constructing the precision matrix Q of the GMRF. That is, if we know what happens nearby, we can ignore everything that lies further away. Consider

$$x_i \perp x_j \mid x_{-(i,j)} \Longrightarrow Q_{i,j} = 0. \tag{22}$$

That is, element (i, j) of Q is zero if the process at location i is conditionally independent of a process at j given the process at all locations except $\{i, j\}$. Figure 4 illustrates the concept of the GMRF.

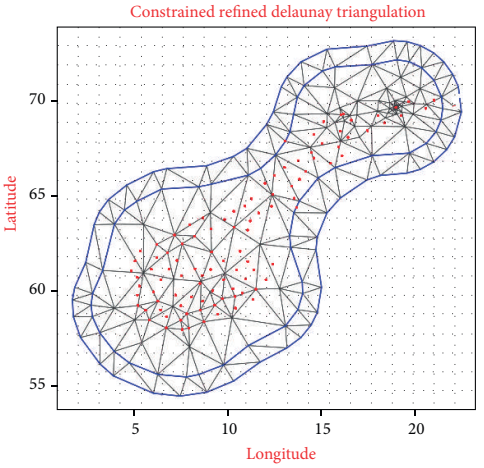


Figure 4: Mesh construction on top of which the GMRF is built. Red circles represent the configured reference stations. The mesh in this figure is used to construct the neighborhoods which are key to specifying GMRF conditional independence structure.

The sparse precision matrix makes the GMRF computationally effective, but it is difficult to construct reasonable precision matrices. As a conclusion, the GMRF is a Gaussian field with a sparse precision matrix $Q = \Sigma^{-1}$. For an excellent description of the theory and applications of GMRF, the reader is referred to, for example, Rue and Held [15].

The integrated nested Laplace approximation (INLA) method developed by Håvard Rue is used to implement the GMRF (see Table 3 for more information).

6. Kriging Interpolator

The kriging interpolator is a linear spatial interpolation algorithm and is primarily used in geostatistics. In recent years, however, the interpolator has been applied in many new areas, such as geophysics and climate data analysis.

Given the observations $\{Z(s_i)\}_{i=1}^N$, we want to predict the value of $Z(s_0)$ where no observations have been made. Our goal is to find an estimator $\widehat{Z}_0 = \widehat{Z}(s_0) = \sum_{i=1}^N w_i Z(s_i)$ such that the following requirements are met.

(i) Unbiasedness. This means that $\mathbb{E}(Z(s_0)) = \mathbb{E}(Z_0)$ and is accomplished if $\sum_{i=1}^N w_i = 1$ and the mean is stationary.

(ii) Minimum Prediction Variance. We make some assumptions about the mean value of the random field $Z(s)$. If the mean is unknown but constant across the entire region of interest, we have *ordinary kriging*. Otherwise, the method is known as *simple kriging*.

Any estimator that meets the conditions of unbiasedness and minimum prediction variance is said to be a BLUP

(best linear unbiasedness predictor). Let us examine the components of the MSPE (mean square prediction error). Consider

$$\begin{aligned}\sigma_{S_0}^2 &= \operatorname{var}\left(Z_0 - \widehat{Z}_0\right)\\ &= \operatorname{var}\left(Z_0\right) + \operatorname{var}\left(\widehat{Z}_0\right) - 2\operatorname{cov}\left(Z_0, \widehat{Z}_0\right)\\ &= \sigma^2 + \operatorname{var}\left(\sum_{i=1}^{N} w_i Z_i\right) - 2\operatorname{cov}\left(Z_0, \sum_{i=1}^{N} w_i Z_i\right)\\ &= \sigma^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \operatorname{cov}\left(Z_i, Z_j\right) - 2\sum_{i=1}^{N} w_i \operatorname{cov}\left(Z_i, Z_0\right)\\ &= \sigma^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j C_{ij} - 2\sum_{i=1}^{N} w_i C_{i0}.\end{aligned} \tag{23}$$

We want to minimize $\operatorname{var}[Z(s_0) - \widehat{Z}(s_0)]$ subject to the constraint $\sum_{i=1}^{N} w_i = 1$.

The procedure is well defined by the method of Lagrange multipliers. Form the Lagrangian L,

$$L = \operatorname{var}\left(Z\left(s_0\right) - \widehat{Z}\left(s_0\right)\right)^2 - 2\lambda\sum_{i=1}^{N}\left(w_i - 1\right). \tag{24}$$

We then take the partial derivatives of L with respect to the weights w_i and to λ, set the equations to be equal to zero, and solve them; we get

$$\begin{bmatrix} w_1 \\ \vdots \\ w_N \\ \lambda \end{bmatrix} = \begin{bmatrix} C_{11} & \cdots & C_{1N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C_{1N} & \cdots & C_{NN} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} C_{10} \\ \vdots \\ C_{N0} \\ 1 \end{bmatrix}. \tag{25}$$

Equation (25), which is the kriging equation, is used to compute the weights. The computation of weights is based on the covariances among locations in the sample (region of interest) and the covariances between sample locations and the location to be predicted. To be specific.

(1) Covariances among the locations in the sample:

$$C_{ij} = C\left(s_i, s_j\right) = \operatorname{Cov}\left(Z_i, Z_j\right) \quad \forall i, j. \tag{26}$$

The covariance matrix of the N sample values read

$$\mathbf{C} = \begin{bmatrix} C_{11} & \cdots & C_{1N} \\ \vdots & \ddots & \vdots \\ C_{1N} & \cdots & C_{NN} \end{bmatrix}. \tag{27}$$

(2) Covariances between the sample locations and the prediction point:

$$C_{i0} = C\left(s_i, s_0\right) = \operatorname{Cov}\left(Z\left(s_i\right), Z\left(s_0\right)\right) \quad \forall i. \tag{28}$$

The vector of covariances between the sample locations and the prediction point read

$$\mathbf{c} = \begin{bmatrix} C_{10} & C_{20} & \cdots & C_{N0} \end{bmatrix}^{\mathrm{T}}. \tag{29}$$

Equation (25) becomes

$$\begin{bmatrix} \mathbf{w} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{C} & \mathbf{1} \\ \mathbf{1}^{\mathrm{T}} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{c} \\ 1 \end{bmatrix}, \tag{30}$$

where $\mathbf{w}$ is $1 \times N$ vector of weights and $\mathbf{1} = [1 \cdots 1]$ is a vector of the same dimensions.

6.1. Directional Effects. Another form of nonstationarity lies in the covariance structure. One specific way to relax the stationarity assumption is to allow directional effects. For instance, the correlation decay rate at increasing distances may be allowed to depend on the relative orientation between pairs of locations. The simplest form of directional effects in the covariance structure is called *geometrical anisotropy*. This arises when a stationary covariance structure is transformed by a differential stretching and rotation of the coordinate axes. Hence, geometrical anisotropy is defined by two additional parameters. Algebraically, a model with geometrical anisotropy in spatial coordinates $x = (x_1, x_2)$ can be converted to a stationary model in coordinates $y = (y_1, y_2)$ by the transformation

$$\begin{pmatrix} y_1 & y_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} \cos\left(\psi_A\right) & -\sin\left(\psi_A\right) \\ \sin\left(\psi_A\right) & \cos\left(\psi_A\right) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \psi_R^{-1} \end{pmatrix}. \tag{31}$$

ψ_A is called the *anisotropy angle* and $\psi_R > 1$ the *anisotropy ratio*. The direction with the slowest correlation decay is called the *principal axis*.

6.2. Choice of Covariance Function. The spatial correlation between measurements at different locations is described by the semivariogram functions:

$$\begin{aligned}\gamma(h) &= \frac{1}{2}\operatorname{var}\left(Z\left(s_i\right) - Z\left(s_j\right)\right)\\ &= \frac{1}{2}\left\{\operatorname{var}\left(Z\left(s_i\right)\right) + \operatorname{var}\left(Z\left(s_j\right)\right)\right.\\ &\qquad \left. -2\operatorname{cov}\left(Z\left(s_i\right), Z\left(s_j\right)\right)\right\}\\ &= C(0) + C(h),\end{aligned} \tag{32}$$

where $C(0)$ is the variance and $C(h)$ is the covariance. The variogram and the covariance contain the same information and can be used interchangeably.

In this study, the spatial correlation function $C(h)$ is defined by the Matérn function and is given by

$$C(h) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}\left(\kappa\|h\|\right)^{\nu} K_{\nu}\left(\kappa\|h\|\right). \tag{33}$$

$h = \|s_i - s_j\| \in \mathbb{R}^+$ is the Euclidean spatial distance between locations. K_{ν} is the modified Bessel function of

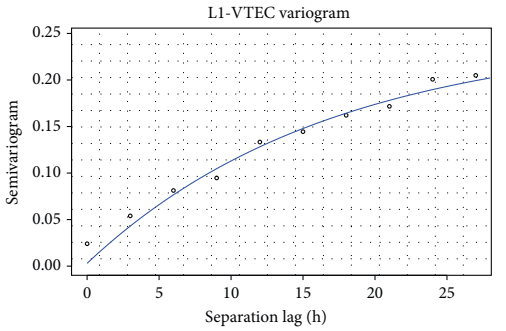


Figure 5: An example of the Matérn covariance function, using data with moderate ionosphere activity. Empirical variogram (small circles) and Matérn model fitted by maximum likelihood (solid line). The parameter values used were: smoothness $\kappa = 0.5$, partial sill $\sigma^2 = 0.185$, and range parameter $\phi = 4.05$.

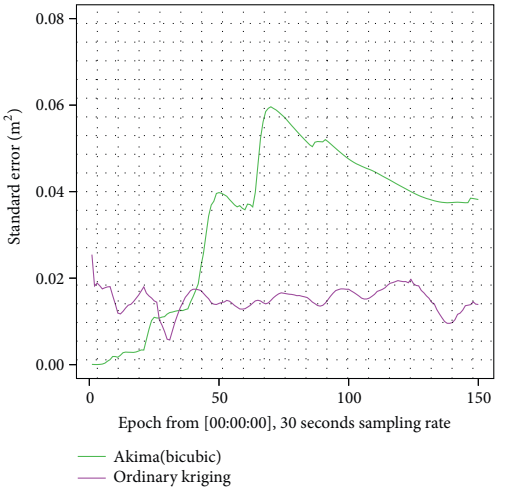


Figure 6: The ordinary kriging algorithm (BLUP, pink curve) has a stable and minimum variance compared to the AKIMA algorithm (green curve). Normal ionospheric activity (year 2013, doy: 153), 110 reference stations.

the second kind [16], the order $\nu > 0$ measures the degree of the smoothness of the process, and κ is the scaling parameter related to the distance of decorrelation (dependency becomes almost 0).

$C(h)$ is obtained from spectral densities [17, p. 31] of the form

$$f(\omega) = \phi\left(\alpha^2 + \omega^2\right)^{-(\nu+1/2)} \qquad \nu > 0,\ \phi > 0,\ \alpha > 0. \quad (34)$$

Figure 5 shows the empirical semivariogram $\gamma(h)$ with the Matérn covariance function, which fits the L1-VTEC data well. It also works in a wide range of circumstances; including low, moderate, and high ionospheric activities, tested with several different reference station configurations, more specifically 75, 100, and 115 stations.

6.3. Computation of the Inverse Matrix. The kriging equation (25) requires the inverse of the covariance matrix to be computed, and this is detrimental to the performance of the algorithm for large data sets. The operation may occasionally even fail to invert the matrix. Numerical methods with optimization algorithms will help us avoid this, for instance, factorization methods, ill-conditioned test, and other suitable methods.

7. Performance Parameters

In order to carry out the performance analysis of each individual algorithm, an averaging weighted reference signal X was constructed. It is defined as a linear combination of values generated by algorithms with different weights, that is, $X = \sum_{i=1}^{5} w_i \mathrm{Alg}_i$, under the normalization constraint $\sum_{i=1}^{5} w_i = 1$. Five algorithms are involved to construct the reference signal X.

The weights are chosen according to algorithm performance measured in terms of minimum value and stability of variance, functionality, correctness, and processing time. Figure 6 shows the variance of two algorithms. We see that ordinary kriging has a minimum and stable variance; therefore its weight is higher than for the Akima bicubic spline. Table 1 summarizes the weight assignment for the algorithms.

Table 1: Definition of algorithm weights.

Algorithm type	Weights	Comments
OK	$w_1 = .25$	Produces trustworthy results with MCF
TPS	$w_2 = .25$	A real competitor of kriging, always delivers, even when the covariance structure is poor
GMRF	$w_3 = .25$	Trustworthy results, like OK and TPS
Akima	$w_4 = .15$	Handles small variations perfectly. The weight is reduced compared to OK, GMRF, and TPS
IDW	$w_5 = .10$	No covariance structure is preserved. The assigned weight is reduced compared to OK, GMRF, TPS, and Akima

7.1. Quality of Service Parameters Definitions. For each one of the quality of service (QoS) parameters whose values are negotiable, the worst case performance must be specified. In some cases, the minimum or the maximum values are preferable, in other cases the averaged value.

The criteria chosen for performance evaluation in this paper are based on comparing the reference signal X to the

Table 2: Quality of service (QoS) parameter definitions.

Parameters	Explanations
MAE	Mean absolute error
MSE	Mean square error
RMSE	Root mean square error
NSE	Nash-Sutcliffe efficiency
KGE	Kling-Gupta efficiency
R	Pearson correlation coefficient
β	Changes in locations (mean)
α	Changes in scale (variance)
CT	Computation time is set to 5 seconds. In NRTK processing the values most used are: ionosphere: 10 seconds geometrical: 15–20 seconds

output Y_k from algorithm k. Analysis is based on statistical monitoring and detecting the changes in spatial location, scale, and level. The full list is given in Table 2.

The required length of time series before we can carry out the goodness of fit is a critical parameter. With the data sets used for testing, values in the range of 100–200 epochs were acceptable.

All algorithms compete about the QoS, the one with highest score is selected as the winner, and the corrections from this algorithm are used. 12 locations (can be regarded as VRS) are chosen inside the CPOS RTK network for testing, and one location is chosen randomly for each run to compute the QoS. The mathematical definitions of the QoS parameters are given in Table 2.

(i) Mean Absolute Error (MAE). MAE measures the average absolute error and is defined below. Ideally, this value should be as small as possible. Consider

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|X_i - Y_i\right|. \tag{35}$$

(ii) Mean Square Error (MSE). This measures the average squared error and is defined below. This value should also be as close to zero as possible. Consider

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - Y_i\right)^2. \tag{36}$$

(iii) Root Mean Square Error (RMSE). RMSE between reference signals X_i and Y_i gives the standard deviation of the algorithm prediction error and minimum value is preferable. Consider

$$\mathrm{RMSE} = \left(\frac{1}{N}\sum_{i=1}^{N}\left(X_i - Y_i\right)^2\right)^{1/2}. \tag{37}$$

(iv) Nash-Sutcliffe Efficiency (NSE) [18]. NSE determines the relative magnitude of the noise variance compared to the observed data variance. Unity means a perfect match, and zero means that algorithm predictions are as accurate as the mean of the observed information, while negative values imply that the observed mean is better than the predicted one. Consider

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(X_i - Y_i\right)^2}{\sum_{i=1}^{N}\left(X_i - \overline{X}_i\right)^2}. \tag{38}$$

(v) Kling-Gupta Efficiency (KGE) [19]. KGE was developed by Gupta et al. as a goodness of fit that decomposes the NSE to facilitate the analysis of correlation, bias, and variability. Consider

$$\mathrm{KGE} = 1 - \left\{(R-1)^2(\alpha-1)^2(\beta-1)^2\right\}^{1/2}. \tag{39}$$

Three components are involved in computation of this index.

(a) R is the Pearson product moment correlation coefficient, which ideally should tend to unity. This quantity is defined by the expression

$$R = \frac{N\left(\sum X_i Y_i\right) - \left(\sum X_i\right)\left(\sum Y_i\right)}{\left\{\left[N\sum\left(X_i\right)^2 - \left(\sum X_i\right)^2\right]\left[N\sum\left(Y_i\right)^2 - \left(\sum Y_i\right)^2\right]\right\}^{1/2}}. \tag{40}$$

(b) β represents the change in locations. This index is defined as the ratio between distribution locations (means), and the ideal value is unity. Consider

$$\beta = \frac{\mu_O}{\mu_C}. \tag{41}$$

(c) Variability ratio (VR) represents changes in scale (variances). This index is defined as the ratio between distribution standard deviations, and the ideal value is again unity. Consider

$$\mathrm{VR} = \frac{\sigma_C/\mu_C}{\sigma_O/\mu_O}. \tag{42}$$

(vi) Computation Time (CT). Algorithms with high quality data and minimum CT are preferable.

(vii) Coefficient of Determination (R2). $0 \leqslant R2 \leqslant 1$ and gives the portion of the variance of one variable that is predictable from the other variable.

(viii) Spearman Correlation Coefficient (ρ). $-1 \leqslant \rho \leqslant 1$ is a nonparametric test used to measure the degree of associations between two variables.

8. Implementation and Analyses

Packages used in the implementation are downloaded from the Comprehensive R Archive Network (CRAN). Table 3 gives a full description of each package.

Table 3: Implementation packages.

Algorithm	Package name	Download	Comments
Kriging	geoR	CRAN	Prediction by OK (i) AIC computation (ii) BIC computation (iii) MLE estimation (iv) WLS estimation (v) Anisotropy
GMRF	INLA	CRAN	Prediction by GMRF
TPS	rgcvpack	CRAN	Prediction by TPS (i) GCV to estimate λ
AKIMA	akima	CRAN	Bilinear interpolation Bicubic interpolation
IDW	gstat	CRAN	IDW interpolation
Normality	Moments	CRAN	Jarque-Bera normal test
Smoothing	KernSmooth	CRAN	Wand Kernel smoothing

8.1. Data Preprocessing. The data used in this investigation is ionospheric data obtained from the Norwegian CPOS RTK network. The TEC values are generated using the GFA (geometry-free approach) algorithm. The algorithm takes two steps to process.

(a) From IGS [20], the global ionospheric model (GIM) is available. The ionospheric path delay and differential code biases (DCBs) for each satellite are retrieved from the IONEX (IONosphere map EXchange) format. This information is used to estimate the hardware biases of the reference stations, by using the code observable.

(b) From the previous step, we then use the biases in phase measurement and compute the ionospheric delay.

The procedure is described in more detail in [21].

8.2. Interpretation of Results. The test results show that ordinary kriging with the Matérn covariance function is the most appropriate choice under normal circumstances and produces a smooth solution with acceptable accuracy. The Matérn covariance function is well-behaved even for nonstationary fields and is governed by three parameters: location, scale, and shape.

Stein [17, p. 12] recommended the use of the Matérn model due to its flexibility (ability to model the smoothness of physical processes and possibility to handle nonstationarity).

The processing time of this algorithm increases as the number of observations increase. Another approach is to exclude observations that are far away from the interpolation points and use only a subset of the data for the interpolation. This approach is called *local kriging*.

In order to increase the convergence of OK, we have incorporated the bootstrapping algorithm on historical data to get a very good guess for initial values. Figure 9 illustrates the concept.

One major obstacle of this algorithm is the computation of the inverse matrix in the kriging equation (25). Using the numerical approximation of the inverse matrix, the computation time will improve considerably, as mentioned previously in Section 6.3.

The WLSE (weighted least square estimation) algorithm is preferable to maximum likelihood or restricted maximum likelihood and works in most cases, regardless of the distribution of the observations. If the observations have a Gaussian distribution, the WLS and ML/REML yield the same results.

We are often faced with a nonstationary process where we are interested in estimating the spatial covariance for the entire random field. Guttorp and Sampson [22] proposed a two-step approach for solving this problem, a nonparametric algorithm to estimate the spatial covariance structure for the entire random field without assuming stationarity. The interested reader is referred to [23, pp. 93–95].

When the covariance structure preserves sparsity, numerical approximation methods are preferable to all other methods, as they require less memory and computation time.

TPS algorithm is preferred when performing smoothing rather than interpolating data.

8.3. Algorithms Delay Comparison. In this subsection, the delays caused by different algorithms are investigated. The test results are shown in Figure 7. The number of observations varied between 700 and 3000 ionospheric piercing points (IPPs). The data used for the plot are the first 255 epochs with a resolution of 30 seconds from high ionospheric activities, from the year 2013 and day of year 152.

The delay caused by local methods is shown on the right of Figure 7 and is much lower compared to global methods.

The GMRF has the highest delay over the OK and the TPS. The only challenge of TPS is to select a good smoothing parameter, λ. The modified cross-validation, the generalized cross-validation, and robust GCV all work well.

The IDW methods are local interpolation techniques and use only a subset of the data set to perform the interpolation. The benefit of these methods is the reduced computation time.

8.4. QoS Results. Statisticians are usually more interested in smoothing data than interpolating it. When the data is noisy, the TPS smoothing scheme works best. One major advantage of this approach is that once the curve that represents the variation in the data is constructed, we can retrieve the value at any other location without reinterpolating the data set.

Figure 8 shows the result for an arbitrary prediction point with coordinates $(\text{lon}, \text{lat}) = (5.0, 57.0)$. The reference signal X is compared to the predicted values generated by the ordinary kriging algorithm with MCF. The computed quality of service parameters (QoS) are presented below the plot.

The results are summarized in Table 4 where the QoS parameters are provided for each of the interpolation algorithms tested. An arbitrary epoch has been picked for the test. High scores are highlighted in bold font. The result shows that the ordinary kriging has the best performance. The TPS comes in second place and is the only real competitor to the ordinary kriging for this case. As kriging has the best performance, the corrections from this algorithm will be used to generate synthetic observations of the user in the field. This

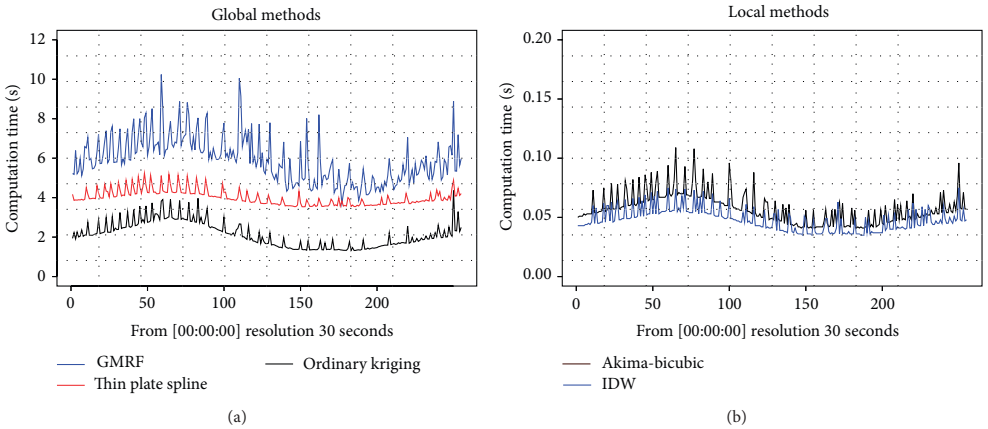


Figure 7: The left panel shows the CT from global methods, the GMRF (blue curve), TPS (red), and OK (brown). The right panel shows the computation time (CT) caused by the local interpolation methods, the AKIMA (brown) and IDW (blue).

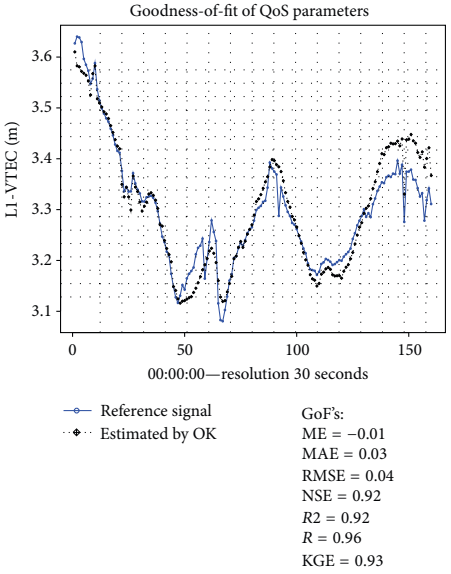


Figure 8: QoS computation. Time series of predicted values by ordinary kriging (black curve) and the reference signal (blue curve). Goodness of fit parameters are given below the plot.

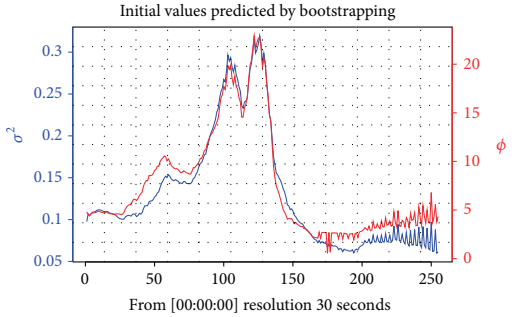


Figure 9: Bootstrapping to accelerate the convergence of OK. Partial sill (σ^2, blue curve) and range parameter (ϕ, orange curve).

Table 4: Computed QoS by different algorithms.

QoS Parameters	Akima bilinear	Akima bicubic	TPS	GMRF	Kriging
ME	−0.050	0.030	**0.010**	−0.030	0.020
MAE	0.090	0.300	**0.070**	0.090	0.080
MSE	**0.020**	0.410	**0.020**	**0.020**	**0.020**
RMSE	0.150	0.640	**0.130**	0.140	0.140
NSE	0.660	0.440	**0.790**	0.480	0.740
$R2$	0.780	0.660	0.820	**0.840**	0.780
ρ	.920	.930	.940	**.950**	**.950**
KGE	0.925	0.589	0.911	0.955	**0.976**
R	0.958	0.829	0.953	0.959	**0.980**
β	0.997	1.002	0.994	1.005	**1.001**
α	0.937	0.626	1.076	1.016	**1.013**

comparison to determine the best interpolation algorithm is performed for each epoch.

8.5. Practical Considerations

8.5.1. Ordinary Kriging. The library geoR from CRAN (Comprehensive R Archive Network) is used to implement the spatial interpolation. In order to produce correct results, the data analyst must handle many small exceptions that may otherwise result in unexpected output.

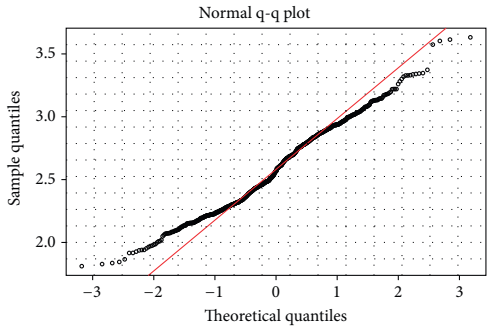


Figure 10: The quantile-quantile (q-q) plot for visualizing normality test. Jarque-Bera test failed and maximum likelihood is not used to estimate parameters.

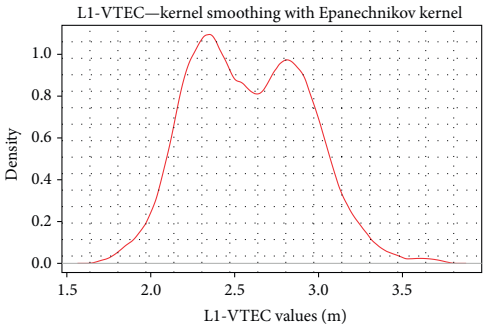


Figure 11: Nonparametric smoothing with an Epanechnikov kernel is used to determine the L1-VTEC distribution. The distribution is not Gaussian. The weighted least square is used in this case to estimate parameters.

(i) Performance Analysis. The variance of the estimated grids are analyzed. If the variance is very small, this ensures stability of the algorithm.

(ii) Parameter Estimations. In order to increase/accelerate the convergence of ordinary kriging, we have incorporated the bootstrapping algorithm on historical data to get very good estimates of the initial values. Figure 9 shows the first 255 estimates for the parameters partial sill (σ^2) and range parameter (ϕ), for moderate ionospheric activity, and the network configuration with 75 reference receivers.

(iii) Anisotropy. The optimal estimated kriging weights (negative weights) and variances are very sensitive to anisotropy. Our aim is to ensure that the spatial process does not depend on direction. The geometry anisotropy correction is applied by transforming a set of coordinates according to the geometric anisotropy parameters. The package gstat does not provide automatic fitting of anisotropy parameters, while the *geoR* package transforms/backtransforms a set of coordinates according to the estimated geometric anisotropy parameters.

(iv) Normality Test. The MLE (maximum likelihood estimation) procedure requires that the observations are Gaussian distributed, this assumption is violated in most cases. Therefore the Jarque-Bera test is used as a test of the normality and is based on the third and fourth moments of a distribution, called skewness and kurtosis coefficients; the interested is referred to [24]. If the test fails, the weighted least square method is used to estimate the parameters. Figures 10 and 11 from a configuration with 100 sites and high ionospheric activity confirm that the L1-VTEC distribution is not normally distributed.

Based on the tests and checks mentioned above, the ordinary kriging is assigned a weight of 0.25 when computing the QoS values.

8.5.2. Test Results. Figure 12 shows the smoothed curve generated by TPS software with GCV. Once the curve is determined, we can easily retrieve any value inside the coverage area without extra computation compared to other

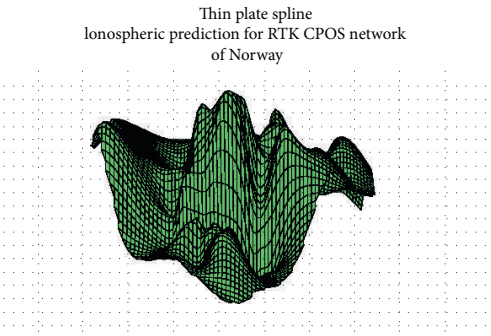


Figure 12: L1-VTEC variation surface generated by the TPS algorithm. The modes are well represented in this case.

interpolation methods. In addition, this describes the variation very well.

Once the smoothing parameter λ is determined by GCV, TPS is the real competitor of the kriging algorithm and the weight assigned to it is 0.25.

9. Conclusion

A significant improvement of the rover position estimation can be achieved by applying the right interpolation/smoothing algorithm at the NRTK interpolation segment. This will reduce the information loss under prediction of the user error level and will provide high quality of virtual reference station data from epoch to epoch.

Five methods have been suggested to generate the rover correction. The study shows that the kriging interpolator, the ordinary kriging with the Matérn covariance function, is the most appropriate choice for weighted spatial linear interpolation, while TPS is a strong competitor of OK when the aim is to smooth, not to interpolate, the data. After performing matrix sparsity tests, the GMRF is computationally effective,

Table 5: CPOS network reference station characteristics.

Site name	4chars	ID	Lon	Lat	Receiver	Antenna
Aas	aasc	189	10.78	59.66	NETR9	TRM57971.00
Adamselv	adac	147	26.70	70.41	NETR9	TRM57971.00
Akrahamn	akrc	13	5.19	59.26	NETR9	TPSCR3_GGD
Alesund	ales	135	6.20	62.48	NETR8	TRM55971.00
Alta	altc	70	23.30	69.98	NETR5	TRM55971.00
Alvdal	alvc	51	10.63	62.11	NETR5	TRM55971.00
Andenes	ande	69	16.13	69.33	NETR5	TRM55971.00
Ardal	arda	177	7.70	61.24	NETR9	TRM57971.00
Arendal	arec	27	8.71	58.41	NETR9	TRM55971.00
Arnes	arnc	99	11.48	60.12	NETR5	TRM55971.00
Asane	asac	169	5.34	60.47	NETR9	TRM57971.00
Balsfjord	balc	71	19.23	69.24	NETR5	TRM55971.00
Bardu	barc	64	18.35	68.86	NETR5	TRM55971.00
Batsfjord	batc	132	29.73	70.64	NETR5	TRM57971.00
Bergen	brgs	150	5.27	60.29	NETR8	TRM55971.00
Berkak	berc	23	10.01	62.83	NETR5	TRM55971.00
Bjarkoy	bjac	63	16.57	69.00	NETR5	TRM55971.00
Bjorli	bjoc	127	8.20	62.26	NETR9	TRM57971.00
Bjugn	bjuc	61	9.81	63.77	NETR5	TRM55971.00
Bleikvassli	blec	40	13.81	65.89	NETR5	TRM55971.00
Bo	boec	2	14.47	68.68	NETR5	TRM55971.00
Bodo	bod3	121	14.43	67.29	NETR8	TRM55971.00
Bomlo	bmlc	142	5.20	59.80	NETR5	TRM55971.00
Breivikbotn	brec	134	22.30	70.59	NETR5	TRM57971.00
Bygland	bygc	28	7.80	58.83	NETR5	TRM55971.00
Dagali	dgls	160	8.50	60.42	NETR8	TRM55971.00
Dombas	doms	140	9.11	62.07	NETR8	TRM55971.00
Donna	donc	90	12.47	66.10	NETR5	TRM55971.00
Drevsjo	drec	144	12.03	61.88	NETR5	TRM55971.00
Eidsvoll	eids	176	11.26	60.33	NETR9	TRM57971.00
Etne	etne	178	5.93	59.66	NETR9	TRM57971.00
Fauske	faus	183	15.39	67.26	NETR9	TRM57971.00
Fet	fetc	173	11.17	59.92	NETR9	TRM57971.00
Finnoy	fioy	180	5.87	59.17	NETR9	TRM57971.00
Finnsnes	finc	65	17.99	69.23	NETR5	TRM55971.00
Flisa	flic	98	12.01	60.61	NETR5	TRM55971.00
Floro	floc	42	5.04	61.60	NETR5	TRM55971.00
Folling	folc	91	11.62	64.12	NETR5	TRM59800.00
Forde	forc	168	5.89	61.46	NETR9	TRM57971.00
Fredrikstad	frec	143	10.94	59.21	NETR9	TRM57971.00
Froya	froc	60	8.66	63.87	NETR5	TRM55971.00
Gjora	gjoc	12	9.11	62.55	NETR9	TRM57971.00
Gloppen	gloc	41	6.19	61.77	NETR5	TRM55971.00
Grong	groc	3	12.31	64.46	NETR5	TRM55971.00
Hammerfest	hamc	137	23.66	70.67	NETR5	TRM57971.00
Hansnes	hanc	72	19.63	69.97	NETR5	TRM55971.00
Hardbakke	harc	38	4.84	61.08	NETR9	TPSCR3_GGD
Haukeli	hauc	79	7.20	59.81	NETR5	TRM55971.00
Havoysund	havc	141	24.66	71.00	NETR5	TRM57971.00
Hedalen	hedc	31	9.74	60.61	NETR5	TRM55971.00
Heggenes	hegc	50	9.07	61.15	NETR5	TRM55971.00
Hellesylt	helc	18	6.76	62.04	NETR5	TRM55971.00
Hemne	hemc	59	9.08	63.29	NETR5	TRM55971.00
Holasen	holc	45	11.18	63.70	NETR5	TRM55971.00
Honefoss	hfs4	1	10.25	60.14	NETR5	TRM59800.00

Table 5: Continued.

Site name	4chars	ID	Lon	Lat	Receiver	Antenna
Honningsvag	hons	111	25.96	70.98	NETR8	TRM59800.00
Horten	horc	158	10.47	59.41	NETR9	TRM57971.00
Hustad	husc	15	7.14	62.97	NETR5	TRM55971.00
Ibestad	ibec	153	17.18	68.78	NETR9	TRM55971.00
Innfjorden	innc	17	7.55	62.50	NETR5	TRM55971.00
Jostedalen	josc	167	7.28	61.66	NETR9	TRM57971.00
Karasjok	karc	139	25.52	69.47	NETR5	TRM57971.00
Kautokeino	kaus	112	23.02	69.02	NETR8	TRM59800.00
Kirkenes	kirc	129	30.04	69.73	NETR5	TRM57971.00
Kjopsvik	kjoc	82	16.39	68.10	NETR5	TRM55971.00
Kobbelv	kobc	83	15.89	67.58	NETR5	TRM55971.00
Kongsvinger	konc	172	12.00	60.19	NETR9	TRM57971.00
Koppang	kopc	52	11.04	61.57	NETR5	TRM55971.00
Kristiansand	krss	185	7.91	58.08	NETR8	TRM55971.00
Kristiansund	krsu	162	7.73	63.11	NETR9	TRM57971.00
Kvanangsbotn	kvac	73	22.06	69.72	NETR5	TRM55971.00
Kvikne	kvic	122	10.32	62.56	NETR5	TRM57971.00
Kyrkjebo	kyrc	58	5.90	61.16	NETR5	TRM57971.00
Lakselv	lakc	138	24.96	70.05	NETR5	TRM57971.00
Lauvsnes	lauc	14	10.90	64.50	NETR5	TRM55971.00
Leikanger	leic	6	6.86	61.18	NETR5	TRM55971.00
Leknes	lekc	136	13.61	68.15	NETR5	TRM57971.00
Leksvik	leks	186	10.63	63.67	NETR9	TRM57971.00
Lierne	liec	30	13.65	64.40	NETR5	TRM59800.00
Lillehammer	lilc	32	10.44	61.14	NETR5	TRM55971.00
Lindas	linc	37	5.16	60.73	NETR9	TPSCR3_GGD
Lista	lstc	24	6.69	58.09	NETR9	TPSCR3_GGD
Lodingen	lodc	62	15.99	68.41	NETR9	TRM55971.00
Lofoten	lofs	119	13.04	67.89	NETR8	TRM55971.00
Lom	lomc	126	8.57	61.84	NETR8	TRM57971.00
Lonsdal	lonc	88	15.46	66.74	NETR5	TRM55971.00
Loppa	lopc	74	22.35	70.24	NETR5	TRM55971.00
Loten	lotc	163	11.35	60.82	NETR9	TRM57971.00
Luroy	lurc	48	13.01	66.51	NETR5	TRM55971.00
Lysefjorden	lysc	9	6.39	59.03	NETR9	TRM55971.00
Mare	mare	184	11.43	63.93	NETR9	TRM57971.00
Maurset	mauc	124	7.33	60.41	NETR5	TRM57971.00
Maze	mazc	149	23.67	69.45	NETR9	TRM57971.00
Mebonden	mebc	46	11.03	63.23	NETR5	TRM55971.00
Mehamn	mehc	133	27.85	71.04	NETR5	TRM55971.00
Mo i Rana	moic	49	14.14	66.31	NETR5	TRM55971.00
Moelv	moec	174	10.70	60.93	NETR9	TRM57971.00
Molde	mldc	159	7.15	62.73	NETR9	TRM57971.00
Moldjord	molc	84	14.57	67.01	NETR5	TRM55971.00
Mosjoen	mosj	188	13.20	65.84	NETR9	TRM55971.00
Myre	myrc	67	15.09	68.91	NETR5	TRM55971.00
Mysen	mysc	97	11.33	59.57	NETR5	TRM55971.00
Namsos	nams	187	11.51	64.47	NETR9	TRM57971.00
Narvik	narc	66	17.43	68.44	NETR5	TRM55971.00
Olderdalen	oldc	75	20.53	69.60	NETR5	TRM55971.00
Opera	opec	4	10.75	59.91	NETR5	TRM55971.00
Ornes	ornc	47	13.73	66.86	NETR9	TRM55971.00
Oslo	osls	175	10.37	59.74	NETR8	TRM55971.00
Osterbo	ostc	81	7.51	60.83	NETR5	TRM55971.00
Overbygd	ovec	152	19.29	69.03	NETR9	TRM55971.00

Table 5: Continued.

Site name	4chars	ID	Lon	Lat	Receiver	Antenna
Pasvik	pasc	148	29.69	69.37	NETR9	TRM57971.00
Porsgrunn	prgc	171	9.66	59.14	NETR9	TRM57971.00
Portor	porc	7	9.43	58.80	NETR5	TPSCR3_GGD
Prestasen	prec	10	6.25	59.49	NETR9	TPSCR3_GGD
Rauland	rauc	55	8.05	59.70	NETR9	TRM55971.00
Rena	renc	33	11.37	61.14	NETR5	TRM55971.00
Roan	roac	92	10.30	64.21	NETR5	TRM55971.00
Roros	rorc	123	11.39	62.58	NETR5	TRM55971.00
Rosendal	rosc	35	6.01	59.98	NETR9	TPSCR3_GGD
Royrvik	royc	19	13.53	64.90	NETR5	TRM59800.00
Sandvika	svic	151	10.52	59.89	NETR9	TRM55971.00
Seljord	selc	43	8.63	59.49	NETR5	TRM55971.00
Sirevag	sirc	11	5.79	58.50	NETR9	TPSCR3_GGD
Skaland	skac	154	17.30	69.45	NETR9	TRM57971.00
Stjordal	stjc	170	10.92	63.47	NETR9	TRM57971.00
Sulitjelma	sulc	86	16.08	67.12	NETR9	TRM55971.00
Sveindal	svec	26	7.47	58.49	NETR5	TRM55971.00
Svolvar	svoc	87	14.56	68.23	NETR5	TRM55971.00
Tana	tanc	131	28.19	70.20	NETR5	TRM55971.00
Terrak	terc	57	12.38	65.09	NETR5	TRM55971.00
Tingvoll	tinc	16	8.21	62.91	NETR5	TRM55971.00
Tjome	tjmc	68	10.40	59.13	NETR5	TRM55971.00
Tonstad	tnsc	34	6.71	58.67	NETR5	TRM55971.00
Tregde	tgde	96	7.55	58.01	NETR9	AOAD/M_T
Treungen	trec	25	8.52	59.02	NETR5	TRM55971.00
Trofors	troc	39	13.39	65.54	NETR5	TRM55971.00
Tromso	tro1	115	18.94	69.66	NETR8	TRM59800.00
Trondheim	trds	125	10.32	63.37	NETR8	TRM55971.00
Trysil	trys	145	12.38	61.42	NETR8	TRM55971.00
Tyin	tyic	89	8.23	61.18	NETR5	TRM55971.00
Tysvar	tysv	179	5.40	59.43	NETR9	TRM57971.00
Ulefoss	ulec	8	9.28	59.28	NETR5	TPSCR3_GGD
Ulsak	ulsc	54	8.62	60.84	NETR5	TRM55971.00
Vadso	vadc	130	29.74	70.08	NETR5	TRM57971.00
Valle	valc	78	7.48	59.25	NETR5	TRM55971.00
Vardo	vars	110	31.03	70.34	NETR8	TRM59800.00
Vega	vegs	118	11.96	65.67	NETR8	TRM59800.00
Veggli	vegc	44	9.17	60.04	NETR9	TRM55971.00
Vikna	vikc	56	11.24	64.86	NETR5	TRM55971.00
Vinstra	vinc	53	9.75	61.60	NETR5	TRM55971.00
Volda	vold	181	6.08	62.14	NETR9	TRM57971.00
Voss	vssc	128	6.42	60.63	NETR9	TPSCR3_GGD
Skjervoy	skjc	76	20.98	70.03	NETR5	TRM55971.00
Skollenborg	skoc	5	9.68	59.62	NETR9	TPSCR3_GGD
Skreia	skrc	29	10.93	60.65	NETR5	TRM55971.00
Smola	smol	182	7.96	63.51	NETR5	TRM55971.00
Smorfjord	smrc	146	24.95	70.52	NETR9	TRM57971.00
Stadt	stac	20	5.32	62.12	NETR5	TRM55971.00
Stavanger	stas	165	5.60	59.02	NETR8	TRM55971.00
Steigen	stgc	164	15.02	67.78	NETR5	TRM55971.00

requires less memory, and produces good results as TPS and OK.

For local methods the covariance structure is in general not conserved. For gentle variation in data, the Akima with bicubic method is an appropriate choice because it is the real spline method. While IDW is stable, it is inaccurate and in addition does not conserve the covariance structure of the process under study.

One major benefit of these techniques is that there is no need for any prior estimation of the spatial dependency, as in the case of Bayesian analysis (e.g., Kalman filter).

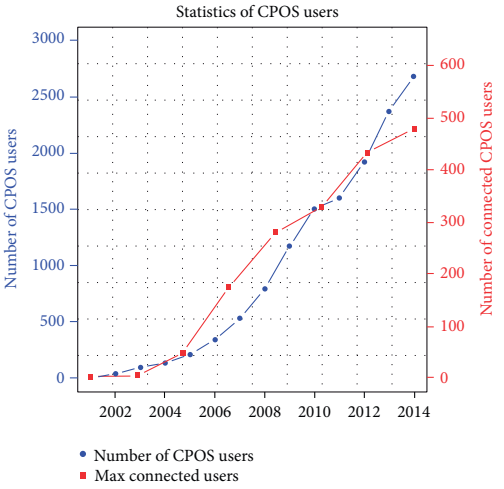


Figure 13: Number of Norwegian CPOS RTK users (blue curve) increased from 1 to ~3000 and the users connected to the system at the same time (red curve) increased from 1 to ~500.

10. Discussions

(1) As we mentioned in the Introduction, processing large data sets is a challenge of the future, and our suggestion for how to handle this is formulated as follows. First of all, we already have enough mathematical tools to do the job, so we do not need to develop new ones. These tools can be considered as elementary building blocks in the hands of the data analyst/modeler. The main challenge is to know that the right tools exist, what they can do for us, what their strengths and weaknesses are, and how to combine them in appropriate ways to describe the observed variations as well as possible. Figure 13 shows the number of users connected at the same time and the historical data of the users using the CPOS services in Norway. Both curves increase exponentially in a period of one decade, and if the development continues to follow the same pattern, the existing tools will not be sufficient to process the large data sets.

(2) Data quality and quantity are important to perform reliable statistical analysis, and elementary checks are necessary before starting analysis.

(3) In geodesy and geophysical data analysis, the Gauss-Markov model and Kalman Filter are often considered when modeling and when state estimation is necessary. Since new navigation satellite system (e.g., Galileo, Beidou) in addition to the old GPS and GLONASS becomes operation, massive data sets need to be processed in real-time, so we are experiencing a computational paradigm shift.

(4) To avoid information loss between the correction and interpolation segments, picking the right algorithm for the job is essential for improving the user position errors.

Appendix

CPOS Station Characteristics

See Table 5.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank the SATREF group at the Norwegian Mapping Authority for their effort for providing the data, especially Tor O. Trondersk. Without their effort, this investigation would not have been possible. They would also like to thank Ole Ostevdal from the Norwegian University of Life Sciences for his software to generate absolute ionospheric vertical error that are used in this investigation. The software is modified in order to meet their computation needs. The international GNSS Service (IGS) is acknowledged for providing geodetic infrastructure and geodetic products used in this work.

References

[1] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance*, Ganga-Jamuna Press, 2006.

[2] H. Landau, U. Vollath, and X. Chen, "Virtual reference station systems," *Journal of Global Positioning Systems*, vol. 1, no. 2, pp. 137–143, 2002.

[3] F. Takac and O. Zelzer, "The relationship between network RTK solutions MAC, VRS, PRS, FKP and i-MAX," in *Proceedings of the 21st International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS '08)*, pp. 348–355, 2008.

[4] N. A. C. Cressie, *Statistics for Spatial Data*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York, NY, USA, 1993.

[5] J. Dennis and J. J. More, "Quasi-Newton methods, motivation and theory," *SIAM Review*, vol. 19, no. 1, pp. 46–89, 1977.

[6] M. A. Luersen and R. le Riche, "Globalized nelder-mead method for engineering optimization," *Computers and Structures*, vol. 82, no. 23–26, pp. 2251–2260, 2004.

[7] W. J. Gordon and J. A. Wixom, "Shepard's method of 'metric interpolation' to bivariate and multivariate interpolation," *Mathematics of Computation*, vol. 32, no. 141, pp. 253–264, 1978.

[8] R. J. Renka and R. Brown, "Algorithm 792: accuracy tests of ACM algorithms for interpolation of scattered data in the plane," *ACM Transactions on Mathematical Software*, vol. 25, no. 1, pp. 78–94, 1999.

[9] H. Akima, "A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points," *ACM Transactions on Mathematical Software*, vol. 4, no. 2, pp. 148–159, 1978.

[10] H. Akima, "Algorithm 761: scattered-data surface fitting that has the accuracy of a cubic polynomial," *ACM Transactions on Mathematical Software*, vol. 22, no. 3, pp. 362–371, 1996.

[11] C. H. Reinsch, "Smoothing by spline functions," *Numerische Mathematik*, vol. 10, pp. 177–183, 1967.

[12] J. Duchon, "Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces," *Analyse Numerique*, vol. 10, pp. 5–12, 1976.

[13] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, USA, 2010.

[14] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society Series B: Methodological*, vol. 48, no. 3, pp. 259–302, 1986.

[15] H. Rue and L. Held, *Gaussian Markov Random Fields: Theory and Applications*, CRC Monographs on Statistics & Applied Probability, CRC Press, London, UK, 2005.

[16] M. Abramowitz, I. A. Stegun, and D. Miller, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, vol. 32, Dover Publications, 1965.

[17] M. L. Stein, *Interpolation of Spatial Data*, Springer Series in Statistics, Springer, New York, NY, USA, 1999.

[18] J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models. Part I. A discussion of principles," *Journal of Hydrology*, vol. 10, no. 3, pp. 282–290, 1970.

[19] H. V. Gupta, H. Kling, K. K. Yilmaz, and G. F. Martinez, "Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling," *Journal of Hydrology*, vol. 377, no. 1-2, pp. 80–91, 2009.

[20] J. M. Dow, R. E. Neilan, and C. Rizos, "The international GNSS service in a changing landscape of global navigation satellite systems," *Journal of Geodesy*, vol. 83, no. 3-4, pp. 191–198, 2009.

[21] N. Kjorsvik, O. Ovstedal, and J. G. Svendsen, "Estimating local ionosphere parameters in Norway," in *Proceedings of the 15th International Technical Meeting of the Satellite Division of the Institute Of Navigation (ION GPS '02)*, pp. 2304–2310, 2002.

[22] P. Guttorp and P. D. Sampson, "Methods for estimating heterogeneous spatial covariance functions with environmental applications," in *Handbook of Statistics XII: Environmental Statistics*, pp. 663–690, 1994.

[23] N. D. Le and J. V. Zidek, *Statistical Analysis of Environmental Space-Time Processes*, Springer Series in Statistics, Springer, New York, NY, USA, 1st edition, 2006.

[24] C. M. Jarque and A. K. Bera, "A test for normality of observations and regression residuals," *International Statistical Review*, vol. 55, no. 2, pp. 163–172, 1987.

Paper #3

Reference

M. Ouassou, O. Kristiansen, J.G.O. Gjevestad, K.S. Jacobsen & Y.L. Andalsvik
Estimation of scintillation indices:
a novel approach based on local kernel regression methods
International Journal of Navigation and Observation, 3582176 (2016)

Contribution

M. Ouassou conceived the idea, performed the numerical calculations, and wrote the majority of the manuscript. All authors contributed to the discussion of the results and revisions of the manuscript.

Y. L. Andalsvik generated the test data from the Norwegian Regional Ionospheric Scintillation Network.

K. S. Jacobsen involved in validating the raw scintillation data and contributed to section 1.

J.G.O. Gjevestad and O. Kristiansen involved in discussion, support and conclusion.

J. A. Ouassou involved to derive the equations in Section 4.4: Analytical Analysis of KDE and also acknowledged for useful discussions, feedback, and proofreading.

Hindawi Publishing Corporation
International Journal of Navigation and Observation
Volume 2016, Article ID 3582176, 18 pages
http://dx.doi.org/10.1155/2016/3582176

Research Article

Estimation of Scintillation Indices: A Novel Approach Based on Local Kernel Regression Methods

Mohammed Ouassou,[1,2] Oddgeir Kristiansen,[1,2] Jon G. O. Gjevestad,[2,3] Knut Stanley Jacobsen,[1] and Yngvild L. Andalsvik[1]

[1]*Norwegian Mapping Authority, Geodetic Institute, 3511 Hønefoss, Norway*
[2]*Department of Mathematical Sciences and Technology, NMBU, 1432 Akershus, Norway*
[3]*Department of Electronics and Telecommunications, NTNU, 7491 Trondheim, Norway*

Correspondence should be addressed to Mohammed Ouassou; mohammed.ouassou@statkart.no

Received 25 January 2016; Revised 25 April 2016; Accepted 9 May 2016

Academic Editor: Sandro M. Radicella

We present a comparative study of computational methods for estimation of ionospheric scintillation indices. First, we review the conventional approaches based on Fourier transformation and low-pass/high-pass frequency filtration. Next, we introduce a novel method based on nonparametric local regression with bias Corrected Akaike Information Criteria (AICC). All methods are then applied to data from the Norwegian Regional Ionospheric Scintillation Network (NRISN), which is shown to be dominated by phase scintillation and not amplitude scintillation. We find that all methods provide highly correlated results, demonstrating the validity of the new approach to this problem. All methods are shown to be very sensitive to filter characteristics and the averaging interval. Finally, we find that the new method is more robust to discontinuous phase observations than conventional methods.

1. Introduction

Global Navigation Satellite Systems (GNSS) are based on satellite signals being transmitted to receivers on the ground. These signals must pass through the ionosphere on the way, where the variation in electron density causes undesirable fluctuations in the observed signal. This distortion, known as *scintillation*, can affect both the amplitude and phase of GNSS signals. The relevant electron density variations range from decameters to kilometers in size [1–3].

In the auroral zones, it is much more common for phase scintillation than amplitude scintillation to be the dominant effect (e.g., [4–6]). Scintillation is a large challenge for navigational systems, as it can disturb not only single-receiver systems but also networked systems [3, 4, 7–9]. In a worst-case scenario, the user receiver can end up with a loss of signal lock, which causes discontinuities in the phase measurements. This is known as a *cycle-slip*.

A good understanding of the ionosphere morphology can aid the development of suitable mitigation algorithms, that is, algorithms that assign weights to observations based on the scintillation distortion of each satellite link.

To be more specific, a more realistic stochastic model for GNSS observables would have to take into account the variance caused by scintillation. This would avoid biased solutions. Today, the stochastic model includes the following: correlation among observations [10]; satellite elevation dependency modeled by exponential function [11]; temporal and cross-correlations [12]; and multipath detection and monitoring [13]. That is, a suitable robust weighting algorithm that reduces the influence of the satellite exposed by scintillation will look similar to the one proposed by Eueler and will enhance the ability to resolve the carrier-phase ambiguity and to improve the stochastic model for GNSS processes.

This is most commonly gauged by the *phase scintillation index* σ_ϕ, which is simply defined as the standard deviation of the detrended carrier phase ϕ over some period of time. A critical step in the calculation of σ_ϕ is a frequency filter, which is used to separate the high-frequency ionospheric distortions from the low-frequency distortions due to, for example,

multipath interference. The scintillation indices are usually calculated for one-minute intervals and then processed with a Butterworth sixth-order high-pass filter [14–23]. However, the estimated value of σ_ϕ can be highly sensitive to the cutoff frequency of the filter. While the most common cutoff frequency is 0.1 Hz, it has been shown that this is suboptimal at high latitudes and that higher values like 0.3 Hz might yield better results [24]. The variation in the carrier phase of GNSS signals is also commonly quantified using the rate of total electron content index (ROTI) [25–27]. Several other indices for quantifying phase variations have also been proposed [24, 28–32]. It is worth noting that all of these indices are directly related to the scintillation signal variance.

Some scintillation indices use methods based on wavelet transforms for the filtration. A major advantage of wavelet filters is that they manage to preserve local features in the signal, thus avoiding misinterpretation that can occur when using standard filter approaches [24, 33–36]. While such wavelet filters may yield better results than conventional filters, the computational load can become very high, especially when processing data with high sampling rates (50+ Hz). The improvement that can be achieved by substituting a wavelet filter for the conventional filters must therefore be weighed against the computational load required, especially if the algorithm is intended for real-time applications.

The primary focus of this paper is on nonparametric local regression with bias Corrected Akaike Information Criteria as a viable alternative for the computation of reliable scintillation indices. In addition, we perform a statistical analysis of the scintillation indices and show that they can be well described by the Nakagami and Frechet distributions. The paper is organized as follows.

Section 2: it includes detailed description of the data sets used in this paper; Section 3: it deals with digital filter construction of low-pass filter algorithms for amplitude scintillation index computation and high-pass filter algorithms for phase scintillation computation. Filter types discussed are the Chebyshev, Butterworth, and Elliptic filters. These algorithms are used to validate the new approach; Section 4: it tackles nonparametric regression with kernel smoothing; Section 5: this section presents how the scintillation indices are computed by the conventional methods and the new approach; Section 6: it shows statistical analysis of scintillation indices, including distributions, sensitivity, and correlation analysis of implemented algorithms; Section 7: it presents generalization of the preceding sections; Section 8: it includes conclusion and applications of the results.

2. Test Data

The data used for the investigation herein was obtained from three different NRISN observation sites in the northernmost parts of Norway. These sites are highlighted with an italic typeface in Table 1 and encircled in Figure 1. The observational data is taken from day 50 of year 2014, which according to the I95 index corresponds to relatively high ionospheric activity. (The *I95 index* [37] is used for the classification of ionospheric activity by CPOS software.) For more information about the ionospheric activity levels, see Figures 15 and 16. The NRISN sites are equipped by dual-frequency Septentrio PolaRxS receivers tracking GPS and GLONASS L1 and L2 frequencies with sampling rate of 100 Hz and are capable of tracking up to 12 satellites for each system. The baselines vary between 176 and 1988 km, and the height difference between sites is around 400 m. To give a full picture of the baseline lengths and the height difference between sites above the ellipsoid WGS84, Tables 2 and 3 are provided.

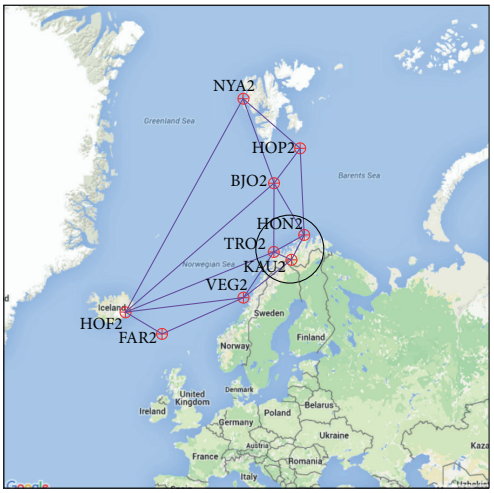


Figure 1: Sites characteristics of Norwegian Ionospheric Scintillation Network.

To accurately determine the relative performances of the various algorithms, it is important to test different values for the filter parameters. This also allows us to determine the optimal filter parameters for the computation of the amplitude and phase variances for GNSS signals. In this case, the relevant parameters are the satellite elevation angle, cutoff frequency, and averaging period. The algorithms were tested with a cutoff frequency of 0.1 ± 0.05 Hz for the signal amplitude and 0.3 ± 0.1 Hz for the phase and an averaging interval of 30, 60, or 90 seconds.

3. Digital Filters Construction

To study the influence of scintillation on GNSS signals, the standard deviations of the amplitude fluctuations S_4 and phase fluctuations σ_ϕ must be computed. To achieve this goal, digital filters that satisfy our requirements have to be constructed. For this, we require both a suitable low-pass filter to compute the amplitude scintillation index S_4 and a high-pass filter to compute the phase scintillation index σ_ϕ.

Digital Finite Impulse Response (FIR) and Infinite Impulse Response (IIR) filters are chosen and used to construct sixth-order Butterworth, Chebyshev, and Elliptic low-pass filters. Further details on the design and implementation of digital FIR/IIR filters are well described in [38–41]. Once

Table 1: List of NRISN observation sites.

Site	4-char	ID	Receiver type	Antenna type
Tromso	*TRO2*	1001	Septentrio PolaRxS	MCGNSS Septentrio
Vega	VEG2	1002	Septentrio PolaRxS	PolaNt_MC
Ny-Ålesund	NYA2	1003	Septentrio PolaRxS	JAV_RING-ANT_G3T
Hofn	HOF2	1010	Septentrio PolaRxS	TRM41249.00
Færøyene	FAR2	1012	Septentrio PolaRxS	LEIAT504GG
Kautokeino	*KAU2*	1006	Septentrio PolaRxS	TRM559800.00
Honningsvag	*HON2*	1005	Septentrio PolaRxS	TRM559800.00
Bjørnøya	BJO2	1008	Septentrio PolaRxS	TRM41249.00
Hopen	HOP2	1009	Septentrio PolaRxS	TRM41249.00

Table 2: List of NRISN distance differences (km).

Sites	TRO2	VEG2	NYA2	HOF2	FAR2	KAU2	HON2	BJO2	HOP2
TRO2	X	533.5	1053	1578	1433	176	301	540	787
VEG2	—	X	1476	1278	1004	603	822	1018	1288
NYA2	—	—	X	1846	1982	1152	970	526	411
HOF2	—	—	—	X	493.5	1734	1849	1715	1954
FAR2	—	—	—	—	X	1559	1728	1720	1988
KAU2	—	—	—	—	—	X	245	626	837
HON2	—	—	—	—	—	—	X	454	617
BJO2	—	—	—	—	—	—	—	X	280
HOP2	—	—	—	—	—	—	—	—	X

Table 3: NRISN ellipsoidal station heights (WGS84).

Site	Height (m)
Tromso	132.4342
Vega	56.4121
Ny-Ålesund	81.2904
Honningsvag	73.8241
Hopen	37.4744
Hofn	82.8036
Færøyene	110.2158
Kautokeino	413.3949
Bjørnøya	53.6580

the filter specifications have been determined, the next step is to compute the corresponding filter coefficients in the frequency domain. This is usually done with the computationally efficient Parks-McClellan algorithm. A detailed description of the algorithm is beyond the scope of this paper. The interested reader is referred to [41, Section 7.3]. Since an ideal low-pass filter is impossible to obtain, we often use approximations. In this paper, three different approximations are considered, namely, the Butterworth, Chebyshev, and Elliptic filters.

3.1. Cutoff Frequency Determination. A raw GNSS scintillation signal is a nonstationary signal that is affected by the Doppler shift caused by satellite-receiver relative motion (trend), the slowly varying background ionosphere, and scintillation. In order to study the scintillation, trends and slow fluctuations in the raw signal should be removed; that is, any signal components that do not originate from scintillation must be excluded. There exist many methods for carrying out this detrending operation, for example, linear regression by fitting a straight line to the original data or the ratio and difference detrending methods. Detrending is the statistical operation performed to remove the long-term change in the mean function and is equivalent to high-pass filter. That is, the variance at lower frequencies is reduced compared to variance at high frequencies.

In general, the data detrending is a preprocessing step to achieve stationarity of the observed signal. The next step is then to analyze the signal statistically, by, for example, computing the first and the second moments of the signal, namely, the phase scintillation index σ_ϕ and amplitude scintillation index S_4. To perform a proper detrending, the cutoff frequency must be selected properly. The most important part of the scintillation spectrum is around the Fresnel frequency. It also extends to higher frequencies, but with decreasing power per bandwidth. The Fresnel frequency is $f_F = v/\sqrt{2\lambda z}$ [3, 30], where v is the relative velocity between the satellite and ionosphere, λ is the wavelength of the signal (~19 cm), and z denotes the distance between the receiver and the ionosphere (~350 km at 90 degrees elevation). Ideally, the cutoff frequency should be slightly below the Fresnel

frequency but still high enough to remove the unwanted signal components. Also, it should be a constant, so that the scintillation values are comparable. For the case of GNSS scintillation data, a Butterworth filter with a constant cutoff frequency of 0.1 Hz is the most common approach, but it has been shown that this value is inappropriate at high latitudes [16, 24].

3.2. Butterworth Approximation. The high-order Butterworth approximation [40, p. 264] satisfies our needs, as the filter has a low ripple and minimum transition band. Due to these characteristics, the Butterworth sixth-order filter is the most used filter in computation of scintillation indices S_4 and σ_ϕ. The cutoff frequency f_c is user defined parameter, by default set to 0.1 Hz for low-pass filters and 0.3 Hz for high-pass filters. The derivation of such indices is usually carried out using Algorithm 1, that is, fast convolution using a fast Fourier transform (FFT) and multiplication in the frequency domain. For each Fourier coefficient, the magnitude of the frequency response of nth-order Butterworth filter is given by the following equation:

$$|H(f)| = \frac{1}{\sqrt{1 + (f/f_c)^{\pm 2n}}}. \tag{1}$$

Note the ± sign in the exponent, which toggles whether the filter exhibits a low-pass or high-pass behavior. Also, the filter gain has been normalized to unity in the region with minimum attenuation, that is, when either $f \ll f_c$ or $f \gg f_c$.

3.3. Chebyshev Approximation. The Chebyshev approximation [41, p. 27] has passband ripple that can be dumped and a remarkable transition band. Based on these qualities, its chosen to approximate the ideal low-pass filter with the user specification parameters. However, unlike the Butterworth case, the Chebyshev filter has ripples in the passband and rapid transition and is able to satisfy the user specification with lower order than Butterworth. The phase response is not linear as the Butterworth case.

The magnitude of the frequency response of an nth-order Chebyshev filter is

$$|H(f)| = \frac{1}{\sqrt{1 + \epsilon^2 T_n^2 (f/f_c)}}, \tag{2}$$

where ϵ determines the ripple magnitude and is less than 1. T_n is a Chebyshev polynomial given by the expression

$$T_n(x) = \begin{cases} \cos\left[n \cos^{-1}(x)\right], & \text{if } x \le 0, \\ \cosh\left[n \cosh^{-1}(x)\right], & \text{if } x > 0. \end{cases} \tag{3}$$

3.4. Elliptic Approximation. The Elliptic filters [40, p. 275] known as *Cauer filters* yield smaller transition bandwidths than Butterworth or Chebyshev filters for any given order but are equiripple in both the passbands and stopbands. In general, Elliptic filters meet a given set of performance specifications with the lowest order of any filter type.

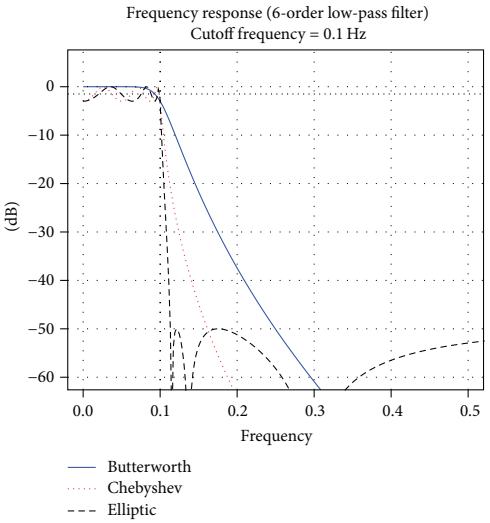


Figure 2: Typical low-pass digital filter characteristics used to compute the amplitude scintillation.

For Elliptic filters, the normalized cutoff frequency f_c is a number between 0 and 1, where 1 corresponds to half the sampling frequency (Nyquist frequency). The filter has the smallest transition band of the two approaches described so far but is very complicated analytically.

The magnitude of the frequency response of an nth-order Elliptic filter is

$$|H(f)| = \frac{1}{\sqrt{1 + \epsilon^2 E_n^2 (f/f_c)}}, \tag{4}$$

where ϵ controls the deviation in the passband and E_n is the Jacobian Elliptic function [42, pp. 567–588] of order n.

3.5. Implementation. In this study, we have implemented digital filters of types FIR and IIR. The frequency response for all three filter types are given, respectively, by Figures 2 and 3. These figures show the difference between the filter types (Butterworth, Elliptic, and Chebyshev).

4. Nonparametric Regression

In contrast to other data modeling methods, for instance, parametric and semiparametric, a nonparametric regression makes no assumptions about the shape of the distribution function. This class of modeling lets the data speak for themselves and they open the way for a new model by their flexibility.

Let us now suppose that we have a vector of observations $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ sampled from an unknown density function $f(x)$ and that our aim is to estimate $f(x)$ and display

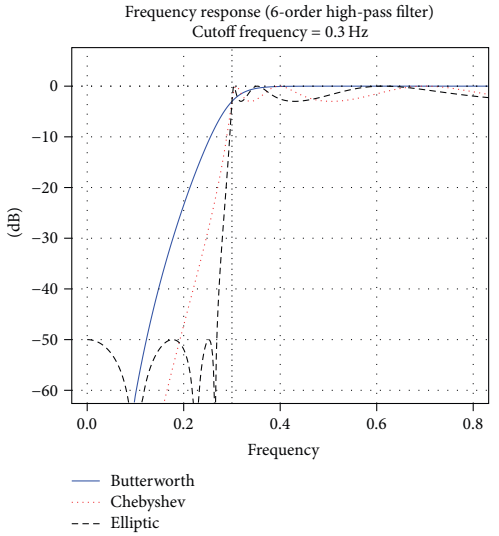


Figure 3: Typical high-pass digital filter characteristics used to compute the phase scintillation.

it graphically. The density estimation function is then defined as

$$\widehat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left[\frac{(x_i - x)}{h}\right], \tag{5}$$

where h is the *smoothing parameter* and $K(\cdot)$ is the kernel function. This kernel cannot be arbitrarily chosen but needs to satisfy three requirements: it must be nonnegative, L_2 normalized, and symmetric around the origin.

There are lots of different kernel functions that satisfy these requirements, for example, the uniform kernel, quartic kernel, Gaussian kernel, and Epanechnikov kernel. The precise choice of kernel function is in practice less important than the choice of smoothing parameter h.

4.1. MSE and MISE. We wish to investigate the performance of the kernel density estimation (KDE) at a single point or over the whole real line and find out how close our estimator is to its target. The Mean Square Error (MSE) and Mean Integrated Square Error (MISE) can be used to measure such performance or efficiency of nonparametric methods.

The MSE when estimating $f(x)$ using the estimator $\widehat{f}(x)$ at point x is given by

$$\begin{aligned}\mathrm{MSE}\left\{\widehat{f}(x)\right\} &= E\left[\widehat{f}(x) - f(x)\right]^2 \\ &= \mathrm{var}\left\{\widehat{f}(x)\right\} + \left\{\mathrm{bias}\left(\widehat{f}(x)\right)\right\}^2.\end{aligned} \tag{6}$$

We see that the MSE can be decomposed into two parts, namely, variance and bias terms. In contrast to parametric models, the bias is ignored. The MISE is defined as

$$\mathrm{MISE}\left\{\widehat{f}(x)\right\} = \int \mathrm{MSE}\left\{\widehat{f}(x)\right\} \mathrm{d}x. \tag{7}$$

4.2. Asymptotic Statistical Properties of KDE. In this section, we will derive asymptotic approximations for the bias and variance of the KDE. These terms are necessary to derive the optimal smoothing parameter h_{opt}. For these derivations, it will be useful for defining the second moment $\mu_2(K)$ and norm $\|K\|$ of the KDE:

$$\begin{aligned}\mu_2(K) &= \int_{-\infty}^{\infty} u^2 K(u)\,\mathrm{d}u, \\ \|K\|_2^2 &= \int_{-\infty}^{\infty} \{K(u)\}^2\,\mathrm{d}u.\end{aligned} \tag{8}$$

4.2.1. Derivation of Bias Approximation. We will start by deriving an expression for the expectation value of the kernel density $\widehat{f}_h(x)$:

$$\begin{aligned}E\left\{\widehat{f}_h(x)\right\} &= \frac{1}{nh}\sum_{i=1}^{n} E\left[K\left(\frac{x - X_i}{h}\right)\right] \\ &\approx f(x) + \frac{1}{2}h^2 f''(x)\,\mu_2(K).\end{aligned} \tag{9}$$

The approximation of $E\{\widehat{f}_h(x)\}$ is obtained by using the Taylor expansion up to the second terms in h and letting $h \to 0$. By the definition of bias, this result yields

$$\begin{aligned}\mathrm{bias}\left\{\widehat{f}_h(x)\right\} &= E\left[\widehat{f}_h(x)\right] - f(x) \\ &\approx \frac{1}{2}h^2 f''(x)\,\mu_2(K).\end{aligned} \tag{10}$$

Some facts about the bias of $\widehat{f}_h(x)$ are as follows:

(1) $\mathrm{bias}\{\widehat{f}_h(x)\} \sim o(h^2)$. Larger values of h will characterize the process as oversmoothing.

(2) The sign and the direction are decided by $f''(x)$.

(3) If we include more terms in the Taylor series expansion, we will reduce the error term.

4.2.2. Derivation of Variance Approximation. The variance of stochastic variable X is defined as $\mathrm{var}(X) = E(X^2) - E(X)^2$ and applying the Taylor expansion of $f(x + sh)$ and $\int_{-\infty}^{\infty} sK(s)^2\mathrm{d}s = 0$ (from the symmetry of the kernel):

$$\begin{aligned}\mathrm{var}\left\{\widehat{f}_h(x)\right\} &= \frac{1}{n^2}\,\mathrm{var}\left[\frac{1}{h}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)\right] \\ &= \frac{1}{nh}\|K\|_2^2 f(x) + o\left(\frac{1}{nh}\right).\end{aligned} \tag{11}$$

Some facts about the variance of $\widehat{f}_h(x)$ are as follows:

(1) $\mathrm{var}\{\widehat{f}_h(x)\} \sim o(h^{-1})$. Small values of h will characterize the process as undersmoothing.

(2) nh gives the number of the observations inside the processing window.

(3) The variance increases with the magnitude of $f(x)$.

4.2.3. Bandwidth Selection of KDE. As mentioned earlier, the bandwidth selection is more important than the choice of the kernel. This section is devoted to computing the optimal smoothing parameter h_{opt}. We start by evaluating the expressions given by (6) and (7):

$$\begin{aligned}\mathrm{MSE}\left\{\widehat{f}_h(x)\right\} &= E\left[\widehat{f}_h(x) - f(x)\right]^2 \\ &\approx \frac{h^4}{4}\left\{f''(x)\,\mu_2(K)\right\}^2 \\ &\quad + \frac{1}{nh}\|K\|_2^2 f(x),\end{aligned} \tag{12}$$

$$\begin{aligned}\mathrm{MISE}\left\{\widehat{f}_h(x)\right\} &= E\left\{\int_{-\infty}^{\infty}\left[\widehat{f}_h(x) - f(x)\right]^2 \mathrm{d}x\right\} \\ &\approx \frac{h^4}{4}\left\{f''(x)\,\mu_2(K)\right\}^2 + \frac{1}{nh}\|K\|_2^2 \\ &\quad + o\left((nh)^{-1}\right) + o\left(h^4\right)\end{aligned} \tag{13}$$

for $h \to 0$ and $(nh) \to \infty$, and the Asymptotic Mean Integrated Square Error (AMISE) reads

$$\mathrm{AMISE}\left\{\widehat{f}_h(x)\right\} = \frac{h^4}{4}\left\{f''(x)\,\mu_2(K)\right\}^2 + \frac{1}{nh}\|K\|_2^2. \tag{14}$$

The exact computation of MISE is given in Wand and Jones [43, p. 24].

The optimal smoothing parameter h_{opt} is found by minimizing the expression given by (14):

$$\begin{aligned}&\frac{\partial\left[\mathrm{AMISE}\left(\widehat{f}_h(x)\right)\right]}{\partial h} \\ &\quad = \frac{\mathrm{d}}{\mathrm{d}h}\left\{\frac{h^4}{4}\left[f''(x)\,\mu_2(K)\right]^2 + \frac{1}{nh}\|K\|_2^2\right\};\end{aligned} \tag{15}$$

set (15) to be equal to 0, and solve; we get

$$h_{\mathrm{MISE}} \sim \left\{\frac{\|K\|_2^2}{\left[f''(x)\,\mu_2(K)\right]^2 n}\right\}^{1/5} \sim o\left(n^{-1/5}\right). \tag{16}$$

Inserting this expression of h_{MISE} into (13), we get

$$\begin{aligned}\mathrm{MISE}\left\{\widehat{f}_h(x)\right\} &\sim \frac{5}{4}\left\{f''(x)\,\|K\|^4\,\mu_2(K)^2\right\}^{1/5} n^{-4/5} \\ &\sim o\left(n^{-4/5}\right).\end{aligned} \tag{17}$$

Several smoothing parameter selection procedures exist, for instance, plug-in methods (Section 4.2.3), focused information criteria (FIC) [44, p. 145], cross-validation, penalizing functions, and Akaike Information Criteria (AIC). The interested reader is referred to [45, Chap. 5].

4.3. Kernel Methods for Nonparametric Regression. In this section, we try to put all pieces together to construct a class of kernel-type regression estimators known as *local polynomial kernel estimators*. The main idea is to estimate the regression function at a particular point x by locally fitting a pth degree polynomial to the data by employing the weighted least square techniques. The weights are chosen to the height of the kernel function centered about the point x.

The steps needed to derive the expression of the local polynomial kernel estimator are as follows.

(i) Local polynomial definition: Let p be the order of the polynomial we fit at point x to estimate $\widehat{y} = \widehat{f}(x)$:

$$P(x) = \beta_0 + \beta_1(X_i - x) + \cdots + \beta_p(X_i - x)^p. \tag{18}$$

(ii) Weights definition: the weights are given by the kernel density function

$$K_h(X_i - x) = \frac{1}{h}K\left(\frac{X_i - x}{h}\right). \tag{19}$$

(iii) Weighted least square: the value of the estimate at a point x is β_0, where $\widehat{\beta}$ minimize the expression:

$$S(\beta) = \sum K_h(X_i - x)\left(Y_i - P(x)\right)^2. \tag{20}$$

The weighted least square solution in matrix form read $\widehat{\beta} = (X^T W X)^{-1} X^T W Y$, where Y is $(n \times 1)$ vector of responses. The design matrix X read

$$\mathbf{X} = \begin{bmatrix} 1 & (X_1 - x) & \cdots & (X_1 - x)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (X_n - x) & \cdots & (X_n - x)^p \end{bmatrix}. \tag{21}$$

The weight matrix W is $(n \times n)$ diagonal matrix of weights $w_i = K_h(X_i - x)$:

$$\mathbf{W} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n \end{bmatrix}. \tag{22}$$

The value of the estimate at a point x is the intercept coefficient β_0 of the local fit of $P(x)$:

$$\widehat{y} = e^T\left(X^T W X\right)^{-1} X^T W Y, \tag{23}$$

where $e^T = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$ is a vector of dimensions $((p+1) \times 1)$.

Some important remarks are as follows.

(1) For $p = 0$, we fit a constant function locally and this estimator is known as the Nadaraya-Watson kernel estimator:

$$\widehat{f}_{\mathrm{NW}}(x) = \frac{\sum_{i=1}^{n} K_h(X_i - x) Y_i}{\sum_{i=1}^{n} K_h(X_i - x)}. \tag{24}$$

(2) For $p = 1$, the estimator function corresponds to the Priestley-Chao kernel estimator

$$\widehat{f}_{\mathrm{PC}}(x) = \frac{1}{h}\sum_{i=1}^{n}(x_i - x_{i-1}) K\left(\frac{x - x_i}{h}\right) y_i. \tag{25}$$

(3) Final product: the key is the optimal smoothing parameter h_{opt}. To compute the phase scintillation index σ_ϕ, the steps are as follows:

(i) Find the smoothing parameter h_{opt}. In this study, we have employed the AIC and the bias Corrected AIC (AICC) for smoothing parameter selection [46, Chap. 3]. This operation is repeated for each computation of σ_ϕ.

(ii) Compute the weights based on the height of the chosen kernel.

(iii) Locally fit a pth degree polynomial to the data by using the weighted least square technique.

4.4. Analytical Analysis of KDE. We will now derive an analytical relationship between the conventional filters and the new approach based on kernel smoothing. With the conventional approach, we apply a Fourier transform to the function $g(x)$, multiply it by the filter transfer function $H(f)$ in the frequency domain, and then apply the inverse Fourier transform to transform the filtered function back to its original domain. According to the convolution theorem of Fourier analysis, this can equivalently be written as a convolution of the function $g(x)$ with the inverse Fourier transform $h(x) = \mathcal{F}^{-1}\{H(f)\}$ of the filter transfer function:

$$\mathcal{F}^{-1}\{\mathcal{F}\{g(x)\} H(f)\} = g(x) * h(x). \tag{26}$$

For simplicity, we will limit the analysis to a first-order low-pass filter of the Butterworth type, as obtained by setting $n = 1$ and choosing a positive exponent in (1):

$$H(f) = \frac{1}{\sqrt{1 + (f/f_c)^2}}. \tag{27}$$

Taking the inverse Fourier transform of this, we get

$$h(x) = \frac{f_c}{\pi} K_0[f_c |x|], \tag{28}$$

where $K_0(\cdot)$ is the modified Bessel function of the second kind. Substituting back into (26), we get

$$\mathcal{F}^{-1}\{\mathcal{F}\{g(x)\} H(f)\} = \frac{f_c}{\pi} g(x) * K_0[f_c |x|]. \tag{29}$$

Writing the convolution explicitly out, we find that the first-order Butterworth filter is equivalent to a kernel smoothing using a Bessel function as the kernel:

$$\begin{aligned}&\mathcal{F}^{-1}\{\mathcal{F}\{g(x)\} H(f)\} \\ &\quad = \frac{f_c}{\pi}\int g(x') K_0\left[f_c \left|x - x'\right|\right] \mathrm{d}x'.\end{aligned} \tag{30}$$

From this result, we see that the cutoff frequency f_c of the Butterworth filter is directly equivalent to the smoothing parameter h in KDE equation (5). The modified Bessel function of the second kind can be written in integral form as

$$K_0(z) = \int_0^\infty \frac{\cos(zt)}{\sqrt{1 + t^2}} \mathrm{d}t. \tag{31}$$

It is also worth noting that

$$\frac{1}{\pi}\int_{-\infty}^{\infty} K_0(z)\, \mathrm{d}z = 1, \tag{32}$$

which means that the Bessel kernel is a proper density function, but with a pole at the origin. We have established a direct analytical link between the conventional Butterworth filter approach and the new kernel smoothing approach. In this light, the nonparametric filter approach can be considered as a generalization of the conventional filters to other kernels, with the conventional filters as a special case.

4.4.1. Higher Order Approximation. For higher order n, the computation of the Fourier transforms of (1), (2), and (4) is very complex analytically. For future work, the numerical methods and Monte Carlo simulation may be considered as suitable techniques for analyzing the relationship between the conventional filters and the new approach based on nonparametric local regression.

4.4.2. Scintillation Cutoff Frequency Determination. We have established the link between the optimal smoothing parameter $1/h_{\mathrm{opt}} = B_w$ and the cutoff frequency f_c of associated filter equations (30) and (5). As a result, we can define the filter link function implicitly by $\Psi = F(f_c, B_w) = 0$. The relation between f_c and B_w is given by the following equation:

$$\begin{aligned} f_c &= \kappa B_w \\ \text{or } B_w &= \kappa f_c, \end{aligned} \tag{33}$$

where κ is a proportionality constant.

5. Scintillation Indices Computation

This section presents the methods used to compute the amplitude scintillation index S_4 and phase scintillation index σ_ϕ, that is, to suppress all contributions but the scintillation effects in the signal variance. Following the mathematical formulation and notation suggested by Kaplan et al. [47, chap.

6.4], a simplified model for a received signal in the absence of scintillation is

$$r(t) = \sqrt{2P}s(t)\cos(\omega t + \phi) + n(t), \quad (34)$$

where P is the received signal power, $r(t)$ is the observed signal, $s(t)$ is the normalized transmitted signal, $n(t)$ is the signal noise, t is the time, ω is the carrier frequency, and ϕ is a phase offset. Scintillation causes a perturbation to both the received signal amplitude and the phase. Thus, in the presence of scintillation, the model may be extended as

$$r(t) = \sqrt{2P\delta P}s(t)\cos(\omega t + \phi + \delta\phi) + n(t), \quad (35)$$

where δP is a positive number that parametrizes signal attenuation, while $\delta\phi$ represents fluctuations in the phase offset.

Figure 4 shows the steps necessary to compute the indices discussed above.

The first step in the processing chain is the data cleaning, which includes detecting and pruning outliers and handling discontinuities caused by cycle-slips and missing observations in the data set. Next, we have to detrend the data set. This means that we construct a time series of the received signal, estimate the trend, and remove this from the signal. This will remove disturbances due to, for example, the Doppler shift from the relative motion of satellite and receiver. Next, we pass the signal through either a high-pass or low-pass filter, in order to isolate the scintillation components in the signal. Finally, the last step is the computation of the scintillation indices, which is the interesting part.

5.1. Filtering Techniques. Using digital filters to extract specific signal components is an important technique in many practical applications. For instance, we need the low-frequency components of the signal strength to calculate the amplitude scintillation S_4 and the high-frequency components of the signal phase to calculate σ_ϕ.

There are a variety of methods that can be used to design the digital filters. One commonly used method is the bilinear transform. The method uses the analog filter approximation functions that have been developed and translate the filter coefficients in a way that will make them usable for discrete systems, the IIR filter. That is, the output of the filter will depend on the previous values of the output as well as on the past and the current values of the input. Such a process is known as the autoregressive moving average (ARMA) process.

Other methods use frequency response of the desired filter to directly determine the digital filter coefficients. These are known as Finite Impulse Response (FIR) filters and are implemented as moving average filters. Once the FIR filter coefficients have been determined, the output signal is simply the convolution of the input signal with the filter coefficients.

One efficient way to implement this approach computationally is that the convolution is performed through multiplication in the frequency domain. This is done by first using a fast Fourier transform (FFT) and then multiplying the filter transfer function with the input signal and performing the inverse transform. This is described by Algorithm 1, which assumes that the filter coefficients have already been determined.

Figure 4: Building blocks of scintillation indices computation.

Algorithm 1 (filtering with FFT convolution). (1) Apply the Fourier transform to the filter coefficients $h(n)$:

$$H(z) = \mathrm{FFT}[h(n)]. \quad (36)$$

(2) Apply the Fourier transform to the input signal $x(n)$:

$$X(z) = \mathrm{FFT}[x(n)]. \quad (37)$$

(3) Multiply the complex sequences elementwise:

$$Y(z) = H(z)X(z). \quad (38)$$

(4) Apply the inverse Fourier transform to transform the filtered function back to its original domain:

$$y(n) = \mathrm{FFT}^{-1}[Y(z)]. \quad (39)$$

(5) Normalization: The IFFT coefficients are the complex conjugates of the Discrete Fast Transform (DFT) coefficients. The output signal has to be scaled down by the length of the signal.

5.2. Computation of the Amplitude Scintillation Index S_4. The amplitude scintillation index S_4 is defined as the standard deviation of the signal power normalized to the mean value over the interval of interest, usually a 1-minute interval.

In order to compute this index, the necessary steps are as follows:

(I) Computation of the total amplitude index S_{4_T}. The index S_{4_T} is defined by the expression

$$S_{4_T} = \sqrt{\frac{\langle SI^2\rangle - \langle SI\rangle^2}{\langle SI\rangle^2}}, \quad (40)$$

where SI is the detrended power or satellites signal intensity.

(II) Correction of S_4 by removing the ambient noise: S_{4_T} can have large values due to the ambient noise, and this must be removed. The removal process is done by estimating the average carrier-to-noise ratio C/N_0 density over the entire evaluation interval and using the estimate to compute S_4 due the ambient noise. If the signal-to-noise (S/N) density is known, the predicted S_{4_N} due to the ambient noise [48] is given by the expression

$$S_{4_N} = \sqrt{\frac{100}{S/N}\left[1 + \frac{500}{19\,S/N}\right]}. \quad (41)$$

Replacing S/N with the 60-second estimate $\widehat{S}/\widehat{N}$, we obtain an estimate of S_4 due to the noise S_{4_N}. Subtracting the square of this value from the square of (40) yields the revised estimate of S_4:

$$S_4 = \sqrt{\frac{\left\langle SI^2 \right\rangle - \langle SI \rangle^2}{\langle SI \rangle^2} - \frac{100}{\widehat{S}/\widehat{N}}\left[1 + \frac{500}{19\,\widehat{S}/\widehat{N}}\right]}. \quad (42)$$

(III) Detrending the signal to achieve stationarity: Detrending the raw signal intensity is accomplished by dividing the original time series by the filtered one. The effect of this is the removal of the low-frequency variance. The ratio method is attractive because it is dimensionless; however, it cannot be used if the data contains negative values, and it can also become problematic if the fitted line crosses zero.

Detrending by the difference method produces stationarity of the signal under study. The first difference $\nabla x_t = x_t - x_{t-1}$ eliminates a linear trend, while the second difference $\nabla^2 x_t = x_t - 2x_{t-1} + x_{t-2}$ can eliminate a quadratic trend and so on. In this study, both methods are used and they produce almost the same result; the first difference method has a little higher range than the ratio method.

In order to classify the scintillation activity indicated by S_4, four categories [49] are defined: $S_4 \leq 0.25$ is quiet, $S_4 \in (0.25, 0.50]$ is moderate, $S_4 \in (0.50, 1.00]$ is disturbed, and $S_4 > 1.00$ is severe.

(IV) Practical considerations: A full understanding of the difference between S/N_0 and C/N_0 is necessary and useful for the users of the GNSS receivers. $S/N_0$is usually expressed in dB; it refers to the ratio of the signal power and noise power in the given bandwidth while C/N_0 is expressed in dB-Hz; it refers to carrier power and the noise power per unit bandwidth.

The signal strength indicator measured by the receiver is C/N_0, the carrier-to-noise density measured in dB-Hz. In the computation of S_{4_T}, the signal-to-noise density S/N_0 is used. The relation between these two quantities [50] is given by the expression

$$S/N_0 = 10^{C/10N_0}. \quad (43)$$

5.3. Computation of Scintillation Index σ_ϕ. Computation of the phase scintillation σ_ϕ is straightforward, as the procedure is very similar to the computation of S_4. The main difference exists in data filtering; for σ_ϕ high-pass filter, usually a Butterworth filter with cutoff frequency of 0.3 Hz is used while S_4 is as mentioned in Section 3.2.

For typical FIR, IIR, and FFT convolution, the filter transfer function given by (1) is used. The sign in the exponent is negative, which results in the high-pass behavior. Detrending is a necessary step before carrying out the filtering. This is accomplished by any method that is able to produce a stationary time series. In this study, the ratio method is used for data detrending. The next step is filtering. Algorithm 1 described previously is used.

For the case of the new approach, the details are given in Section 4.3 and are summarized as follows:

(I) Preprocessing step requires handling outliers, missing observations, and cycle-slip detection and repair.

(II) Choose a kernel density function to compute the weights.

(III) Model selection with AIC/AICC is to determine the optimal smoothing parameter h_{opt}. The key is h_{opt}.

(IV) Compute the weights based on the height of the chosen kernel centered about the point of interest.

(V) locally fit a pth degree ($p = 0, 1, or\ 2$) polynomial to the data by using the weighted least square technique.

(VI) Compute σ_ϕ by taking the standard deviation of the residuals.

6. Statistical Analysis

The main objective of this section is to define the distributions of phase and amplitude scintillation indices (S_4 and σ_ϕ), correlation, and sensitivity analysis of the implemented filtering algorithms. Focus will be directed to the new approach, the nonparametric local regression with smoothing parameter selection. Outliers detection and repair, discontinuities, and missing observations are handled in data preparation.

6.1. Statistical Distribution of the Scintillation Index S_4. The power fluctuations of δP given in (35) are generally modeled as a Nakagami m-distribution. The probability density function (PDF) of this distribution is given by (B.1) with mean value of one and variance $1/m$.

Due to the characteristics of the Nakagami distributions, $S_4 = \sqrt{1/m}$ will not exceed $\sqrt{2}$ [47, p. 296]. Figure 5 shows the computed S_4 index for an arbitrary GPS satellite and Figure 6 shows the distribution of S_4.

6.2. Distribution of σ_ϕ. This subsection is devoted to the computation of phase scintillation σ_ϕ. Figure 7 shows the computed phase scintillation index for GPS satellite. Clearly, the heavy tailed distribution family is the appropriate choice. Here, we found that Frechet distribution equation (B.5) models the empirical data very well.

The motivations for considering the log-normal as an approximation model of the distribution of σ_ϕ are as follows: First, the log-normal distribution is a positive real-valued function with a heavy tail that can describe the presence of extreme variability in the data. Secondly, the distribution is simple and practical and provides a very good fit to the data. Finally its variance is scaling.

Figure 8 shows the distribution of phase scintillation indices for an arbitrary GLONASS satellite, PRN 08. Clearly, the distribution follows the heavy tailed distribution (the Frechet and log-normal ones are appropriate fit).

6.3. Correlation Analysis. Often, we like to measure the linear predictability of one signal x_t from another one y_t. Assuming

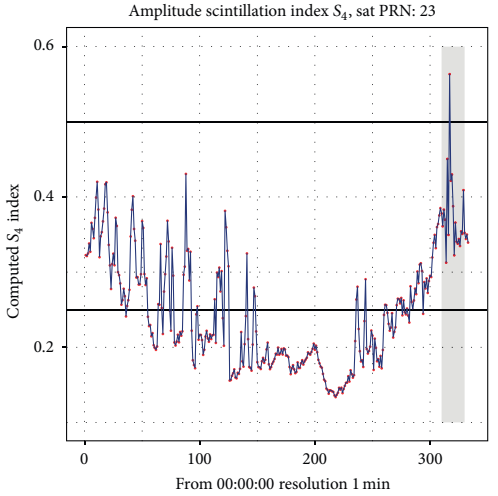


FIGURE 5: Amplitude scintillation index, S_4 computed for GPS satellite PRN 23, year 2014 and DOY 50, site Tromso. False scintillation(s) due to low elevation angle(s) is (are) given by shadowed rectangle in the figure.

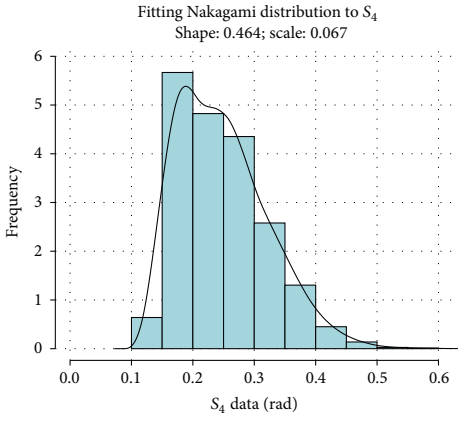


FIGURE 6: Distribution of the amplitude scintillation index S_4 computed for GLONASS satellite PRN 13, year 2014 and DOY 50, observed at site Tromso.

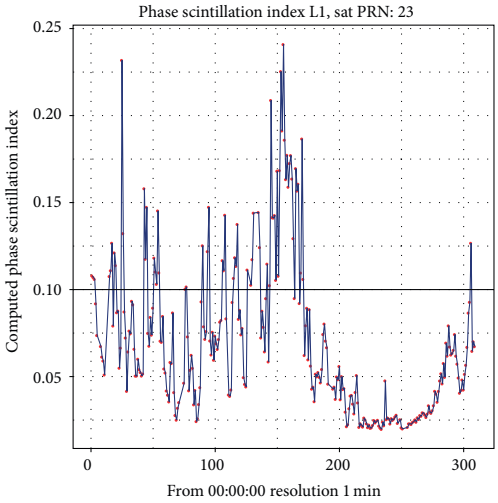


FIGURE 7: Phase scintillation index, σ_ϕ for GPS satellite PRN 23, year 2014 and DOY 50, site Tromso.

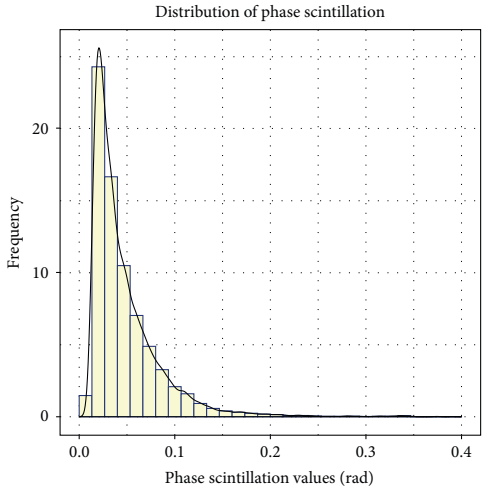


FIGURE 8: Empirical distribution of phase scintillation for an arbitrary GLONASS satellite, PRN 08. Heavy tailed Frechet distribution is an appropriate fit to phase scintillation.

that the variance of both time series is finite, the cross-correlation function [51, p.23] can be used for this purpose. This function is defined as

$$\rho_{xy}(s,t) = \frac{\gamma_{x,y}(s,t)}{\sqrt{\gamma_x(t,t)\,\gamma_y(s,s)}}, \tag{44}$$

where $\gamma_{x,y}(s,t) = E[(x_t - \mu_{xt})(y_s - \mu_{ys})]$ is the cross-covariance function between two time series, x_t and y_s, respectively. $E(\cdot)$ is the expectation operator. The Cauchy-Schwartz inequality implies $|\gamma_{x,y}(s,t)|^2 \le \gamma_x(t,t)\gamma_y(s,s)$.

The phase scintillation index σ_ϕ is the dominant disturbance at high latitudes and is obtained by high-pass filtering. For this reason, we have chosen two implementations of high-pass filtering algorithms (FFT convolution and IIR) for the correlation analysis. All plots have in common the cutoff frequency $f_c = 0.3$ Hz and the averaging interval $\langle \overline{T} \rangle = 1$

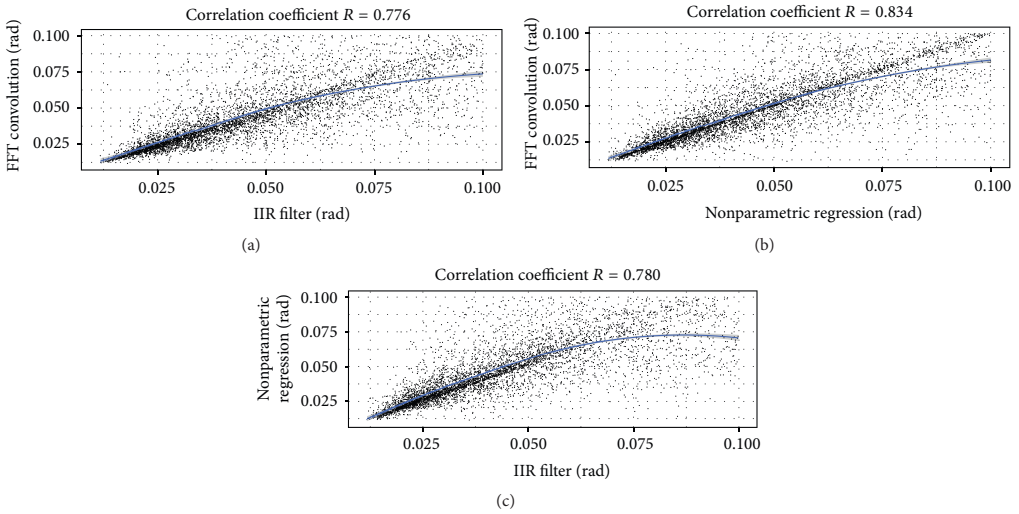


Figure 9: (a) shows the correlation between the Butterworth sixth-order filter and the IIR filter. (b) shows the correlation between the Butterworth sixth-order filter and an advanced nonparametric regression with kernel smoothing. Finally, (c) shows the correlation between the Butterworth sixth-order filter and nonparametric regression. The plots show high correlation between filtering algorithms, up to 83.4% for $\sigma_\phi \leq 0.1$ rad.

minute. Three scenarios are presented below, using different threshold values Th $\in \{0.1, 0.15, 0.25\}$ rad.

6.3.1. Scenario 1: $\sigma_\phi \leq 0.1$ rad. For this scenario, we drop all scintillations indices σ_ϕ above a threshold of 0.1 rad. Figure 9 shows that the algorithms are highly correlated, and the correlation coefficient ρ varies between 78% and 83%. The new approach, that is, the nonparametric regression with Akaike information criterion (AIC) [51, pp. 53-54] model selection, performs very well. The algorithm does not take into consideration any knowledge of physical problems, for instance, multipath, Doppler effects due to satellite motion, and cutoff frequency selection. The key is the smoothing parameter h_{opt} that determines the trade-off between the variance and the bias terms.

In this scenario, we have dropped all observations above the threshold value Th = 0.1 rad. These observations are classified as extreme and correspond to severe ionospheric activities.

6.3.2. Scenario 2: $\sigma_\phi \leq 0.15$ rad. Figure 10 shows that the algorithms are still highly correlated and the correlation coefficient ρ has decreased to values between 77.2% and 79.3%. The main reason for this reduction is the inclusion of more extremal events that are classified as severe ionospheric activities. Before the vertical lines indicated in Figure 10, the algorithms are highly correlated. After the lines, decorrelation appears between the two algorithms.

6.3.3. Scenario 3: $\sigma_\phi \leq 0.25$ rad. In this scenario, we have included all σ_ϕ values ($\sigma_\phi > 0.25$ rad are treated as outliers and removed). Figure 11 shows that the correlation coefficient has dropped to values between 69 and 74%. The reason is the inclusion of the extremal events classified as severe.

The vertical line is used to point out/distinguish two classes; the one to the left is highly correlated while the other is uncorrelated and corresponds to the extreme ionospheric activities and possibly noise in the signal.

6.4. Sensitivity Analysis. In order to carry out the sensitivity analysis of the implemented algorithms, a reference algorithm has to be chosen. The most common one used is high-pass filter Butterworth with FFT convolution, order 6, and cutoff frequency of 0.3 Hz.

Let the function $F = f(x, y, z)$ represents the phase scintillation index σ_ϕ, depending on parameters (x, y, z), where x and y are the cutoff frequency and the averaging interval parameters, respectively. $z = g(\alpha_1, \alpha_2, \ldots, \alpha_k)$ represents the remaining parameters, for instance, cycle-slip detection and repair and handling outliers/glitch. We will ignore the parameter z to facilitate analysis.

In order to carry out the sensitivity analysis, keeping all parameters constant and varying one parameter at a time are common. This is done by the following equations:

$$\left|\frac{\partial F}{\partial x}\right|_{y=\text{const}}, \quad \left|\frac{\partial F}{\partial y}\right|_{x=\text{const}}. \tag{45}$$

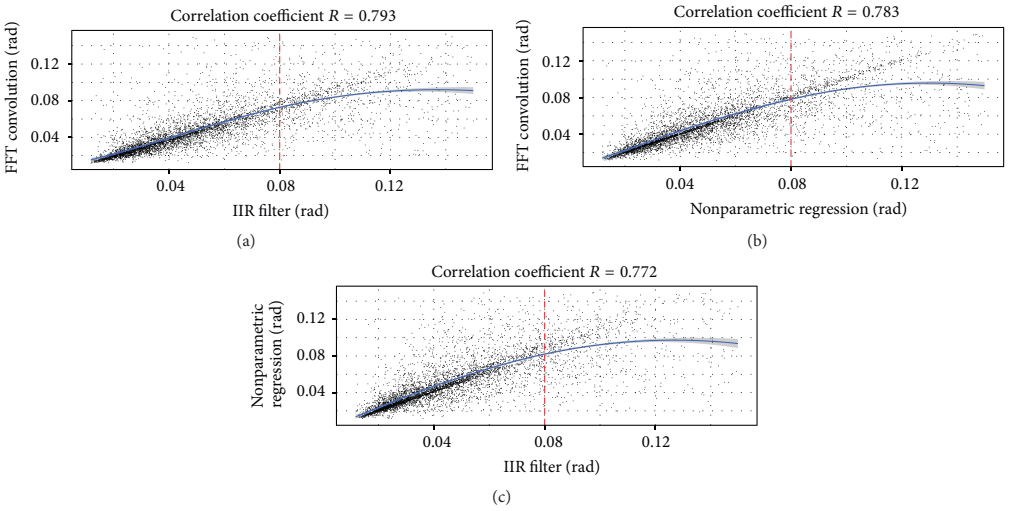


Figure 10: (a) shows the correlation between the Butterworth sixth-order filter and the IIR filter. (b) shows the correlation between the Butterworth sixth-order filter and an advanced nonparametric regression with kernel smoothing. Finally, (c) shows the correlation between the Butterworth sixth-order filter and nonparametric regression. The plots show high correlation between filtering algorithms, up to 78% for $\sigma_\phi \leq 0.15$ rad.

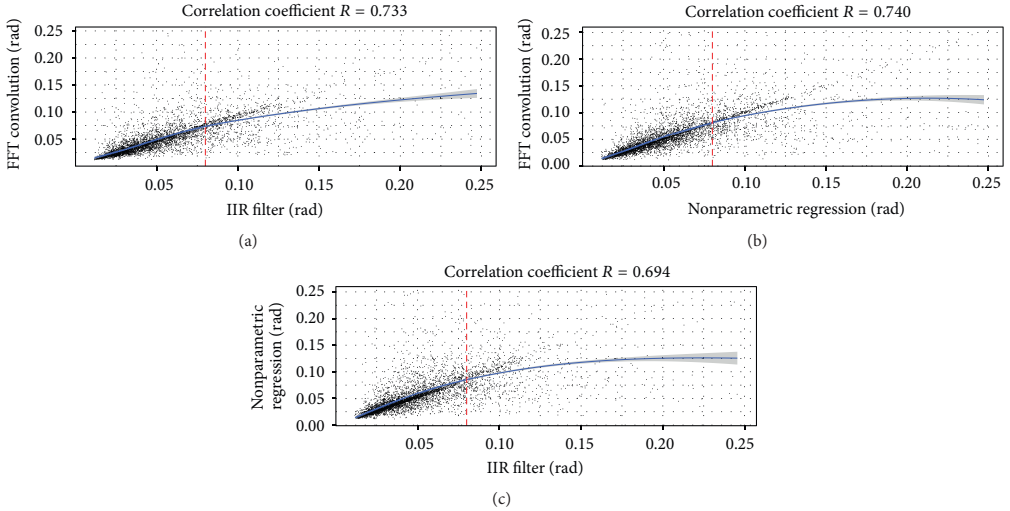


Figure 11: (a) shows the correlation between the Butterworth sixth-order filter and the IIR filter. (b) shows the correlation between the Butterworth sixth-order filter and an advanced nonparametric regression with kernel smoothing. Finally, (c) shows the correlation between the Butterworth sixth-order filter and nonparametric regression. The plots show high correlation between filtering algorithms, up to 74% for $\sigma_\phi \leq 0.25$ rad.

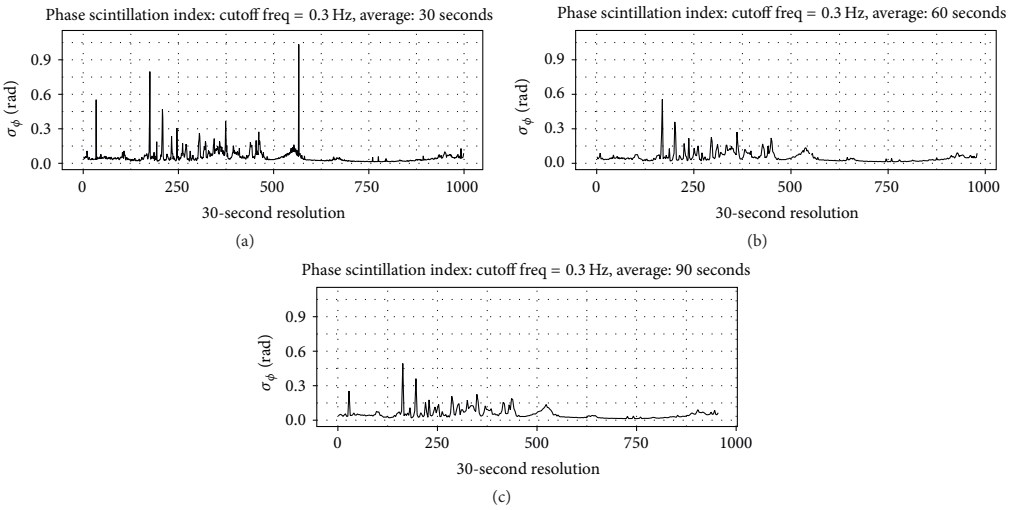


Figure 12: The panels show time series of σ_ϕ calculated from filtered signal phase, using different settings for the σ_ϕ calculation. The data is from GLONASS satellite 25, year 2014, DOY 50, site Honningsvag. (a), (b), and (c) have been calculated for each 30-second, 60-second, and 90-second interval, respectively.

6.4.1. Scenario 1: Cutoff Frequency Constant. In this scenario, we kept the cutoff frequency $f_c = 0.3$ Hz constant and varied the average interval $\langle \overline{T} \rangle \in \{30, 60, 90\}$ seconds.

Figure 12 shows the results. The general shape of the time series is very similar, but some of the modes/spikes are not present in all of them. In general using a longer averaging interval smooths the time series, broadening the spikes, reducing their magnitude, and causing some of the spikes to vanish. The longest interval time series does not appear to offer any significant advantage over the middle interval time series. The 30 sec interval time series contains more fine structure than the 60 sec interval time series. In conclusion, the averaging interval should not exceed 60 seconds.

6.4.2. Scenario 2: Averaging Interval Constant. In this scenario, the averaging interval is constant and equal to 60 seconds while varying the cutoff frequency $f_c \in \{0.2, 0.3, 0.4\}$ Hz.

Figure 13 shows the results. The general shape of the time series is very similar, but some of the modes/spikes are not present in all of them. In general using a higher cutoff frequency results in a reduction of the scintillation index value throughout the time series. At time ≈ 340, there is a large spike which is significantly reduced when changing the cutoff frequency from 0.2 to 0.3 Hz and vanishes for a cutoff frequency of 0.4 Hz. The phase scintillation values are clearly strongly dependent on the cutoff frequency, even for a modest change such as the one seen here. This is a known weakness in the standard techniques of σ_ϕ calculation (e.g., [52]).

Taking into account the first and the second sensitivity analysis, the acceptable range for the computation of σ_ϕ is $f_c \in [0.1, 0.3]$ and $\langle \overline{T} \rangle \in [30, 60]$. These ranges give the maximum variations in the data.

7. Implementation and Analysis

7.1. Data Collection and Analysis. As mentioned in the introduction, the data used in this investigation are obtained from the Norwegian Ionospheric Scintillation Network. The reference stations used are encircled as shown in Figure 1.

For implementation and software packages used to compute the scintillation indices S_4 and σ_ϕ, the interested reader is referred to Appendix C.

7.2. Interpretation of Results. The test results show that all detrending filtering techniques produce almost the same results. The main difference exists in statistical methods used in data editing. That is, how outliers are detected and removed, the missing observations, and glitch detection and removal.

The new approach is robust against outliers, discontinuity, and missing observations. Figure 14 shows how algorithms (FFT convolution, IIR, and local kernel regression) handle a discontinuity with a magnitude of 0.6 rad. The new approach handles perfectly the discontinuity and avoids generating a false scintillation due to discontinuity.

7.3. Remarks. Some remarks and suggestions are as follows:

(I) In order to capture false scintillation due to low satellite elevation angles, the cutoff elevation angle

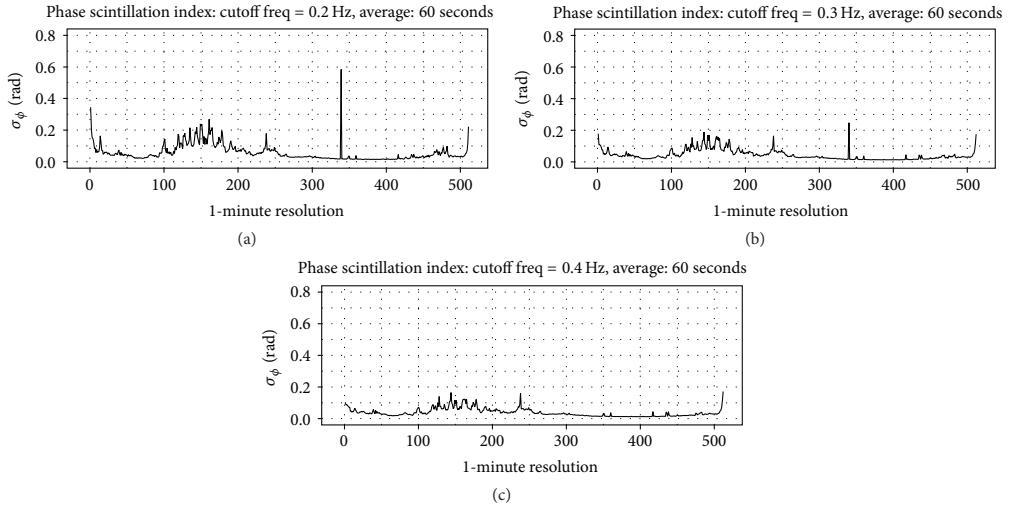


FIGURE 13: The panels show time series of σ_ϕ calculated from filtered signal phase, using different settings for the high-pass filter. The data is from GPS satellite 17, year 2014, DOY 50, site Honningsvag. (a), (b), and (c) have been calculated using a cutoff frequency of 0.2, 0.3, and 0.4 Hz, respectively.

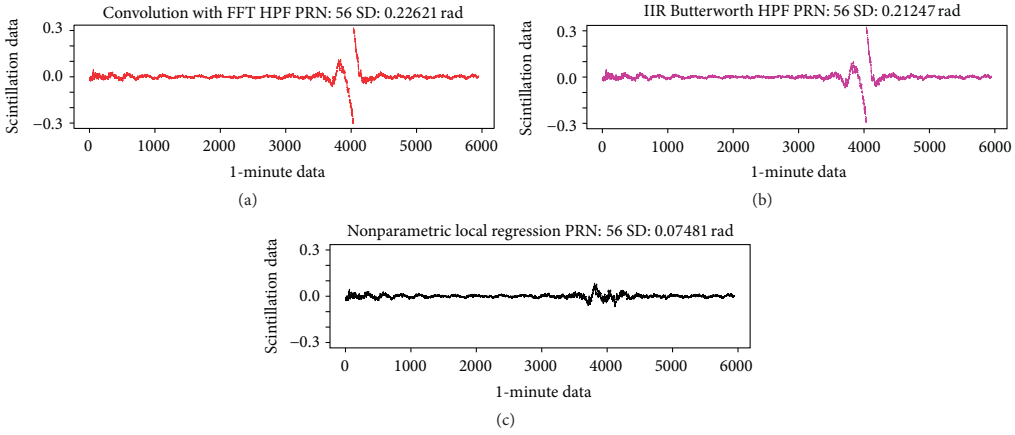


FIGURE 14: The panels show 1 minute of phase data that contains a discontinuity, filtered using different techniques. (a) has been filtered using FFT convolution filter. (b) has been filtered using a IIR Butterworth filter. (c) has been filtered using the new approach.

was set to zero. Figure 5 shows the events which are presented by shadowed rectangle.

(II) In this paper, we have computed scintillation indices by means of different detrending and filtering techniques. Identifying the weaknesses and the strengths of each method, we can compute a reliable index that can be used in further analysis, for instance, to carry out the classification of the level of the ionosphere disturbances.

(III) Nonparametric local regression with optimal smoothing parameters can be used as scintillation indices computation.

(IV) As can be seen in Figure 14, the new approach handles discontinuities very well, while the other filter techniques generate artificially high σ_ϕ values. The new approach computes $\sigma_\phi = 0.0748$ rad while the other algorithms compute values of 0.226 and 0.212 rad,

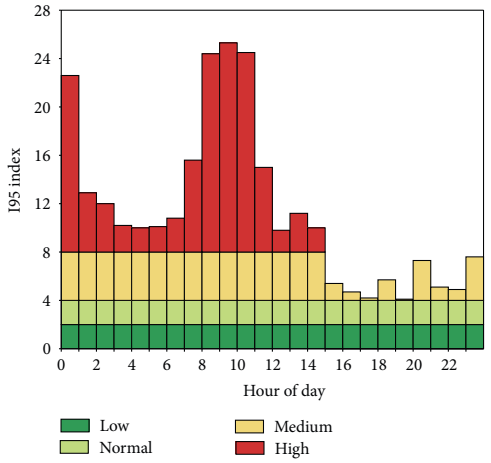


Figure 15: I95 index generated by the CPOS software for north of Norway.

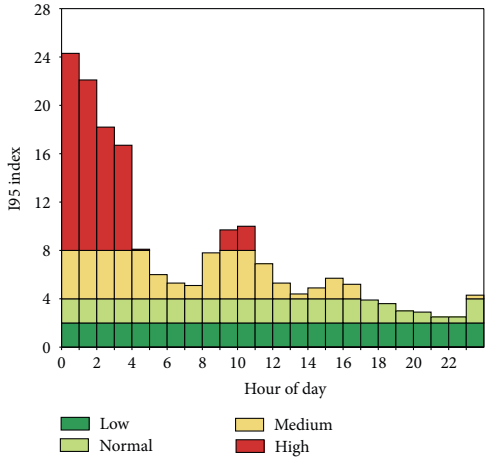


Figure 16: I95 index generated by the CPOS software for south of Norway.

respectively. Using the traditional techniques this discontinuity would be falsely reported as a scintillation event.

8. Conclusion

We have shown that the new approach based on local regression with kernel smoothing and with AIC/AICC for bandwidth parameter selection can be used for computation of scintillation indices for high latitude data. The new approach can be analytically related to the existing filtering methods and is shown to produce highly correlated results with the traditional approaches. However, the new approach shows superior handling of discontinuities.

We have shown that applying the right detrending and filtering techniques to the scintillation data at high latitude one can obtain reliable scintillation indices (S_4 and σ_ϕ). Clearly, the studies show that the phase scintillations are dominant for these data sets.

For the derivation of S_4, the digital filter FIR (Butterworth, Chebyshev, and Elliptic) and FFT convolution used to implement the low-pass filter work well with minor differences. For σ_ϕ, the digital filter IIR (Butterworth and Elliptic) and FFT convolution work well. The difference exists in statistical methods used to compute standard deviations and how outliers, glitches, and missing observations are handled. Poor methods can bias the estimation process. In addition, the study shows that the derivation of scintillation indices S_4 and σ_ϕ is sensitive to averaging interval and the cutoff frequency. Figure 5 shows the false scintillation for S_4. In this case, all elevation angles are used under computation.

For the data sets classified as high ionospheric activity, we have defined the empirical distribution of the scintillation indices S_4 and σ_ϕ for the Norwegian Regional Ionospheric Scintillation Network, located at high latitude (61.99°N to 70.98°N). Heavy tailed Frechet/log-normal distribution (B.5) is a good model for phase scintillation σ_ϕ index as Figure 8 confirms. The distribution of S_4 follows the Nakagami distribution, given by (B.1) and shown by Figure 6.

Appendix

A. Ionosphere Classification

The Index 95 values are computed by the Trimble software running the CPOS network and is used to classify the ionosphere activity. I95 classification is defined as follows:

(i) $[0, 2]$: ionospheric disturbances have no influences.

(ii) $(2, 4]$: ionospheric disturbances are classified as weak.

(iii) $(4, 8]$: ionospheric disturbances are classified as strong.

(iv) $(8, \infty)$: ionospheric disturbances are classified as severe.

Figures 15 and 16 show the ionosphere activities for the data sets analyzed (y: 2014, DOY: 50) which are classified as strong to severe.

B. Empirical Distributions of S_4 and σ_ϕ

This section gives some distributions employed in communication systems in order to characterize the statistics of the signals transmitted through multipath fading channels. These distributions are relevant for scintillation data as well.

B.1. The Nakagami m-Distribution. A random variable X is said to have a Nakagami m-distribution if, for some $m \geq 1/2$ and $\nu \geq 0$, its probability density function (pdf) is given by

$$f_X(x) = \frac{2}{\Gamma(m)} \left(\frac{m}{\nu}\right)^m x^{(2m-1)} e^{-(mx^2)/\nu}, \quad \text{(B.1)}$$

where ν is the mean square value of X and is defined as $\nu = E(X^2)$ and $E(\cdot)$ is the expectation operator. For the parameter m, the ratio of the moments is defined by $m = \nu^2/E(X-\nu)^2$.

For $m = 1$, we obtain the Rayleigh density function, given by expression (B.2).

B.2. Rayleigh Distribution. A random variable X is said to have a Rayleigh distribution if its probability density function (pdf) is given by

$$f_X(x) = \frac{x}{\sigma^2} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}, \quad x \geq 0. \quad \text{(B.2)}$$

B.3. The Log-Normal Distribution. A random variable X is said to have a log-normal distribution if its probability density function (pdf) is given by

$$f_X(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\dfrac{\ln^2(x/x_m)}{2\sigma}\right\}, & x \geq 0, \\ 0, & \text{otherwise}, \end{cases} \quad \text{(B.3)}$$

where x_m is the median of X and σ^2 is the variance. The log-normal distribution is characterized by the parameter ρ known as the mean-to-median ratio and is given by $\rho = E(X)/X_m$.

Alternatively, log-normal pdf can be written as

$$f_X(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\dfrac{(\ln x - \ln x_m)^2}{2\sigma}\right\}, & x \geq 0, \\ 0, & \text{otherwise}. \end{cases} \quad \text{(B.4)}$$

B.4. Frechet Distribution. A random variable X is said to have a Frechet distribution if for parameter $\theta = (a, b, \alpha)$, respectively, defining the scale, location, and the shape parameters, the distribution is given by the expression

$$f_X(x) = \begin{cases} 0, & x \leq b, \\ \exp\left\{-\left(\dfrac{x-b}{a}\right)^{-\alpha}\right\}, & x > b. \end{cases} \quad \text{(B.5)}$$

C. Implementation and Software Packages

The software used to generate the scintillation indices S_4 and σ_ϕ is implemented in the programming languages C, C++, and R. All programs are configurable and generate a log-file to report all events under processing.

(1) C++ configurable process decodes Septentrio PolaRxS message types 4027 (measurements blocks) and 4046 (correlation values) and produces a suitable matrix format that is easy to process with Matlab, Python, or R. The decoded messages are well described in the SBF reference guide [53].

(2) Convolution with FFT: C++ program reads raw scintillation data in matrix format and computes the indices S_4 and σ_ϕ. This module includes the detrending methods varying between fitting a straight line, ratio method, and more advanced techniques such as nonparametric regression with information criteria for model selection.

(3) Median filter: C++ program removes the glitches (spikes) from scintillation indices data.

(4) Digital filters: R package, `signal`, is downloaded from the Comprehensive R Archive Network (CRAN). The package is used to implement FIR (Butterworth, Elliptic, and Chebyshev) and IIR (Butterworth, Elliptic) filters.

(5) Outliers: R package `outliers` downloaded from CRAN is used to detect and remove outliers.

(6) Local regression: R package `locfit` downloaded from CRAN is used to carry out filtering via nonparametric regression with AIC for model selection.

(7) Nonparametric analysis of covariance for regression: R package `fANCOVA` downloaded from CRAN is used to carry out filtering via nonparametric regression with AICC for model selection.

Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

Acknowledgments

This work is funded by the Norwegian Mapping Authority (NMA), part of Ph.D. program and cooperation with the Norwegian University of Life Sciences (NMBU). The authors would also like to thank Anna B. O. Jensen and Jabir A. Ouassou for their constructive discussions and proofreading of this paper.

References

[1] S. Basu, S. Basu, E. MacKenzie, W. R. Coley, J. R. Sharber, and W. R. Hoegy, "Plasma structuring by the gradient drift instability at high latitudes and comparison with velocity shear driven processes," *Journal of Geophysical Research: Space Physics*, vol. 95, no. 6, pp. 7799–7818, 1990.

[2] J. S. Hey, S. J. Parsons, and J. W. Phillips, "Fluctuations in cosmic radiation at radio-frequencies," *Nature*, vol. 158, no. 4007, p. 234, 1946.

[3] P. M. Kintner, B. M. Ledvina, and E. R. de Paula, "GPS and ionospheric scintillations," *Space Weather*, vol. 5, no. 9, 2007.

[4] M. Aquino, F. S. Rodrigues, J. Souter, T. Moore, A. Dodson, and S. Waugh, "Ionospheric scintillation and impact on GNSS users in Northern Europe: results of a 3 year study," *Space Communications*, vol. 20, no. 1-2, pp. 17–29, 2005, http://dl.acm.org/citation.cfm?id=1498858.1498862.

[5] S. Skone, M. Feng, R. Tiwari, and A. Coster, "Characterizing ionospheric irregularities for auroral scintillations," in *Proceedings of the 22nd International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS'09)*, pp. 2448–2455, Savannah, Ga, USA, September 2009.

[6] R. Tiwari, F. Ghafoori, O. Al-Fanek, O. Haddad, and S. Skone, "Investigation of high latitude ionospheric scintillation observed in the Canadian region," in *Proceedings of the 23rd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS '10)*, pp. 349–360, Portland, Ore, USA, September 2010.

[7] M. Aquino, A. Dodson, J. Souter, and T. Moore, "Ionospheric scintillation effects on GPS carrier phase positioning accuracy at auroral and sub-auroral latitudes," in *Dynamic Planet, International Association of Geodesy Symposia*, P. Tregoning and C. Rizos, Eds., vol. 130, pp. 859–866, Springer, Berlin, Germany, 2007.

[8] K. S. Jacobsen and S. Schäfer, "Observed effects of a geomagnetic storm on an RTK positioning network at high latitudes," *Journal of Space Weather and Space Climate*, vol. 2, article A13, 2012.

[9] J. Seo, T. Walter, and P. Enge, "Availability impact on GPS aviation due to strong ionospheric scintillation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 3, pp. 1963–1973, 2011.

[10] P. J. Teunissen, N. F. Jonkman, and C. C. Tiberius, "Weighting GPS dual frequency observations: bearing the cross of cross-correlation," *GPS Solutions*, vol. 2, no. 2, pp. 28–37, 1998.

[11] H.-J. Eueler and C. C. Goad, "On optimal filtering of GPS dual frequency observations without using orbit information," *Bulletin Géodésique*, vol. 65, no. 2, pp. 130–143, 1991.

[12] C. C. J. M. Tiberius, *Recursive Data Processing for Kinematic GPS Surveying*, Newtherlands Geodetic Commission, 1998.

[13] A. Wieser and F. K. Brunner, "An extended weight model for GPS phase observations," *Earth, Planets and Space*, vol. 52, no. 10, pp. 777–782, 2000.

[14] L. Alfonsi, L. Spogli, G. De Franceschi et al., "Bipolar climatology of GPS ionospheric scintillation at solar minimum," *Radio Science*, vol. 46, no. 3, 2011.

[15] Y. Bëniguel, V. Romano, L. Alfonsi et al., "Ionospheric scintillation monitoring and modelling," *Annals of Geophysics*, vol. 52, no. 3-4, pp. 391–416, 2009.

[16] B. Forte, M. Materassi, L. Alfonsi, V. Romano, G. De Franceschi, and P. Spalla, "Optimum parameter for estimating phase fluctuations on transionospheric signals at high latitudes," *Advances in Space Research*, vol. 47, no. 12, pp. 2188–2193, 2011.

[17] T. W. Garner, R. B. Harris, J. A. York, C. S. Herbster, C. F. Minter, and D. L. Hampton, "An auroral scintillation observation using precise, collocated GPS receivers," *Radio Science*, vol. 46, no. 1, Article ID RS1018, 2011.

[18] A. K. Gwal and A. Jain, "GPS scintillation studies in the arctic region during the first winter-phase 2008 Indian Arctic Expedition," *Polar Science*, vol. 4, no. 4, pp. 574–587, 2011.

[19] Y. Jiao, Y. T. Morton, S. Taylor, and W. Pelgrum, "Characterization of high-latitude ionospheric scintillation of GPS signals," *Radio Science*, vol. 48, no. 6, pp. 698–708, 2013.

[20] Y. Jin, J. I. Moen, and W. J. Miloch, "GPS scintillation effects associated with polar cap patches and substorm auroral activity: direct comparison," *Journal of Space Weather and Space Climate*, vol. 4, article A23, 2014.

[21] J. Kinrade, C. N. Mitchell, P. Yin et al., "Ionospheric scintillation over Antarctica during the storm of 5-6 April 2010," *Journal of Geophysical Research: Space Physics*, vol. 117, no. 5, 2012.

[22] G. Li, B. Ning, Z. Ren, and L. Hu, "Statistics of GPS ionospheric scintillation and irregularities over polar regions at solar minimum," *GPS Solutions*, vol. 14, no. 4, pp. 331–341, 2010.

[23] C. N. Mitchell, L. Alfonsi, G. De Franceschi, M. Lester, V. Romano, and A. W. Wernik, "GPS TEC and scintillation measurements from the polar ionosphere during the October 2003 storm," *Geophysical Research Letters*, vol. 32, no. 12, pp. 1–4, 2005.

[24] B. Forte, "Optimum detrending of raw GPS data for scintillation measurements at auroral latitudes," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 67, no. 12, pp. 1100–1109, 2005.

[25] Y. L. Andalsvik and K. S. Jacobsen, "Observed high-latitude GNSS disturbances during a less-than-minor geomagnetic storm," *Radio Science*, vol. 49, no. 12, pp. 1277–1288, 2014.

[26] K. S. Jacobsen, "The impact of different sampling rates and calculation time intervals on ROTI values," *Journal of Space Weather and Space Climate*, vol. 4, p. A33, 2014.

[27] X. Pi, A. J. Mannucci, U. J. Lindqwister, and C. M. Ho, "Monitoring of global ionospheric irregularities using the Worldwide GPS Network," *Geophysical Research Letters*, vol. 24, no. 18, pp. 2283–2286, 1997.

[28] A. Ahmed, R. Tiwari, H. J. Strangeways, S. Dlay, and M. G. Johnsen, "Wavelet-based analogous phase scintillation index for high latitudes," *Space Weather*, vol. 13, no. 8, pp. 503–520, 2015.

[29] N. Jakowski, C. Borries, and V. Wilken, "Introducing a disturbance ionosphere index," *Radio Science*, vol. 47, no. 4, 2012.

[30] S. C. Mushini, P. T. Jayachandran, R. B. Langley, J. W. MacDougall, and D. Pokhotelov, "Improved amplitude- and phase-scintillation indices derived from wavelet detrended high-latitude GPS data," *GPS Solutions*, vol. 16, no. 3, pp. 363–373, 2012.

[31] R. Tiwari, A. Ahmed, H. J. Strangeways et al., "Auroral monitoring index using a network of GNSS receivers," in *Proceedings of the 28th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS '15)*, pp. 3779–3786, Tampa, Fla, USA, 2015.

[32] R. Tiwari, H. J. Strangeways, S. Tiwari, and A. Ahmed, "Investigation of ionospheric irregularities and scintillation using TEC at high latitude," *Advances in Space Research*, vol. 52, no. 6, pp. 1111–1124, 2013.

[33] B. Forte, "On the relationship between the geometrical control of scintillation indices and the data detrending problems observed at high latitudes," *Annals of Geophysics*, vol. 50, no. 6, pp. 699–706, 2007.

[34] M. Materassi and C. N. Mitchell, "Wavelet analysis of GPS amplitude scintillation: a case study," *Radio Science*, vol. 42, no. 1, Article ID RS1004, 2007.

[35] J.-L. Starck, F. D. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis*, Cambridge University Press, 1998.

[36] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bulletin of the American Meteorological Society*, vol. 79, no. 1, pp. 61–78, 1998.

[37] L. Wanninger, "Ionospheric disturbance indices for RTK and network RTK positioning," in *Proceedings of the 17th International Technical Meeting of the Satellite Division of The Institute*

of Navigation (ION GNSS '04), pp. 2849–2854, LongBeah, Calif, USA, September 2004, http://www.wasoft.de/lit/ion04f64.pdf.

[38] J. M. De Freitas, *Digital Filter Design Solutions*, Artech House Microwave Library, Artech House, 2005.

[39] P. Embree and D. Danieli, *C++ Algorithms for Digital Signal Processing*, Pearson Education, 1998.

[40] R. Kuc, *Introduction to Digital Signal Processing*, BS Publications, Hyderabad, India, 1982.

[41] L. D. Thede, *Practical Analog and Digital Filter Design*, Artech House Microwave Library, Artech House, Norwood, Mass, USA, 2005.

[42] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, Dover Books on Mathematics, Dover, 2012, https://books.google.no/books?id=KiPCAgAAQBAJ.

[43] M. P. Wand and M. C. Jones, *Kernel Smoothing*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1994, https://books.google.no/books?id=GTOOi5yE008C.

[44] G. Claeskens and N. L. Hjort, *Model Selection and Model Averaging*, vol. 27 of *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, New York, NY, USA, 2008.

[45] W. Härdle, *Applied Nonparametric Regression*, Econometric Society Monographs, Cambridge University Press, 1990, https://books.google.no/books?id=8TprPX-7PSMC.

[46] E. Kaplan and C. Hegarty, *Understanding GPS, Principles and Applications*, Artech House Mobile Communications Series, Artech House, 2nd edition, 2005.

[47] E. Kaplan, C. Hegarty, and G. P. S. Understanding, *Principles and Applications*, Artech House Mobile Communications Series, Artech House, 2nd edition, 2005.

[48] A. J. Van Dierendonck, J. Klobuchar, and Q. Hua, "Ionospheric scintillation monitoring using commercial single frequency C/A code receivers," in *Proceedings of the 6th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS '93)*, pp. 1333–1342, September 1993.

[49] M. Abdullah, A. F. M. Zain, Y. H. Ho, and S. Abdullah, "TEC and scintillation study of equatorial ionosphere: a month campaign over Sipitang and Parit Raja stations, Malaysia," *American Journal of Engineering and Applied Sciences*, vol. 2, no. 1, pp. 44–49, 2009.

[50] A. Joseph, *Measuring GNSS Signal Strength*, GNSS Solutions, 2010.

[51] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications: With R Examples*, Springer Texts in Statistics, Springer, New York, NY, USA, 2nd edition, 2006.

[52] B. Forte and S. M. Radicella, "Problems in data treatment for ionospheric scintillation measurements," *Radio Science*, vol. 37, no. 6, pp. 8-1–8-5, 2002.

[53] S. Team, *SBF Reference Guide, Version 1*, Septentrio Satellite Navigation, 2012.

.

Paper #4

Reference

M. Ouassou, A. B. O. Jensen & J. G. O. Gjevestad
Network real-time kinematic data integrity by means of multivariate statistical analysis
Submitted to the International Journal of Navigation and Observation: 06/11/2016

Contribution

M. Ouassou conceived the idea, performed the numerical calculations, and wrote the majority of the manuscript. All authors contributed to the discussion of the results and revisions of the manuscript.

J.G.O. Gjevestad involved in discussion, support and conclusion.

A.B.O. Jensen wrote the majority of section 1 and involved in the validation process.

J.A. Ouassou are also acknowledged for useful discussions, feedback, and proofreading.

The main core of this dissertation. Quality check of the correctness of the information provided by NRTK systems to the users are monitored. The average corrections and variances fields are determined from data. Global and local tests statistics are constructed, and variance-covariance matrices are checked for Heywood case.

The carrier-phase and code statistics are predicted by employing the stochastic generalized linear model (SGLM).

The NRTK data integrity algorithms are designed, implemented and tested. The rover receiver's position will be improved by applying the NRTK data integrity.

International Journal of Navigation and Observation manuscript No.
(will be inserted by the editor)

Network real-time kinematic data integrity by means of multivariate statistical analysis

Mohammed Ouassou · Anna B. O. Jensen · Jon G. O. Gjevestad

Received: date / Accepted: date

Abstract We introduce a novel approach to the computation of network real-time kinematic (NRTK) data integrity, which can be used to improve the position accuracy for a rover receiver in the field. Our approach is based on multivariate statistical analysis and Stochastic Generalized Linear Model (SGLM). The network average error corrections and the corresponding variance fields are computed from the data, while the squared Mahalanobis distance (SMD) and Mahalanobis depth (MD) are used as test statistics to detect and remove data from satellites that supply inaccurate data. The covariance matrices are also inspected and monitored to avoid the Heywood effect, i.e. negative variance generated by the processing filters. The quality checks were carried out at both the system and user levels in order to reduce the impact of extreme events on the rover position estimates. The SGLM is used to predict the user carrier-phase and code error statistics. Finally, we present analyses of real-world data sets to establish the practical viability of the proposed methods.

Keywords network RTK (NRTK) · integrity monitoring · multivariate statistics.

Mohammed Ouassou
Norwegian Mapping Authority, Geodetic Institute
3511 Hønefoss, Norway
Tel.: +47 32 11 83 76
E-mail: mohammed.ouassou@statkart.no

Anna B. O. Jensen
KTH Royal Institute of Technology 10044 Stockholm, Sweden.
Tel.: +46 8-790 7353
E-mail: anna.jensen@abe.kth.se

Jon G. O. Gjevestad
Dept. of Mathematical Sciences and Technology, NMBU
Tel.: +47 73 59 14 13
E-mail: jon.glenn.gjevestad@nmbu.no

1 Introduction

An integrity service is a set of procedures used to check the correctness of the information provided by a system. Such services are already implemented in safety of life navigation augmentation systems such as WAAS, EGNOS, GAGAN and others.

There are also other types of integrity algorithms, for instance GNSS receiver-based integrity monitoring known as Receiver Autonomous Integrity Monitoring (RAIM) and Fault Detection and Exclusion (FDE) algorithms (Grewal et al, 2007; Kaplan and Hegarty, 2006; Ramjee and Ruggieri, 2005). These algorithms identify satellites with bad observations using a least-squares method, and then exclude them from the solution. However, RAIM and FDE were developed as pseudo-range residual data analysis algorithms for GNSS safety-critical applications, such as e.g. the approach phase of flight. For high-accuracy applications, an extension of pseudo-range RAIM (PRAIM) known as carrier-phase based RAIM (CRAIM) was proposed by Feng et al (2009).

Data quality checks and integrity monitoring techniques has been a research topic for many years in geodesy, surveying and navigation. For instance, Baarda (1968) developed a test procedure for use in geodetic networks, which has been used to check data against outlying observations in many different applications, for instance the analysis of the deformation problem in geodesy (Kok, 1982). The DIA procedure (Teunissen, 1990) can be applied to any set of GNSS observation equations, such as the GPS quality control (Kleusberg and Teunissen, 1998), geodetic networks (Teunissen, 1985) or integrated navigation system (Teunissen, 1985). Another approach to error modeling is to perform a reliability and quality control procedure (Kuusniemi et al, 2007), using good statistical methods for the analysis (Leick, 2015).

In recent years, mobile phones have also emerged as a new market for GNSS applications. Quality control for

handset-based users is already in demand. For instance, Trimble introduced the CenterPoint RTX system, which offers real-time position estimation and coordinates integrity via a mobile app (Trimble pivot), including an analysis of the ionosphere activity and network status (Chen et al, 2011; Leandro et al, 2011).

The users of high accuracy GNSS NRTK positioning systems have requested the development of data integrity for a long time. In this article, we consider how such a service can be designed and implemented, which can be of interest to both the NRTK service providers and their users.

The NRTK processing chain can be summarized as follows. The first step is to collect raw measurements from the network of reference stations, solve for the ambiguities within the reference network, and generate error estimates. Then an interpolation/smoothing scheme is applied to generate the NRTK corrections for the user location. For information on how to avoid loss of information under interpolation of NRTK data, the interested reader is referred to Ouassou et al (2015). The NRTK corrections are then transmitted to users who can perform real-time positioning with an accuracy at the cm-level. Several NRTK techniques exist, and the most common used ones are the Master Auxiliary Concept (MAC) (Euler et al, 2001; Takac and Zelzer, 2008), the Virtual Reference Station (VRS) concept (Landau et al, 2002), and the FKP techniques (Wübbena et al, 1996). However, we limit the discussion in this paper to the Network Adjustment (NetAdjust) method developed by Raquet (1998).

Figure 1 shows the structure of the NRTK processing chain. The new data integrity segment (red box) is the main focus of this article. At the system level, the integrity service is driven by a three-step process, where the average correction field and associated variances are generated by constructing time series with a sliding window. The size of the sliding window is set to the correlation length, i.e. the time span for which the observations can be considered completely decorrelated. As described in Section 5, we use two Mahalanobis metrics (SMD and MD) to detect extremal events, and use the t-distribution as a local identification test. For adaptation, we can either send the satellite identities to the rover, or just ignore them and abstain from sending the corrections.

The reason for using this type of metrics is that when using the SMD approach, the explanatory observations are those that lie far from the bulk of the data. The computed metric values may then be compared with quantiles of the χ^2-distribution with $p-1$ degrees of freedom, where p is the number of common satellites used by the filters. Another important characteristic of the metric is that there exists a unique mapping to the diagonal of the prediction matrix shown in Eq. (14) (Rousseeuw and Leroy, 2003, p. 224). For more information about the properties and benefits of SMD-based approaches, please consult Rousseeuw and Leroy (2003); Somesh Dasgupta (1995); Timm (2007). MD-based approaches are similarly described by Djauhari and Umbara (2007); Liu et al (2003); Mosler (2013).

At the user level, the raw phase observations can be inspected to ensure that only high-quality observations are included in the analysis, and this can be accomplished using the Danish method (Wieser and Brunner, 2000). The main reason for choosing the Danish method is that ordinary least-squares methods are sensitive to outliers. Unfortunately, most estimators that are robust to outliers are only applicable to uncorrelated data sets, while e.g. double-difference carrier phase observables and network baseline vectors are examples of the abundant correlated observables in GNSS systems (Leick, 2015). However, a straight-forward solution to this problem is to decorrelate the original data set using e.g. a Mahalanobis transformation, and then apply well-known robust estimation methods for uncorrelated data to the results. Various such schemes exist that provide a certain resistance against outlying observations and reduce their influences on the estimation process. Additional benefits is that the method guarantees convergence, and can automatically locate and eliminate errors. For more information, see for instance Leick (2015).

Finally, the residuals of the baseline and corresponding variances are used to predict the position error. The focus is directed to the double-difference error covariance matrix, which will be used to construct the relevant prediction function. The covariance matrices at both the system and user levels are continuously inspected for Heywood cases (Heywood, 1931), i.e. anomalous generation of negative variance. The validation procedure is carried out by excluding all suspicious satellites from the position computation.

In order to evaluate our proposed integrity method, we use a data sample from the Norwegian GNSS network, which is described in detail in section 2. The NetAdjust method is briefly discussed in section 3. The architecture of the proposed integrity solution is then presented in section 4. After that, the network correction integrity is discussed in section 5, rover observation integrity in section 6, and relative positioning integrity in section 7. Finally, the implementation and analysis are presented in section 8, and the discussion and conclusion in section 9.

2 Test data

The data sample used to evaluate our proposed method was provided by the Norwegian RTK network known as CPOS, which is operated by the Norwegian Mapping Authority (NMA). The test area is from the Rogaland region in the south west of Norway. Reference receivers are equipped with *Trimble NetR9* receivers, tracking GPS and GLONASS satellite signals. Baselines vary between 35 to 112 km and the height difference between the sites is about 225 m. Tables 1, 2

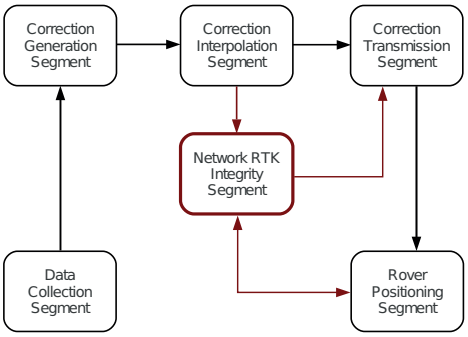


Figure 1: Extension of network real-time kinematics segments with a new service known as the NRTK data integrity segment (red rectangle).

and 3 give a description of sub-network while Figure 2 shows the location of reference receivers.

The NRTK software modules are executed at the same rate, of one second interval. Once every ten seconds, the network modules generate the user corrections. The updating rate was chosen intentionally and corresponds to the optimal update rate of the network corrections, dispersive and non-dispersive, respectively. The module can be interpreted as a discrete event model. The user position is computed once every second.

Many tests have been carried out in this research. For the manuscript, we have used data from DOY 152, 2013 to illustrate the NRTK concept. For Network results shown in this paper, approximately five and a half hours data is used and for baseline processing and rover position computations, approximately two hours of data is used.

Table 1: Sub-network reference receivers characteristics

Site	4-chars ID	Receiver type	Antenna type
Tonstad	TNSC	TRIMBLE NETR9	TRM55971.00
Sirevag	SIRC	TRIMBLE NETR9	TPSCR3_GGD
Stavanger	STAS	TRIMBLE NETR9	TRM55971.00
Akrahamn	AKRC	TRIMBLE NETR9	TPSCR3_GGD
Lysefjorden	LYSC	TRIMBLE NETR9	TRM55971.00
Prestaasen	PREC	TRIMBLE NETR9	TPSCR3_GGD

3 Network adjustment method

As mentioned in Section 1, several NRTK techniques exist as described in for instance Euler et al (2001); Landau et al (2002); Takac and Zelzer (2008); Wübbena et al (1996). The integrity monitoring algorithms developed and described in the remainder of this paper works independent of the method used for generation of the NRTK corrections.

Table 2: Distances in sub-network [km]

Sites	TNSC	SIRC	STAS	AKRC	LYSC	PREC
TNSC	X	56.32	75.00	109.60	44.61	95.23
SIRC	-	X	58.38	91.26	68.50	112.96
STAS	-	-	X	35.83	45.72	64.41
AKRC	-	-	-	X	73.51	65.60
LYSC	-	-	-	-	X	51.45
PREC	-	-	-	-	-	X

Table 3: Reference receiver coordinates - Euref89 XYZ coordinates in meter.

Sites	X	Y	Z
TNSC	3302221.359	388315.600	5424777.872
SIRC	3323397.670	336993.537	5415277.838
STAS	3275753.912	321110.865	5445041.883
AKRC	3254758.852	295601.453	5458918.670
LYSC	3269684.205	366420.447	5446037.395
PREC	3227088.927	353649.666	5471909.728

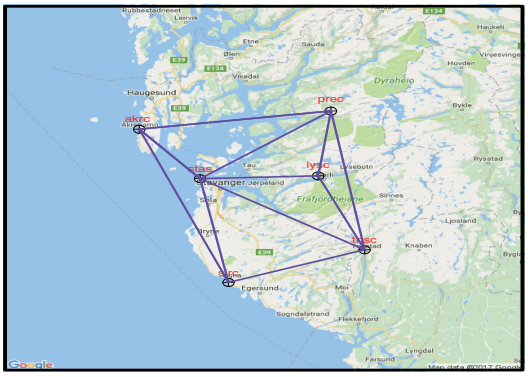


Figure 2: Test area used in this investigation, from Rogaland region in Norway. Composed of six reference receivers.

Our proposed NRTK data integrity concept is built on top of existing NRTK services. However, the computation of the correction field depends strongly on the method employed. For instance, it is essential whether the data itself is un-, single-, or double-differenced. The output from these filters are the dispersive and geometric biases, which can be provided either as one component or as separate components. For further analysis, the correction field has to be explicitly constructed, and their covariance matrices have to be examined closely. In addition, the filters variance-covariance matrices has to be inspected for Heywood cases. However, our method is independent of the approach and linear combinations used to generate these biases, and whether they are decomposed or not.

For derivation and development of the integrity and quality control algorithms we need a test case and we have based our work on the conceptual approach as given by the NetAdjust method (Raquet and Lachapelle, 1999; Raquet, 1998). Most of the NRTK techniques mentioned above are developed commercially and details about these are not readily available. But the NetAdjust method is well-described in literature, it is therefore suitable as a starting point for our work, and we provide a review of the method in the following.

The network adjustment method known as *NetAdjust* uses least-squares collocation techniques (Hofmann-Wellenhof and Moritz, 2006, Chap. 10) to compute the user network corrections. The Danish mathematician Torben Krarup (Borre, 2006) was the first to build the theoretical foundation for this new concept, namely the *collocation methods*. Since then, the method has been considered by geodesists as an algorithm for performing geodetic computations. For statisticians, this method is also known as *kriging*, a spatial linear interpolation algorithm named after the South African mining engineer D. G. Krige (Schabenberger and Gotway, 2004, p. 216). In this paper, we will refer to such collocation methods as kriging.

The NetAdjust method is based on carrier-phase double-difference techniques. Taking the difference between the original observation signals allows us to eliminate or reduce the effect of uncanceled differential biases. In addition, the technique has the advantage of a reduction in both the measurement and parameter count. One need not to include the entire set of double-difference measurements because it contains redundant information. In the case of double-difference observations, receiver and satellite clock errors are eliminated, i.e. the results become independent of the receiver and satellite clock biases. In this work, the effect of residual double-differenced code and phase hardware delays is assumed to be negligible.

The overarching philosophy behind the NetAdjust design can be summarized as follows (Raquet, 1998):

1. *Main equation*:
$$\Delta\nabla\ell = \underbrace{\Delta\nabla\delta\ell}_{\text{first-term}} + \underbrace{\Delta\nabla N}_{\text{second-term}} \tag{1}$$
Note that $\Delta\nabla$ is the double-difference operator and $\Delta\nabla\ell$ is the double-difference carrier-phase measurements, after subtracting range observables and the troposphere delay. This equation states that after correcting for double-difference ambiguity $\Delta\nabla N$, this is equivalent to the double-difference error $\Delta\nabla\delta\ell$, which is composed of residual atmospheric effects (ionosphere and troposphere), residual effects of the satellite position errors, as well as residual effects of multipath, receiver noise, antenna phase center variation, etc.
2. *NetAdjust signature*: Regardless of what ambiguity resolution algorithm one uses, the resolution is improved when the GNSS errors are minimized. This can be accomplished by reducing the uncertainties in the first term of Eq. (1), which facilitates the estimation of the second term, yielding improved ambiguity resolution.
3. *Error characteristics*: The NetAdjust method describes the error as a function of the position.
4. *Optimization*: Given the network measurements minus range observables and troposphere delay, one can estimate the differential measurement error δl that minimizes the total variance. The optimal estimator is determined using a Bayesian method, i.e. selecting a suitable loss function $L(\cdot)$ and thus an appropriate Bayes risk function $B(\cdot) = \mathbb{E}[L(\cdot)]$, where $\mathbb{E}$ is the expectation operator. For more details, e.g. Berger (1985) offers an elegant explanation of decision theory and Bayesian analysis.
5. *Prediction*: Least-squares collocation is a statistical estimation method that combines least-squares adjustment and prediction methods. The NetAdjust method uses the least-squares covariance analysis for accuracy prediction, i.e. to predict the carrier-phase error statistics for a given network configuration. For more details of this technique, the reader is referred to e.g. Pullen et al (1995).

We will now provide a brief discussion of the mathematical details of the method. We assume that the relationship between the parameter vector $\mathbf{x}$ and observation vector $\mathbf{Y}$ is a simple linear model $\mathbf{Y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, where $\mathbf{e}$ is an error vector. The Bayesian optimal estimator $\hat{\mathbf{x}}_{\text{opt}}$ with quadratic loss function is then obtained by minimizing the Bayes risk $B(\mathbf{x}) = \mathbb{E}\big[\|\mathbf{x} - \hat{\mathbf{x}}\|^2\big]$, thus yielding

$$\hat{\mathbf{x}}_{\text{opt}} = \mathbf{C}_{\mathbf{xY}}\mathbf{C}_{\mathbf{Y}}^{-1}\mathbf{Y}, \tag{2}$$

where $\mathbf{C}_{\mathbf{Y}}$ is the covariance matrix between sample locations, and $\mathbf{C}_{\mathbf{xY}}$ the covariance matrix between sample and prediction locations. This is also known as the *kriging equation*, and is used to compute the weights $\mathbf{W} = \mathbf{C}_{\mathbf{xY}}\mathbf{C}_{\mathbf{Y}}^{-1}$. To be more specific:

1. The elements of the covariance matrix $\mathbf{C}_{\mathbf{Y}}$ for the locations $\mathbf{Y}$ in the sample are defined as:
$$\forall i,j: \quad \big[\mathbf{C}_{\mathbf{Y}}\big]_{ij} = \text{Cov}(Y_i, Y_j)\,. \tag{3}$$
2. The elements of the covariance matrix $\mathbf{C}_{\mathbf{xY}}$ between the prediction points $\mathbf{x}$ and the sample locations $\mathbf{Y}$ are:
$$\forall i,j: \quad \big[\mathbf{C}_{\mathbf{xY}}\big]_{ij} = \text{Cov}(x_i, Y_j)\,. \tag{4}$$
3. The NetAdjust estimator $\hat{\mathbf{x}}_{\text{opt}}$ is the optimal minimum variance error estimator. Note that Eq. (2) can also be written in the simple form $\hat{\mathbf{x}}_{\text{opt}} = \mathbf{W}\mathbf{Y}$, which is a linear function of the observation vector $\mathbf{Y}$, and takes into consideration the covariance structure of the problem when estimating the weight matrix $\mathbf{W}$.

Computationally, the bottleneck when calculating the weight matrix $\mathbf{W}$ is the matrix inversion $\mathbf{C}_{\mathbf{Y}}^{-1}$. If the covariance matrix is large, the matrix inversion can become very time consuming. Moreover, if the matrix is ill-conditioned, there is also a risk of negative variance generation (Heywood, 1931).

NetAdjust uses the kriging equation [Eq. (2)] to compute the network corrections. The corrections are then transmitted to the user, and involves the users themselves in the position computation process. For more details, the reader is referred to Raquet (1998).

4 NRTK Integrity Design

In this section, we first briefly introduce classical RTK data processing schemes. We then follow up with a discussion of the advantages of NRTK systems, which extend the classical schemes through a network of reference receivers. We then discuss a further extension of NRTK systems with a novel and currently unavailable layer, namely the *NRTK Quality Control* or *data integrity layer*, referred to as the network RTK integrity segment in Figure 1.

Figure 3 shows the high-level functional decomposition of the NRTK data integrity, where the quality control is performed at both the system and user levels. Different processing schemes can be used to generate the user corrections: un-, single-, or double-differenced (Wübbena and Willgalis (2001), Zebhauser et al (2002), Dach et al (2015)). The user corrections may optionally be further decomposed into dispersive and geometric contributions based on their frequency-dependence. Our aim is to identify the exact locations in the NRTK data processing chain where data quality ought to be inspected and diagnosed. The result of this analysis should be a list of suspicious satellites that generate anomalous data.

At the network level, a check barrier is implemented to check the quality of the user corrections and the uncertainty provided by the covariance matrices. This check guarantees high quality for a simulated reference receiver, known as a virtual reference receiver (VRS) or computation point (CP). This check is referred to as network data integrity. The curved line of left panel in the Figure 3 indicates the output for this computation point. The next check barrier is at the baseline level, where the local data integrity is handled. The raw rover observation data is inspected by the variance weighting algorithm (i.e. the Danish method). The covariance matrix can then be analyzed at the double-difference level to check for stability. The relative positioning between the computation and rover points is handled at this level, as shown in the middle panel of Figure 3. Finally, the last check barrier is the inspection of the rover position accuracy and the construction of the error ellipse.

Other NRTK methods typically use two filters to compute the user corrections. The first filter uses an ionosphere-free linear combination to compute the geometric corrections, i.e. corrections for distortions caused by the troposphere, satellite position errors, and clocks. The advantage of this method is that the ionosphere path delay is practically eliminated. The second filter uses geometry-free linear combinations to estimate the ionospheric corrections. The advantage of this method is that it is independent of receiver clocks and geometry, and contains the ionospheric path delay and initial phase ambiguity. Regardless of the method, an average error level must be determined, and the statistical procedure and test statistics are similar for both approaches.

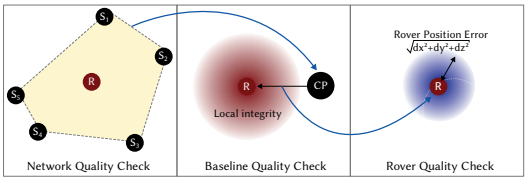


Figure 3: Check barriers of the network RTK data integrity. The left panel shows a network with five reference receivers S_i and a rover R. The middle panel shows the baseline quality check. The right panel shows the rover position error.

4.1 Network corrections quality check

Network real-time data processing is a pure spatio-temporal process, since data is continuously recorded at different stations, and the analysis has to account for both spatial and temporal correlations in the observation data. First of all, the observations at each station have intrinsic correlations when they are in geographical proximity. Additional correlations are introduced by both differencing schemes (El-Rabbany, 1994) and network processing (Leick, 2015). All of these effects have to be considered in a rigorous spatio-temporal analysis.

One way to treat the spatial part of the correlations, is to perform a Cressie decomposition (Cressie, 1993, Chap. 3):

$$R(\mathbf{s},t) = M(\mathbf{s},t) + V(\mathbf{s},t); \tag{5}$$

$R(\mathbf{s},t)$: real signal;
$M(\mathbf{s},t)$: mean function, trend, or large-scale variation;
$V(\mathbf{s},t)$: variance function, or small scale variation;
$(\mathbf{s},t)$: spatial and temporal variables.

The mean function $M(\mathbf{s},t)$ is calculated using standard GNSS processing techniques, including the detection and mitigation of GNSS error sources. These errors include models for

the signal path delays caused by e.g. tropospheric or ionospheric activity. However, no model is perfect, and there will always be room for improvements. Challenges in estimating this mean function include mapping out the covariance structure of the network, handling non-stationarity, handling non-Gaussian processes, and constructing models that are computationally efficient for large-scale data processing.

The variance function $V(\mathbf{s},t)$ is actually just the uncertainty of the network correction field. Although it seemingly plays a lesser role compared to the mean function $M(\mathbf{s},t)$, the importance of the variance function $V(\mathbf{s},t)$ cannot be overemphasized. This is because it can be used as a feedback control component when estimating $M(\mathbf{s},t)$, where one monitors undetected anomalies in $V(\mathbf{s},t)$ and attempt to compensate for its weaknesses. Thus, the variance function can be used to inform users in the field when the network corrections cannot be trusted, which is what we refer to as a data integrity. The main objective is to allow only satellites with high-quality data to be involved in the generation of the correction of the computation points, as discussed in more detail in Section 5.

4.2 Integrity of raw carrier-phase data

Figure 3 illustrates the importance of local data integrity. The NetAdjust system constructs high quality computation point(s) using data from the reference receivers. If the rover raw carrier phase observations have not been inspected for signal diffraction, multipath interference, and possibly also scintillation, the result of the double-difference baseline processing will be biased. Robust estimation techniques reduce the influence of outliers on the result. The distorted signals of the cases mentioned above, are not really outliers but biased observations.

Outliers are usually not just biased observations, but rather artifacts caused by filters, instruments, and other chaotic phenomena. They significantly deviate from the distribution of regular observations, and this makes them straight-forward to eliminate. In contrast, data distortion caused by multipath, scintillation, etc. result in biased observations that still resemble regular data, and these data points are much more challenging to detect in real-time.

Nevertheless, in cases where the bias itself is not explicitly modeled, one must take care to assign lower weights to these biased observations to prevent them from skewing the results. The combination of carrier-phase signal-to-noise ratios and the double-difference phase residuals is discussed in Section 6.

4.3 Baseline data integrity

The output from the baseline computations are the widelane double-difference carrier-phase residuals and the corresponding error covariance matrix. These parameters are combined in an appropriate way to predict the carrier-phase and code error statistics. This topic is the subject of the Section 7. The methods used in this subsection are summarized in Wieser and Brunner (2000).

5 Integrity for network corrections

The NetAdjust method as well as other NRTK methods can use widelane double-difference observations to generate the user corrections. In this paper, we aim to construct the corrections and corresponding variance fields on a satellite-by-satellite basis. This includes both test statistics and a determination of the temporal correlation length of observation combinations, which has to be computed from the observed data. For this purpose, we employ powerful methods from multivariate statistical data analysis for detection, identification and adaptation procedures, which produces a list of satellites that generate anomalous data.

Global tests are needed to assess whether a set of observations includes errors or not, while local tests are needed to identify the main reasons behind the failure of the global test. We have two candidates for global test statistics, and t distribution for local test statistics. For adaptation, the corrections from high residual values and variances are flagged for exclusion, and are thus not involved in the solution computation.

Using the theory of excursion probability (Adler and Taylor, 2009, Chap. 4), one can construct an optimal alarm condition for NRTK systems:

$$P_{\text{exc}} = \mathbb{P}\left\{ \sup_{\mathbf{s}\in S,\, t\in T} \int G(\mathbf{s},t)\, \mathrm{d}\mathbf{s}\, \mathrm{d}t \geq \text{Th} \right\}, \tag{6}$$

where *sup* stands for supremum (least upper bound), S and T are the spatial and the temporal spaces, while $G(\mathbf{s},t)$ is an empirical Green function that is constructed from the data. Our main concern is directed to those extremal events of the correction field that exceed some chosen threshold Th. When solving an optimization problem, one tries to solve the inherent conflict between accuracy and some heuristic cost function in the best possible way. These level-crossing events can bias the position solution of the rover. The next sections will be focused on constructing the components of $G(\mathbf{s},t)$.

5.1 Network average error levels

This section is devoted to construction of the average error level for each satellite observed at each configured reference receiver in network. The technique of multivariate statistical analysis is used for this purpose.

5.1.1 Time series

Let $Y = \{Y_{ijk}\}$ be observations, where $i = 1,\ldots,n_{\text{rec}}$ are the reference receivers, $j = 1,\ldots,n_{\text{sat}}$ are the satellites observed at each site i, and k is size of the moving window. The size of the moving window is equal to the correlation length of the observations used. According to Schön and Brunner (2008), this correlation length is in the range of 300–600 seconds in the widelane case. Odolinski Odolinski (2012) presented two methods to estimate the correlation length, and found ~17 min for the horizontal component, and ~37 min for the vertical one. In any case, the larger the moving window, the lower the correlation separation time.

The correlation time can also vary depending on the baseline length. For example, for short baselines of only a few kilometers, we expect only multipath errors and internal receiver effects to be relevant, and that these two factors will determine the correlation time. However, for longer baselines, larger correlation times can be expected if any residual atmospheric delays still remain.

We can describe Y as a matrix-valued sequence of length k, describing the dynamics of the network correction field $G(\mathbf{s},t)$. Figure 4 shows the constructed average error level for ionospheric corrections in a network of six receivers. The geometry-free linear combination $L_4 = L_1 - L_2$ is used to generate the data presented in Figure 4. This observation cancels out all the geometry information leaving only the ionosphere effects and initial phase ambiguities. It is commonly used for the estimation of the ionosphere path delay. In the plot the variation of different receivers is shown. Three sites contribute with an equal average error level (top curves), the next two contribute almost equally too (middle), but the final one is distinct from all the other ones (bottom).

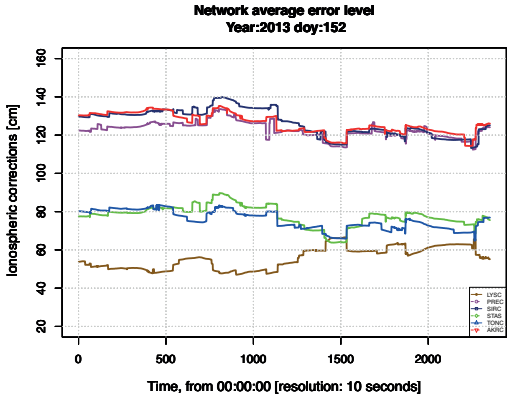


Figure 4: Computed ionospheric average error level for a configured network. Year 2013, DOY 155.

5.1.2 Missing observations

In order to compute the mean, median, and corresponding covariance matrices of Y_i on satellite-by-satellite basis, the constructed time series need to have the same length. In practice, this will of course be nearly impossible, so we need to perform a procedure known as *data imputation*. For this, one can apply an expectation-maximization data augmentation algorithm, such as the one proposed by Dempster et al (1977).

5.2 Global and local test statistics

The empirical stochastic correction field $G(\mathbf{z})$ can be regarded as a function of $\mathbf{Y}_i$, where $\mathbf{z} = (\mathbf{s},t)$ is a 4-dimensional vector in space $\mathbf{s}$ and time t. We will assume that it is a Gaussian field with a p-dimensional probability density function $f(\mathbf{z})$, which is parametrized by a mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$:

$$\begin{aligned} f(\mathbf{z}) &= |2\pi|^{-p/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{z}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{z}-\boldsymbol{\mu})\right\} \\ &= T_A(\mathbf{z})\,\exp\left\{-\frac{1}{2}T_B(\mathbf{z})\right\} \end{aligned} \tag{7}$$

where the notation $|\cdot|$ refers to the matrix determinant, and the functions T_A and T_B are defined respectively by the expressions $|2\pi|^{-p/2}|\boldsymbol{\Sigma}|^{-1/2}$ and $(\mathbf{z}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{z}-\boldsymbol{\mu})$. T_A and T_B are elementary building blocks of the test statistics used in this article.

Our check algorithm is a three-step process, composed of Detection, Identification, and Adaptation. Extremal crossing events can be detected using the global test statistic given by Eqs. (11) and (12). Let our current correction vector for reference receiver i be denoted $\mathbf{x}_i$. If we are interested in measuring how far the observation $\mathbf{x}_i$ is from the mean $\boldsymbol{\mu}_i$, then a Euclidean metric, Eq. (8) performs well mathematically, but is sensitive to the specific units of measurements.

$$E_i = (\mathbf{x}_i - \boldsymbol{\mu}_i)^{\mathrm{T}}(\mathbf{x}_i - \boldsymbol{\mu}_i) \tag{8}$$

One may therefore wonder if there is a more informative way, particularly in a statistical sense, to measure if the distance $\mathbf{x}_i$ is far from the mean $\boldsymbol{\mu}_i$. One such metric is given by the *squared Mahalanobis distance* (SMD) defined in Eq. (9), which accounts for the correlations between the observations and measures the distance in units of standard deviations.

$$M_i = (\mathbf{x}_i - \boldsymbol{\mu}_i)^{\mathrm{T}}\,\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_i) \tag{9}$$

An alternative metric is the *Mahalanobis depth* (MD):

$$m_i = \frac{1}{1+M_i} \tag{10}$$

This time, we measure how far the observations $\mathbf{x}_i$ are from the median, and we note that large values of m_i correspond to values of x_i that are deep inside the distribution.

5.2.1 Definition of global test statistics

In order to detect when extremal events occur, we need some kind of global statistical tests. For this purpose, we have chosen two test statistics:

$$T_1(\mathbf{z}_i) = (\mathbf{z}_i - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}^{-1} (\mathbf{z}_i - \boldsymbol{\mu}_i) \tag{11}$$

$$T_2(\mathbf{z}_i) = \frac{1}{1 + T_1(\mathbf{z}_i)} \tag{12}$$

where $\mathbf{z}_i$ is the correction vector observed at reference receiver i at time epoch t. Note that T_1 and T_2 follow the multivariate χ^2-distribution and its inverse.

Figures 5 and 6 show how global tests can detect the extremal events caused by network corrections. The plots are provided as functions on time. We see that both the tests are capable of detecting the same events – but while the SMD detects the *maxima* that exceed the threshold value T_h, the MD detects the *minima* in the data set. Note that this approach is based on the median vector, and not the less robust mean vector.

For SMD, the threshold T_h and level of significance α was set to 15 and 90%, respectability in this test, and correspond to $\chi^2_9(.10) \approx 15$. The subscript 9 corresponds to degree of freedom (i.e. average of observed satellites). In contrast to the MD case, the threshold T_h was set to $1/16$ in this test.

The resolution is set to 10 seconds intentionally and corresponds to the optimal update rate of the network corrections.

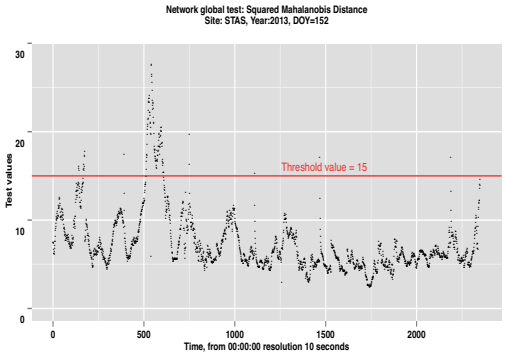


Figure 5: Sample of SMD based on 2500 epochs with a resolution of 10 seconds. Red horizontal line shows the rejection region of the test.

With one samples from a univariate normal distribution, the variability of the sample variance S^2 is governed by the chi-squared distribution χ^2. This distribution also holds an important role in multivariate statistics (Johnson and Wichern, 2002, Chap. 4). To see this, let us first define $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, i.e. $\mathbf{X}$ is a normally distributed random variable with a mean vector $\boldsymbol{\mu}$ and a positive-definite covariance matrix $\boldsymbol{\Sigma}$. We denote the SMD of this variable as $M(\mathbf{X}) = (\mathbf{X} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu})$. It can then be shown that:

(i) $M(\mathbf{X}) \sim \chi^2_p$, meaning that the SMD follows a chi-squared distribution with p degrees of freedom.

(ii) There is a probability $(1 - \alpha)$ for an observation to be within the ellipsoid defined by $M(\mathbf{X}) \leq \chi^2_p(\alpha)$. We therefore use the index $\chi^2_p(\alpha)$ as the appropriate threshold value.
Here, $\chi^2_p(\alpha)$ refers to the quantiles of the chi-squared distribution with p degrees of freedom, where $(p + 1)$ is the number of satellites used in the computation. The argument α is the level of significance (e.g. 99%), and defines the rejection level of the crossing events. Note that this is different from the false alarm rate, which instead refers to error *type I* (Shanmugan and Breipohl, 1988, p. 346).

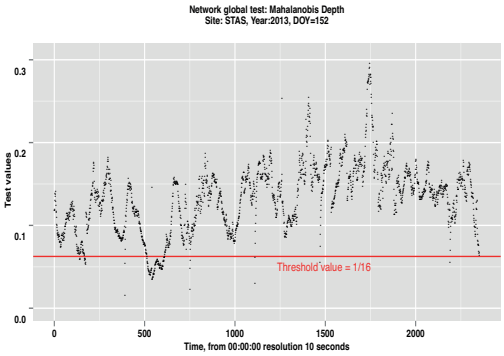


Figure 6: Sample of MD based on 2500 epochs with a resolution of 10 seconds. Red horizontal line shows the rejection region of the test.

If we combine the MD [Eq. (10)] with the median $\boldsymbol{\mu}_{\text{med}}$, we can interpret $G(\mathbf{z})$ as the median correction field. On the other hand, combining the SMD [Eq. (9)] with the mean $\boldsymbol{\mu}$, the correct interpretation of $G(\mathbf{z})$ is the mean correction field. The accuracy of this method is measured by the expected variance with respect to a certain distribution. This means that the standard deviation field $F(\mathbf{z})$ has to be determined. Note that the standard deviation of the widelane observation combinations depends on the standard deviations of the original L_1 and L_2 signals, which again vary with e.g. the receiver type and antennas used for the observations. For a summary of the most common linear combinations of carrier phases and the corresponding variances, see e.g. (Seeber, 2003, Tab. 7.7). These procedures are similar to the ones

used for the corrections field itself; at each reference receiver, the standard deviation of each observed satellite has to be investigated with respect to $F(\mathbf{z})$.

5.2.2 Interpretation of the global tests

The SMD $M(\mathbf{z})$ is a statistical metric that measures the squared distance between some point $\mathbf{z}$ and the population mean $\boldsymbol{\mu}$. One way to understand this metric $M(\mathbf{z}) = (\mathbf{z} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{z} - \boldsymbol{\mu})$, is that it is similar to the Euclidean metric $E(\mathbf{z}) = (\mathbf{z}-\boldsymbol{\mu})^{\mathrm{T}}(\mathbf{z}-\boldsymbol{\mu})$, but deformed by the covariance structure $\boldsymbol{\Sigma}^{-1}$ of the data. This has two important consequences which render $M(\mathbf{z})$ more useful than $E(\mathbf{z})$ for our purposes:

(i) Even though some components of $\mathbf{z}$ have a larger variance than others, they can contribute equally to the SMD;
(ii) Two highly correlated random variables will contribute to the SMD more than two uncorrelated randoms variables.

In order to use the inverse of the covariance matrix $\boldsymbol{\Sigma}^{-1}$ properly, these steps are recommended in practical implementations:

(i) Standardizes all the variables, that is, transform the random variable $\boldsymbol{Z}$ into p independent standard normal random variables $\boldsymbol{X}$.
(ii) One can eliminate the correlation effects by performing a variable transformation $\mathbf{x} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{z} - \boldsymbol{\mu})$, since this results in $\mathbf{x} \sim N_p(\mathbf{0}, \mathbf{I}_p)$ having a trivial normal distribution with zero mean and a diagonal covariance structure. The SMD can then be calculated as if $\mathbf{z}$ is transformed into p independent random variables (i.e. the elements of $\mathbf{x}$), where each variable follows a standard normal distribution.

5.2.3 Definition of local test statistics

The next step in the investigation process is the identification of influential residuals, and the assessment of their effects on various aspects of the analysis.

Considering the general linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{y}$ is a vector of response variable, $\mathbf{X}$ is the design matrix, $\boldsymbol{\beta}$ is a vector of unknown coefficients to be estimated, and $\boldsymbol{\epsilon}$ is a vector of random disturbances. Applying a least-squares parameter estimation, we find:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} \tag{13}$$

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} = \mathbf{P}\mathbf{y} \tag{14}$$

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = (\mathbf{I} - \mathbf{P})\boldsymbol{\epsilon} \tag{15}$$

The error vector $\mathbf{e}$ can then be considered as a reasonable substitute of $\boldsymbol{\epsilon}$. Note the error vector $\mathbf{e}$ depends strongly on the prediction matrix $\mathbf{P}$. It is also required that the design matrix

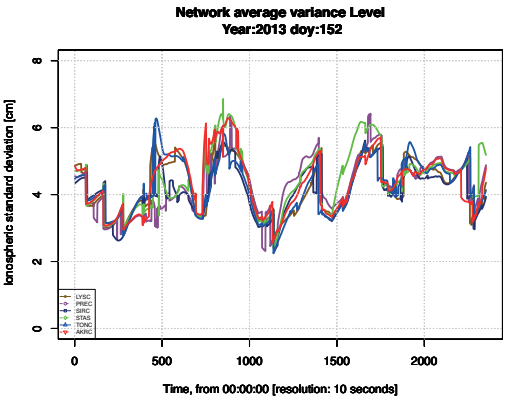


Figure 7: Standard deviations for the reference receivers in the network. This plot shows that the variance at each site behaves in the same way.

$\mathbf{X}$ is homogeneous, meaning that the diagonal elements of $\mathbf{P}$ are equal, while the off-diagonal elements are reasonably small. For these reasons, it is preferable to use a transformation of the ordinary residuals for diagnostic purposes. That is, instead of using the error vectors $\mathbf{e}_i$, one may use the reduced error vectors $\mathbf{T}_i$, where σ_i is the standard deviation of the i'th residual.

$$\mathbf{T}_i = \frac{\mathbf{e}_i}{\sigma_i}. \tag{16}$$

In this research paper, we restrict the local test statistics to the normal distribution and t-distribution. Both tests are used interchangeably, and we find that they produce nearly identical results. Interested readers are referred to Ref. (Chatterjee and Hadi, 2009, Chap. 4) for a discussion of other tests that can be constructed for this purpose.

5.2.4 Variance Monitoring

It is critical to monitor the variance of each satellite when performing GNSS NRTK calculations. For an example of how the variance changes for reference receivers, see Figure 7.

5.2.4.1 Generalized variance According to the large sample theory, it is clear that the correction field should be well-described by a multivariate normal distribution known as a Gaussian field. This means that the distribution should converge to this regardless of the parent population we sample from.

If we take a close look at the probability density function given by Eq. (7), it contains the prefactor $|\boldsymbol{\Sigma}|$, which is also known as the *generalized variance* (GV) and provides a way of writing the information on all variances and covariances as

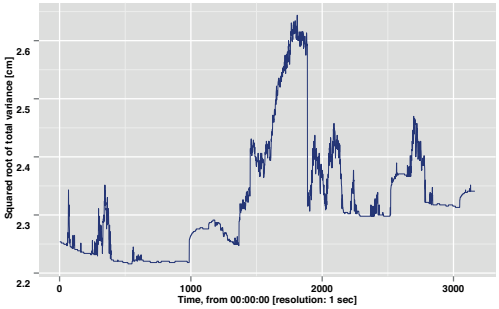


Figure 8: Computed squared total variance for baseline of 41 km between reference receivers HFSS and SAND. Year=2013, DOY=152.

a single judging number. The drawback is that the GV does not contain any information on the orientation of the pattern.

The covariance matrix contains a lot of information: the diagonal describes the variance of each observed satellite, while the off-diagonal corresponds to the covariance between them. When the generalized variance is computed, all directional information contained in the structure of the matrix is discarded. In other words, the covariance matrix is distilled down to a single number, which we can heuristically treat as the "generalized variance" of the system. In this paper, our goal is to monitor the variation of the generalized variance itself. We therefore form a time-series from the generalized variance of the sample covariance matrix $\mathbf{S}$, and study its variations on an epoch-by-epoch basis.

We will define a new stochastic variable $y_i = |\mathbf{S}_i|$, where $\mathbf{S}$ is the sample covariance matrix. We can then construct a time-series for these y_i, and thus monitor the variations over time.

5.2.4.2 Total variance Given the sample covariance matrix $\mathbf{S}$, we may define the total variance as $z = \text{tr}\,(\mathbf{S})$. This definition can be intuitively understood, since variance is an additive quantity, and the diagonal of the covariance matrix contains the variance of each component of the random variable. If we then construct a time-series for the observable quantity z, we can directly monitor how the total variance changes on an epoch-by-epoch basis. The total variance is always attractive to investigate due to the following facts:

– For any estimator $e(Y)$ of type Linear Unbiased Minimum Variance (LUMV). The following expression (Raquet, 1998, p. 54) holds.

$$\begin{aligned}\mathbf{E}(\mathbf{x} - e(Y))^2 &= \mathbf{E}[(\mathbf{x} - e(Y))^T(\mathbf{x} - e(Y))] \\ &= \text{trace}\left\{\mathbf{E}[(\mathbf{x} - e(Y))(\mathbf{x} - e(Y))^T]\right\} \quad (17)\end{aligned}$$

The left expression of the Eq. (17) is the Bayesian risk with quadratic loss function, while the right side is the total variance given by covariance of the estimator $e(Y)$.

– The optimization of Kalman filter (Brown and Hwang, 1997, pp. 216–217) is the minimization of trace of the error covariance matrix of the state vector $\mathbf{x}$.

5.2.5 Link function definition

Construction of the prediction function of the rover position error is directly linked to the total variance of the error covariance matrix $\boldsymbol{C}_{\mathbf{err}}$.

Our proposed model is the Stochastic Generalized Linear Model (SGLM). The GLM model was proposed by McCullagh and Nelder (1989), and is an extension of the classical linear model (LM) with additional component known as a linear predictor $\Psi = g(.)$. Function $g(.)$ is the link function.

Let Eq. (18) be the double-difference (DD) observation model of the baseline between the rover receiver and the computation point.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{A}\boldsymbol{a} + \boldsymbol{\epsilon} \quad (18)$$

The random component $\mathbf{y}$ of the SGLM may come from any exponential family distribution rather than a Gaussian distribution as in case of a LM. $\mathbf{y}$ is a vector of Observed Minus Computed (OMC) values; $\boldsymbol{\beta}$ is a vector of all parameters except the DD ambiguities; $\boldsymbol{a}$ is a vector of unknown DD ambiguity parameters, $\mathbf{X}$ and $\mathbf{A}$ are design matrices.

The systematic component in GLM is computed by the covariates of $\mathbf{X}$, that is $\Psi = \mathbf{X}\boldsymbol{\beta}$. In our case, this component is linked to the uncertainty of the model.

$$\Psi = g(\mathbf{trace}(\boldsymbol{C}_{\mathbf{err}})) = \left\{\frac{1}{p}\sum_{1}^{p} c_{ii}\right\}^{1/q} \quad (19)$$

where p i the number of satellites, c_{ii} are diagonal elements of the covariance matrix $C_{\mathbf{err}}$, and $q \in \{1, \cdots, p\}$ is a parameter. For $q = 2$, $\Psi(.)$ function is the Root Mean Square (RMS) of the diagonal elements of $C_{\mathbf{err}}$.

The link function $\Psi(.)$ is stochastic due to the facts that it is a function of uncertainties of the model. A realistic definition of $\Psi(.)$ can be any monotonic differentiable function. Since Ψ relates the linear predictor to the expected average variance, various forms of $\Psi(.)$ are given in (McCullagh and Nelder, 1989, p. 31).

Figure 8 shows the computed generalized variance for a baseline of 41 km, while Figure 11 shows the predicted square root of the average variance. In this case $q = 2$ and is the predicted RMS.

6 RTK user level phase observable integrity

One common problem with GNSS systems is that some satellite signals arrives at the user receivers with damaged data

due to factors such as low signal quality, low elevation angle, multipath interference, diffraction, or scintillation. It is therefore important to inspect the raw observation data, so that signals suffering from such problems can be discarded from the processing chain at an early stage. It is especially critical that this inspection is performed before the widelane double-difference processing of the baseline.

Since GNSS users often find themselves in places with limited quality satellite signals, the optimal approach is to help these users discard the low-quality satellite data in the field, without requiring further assistance from NRTK systems that may also suffer from limited signal quality. Therefore, the raw phase observations at the users location have to be investigated for the error sources discussed above, before one proceeds with any processing of the data. In practice, this always results in some kind of trade-off between satellite geometry and accuracy. This is because if data from satellites with low elevation are included in the processing, this generally increases noise and systematic errors due to the long signal path through the ionosphere and troposphere.

Several weighting schemes based on the measured carrier-to-noise power density ratio r can be used to model this random error and the relevant distortions. Langley (1997) showed that the standard deviation of phase observations in the phase-locked loop (PLL) of a GPS receiver is a function of carrier-to-noise ratio r, bandwidth B_w, and carrier frequency f_c. Moreover, according to the SIGMA-ϵ weight model Wieser and Brunner (2000), the ratio r can be linked to the variance of the phase measurements using some empirical coefficients β_i. The model reads:

$$\sigma^2_{\phi,i} = \alpha_i + \beta_i \times 10^{-r/10} \tag{20}$$

where $\sigma_{\phi,i}$ is the standard deviation of the undifferenced carrier-phase observation, α_i and β_i are the model parameters, and i is an index that determines the receiver type, antenna type, and frequency. Note, however, that Eq. (20) has a well-known drawback: the detection process is delayed.

This is because observations become biased when subjected to local disturbances such as multipath interference, diffraction, or scintillation. The detection of level changes caused by increasing variance, takes a time to be detected by applying the function given by Eq. (20), and the Danish method is very sensitive to small level changes.

The ameliorations are therefore carried out by the Danish method (Wieser and Brunner, 2000) in this work, because this is a robust estimator based on iterative least squares reweighting algorithm.

7 Baseline integrity

The last step in the NRTK integrity scheme is a three-step baseline computation. At the first step, we require that the double-difference ambiguity between the computation point and rover receiver is correctly resolved. For short baselines < 20 km, this can be done using for instance an algorithm developed by Ming et al (1994). Figure 9 shows the convergence of the ambiguity and the estimated double differenced ionospheric delay as function of time. The weighting scheme proposed by Wieser and Brunner (2000) combines the information inherent in the ratio r, and the double-difference residuals are then used for the local data integrity calculations. With *local*, we here mean scintillation (Ouassou et al, 2016), multipath interference, or any other environmental disturbances that affect the rover receiver. The results show that the proposed scheme significantly improves the precision of the positioning service.

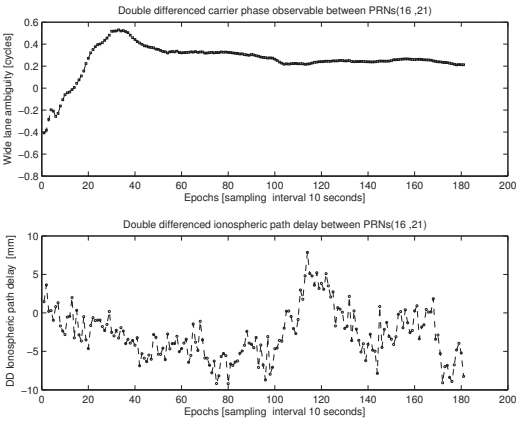


Figure 9: Short baseline ambiguity resolution, year= 2014, DOY=85, and baseline length ~ 1 km between HFSS and mobile receiver MHFS. (a) Upper panel shows the convergence of the widelane ambiguity resolution in double difference level. (b) Lower panel shows the corresponding ionospheric path delay.

In Section 5, a computation point is constructed corresponding to the average error level of the sub-network of reference receivers, while in Section 6, the carrier-phase observables are checked against outliers. After these calculations, it is appropriate to combine both quality control in the form of model residual minimization, ambiguity in the form of time-to-fix, and finally the user accuracy. The next step is the analysis of the double-difference residuals and the corresponding error covariance matrix. Test statistics similar to the ones introduced for network data integrity, Eqs. (11) and (12) are also suitable for baseline processing. Figure 10 shows the results of the global tests used in the detection process. The shadowed rectangle is caused by the occurrence of negative variance in covariance processing matrix, known as Heywood case (Heywood, 1931).

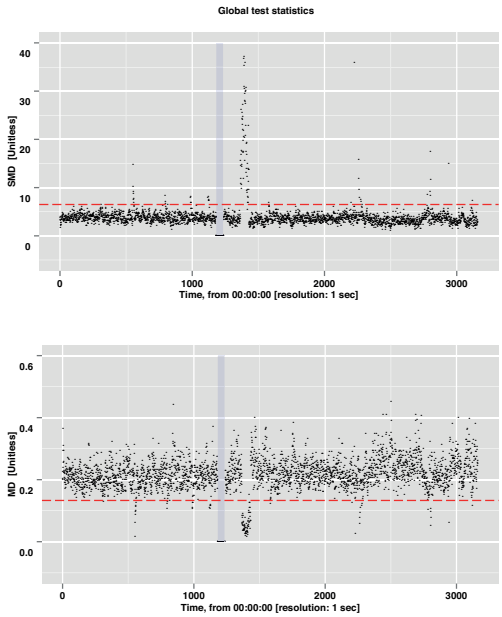


Figure 10: Sample of global test statistics for baseline based on 3200 epochs with a resolution of 1 sec. Dashed red horizontal line determines the rejection region of the test. Year= 2013, DOY=152.

The upper and lower panels of the Figure 10 show the SMD and MD test results. The values used in the detection process are 6.5 and .133, respectively. These values correspond to the critical quantile of the Chi-squared distribution (χ^2_p), where $p = 12$ and correspond to the number of observed satellites used in the computation at $\alpha = 90\%$ significance level.

In addition, a prediction function is obtained by using the SGLM to predict the user carrier-phase error and code statistics.

The last and final step is the computation of the user position standard deviation, and a comparison of the results obtained before and after the improvement, while conserving the geometry of the setup.

8 Implementation and analysis

In order to carry out the performance analysis of NRTK methods, and predict the carrier-phase and code statistics, an averaging variance level of the baseline processing is constructed. Figure 11 shows the predicted RMS from the double-difference error covariance matrix as a function of

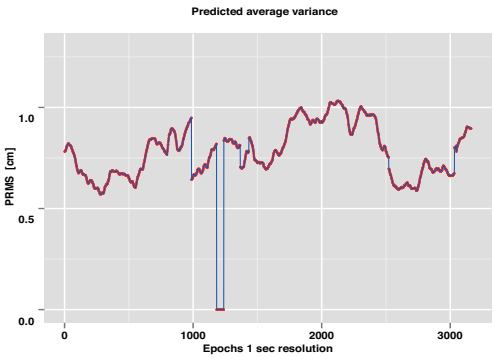


Figure 11: Predicted RMS computed from the double-difference error covariance matrix. Data used in this investigation are from a baseline of ~ 41 km. Year=2013, DOY=152.

time. The discontinuities are caused by the reference satellite changes when resolving the ambiguities.

8.1 Validation of NRTK integrity

Validation is a complex and challenging process to implement correctly, and careful planning is required in order to define appropriate validation procedures. In order to validate the implemented algorithms at both the system and the user levels, a side-by-side comparison of the candidates has to be conducted. According to Kaplan and Hegarty (2006), the accuracy of a GPS solution is proportional to the product of a geometry factor and a pseudorange error factor. The geometric error factor can be described by the Dilution of Precision (DOP) parameter, while the pseudorange error factor is the User Equivalent Range Error (UERE), so one can say that the position error is proportional to DOP × UERE. Thus, high values of either the DOP or UERE will result in a poor positioning accuracy.

The first step of such a validation procedure, is to compute the quality of the rover position errors $\mathbf{\Delta_{enu}} = (\Delta e, \Delta n, \Delta u)$ relative to the standard deviations $\sigma_{\mathbf{\Delta_{enu}}}$, and to calculate the DOP without enabling the mechanisms of NRTK data integrity. The next step is to enable the network data integrity quality check and produce a list of all detected satellites on an epoch-by-epoch basis. This list is read by a software program within observations from RINEX files, excluding all data from satellites mentioned in the list, and produce new RINEX files. After that, the first step is repeated again. The geometry expressed by DOP and standard deviation of the rover position error are then re-computed, and the results may then be compared. For an illustration of results of this processing, see figures 12 and 13.

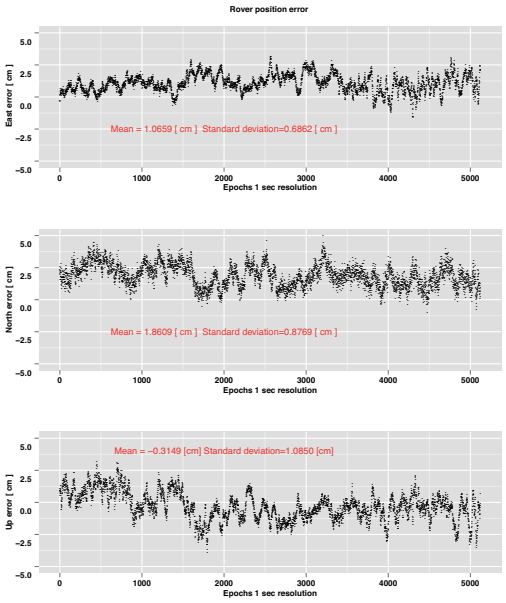


Figure 12: Rover position error as function of time without enabling the quality check procedures.

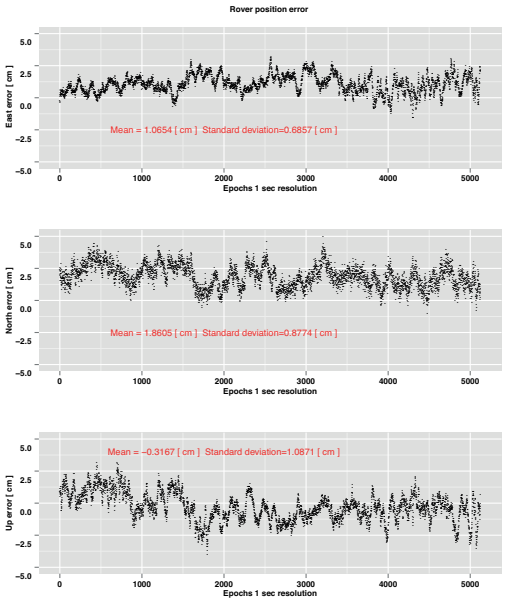


Figure 13: Rover position error as function of time after removing satellites with bad data on an epoch basis. The quality check procedures are enabled at network level.

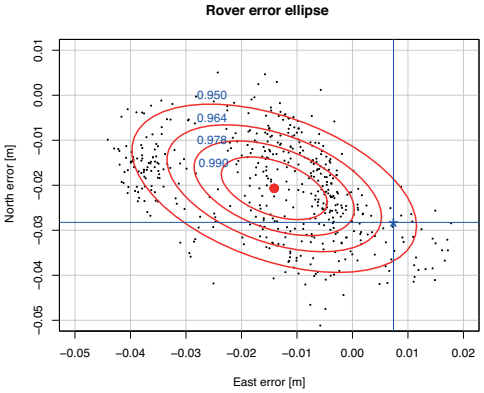


Figure 14: Error ellipse displaying the rover position error in the horizontal plane. The center of the ellipse is displayed by the red point and the actual user location is given by the intersection between the horizontal and vertical blue lines. Each ellipse corresponds to the probability of acceptance of the null hypothesis H_0.

8.2 Rover position error

The final product is to plot the rover position error in the horizontal plane on the receiver display. The user may then choose to either accept or reject the measurement results for the present epoch based on user requirements to acceptable error ellipse or standard deviation of total position error as illustrated in figures 14 and 15. Ideally, there should be no need for re-evaluating the quality of the measurements, potentially saving time for the end-user.

The position error vector is usually defined in a Cartesian coordinate system, i.e. $\boldsymbol{\Delta}_1 = (X, Y, Z)$. However, in practice, it is much more convenient to analyze the covariance matrices in a local topocentric coordinate system, i.e. $\boldsymbol{\Delta}_2 = (E, N, U)$ where the coordinates are given as east, north, and height (up). The transformation between these coordinate systems (Rogers, 2003, p. 48) is then given by the orthogonal matrix T.

$$\boldsymbol{T} = \begin{bmatrix} \sin(N) & \cos(N) & 0 \\ \sin(E)\cos(N) & \sin(E)\sin(N) & \cos(E) \\ \cos(E)\cos(N) & \cos(E)\sin(N) & \sin(E) \end{bmatrix} \tag{21}$$

In addition, the covariance matrix $\boldsymbol{C}_{\mathrm{XYZ}}$ is expressed in $\boldsymbol{\Delta}_1$ coordinates and our aim is to construct the user error ellipse in a topocentric coordinate system $\boldsymbol{\Delta}_2$. Applying the covariance propagation law reads

$$\boldsymbol{C}_{\mathrm{ENU}} = \boldsymbol{T}\, \boldsymbol{C}_{\mathrm{XYZ}}\, \boldsymbol{T}' \tag{22}$$

The constructed error ellipse in the horizontal plane in a topocentric coordinate system is illustrated by the Figure 14.

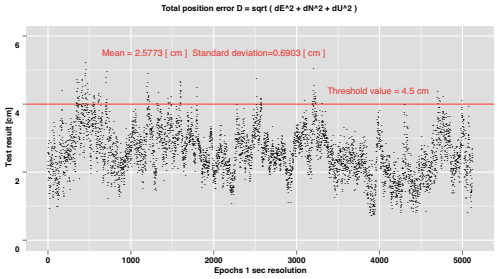


Figure 15: Standard deviation of the rover total position error as function of time. The horizontal line signals the crossing of the extremal events and separates the acceptance and the rejection regions.

The number of observations displayed in the figure, corresponds to the correlation length of the observations combinations used to compute the rover positions. In this test is set to 300 seconds.

Figure 15 shows the error radius given by the expression $D = \sqrt{((\delta e)^2 + (\delta n)^2 + (\delta u)^2)}$ with threshold value $T_h = 4.5$ cm.

The values of $T_h = \chi^2_p(.05) = 4.5$ correspond to $p = 11$ at $\alpha = 0.95\%$ significance level. On average, the common observed satellites between the rover and the base receivers in this test is eleven satellites.

9 Conclusion and discussions

An improvement of the rover position estimation can be achieved by applying procedures for integrity monitoring at the system and user levels in network RTK. In this paper we have presented an approach based on multivariate statistics, where the network average error corrections and the corresponding variance fields are computed from the raw data, while the squared Mahalanobis distance (SMD) and Mahalanobis depth (MD) are used as test statistics to detect and remove inaccurate data. Quality checks are carried out at both the network system level and at the rover user level in order to reduce the impact of extreme events on the rover position estimates. The Stochastic Generalized Linear Model (SGLM) is proposed and used to predict the rover carrier-phase and code error statistics.

The methods tested makes it possible to identify satellites with bad data so this can be eliminated or down weighted in the positioning process leading to an improvement in the rover position from epoch to epoch. Tests carried out as described in the paper show that there is indeed an improvement in the rover position after applying the new method.

It is expected that the suggested approach will reduce the number of wrong or inaccurate rover positions encountered by NRTK users in field, which subsequently will lead to a more efficient work flow for NRTK users.

All test results shown in this paper are based on GPS data only, but the algorithms will work just as well with data from e.g. GLONASS or Galileo satellites.

More tests will be carried out in the future by including other constellations for instance GLONASS and Galileo.

Discussion and considerations on implementation

1. **Benefit from NRTK data integrity**:
   Network RTK data integrity helps the user in the field. To benefit from the NRTK data integrity, use of the new RTCM 3.x (Commission and Services, 2006) message types is recommended. From network data integrity, the anomaly list is produced and suspicious satellites are sent to the rover. The rover software must also be upgraded to be able to decode and use the data properly. This task requires a new software module to be implemented in the rover. Figure 16 illustrates the concept.
2. **NRTK data integrity block diagram**
   Figure 16 shows the NRTK data integrity block diagram exemplified in a case where both GPS and GLONASS are used. The anomaly list is produced, packed and transmitted in RTCM 3.x format to the user rover. The software in the rover decodes the messages and excludes data from the given satellite(s) in the solution computation. The double difference error covariance matrix is used to estimate the user position, and an error ellipse (Figure 14) can be constructed and e.g. displayed to the user.
3. **Data exclusion and processing**:
   In order to test the concept, we have excluded approximately 0.1% of bad data from the computation and we have processed only the GPS. The new concept will also work properly in case of using various GNSS systems for instance GPS and GLONASS. This exclusion caused the change in both the location (mean) and the shape (variance) of the target distribution (see the Figures 12 and 13). We have computed the standard deviation of the rover position while keeping the mean value computed before enabling the quality check procedures. The result shows that the standard deviations of $(\delta e, \delta n, \delta u)$ drop from $= (6.859, 8.776, 10.872)$ to $(6.857, 8.774, 10.870)$ mm. This shows that there is indeed an improvement in the rover position accuracy.
   Excluding the data rises some important issues that need to be considered. To evaluate the results of the test two hypotheses can be set:

   Alternative 1: Two population mean comparison with different variance-covariance matrices

   $$H_0 : \mu_1 = \mu_2, \qquad \Sigma_1 \neq \Sigma_2 \tag{23}$$

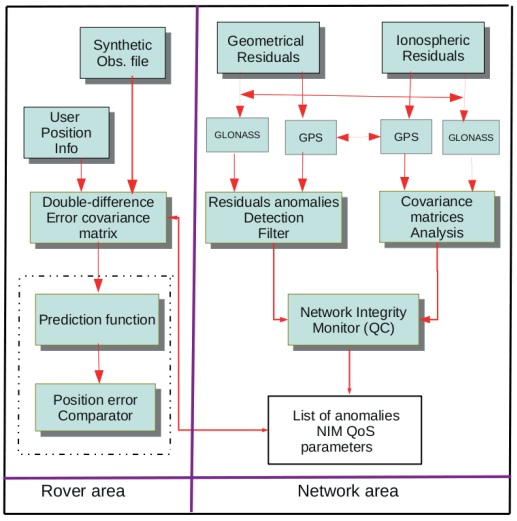


Figure 16: Design blocks of the network RTK data integrity. (a) The right panel shows the network integrity monitoring (NIM) quality of service (QoS) parameters generation. (b) The left panel shows data processing at the rover receiver.

This is unsolved problem in statistics and is know as *"Fisher-Behrens Problem"*.

Alternative 2: Two population mean comparison with equal variance-covariance matrices

$$H_0 : \mu_1 = \mu_2, \qquad \Sigma_1 = \Sigma_2 \tag{24}$$

This requires first to test for equal variance-covariance, before testing for equal means.
In addition, we have excluded only one satellite in the detection step. If there is more than one suspicious satellite, say two or three satellites with bad data, only one satellite with high value is removed.

4. **Performance analysis**: The performance analysis of our NRTK data integrity is measured in terms of carrier-phase and code error statistics at the user location (position domain). The SGLM is used for this purpose.
5. **Ambiguity resolution**: Key for precise positioning is correct determination and validation of the carrier phase ambiguity resolution. Often, this task is carried out by a Kalman filter (Brown and Hwang, 1997, Figure 5.8). Kalman gain K_k is involved in the computation of state vector update $\hat{x}_k^+ = \hat{x}_k^- + K_k(z_k - H_k\hat{x}_k^-)$ and the corresponding error covariance matrix $\hat{P}_k^+ = (I - K_k H_k)\hat{P}_k^-$. Therefore, $\hat{P}_k^+$ must be inspected for Heywood case and K_k must be monitored correctly to avoid the filter instability.

A Appendices

A.1 software development tool

Various computer programs have been developed to process the GNSS data and generate the figures in this paper. Software modules are classified into 3 categories, namely the network, baseline and rover receiver respectively.

1. *Network data processing*:
   - *NMA network SW*: This module is used to generate the NRTK corrections on a satellite-by-satellite basis. Data used in this test is from Rogaland region, year 2013, day of year 152 and classified as high ionosphere activity.
   - *Parsing the generated corrections*: A new C++ module is developed to parse and generate corrections, and it produces a suitable matrix format that is easy to process with Matlab, Python, or R.
     The corrections are ionospheric and geometrical (troposphere, clocks and orbit errors), obtained by forming respectively geometry-free and ionosphere-free linear combinations of the observables.
   - *Satellite data exclusion*: A C++ program is developed to exclude satellite(s) data on an epoch basis.
   - *Plots generation*: Various R scripts are developed and used to generate the figures 3 – 6 and 16.
2. *Baseline data processing*
   - *RTKLIB*: Open source program package for GNSS positioning developed by Takasu is used for experimentation (Tomoji and Yasuda, 2009).
   - *Baseline processing*: A C program based on RTKLIB developed and used to process baseline of different length. The output are the residuals and the variance-covariance matrix.
   - *Matlab script*: Ambiguity resolution for baseline $\leq$ 20 km, developed to produce the figure 8.
   - *R scripts*: Scripts are developed to generate figures 9 – 10. Data used for this investigation are from a baseline of $\sim$ 41 km between HFSS and SAND, year=2014, and day-of-year= 85.
3. *Rover data processing*: R scripts are developed and used to generate the figures 12 – 14.

Acknowledgments

The international GNSS Service (IGS) is acknowledged for providing geodetic infrastructure and geodetic products used in this work.

The authors would like to thank Tor O. Dahlø from the SATREF group at the Norwegian Mapping Authority for providing the GNSS data. Without his effort, this investigation would not have been possible.

Prof. John Raquet is acknowledged for kindly making parts of the source code for NetAdjust available.

The authors also thank the editor and the reviewers for constructive feedback that helped to improve the original manuscript.

The authors would also like to thank Jabir A. Ouassou for constructive discussions and proofreading of this paper.

References

Adler RJ, Taylor JE (2009) Random Fields and Geometry. Springer Monographs in Mathematics, Springer New York

Baarda W (1968) A Testing Procedure for Use in Geodetic Networks. v. 2, no. 5, Netherlands Geodetic Commission

Berger JO (1985) Statistical Decision Theory and Bayesian Analysis. Springer Series in Statistics, Springer

Borre K (2006) Mathematical Foundation of Geodesy: Selected Papers of Torben Krarup. Springer Berlin Heidelberg

Brown RG, Hwang PYC (1997) Introduction to random signals and applied Kalman filtering: with MATLAB exercises and solutions. v. 1, Wiley

Chatterjee S, Hadi AS (2009) Sensitivity Analysis in Linear Regression. Wiley Series in Probability and Statistics, Wiley

Chen X, Timo A, Cao W, Ferguson K, Grünig S, Gomez V, Kipka A, Köhler J, Landau H, Leandro R, Gang Lu RS, Z NT (2011) Trimble RTX, an Innovative New Approach for Network RTK. Trimble TerraSat GmbH, Germany

Commission RT, Services fM (2006) RTCM Standard 10403.1 for Differential GNSS (Global Navigation Satellite Systems) Services: Version 3. Radio Technical Commission for Maritime Services

Cressie NAC (1993) Statistics for Spatial Data, revised edn. Wiley Series in Probability and Mathematical Statistics, Wiley

Dach R, Lutz S, Walser P, Fridez P (2015) Bernese GNSS Software, 5th edn. Astronomical institute, University of Bern, Bern

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. Journal of the royal statistical society, series B 39(1):1–38

Djauhari M, Umbara R (2007) A redefinition of Mahalanobis depth function. Malaysian Journal of Fundamental and Applied Sciences 3:150–157

El-Rabbany AE (1994) The effect of physical correlation on the ambiguity resolution and accuracy estimation in GPS differential positioning. Department of Geodesy and Geomatics Engineering, University of New Brunswick 32(141)

Euler HJ, Keenan CR, Zebhauser BE, Wübbena G (2001) Study of a Simplified Approach in Utilizing Information from Permanent Reference Station Arrays. In: Proceedings of the National Technical Meeting of the Satellite Division of the Institute of Navigation, (ION GPS 2001), vol 104, pp 371–391

Feng S, Ochieng W, Moore T, Hill C, Hide C (2009) Carrier phase-based integrity monitoring for high-accuracy positioning. GPS Solutions 13(1):13–22, DOI 10.1007/s10291-008-0093-0

Grewal MS, Weill LR, Andrews AP (2007) Global Positioning Systems, Inertial Navigation, and Integration, vol 14, 2nd edn. John Wiley & Sons, Inc.

Heywood HB (1931) On Finite Sequences of Real Numbers. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character pp 486–501

Hofmann-Wellenhof B, Moritz H (2006) Physical Geodesy. Springer Vienna

Johnson RA, Wichern DW (2002) Applied Multivariate Statistical Analysis. Prentice Hall

Kaplan ED, Hegarty CJ (2006) Understanding GPS Principles and Applications, 2nd edn. 2006 ARTECH HOUSE, INC

Kleusberg A, Teunissen PJG (1998) GPS for geodesy. Environmental science, Springer

Kok JJ (1982) Statistical analysis of deformation problems using Baarda's testing procedures. In: Forty years of Thought, Delft 2:470–488

Kuusniemi H, Wieser A, Lachapelle G, Takala J (2007) User-level reliability monitoring in urban personal satellite -navigation. IEEE Transactions on Aerospace and Electronic Systems 43(4)

Landau H, Vollath U, Chen X (2002) Virtual Reference Station Systems. Journal of Global Positioning Systems 1(2):137–143, DOI 10.5081/jgps.1.2.137

Langley RB (1997) GPS Receiver System Noise. GPS Solutions pp 40–45

Leandro R, Landau H, Nitschke M, Glocker M, Seeger S, Xiaoming C, Deking A, Zhang MBF, Kendall Ferguson, Ralf Stolz NT, Gang Lu, Timo Allison, Markus Brandl, Victor Gomez, Wei Cao AK (2011) RTX Positioning: The Next Generation of cm-accurate Real-Time GNSS Positioning. In: ION GNSS 2011, Portland, OR, pp 1460–1475

Leick A (2015) GPS Satellite Surveying. John Wiley and Sons, Incorporated

Liu RY, Serfling RJ, Souvaine DL (2003) Data Depth: Robust Multivariate Analysis, Computational Geometry, and Applications. DIMACS series in discrete mathematics and theoretical computer science, American Mathematical Soc.

McCullagh P, Nelder JA (1989) Generalized Linear Models, Second Edition. Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis

Ming Y, Clyde G, Schaffrin B (1994) Real-time On-the-Fly Ambiguity Resolution Over Short Baselines in the Presence of Anti-Spoofing. Proceedings of the 7th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1994) pp 519–525

Mosler K (2013) Depth Statistics. In: Robustness and Complex Data Structures, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 17–34, DOI 10.1007/978-3-642-35494-6{_}2

Odolinski R (2012) Temporal correlation for network RTK positioning. GPS Solution 16:147–155, DOI 10.1007/s10291-011-0213-0

Ouassou M, Jensen ABO, Gjevestad JGO, Oddgeir K (2015) Next Generation Network Real-Time Kinematic Interpolation Segment to Improve the User Accuracy. International Journal of Navigation and Observation 2015:1–15, DOI 10.1155/2015/346498

Ouassou M, Kristiansen O, Gjevestad JGO, Jacobsen KS, Andalsvik YL (2016) Estimation of Scintillation Indices: A Novel Approach Based on Local Kernel Regression Methods. International Journal of Navigation and Observation 2016:1–18, DOI 10.1155/2016/3582176

Pullen S, Enge P, Parkinson B (1995) A New Method for Coverage Prediction for the Wide Area Augmentation System (WAAS)

Ramjee P, Ruggieri M (2005) Applied satellite navigation using GPS, GALILEO, and augmentation systems. Boston ; London : Artech House, 2005

Raquet J, Lachapelle G (1999) Development and Testing of a Kinematic Carrier-Phase Ambiguity Resolution Method Using a Reference Receiver Network 1. Navigation 46(4):283–295, DOI 10.1002/j.2161-4296.1999.tb02415.x

Raquet JF (1998) UCGE Reports Number 20116 Development of a Method for Kinematic GPS Carrier-Phase Ambiguity Resolution Using Multiple Reference Receivers By (20116)

Rogers RM (2003) Applied Mathematics in Integrated Navigation Systems. No. v. 1 in AIAA education series, American Institute of Aeronautics and Astronautics

Rousseeuw PJ, Leroy AM (2003) Robust Regression and Outlier Detection. Wiley Series in Probability an, Wiley

Schabenberger O, Gotway CA (2004) Statistical Methods for Spatial Data Analysis. Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis

Schön S, Brunner FK (2008) A proposal for modelling physical correlations of GPS phase observations. Journal of Geodesy 82(10):601–612, DOI 10.1007/s00190-008-0211-3

Seeber G (2003) Satellite Geodesy, 2nd edn. Walter de Gruyter GmbH & Co

Shanmugan KS, Breipohl AM (1988) Random Signals: Detection, Estimation and Data Analysis. Wiley, URL https://books.google.no/books?id=dvlSAAAAMAAJ

Somesh Dasgupta (1995) The evolution of the D^2 statistic of the Mahalanobis. Indian journal of pure applied Math 26(6):485–501

Takac F, Zelzer O (2008) The Relationship Between Network RTK Solutions MAC, VRS, PRS, FKP and i-MAX. Proceedings of the 21st International Technical Meeting of the Satellite Division of the Institute Of Navigation (ION GNSS 2008) pp 348–355

Teunissen P (1990) An integrity and quality control procedure for use in multi sensor integration. Proceedings of the 3rd international technical meeting of the satellite division of the institute of navigation (ION GPS 1990), Colorado Spring pp 513–522

Teunissen PJG (1985) Quality Control in Geodetic Networks. In: Optimization and Design of Geodetic Networks, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 526–547, DOI 10.1007/978-3-642-70659-2{_}18

Timm NH (2007) Applied Multivariate Analysis. Springer Texts in Statistics, Springer New York

Tomoji T, Yasuda A (2009) Development of the low-cost RTK-GPS receiver with an open source program package

Wieser A, Brunner FK (2000) An extended weight model for GPS phase observations. Earth Planets Space 52:777–782

Wübbena G, Willgalis S (2001) State Space Approach for Precise Real Time Positioning in GPS Reference Networks. In: Proceeding of international Symposium on kinematic Systems in Geodesy, Geomatics and Navigation, KIS-01, Banff, Canada

Wübbena G, Bagge A, Seeber G, Boeder V, Hankemeier P (1996) Reducing distance dependent errors for real-time precise DGPS applications by establishing reference station networks. In: Proceedings of Ion Gps, vol 9, pp 1845–1852

Zebhauser B, Euler H, Keenan C, Wübbena G (2002) A novel approach for the use of information from reference station networks conforming to RTCM V2. 3 and future V3. 0. In: Proceedings ofION NTM, San Diego, CA, USA, pp 28—-30

Paper #5

Reference

M. Ouassou, B. Natvig, Anna B. O. Jensen & J. I. Gåsemyr
Reliability Analysis of Network Real-Time Kinematic
Submitted to the International Journal of Navigation and Observation: 30/11/2016

Contribution

M. Ouassou conceived the idea, performed the numerical calculations, derivation of a novel algorithm *Penalized honored average standard deviation*, and wrote the majority of the manuscript. Multi-state reliability tool is used for the first time to carry out the reliability of NRTK data processing. All authors contributed to the discussion of the results of the manuscript.

B. Natvig and J. I. Gåsemyr contributed to the selection of the right investigation tool (multi-state reliability). Contributed to section 4.

A. B. O. Jensen re-write sections 1 and 3.1.

International Journal of Navigation and Observation manuscript No.
(will be inserted by the editor)

Reliability Analysis of Network Real-Time Kinematic

Mohammed Ouassou · Bent Natvig · Anna B. O. Jensen · Jørund I. Gåsemyr

Received: date / Accepted: date

Abstract In this paper, the multi-state reliability theory was applied to the network real-time kinematic (NRTK) data processing chain, where the quality of the network corrections, baseline residuals and the associated variance-covariance matrices are considered as the system state vectors. The state vectors have direct influence on the rover receiver position accuracy. The penalized honored stochastic averaged standard deviation (PHSASD) is used to map the NRTK sensitive data, represented by the states vectors to different levels of performance. The study shows that the improvement is possible by identification of critical components in the NRTK system and implementation some parallelism that makes the system more robust.

Keywords Global Navigation Satellite Systems · Network Real-Time Kinematic · Multi-state system · Reliability Evaluation · optimization

Mohammed Ouassou
Norwegian Mapping Authority, Geodetic Institute
3511 Hønefoss, Norway
Tel.: +47 32 11 83 76
E-mail: mohammed.ouassou@statkart.no

Bent Natvig
UiO, Department of Mathematics, Norway
Tel.: +47 22 855872
E-mail: bent@math.uio.no

Anna B. O. Jensen
KTH Royal Institute of Technology 10044 Stockholm, Sweden.
Tel.: +46 8-790 7353
E-mail: anna.jensen@abe.kth.se

Jørund I. Gåsemyr
UiO, Department of Mathematics, Norway
Tel.: +47 22855960
E-mail: gaasemyr@math.uio.no

1 Introduction

High accuracy positioning with GNSS is carried out using both code and carrier phase data from the GNSS satellites. To obtain position accuracies at the cm or mm level using carrier phase data, an important part of the data processing is to estimate the initial oscillator phase offset, the so-called ambiguity, for each receiver-satellite pair. Resolution of ambiguities requires that the influence of most errors sources in the positioning process is reduced to the cm-level, and high accuracy positioning is therefore often done in a relative mode where the position of a GNSS receiver located in an unknown position (the rover) is determined relative to one or more reference stations located in positions known on beforehand [7].

With relative carrier phase based GNSS positioning the effects of the distance dependent error sources such as uncertainties in satellite positions, and atmospheric effects on the satellite signals induced by the ionosphere and troposphere are reduced. Also the effects of satellite and receiver clock errors in the positioning process are reduced by relative positioning, and all this in combination makes it possible to resolve the ambiguities and thereafter obtain positions for the rover at the cm or mm level.

For high accuracy GNSS positioning in real time, the real time kinematic (RTK) technique has been developed. Traditionally this is based on a reference station transmitting data to the rover where the data is used in estimation of the position of the rover in a relative or differential mode [14].

Using a network of reference stations for RTK, the so called NRTK technique, provides the opportunity for applying more advanced algorithms for estimation of the distance dependent errors within the network, and thereby possibilities for providing a more robust service. Such operational NRTK services exists in many regions and countries today, and have

become an indispensable tool in high accuracy navigation and surveying.

A brief description of the NRTK functionality is as follows: The first step is collection of raw observations from the network of reference stations, solve for the ambiguities within the reference network, and generate error estimates. Then an interpolation/smoothing scheme is applied to generate the NRTK corrections for the user location. For information on how to avoid a loss of information under interpolation of NRTK data, the interested reader is referred to [19].

The NRTK corrections are then transmitted to rover receivers. Several NRTK techniques exist and the most commonly used at present are for instance the Master Auxiliary Concept (MAC) [3, 24], the Virtual Reference Station (VRS) concept [13], the FKP techniques [27], as well as the Network Adjustment (NetAdjust) concept developed by John F. Raquet [20, 21].

Multi-state system theory has been a research topic for many years. For instance, extension of the system from two-state to multi-state reliability [10, 29] to compute the mean performance level at any given time t, stochastic evaluation and bound computation of multi-state coherent systems [18]. Application of reliability analysis to GNSS data processing [14], a comparative GNSS reliability analysis [8], reliability analysis under GNSS weak signals [23], accuracy and reliability of multi-GNSS real-time precise positioning [15] and robust reliability testing in case of signal degradation environment [1, 12].

In many practical situations, the system and the corresponding components could take different performance levels ranging from the perfect functioning state to the complete failure state. The multi-state reliability can help us to understand and solve practical problems, for instance the NRTK data processing and algorithms selection.

In order to apply reliability analysis to a NRTK system, the starting point is the decomposition of the block diagrams of NRTK processing chain into simple components, and computation of the system reliability. Figure 1 shows three levels of data processing modules, the network, the baseline and the rover receiver modules, where R denotes rover, S denotes reference station, and (dX, dY, dZ) denote errors in the position X, Y, Z coordinates respectively. Based on these levels, we can build the reliability block diagrams for the NRTK processing chain and compute the reliability for the entire system.

The rest of the paper is organized as follows:

Sec. 2: Brief introduction to the traditional reliability theory, deterministic and stochastic reliability.
Sec. 3: NRTK blocks diagram determination and reliability computations.
Sec. 4: Multi-state reliability theory applied to NRTK processing chain.
Sec. 5: This section presents some test results.

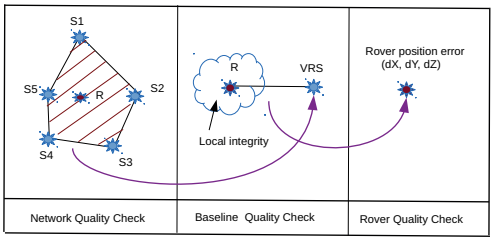


Figure 1: NRTK processing chain. The left panel shows NRTK module that produces the network corrections. The curved line indicates the output generated by the network module. Middle panel shows the local and the baseline processing module. Right panel shows the rover position solution module

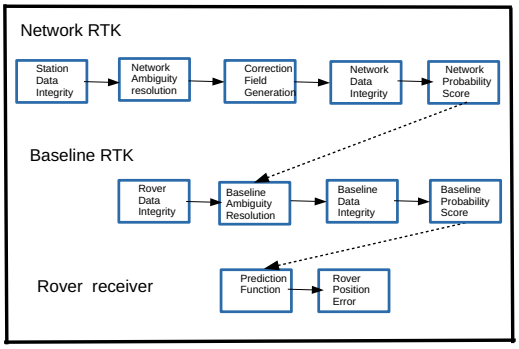


Figure 2: Network real-time kinematic processing chain main modules. Dotted lines describe the components dependency from different levels.

Sec. 6: Procedures used to validate the NRTK system reliability are introduced in this section.
Sec. 7: Discussions and conclusion.

Test data used in this investigation is described in Appendix A.1.

2 Reliability analysis

The aim of this section is to introduce the basic of the traditional reliability theory and decompose the block diagrams of NRTK processing chain into simple components and compute the system reliability. The description of the structural relationship between the components and the system must be defined. Figure 2 illustrates the concept and each main block will be treated separately in the coming sections.

2.1 Structure functions

In order to construct the entire reliability block diagram of the NRTK data integrity, it's necessary to define the elementary building blocks as Bernoulli indicator function [22, pp. 28–29]. The function will signal if a unit or system is functioning or not.

A random variable X is said to be a Bernoulli random variable if its probability function is given by Equation (1). The indicator and structure functions are given below by following the notation given by Natvig [16].

$$X_i = \begin{cases} 1\,, & \text{success with probability } p \\ 0\,, & \text{failure with probability } (1-p) \end{cases} \tag{1}$$

2.1.1 Bernoulli indicator function

Let x_i denote the indicator function of component, i. Then we have:

$$x_i = \begin{cases} 1\,, & \text{if the i'th component is functioning} \\ 0\,, & \text{otherwise} \end{cases} \tag{2}$$

2.1.2 System structure function

Let the state vector $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ give which components are functioning and which are not.

Let $\phi(\boldsymbol{x})$ denote the Bernoulli indicator of the state vector, $\boldsymbol{x}$. Then we have:

$$\phi(\boldsymbol{x}) = \begin{cases} 1\,, & \text{if the system is functioning} \\ 0\,, & \text{otherwise} \end{cases} \tag{3}$$

$\phi(\boldsymbol{x})$ is the structure function of the system.

2.1.3 Series structure function

The series structure function $\phi_s(\boldsymbol{x})$ works *if and only if all components* of the state vector $\boldsymbol{x}$ are functioning. The series structure function $\phi_s(\boldsymbol{x})$ reads:

$$\phi_s(\boldsymbol{x}) = \min(x_1, x_2, \cdots, x_n) = \Pi_{i=1}^{n} x_i \tag{4}$$

2.1.4 Parallel structure function

The parallel structure function $\phi_p(\boldsymbol{x})$ works *if and only if at least one* of the components of the state vector $\boldsymbol{x}$ is functioning. The parallel structure function $\phi_p(\boldsymbol{x})$ reads:

$$\begin{aligned} \phi_p(\boldsymbol{x}) &= \max(x_1, x_2, \cdots, x_n) \\ &= 1 - \Pi_{i=1}^{n}(1 - x_i) \\ &= \amalg_{i=1}^{n} x_i \end{aligned} \tag{5}$$

where $\amalg$ is read "ip" and denotes the parallel coupling operator.

2.1.5 k-out-of-n structure function

A system composed of n components which is functioning *if and only if at least* k components are functioning is called a *k-out-of-n structure*.

$$\phi_{k,n}(\boldsymbol{x}) = \begin{cases} 1\,, & \text{if } \sum_{i=1}^{n} x_i \geq k \\ 0\,, & \text{otherwise} \end{cases} \tag{6}$$

Note that the structure functions given by equations (4) and (5) are an *n-out-of-n structure*, and an *1-out-of-n structure*, respectively.

Other structure functions exist for instance the bridge Natvig [16, p. 12].

2.2 System reliability computation

The structure functions are defined, now is time to compute the system reliability.

We move from the deterministic model to the stochastic one by introducing the random variables. Some notation is needed to represent the state vector, structure function and the reliability. We follow the notations given by Høyland and Rausand [9, Chaps. 3-5].

We denote the state variables of the n independent component at time t by

$$X_1(t), X_2(t), \cdots, X_n(t)$$

The corresponding state vector and the structure function are denoted, respectively, by

$$\boldsymbol{X}(t) = (X_1(t), X_2(t), \cdots, X_n(t)) \quad \text{and} \quad \phi(\boldsymbol{X}(t))) \tag{7}$$

The probabilities of interest are presented by:

$$P(\boldsymbol{X}_i(t) = 1) = p_i(t) \quad \text{for} \quad i = 1, 2, \cdots, n \tag{8}$$

$$P(\phi(\boldsymbol{X}(t) = 1) = p_s(t) \tag{9}$$

where $p_i(t)$ is the component reliability while $p_s(t)$ is the system reliability. Assuming that the components are independent, then the computation of the reliability of the state vector $\boldsymbol{X}(t)$ and the system $\phi(\boldsymbol{X}(t)))$ at time t are defined as the expectation operator.

$$\begin{aligned} \mathbb{E}(X_i(t)) &= 0\,.\,P(X_i(t) = 0) + 1\,.\,P(X_i(t) = 1) \\ &= p_i(t) \quad \text{for} \quad i = 1, 2, \cdots, n \end{aligned} \tag{10}$$

Let R denote the system reliability, then we have:

$$\begin{aligned} \mathbb{E}(\phi(\boldsymbol{X}(t))) &= p_s(t) \\ &= R(p_1(t), p_2(t), \cdots, p_n(t)) \\ &= R(\boldsymbol{p}(t)) \end{aligned} \tag{11}$$

To avoid confusion, $p_i = P(X_i(t) = 1)$ is the probability of functioning of the i'th component and referred to as the component reliability.

2.2.1 Reliability of series structures

The reliability function $R_s(\boldsymbol{p})$ of the series system of n independent components is given by the expression:

$$\begin{aligned} R_s(\boldsymbol{p}) &= P(\phi(\boldsymbol{X}(t)=1)) \\ &= P\{X_i(t)=1 \quad \text{for all} \quad i=1,2,\cdots,n\} \\ &= \Pi_{i=1}^{n} p_i(t) \end{aligned} \tag{12}$$

If all components have the same $p(t)$, the Equation (12) becomes $\{p(t)\}^n$. For $n=5$ and $p=0.99$, then the reliability $R_s(\boldsymbol{p})=0.951$.

An important remark is that the reliability of a series structure is at most as reliable as the least reliable component, that is $R_s(\boldsymbol{p}) \leq \min_i(p_i(t))$.

2.2.2 Reliability of parallel structures

The reliability function $R_p(\boldsymbol{p})$ of the parallel system of n independent components is given by the expression:

$$\begin{aligned} R_p(\boldsymbol{p}) &= P(\phi(\boldsymbol{X}(t)=1)) \\ &= P\{X_i(t)=1 \quad \text{for some} \quad i=1,2,\cdots,n\} \\ &= 1-P\{X_i(t)=0 \quad \text{for all} \quad i=1,2,\cdots,n\} \\ &= 1-\Pi_{i=1}^{n}(1-p_i(t)) \end{aligned} \tag{13}$$

If all components have the same $p(t)$, the Equation (13) becomes $1-\{1-p(t)\}^n$. For $n=5$ and $p=0.99$, then the reliability $R_p(\boldsymbol{p})=1$.

An important remark is that the reliability of a parallel structure is at least as reliable as the most reliable component, that is $R(\boldsymbol{p}) \geq \max_i(p_i(t))$.

Details on how to compute the reliability of parallel structure in general is given in Appendix A.3.

2.2.3 Reliability of k-out-of-n structures

The reliability function $R_{(k-n)}(\boldsymbol{p})$ of the k-out-of-n system of n independent components with equal probability $p_i(t)=p$ is given by the expression:

$$\begin{aligned} R_{(k-n)}(\boldsymbol{p}) &= P(\phi(\boldsymbol{X}(t)=1)) \\ &= P\Big\{\sum_{i=1}^{n} X_i(t) \geq k\Big\} \\ &= \sum_{i=k}^{n} \binom{n}{i} \cdot p^i(1-p)^{n-i} \end{aligned} \tag{14}$$

3 Reliability in the NRTK processing chain

The aim of this section is to determine the structure functions, NRTK module's reliability, and the corresponding block diagrams.

3.1 Considerations around NRTK data processing

Some considerations around the NRTK data integrity are introduced below as a background for the design process and to ease the discussions in the following sections. For more information on GNSS data processing, the reader is referred to [7, 11, 14, 28]. The key to precise positioning is the correct ambiguity resolution and validation. With ambiguities resolved to wrong integer numbers, there will be offsets in the position solution, and with float ambiguities (ambiguities that are not fixed to integer values) the position solution is in-accurate and also very unstable and sensitive to changes in satellite geometry.

Good satellite-receiver geometry, as for instance expressed by the so-called DOP factor (dilution of precision) is important to perform successful ambiguity resolution and achieve centimeter level accuracy in real time.

Spatio-temporal models that describe well the variations of the spatially correlated errors in the corrections field is also an important key for reliable NRTK positioning.

Robust estimation algorithms to handle large data sets are also a key factor becoming more important in the future as observations from several GNSS systems to a larger degree will be combined in one processing loop. Today, most NRTK systems operate with data from the American GPS and the Russian GLONASS system. Including data from the European Galileo as well as the Chinese Beidou systems in NRTK operations will soon be the norm for most NRTK services. With satellites from more GNSS systems being available the satellite-receiver geometry on the rover side is improved. This is especially important when the user is operating in constricted environments such as narrow street canyons or forest areas.

3.2 NRTK corrections reliability analysis

The main function of the NRTK is to provide the rover in the field with high quality corrections on an epoch-by-epoch basis.

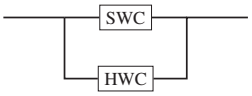
From Figure 2, five modules $M_{N,i}$ are defined for the network and the reliability of each module will be evaluated. $M_{i,j}$ corresponds to level $i \in \{N,B,R\}$ and module $j \in \{1,2,\cdots,5\}$, where N stands for network, B for baseline and R for rover receiver. For instance the module $M_{N,1}$ corresponds to the first network module, the Station Data Integrity, as shown in Figure 2, and $M_{R,1}$ denotes the first of the rover modules, i.e. the Prediction Function as shown in Figure 2.

3.2.1 Reference receivers data integrity:

Generation of high quality raw observations at reference receivers require a reliable hardware (HWC) and components (SWC). Expensive hardwares and sophisticated algorithms

are keys to achieve this goal. Figure 3 shows the concept and each component will be treated separately in the next sections.

Figure 3: Main block diagram of module $M_{N,1}$



Let software and hardware components be represented by the modules $M_{N,1,X}$ and $M_{N,1,Y}$, respectively.

Software component of module $M_{N,1}$ definition

SWC requires an ensemble of sequential checks on raw observations. This includes:

- Let $x_{1,1}$ denote the satellite data integrity algorithm. This algorithm will discard the measurements from unhealthy satellite(s) or from satellite(s) for which we do not have the orbital data.
- Let $x_{1,2}$ denote controlled cycle-slip algorithm. This task requires investigation of carrier-phase discontinuities by examination of loss of lock (LLI) indicator and signal-to-noise ratio (SNR) flags.
- Let $x_{1,3}$ denote uncontrolled cycle-slips in the observations. The algorithm uses the observation combinations for this purpose. The interested reader is referred to [28, pp. 95–101].
- Let $x_{1,4}$ denote the reference receiver clock offset reset algorithm. Continual corrections are carried out to reduce the effect of the jump. The receiver clock offset (jump with ±1 ms) must be detected and corrected because they cause jump in carrier-phase.
- Let $x_{1,5}$ denote outliers detection and repair algorithm.
- Let $x_{1,6}$ denote the low elevation angle. Algorithm that prune satellite(s) based on their low elevation angle.
- Let $x_{1,7}$ denote the minimum observations required to generate the corrections. At least 4 observation types are needed (L_1, L_2, P_1 and P_2). For more information about the observation types provided by satellites, the reader is referred to [5].
- Let $x_{1,8}$ denote the reference receiver clock stability algorithm.
- Let $x_{1,9}$ denote the multipath mitigation algorithm.
- Let $x_{1,10}$ denote the reweighting algorithm. All units are parallel coupled.
  1. Let $x_{[1,10,1]}$ denote a low elevation reweighting algorithm.
  2. Let $x_{[1,10,2]}$ denote a scintillation reweighting algorithm.
  3. Let $x_{[1,10,3]}$ denote a signal-to-noise reweighting algorithm.

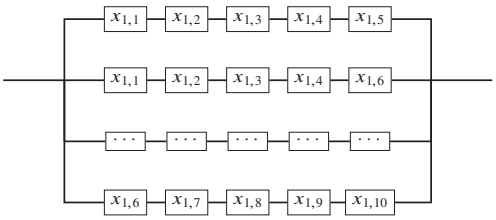
  Figure 5 shows the block diagram of the network reweighting algorithm component $x_{1,10}$.

Block diagram and reliability of software component

The structure function of the software component is well described by the 5-out-of-10 structure function. This means that the more algorithms check the more reliable raw observations become.

In order to produce a reliable raw observations of high quality, it is necessary to perform at least five checks from a total of ten. With ten algorithms check, we can generate a high quality raw observations while five algorithms check will produce an acceptable level raw observations. The selection of the algorithms is independent of the order.

Figure 4: Block diagram of software component of network module $M_{N,1}$



The structure function $\Phi(\boldsymbol{x})$ of Figure 4 is given by the expression

$$\begin{aligned}\Phi(\boldsymbol{x}) = \max(&\min(x_{1,1}, x_{1,2}, x_{1,3}, x_{1,4}, x_{1,5}),\\ &\min(x_{1,1}, x_{1,2}, x_{1,3}, x_{1,4}, x_{1,6}),\\ &\vdots\\ &\min(x_{1,6}, x_{1,7}, x_{1,8}, x_{1,9}, x_{1,10}))\end{aligned} \tag{15}$$

Assuming that the individual algorithms are independent with equal probability $p(t)$, then the reliability of software component is

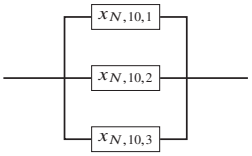
$$p_s(t) = P(\boldsymbol{y}(t) \geq 5) = \sum_{x=5}^{10} \binom{10}{x} p^x (1-p)^{(10-x)} \tag{16}$$

Due to the fact that $\boldsymbol{y}(t) \sim \text{binom}(n, p(t))$.

The block diagram of the $x_{N,1,10}$ is given by The corresponding reliability reads

$$\begin{aligned}p_{N,1,10} = &(p_{N,10,1} + p_{N,10,2} + p_{N,10,3}) -\\ &(p_{N,10,1}p_{N,10,2} + p_{N,10,1}p_{1,10,3} + p_{N,10,2}p_{N,10,3}) +\\ &p_{N,10,1}\, p_{N,10,2}\, p_{N,10,3}\end{aligned} \tag{17}$$

Figure 5: Block diagram of the network unit $x_{1,10}$



Hardware component of module $M_{N,1}$

The quality of the GNSS receiver, firmware robustness, GNSS antenna and choke ring quantify the hardware component. Figure 6 shows the block diagram of the hardware component of the network module $M_{N,1}$.

The elements of the hardware component of the network module $M_{N,1}$ are:

- Let $y_{1,1}$ denote the GNSS receiver type
- Let $y_{1,2}$ denote the rover software known as firmware to decode the GNSS signals.
- Let $y_{1,3}$ denote the GNSS antenna type.
- Let $y_{1,4}$ denote the choke ring that allows better reception of low elevation angle GPS satellites and improved multipath rejection.
- Let$y_{1,5}$ denote duplicated system (as discussed in the next section).

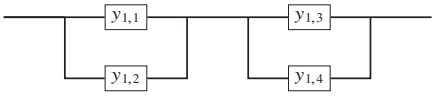


Figure 6: Block diagram of hardware component of module $M_{N,1}$. The left block describes the GNSS receiver and the right is for the GNSS antenna
,

The structure function $\Phi(\boldsymbol{y})$ is given by the expression:

$$\begin{aligned}\Phi(\boldsymbol{y}) &= \min(\max(y_{1,1}, y_{1,2}), \max(y_{1,3}, y_{1,4})) \\ &= (y_{1,1} \amalg y_{1,2})(y_{1,3} \amalg y_{1,4}) \\ &= (y_{1,1} + y_{1,2} - y_{1,1}y_{1,2})(y_{1,3} + y_{1,4} - y_{1,3}y_{1,4}) \end{aligned} \quad (18)$$

The reliability of the hardware of the module $M_{N,1}$ reads

$$p_s(t) = (p_{1,1} + p_{1,2} - p_{1,1}p_{1,2})(p_{1,3} + p_{1,4} - p_{1,3}p_{1,4}) \quad (19)$$

Hardware component improvement

In order to ensure continuous raw data delivery at the station, a duplicated system is recommended. The drawback is the financial issues. HWC reliability reads:

$$p(t) = p_s(t) + p_s(t) - [p_s(t)]^2 \quad (20)$$

3.2.2 Network ambiguity resolution

As mentioned in Sect. 3.1. The key for precise positioning is the correct ambiguity resolution and validation. The module $M_{N,2}$ is composed of two main parts, the first part is the ambiguity processing algorithms performance and the second part is the statistical ambiguity quality indicators derived from the processing chain of the ambiguity.

The components of the module $M_{N,2}$ are:

- Let $x_{2,1}$ denote the float solution of the ambiguity obtained via least square or Kalman filter.
- Let $x_{2,2}$ denote the LAMBDA method [25] applied to a float solution to reduce the search space and to obtain a fix solution.
- Let $x_{2,3}$ denote the validation procedures to validate the final solution.
- Let $x_{2,4}$ denote the administration of the ambiguities.
- Let $x_{2,5}$ denote the success rate of the ambiguities resolution.
- Let $x_{2,6}$ denote the ambiguity dilution of precision (ADOP) [26]. ADOP measures the precision of the ambiguities, and can be viewed as a quality indicator.
- Let $x_{2,7}$ denote time to fix.

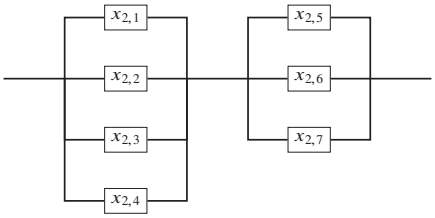


Figure 7: Network ambiguity resolution module $M_{N,2}$. The first block of this figure represents the ambiguity processing algorithms and the second block is the ambiguity quality indicators.

Structure function of $M_{N,2}$

The structure function of the module $M_{N,2}$ reads

$$\begin{aligned}\Phi(\boldsymbol{x}) &= \min(\max(\Phi_1(x)), \max(\Phi_2(x))) \\ &= \min(\max(x_{2,1}, x_{2,2}, x_{2,3}, x_{2,4}), \max(x_{2,5}, x_{2,6}, x_{2,7})) \\ &= [1 - \Pi_{i=1}^{4}(1 - x_{2,i})][1 - \Pi_{i=5}^{7}(1 - x_{2,i})] \end{aligned} \quad (21)$$

Reliability of $M_{N,2}$

Reliability of the module $M_{N,2}$ reads

$$p_s(t) = \left\{1 - \Pi_{i=1}^{4}(1 - p_{2,i})\right\}\left\{1 - \Pi_{i=5}^{7}(1 - p_{2,i})\right\} = \min(P_1, P_2) = P_1 P_2 \quad (22)$$

dropping the index 2 in the expression of $p_s(t)$ and compute P_1 and P_2.

$$P_1(t) = \sum_{i=1}^{4} p_i - \Pi_{i=1}^{4} p_i - \left\{p_1p_2 + p_1p_3 + p_1p_4 + p_2p_3 + p_2p_4 + p_3p_4\right\} + \left\{p_1p_2p_3 + p_1p_2p_4 + p_1p_3p_4 + p_2p_3p_4\right\} \quad (23)$$

$$P_2(t) = \sum_{i=5}^{7} p_i + \Pi_{i=5}^{7} p_i - \left\{p_5p_6 + p_5p_7 + p_6p_7\right\} \quad (24)$$

3.2.3 Network corrections quality

The quality of the network corrections depends on various parameters, for instance the estimation algorithms, network status (sparse/dense), the covariance functions and the smoothing/interpolation algorithms. The components of the module $M_{N,3}$ are:

- Let $x_{3,1}$ denote the network reference receivers separation. Dense network is attractive because the network corrections are better estimated with short distances between reference receivers.
- Let $x_{3,2}$ denote the quality of the estimation algorithm used to estimate the parameter vector $\theta \in \Theta$.
- Let $x_{3,3}$ denote the quality of the covariance function used to model the network correlation errors.
- Let $x_{3,4}$ denote the quality of the interpolation algorithm used to generate the user corrections. Parallel interpolation algorithms will enhance the quality of the user corrections generation and avoid the information loss.

Block diagram of $M_{N,3}$

The block diagram of the module $M_{N,3}$ is given by the Figure 8.

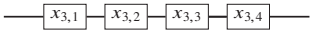


Figure 8: Block diagram of network correction quality module $M_{N,3}$

Structure function of $M_{N,3}$

The structure function of the module $M_{N,3}$ reads

$$\Phi(\boldsymbol{x}) = \min(x_{3,1}, x_{3,2}, x_{3,3}, x_{3,4}) = \Pi_{i=1}^{4} x_{3,i} \quad (25)$$

Reliability of $M_{N,3}$

Dropping the index 3, the reliability of $M_{N,3}$ reads

$$p_s(t) = \Pi_{i=1}^{4} p_{3,i} \quad (26)$$

Amelioration potential of $M_{N,3}$

Our aim is to provide the user in the field with high quality corrections on an epoch-by-epoch basis. The interpolation/smoothing algorithm plays a central role. We can implement different parallel interpolation/smoothing algorithms that compete about the quality of service parameters. The corrections will be sent from the algorithm with higher score. For more information on this topic, the interested reader is referred to [19].

In this case, replacing the component $x_{3,4}$ with a parallel structure function $\phi(x_{3,4}) = max(y_1, y_2, \cdots, y_5)$, and the computation of the new block diagram is straightforward. Let q_i denote the functioning probability of unit y_i (Figure 9)

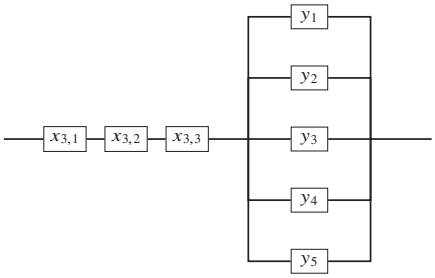


Figure 9: Improved block diagram of module $M_{N,3}$. The component $x_{3,4}$ is replaced by a parallel structure function $\phi(y)$

for $i = 1, 2, \cdots, 5$. Then the reliability of $x_{3,4}$ reads

$$p_{s_{(3,4)}}(t) = \sum_{i=1}^{5} q_i + \Pi_{i=1}^{5} q_i - \left\{q_1q_2 + q_1p_3 + q_1p_4 + q_1p_5 + q_2q_3 + q_2q_4 + q_2q_5 + q_3q_4 + q_3q_5 + q_4q_5\right\} + \left\{q_1q_2q_3 + q_1q_2q_4 + q_1q_2q_5 + q_2q_3q_4 + q_2q_3q_5 + q_3q_4q_5\right\} - \left\{q_1q_2q_3q_4 + q_1q_2q_3q_5 + q_2q_3q_4q_5\right\} \quad (27)$$

The improved reliability of module $M_{N,3}$ reads

$$p_s = p_{s_{(3,4)}} \ \Pi_{i=1}^{3} p_{3,i} \tag{28}$$

3.2.4 Network data integrity

The module $M_{N,4}$ responsibility is to carry out the quality control on the corrections field and the corresponding variance-covariance matrices. This includes:

- Let $x_{4,1}$ denote global test statistics to detect any extremal events that can bias the rover position. Corrections field investigation.
- Let $x_{4,2}$ denote the inspection of variance-covariance matrices for Heywood effects algorithm [6].
- Let $x_{4,3}$ denote the application of the imputation algorithm to compute the statistics.
- Let $x_{4,4}$ denote the total variance monitoring algorithm.
- Let $x_{4,5}$ denote the generalized variance monitoring algorithm.

Serial coupling is the appropriate choice for the module $M_{N,4}$ and the block diagram is given by the Figure 10.

Block diagram of $M_{N,4}$

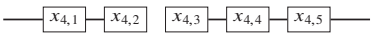


Figure 10: Block diagram of module $M_{N,4}$

Structure function of $M_{N,4}$

The structure function of the module $M_{N,4}$ reads

$$\Phi(\boldsymbol{x}) = \min(x_{4,1}, x_{4,2}, x_{4,3}, x_{4,4}, x_{4,5}) = \Pi_{i=1}^{5} x_{4,i} \tag{29}$$

Reliability of $M_{N,4}$

Reliability of the module $M_{N,4}$ reads

$$p_s(t) = \Pi_{i=1}^{5} p_{4,i} \tag{30}$$

3.2.5 Network probability score

The module $M_{N,5}$ computes the network quality indicators in terms of the successfully ambiguities resolution, and the quality of the network corrections. This is the first state vector of the system under investigation.

- Let $x_{5,1}$ denote quality indicator for the corrections field.
- Let $x_{5,2}$ denote quality indicator for the uncertainty of corrections field.
- Let $x_{5,3}$ denote quality indicator for the ambiguities expressed by ADOP from Sect. 3.2.2.
- Let $x_{5,4}$ denote the number of common satellites used in the computation.
- Let $x_{5,5}$ denote the number of rejected satellites from computation.

Serial coupling is the appropriate choice for the module $M_{N,5}$ and the block diagram is represented by the Figure 11.

Block diagram of $M_{N,5}$

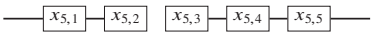


Figure 11: Block diagram of module $M_{N,5}$

Structure function of $M_{N,5}$

The structure function of $M_{N,5}$ reads

$$\Phi(\boldsymbol{x}) = \min(x_{5,1}, x_{5,2}, x_{5,3}, x_{5,4}, x_{5,5}) = \Pi_{i=1}^{5} x_{5,i} \tag{31}$$

Reliability of module $M_{N,5}$

The reliability of $M_{1,5}$ reads

$$p_s(t) = \Pi_{i=1}^{5} p_{5,i} \tag{32}$$

3.3 NRTK Baseline reliability analysis

This is the second level of the NRTK data processing. The corrections generated by the network are involved to generate the computation point (CP), then the unknown rover coordinates are determined relative to the computation point. This method is known as the relative positioning technique.

Similarity between the baseline and the network data processing exists as we see in the coming sub-sections.

3.3.1 Rover receiver data integrity

The module $M_{B,1}$ is similar to the module $M_{N,1}$ defined in Sect. 3.2.1. The quality of the raw observations collected by the rover receiver depends strongly on the statistics methods used to check for anomalies. All variables defined in the module $M_{N,1}$ are applicable to the module $M_{B,1}$.

3.3.2 Baseline ambiguity resolution

The module $M_{B,2}$ is similar to the module $M_{N,2}$ defined in Sect. 3.2.2. All variables defined for the module $M_{N,2}$ are applicable to the module $M_{B,2}$.

3.3.3 Baseline data integrity

The module $M_{B,3}$ is similar to the module $M_{N,4}$ defined in Sect. 3.2.4. All variables defined for the module $M_{N,4}$ are applicable to the module $M_{B,3}$.

3.3.4 Baseline probability score

The module $M_{B,4}$ is similar to the module $M_{N,5}$ defined in Sect. 3.2.5. All variables defined for the module $M_{N,5}$ are applicable to the module $M_{B,4}$.

3.4 NRTK rover reliability analysis

The investigation of the rover position error is the final check. The quality is measured in terms of standard deviations of the topocentric coordinates $(\delta e, \delta n, \delta u)$.

3.4.1 Rover prediction function

This module $M_{R,1}$ uses information from the double-difference variance-covariance matrix to compute the prediction of the position error. The number of satellites used in the computation is considered as a parameter.

- Let $x_{3,1}$ denote inspection of the main diagonal of variance-covariance matrix.
- Let $x_{3,2}$ denote the total variance monitoring algorithm.
- Let $x_{3,3}$ denote the generalized variance monitoring algorithm.

3.4.2 Rover position error

This module $M_{R,2}$ computes the standard deviations of the rover position error $(\delta e, \delta n, \delta u)$ and assign a final score to determine the state of the rover receiver accuracy.

4 Multi-state reliability analysis

Since the rover position accuracy cannot be represented by a binary system with two performance level states as functioning or failed, the multi-state system (MSS) approach is chosen to deal with situations where more than two levels of performance are considered. The material used to construct this section is from Natvig [17].

4.1 Definition of NRTK performance levels

Based on the values computed from the score modules $M_{N,5}$ (Sect. 3.2.5) and $M_{B,4}$ (Sect. 3.3.4), respectively, a single judging number is assigned to determine the performance level of the rover position accuracy.

The states represent level of performance ranging from the perfect functioning level *perfect* down to the complete failure level *catastrophic*. Five states are defined for a NRTK system, namely *start*, *perfect*, *acceptable*, *rejected* and *catastrophic* states.

The rover position accuracy is well described by the state diagrams given by figures 12 and 13, respectively.

1. State **4**: Perfect functioning level. No complications. The user requirements regarding the position accuracy are satisfied.
2. State **3**: Acceptable position accuracy. Minor complications.
3. State **2**: Rejectable position accuracy. Major complications due to the atmosphere, multipath or algorithms failure. The user requirements are not satisfied.
4. State **1**: Catastrophic state. The NRTK system is down, and not delivering the corrections to the user in the field.
5. State **0**: Start state. The process always starts at this state and can reach any other states.

Note that the states **0**,**4**,**3**,**2** are transient states while the state **1** is absorbing state.

Note that the probabilities $P_{i|j}$ must be computed from real data.

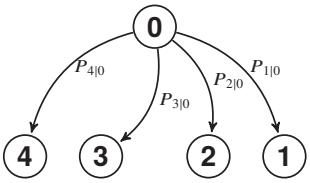


Figure 12: Start state of the rover position accuracy

4.2 Penalized honored stochastic averaged variance

Based on state vectors data, our aim is to construct the NRTK state diagram from the network and baseline. The variance covariance matrix (VCM) is considered as the state vector and the average variance with respect to the number of observed satellites (n_{sat}) is used to compute the quality indicator on an epoch by epoch basis. In addition, the number of rejected satellites (n_{rej}) and the geometry factor (DOP) are used to penalize/honor the average variance.

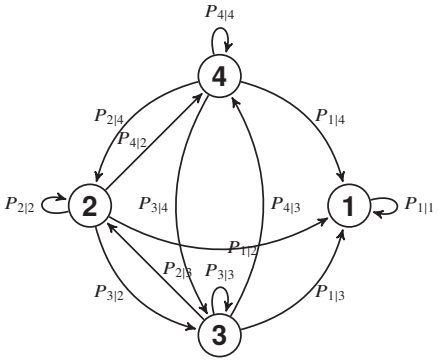


Figure 13: State diagram of the rover position accuracy

4.2.1 Penalized average variance component

The total average variance computed from the VCM shall be penalized in case of rejection of satellite(s) with bad data and causes the increase of DOP indicator. The penalized function shall look similar to:

1. Penalized least square (PLS) proposed by Green and Silverman [4, p. 5].

$$S(g) = \sum_{i=1}^{n} \left\{ y_i - g(t_i) \right\}^2 + \alpha \int_a^b \left\{ g''(x) \right\}^2 dx \qquad (33)$$

   where α is the smoothing parameter and defines the rate of change between the residuals and local variations. Anyway, minimizing $S(g)$ gives the best compromise between smoothness and goodness-of-fit.
   Large value of α will make the penalty term more in action while small value the first term will be the main contribution.
2. Information criteria type penalizing the model complexity. Denote by M the model to be investigated and dim(M) is the length of its parameter vector $\boldsymbol{\theta}$.
   The Akaike's information criterion (AIC) [2, Chap. 2].

$$\text{AIC(M)} = 2 \text{ log-likelihood}_{\max}(M) - 2 \text{ dim(M)} \qquad (34)$$

   The Bayesian information criterion (BIC) of Schwarz (1978) takes the form of a penalised log-likelihood function where the penalty is equal to the logarithm of the sample size times the number of estimated parameters in the model [2, Chap. 3].

$$\text{BIC(M)} = 2 \text{ log-likelihood}_{\max}(M) - (\log n) \text{dim(M)} \qquad (35)$$

4.2.2 Honored average variance component

Detection and rejection of satellite(s) with bad data is a good thing. The check algorithms shall be honored as long as the DOP values remain in the acceptable region.

The value of the horizontal dilution of precision (HDOP) is expected to be less or equal 2.0, that is HDOP $\leq$ 2.0.

Forming a new stochastic variable $T_j = \text{HDOP}_j - \mu_{\text{HDOP}}$, then we can monitor the values of T_j over time.

Note that μ_{HDOP} corresponds to the mean value of the HDOP in time span Δt.

4.2.3 Balanced average variance

Our aim is to put together the pieces defined in 4.2.1 and 4.2.2, respectively and find a way to balance between the satellite(s) rejection n_{rej} and the HDOP value.

The exponential reweighting algorithm type is an option. The algorithm places more importance to more recent data by discounting older data in an exponential manner.

Let $k_j = \dfrac{(n_{sat,j} - n_{rej,j}) - n_{\text{const}}}{n_{sys}}$, where $n_{\text{const}} \geq 8$ is the user defined parameter and corresponds to the minimum number of satellites required to compute a reliable solution and preserve a good HDOP value, and $n_{sys} = 36$ is the total satellites in GPS constellation.

A suitable stabilization factor $\eta \in [0,1]$ is chosen such that the penalized honored average variance (σ_{phav}) reads

$$\sigma_{\text{phasd}}(\eta,k,T) = \underbrace{\left\{ \frac{1}{n_{sat}} \sum_{i=1}^{n_{sat}} c_{i,i} \right\}^{\frac{1}{2}}}_{\text{first-term}} + \underbrace{\frac{1}{n_j} \sum_{j=1}^{n_j} \left\{ \eta k_j + (1-\eta) T_j \right\}}_{\text{second-term}} \qquad (36)$$

where n_{sat} is the number of satellites with valid data, n_{rej} is the number of rejected satellites by the algorithms, $n_j$is the window size which is user defined.

4.2.4 Penalized honored average variance validation

The parameter vector of the Equation (36) is $\boldsymbol{\theta} = (\eta, k, T)$. Our aim is to study the variation of the second term of Equation (36) and try to get some valid answers.

Full details of the penalized honored average standard deviation algorithm is given in Appendix A.2.

4.3 NRTK residuals contribution

NRTK residuals generated by the network and baseline data processing are considered as state vectors and will be to construct the state diagram.

The procedure is defined as follow:

– choose the time window $\Delta t = 10$ seconds.

– compute the standard deviation of the residuals σ_{res}. Figure 18 shows the concept.
– choose a suitable strategy to map the computed values of σ_{res}.

5 NRTK reliability results

The aim of this section is to present the results from the analysis. The level of performance ranging from the perfect functioning level *perfect* down to the complete failure level *catastrophic* shall be determined from the data.

5.1 Horizontal dilution of precision (HDOP)

The geometry of the visible satellites is considered as an important factor in achieving high quality results especially for point positioning and kinematic surveying. Anyway, the geometry changes with time due to the relative motion of the user and satellites. A measure of the instantaneous geometry is the dilution of precision (DOP) factor.

The DOP values are computed from the variance-covariance matrix in the ECEF coordinate system and converted to the topocentric local coordinate system with its axes along the local north, east, and up (i.e., vertical) by rotational matrix R by applying the law of covariance propagation.

The DOP value can be defined in various ways; PDOP value in the local system is identical to the value in the global system. In addition to the PDOP, two further DOP definitions are used. HDOP, the dilution of precision in the horizontal position, and VDOP, denoting the corresponding value for the vertical component. The interested reader is referred to (Hoffmann-Wellenhof et al, 2008) [7, pp. 262-270].

$$GDOP = \sqrt{\sigma_e^2 + \sigma_n^2 + \sigma_u^2 + \sigma_t^2} \tag{37}$$

$$PDOP = \sqrt{\sigma_e^2 + \sigma_n^2 + \sigma_u^2} \tag{38}$$

$$HDOP = \sqrt{\sigma_e^2 + \sigma_n^2} \tag{39}$$

$$VDOP = \sqrt{\sigma_u^2} \tag{40}$$

Acceptable horizontal DOP value is HDOP $\leq$ 2.0. Figure 14 shows the computed HDOP for the analyzed data set. In addition, Figure 15 shows the viewed satellites. Clearly, the number varies between 8 and 10 satellites.

5.2 Rover level of performance prediction

We will predict the rover level of performance ranging from the perfect functioning level *perfect* down to the complete failure level *catastrophic*. Three classification lines are chosen in order to separate the computed average standard deviations using Equation (36) into four decisions regions based on the values of $\sigma_{\text{pred}} \in \{0.2, .03, .4\}$. Figure 16 shows the concept.

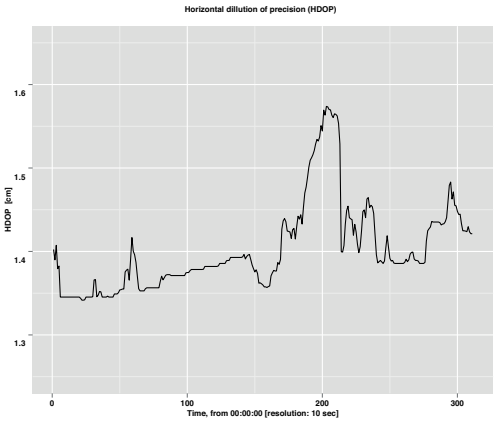


Figure 14: Computed HDOP values. Baseline of ~ 41 km, year:2014, DOY: 85.

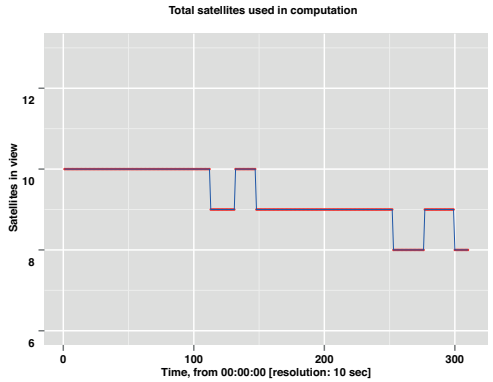


Figure 15: Number of satellites used in the computation. Baseline of ~ 41 km, year:2014, DOY: 85. Plotted as red dots connected by blue lines.

The computation of σ_{pred} using the Equation (36) proceeds as follow:

– Based on the sliding window size Δt, form a data matrix from the baseline residuals. Compute the variance-covariance matrix VCM_{res} and get and sort in ascending order the diagonal elements D.

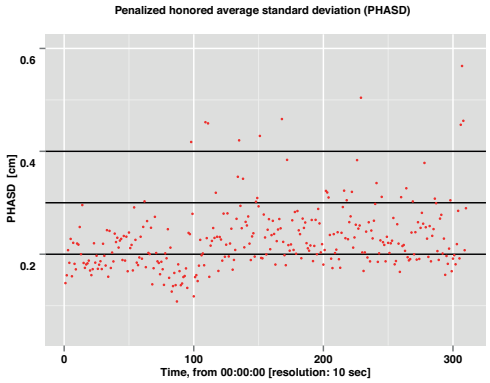


Figure 16: Computed penalized honored average standard deviation. Horizontal lines are used for classification. Baseline of ∼ 41 km, year:2014, DOY: 85.

– Compute the averaged standard deviation, the first term of Equation (36)

$$\sigma_{\text{avg}} = \left\{ \frac{1}{n_{sat}} \sum_{i=1}^{n_{sat}} c_{i,i} \right\}^{\frac{1}{2}} \tag{41}$$

– Generate a random number $n_{\text{rej}} \in \{0,1,2\}$, and compute K_j

$$K_j = \frac{(n_{\text{obs}} - n_{\text{rej}} - n_{\text{const}})}{n_{\text{tot}}} \tag{42}$$

– Compute the second term of Equation (36)

$$\sigma_{\text{pen}} = \frac{1}{n_j} \sum_{j=1}^{n_j} \left\{ \eta k_j + (1-\eta) T_j \right\} \tag{43}$$

where T_j is the HDOP computed from the solution variance-covariance matrix. The computed HDOP values are shown in Figure 14.

6 NRTK Reliability Validation

The aim of this section is to introduce the procedures used to validate the NRTK system reliability. The level of performance ranging from the perfect functioning level *perfect* down to the complete failure level *catastrophic* shall determined from data.

6.1 NRTK state diagram definition

As we mentioned in the introduction, the key of the NRTK method is the measurement of the distance-dependent errors.

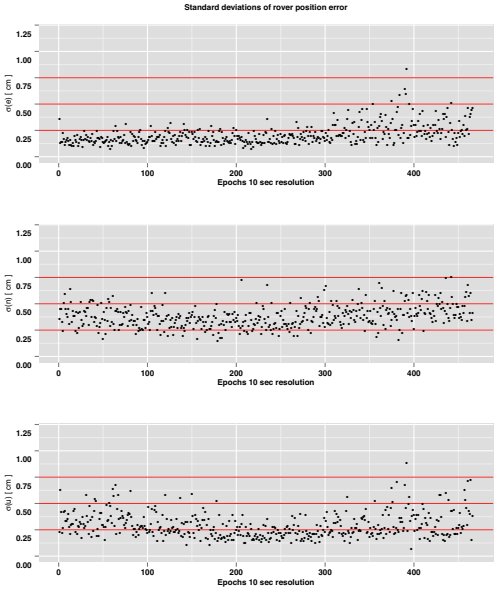


Figure 17: Standard deviations of the rover position errors $(\sigma_e, \sigma_n, \sigma_u)$ in the topocentric coordinate system. Horizontal lines are used for classification. Baseline of ∼ 1 km, year:2013, DOY: 152.

The variations in the ionospheric and tropospheric fields are assumed constant in a period of time $\Delta t < 10$ s.

Since we have access to the rover position error in the topocentric coordinates system $(\delta_e, \delta_n, \delta_u)$, the state diagram is computed as follow:

– Choose the time window $\Delta t = 10$ seconds.
– For each component, compute the standard deviation of the rover position error $\sigma_p = (\sigma_e, \sigma_n, \sigma_u)$. Figure 17 shows the concept.
– Based on the user requirements and the computed values of σ_p, map this value to the performance defined levels, namely the states $S = \{0,1,2,3,4\}$.
– Compute transitions probabilities $P_{i|j}$ from data. This task is accomplished by counting the frequencies and computing the associated probability, that is $p_s = \dfrac{N_s}{N_T}$ where N_s is the number of time we are visiting the state s and N_T is the total number of events.

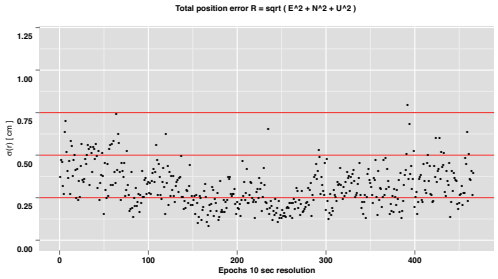


Figure 18: Standard deviation of the rover position error. Horizontal lines are threshold values. Baseline of ~ 1 km, year:2013, DOY: 152.

6.2 States transition probabilities

We have computed the transition states probability based on Figure 18, where the threshold T_h values are defined respectively $T_h \in \{0.25, .5, .75\}$ cm. Note that $\sum_{j=1}^{4} P_{i,j} = 1$, for $i = 1, \ldots, 4$.

$$\begin{array}{c} \\ \text{Next state} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Current State} & & & \\ 4 & 3 & 2 & 1 \end{array} \\ \begin{bmatrix} 0.596 & 0.378 & 0.024 & 0.001 \\ 0.234 & 0.668 & 0.098 & 0.001 \\ 0.077 & 0.69 & 0.231 & 0.001 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{l} 4 \\ 3 \\ 2 \\ 1 \end{array} = \boldsymbol{P}$$

The limiting distribution is obtained by matrix multiplication of the transition matrix:

$$\lim_{n\to\infty} \boldsymbol{P}_{ij}^{n} = \pi_j, \quad j \geq 0 \tag{44}$$

$$\boldsymbol{P}^{(50)} = \begin{pmatrix} 0.3312 & 0.5443 & 0.0798 \\ 0.3321 & 0.5459 & 0.0800 \\ 0.3315 & 0.5449 & 0.0799 \end{pmatrix}$$

We see that each row of $\boldsymbol{P}^{(50)}$ has almost identical entries, this confirms that Equation (44) converges to some values as $n \to \infty$. It seems that the existence of a limiting probability that the process p will be in state j after a large number of transitions n, and the value is independent of the initial state.

For this data set, the probability of being in state 4 is 33.12%, state 3 is 54.59% and state 2 is 8.0%.

6.3 State diagram

Information from the network (Sect. 3.2) and the baseline (Sect. 3.3) are combined in such way that a single judging number is mapped to the performance levels with states $S = \{0, 1, 2, 3, 4\}$. and the transitions probabilities $P_{i|j}$ are obtained. This task is accomplished by counting the frequencies and computing the associated probability, that is $p_s = \dfrac{N_s}{N_T}$ where N_s is the number of times we are visiting the state s and N_T is the total number of events.

In order to define a single judging number, various schemes are considered, for instance an averaging scheme or to assign different weights to each component.

7 Discussions and conclusion

An improvement of the rover position accuracy can be achieved by applying procedures for multi-state reliability analysis at the system and user level in NRTK. More concretely, the network corrections, baseline residuals and the associated variance-covariance matrices are considered as the system states and have a direct influence on the rover position accuracy.

The use of the multi-state reliability analysis will help us to get some concrete answers to the following problems.

- can we trust the corrections provided by NRTK to the user ?
- at which level ?
- what are the amelioration potential ?

The weaknesses and the strengths of the system have to be identified and the amelioration potential can be achieved by modifying the serial critical components coupling into paralleled one with a cost effectiveness.

The methods tested make it possible to identify the NRTK critical component with bad data so this can be eliminated or down weighted in the positioning process leading to an improvement in the rover position from epoch to epoch.

It is expected that the suggested approach will reduce the number of wrong or inaccurate rover positions encountered by NRTK users in field, which subsequently will lead to a more efficient work flow for NRTK users.

Discussions

The rover position accuracy is well described by the state diagram. Based on the values computed by the score modules from the NRTK and the baseline, a single judging number is assigned that determines the performance rate of the rover position accuracy $\rho = (e^2 + n^2 + u^2)^{1/2}$, measured in terms of standard deviations σ_ρ.

The computation of probability of the rover position accuracy in time span $\Delta t = 10$ seconds is carried out as follows:

1. Network RTK corrections score: Based on algorithms efficiency defined in Section 3.2, a probability $p_1 \in [0, 1]$ is assigned to the quality of the corrections.

2. Baseline score: Based on algorithms efficiency defined in Section 3.3, a probability $p_3 \in [0,1]$ is assigned to the quality of the baseline residuals.
3. Rover raw observations score: Based on algorithms efficiency used to edit the rover raw observations, a probability $p_2 \in [0,1]$ is assigned to the quality of the rover raw observations.

The validation process is carried out by computing the standard deviations of the rover position error $(\sigma_n, \sigma_e, \sigma_u)$ of topocentric coordinates. A single number, σ_ρ is assigned and the corresponding state is obtained. Equation (45) shows the mapping used in this investigation.

$$\begin{aligned} F\colon [0,1]\times[0,1]\times[0,1] &\longmapsto [0,1] \\ F(p_1,p_2,p_3) &\longmapsto p \in [0,1] \end{aligned} \tag{45}$$

A Appendices

A.1 Test data

Data used in this investigation is from the Norwegian RTK network known as CPOS operated by the Norwegian Mapping Authority (NMA). The test area is from the Rogaland region in the south west of Norway. Reference receivers are equipped with *Trimble Net9* receivers, tracking GPS and GLONASS satellite signals. Baselines vary between 35 to 112 km and the height difference between the sites is about 225 m. Tables 1, 2 and 3 give a full description of sub-network while Figure 19 shows the location of reference receivers. The data used for testing was collected on day of year (doy) 152 in 2013 and doy 85 in 2014 respectively.

Table 1: Sub-network reference receivers characteristics

Site	4-chars ID	Receiver type	Antenna type
Tonstad	TNSC	TRIMBLE NETR9	TRM55971.00
Sirevag	SIRC	TRIMBLE NETR9	TPSCR3_GGD
Stavanger	STAS	TRIMBLE NETR9	TRM55971.00
Akrahamn	AKRC	TRIMBLE NETR9	TPSCR3_GGD
Lysefjorden	LYSC	TRIMBLE NETR9	TRM55971.00
Prestaasen	PREC	TRIMBLE NETR9	TPSCR3_GGD

Table 2: Distances in sub-network [Km]

Sites	TNSC	SIRC	STAS	AKRC	LYSC	PREC
TNSC	X	56.32	75.00	109.60	44.61	95.23
SIRC	-	X	58.38	91.26	68.50	112.96
STAS	-	-	X	35.83	45.72	64.41
AKRC	-	-	-	X	73.51	65.60
LYSC	-	-	-	-	X	51.45
PREC	-	-	-	-	-	X

Table 3: Reference receiver coordinates - Euref89 XYZ

Sites	X	Y	Z
TNSC	3302221.359	388315.600	5424777.872
SIRC	3323397.670	336993.537	5415277.838
STAS	3275753.912	321110.865	5445041.883
AKRC	3254758.852	295601.453	5458918.670
LYSC	3269684.205	366420.447	5446037.395
PREC	3227088.927	353649.666	5471909.728

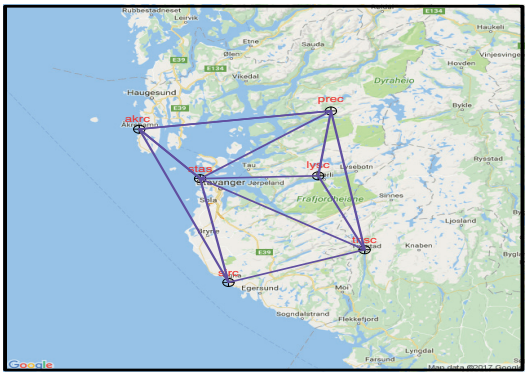


Figure 19: Test area used in this investigation, from Rogaland region. Composed of 6 reference receivers

A.2 Penalized honored average standard deviation

$$\sigma_{\text{phasd}}(\eta,k,T) = \underbrace{\left\{\frac{1}{n_{sat}}\sum_{i=1}^{n_{sat}} c_{i,i}\right\}^{\frac{1}{2}}}_{\text{first-term}} + \underbrace{\frac{1}{n_j}\sum_{j=1}^{n_j}\left\{\eta k_j + (1-\eta)T_j\right\}}_{\text{second-term}} \tag{46}$$

Algorithm recipes

The recipes of the algorithm given by the Equation (46) read:

1. Compute the effective local coverage (ELC) is defined as $K_j = \dfrac{n_{\text{obs}} - n_{\text{rej}} - n_{\text{const}}}{n_{\text{sys}}}$ where
   - n_{obs}: total number of satellites used in the estimation process.
   - n_{rej}: total number of satellites rejected by the algorithms.
   - n_{const}: user defined parameter. Default is set to 8.
   - n_{sys}: Total number of satellites in the GNSS constellation. For the GPS, the value is set to 36.
2. Position domain quality indicator T_j and is defined as the HDOP.
$$T_j = \left\{\sigma_n^2 + \sigma_e^2\right\}^{1/2} \tag{47}$$
3. Stabilization factor η combines and balances between T_j and K_j. A reasonable combination is to binomial/exponential trial:
$$\frac{1}{n_j}\sum_{j_1}^{n_j}(\eta k_j + (1-\eta)T_j) \tag{48}$$

where n_j is the window size and is user defined. Default is set to 10.

4. Operation level. The second term in Equation (46) shall operate on the same level as the first term and shall have the same unit. This is accomplished by adjustment of parameters η, K_j andT_j.

5. Stabilization factor η can be implemented by the Danish method.

A.3 Reliability Computation technique

In this section we present how the computation of the reliability of parallel structure is carried out step by step.

A.3.1 Block diagram

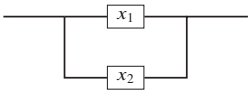


Figure 20: Block diagram of two parallel components

A.3.2 Structure function

The structure function $\Phi(\boldsymbol{x})$ of the Figure 20 reads

$$\Phi(\boldsymbol{x}) = \max(x_1, x_2) = x_1 + x_2 - x_1 x_2 \tag{49}$$

The corresponding reliability reads

$$R(\Phi(x)) = p_1 + p_2 - p_1 p_2 \tag{50}$$

In case of more than two components, the computation of the reliability function $R(\Phi(x))$ is straightforward. The first step is to divide the whole system components into two main components and applying the formula given by Equation (50). The next step is to substitute each individual reliability function with the corresponding terms. The last step is a simple calculations.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The international GNSS Service (IGS) is acknowledged for providing geodetic infrastructure and geodetic products used in this work. The authors would also like to thank Jon G. Gjevestad for proofreading of this paper.

References

1. Angrisano, A., Gioia, C., Gaglione, S., del Core, G.: GNSS Reliability Testing in Signal-Degraded Scenario. International Journal of Navigation and Observation **2013**, 1–12 (2013). DOI 10.1155/2013/870365. URL https://www.hindawi.com/journals/ijno/2013/870365/
2. Claeskens, G., Hjort, N.L.: Model Selection and Model Averaging. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press (2008). URL https://books.google.no/books?id=sjQFwB2XEFAC
3. Euler, H.J., Keenan, C.R., Zebhauser, B.E., Wübbena, G.: Study of a Simplified Approach in Utilizing Information from Permanent Reference Station Arrays. In: Proceedings of the National Technical Meeting of the Satellite Division of the Institute of Navigation, (ION GPS 2001), vol. 104, pp. 371–391 (2001)
4. Green, P.J., Silverman, B.W.: Nonparametric Regression and Generalized Linear Models: A roughness penalty approach. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis (1993). URL https://books.google.no/books?id=-AIVXozvpLUC
5. Gurtner, W., Estey, L.: RINEX: The Receiver Independent Exchange Format Version 2.11 (2007)
6. Heywood, H.B.: On Finite Sequences of Real Numbers. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character pp. 486–501 (1931)
7. Hofmann-Wellenhof, B., Lichtenegger, H., Wasle, E.: GNSS – Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and more. Springer Vienna (2008). URL https://books.google.no/books?id=Np7y43HU{_}m8C
8. Hossam-E-Haider, M., Tabassum, A., Shihab, R.H., Hasan, C.M.: Comparative analysis of GNSS reliability: GPS, GALILEO and combined GPS-GALILEO. In: 2013 International Conference on Electrical Information and Communication Technology (EICT), pp. 1–6. IEEE (2014). DOI 10.1109/EICT.2014.6777835. URL http://ieeexplore.ieee.org/document/6777835/
9. Hoyland, A., Rausand, M.: System reliability theory: models, statistical methods, and applications, *Wiley Series in Probability and Statistics - Applied Probability and Statistics Section*, vol. 396. Wiley (2004)
10. Jianan Xue, Kai Yang: Dynamic reliability analysis of coherent multistate systems. IEEE Transactions on Reliability **44**(4), 683–688 (1995). DOI 10.1109/24.476002. URL http://ieeexplore.ieee.org/document/476002/
11. Kaplan, E., Hegarty, C.: Understanding GPS: Principles and Applications, Second Edition. Artech House mobile communications series. Artech House (2005)
12. Kuusniemi, H., Wieser, A., Lachapelle, G., Takala, J.: User-level reliability monitoring in urban personal satellite-navigation. IEEE Transactions on Aerospace and Electronic Systems **43**(4), 1305–1318 (2007). DOI 10.1109/TAES.2007.4441741
13. Landau, H., Vollath, U., Chen, X.: Virtual Reference Station Systems. Journal of Global Positioning Systems **1**(2), 137–143 (2002). DOI 10.5081/jgps.1.2.137
14. Leick, A.: GPS Satellite Surveying. John Wiley and Sons, Incorporated (2015)
15. Li, X., Ge, M., Dai, X., Ren, X., Fritsche, M., Wickert, J., Schuh, H.: Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BeiDou, and Galileo. Journal of Geodesy **89**(6), 607–635 (2015). DOI 10.1007/s00190-015-0802-8. URL http://link.springer.com/10.1007/s00190-015-0802-8
16. Natvig, B.: Reliability analysis with technological applications. Department of Mathematics, University of Oslo (1998). Third edition
17. Natvig, B.: Multistate Systems Reliability Theory with Applications. Wiley Series in Probability and Statistics.

Wiley (2011). DOI 10.1002/9780470977088. URL https://books.google.no/books?id=ua62DAEACAAJ
18. Ohi, F.: Stochastic Evaluation of Multi-State Coherent Systems. Omohi College, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, 466-8555, Japan **1802**, pp. 200–206 (2012)
19. Ouassou, M., Jensen, A.B.O., Gjevestad, J.G.O., Kristiansen, O.: Next Generation Network Real-Time Kinematic Interpolation Segment to Improve the User Accuracy. International Journal of Navigation and Observation **2015**, 1–15 (2015). DOI 10.1155/2015/346498. URL http://www.hindawi.com/journals/ijno/2015/346498/
20. Raquet, J., Lachapelle, G.: Development and Testing of a Kinematic Carrier-Phase Ambiguity Resolution Method Using a Reference Receiver Network 1. Navigation **46**(4), 283–295 (1999). DOI 10.1002/j.2161-4296.1999.tb02415.x. URL http://doi.wiley.com/10.1002/j.2161-4296.1999.tb02415.x
21. Raquet, J.F.: UCGE Reports Number 20116 Development of a Method for Kinematic GPS Carrier-Phase Ambiguity Resolution Using Multiple Reference Receivers By (20116) (1998)
22. Ross S.M.: Introduction to probability models, 8 edn. Academic Press (2003). URL https://books.google.no/books?id=wlI9PgAACAAJ
23. Satyanarayana, S., Borio, D., Lachapelle, G.: C/N0 estimation: design criteria and reliability analysis under global navigation satellite system (GNSS) weak signal scenarios. IET Radar, Sonar & Navigation **6**(2), 81 (2012). DOI 10.1049/iet-rsn.2011.0164. URL http://digital-library.theiet.org/content/journals/10.1049/iet-rsn.2011.0164
24. Takac, F., Zelzer, O.: The Relationship Between Network RTK Solutions MAC, VRS, PRS, FKP and i-MAX. Proceedings of the 21st International Technical Meeting of the Satellite Division of the Institute Of Navigation (ION GNSS 2008) pp. 348–355 (2008)
25. Teunissen, P.J.G., De Jong, P.J., Tiberius, C.C.J.M.: Performance of the LAMBDA Method for Fast GPS Ambiguity Resolution. Navigation **44**(3), 373–383 (1997). DOI 10.1002/j.2161-4296.1997.tb02355.x. URL http://doi.wiley.com/10.1002/j.2161-4296.1997.tb02355.x
26. Teunissen, P.J.G., Odijk, D.: Ambiguity Dilution of Precision: Definition, Properties and Application. Delft Geodetic Computing Centre (LGR) Faculty of Geodesy Delft University of Technology The Netherlands (1997)
27. Wübbena, G., Bagge, A., Seeber, G., Boeder, V., Hankemeier, P.: Reducing distance dependent errors for real-time precise DGPS applications by establishing reference station networks. In: Proceedings of Ion Gps, vol. 9, pp. 1845–1852 (1996)
28. Xu, G.: GPS: Theory, Algorithms and Applications. Springer Berlin Heidelberg (2007). URL https://books.google.no/books?id=peYFZ69HqEsC
29. Yingkui, G., Jing, L.: Multi-State System Reliability: A New and Systematic Review. Procedia Engineering **29**, 531–536 (2012). DOI 10.1016/j.proeng.2011.12.756. URL http://linkinghub.elsevier.com/retrieve/pii/S1877705811065933

.

Errata

Table 1: Correcting formal errors in the PhD thesis

Page number	Paragraph	Changed from	Changed to
p.32	2.5.2, Eq.(2.41)	$\hat{\boldsymbol{z}} = \boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{z} = \boldsymbol{P}\boldsymbol{y}$	$\hat{\boldsymbol{z}} = \boldsymbol{H}(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{z} = \boldsymbol{P}\boldsymbol{z}$
p.33	2.5.3, Eq.(2.44)	$\frac{1}{2}[\boldsymbol{z}-\boldsymbol{H}\boldsymbol{W}\boldsymbol{x}]^T[\boldsymbol{z}-\boldsymbol{H}\boldsymbol{W}\boldsymbol{x}]$	$\frac{1}{2}[\boldsymbol{z}-\boldsymbol{H}\boldsymbol{x}]^T\boldsymbol{W}[\boldsymbol{z}-\boldsymbol{H}\boldsymbol{x}]$
p.33	2.5.3	$\hat{\boldsymbol{x}}_W = (\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{z}$	$\hat{\boldsymbol{x}}_W = (\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{W}\boldsymbol{z}$
p.33	2.5.3, Eq.(2.45)	$\frac{1}{2}[\boldsymbol{z}-\boldsymbol{H}\boldsymbol{G}\boldsymbol{x}]^T[\boldsymbol{z}-\boldsymbol{H}\boldsymbol{G}\boldsymbol{x}]$	$\frac{1}{2}[\boldsymbol{z}-\boldsymbol{H}\boldsymbol{x}]^T\boldsymbol{G}[\boldsymbol{z}-\boldsymbol{H}\boldsymbol{x}]$
p.38	2.7.1, Eq.(2.55)	$\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_{k-1})$	$\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)$
p.58	3.2.2	Adding	GPS new civil signals paragraph, p.59
p.60	3.3.1, Eqs.(3.5, 3.6)	α	β, p.61
p.60	3.3.1, Eq.(3.7)	Removing empty equation Eq. (3.7)	Equations number will be decreased by 1 from p. 61, Chap. 3.
p.71	3.4.1, Eq. (3.22)	Adding	ϵ_ϕ consists of noise and unmodeled effects, and mainly accounts for the remaining biases (e.g., carrier-phase offset, variation and wind-up, receiver noise, etc.), p.72 Eq.(3.22) becomes Eq.(3.21).
p.73	3.5, Eq.(3.24)	Adding	$\lambda^i N^i_{ab}$ to Eq.(3.23)
p.73	3.5, Eq.(3.26)	Adding	$\lambda^{ij} N^{ij}_{ab}$ to Eq.(3.25), p. 74
p.87	3.8, Eq.(3.58)	Removing	N_1 from Eq.(3.57), p. 88
p.87	3.8, Eq.(3.59)	Removing	N_2 from Eq.(3.58), p. 88
p.91	3.9	Remove Eqs. (3.73) and (3.74)	Refer to Eqs. (3.20) and (3.21) for consistency.
p.93	4.1	Adding explanation to Eq. (4.1)	where the site 1 and satellite 1 are chosen as reference site and satellite, respectively (A=1, i=1), p. 94
p.98	4.1.3, Eq.(4.16)	var($\Delta\nabla\Phi$)	$\sqrt{\text{var}(\Delta\nabla\Phi)}$, p. 99
p.159	Other contribution	Removed	Suggested by the opponents.

ISBN: 978-82-575-1488-4
ISSN: 1894-6402

Kartverket

Kartverksveien 21
3511 Hønefoss
+47 32 11 81 00
www.kartverket.no

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

97312 / ANDVORDGRAFISK.NO