

## **Report**

### ***Research and Development Project for Geodetic Deformation Monitoring***

Contribution to the research project:

“Long-term study on anomalous time-dependent subsidence in the Wadden Sea Region”

Sami Samiei-Esfahany  
Department of Geoscience and Remote Sensing  
Delft University of Technology, Netherlands

Hermann Bähr  
Nederlandse Aardolie Maatschappij B.V.

<i>EP Document Nummer.:</i>	<b>EP201505216980</b>		
<i>Revision No.:</i>	<b>2.0</b>		
<i>Document Date:</i>	<b>26-06-2015</b>	<i>Security:</i>	<b>Restricted</b>

The copyright of this document is vested in Nederlandse Aardolie Maatschappij B.V., Assen, The Netherlands. All rights reserved. Neither the whole, nor any part of this document may be reproduced, stored in any retrieval system or transmitted in any form or by any means (electronic, mechanical, reprographic, recording or otherwise) without the prior written consent of the copyright owner.



## **Executive Summary**

This study proposes considerable improvements to the current NAM geodetic processing workflow with demonstrated benefit for geomechanical modelling. Their consideration will add value to the development of a testing framework for candidate hypotheses in geomechanical modelling. The proposed stochastic models also provide opportunities for a substantiated optimisation of the current survey design.

The NAM geodetic processing workflow has been examined in the context of geomechanical modelling with focus on observations from levelling and InSAR data. Together with GPS, these techniques play a complementary role in subsidence modelling due to their specific properties.

In this report, we investigated potential improvements for preparing subsidence measurements from these techniques for geomechanical model calibration. As an essential prerequisite for testing candidate hypotheses on geomechanical models, appropriate stochastic models for geodetic datasets have been proposed. It was shown that taking into account correlation structures significantly improves the precision of geomechanical model predictions. Furthermore, several effective measures to optimise the model calibration workflow have been identified, aiming at minimising uncertainties and biases due to simplifications and not validated assumptions.

### **Stochastic modelling (Chapter 2):**

The stochastic model for geodetic data that is currently deployed for geomechanical calibration at NAM is simplified and incomplete, because it neglects correlations and only accounts for the uncertainty of the measurement itself. However, when identifying subsidence due to gas extraction as signal of interest, any deformation caused by shallow sources should be considered noise and thus included into the uncertainty model. Complementing the technique-related measurement noise by this so-called idealisation noise is the major innovation of this study. Idealisation noise is by far the most dominant noise component and thus undeniably significant. However, its quantification in parametric models is much less reliable than what is known on measurement noise.

In the case of both levelling and InSAR, we proposed specific state-of-the-art models to describe the measurement noise. Furthermore, we proposed a Monte Carlo approach to account for the processing induced correlation resulting from filtering out the atmospheric component in InSAR data.

An idealisation noise model has been derived from levelling surveys outside the influence area of gas production where shallow deformation effects can be isolated from deep source subsidence. The spatio-temporal stochastic analysis of the selected area shows the presence of both temporal non-stationary and spatially correlated idealisation noise components in the levelling dataset. The obtained model parameters for idealisation noise are deemed representative for the Wadden area but not for upcountry regions with different soil properties. We proposed to also use this model to approximate idealisation noise in InSAR data.

### **Optimal output level (Chapter 3):**

Based on simulations with a simple geomechanical model, the optimal interface between geodetic data processing and geomechanical modelling workflow has been

investigated. Recommendations aim at maximising modelling efficiency while minimising not fully validated assumptions and computational complexity. The following conclusions were drawn from the output level study. Biases in geomechanical calibration can be mitigated by selecting the output level of spatially and temporally differenced instead of pseudo-absolute subsidence measurements. A demonstrated bias can be easily avoided by calibrating model predictions against InSAR observations in the original line-of-sight (LOS) geometry instead of relying on the incorrect assumption of purely vertical ground deformation.

We showed that the currently implemented simplified processing compromises the precision of the geomechanical model parameters. Thus, the uncertainty of model predictions can be significantly reduced by taking covariances into account for geodetic data. A similar effect has the use of multiple reference points and multiple reference epochs for the individual double difference observations. This measure has also the potential to mitigate biases due to the choice of a single reference point.

Finally, the simulation study showed that in the case of using the full covariance matrix in the inversion, atmospheric filtering does not have a significant effect on the modelling quality. In that sense, not removing the atmospheric signal from InSAR data is worth consideration for the geomechanical modelling as it may have some considerable advantages for bias mitigation.

### **Outlier handling (Chapter 4):**

Considering outlier handling in the geodetic processing workflow is most relevant for levelling data, in which unavoidable human errors regularly cause huge discrepancies. Rigorous outlier handling, however, requires reliable knowledge on geomechanical model uncertainties. Since this is not available, we proposed to focus on the very obvious outliers that can be identified without that knowledge, using a pragmatic approach. For InSAR, outliers are generally a minor issue due to the high spatio-temporal sampling in combination with the smooth signal of interest. We propose to address a subclass of InSAR outliers that may become critical for geomechanical modelling. For both levelling and InSAR, a sensitivity analysis in operational modelling can help quantifying the actually impact of outliers.

### **InSAR Data volume (Chapter 5):**

For InSAR, we proposed an approach for data reduction in order to cope with the large data volume. We proposed simple binned averaging for data reduction in the time domain and hierarchical k-means clustering for data reduction in the space domain. Both approaches can be formulated as linear transformation. This is beneficial as it allows for rigorous linear error propagation. Replacing the currently deployed ad-hoc approach of resampling InSAR observations to levelling benchmark locations, the new approach better exploits the full potential of the technique.

### **Data integration (Chapter 6):**

To integrate datasets from different measurement techniques into the workflow of geomechanical model calibration, we proposed not to combine or collocate them prior to modelling. To avoid interpolation artifacts, techniques should be introduced separately into the modelling. The conceptual mathematical framework for joint geomechanical inversion of all the geodetic techniques was given.

**Basic recommendations:**

1. Use the proposed stochastic model of geodetic data in the form of full noise covariance matrices in geomechanical modelling.
2. Use double differences with multiple reference epochs and points as an optimal interface between geodetic data and geomechanical modelling.
3. In case of InSAR, use line of sight (LOS) data directly in the geomechanical modelling and do not convert the data to vertical. Instead, geomechanical models should be evaluated in LOS direction.
4. Use the proposed approach for outlier detection and removing only the most obvious levelling outliers
5. Reduce the InSAR data volume by the proposed spatio-temporal averaging methods.
6. Consider not removing the atmospheric signal from InSAR data as the preferred option for geomechanical modelling.



## Table of Contents

<b>Executive Summary.....</b>	<b>3</b>
<b>1. Introduction.....</b>	<b>9</b>
<b>2. Stochastic Modelling .....</b>	<b>11</b>
2.1. Levelling measurement noise .....	12
2.2. InSAR measurement noise .....	14
2.2.1. InSAR decorrelation/scattering noise $Q_n$ .....	14
2.2.2. InSAR atmospheric noise $Q_w$ .....	15
2.3. Idealisation noise .....	19
2.3.1. Discussion on the idealisation noise model .....	26
2.4. Propagation of noise models through InSAR processing.....	28
2.5. Summary and conclusions .....	30
<b>3. Output level study.....</b>	<b>33</b>
3.1. Output level study: generic methodology .....	34
3.2. Output level Study: Levelling .....	36
3.2.1. Output level overview .....	36
3.2.2. Simulation setting .....	38
3.2.3. Results .....	39
3.2.4. Levelling output level study: discussion and conclusion.....	43
3.3. Output level Study: InSAR .....	46
3.3.1. Output level overview .....	46
3.3.2. Simulation setting .....	47
3.3.3. Results .....	48
3.3.4. InSAR output level study: discussion and conclusion .....	54
<b>4. Outlier Handling.....</b>	<b>57</b>
4.1. Introduction .....	57
4.2. Outlier handling: levelling data.....	58
4.2.1 Hypothesis testing for gross error detection in double difference time series .	59
4.3. Outlier handling: InSAR data.....	62
4.4. Outlier handling: summary.....	63
<b>5. InSAR data reduction .....</b>	<b>65</b>
5.1. Introduction .....	65
5.1. InSAR data reduction in time .....	66
5.2. InSAR data reduction in space.....	66
5.2.1 Quadtree decomposition .....	67
5.2.2 Hierarchical K-means clustering.....	68
5.3. Summary/Conclusion .....	69
<b>6. Data integration.....</b>	<b>71</b>
<b>A.1. Appendix 1- Demo: Simulation of deformation for InSAR observations.....</b>	<b>75</b>
<b>A.2. Appendix 2- Aspects of the current NAM geodetic data processing workflow</b>	<b>81</b>
<b>Bibliography .....</b>	<b>83</b>



## 1. Introduction

This report provides the results of the geodetic subproject of the long-term study on anomalous time-dependent subsidence in the Wadden Sea region. Geodetic measurements of land subsidence are one of the key components of the subsidence modelling/prediction workflow. This study investigates how to correctly and optimally exploit geodetic data in this process, aiming for improvements to the current implementation at NAM (see Appendix 2). The present report presents results on different aspects of geodetic data processing/exploitation as follows.

- **Stochastic modelling (Chapter 2):** One of the main research questions is how to describe the stochasticity of geodetic data, or in other words what the stochastic model (or noise model) of geodetic data is. Such a stochastic model is an indispensable requirement for hypothesis testing as well as the research presented in later chapters. Although we can find different stochastic models in literature and technical studies, we cannot find substantiated support to prefer one model to another in the project context. Furthermore, for modern geodetic techniques like time series InSAR (Interferometric Synthetic Aperture Radar), the available knowledge on potential noise components and their variability in time and space is limited.

We should note that the major challenge is not so much quantifying the uncertainty of the measurement itself but rather parameterising the so-called idealisation noise. This component of the stochastic model accounts for any real deformation that is not associated with the deep source deformation induced by gas production. Considering the latter the signal of interest, any deformation due to shallow sources should be included in the stochastic model.

The goal in this report is to review the stochastic properties of levelling and InSAR subsidence measurements, to get more insight in possible spatio-temporal variability of different noise components, and finally to propose a clear and coherent step-wise procedure to construct the required stochastic model.

- **Optimal output level (Chapter 3):** Geo-mechanical modelling/calibration can take place using different levels of geodetic processing, ranging from raw observations to elaborately tailored products. The more processing steps are involved, the more a priori knowledge is needed and the more complicated becomes the stochastic model. On the other hand, geomechanical model computations become more complicated if the model is matched with geodetic data at a lower processing level. The main question here is what the best output level of geodetic data is to be used in the modelling considering

transparent use of a priori knowledge, simpler structure of the stochastic model and modelling efficiency/optimality.

- **Outlier handling (Chapter 4):** Similar to any kind of measurements, geodetic datasets may include outlier observations mainly due to human or processing errors. As outliers are described neither by the stochastic model nor by the geo-mechanical model, they may introduce a bias in the final subsidence model. The objective here is to propose an efficient approach to detect the potential outliers in geodetic datasets.
- **Data volume (Chapter 5):** Handling large data-volumes is an important aspect especially for InSAR data. Output of the InSAR processing is a 3D spatio-temporal dataset, i.e., deformation time series for a large number of radar targets. Contrary to levelling datasets, which typically contain hundreds of benchmarks and a limited number of epochs, InSAR datasets often comprise tens of thousands of points and many epochs. Associated covariance matrices result in computationally expensive geomechanical modelling or are even too large to work with. The main objective here is to propose a data reduction approach for InSAR data in both time and space with clear/plausible assumptions.
- **Data integration (Chapter 6):** Geodetic subsidence datasets include data from three different geodetic techniques, i.e. levelling, InSAR, and GPS. How to properly integrate all different datasets together with their stochastic models into the subsidence modelling process is addressed in this part.

## 2. Stochastic Modelling

The main objective of the stochastic modelling here is to describe the precision, i.e. “noise” variability/dispersion, of geodetic data (Levelling and InSAR) using a clear mathematical formulation. The main focus is on the first and the second central moment of noise components. Note that InSAR and levelling datasets are spatio-temporal, i.e., they include deformation time series for a large number of levelling benchmarks or InSAR persistent scatterers, respectively. The final goal is to propose an analytical formulation and a step-wise algorithm to calculate the variances of all spatio-temporal deformation measurements and covariances among them.

As the term “noise” may be interpreted loosely, in order to avoid misinterpretation, we will give a brief introduction to what we mean with this term in the context of this project.

Here, the term “noise” subsumes all signal components in geodetic observations that are not related to the signal of interest. In our case, the signal of interest is the deformation induced by deep mechanisms such as reservoir compaction or aquifer depletion. Based on this definition, we encounter two different kinds of noise components. One is the random error of the measurements themselves, called “measurement noise”, and another one covers any other kind of deformation signal rather than the signal of interest. We call the latter “idealisation noise”. This term is introduced based on the concept of “idealisation precision”. In classical surveying, the idealisation precision describes how well the location of a topographic feature can be realized in the terrain. For example, the corner of a building can be identified more sharply than the middle of a river and so it has a higher idealisation precision. The same concept was introduced in the context of subsidence monitoring by (Ketelaar 2009) when the signal of interest is contaminated by other kinds of deformation, so it cannot be identified precisely. With this definition, the dispersion (variance and covariances) of idealisation noise components is the measure of idealisation precision, and the dispersion of measurement noise is the measure of measurements precision. To summarize these definitions in a mathematical form, assume  $y$  is the vector of geodetic observations. In generic form, the vector  $y$  can be written as the summation of three contributions:

$$y = S + d + e$$

where  $S$  is the true and unknown value of the signal of interest,  $d$  is any other kind of deformation rather than the signal of interest (idealisation noise), and  $e$  is the measurement noise. The dispersion of the noise components is given by their covariance matrices as:

$$D\{e\} = Q_e, D\{d\} = Q_d$$

where  $D\{\cdot\}$  is the dispersion operator. The goal of this chapter is to introduce a step-wise algorithm to calculate  $Q_e$  and  $Q_d$  for leveling and InSAR datasets.

**Chapter outline:** This chapter consist of four parts. The sections 2.1 and 2.2 will address the stochastic model for levelling and InSAR measurement noise, respectively. The idealisation noise is explained in section 2.3. Finally, the propagation of error sources through InSAR processing is addressed in section 2.4.

## 2.1. Levelling measurement noise

Levelling measurement noise is a noise component with a very well known model in geodesy. The typical model of measurement noise assumes that all the measured height differences in each levelling campaign are uncorrelated. Furthermore the variance of a measured height difference between point  $i$  and  $j$  is linearly dependent on the length of the trajectory  $l_{ij}$ :

$$\sigma_{h_{ij}}^2 = \sigma_m^2 l_{ij}$$

where  $\sigma_m$  (in mm/ $\sqrt{\text{km}}$ ) is a scaling factor that can be specified per epoch. From twelve major levelling campaigns in the northern part of the Netherlands, using variance component estimation, we derive an average value of 0.86 mm/ $\sqrt{\text{km}}$  for  $\sigma_m$ . The final covariance matrix of the vector  $y_k$ , which subsumes all measured height differences at epoch  $t_k$ , can be constructed as:

$$Q_{y_k} = \sigma_m^2 \begin{bmatrix} l_{01} & & 0 \\ & l_{12} & \\ 0 & & \ddots \\ & & l_{ij} \end{bmatrix}$$

This covariance matrix is propagated to the estimated heights (i. e., height differences with respect to a single reference point) per epoch as:

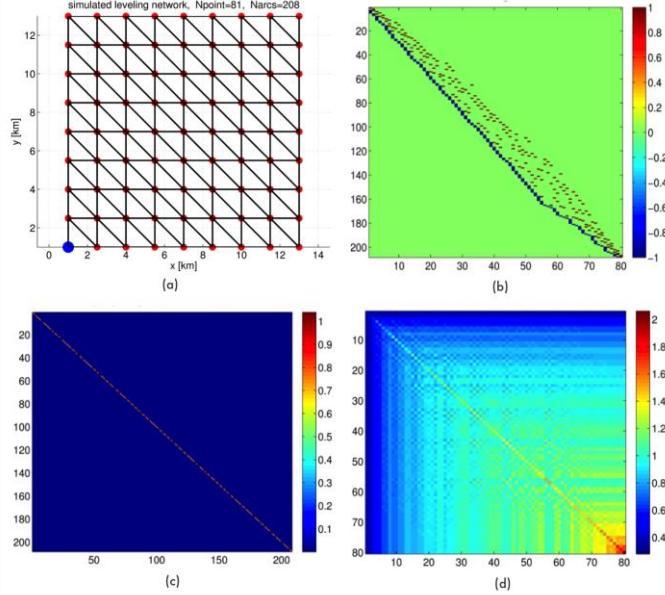
$$Q_{H_k} = (A_k Q_{y_k}^{-1} A_k)^{-1}$$

where  $A_k$  is the levelling network design matrix for the respective epoch. In this way, we construct  $Q_H$  for all the epochs. Note that  $Q_{H_k}$  is the covariance matrix of single differences (height differences of epoch  $t_k$ ) and should be further propagated to double differences (differences between two epochs) using the linear error propagation law. For example for epochs  $t_1$ ,  $t_2$ , and  $t_3$ :

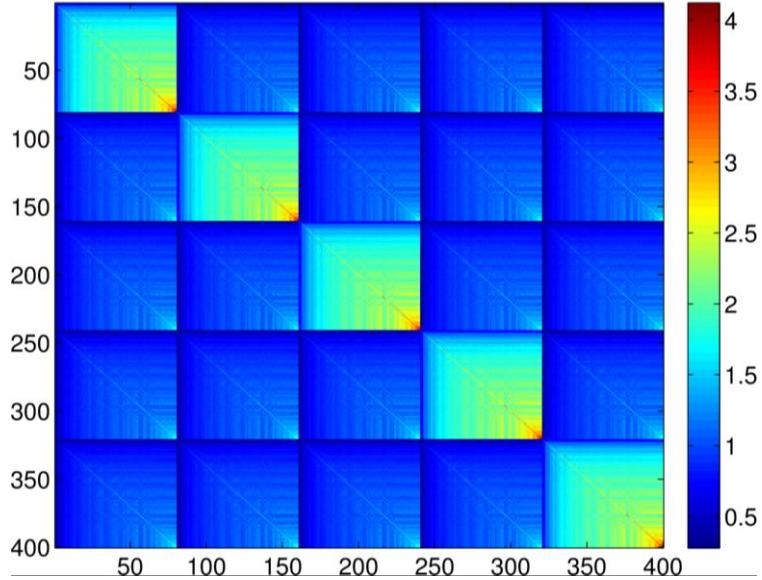
$$\begin{bmatrix} Q_{m_{12}} & Q_{m_{12},m_{13}} \\ Q_{m_{13},m_{12}} & Q_{m_{13}} \end{bmatrix} = \begin{bmatrix} -I & I & 0 \\ -I & 0 & I \end{bmatrix} \begin{bmatrix} Q_{H_1} & & \\ & Q_{H_2} & \\ & & Q_{H_3} \end{bmatrix} \begin{bmatrix} -I & -I \\ I & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} Q_{H_1} + Q_{H_2} & Q_{H_1} \\ Q_{H_1} & Q_{H_1} + Q_{H_3} \end{bmatrix}$$

where  $I$  is the identity matrix,  $Q_{m_{12}}$  is the covariance matrix of double differences between epochs 1 and 2, and  $Q_{m_{12},m_{13}}$  is contains the cross-covariances between the double differences of two pairings of epochs. Using this error propagation, the final covariance matrix of levelling measurement noise ( $Q_M$ ) is constructed. Figure 2.1 demonstrates the measurement noise covariance matrix construction. Figure 2.1a is a stylized levelling network for two epochs  $t_1$  and  $t_2$  with the same design matrices  $A_1$  and  $A_2$  visualized in Figure 2.1b. Then the Figures 2.1c and 2.1d are the  $Q_{y_1} = Q_{y_2}$  and  $Q_{H_1} = Q_{H_2}$  respectively. Note that the observations are ordered based on their distance to the reference point. This is the main reason why the variances in

Figure 2.1d are increasing. Assuming six epochs with the same design matrix, all the  $Q_{H_k}$  matrices are propagated to  $Q_M$ , which has been visualized in the Figure 2.2.



*Figure 2.1: Demonstration of measurement noise covariance matrix construction for a stylized levelling network: a) levelling network configuration for two epochs  $t_1$  and  $t_2$ , red points are benchmark locations, black arcs are levelling paths between benchmarks, and the blue point is the network reference benchmark, b) design matrix  $A_1 = A_2$ , c) diagonal covariance matrix of the observed height differences  $Q_{y_1} = Q_{y_2}$ , d) covariance matrix of estimated heights with respect to the reference point  $Q_{H_1} = Q_{H_2}$ . Benchmarks are ordered based on the distance to the reference point. Colorbars are in  $\text{mm}^2$ .*



*Figure 2.2: The final measurement noise covariance matrix  $Q_M$  for the stylized network of Figure 2.1 and six levelling epochs (five double difference epochs with respect to the first epoch). Colorbars is in  $\text{mm}^2$ .*

## 2.2. InSAR measurement noise

Error sources associated with InSAR measurements can generally be categorized into four components (Hanssen. 2001) as following:

1. Decorrelation effect induced by different error sources such as thermal noise, scattering mechanism, resampling and coregistration errors,
2. Atmospheric effects including turbulent mixing and vertical stratification,
3. Orbital errors and
4. Unwrapping errors.

The main InSAR measurement noise components considered in this study are decorrelation noise and atmospheric effects due to turbulent mixing. Atmospheric errors due to vertical stratification affects only mountainous terrains, is strongly correlated with topography and insignificant in flat areas such as the northern Netherlands. Hence, this effect is not considered in our noise model. Orbital effects are assumed deterministic; they are usually estimated and removed during InSAR processing. Unwrapping errors are considered as outliers in InSAR results and will be addressed later under the subject InSAR outlier handling.

Assuming that  $Q_n$  and  $Q_w$ , are covariance matrices of decorrelation noise and atmospheric effects respectively, we need the final stochastic model in the form of  $Q_{M_{InSAR}} = Q_n + Q_w$  for InSAR time series. Note that this noise model describes error sources of unwrapped time series without any smoothing/filtering applied on InSAR data. In the case of filtering (e.g. atmospheric filtering), this noise model should be propagated through filtering steps in order to construct the stochastic model for final filtered data. In this section we address only how to construct  $Q_n$  and  $Q_w$ . Later, in section 2.4, it will be explained how this model should be propagated further through InSAR filtering steps.

### 2.2.1. InSAR decorrelation/scattering noise $Q_n$

The significance of scattering and thermal noise for persistent scatterers is very well studied in the InSAR community. In addition to some theoretical models given in literature, some experimental validation also estimates the uncertainty component due to scattering noise (Marinkovic 2008, Ferretti. 2007). Based on these studies, the variance of PS scattering noise for double-difference measurements ranges from  $\sim 2.5 \text{ mm}^2$  for perfect point scatterers (such as big corner reflectors) to  $\sim 25 \text{ mm}^2$  for ordinary points (points with a larger variance are generally disregarded in the InSAR processing). Although these numbers can be used as an approximate value, in practice we are dealing with large numbers of PS with different signal-to-noise ratio and scattering behaviour. Here, we propose to use a noise model that is based on the amplitude dispersion index (Ferretti, Prati and Rocca 2001). For InSAR stacks with a sufficient number of images (i.e.,  $>25$ ; this criterion is usually fulfilled for InSAR datasets in northern Netherlands), this index is a good approximation of the phase standard deviation.

The relation of the normalized amplitude dispersion  $D_a$  and phase standard deviation  $\sigma_\varphi$  (of one PS in one image) is defined as:

$$D_a = \frac{\sigma_a}{\mu_a} \approx \sigma_\varphi,$$

where  $\sigma_a$  is the temporal standard deviation of the radar amplitude, and  $\mu_a$  is the temporal mean of the amplitude for a certain PS. This relation is demonstrated in Figure 2.3 using a simulation for a stack of 70 images. We can see that small phase standard deviations (as expected for PS) can be approximated well by the normalized

amplitude dispersion index. The phase standard deviation is in radian units and is converted to meters as:

$$\sigma_n = \frac{\lambda}{4\pi} \sigma_\varphi ,$$

where  $\sigma_n$  is the scattering noise standard deviation in meters and  $\lambda$  is the radar wavelength (typical: 0.056 m).

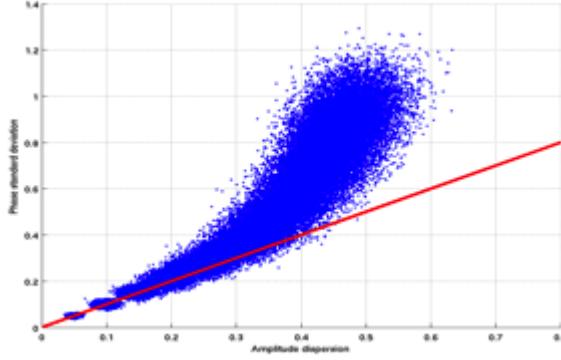


Figure 2.3: Simulated scatter plot of the relation between amplitude dispersion and phase standard deviation for a stack of 70 images.

The amplitude dispersion index is a standard product of InSAR processing software and can be used as an approximation of the phase standard deviation of one PS in one SAR image. This standard deviation should be propagated to double-difference InSAR measurements using the linear error propagation law. For example, for PS points  $i, j, r, s$  and acquisition times  $t_0, t_1, t_2, t_3$ , double difference variances and covariances are computed as:

$$\begin{aligned} C\{N_{ir}^{t_1 t_0}, N_{ir}^{t_1 t_0}\} &= 4\sigma_n^2 \\ C\{N_{ir}^{t_1 t_0}, N_{ir}^{t_2 t_0}\} &= 2\sigma_n^2 \\ C\{N_{ir}^{t_1 t_0}, N_{jr}^{t_1 t_0}\} &= 2\sigma_n^2 \\ C\{N_{ir}^{t_1 t_0}, N_{jr}^{t_2 t_0}\} &= \sigma_n^2 \\ C\{N_{ir}^{t_1 t_0}, N_{js}^{t_3 t_2}\} &= 0 \end{aligned}$$

where  $C\{\cdot\}$  is the covariance operator, and  $N_{ir}^{t_1 t_0}$  is the scattering noise component of the double difference measurement between points  $i$  and  $r$ , and acquisitions  $t_1$  and  $t_0$ . Using these equations all the elements of the atmospheric covariance matrix  $Q_n$  will be constructed.

### 2.2.2. InSAR atmospheric noise $Q_w$

One of the dominant error sources in InSAR data results from the heterogeneity of the wet component of atmospheric refractivity, causing varying time delays in satellite radar observations. It has been shown that the variability of atmospheric signal between two points increases with distance between the points, and almost linearly with their height difference (Emardson 2003, R. Hanssen 1998, Onn 2006). In this study, we ignore the height dependent variation of the atmospheric effect, because the topographic height variations close to the Wadden Sea are insignificant.

It has been shown by (R. Hanssen 2001) that the atmospheric turbulence effect in InSAR is temporally uncorrelated but spatially correlated and can be described by a power-law scaling behaviour in space following the turbulence theory proposed by (Kolmogorov 1941). Such a power-low behaviour can be effectively described with a covariance model from Matern-family as (Tatarskii 1971, Handcock 1994, Stein 2005, Grebenitcharsky 2005):

$$C(l) = \frac{\sigma_w^2}{2^{(\tau-1)} \Gamma(\tau)} \left( \frac{\sqrt{2\tau} |l|}{L_w} \right)^\tau K_\tau \left( \frac{\sqrt{2\tau} |l|}{L_w} \right),$$

where  $l$  is the spatial distance,  $\sigma_w^2$  is the atmospheric variance factor,  $L_w$  is the spatial range beyond that the turbulence effect becomes practically uncorrelated, and  $\tau$  is the model parameter which controls the spatial smoothness of the turbulence signal.  $\Gamma$  and  $K$  are the Gamma function and the modified Bessel function of the second kind respectively. This covariance model has three unknown values ( $\sigma_w^2$ ,  $L_w$ , and  $\tau$ ), which in principle can be estimated from filtered atmospheric data of PS-InSAR. We should note that this Matern covariance model coincides with well known covariance models for specific choices of the  $\tau$  parameter. For example if  $\tau \rightarrow \infty$ , the Matern model converges to a Gaussian model as

$$C(l) = \sigma_w^2 e^{-0.5 \left( \frac{l}{L_w} \right)^2},$$

or if  $\tau = 1/2$  the Matern model is identical with an exponential model:

$$C(l) = \sigma_w^2 e^{-\left( \frac{l}{L_w} \right)}.$$

Based on the experimental studies in (Liu 2012), the parameter  $\tau$  mainly varies between 2/3 and 5/3 in the Netherlands. So the Gaussian model (with  $\tau = \infty$ ) is far off from reality. However, the exponential covariance model (with  $\tau = 1/2$ ) can be used as an adequate approximation as will be shown. The advantage of the exponential model is that it has one unknown less to be estimated, and at the same time it can reasonably describe the spatial variability of the atmosphere. In order to validate this argument, we simulate five interferometric atmospheric signals with different degrees of variability but all based on the experimental power-law model presented by (R. Hanssen 2001). Figure 2.4 shows these simulations together with their empirical variogram. The modelled exponential variograms are also plotted. We can clearly see that an exponential model can effectively approximate the degree of the spatial variability in the atmospheric signal. So in this study we propose to use the exponential covariance/variogram for InSAR atmospheric noise.

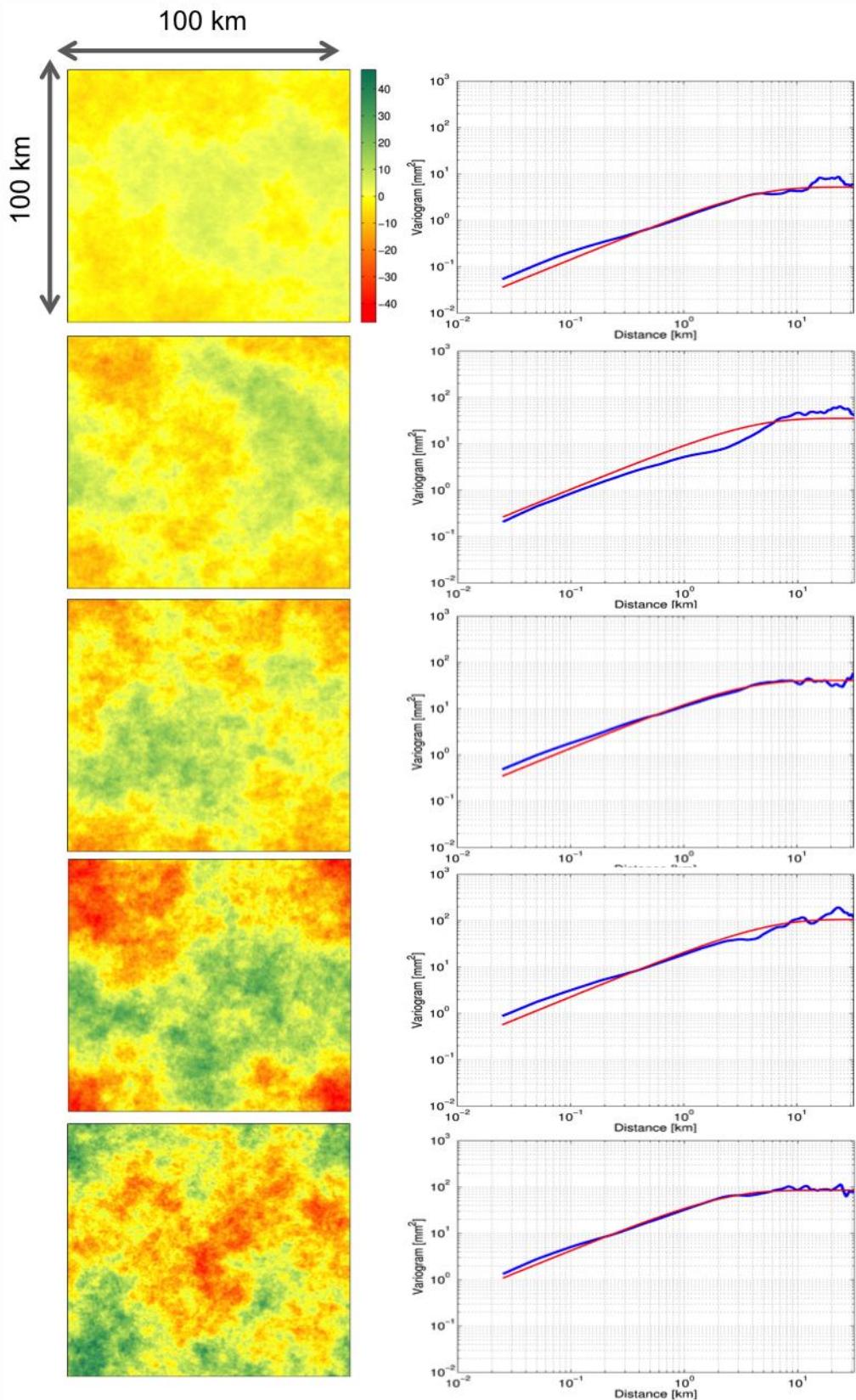
The variation in covariance model parameters (i.e.  $\sigma_w^2$  and  $L_w$ ) depends on the random atmospheric conditions at the time of the radar acquisition for each scene. In practice, when an estimate for the InSAR atmospheric component is available, we recommend to estimate these parameters directly from the data. For simulation and feasibility studies a standard deviation of 4~5 mm and a correlation range of ~10 km are typical numbers to use. Note that the standard deviation  $\sigma_w$  relates to one radar scene. For double difference interferometric observations, this number should be further propagated as follows. If  $W_{ir}^{t_0 t_1}$  is the relative atmospheric noise component of point  $i$  with respect to point  $r$  between epoch  $t_0$  and the epoch  $t_1$ , then based on the exponential variogram model, the variances and covariances between different double difference atmospheric components can be computed as:

$$\begin{aligned} C\{W_{ir}^{t_0 t_1}, W_{ir}^{t_0 t_1}\} &= 2\sigma_{w_{t_1}}^2 + 2\sigma_{w_{t_0}}^2 - 2\sigma_{w_{t_1}}^2 \exp\left(-\frac{l_{ir}}{L_{w_{t_1}}}\right) - 2\sigma_{w_{t_0}}^2 \exp\left(-\frac{l_{ir}}{L_{w_{t_0}}}\right) \\ C\{W_{ir}^{t_0 t_1}, W_{jr}^{t_0 t_1}\} &= \sigma_{w_{t_1}}^2 + \sigma_{w_{t_0}}^2 + \sigma_{w_{t_1}}^2 \exp\left(-\frac{l_{ij}}{L_{w_{t_1}}}\right) + \sigma_{w_{t_0}}^2 \exp\left(-\frac{l_{ij}}{L_{w_{t_0}}}\right) \\ &\quad - \sigma_{w_{t_1}}^2 \exp\left(-\frac{l_{ir}}{L_{w_{t_1}}}\right) - \sigma_{w_{t_0}}^2 \exp\left(-\frac{l_{ir}}{L_{w_{t_0}}}\right) \\ &\quad - \sigma_{w_{t_1}}^2 \exp\left(-\frac{l_{jr}}{L_{w_{t_1}}}\right) - \sigma_{w_{t_0}}^2 \exp\left(-\frac{l_{jr}}{L_{w_{t_0}}}\right) \end{aligned}$$

$$C\{W_{ir}^{t_0 t_1}, W_{jr}^{t_0 t_2}\} = \sigma_{w_{t_0}}^2 + \sigma_{w_{t_0}}^2 \exp(-\frac{l_{ij}}{L_{w_{t_0}}}) - \sigma_{w_{t_0}}^2 \exp(-\frac{l_{ir}}{L_{w_{t_0}}}) - \sigma_{w_{t_0}}^2 \exp(-\frac{l_{jr}}{L_{w_{t_0}}})$$

In some time series InSAR algorithms, the atmospheric effect of the reference epoch (i.e., the master image) is assumed deterministic and estimated during the processing. In that case all the terms related to the reference epoch are zero.

Using these equations all the elements of the atmospheric covariance matrix  $Q_w$  are computed. The final measurement noise covariance matrix will then be  $Q_{M_{InSAR}} = Q_n + Q_w$ , where  $Q_n$  is the covariance matrix of decorrelation/scattering noise. Note that  $Q_{M_{InSAR}}$  is the covariance matrix of the first output level of InSAR data, i.e. the unwrapped InSAR measurement prior to filtering of atmospheric effects. In order to compute the covariance matrix of other output levels, for example measurements after atmospheric filtering,  $Q_{M_{InSAR}}$  should be propagated through the further processing steps.



*Figure 2.4: Five simulated atmospheric effects over an area of 100x100 km, with different signal power based on the experimental power-law model presented by (R. Hanssen). The plots on the right are the variogram plots. Blue lines show the empirical variograms of the simulated signal. Red curves show the modelled exponential variograms. An exponential model can thus effectively approximate the degree of the spatial variability in the atmospheric signal.*

## 2.3. Idealisation noise

At the beginning of this chapter the concepts of “idealisation precision” and “idealisation noise” were introduced. In summary, the idealisation noise is the effect of other deformation sources than the signal of interest. The main sources of idealisation noise are deformation regimes such as (Ketelaar 2009, Houtenbos and Kenselaar 2001):

- Independent motion of individual benchmarks or PS with respect to the foundation layer due to for example structural instabilities, benchmark weight or pile friction. As the benchmark/PS construction settings differ among all benchmarks/PS, this autonomous deformation is assumed spatially uncorrelated. However, as deformation is developing in time, it is temporally correlated. Hence, the noise contribution of this regime shall be referred to as “temporal component”.
- Shallow compaction of the Holocene layer beneath benchmarks/PS due to for example ground water level variation or peat compaction. These mechanisms are assumed to have dependencies in both space and time. Hence, the associated noise contribution shall be referred to as “spatio-temporal component”.

In order to evaluate the spatio-temporal variability of these deformation regimes, we analysed a levelling dataset over an assumedly signal-free area close to the Wadden Sea region. The generic methodology that we have used is divided into the following steps:

1. Signal free data selection: select areas with expected zero deep source subsidence
2. Data preparation and epoch-wise levelling adjustment
3. Compute empirical spatio-temporal variograms
4. Correct the empirical variograms for the effect of measurement noise
5. Fit adequate model to the corrected variograms
6. Propose a final analytical equation to calculate (co)-variances based on the modelled variogram.

**1. Signal-free data selection:** Figure 2.5 shows the study area for the analysis. The assumption that it is unaffected by deep source subsidence is based on the absence of gas fields in the area and an evaluation of subsidence prediction models due to gas production in nearby gas fields. Another constraint was the proximity to the Wadden Sea that makes the area as representative as possible for the area of interest in terms of soil composition and other properties.

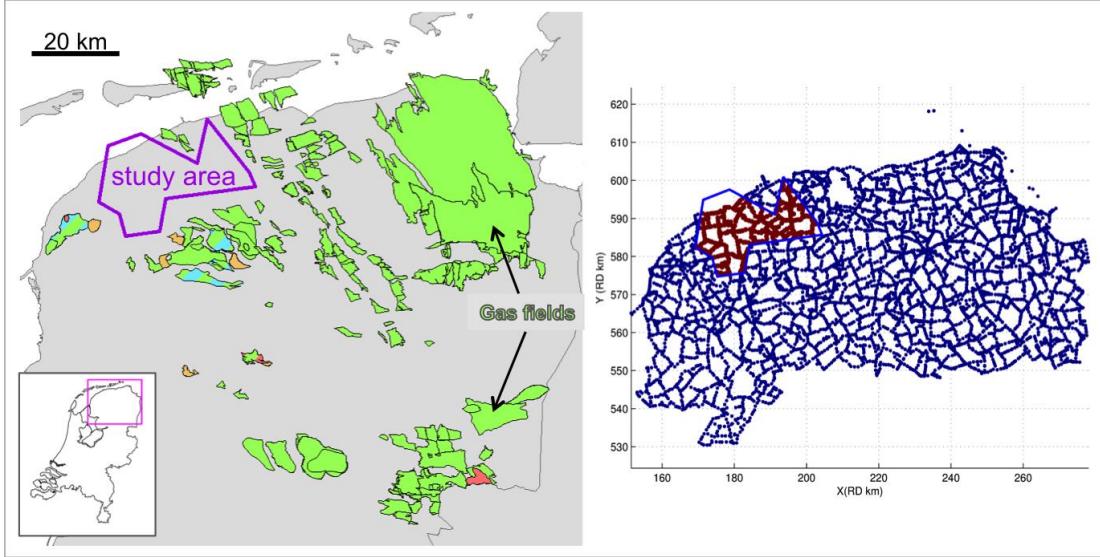


Figure 2.5: Left) selected area for levelling data analysis. Right) the levelling network over the area - Blue dots: levelling benchmarks in the northern Netherlands, Red dots: selected benchmarks

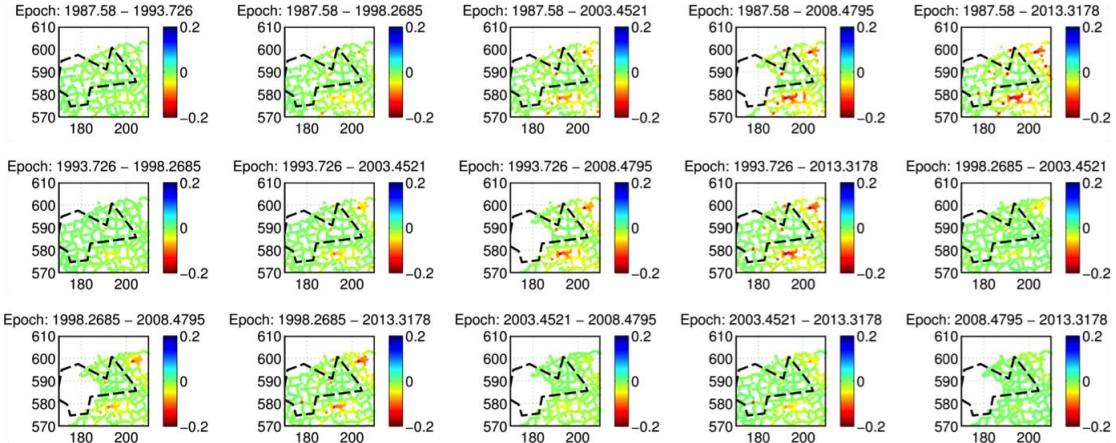


Figure 2.6: all the epoch combinations over the study area (after network adjustment). The color-scale is in meters. X-axis and Y-axis are RD coordinates in km.

**2. Data preparation:** Data from six levelling campaigns over ~25 years are included in the study: 1987, 1993, 1998, 2003, 2008 and 2013.

Network adjustment is applied per epoch with a common datum (i.e. one common reference point) for all the epochs. Then, subtracting pairs of epochs from each other results in 15 sets of double differences as visualized in Figure 2.6.

**3. Computation of empirical variograms:** The spatio-temporal empirical variograms are computed using robust algorithms (Cressie N. 1980, Genton M. 1998). The empirical variograms are averaged per spatial and temporal bins. The results are visualized in Figure 2.7 and 2.8.

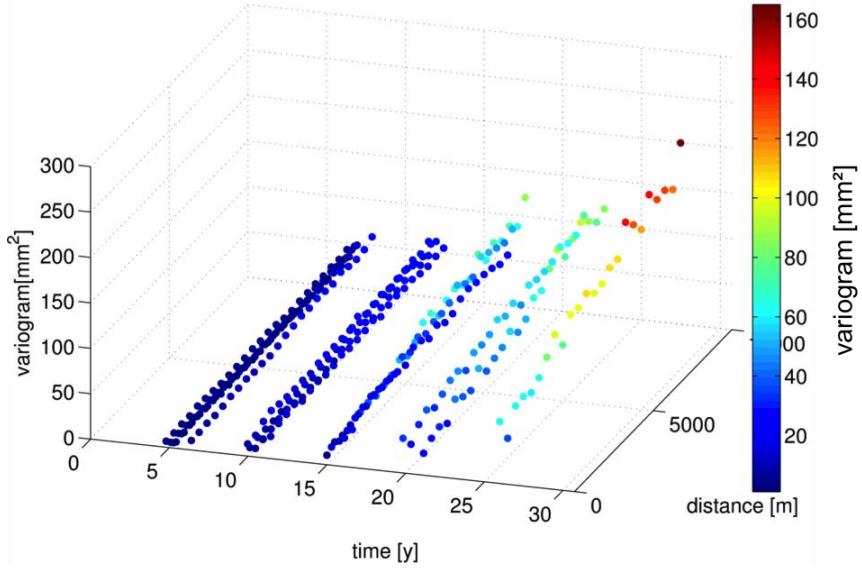


Figure 2.7: empirical spatio-temporal variograms of levelling data over the signal-free area.

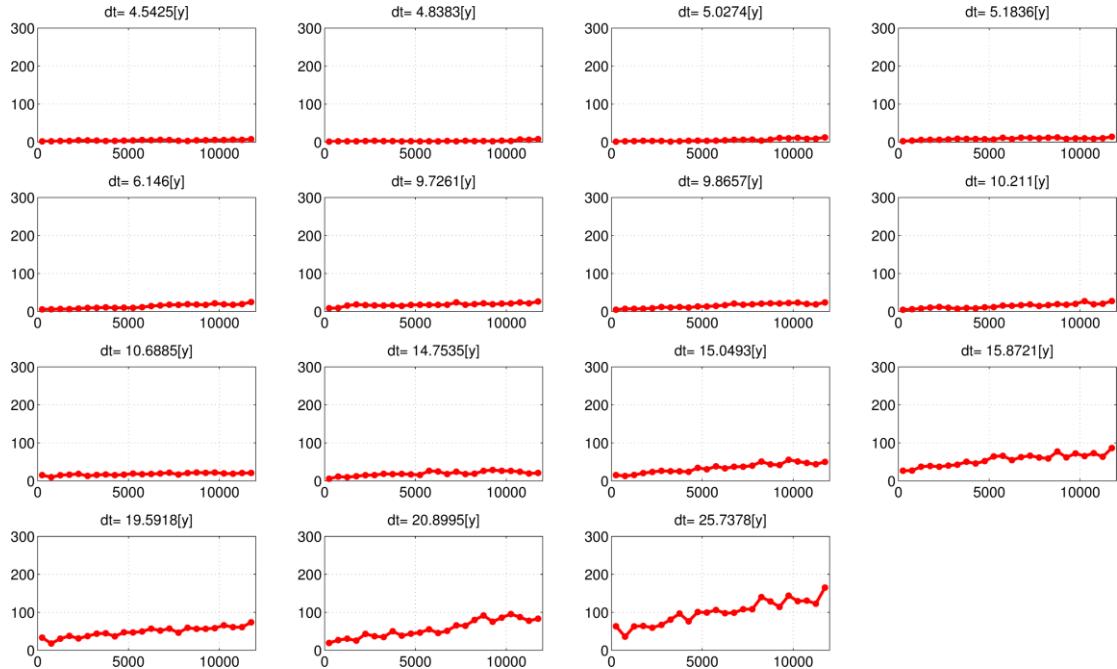


Figure 2.8: Spatial variogram profiles of levelling data over the signal-free area for different temporal lags (from 4.5 to 25.7 years). The X-axes are labelled by the spatial lag in m, and the Y-axes display the variograms in  $\text{mm}^2$

#### 4. Correction for levelling measurement noise:

Based on the measurement noise covariance matrix of section 2.1 and the network configuration of the levelling dataset, we simulated a large number ( $\sim 1000$ ) of realizations of measurement noise in the Monte-Carlo manner, followed by computation of empirical variograms for each realization. Then by averaging all the 1000 variograms, we estimate the effect of the measurement noise on the empirical variogram of the signal-free area (Figure 2.9).

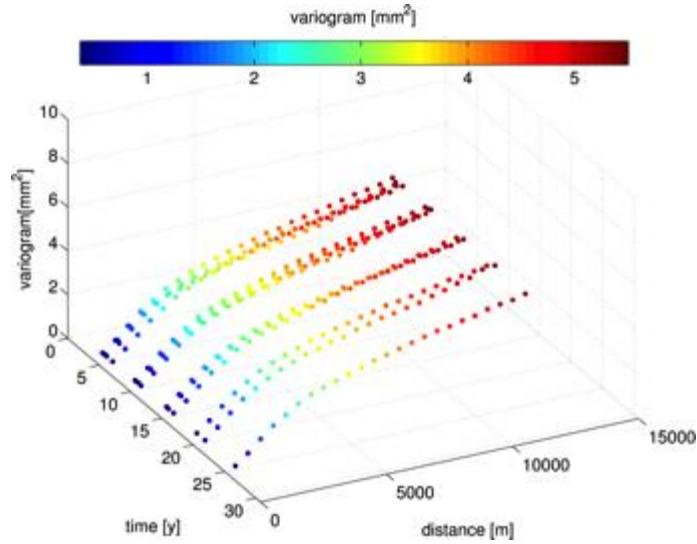


Figure 2.9: empirical spatio-temporal variogram of measurement noise (average of 1000 empirical variograms computed from realizations of measurement noise). Note that the variogram is constant in time, since individual measurement campaigns are independent in terms of measurement noise.

Comparing Figures 2.7 and 2.9, we can see that, as expected, the level of measurement noise is far below that of other noise components. The simulated variogram of measurement noise is finally subtracted from the empirical variogram calculated from levelling data in order to correct for the effect of measurement noise.

**5. Variogram modelling:** By visual/qualitative analysis of Figures 2.7 and 2.8, we recognise two different behaviours:

- A non-stationary signal in time (the variogram is unbounded in time)
- A correlated signal in space

These two behaviours can be related to the above distinguished deformation regimes, i.e. the spatio-temporal component and the temporal component of idealisation noise. To combine these two effects in a generic model we propose the following variogram model:

$$\gamma(\Delta t, \Delta d) = \left( \sigma_s^2 - \sigma_s^2 e^{-\left(\frac{\Delta d}{L}\right)} \right) \Delta t^{p_s} + \sigma_t^2 \Delta t^{p_t}$$

where:

$\Delta t$  and  $\Delta d$  are time difference and spatial distance, respectively

$\gamma$  is the variogram as a function of  $\Delta t$  and  $\Delta d$

$\sigma_s^2$  is the variance of the spatio-temporal component

$L$  is the correlation length of the spatio-temporal component

$p_s$  is the power of the non-stationary signal associated with the spatio-temporal component

$\sigma_t^2$  is the variance of the temporal component

$p_t$  is the power of the non-stationary signal associated with the temporal component.

We should note that the very same model is also proposed by (Houtenbos and Kenselaar 2001) to describe a spatio-temporal and a temporal deformation component in the Netherlands. To fit this model to the empirical variogram, there are five parameters ( $\sigma_s^2$ ,  $L$ ,  $p_s$ ,  $\sigma_t^2$ , and  $p_t$ ) to estimate. We use nonlinear weighted least squares to estimate the parameters. The weights are computed based on the number of samples per bin.

The results of the estimation, which turned out to be reasonably insensitive to their approximate values, are summarized in the following table. We should note that in general, there are different choices available to model the spatial component. Two commonly used models in geostatistics are the exponential ( $\sigma_s^2 - \sigma_s^2 e^{-(\frac{d}{L})}$ ) and the Gaussian  $\sigma_s^2 - \sigma_s^2 e^{-(\frac{d}{L})^2}$  model. In our case, the misfit of the exponential model to the empirical variograms is smaller than the misfit of the Gaussian model. This is the main reason to propose the exponential model rather than the Gaussian for the spatial component of idealisation noise.

	$\sigma_s^2$ [mm <sup>2</sup> /km/(year) <sup>p<sub>s</sub></sup> ]	L [km]	p <sub>s</sub> [-]	$\sigma_t^2$ [mm <sup>2</sup> /(year) <sup>p<sub>t</sub></sup> ]	p <sub>t</sub> [-]
Estimated parameters	0.651	12646	1.66	0.148	1.688

Table 1: Estimated parameters

The fitted models are visualized in Figures 2.10, 2.11, and 2.12.

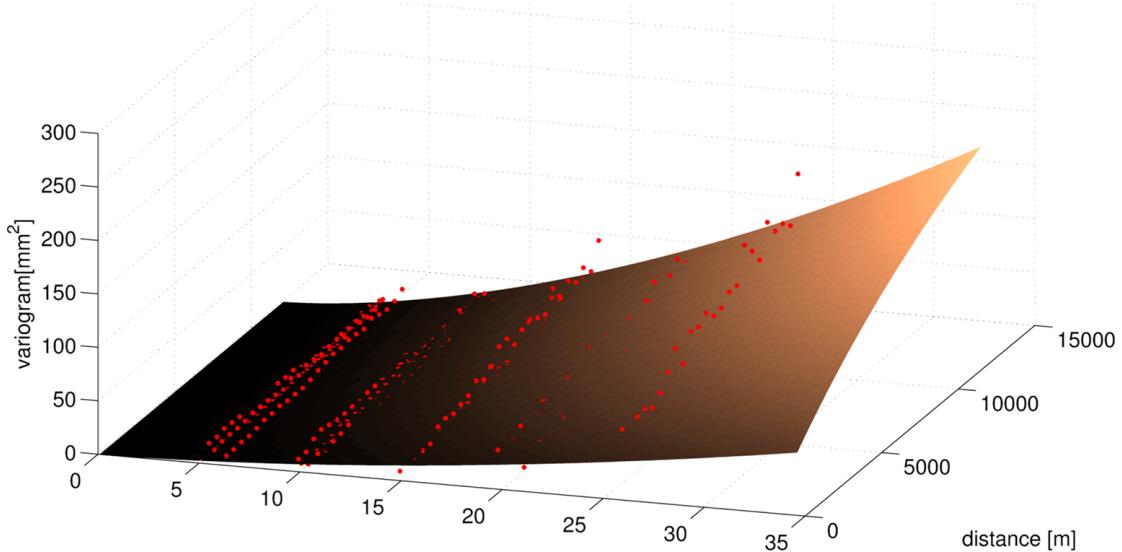


Figure 2.10: Estimated model for the empirical spatio-temporal variogram

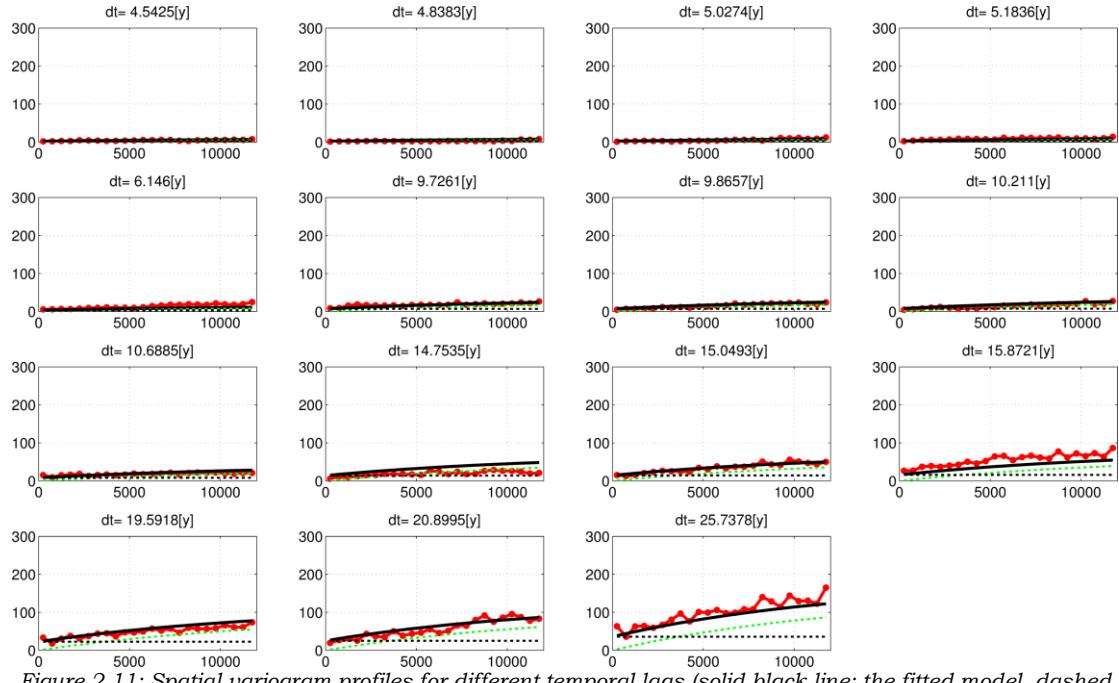


Figure 2.11: Spatial variogram profiles for different temporal lags (solid black line: the fitted model, dashed black line: fitted temporal component, dashed green line: fitted spatio-temporal component).

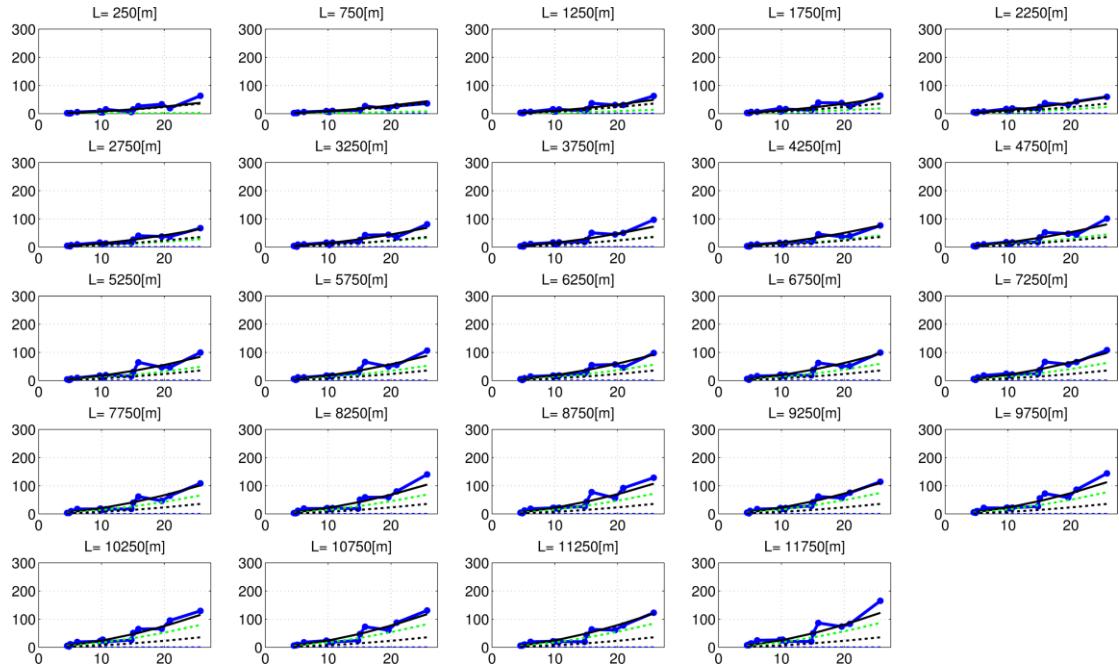


Figure 2.12: Temporal variogram profiles for different spatial lags (solid black line: the fitted model, dashed black line: fitted temporal component, dashed green line: fitted spatio-temporal component).

## 6. Construction of the covariance matrices:

The final covariance matrix of idealisation noise is a summation of the two contributions:

1. Spatio-temporal component of noise ( $Q_s$ )
2. Temporal component of noise ( $Q_t$ )

The final covariance matrix of idealisation noise will then be:  $Q_d = Q_s + Q_t$ .

### **Construction of the covariance matrix of the temporal component ( $Q_t$ )**

If  $T_{ir}^{t_0 t_1}$  is the temporal noise component of point  $i$  with respect to the reference point  $r$  between epoch  $t_1$  and the reference epoch with  $t_0 = 0$ , then the covariance elements will be (Yaglom 1962, A. Houtenbos 2004):

$$C\{T_{ir}^{t_0 t_1}, T_{ir}^{t_0 t_1}\} = 2\hat{\sigma}_t^2 |t_1|^{\hat{p}_t}$$

$$C\{T_{ir}^{t_0 t_1}, T_{jr}^{t_0 t_1}\} = \hat{\sigma}_t^2 |t_1|^{\hat{p}_t}$$

$$C\{T_{ir}^{t_0 t_1}, T_{ir}^{t_0 t_2}\} = \hat{\sigma}_t^2 (|t_1|^{\hat{p}_t} + |t_2|^{\hat{p}_t} - |t_1 - t_2|^{\hat{p}_t})$$

$$C\{T_{ir}^{t_0 t_1}, T_{jr}^{t_0 t_2}\} = \frac{1}{2} \hat{\sigma}_t^2 (|t_1|^{\hat{p}_t} + |t_2|^{\hat{p}_t} - |t_1 - t_2|^{\hat{p}_t})$$

where  $\hat{\sigma}_t^2$  and  $\hat{p}_t$  are the estimated parameters from variogram modelling.

### **Construction of the covariance matrix of the spatio-temporal component ( $Q_s$ )**

If  $S_{ir}^{t_0 t_1}$  is the spatio-temporal noise component of point  $i$  with respect to the reference point  $r$  between epoch  $t_1$  and the reference epoch with  $t_0 = 0$ , then:

$$C\{S_{ir}^{t_0 t_1}, S_{ir}^{t_0 t_1}\} = 2\hat{\sigma}_s^2 |t_1|^{\hat{p}_s} \left(1 - \exp\left(-\frac{l_{ir}}{\hat{L}}\right)\right)$$

$$C\{S_{ir}^{t_0 t_1}, S_{jr}^{t_0 t_1}\} = \hat{\sigma}_s^2 |t_1|^{\hat{p}_s} \left(1 - \exp\left(-\frac{l_{ir}}{\hat{L}}\right) - \exp\left(-\frac{l_{jr}}{\hat{L}}\right) + \exp\left(-\frac{l_{ij}}{\hat{L}}\right)\right)$$

$$C\{S_{ir}^{t_0 t_1}, S_{ir}^{t_0 t_2}\} = \hat{\sigma}_s^2 (|t_1|^{\hat{p}_s} + |t_2|^{\hat{p}_s} - |t_1 - t_2|^{\hat{p}_s}) \left(1 - \exp\left(-\frac{l_{ir}}{\hat{L}}\right)\right)$$

$$C\{S_{ir}^{t_0 t_1}, S_{jr}^{t_0 t_2}\} = \frac{1}{2} \hat{\sigma}_s^2 (|t_1|^{\hat{p}_s} + |t_2|^{\hat{p}_s} - |t_1 - t_2|^{\hat{p}_s}) \left(1 - \exp\left(-\frac{l_{ir}}{\hat{L}}\right) - \exp\left(-\frac{l_{jr}}{\hat{L}}\right) + \exp\left(-\frac{l_{ij}}{\hat{L}}\right)\right)$$

where  $\hat{\sigma}_s^2$  and  $\hat{p}_s$  and  $\hat{L}$  are the estimated parameters from variogram modelling, and  $l_{ir}$  is the distance between points  $i$  and  $r$ . Using these equations, all the elements of the idealisation noise covariance matrix  $Q_d = Q_s + Q_t$  are constructed. The final covariance matrix of levelling or InSAR observations is then constructed as:

$$Q_y = Q_e + Q_d,$$

where  $Q_e$  is the measurement noise covariance matrix (i.e.  $Q_m$  in case of levelling and  $Q_{InSAR}$  in case of InSAR). In order to get a feeling about the spatio-temporal significance of the proposed idealisation noise model, we demonstrate the spatial variogram of the different noise model components for levelling double differences (in Figure 2.13). The spatial variograms are evaluated for time spans of 1, 5, 20, and 50 years. Note that the measurement noise component here is just an arbitrary example. In practice the contribution of levelling measurement noise is adaptive and depends on the levelling network configuration. As a demonstration, Figure 2.14 shows the  $Q_s$  and  $Q_t$  covariance matrices for the stylized levelling network of Figure 2.1 with six epochs at 0, 6, 12, 18, 24, and 30 years (i.e. five double difference epochs with respect to the reference epoch  $t = 0$ ).

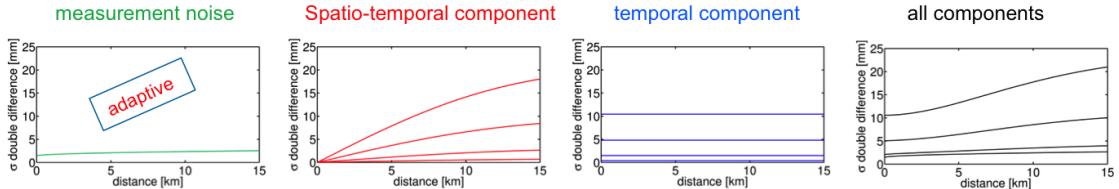


Figure 2.13: Spatial variogram plot of different component of the proposed model evaluated for time spans of 1, 5, 20, and 50 years. Note that the measurement noise component here is just an arbitrary example for a levelling network. In practice the contribution of levelling measurement noise is invariant in time and adaptive to the levelling network configuration.

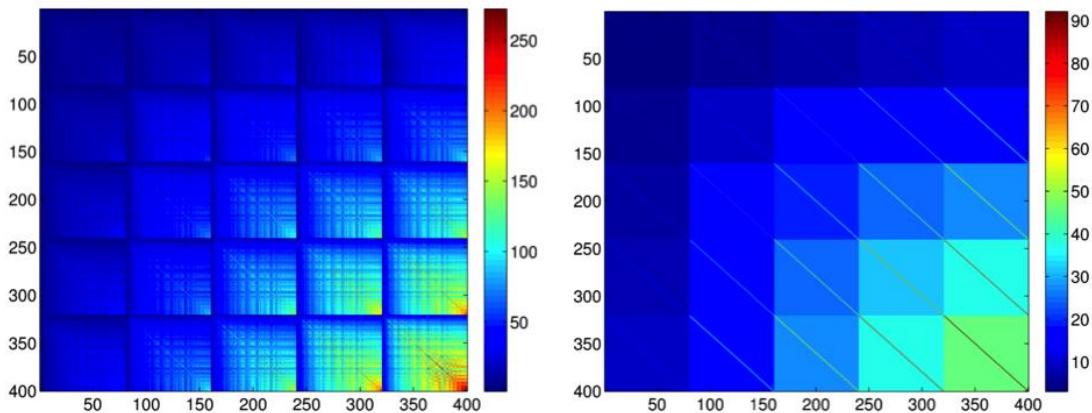


Figure 2.14:  $Q_s$  (left) and  $Q_t$  (right) for the stylized levelling network of Figure 2.1 with six epochs at 0, 6, 12, 18, 24, and 30 years (i.e. five double difference epochs with respect to the reference epoch  $t = 0$ ). Colorbars are in  $\text{mm}^2$ .

### 2.3.1. Discussion on the idealisation noise model

#### Interpretation of the proposed model:

The temporal component  $\sigma_t^2 \Delta t^{p_t}$ , which describes temporally correlated and spatially uncorrelated noise, can be interpreted as representation of autonomous benchmark movements or settlements with respect to the foundation layer due to benchmark weight and pile friction. As the benchmark construction settings differ among all benchmarks, these autonomous deformations are spatially uncorrelated. In general, the power variogram model is a representation of fractional Brownian motion noise. In the case of  $p_t = 1$ , the model reduces to the random walk model.

The spatial component  $\left(\sigma_s^2 - \sigma_s^2 e^{-\left(\frac{\Delta d}{L}\right)}\right) \Delta t^{p_s}$  represents any spatio-temporally correlated motion. The most likely cause for such a motion is shallow compaction of the Holocene layer beneath benchmarks. The exponential spatial variogram model describes the signal with a spatial correlation of range  $L$ . The temporal power model ( $\Delta t^{p_s}$ ) takes the non-stationary behaviour into account. From the physical point of view, since shallow compaction is developing in time, we expect the effect of this signal to be non-stationary.

Based on this analysis we can conclude:

1. There is indication of temporal non-stationary noise in the levelling dataset. This can be explained by independent motion of individual benchmarks, by spatially correlated deformation due to shallow compaction or by both which seems most likely.

2. There is indication of spatially correlated noise in the levelling dataset. This can be explained by either shallow compaction or by deep source deformation in the selected area, whereas the latter is not likely.

### **On the applicability of the proposed model for InSAR data:**

Although the parameters of the proposed idealisation noise model have been estimated from levelling data, we propose here to confer the stochastic deformation properties of levelling benchmarks to InSAR persistent scatterers. The rationale is as follows. If these idealisation noise contributions are induced by shallow deformation (either spatially correlated or not), both levelling benchmarks and PS should be affected in a similar manner by these shallow sources. It could be argued that because the two techniques measure different objects, the effects of shallow deformation on them are not necessarily comparable. We acknowledge that the idealisation noise in InSAR can be more prominent than in levelling, as levelling benchmarks are usually attached to buildings with better foundations. However, for the PS that are associated with buildings, the proposed model should serve as a fair approximation of idealisation noise. For all other PS (on the ground or other objects), the model most likely underestimates the idealisation noise, but is nevertheless still the lower bound approximation. Further classification of PS into high and low points can be beneficial here in order to identify PS which are associated with buildings and high infrastructures.

One could also suggest estimating the same noise parameters from InSAR data over a signal-free area. This was indeed intended at the start of the project. Using InSAR data from an assumedly stable area to deduce a noise model for PS displacements due to shallow sources would certainly have been more straightforward than conferring a noise model derived from levelling data. It would also have provided the opportunity to validate the models from different techniques against each other.

However, during the study, it turned out that the final InSAR deformation results are affected by InSAR processing settings, mainly by the spatio-temporal atmospheric filtering. In other words, the filtering of the atmospheric component alters the spatio-temporal correlation structure of InSAR deformation estimates. Since the stochastic properties turned out to be significantly affected by the filtering step, it is not possible to discriminate and isolate the stochastic properties of shallow deformation effects from filtered InSAR data.

Inference from unfiltered InSAR data may be possible in principle but it is considerably complicated as it would be hampered by the dominant atmospheric noise. As the latter significantly differs per image, its consideration would require a variogram estimation with atmospheric noise parameters per SAR image, and thus the number of variance components scales with the number of images. However, a joint estimation of a large number of variance components is usually a very unstable optimisation algorithm with little chance to converge to a unique solution. In conclusion, finding a more rigorous description for InSAR idealisation noise would be a very challenging task with no straightforward solution that could not be covered within the scope of this project.

We should note that the noise model derived from levelling data applies to shallow deformation in vertical direction only. To confer it to InSAR data, we need to take into account that InSAR observations are made in the line of sight (LOS) direction of the radar satellite. Hence,  $\sigma_t^2$  and  $\sigma_s^2$  are converted to that direction under the assumption that the horizontal component of shallow deformation noise is negligible. This is a reasonable assumption, since shallow deformation mechanisms work mainly in vertical direction. Assuming an incidence angle of  $\theta$ , the converted variances are computed as

$$\sigma_{t\_LOS}^2 = \sigma_t^2 \cos(\theta) \text{ and } \sigma_{s\_LOS}^2 = \sigma_s^2 \cos(\theta).$$

### **On the zero mean and directional properties of idealisation noise components:**

At the first glance, the idea of stochastic modelling of shallow deformation sources via the proposed covariance (variogram) function may suggest that vertical motions in the up and down direction are equally likely. However, we should note that we are dealing with double difference observations. So if all the benchmarks displace into one direction in an absolute sense, the double difference observations are insensitive to the mean (single-difference) displacement of all the benchmarks. So in other words, the deviation of vertical motion with respect to the mean deformation is equally likely to be in the up and down direction.

However, the mean of double difference shallow deformations is not necessarily zero, and its value depends on the choice of the reference point. The fulfilment of the zero-mean property that is implied by the stochastic model depends on the choice of the reference point(s). This has implications for variogram estimation. In the variogram estimation methodology, we have removed the mean deformation of each epoch in order to yield zero-mean double differences. So the empirical variograms are invariant with respect to the choice of reference point.

Note that the mean shallow displacement signal is present in the data but described neither by the stochastic model nor by the geo-mechanical model. In the next chapter we discuss that the non-zero mean effect can be mitigated effectively by choosing multiple reference points.

## **2.4. Propagation of noise models through InSAR processing**

As mentioned before in section 2.2, the proposed covariance model for InSAR data ( $Q_{y_{InSAR}} = Q_{M_{InSAR}} + Q_d$ ) is the covariance matrix of the first level of InSAR data, i.e. unwrapped InSAR measurement prior to filtering of atmospheric effects. In order to compute the covariance matrix of other InSAR output levels (as explained in Section 3.4), for example deformation time series after atmospheric filtering,  $Q_{y_{InSAR}}$  should be propagated through the InSAR filtering step. Evaluating a closed-form analytical error propagation for spatio-temporal atmospheric filtering is a complicated problem. Here, we propose easy and efficient Monte-Carlo error propagation to account for the processing noise induced during filtering. The steps of this approach can be summarized as following:

1. Construct  $Q_{y_{InSAR}}$  for a hypothetical dataset that has the same number of images as the InSAR stack but a limited number of PS
2. Create a large number of noise realizations (e.g. 1000) based on  $Q_{y_{InSAR}}$
3. Apply the atmospheric filtering on all the realizations using the same settings as for the real dataset
4. Compute the empirical covariance matrix from all the filtered realizations
5. Fit an analytical covariance model to the empirical covariance matrix
6. Use the fitted analytical model to fill all the elements of the final covariance matrix  $Q_{y_{InSAR\_filtered}}$ .

In principle this simulation approach can be applied for all the PS locations in InSAR stack, and directly provide the final covariance matrix without applying steps 5 and 6. However, in practice this is not computationally feasible due to the large data volume. This is the main reason behind step 5 to use an analytical model. Furthermore by using admissible analytical covariance/variogram function, we avoid singularity in the empirical covariance matrix due to numerical issues. We use here the same covariance/variogram function that we proposed for idealisation noise. The only difference is that we consider also a nugget effect in order to account for spatio-temporally uncorrelated noise. In principle, this model is quite generic in

the sense that it can describe all the different spatio-temporal noise behaviors such as spatio-temporally uncorrelated (nugget effect), spatio-temporally correlated, and temporally correlated noise contributions.

Here we demonstrate the covariance construction for a stylized InSAR dataset (Figure 2.15.) with 264 PS points and 25 radar acquisitions with a 35 days' time interval. 64 of the PS are 1<sup>st</sup> order PS, which are used for atmospheric spatial interpolation (kriging). For atmospheric filtering we used the methodology of the Delft Persistent Scatterer Interferometry (DePSI) presented in (van Leijen 2014). However, in principle any other algorithm can be used.

Figure 2.16 shows the covariance matrices for different InSAR noise components, that is  $Q_n$ ,  $Q_w$ ,  $Q_t$ , and  $Q_s$ . The summation of all these covariance matrices gives  $Q_{y_{InSAR}}$  (Figure 2.17.left) :

$$Q_{y_{InSAR}} = Q_n + Q_w + Q_s + Q_t.$$

This matrix is propagated through InSAR atmospheric filtering using the proposed Monte-Carlo approach. The final covariance matrix  $Q_{y_{InSAR\_filtered}}$  of the filtered InSAR dataset is visualized in Figure 2.17.right.

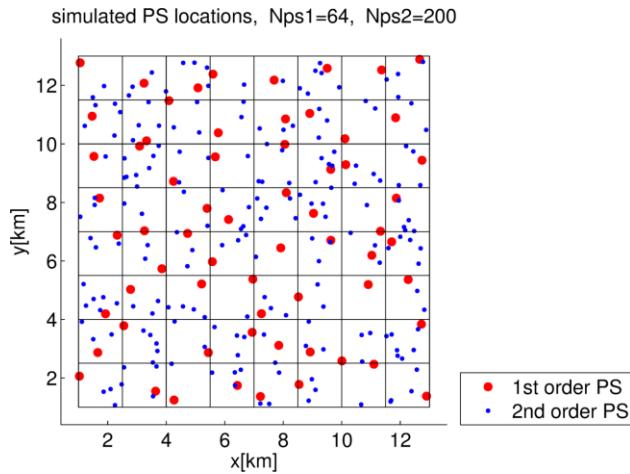


Figure 2.15: Stylized dataset of 1<sup>st</sup> and 2<sup>nd</sup> order persistent scatterers. (In total 264 PS).

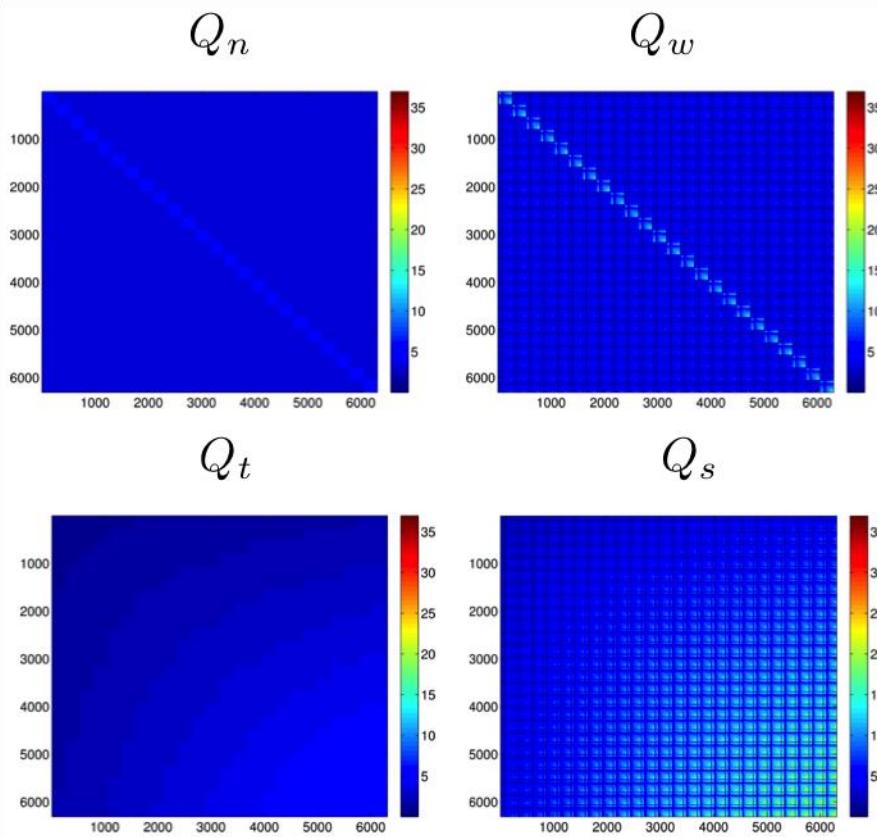


Figure 2.16: Covariance matrices for InSAR noise components for the stylized dataset of Figure 2.15:  $Q_n$  is the covariance matrix of InSAR scattering noise,  $Q_w$  is the covariance matrix of InSAR atmospheric noise,  $Q_t$  is the covariance matrix of the temporal component of idealisation noise (converted to LOS direction), and  $Q_s$  is the covariance matrix of the spatio-temporal component of idealisation noise (converted to LOS direction). Colorbars are in mm.

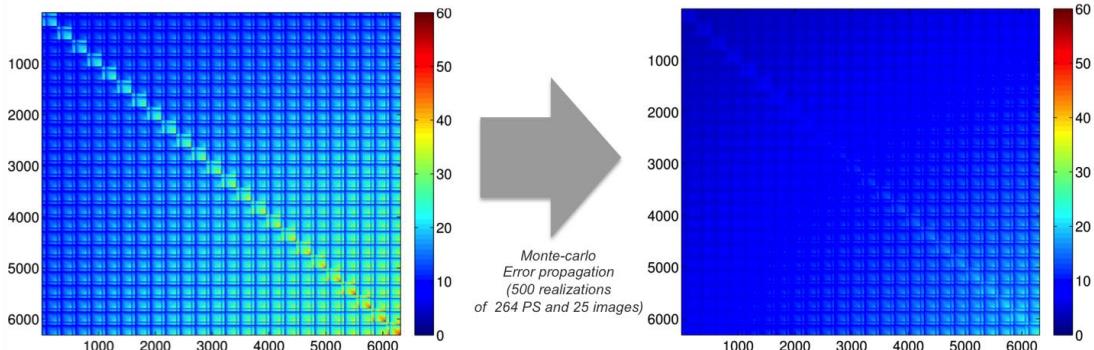


Figure 2.17: Left) Covariance matrix  $Q_{y_{\text{InSAR}}}$  for the stylized dataset of the Figure 2.15 as  $Q_{y_{\text{InSAR}}} = Q_n + Q_w + Q_s + Q_t$ , Right) the final covariance matrix  $Q_{y_{\text{InSAR}}, \text{filtered}}$  of the filtered InSAR dataset as the result of Monte-Carlo error propagation through InSAR atmospheric filtering (Colorbar is in mm<sup>2</sup>).

## 2.5. Summary and conclusions

In this section, we proposed an analytical approach to construct a noise covariance matrix for both InSAR and leveling datasets. The concept of idealisation noise was exploited to explain deformation sources other than the signal of interest that is the deep source deformation. We estimated the parameters of the idealisation noise model using spatio-temporal variogram modelling. The summary of the conclusions of this section are listed as following:

1. There is temporally non-stationary noise in the levelling dataset. This can be explained by independent motion of individual benchmarks, by spatially correlated deformation due to shallow compaction or by both, which seems most likely.
2. There is spatially correlated noise in the levelling dataset. This can be explained by either shallow compaction or by deep source deformation in the selected area, whereas the latter is not likely.
3. The proposed idealisation noise model can reasonably explain the uncertainty of the studied dataset. It may be representative for the coastal regions around the Wadden Sea. For other regions in the Netherlands with harder soil it may still serve as a conservative approximation, whereas the possibility of overestimated uncertainties should be taken into account.
4. We propose to use a state of the art analytical model for InSAR scattering and atmospheric noise components.
5. InSAR processing (especially the atmospheric filtering step) affects the spatio-temporal behavior of the other noise components and introduces spatio-temporal correlation in the final deformation estimates. We propose a Monte Carlo approach to account for the spatio-temporal correlation induced by filtering out the atmospheric component.
6. We propose to use the idealisation noise model derived from levelling data also for InSAR data, assuming that the idealisation noise effect on persistent scatterers is similar with the effect on levelling benchmarks.



### 3. Output level study

Geo-mechanical modelling/calibration can take place using different levels of geodetic processing ranging from raw observations to elaborately tailored products. This is conceptually visualized in Figure 3.1. The main question here is what the best output level of geodetic data is to be used in the modelling based on three main considerations:

1. Transparent use of a priori knowledge: the more processing steps are involved, the more a priori knowledge (including non-validated assumptions) may be needed.
2. Simpler structure of the stochastic model: the more processing steps are involved, the more complicated the stochastic model may become, as all the error sources should be propagated through the processing steps.
3. Modelling efficiency/optimality: geomechanical model computations may become more complicated if the model is compared with geodetic data at a lower processing level.

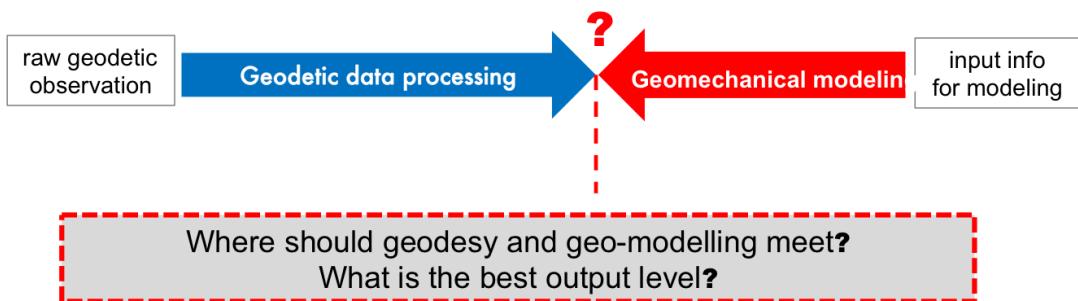


Figure 3.1: Conceptual sketch of the main objective of the output level study: to answer where geodesy and geo-modelling should meet.

This chapter presents the results of our simulation study to find the best output level of InSAR and leveling data for geo-mechanical modelling. Section 3.2 covers the generic methodology that we have used. The results and discussion of the output level study for leveling and InSAR data are presented in sections 3.3 and 3.4 respectively.

### 3.1. Output level study: generic methodology

Our methodology for the output level study can be outlined in four steps:

1. Assume (or initialize) a simple reservoir and geomechanical model with known parameters and construct the subsidence field based on the assumed model.
2. Simulate a large number of realizations of geodetic observations in different output levels by summation of the constructed subsidence field and a large number of geodetic noise realizations.
3. Estimate geomechanical model parameters by searching the parameter space for each set of geodetic observations.
4. Give recommendation on the optimal output level based on a posteriori statistics of model parameters.

In the following, these steps and assumed parameters/models are described in detail.

#### 1. Assumed reservoir model and construction of subsidence field

Even though the shape, properties, and internal pressure variation of real gas reservoirs can be complex, anisotropic and variable, we focused on a simple reservoir/geomechanical model for the purpose of this study. It has the following properties:

- Reservoir shape: disk-shape reservoir at center position  $(x_c, y_c)$  and depth  $D$  with thickness  $H$  and radius  $R$  (Figure 3.2a).
- Pressure variability: uniform pressure change ( $\Delta p$ ) through the reservoir.
- Compaction model: time decay compaction model, with time decay relation between pressure change and compaction. In this model, the volumetric strain  $e_{ii}$  at a point in the reservoir at time  $t$  is the product of pressure change  $\Delta p$  and constrained uniaxial compressibility  $C_m$ , but convolved with an exponential time decay function:

$$e_{ii} = \Delta p C_m *_t \frac{1}{T} \exp\left(\frac{-t}{T}\right),$$

where  $*_t$  is the convolution operator with respect to time, and  $T$  is the time decay constant. Figure 3.2b demonstrates this time decay model for  $T = 0$ , 2, and 8 years.

- Compaction-subsidence relation: Geertsma analytical subsidence model for a disk shape reservoir (Geertsma 1973, Fjaer 1992).

By setting all the model parameters ( $C_m, T, H, R, D, x_c, y_c$ ) and having a pressure change as a function of time  $\Delta p(t)$ , both vertical and horizontal components of subsidence at any point on the surface are calculated using the Geertsma analytical model.

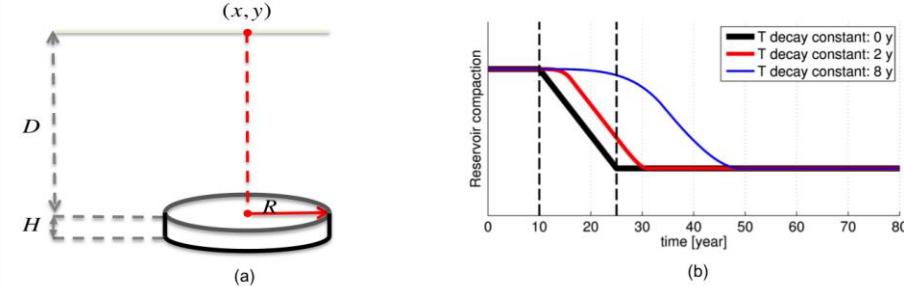


Figure 3.2: a) sketch of the disk-shape reservoir model, b) demonstration of the time decay model with  $\tau = 0, 2, 8$  years. Dashed lines show the start and end of the pressure depletion (or start and end of the production, respectively).

## 2. Simulation of geodetic data.

Simulation of geodetic data generally is done in two separate steps of “noise” and “signal” simulation. As discussed, the “signal” component is simulated based the assumed reservoir model using Geertsma analytical model. However, the simulated signal is transformed based on the configuration of geodetic data. For example in InSAR, the vertical and horizontal signal components are transformed to the LOS direction based on the acquisition geometry. For both levelling and InSAR data, the single-difference subsidence signal is converted to double differences based on the spatio-temporal configuration of the observations. We will give more examples of these simulations for InSAR and leveling in the next sections.

For the simulation of noise realizations, we use the proposed stochastic models for measurement and idealisation noise in levelling and InSAR. Given the spatio-temporal configuration of geodetic data, the covariance matrix of all noise components is computed based on the methodology introduced in chapter 2.

Assuming that noise components have a multivariate normal distribution, the noise is then generated using the constructed covariance matrix. If  $S$  is the vector of the simulated signal and  $\epsilon_i$  is the vector of the  $i$ th noise realization with  $i = 1 \dots N$ , then  $N$  geodetic datasets are computed as  $y_i = S + \epsilon_i$ .

## 3. Geomechanical modelling (Inversion)

For geomechanical modelling, we assume that all the model parameters related to shape/location of the reservoir (e.g.,  $H, R, D, x_c, y_c$ ) and also pressure changes  $\Delta p(t)$  are known. It is a common practice in geomechanical modelling/calibration that the shape and pressure values are fixed, or in other words are computed a priori based on the geological and reservoir models. Then during geomechanical calibration only few geomechanical parameters are calibrated with respect to geodetic data. For the purpose of this study, we assume there are only two unknown parameters in the model: a uniaxial compressibility  $C_m$  and a time decay parameter  $T$ . Note that in practice, there is always a good approximate value for  $C_m$ . Assuming the approximate value to be  $C_{m_0}$ , we introduce a  $C_m$  proportionality factor  $\alpha$  as:  $C_m = \alpha C_{m_0}$ . So the two unknown parameters in our model are  $\alpha$  and  $T$ .

Although the assumed model is linear with respect to  $\alpha$ , this is not the case with respect to the time decay parameter  $T$ . In order to avoid the complexities introduced by nonlinear inversion methods, we use here simple forward modelling by searching the two-dimensional parameter space. The objective function for this inversion is the L2-norm of the misfit between data and model. Let’s assume  $y$  is the vector of observations, and  $G(\alpha, T)$  is the geomechanical model. Then the mathematical model for the inversion can be written as:

$$E\{y\} = G(\alpha, T), \quad D\{y\} = Q_y,$$

where  $E\{\cdot\}$  and  $D\{\cdot\}$  are expectation and dispersion operators, respectively. The optimal L2-norm solution of this model is given by:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{T} \end{bmatrix} = \min_{(\alpha, T)} \|y - G(\alpha, T)\|_{Q_y^{-1}}^2,$$

or if  $e = y - G(\alpha, T)$  then:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{T} \end{bmatrix} = \min_{(\alpha, T)} (e^T Q_y^{-1} e).$$

The solution is obtained from the direct search of the solution space for the minimum of the objective function. Figure 3.3 gives an example of such a search space.

In this study, in order to evaluate the effect of covariance matrix on the final estimates, we consider three different L2-norm objective functions:

- Simple L2-norm of the residuals:

$$\min_{(\alpha, T)} (e^T e)$$

- L2-norm of the residuals in the metric of the weight matrix with weights equal to the inverse of observation variance:

$$\min_{(\alpha, T)} (e^T W_y e), \text{ where } W_y = (\text{diag}(Q_y))^{-1}$$

- L2-norm of the residuals in the metric of the full covariance matrix of the observations:

$$\min_{(\alpha, T)} (e^T Q_y^{-1} e).$$

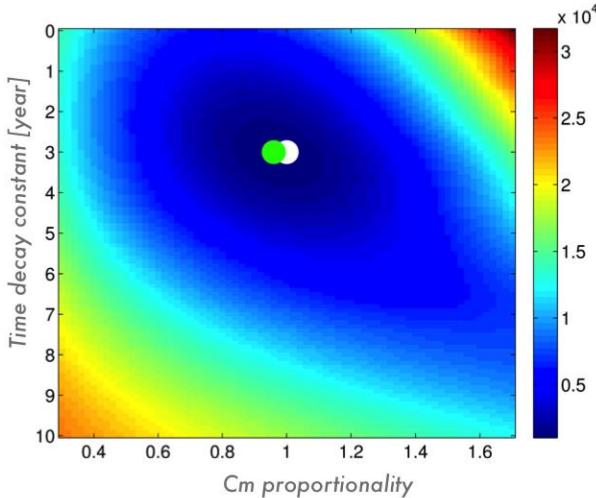


Figure 3.3: Example of the two dimensional search space. The color-scale shows the L2-norm of the residuals:  $\|y - G(\alpha, T)\|_{Q_y^{-1}}^2$ . The white dot shows the true value and the green dot shows the solution with minimum L2-norm.

#### 4. A posteriori analysis

After applying inversion to all  $N$  realizations of geodetic data, we have  $N$  estimates of unknown parameters. Statistical analysis of this large number of estimates can provide us with different quality measures of the inversion. For example, the agreement between the mean value of all the estimates and the true value is a measure of unbiasedness of the inversion. Dispersion (e.g. standard deviation) of the estimates gives the precision of the inversion. The full covariance matrix (or even the two-dimensional probability density function) of the estimated parameters can be computed empirically using all the estimates. Based on these kinds of posterior analyses, we can make formal statements about the effect of using different output levels on the quality of geomechanical modelling, and consequently on subsidence prediction.

In the next two sections, we will present the results of the application of the proposed methodology to both leveling and InSAR data.

## 3.2. Output level Study: Levelling

### 3.2.1. Output level overview

The overview of the output levels for leveling data is:

- **Level10:** Raw levelling campaign observations, that is height difference observations between benchmarks. This level is the direct result of levelling campaigns.
- **Level11:** Adjusted heights per epoch or height differences per epoch with respect to a single reference benchmark (or with respect to a common reference frame defined by all benchmarks or a subset of benchmarks respectively). This level is the result of least squares free network adjustment per epoch.
- **Level12:** Double difference heights with respect to a common reference benchmark and a common reference epoch. This is the result of temporal differentiation of adjusted heights per epoch with respect to the height in a reference epoch.
- **Level13:** Double difference heights with respect to a common reference benchmark but multiple reference epochs. This is the result of temporal differentiation of level 2 data.
- **Level14:** Double difference heights with respect to a common reference epoch but multiple reference benchmarks. This is the result of spatial differentiation of level 2 data.
- **Level15:** Double difference heights with respect to multiple reference epochs and multiple reference benchmarks. This is the result of both spatial and temporal differentiation of level 2 data.

Every level can be transformed to any other level by linear transformation (except for backward transformation to Level 0 and 1). Consequently the covariance matrix of each level can be transformed to any other level by linear error propagation. In this study we disregard the output levels 0 and 1 to be used in geomechanical modelling. The main reason for considering only the data levels after network adjustment is to take advantage of the strong mathematical model of the adjustment step. The usual high redundancy of levelling networks together with the strongly valid constraint of closing loops allows us to effectively adjust the measurement noise components and detect possible outliers in the network. Furthermore, the Level1 data (i.e. single differences in space) was not considered in the study, because the pure (relative) benchmark heights do not have any information about the subsidence or geomechanical parameters and thus not any added value for geomechanical modelling. Therefore, only those output levels are considered in the study that have double-difference observations after network adjustment (i.e., levels 2,3,4, and 5).

The schematic visualization of levels 2 to 5 is given in Figure 3.4. Note that we do not introduce any fake redundancy in levels 3 to 5 with multiple reference points/epochs; i.e. the number of observations is the same for all the levels 2 to 5.

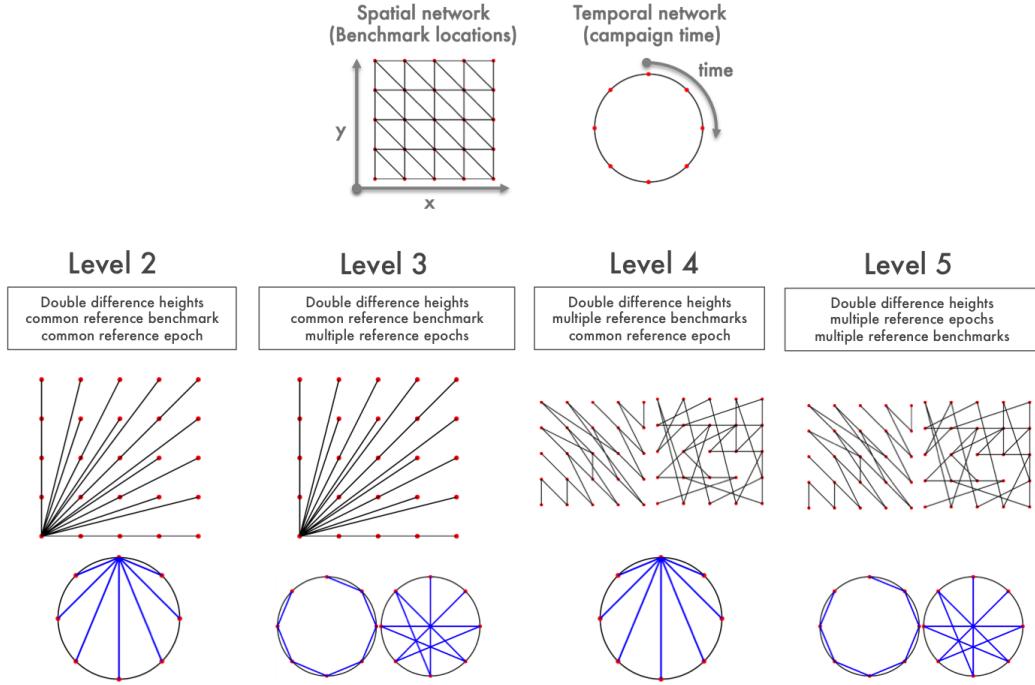


Figure 3.4: The schematic visualization of levelling output levels 2 to 5.

### 3.2.2. Simulation setting

**Reservoir/geomechanical parameters:** We applied the output level study methodology on three different scenarios with respect to production history and leveling temporal sampling. Except for pressure change rate, all the other geomechanical/reservoir parameters are the same for the three scenarios. The parameters used in the simulation are summarized in table 3.1. The reservoir/geomechanical parameters are a very rough approximation of the parameters of the Ameland gas field. Also all the scenarios are evaluated with time decay parameters of 1 to 5 years.

Reservoir radius ( $R$ )	5 km
Reservoir depth ( $D$ )	3 km
Reservoir Thickness ( $H$ )	100 m
Poisson ratio	0.3
Compaction coefficient ( $C_m$ )	$7.4 \cdot 10^{-11} \text{ Pa}^{-1}$
Time decay parameter ( $T$ )	1,2, ..., 5 years
Start of the production	$t=3$ years
End of the production	$t=18$ years
Pressure depletion rate ( $\partial \Delta p / \partial t$ )	(scenario 1) $2.0 \cdot 10^6 \text{ Pa/year}$
	(scenario 2) $2.0 \cdot 10^5 \text{ Pa/year}$
	(scenario 3) $2.0 \cdot 10^6 \text{ Pa/year}$
Observation period	30 years
Sampling interval	6 years
	30 years
	6 years
	12 years
	3 years

Table 3.1: parameters used in the simulation for the leveling output level study

A depicted overview of the production/pressure history and the observation period of all the scenarios is shown in Figure 3.5. The production period is assumed to be 15 years, starting at  $t=3$  years (with respect to the modelling reference time). The pressure is assumed to deplete linearly with time during the production period. In scenarios 1 and 2, we have leveling data every 6 years starting from  $t=0$  till  $t=30$  years. The only difference between scenarios 1 and 2 is the pressure change rate, which is 10 times smaller for scenario2. In scenario 3, we have the same pressure

change as in scenario 1, but leveling data is available only during the early field life from  $t=0$  till  $t=12$  years and with higher temporal sampling of every 3 years.

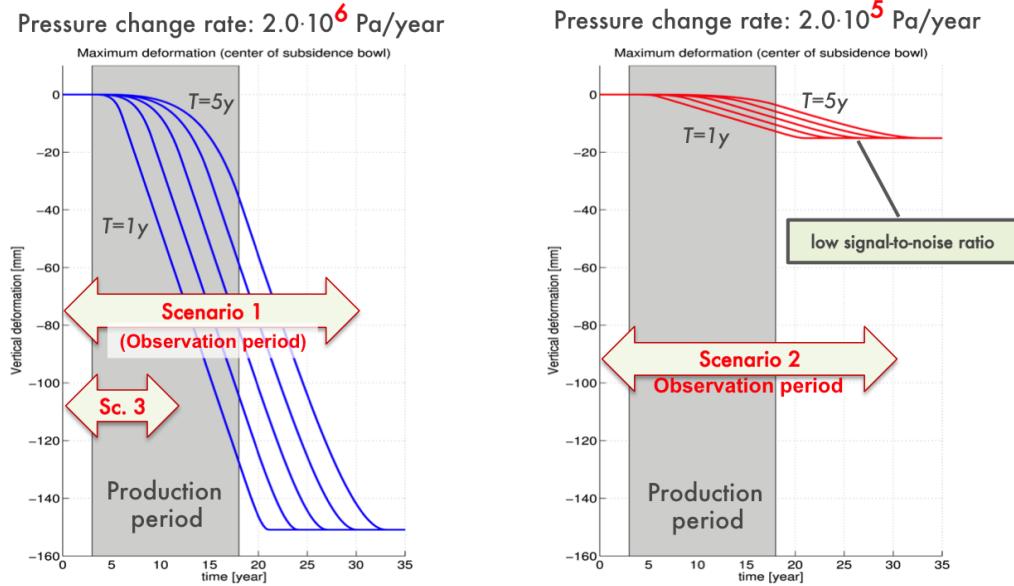


Figure 3.5: The depicted overview of the production/pressure history and the observation period of all the three scenarios.

**Leveling network:** We used the stylized leveling network visualized in Figure 2.1. The noise components are simulated based on the covariance matrices computed from the noise model presented in Sections 2.1 and 2.3.

**Inversion search space:** The Inversion is done by searching the two-dimensional search space of the unknown parameters (proportionality factor  $\alpha$  and time decay parameter  $T$ ). The search space range is [0.5 1.5] for  $\alpha$  and [0 12] years  $T$ . The sampling interval of the search space is 0.02 and 0.2 years for  $\alpha$  and  $T$ , respectively.

### 3.2.3. Results

#### Results (Scenario1)

The results of the inversion application on 400 level2 levelling realizations are summarized in Figure 3.6. The figure shows the posterior standard deviation of the model parameter estimates from all 400 realizations. The results clearly show, as expected, that the precision of the estimated parameters increases if the full covariance matrix is taken into account. The same simulation was also applied on the same data in other output levels. The results for output levels 2, 3, 4 and 5 are comparatively visualized in Figures 3.7 and 3.8 for the time decay parameter and the  $C_m$  proportionality factor respectively. We can see two effects here:

- If the full covariance matrix is taken into account, the precision of the estimated model parameters is identical for all the output levels. This was expected, because the information content of different output levels is identical. By using the full covariance matrix, we always exploit the full information content from each output level, and consequently there is no difference in the precision of the estimates based on different output levels.

However, when we do not use the full covariance matrix (e.g., in practice the precise stochastic model may be not available, or not be completely/correctly known), we compromise the precision of the estimates more or less severely. This effect can be mitigated by using multiple reference points/epochs. This is an important conclusion. More generically this implies that by simply transforming the data to an output level with multiple reference points or epochs, we reduce the sensitivity of the model parameters to the use of covariance information in the modelling.

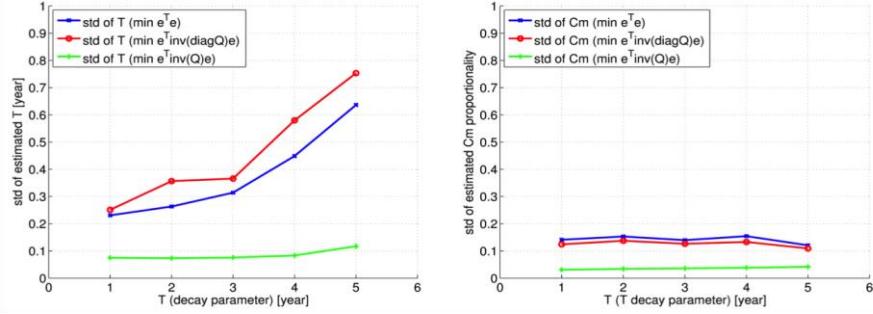


Figure 3.6: The results of the inversion application (scenario 1) on 400 level2 dataset realizations with different time decay parameters (1 to 5 years) with three different objective functions, Left) the posterior standard deviation of time decay estimates, Right) the posterior standard deviation of  $C_m$  proportionality factor estimates. The precision of the estimated parameters improves significantly if the full covariance matrix is taken into account.

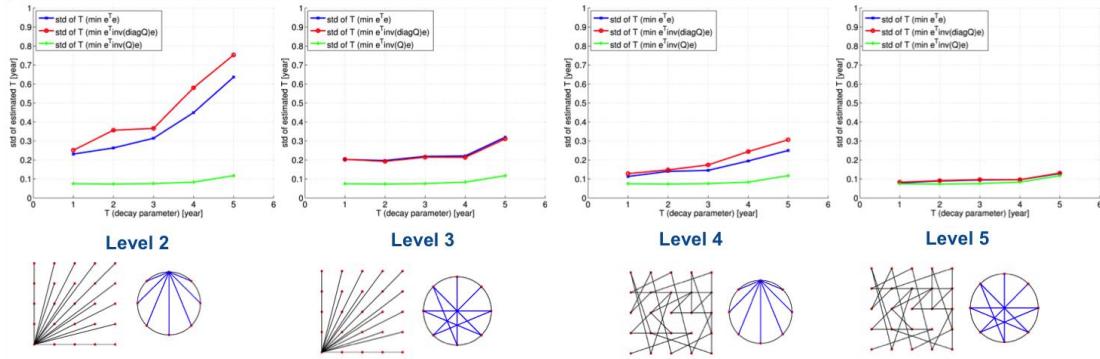


Figure 3.7: The results of the inversion application on 400 dataset realizations of levels 2 to 5 for the time decay parameter (scenario 1). The figure shows the posterior standard deviation of the 400 estimates.

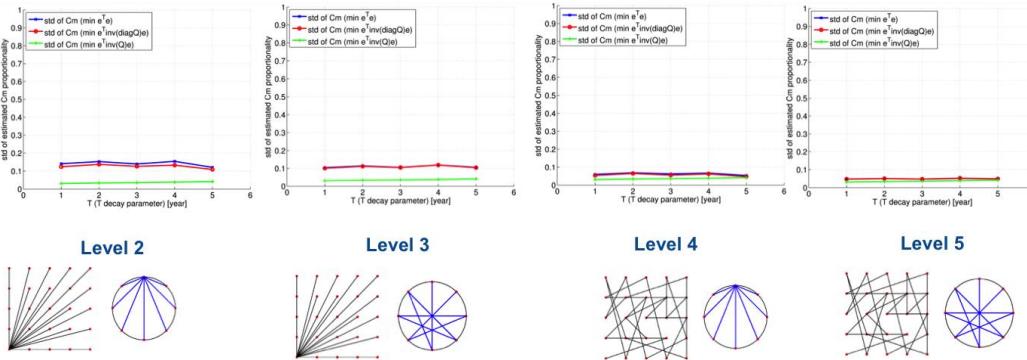


Figure 3.8: The results of the inversion application on 400 dataset realizations of levels 2 to 5 for the  $C_m$  proportionality factor (scenario 1). The figure shows the posterior standard deviation of the 400 estimates.

### Results (Scenario2: low signal to noise ratio)

The results of the inversion of level2 data for this scenario are presented in Figure 3.9. The same as the first scenario, the precision of the estimated parameters improves significantly if the full covariance matrix is taken into account. Note that

while the noise significance for both scenarios 1 and 2 are identical, the pressure change rate (and so the maximum deformation) is ten times smaller for the 2<sup>nd</sup> scenario. Consequently the achievable parameter precision is worse for scenario 2 due to the lower signal-to-noise ratio (compare figures 3.6 and 3.9).

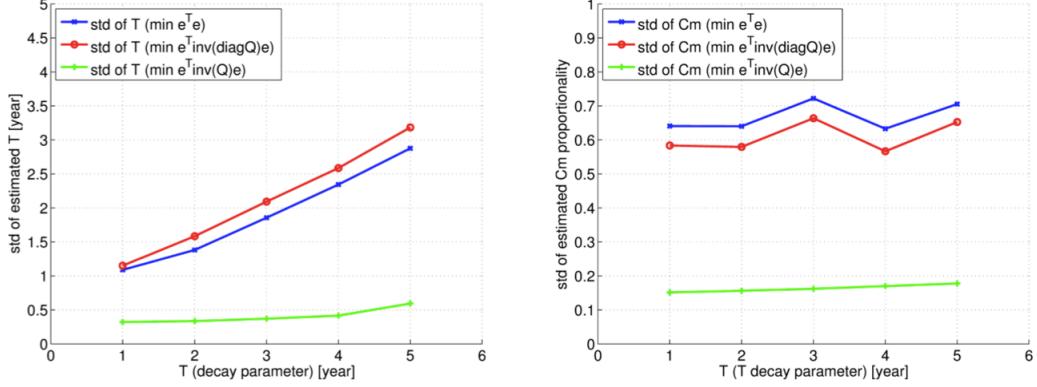


Figure 3.9: The results of the inversion application (scenario 2) on 400 level2 dataset realizations with different time decay parameters (1 to 5 years) with three different objective functions, Left) the posterior standard deviation of time decay estimates, Right) the posterior standard deviation of  $C_m$  proportionality factor estimates. The precision of the estimated parameters improves significantly if the full covariance matrix is taken into account. The achievable parameter precision for scenario 2 is worse than for scenario 1 due to the lower signal-to-noise ratio.

The scenario 2 results for all the output levels are shown in Figures 3.10 and 3.11 for the time decay parameter and the  $C_m$  proportionality factor, respectively. We see the same pattern as in the first scenario but more dominantly in the case of scenario 2. The effect of neglecting stochastic properties in the inversion is much more dominant and can be effectively reduced (or even mitigated) by using the 5<sup>th</sup> output level with multiple reference point/epochs.

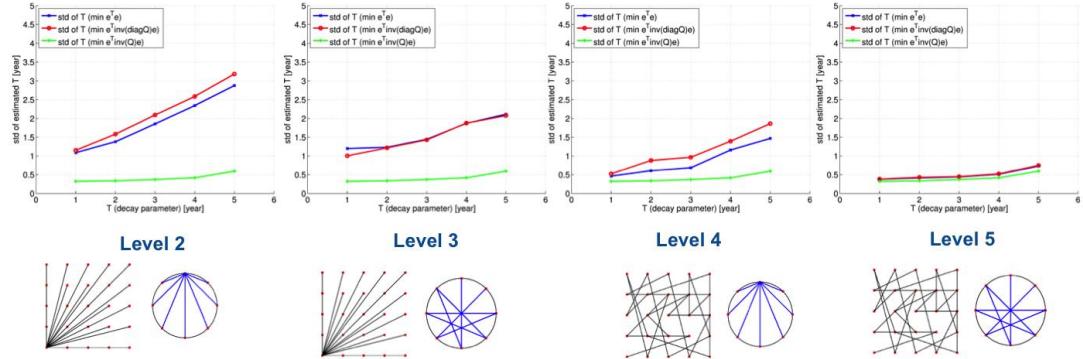


Figure 3.9: The results of the inversion application on 400 dataset realizations of levels 2 to 5 for the time decay parameter (scenario 2). The figure shows the posterior standard deviation of the 400 estimates.

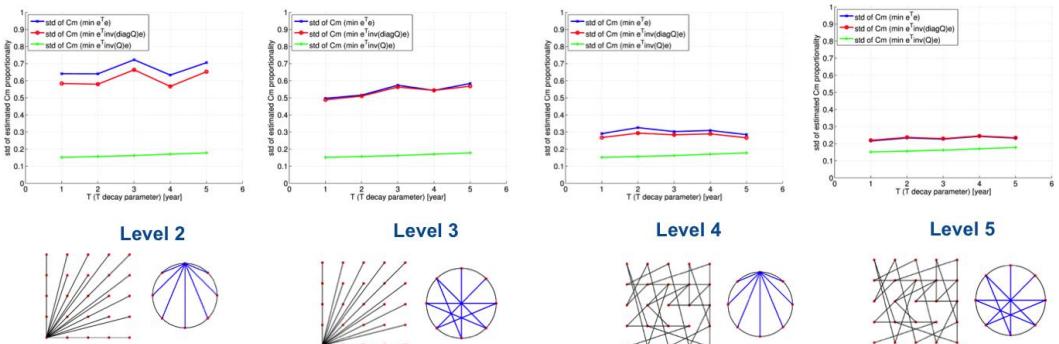


Figure 3.10: The results of the inversion application on 400 dataset realizations of levels 2 to 5 for the  $C_m$  proportionality factor (scenario 2). The figure shows the posterior standard deviation of the 400 estimates.

### Results (Scenario3: early-field life observations)

The results of this scenario for the leveling output level 2 are shown in Figure 3.11. We can see that, for larger time decays ( $T > 2$ ), we are not able to estimate the model parameters precisely. This is due to the fact that for larger time decays, the onset of subsidence is relatively late in the observation period (see Figure 3.5), resulting in a very low signal to noise ratio. We even observe that the precision of the model parameters is not necessarily best when using the most rigorous stochastic model. However, we believe that this observation is an artefact of the limited numerical accuracy of the inversion approach. For the output levels 3,4, and 5, Figures 3.12 and 3.13 suggest the same conclusions as already drawn for scenarios 1 and 2..

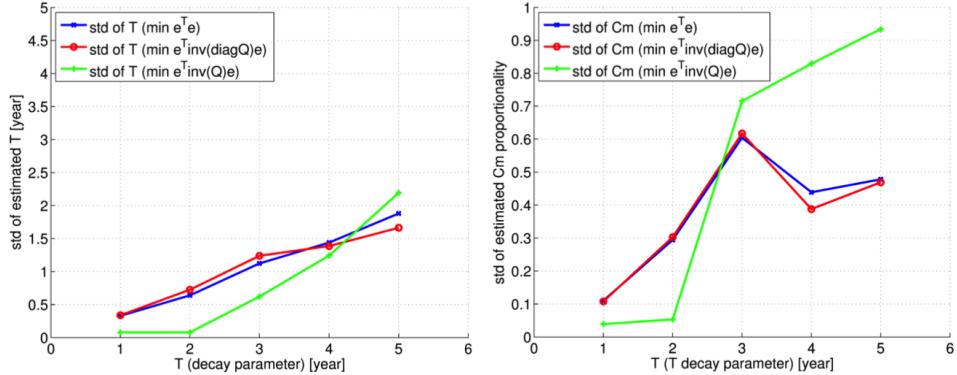


Figure 3.11: The results of the inversion application (scenario 3) on 400 level2 dataset realizations with different time decay parameters (1 to 5 years) with three different objective functions, Left) the posterior standard deviation of time decay estimates, Right) the posterior standard deviation of  $C_m$  proportionality factor estimates. The precision of the estimated parameters improves significantly if the full covariance matrix is taken into account. For larger time decay parameters  $T > 2$ , we are not able to estimate parameters due to very low signal to noise ratio.

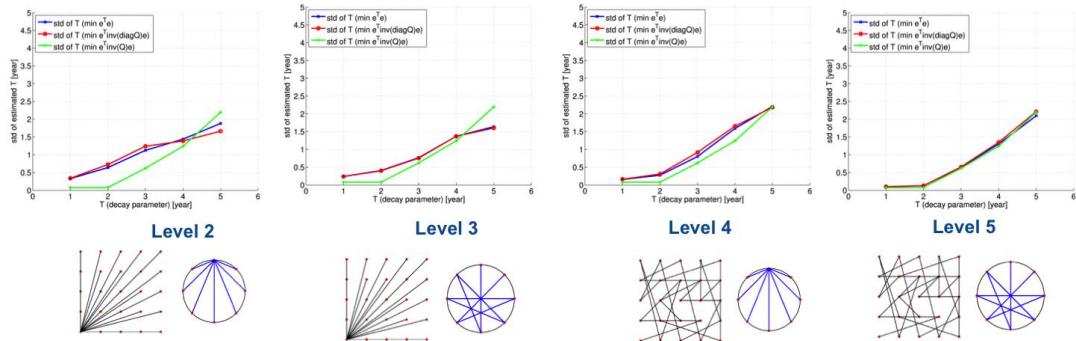


Figure 3.12: The results of the inversion application on 400 dataset realizations of levels 2 to 5 for the time decay parameter (scenario 3). The figure shows the posterior standard deviation of the 400 estimates.

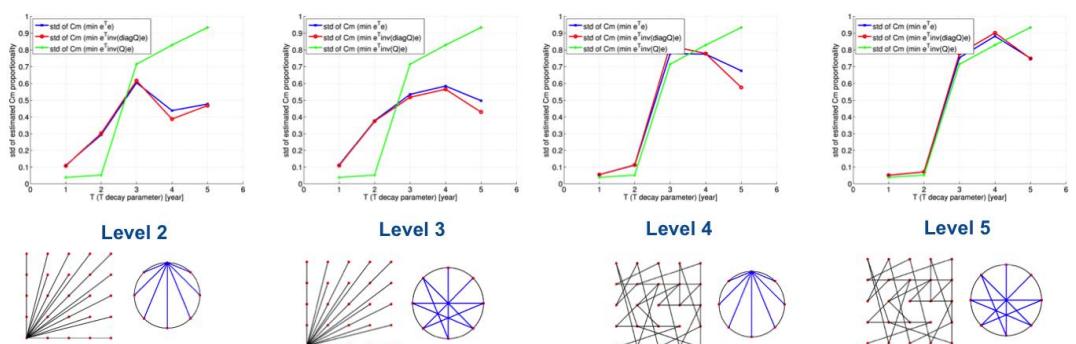


Figure 3.13: The results of the inversion application on 400 dataset realizations of levels 2 to 5 for the  $C_m$  proportionality factor (scenario 3). The figure shows the posterior standard deviation of the 400 estimates.

### 3.2.4. Levelling output level study: discussion and conclusion

#### On the effect of the covariance matrix:

We clearly observed from the leveling output study that using the full covariance matrix in the inversion can significantly improve the modelling precision. We also observed that if the full covariance matrix of observations is used in the inversion there is no difference between the results of the different output levels. Another lesson learned from this experience is that if the covariance matrix is not used, the precision of the estimation can be significantly improved by using multiple reference points and reference epochs. We should note that in practice and for some areas the stochasticity of the data may be not known, and so the covariance matrix may be not available. Therefore we propose to always use (if available) use the full covariance matrix in the modelling. If due to any reason the covariance model is not available or the modeler decided not to use the covariance matrix, it is highly recommended to use multiple reference points and multiple reference epochs in order to mitigate the effect of neglecting the covariance model.

The more insight to the effect of using multiple reference points and epochs can be drawn by looking at the structure of the covariance matrix at different output levels. By using multiple reference point/epochs, we linearly transform the data into new set of observations with the same information content, but with less correlation or with a more homogenous covariance matrix. This is the main reason why the inversion is almost insensitive to the use of the covariance matrix. In other words, this transformation decorrelates the observations to some extent. In an ideal case, if the observations were totally decorrelated, there would be zero sensitivity to covariances, as in fact there is no covariance among data.

Whereas the full decorrelation of observations may be not possible, we showed that by using multiple reference points/epochs, we can effectively (although not completely) decorrelate the data. In order to demonstrate this effect, we visualize a sample covariance matrix for different outputs levels in Figure 3.14. We can see that the structure of the Level5 covariance matrix is most homogenous and least complex: The variances (diagonal elements) are relatively in the same order of magnitude, and the covariance elements are mainly zero or have a homogenous pattern.

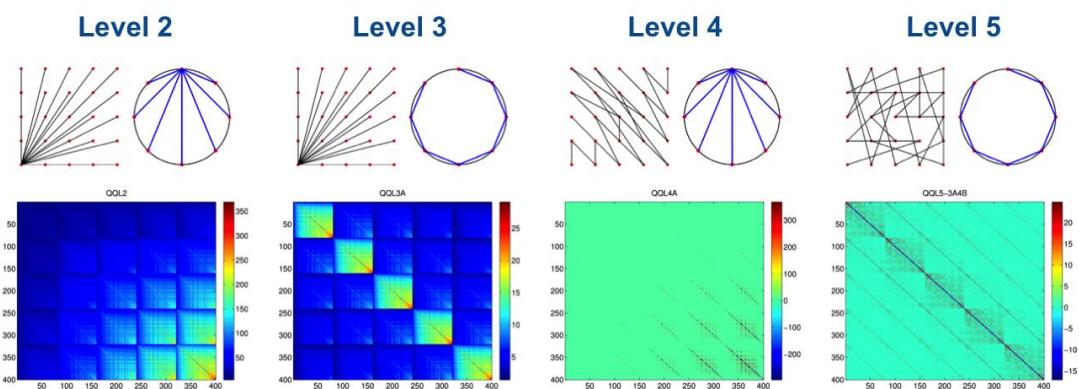


Figure 3.14: The example of the covariance matrix for different outputs levels. The structure of the Level5 covariance matrix is the most homogenous and the least complex

#### Effect on subsidence prediction:

We observed that using the full covariance matrix or using multiple references improves the precision of the estimated geomechanical model parameters. Propagating their precision further shows that likewise also the confidence intervals of subsidence predictions are narrowed. Figures 3.15 and 3.16 demonstrate this

effect. We are using the setting of simulation scenarios 1 and 3 with a time decay of 2 years and predict the subsidence for the years after the observation period based on the inverted parameters from Level2 and Level5. We propagate the full empirical covariance matrix of estimated parameters to the subsidence predication curve and plot the 95% (2 sigma) confidence intervals. We can see in the both scenarios that the increased parameter precision due to consideration of covariances (or due to using an output level with multiple reference points/epochs) can substantially enhance the precision of subsidence predictions.

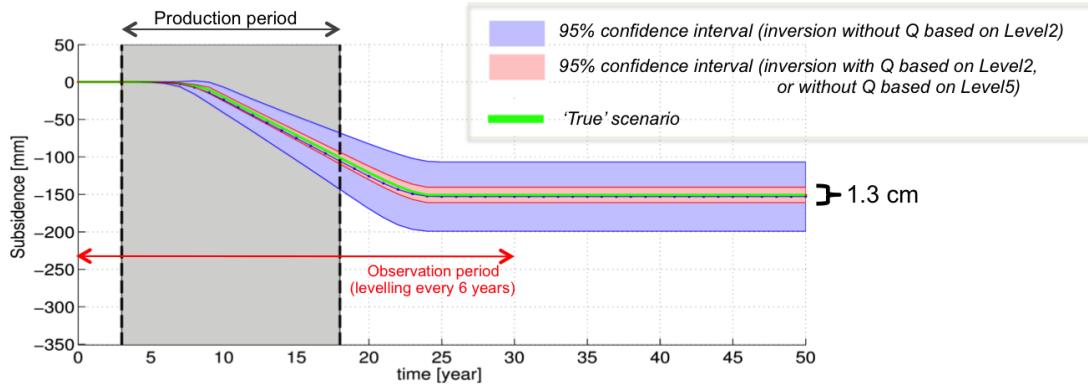


Figure 3.15: Effect of the precision of geomechanical inversion on the confidence interval of subsidence prediction. The prediction is based on the scenario 1 with a time decay of 2 years, based on the inverted parameters from Level2 and Level5. Increased inversion precision can substantially enhance the precision of subsidence predictions or narrow their confidence interval, respectively. Note that the visualization assumes normally distributed predictions, which is not fully rigorous.

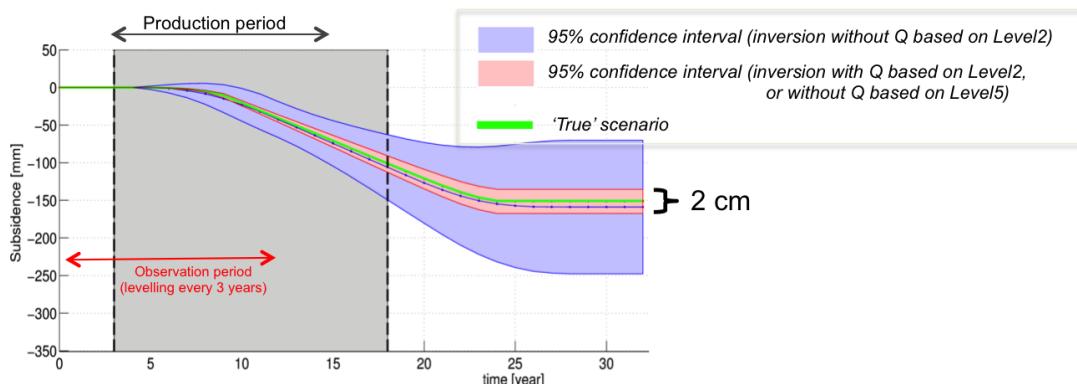


Figure 3.16: Effect of the precision of geomechanical inversion on the confidence interval of subsidence prediction. The prediction is based on the scenarios 2 with time decay of 2 years, and for 20 years after observation period based on the inverted parameters from Level2 and Level5. Increased inversion precision can substantially enhance the precision (or confidence interval) of subsidence predictions.

### On the effect of the zero mean property of noise components:

As discussed in section 2.3.1 the fulfillment of the zero-mean property of double differences that is implied by the stochastic model for idealisation noise depends on the choice of the reference point(s). This has implications for geomechanical modelling. In the output level study, we have simulated the shallow motion of the reference benchmark in order to take into account its effect on the geomechanical model parameters, and we did not observe any significant bias in the inversion based on different output levels. However we should note that in contrast to this output level study, we have only one noise realization in practice, and if we then rely on a single, arbitrarily chosen reference point (Levels 2+3), we risk a bias if this point does not reflect by chance the mean shallow motion of all points. By using

multiple reference points (Levels 4+5), we can sufficiently mitigate any potential bias due to non-fulfillment of the zero-mean property. Against this background and also considering the effect of the covariance matrix, we propose to always use multiple reference points in the geomechanical modelling.

**Conclusions:**

The main conclusion of the leveling output level study can be summarized as:

1. Using the full covariance matrix significantly enhances the precision of estimated geomechanical model parameters and can substantially improve the precision of subsidence predictions.
2. Weighting of individual observations yields no significant improvement with respect to unweighted estimation if covariances are neglected.
3. Using multiple reference points and reference epochs (Level 5) can significantly mitigate the effect of neglecting the covariance matrix in the modelling.
4. Using multiple reference points also reduces the risk of biased estimates due to non-fulfillment of the zero-mean property of double differences.
5. By using Level 5 data, the structure of the covariance matrix of observations will be much simpler. Its effect on the inversion may be even insignificant. For geomechanical modelling, it is recommended not to use a common reference in time and space.

### 3.3. Output level Study: InSAR

#### 3.3.1. Output level overview

The overview of the output levels for InSAR data is given in table 3.2. The schematic summary of these levels is given in Figure 3.17. The output levels 0 and 0w are not considered in the output level study because of the considerable complexity of geomechanical modelling with respect to these levels.

Level	Description	Signal/Noise components	How to be computed (Required a priori knowledge (APK) in the processing)
Level0w	Single-master wrapped (modulo- $2\pi$ ) interferometric phase time series per PS  (w.r.t. one reference point and one reference image )	Line of sight (LOS) deformation (projection of superposition of different deformation mechanism on the satellite LOS), relative heights, atmospheric delays, thermal and scattering noise	Output of InSAR processing
Level0	Single-master unwrapped interferometric phase time series per PS  (w.r.t. one reference point and one reference image )	LOS deformation (projection of superposition of different deformation mechanisms on the satellite LOS), relative heights, atmospheric delays, thermal and scattering noise	Output of spatio-temporal unwrapping (Requires APK on spatio-temporal smoothness of deformation signal)
Level1A	Single-master unwrapped interferometric phase time series per PS (height estimated and excluded)  (w.r.t. one reference point and one reference image )	LOS deformation (projection of superposition of different deformation mechanisms on the satellite LOS), atmospheric delays, thermal and scattering noise	Output of spatio-temporal unwrapping and (residual) height estimation (Requires APK on spatio-temporal smoothness of deformation signal, topographic variability)
Level2A	Single-master LOS deformation time series per PS, excluding atmospheric effect and topography  (w.r.t. one reference point and one reference image )	LOS deformation (superposition of different deformation signals on the satellite LOS), thermal and scattering noise	Spatio-temporal unwrapping + height estimation + atmosphere filtering (Requires APK on, topographic variability , and spatio-temporal behavior of both deformation and atmospheric effect )
Level3A	Single-master <b>vertical</b> deformation time series per PS, excluding atmospheric effect and topography, and <u>converted to vertical assuming zero horizontal deformation</u>  (w.r.t. one reference point and one reference image )	Vertical deformation (superposition of different deformation mechanisms), thermal and scattering noise	LOS to vertical conversion (Requires APK that horizontal deformation is insignificant).
Level1B	Same as level1A but transformed w.r.t. multiple reference images and multiple reference points	Same as Level1A	Spatial and temporal differentiation of Level1A
Level2B	Same as level2A but transformed w.r.t. multiple reference images and multiple reference points	Same as Level2A	Spatial and temporal differentiation of Level2A
Level3B	Same as level3A but transformed w.r.t. multiple reference images and multiple reference points	Same as Level3A	Spatial and temporal differentiation of Level3A

Table 3.2: The overview and description of InSAR output levels.

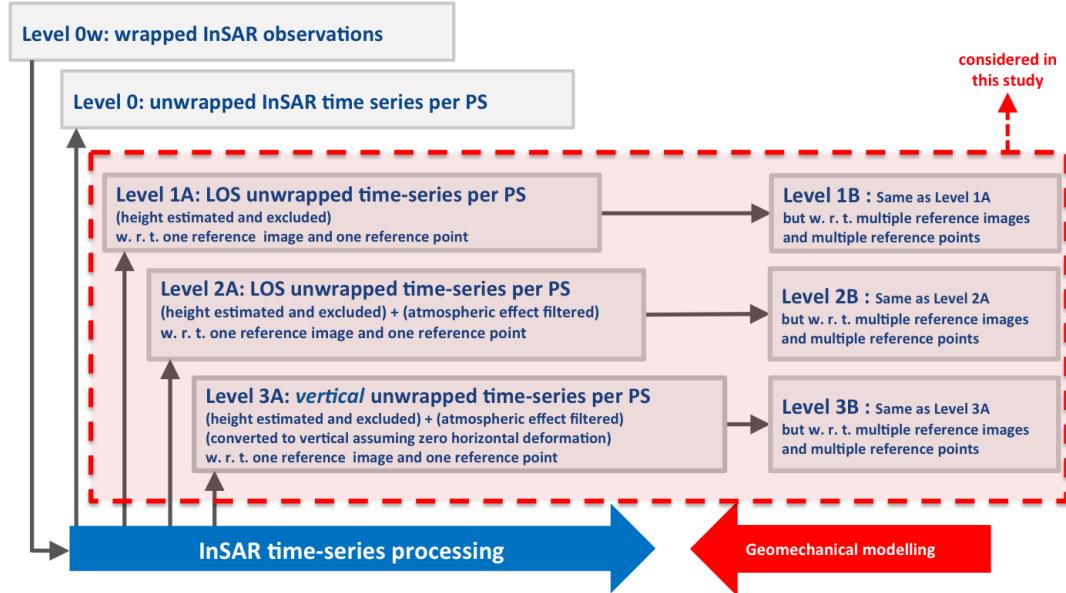


Figure 3.17: The overview of InSAR output levels.

### 3.3.2. Simulation setting

**Reservoir/geomechanical parameters:** We applied the output level study methodology (See section 3.2) on three different scenarios with time decay parameters of 1 to 3 years. All the geomechanical/reservoir parameters are the same for the three scenarios. The parameters used in the simulation are summarized in table 3.3. Similar to leveling output level study, the reservoir/geomechanical parameters are a very rough approximation of the parameters of the Ameland gas field.

Reservoir radius ( $R$ )	5 km		
Reservoir depth ( $D$ )	3 km		
Reservoir Thickness ( $H$ )	100 m		
Poisson ratio	0.3		
Compaction coefficient ( $C_m$ )	$7.4 \cdot 10^{-11} \text{ Pa}^{-1}$		
Time decay parameter ( $T$ )	(scenario 1) 1 years	(scenario 2) 2 years	(scenario 3) 3 years
Start of the production	$t=3$ years		
End of the production	$t=18$ years		
Pressure depletion rate ( $\partial \Delta p / \partial t$ )	$2.0 \cdot 10^6 \text{ Pa/year}$		
Observation period	6.5 years		
Start of the observation period	$t=3$ year (coincident with start of the production)		
Number of SAR images	25 radar images		
Time interval	96 days		
Standard deviation of perpendicular baselines	200 m		
Satellite elevation	800 km		
Incidence angle	23°		
Heading angle	193° (descending orbit)		

Table 3.3: parameters used in the simulation for the InSAR output level study

A depicted overview of the maximum deformation (center of the subsidence bowl) and the observation period of all the scenarios is shown in Figure 3.18. The production period is assumed to be 15 years, starting at  $t=3$  years (with respect to

the modelling reference time). The pressure is assumed to deplete linearly with time during the production period.

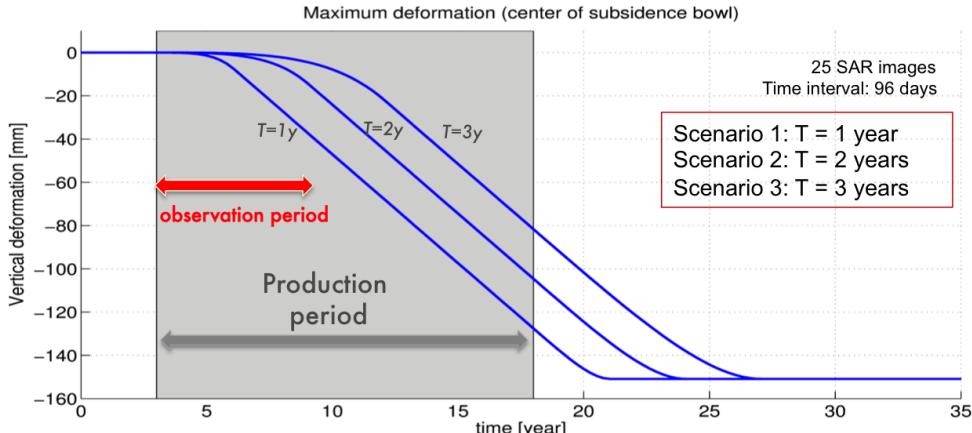


Figure 3.18: The depicted overview of the maximum deformation (at the center of the subsidence bowl) and the observation period of all the scenarios of the InSAR output level study.

**InSAR observations:** The locations of persistent scatterers (PS) are randomly distributed over an area of  $13 \times 13 \text{ km}^2$ . In total 64 1<sup>st</sup> order PS and 200 second order PS are simulated. The 1<sup>st</sup> order PS are used in the processing to estimate and filter the atmospheric effect (the same as DePSI methodology (van Leijen 2014)). An example of the PS locations is given in Figure 2.15. For each simulation 25 radar images with a time interval of 96 days are considered. The steps of InSAR data simulation can be summarized as:

1. Create a network of 1<sup>st</sup> and 2<sup>nd</sup> order PS
2. Construct the full spatio-temporal covariance matrix of InSAR observations for Level1A
3. Simulate observation noise for Level 1A based on the covariance matrix
4. Simulate a 3D deformation signal for Level 1A based on the known geomechanical model, and project the 3D deformation on the radar line of sight (LOS).
5. Compute the observation vector for Level1A (observation = signal + noise)
6. Transform Level 1A data to other levels (linear transformation or atmospheric filtering)
7. Propagate the covariance matrix to other levels using linear or Monte-Carlo error propagation.

The noise components are simulated based on covariance matrices computed from the noise model presented in Sections 2.2 and 2.3. The error propagation for atmospheric filtering is based on the Monte-Carlo approach presented in the Section 2.4. The demonstration of signal simulation for InSAR data is given in Appendix 1.

**Inversion search space:** The inversion is done by searching the two-dimensional search space of the unknown parameters (proportionality factor  $\alpha$ , and time decay parameter  $T$ ). The search space range is [0 2] for  $\alpha$  and [0 12] years for  $T$ . The sampling interval of the search space is 0.02 and 0.1 years for  $\alpha$  and  $T$ , respectively.

### 3.3.3. Results

#### Results (scenario1):

The deformation (signal) time series of all the simulated PS are plotted in Figure 3.19. We can see that the magnitude of the deformation during the observation period (from  $t=3$  till  $t=9$  years) is significant (with a maximum of

$\sim 35$  mm). The results of the inversion of 400 InSAR data realizations are summarized in Figures 3.20-3.24.

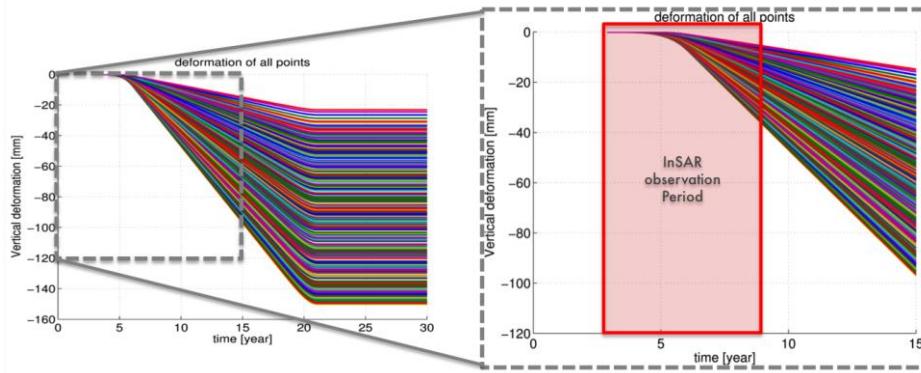


Figure 3.19: The deformation (signal) time series of all the simulated PS for scenario 1.

These plots show the inversion results using three different objective functions as described in section 3.2. The background of the plots shows the normalized empirical probability distribution function of the estimated parameters (i.e., the empirical probability density function (PDF) derived from the 400 estimates). The 1-sigma confidence ellipses of the estimated parameters are also plotted. Note that the wideness of the empirical PDF or the confidence ellipses is representative for the precession/dispersal of the estimated parameters, and the closeness of the PDF mode to the true value is the measure of the biasedness of the inversion.

Figure 3.20 shows the results for the output levels 1A and 2A, i.e. the results before and after atmospheric filtering. We can see that, similar to the leveling study, the precision is increased significantly if the full noise covariance matrices are taken into account in the inversion. The important observation here is that atmospheric filtering does not have a significant effect on the inversion precision when the full covariance matrix is considered in the inversion.

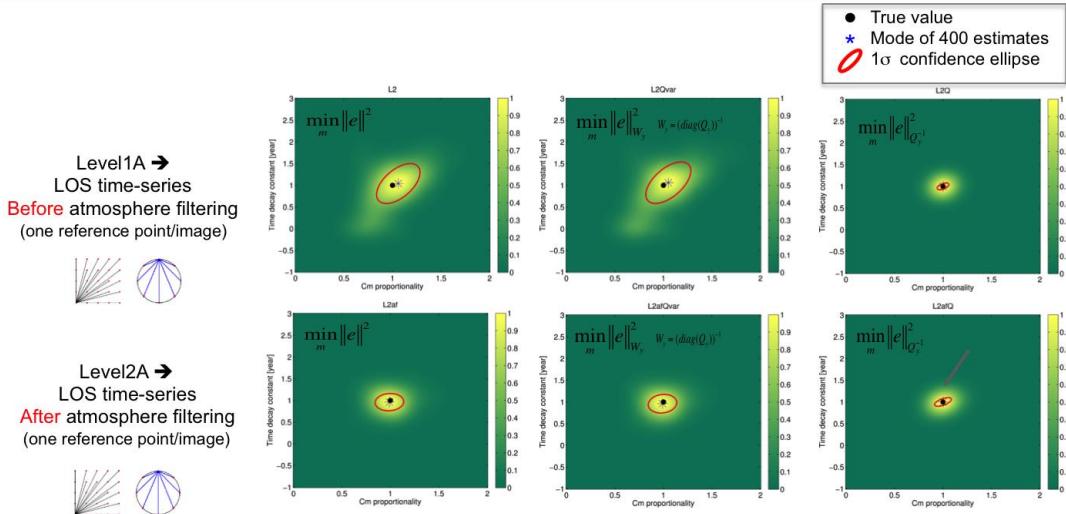


Figure 3.20: the results (scenario 1) for the output levels 1A and 2A using three different objective functions.

The background of the plots shows the normalized empirical probability distribution function of the estimated parameters. Atmospheric filtering does not have a significant effect on the inversion precision when the full covariance matrix is considered in the inversion.

Figure 3.21 compares the results of the output levels 2A and 3A, i.e. the results with and without conversion to vertical. The results demonstrate a significant bias in the inversion due to the neglect of horizontal deformation in the LOS-to-vertical conversion. This is due to the well-known error introduced on InSAR deformation

estimates if we convert them to vertical assuming zero lateral movement (Samiei-Esfahany n.d.).

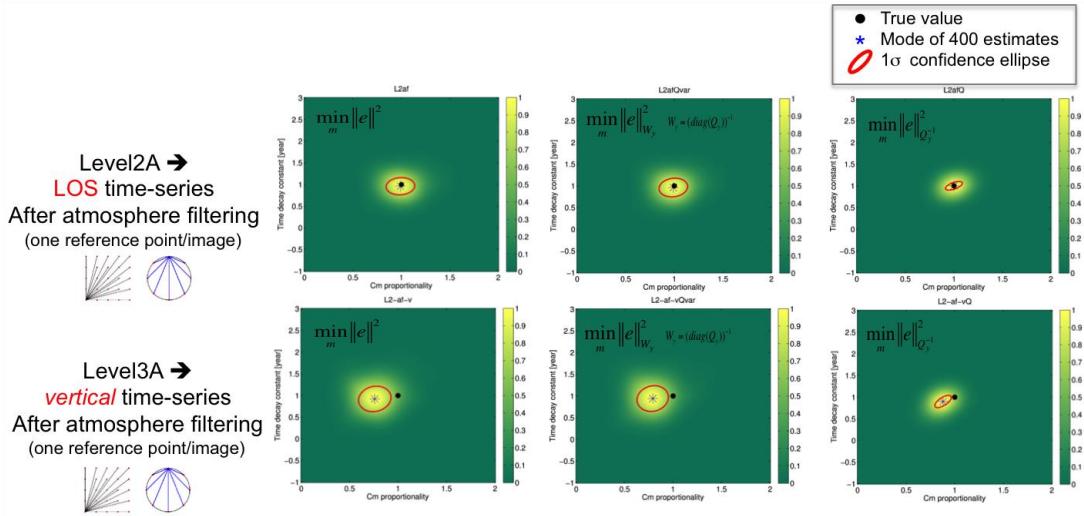


Figure 3.21: the results (scenario1) for the output levels 2A and 3A using three different objective functions. The background of the plots shows the normalized empirical probability distribution function of the estimated parameters. We can see a significant bias in the inversion due to the neglect of horizontal deformation in the LOS-to-vertical conversion.

Figures 3.22 and 3.23 compare the results of output levels with a single reference points/image (1A and 2A) to the results of levels with multiple reference points/images (1B and 2B). By analogy to the output level study on levelling can be concluded that using multiple reference points/images can enhance the precision of model parameters if covariances are neglected.

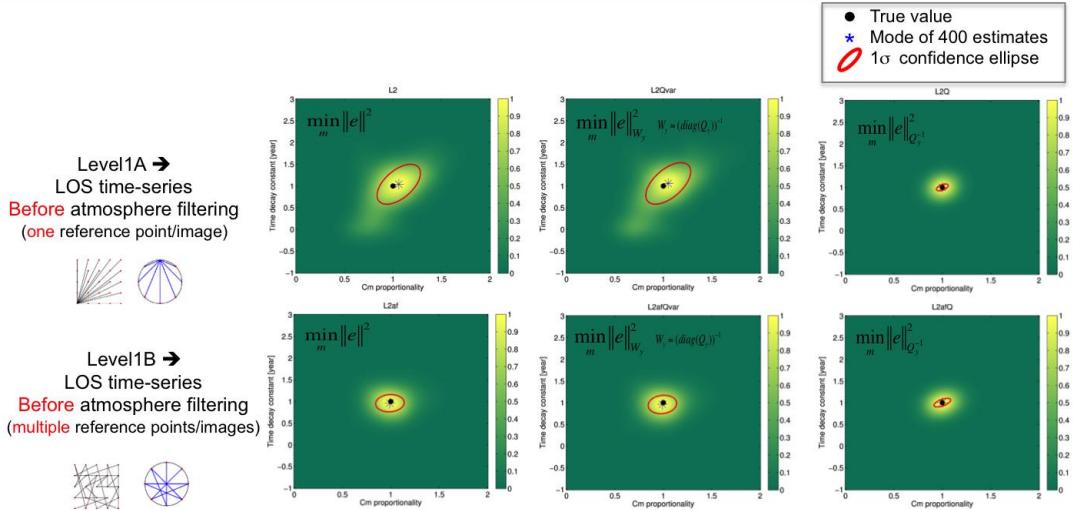


Figure 3.22: the results (scenario1) for the output levels 1A and 1B using three different objective functions. The background of the plots shows the normalized empirical probability distribution function of the estimated parameters. If covariances are neglected, using multiple reference points/images can enhance the precision of model parameters.

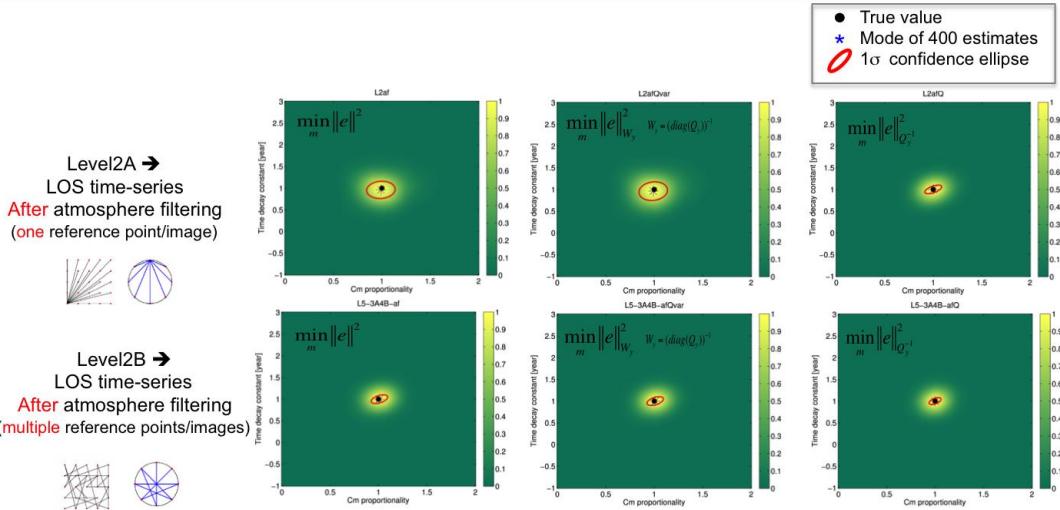


Figure 3.23: the results (scenario1) for the output levels 2A and 2B using three different objective functions.

The background of the plots shows the normalized empirical probability distribution function of the estimated parameters. If covariances are neglected, using multiple reference points/images can enhance the precision of model parameters.

### Results (scenario2):

The deformation (signal) time series of all the simulated PS for the 2<sup>nd</sup> scenario are plotted in Figure 3.24. We can see that, due to the larger time decay parameter, the magnitude of the deformation during the observation period (from t=3 till t=9 years) is much less than the 1<sup>st</sup> scenario, resulting in a much lower signal-to-noise ratio. The results of the inversion of 400 InSAR data realizations in the 2<sup>nd</sup> scenario are summarized in the Figures 3.25-3.28.

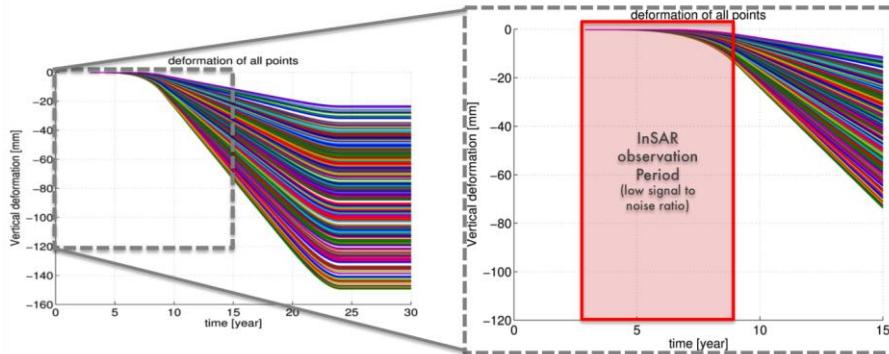


Figure 3.24: The deformation (signal) time series of all the simulated PS for scenario2. The magnitude of the deformation during the observation period is much less than in the 1st scenario, resulting in a much lower signal to noise ratio.

Figure 3.25 shows the results for the output levels 1A and 2A, i.e. the results before and after atmospheric filtering. First of all we can see a significant degradation of parameter precision due to the lower signal-to-noise ratio. As a consequence, we only get reasonable results when the full covariance matrix is used in the inversion,. This outcome underlines the essential importance of using the full covariance matrix for the fields with smaller subsidence (or lower signal-to-noise ratio scenarios). Comparing the dispersion of inverted parameters, we can see that, by analogy to scenario1, atmospheric filtering does not have significant effect on the inversion precision when the full covariance matrix is used in the inversion.

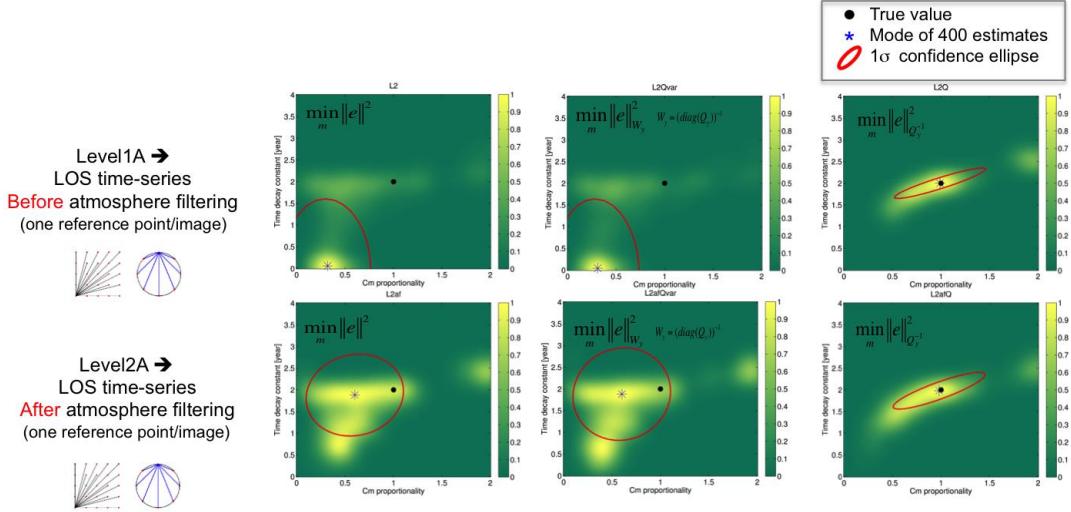


Figure 3.25: the results (scenario2) for the output levels 1A and 2A using three different objective functions.

The background of the plots shows the normalized empirical probability distribution function of the estimated parameters. The parameter precision is significantly worse than in scenario 1 due to the lower signal-to-noise ratio. Only when the full covariance matrix is used, the results are reasonable. Atmospheric filtering does not have a significant effect on the inversion precision when the full covariance matrix is used.

Figure 3.26 compares the results of the output levels 2A and 3A, i.e. the results with and without conversion to vertical. As in scenario 1, we see a significant bias in the inversion due to the neglect of horizontal deformation in the LOS-to-vertical conversion.

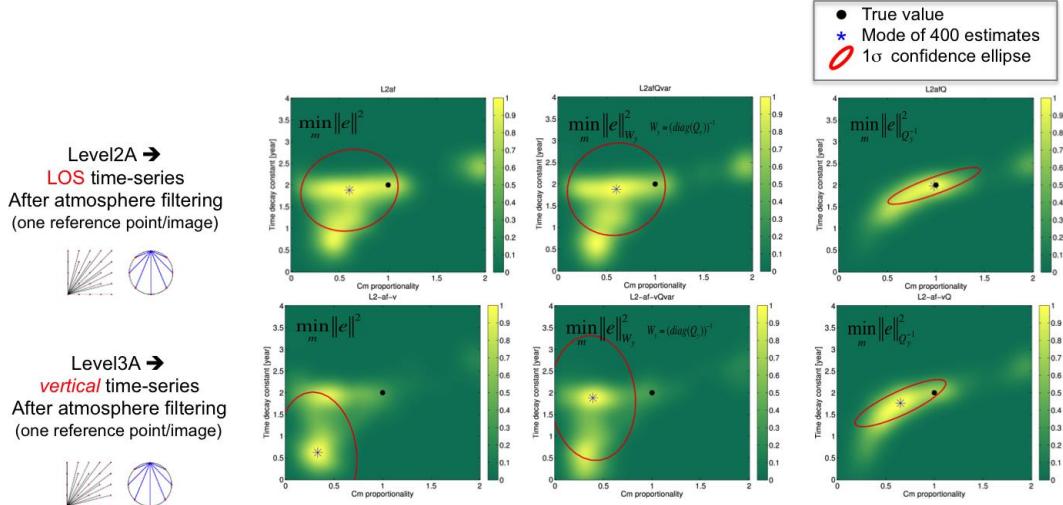


Figure 3.26: the results (scenario2) for the output levels 2A and 3A using three different objective functions.

The background of the plots shows the normalized empirical probability distribution function of the estimated parameters. We can see a significant bias in the inversion due to the neglect of horizontal deformation in the LOS-to-vertical conversion.

Figures 3.27 and 3.28 compare the results of output levels with a single reference points/image (1A and 2A) to the results of levels with multiple reference points/images (1B and 2B). Here, the effect that using multiple reference points/images enhances the precision of model parameters if the covariances are neglected is even more significant than in scenario 1.

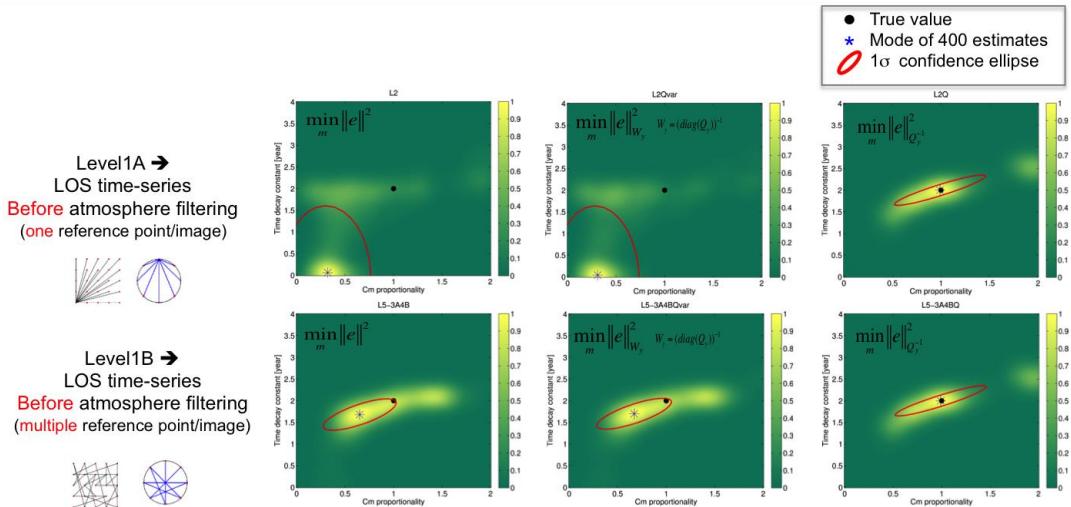


Figure 3.27: the results (scenario2) for the output levels 1A and 1B using three different objective functions.

The background of the plots shows the normalized empirical probability distribution function of the estimated parameters. If covariances are neglected, using multiple reference points/images can enhance the precision of model parameters. Estimation is biased if the covariances are neglected.

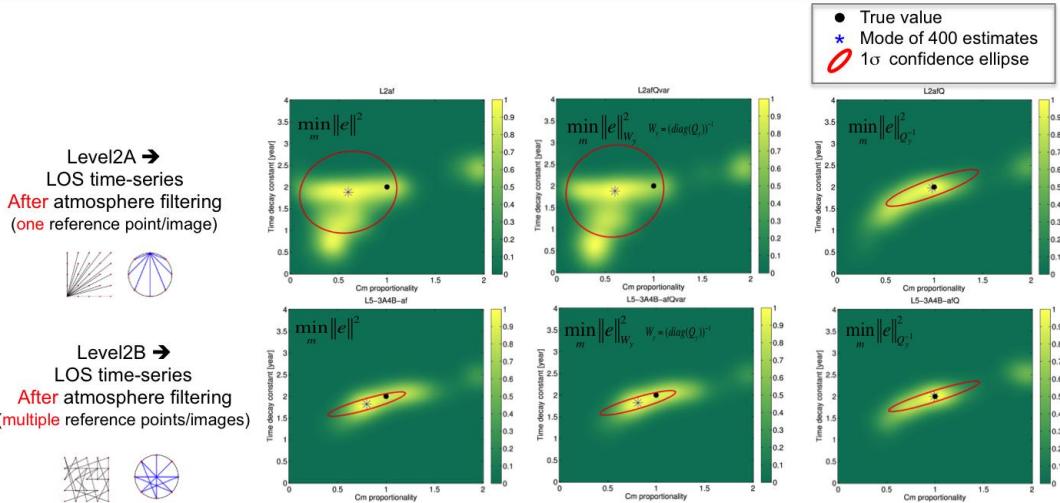


Figure 3.28: the results (scenario2) for the output levels 2A and 2B using three different objective functions.

The background of the plots shows the normalized empirical probability distribution function of the estimated parameters. If covariances are neglected, using multiple reference points/images can enhance the precision of model parameters. Estimation is biased if the covariances are neglected.

### Results (scenario3):

The deformation (signal) time series of all the simulated PS for the 3<sup>rd</sup> scenario are plotted in Figure 3.29. We can see that, due to the time decay of 3 years, the magnitude of the deformation during the observation period is very small (maximum ~5mm in 9 years), resulting in a too low signal-to-noise ratio to get any significant (precise and unbiased) results from the inversion.

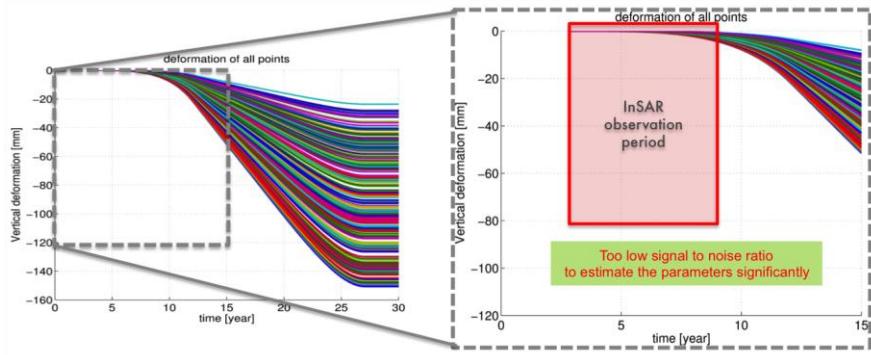


Figure 3.29: The deformation (signal) time series of all the simulated PS for scenario3. The magnitude of the deformation during the observation period is very small, resulting in too low signal- to-noise ratio to estimate the parameters significantly.

### 3.3.4. InSAR output level study: discussion and conclusion

#### On the effect of the covariance matrix:

As in the leveling output level study, we observed the significant importance of using the full covariance matrix in order to get more precise results in the inversion. The results for the 2<sup>nd</sup> scenario with lower signal-to-noise ratio show that this effect is even more significant for the fields with smaller subsidence. We see again that using multiple reference points/images can improve the modelling/prediction precision when information on data stochasticity is not fully available.

#### On the effect of horizontal deformation:

We have observed that the wrong assumption on horizontal deformation in the LOS-to-vertical conversion of InSAR data can introduce a significant bias in the modelling and consequently in subsidence predictions. Although the error introduced by neglecting horizontal deformation in InSAR data is well studied and known in the InSAR community, it is still often neglected for some applications. It is highly recommended, based on this study, to improve the geomechanical modelling methodology by making the inversion from InSAR data in the original LOS geometry.

#### On the effect atmospheric filtering:

All our simulation studies have shown that atmospheric filtering does not have significant effect on the modelling quality if the full (and correct) covariance matrix is used in the inversion. This is an interesting result. The atmospheric filtering is the very standard step in most of the InSAR processing softwares. The filtering always results in more smooth time series, which is more convenient for the user/modeler interpretation. However, atmospheric filtering requires also assumptions and a-priori knowledge about the spatio-temporal behavior of both deformation and atmospheric signal. Wrong assumptions on the temporal smoothness of deformation may introduce error (or bias) in the deformation time series, and consequently in the modelling. Note that in the simulation study, we used correct assumptions, so we do not see any bias in the modelling. However, in practice, this is a likely situation. So regarding the potential bias, it may be beneficial to use unfiltered InSAR data in geomechanical modelling. On the other hand, we should consider that a completely perfect/correct stochastic model for InSAR is not always available. In that case, filtering the atmosphere can improve the results of the modelling (for example see the results in Figure 3.20). It also makes the interpretation of the residuals (misfit between data and the model) easier. So in the end, considering both pros and cons of filtering, we think not removing the atmospheric signal from InSAR data is worth consideration. We recommend trying the inversion with and without atmospheric filtering. In principle the results of the both inversion should not deviate significantly. In case of significant differences in the results, the feedback should be

given to InSAR data providers and the cause of the difference should be further investigated.

### **Conclusions:**

The main conclusions of the InSAR output level study can be summarized as:

1. Conversion to vertical assuming zero horizontal deformation introduces an error in the vertical component of deformation that yields a significant bias in geomechanical modelling.
2. Using the full covariance matrix significantly enhances the precision of the estimated geomechanical model parameters. This effect is very significant for cases with low signal to noise ratio. In these cases, the estimates can even be biased if covariances are neglected.
3. Weighting of individual observations yields no significant improvement with respect to unweighted estimation if covariances are neglected.
4. Using multiple reference points and reference images can significantly mitigate the effect of neglecting the covariance matrix in the modelling and reduce the risk of biased estimates due to non-fulfillment of the zero-mean property of double differences. For geomechanical modelling, it is recommended not to use a common reference in time and space!
5. In the case of using the full covariance matrix in the inversion, atmospheric filtering does not have a significant effect on the modelling quality. Considering that the atmosphere filtering is more demanding with respect to the required a priori knowledge on the deformation signal than accounting for its effect via the stochastic model, we propose to consider not removing the atmospheric signal from InSAR data as the preferred option for geomechanical modelling.



## 4. Outlier Handling

### 4.1. Introduction

Geodetic datasets may include outlier observations, for example due to human or processing errors in levelling campaigns, or due to unwrapping errors in InSAR data. As outliers are described neither by the stochastic model nor by the geomechanical model, they may introduce bias in the final subsidence model. The objective of this section is to propose an efficient approach to detect the potential outliers in geodetic datasets.

Outlier handling is a very well addressed issue in geodesy and geodetic surveying, and there exist different rigorous methodologies to detect and identify outliers in geodetic networks. Most of these methodologies, such as data-snooping (Baarda 1968), are based on iterative hypothesis testing in linear(ized) functional models. These methods involve testing a null hypothesis of a well-known functional relationship between observations and estimated parameters against alternative hypotheses for different types of outliers.

However, there are some obstacles and difficulties associated with the application of such a rigorous approach for geomechanical modelling:

- In geodetic networks or surveying applications, the functional relationship between observations and unknown parameters is often well-defined. In geomechanical modelling, however, these functional relationships are based on deterministic assumptions from geological and reservoir models that are subject to errors. In a rigorous testing framework, we would need to consider these assumptions as observations with associated uncertainties. Therefore any kind of hypothesis testing for outlier detection requires a reasonable stochastic model not only for the geodetic data but also for the subsurface observations, propagated to surface deformation. Such a stochastic model that takes all the model uncertainties into account is not available so far.
- It is also possible to consider a testing framework, in which geodetic observations and geomechanical predictions are compared in the subsidence domain without tracking down model uncertainties to their driving mechanisms. In that case, a simple parametric model can be used to describe uncertainties, which are also referred to as model noise. There are approaches to derive parameters describing the model noise in an iterative manner (e.g. using variance component estimation) from the residuals between the observations and an a priori geomechanical prediction. However,

due to the potentially large imperfections in an a priori geomechanical model, any stochastic model that is estimated from the residuals runs the risk to be substantially biased. So would be any hypothesis testing for outlier detection that is based thereupon.

- Applying geodetic testing concepts is (not necessarily impossible but) usually not straightforward for geomechanical models due to the lacking closed-form functional relationship or due to the highly non-linear models.
- Iterative geodetic outlier identification, such as data snooping, involves time consuming computations. In every iteration, the full estimation/inversion step must be repeated. Such an approach is computationally inefficient for geomechanical inversion, which is usually based on the time consuming forward modelling and parameter-space searching.
- Developing a framework for rigorous hypothesis testing for outlier detection (even if other obstacles can be circumvented) is a lengthy process that keeps us from achieving timely results.

Due to the aforementioned difficulties, we propose to use a pragmatic approach and focus only on the very obvious outliers that can be identified without detailed knowledge about model noise. In the following sections we describe and demonstrate our pragmatic proposal for both levelling and InSAR data.

## 4.2. Outlier handling: levelling data

For the levelling data, we propose two main steps for outlier handling:

1. We detect and remove gross observation errors that are detectable from closed loop conditions within one epoch, using standard geodetic testing methodology (i.e., overall model test and w-test). In this step, we take advantage of the very strong functional model of network adjustment, and also very well studied levelling measurement noise model.
2. We apply hypothesis testing to the individual time series of double differences to detect gross (potentially human) errors that are not detectable from closed loop conditions. These are benchmark misidentifications that create discontinuities in the time series. In this way, instead of computationally expensive testing of a complete spatio-temporal dataset, we test each time series individually, which is a very fast process.

The first step is a very well known, standard, and well addressed experiment in geodetic surveying. So we do not explain it further in this report. First, we elaborate the generic concept of the second step in the next paragraphs, and then in the next subsection, the mathematical formulation of the testing is given.

Given a time series of double differences, we initiate the test by removing the expected signal component due to deep-source deformations. For this purpose, a first order geomechanical model of subsidence can be used to be subtracted from the data. Note that we do not need a perfect model here. As the geomechanical model behaviour is always temporally smooth, any model imperfection can only create temporally correlated components in the residuals. As we only apply the test for gross errors that are often temporally uncorrelated the test is not affected significantly by the model imperfection.

Two types of gross errors are considered in the testing:

1. Identification error: this is a single outlier in the time series, most likely due to benchmark misidentification during one levelling campaign.
2. Disturbance: this is an abrupt jump in the time series, most likely due to repeated benchmark misidentification.

Figure 4.1 shows the stylized concept of disturbances and identification errors.

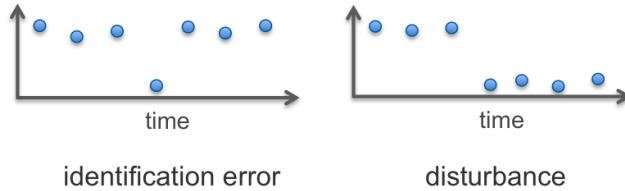


Figure 4.1: the stylized concept of disturbances and identification errors.

The null hypothesis is that the residuals only contain the contributions from measurement and idealisation noise, so the residual behaviour should be consistent with the stochastic model proposed in Chapter 2. We test the null hypothesis against the alternative hypotheses of having gross errors (identification or disturbance) in the residuals with an extremely relaxed test sensitivity. Thus, we prevent the rejection of the null hypothesis due to:

- eventual imperfections of the used geomechanical model,
- neglect of uncertainties of the used geomechanical model,
- uncertainties in the stochastic model for levelling data.

In the next paragraphs, we give a mathematical formulation of the proposed hypothesis testing.

## 4.2.1 Hypothesis testing for gross error detection in double difference time series

**Detection and identification of gross errors:** Assuming  $y$  is a vector of residuals for an individual double difference time series with  $m$  observations, and  $Q_y$  is the covariance matrix of  $y$ , the null hypothesis is defined as:

$$H_0: E\{y\} = 0, \quad D\{y\} = Q_y.$$

This null hypothesis indicates that there is no residual deep source deformation expected (i.e. zero mean expectation), and all the noise contributions together have a dispersion of  $Q_y$ . Three different kinds of alternative hypotheses are considered:

1. Abnormal behavior: There is no gross error, but the residual time series is biased and has an expectation of  $\nabla$ . This is equivalent to the so-called overall model test.

$$H_A: E\{y\} = \nabla, \quad D\{y\} = Q_y$$

2. Identification error: in total,  $m$  alternative hypotheses for an identification error can be defined as

$$H_{li}: E\{y\} = C_{y_i}^T \nabla, \quad D\{y\} = Q_y, \quad i = 1 \dots m,$$

where  $C_{y_i}^T$  is a canonical unit vector as  $C_{y_i} = [0, \dots, 1, \dots, 0]^T$  and  $\nabla$  is a scalar with the value of the identification error.

3. Disturbance: in total,  $m$  alternative hypotheses for a disturbance can be defined as

$$H_{D_i}: E\{y\} = C_{y_i}^T \nabla, \quad D\{y\} = Q_y, \quad i = 1 \dots m$$

where  $C_{y_i}^T$  is a canonical vector as  $C_{y_i} = [0, \dots, 1, 1, \dots, 1]^T$  and  $\nabla$  is a scalar with the value of the identification error.

Considering these three kinds of scenarios, we have in total  $2m-2$  alternative hypotheses:  *$m$  hypotheses for identification errors,  $m-3$  hypotheses for disturbances and one for abnormal behavior*. Note that disturbances after the first or before last epoch would be equivalent with identification errors. So there are  $m-3$  hypotheses for disturbances. We test them against each other and select the most likely scenario based on the B-method of testing (Baarda 1968, Teunissen 2000). As a result we can detect and identify the gross errors in every double difference time series. The next question is what to do with the detected outliers or how to adapt the observation vector to handle the outlier effect.

**Adaptation step:** Based on the outcome of the test, one of the following actions should be taken (See Figure 4.2):

- in case of an identification error, we exclude the corresponding observation from the time series,
- in case of a disturbance, we split the time series at the time of the disturbance,
- and in case of abnormal behaviour, we do not take any action and just flag the time series to have a potentially abnormal behaviour. This hypothesis may be sustained due to imperfections in geomechanical modelling or poorly separable combinations of identification errors and/or disturbances. We flag them just to inform the geomechanical modeller that there is a potential artefact in these particular time series. This information may be useful in the post-analysis and interpretation of the misfit of geomechanical modelling.

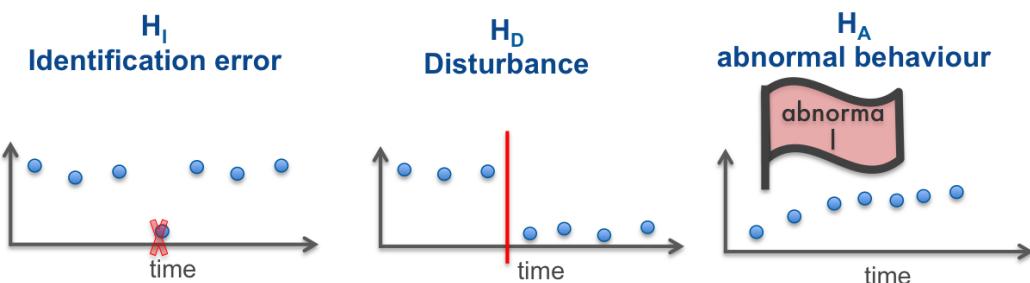
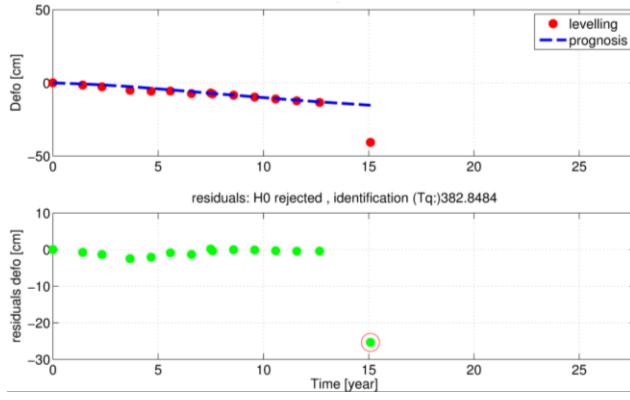


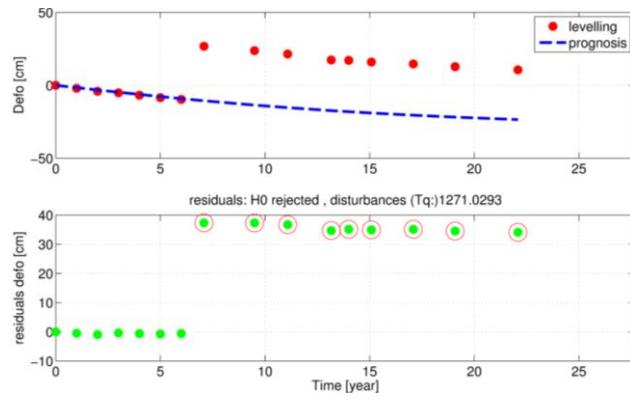
Figure 4.2: The decision for three different kinds of alternative hypothesis: removing the identification error, splitting the time series at disturbance, and flag the abnormal time series.

After detection of all outliers and the corresponding adaptations, the full procedure can be iterated a couple of times with geomechanical modelling until the null hypothesis is accepted. The sensitivity of geomechanical modelling to outliers can be evaluated in the implementation phase by comparatively using two levelling datasets: one complete dataset and one dataset with outliers removed.

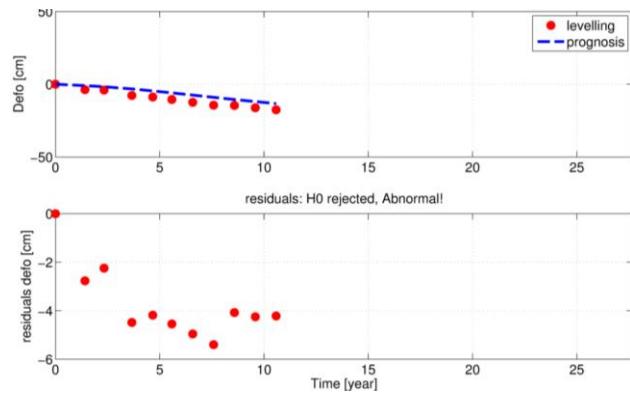
As a demonstration, figures 4.3, 4.4, and 4.5 give three examples of outliers detected by the testing methodology applied on the Ameland levelling dataset.



*Figure 4.3: Arbitrary example of a detected identification error in the Ameland levelling dataset. The red dots show the levelling observations, the blue dashed line is the first order geomechanical model, and the green dots are the residuals (differences between observations and the 1<sup>st</sup> order model). The red circle indicates the detected identification error in the time series.*



*Figure 4.4: Arbitrary example of a detected disturbance in the Ameland levelling dataset. The red dots show the levelling observations, the blue dashed line is the first order geomechanical model, and the green dots are the residuals (differences between observations and the 1<sup>st</sup> order model). The red circles mark all observations after the detected disturbance in the time series.*



*Figure 4.5: Arbitrary example of detected abnormal behavior in the Ameland levelling dataset. The red dots in the upper plot show the levelling observations, the blue dashed line is the first order geomechanical model, and the red dots in the lower plot are the residuals (differences between observations and the 1<sup>st</sup> order model). This time series is flagged to have a potentially abnormal behavior.*

## 4.3. Outlier handling: InSAR data

In InSAR datasets, there are three generic sources of outliers as follows.

1. Incoherent pixels, which are mistakenly detected as PS in the InSAR processing (falsely detected PS)
2. Spontaneous movement of a PS, due to for example very fast movement or collapse of a building.
3. Unwrapping errors in InSAR processing.

Outliers induced by the aforementioned sources can be classified with respect to their representations in time and space.

Part of the outliers are **uncorrelated in space**. This applies to falsely detected PS, spontaneous movement of PS and a subset of unwrapping errors. PS with spatially uncorrelated outliers can be easily identified (and removed) based on their large differences with respect to the neighboring PS in areas with high PS density. Such an identification procedure can be incorporated easily in the InSAR post-processing.

Another class of outliers is **uncorrelated in time**. This applies to temporally isolated unwrapping errors (similar to identification errors in levelling data) and persistent jumps in time series due to unwrapping errors (similar to disturbances in levelling data). To detect these kinds of isolated outliers in time, we propose to apply a similar procedure as proposed for levelling to all the InSAR double-difference time series.

Whereas spatially-uncorrelated outliers in high PS density areas and temporally uncorrelated outliers can be easily identified in InSAR datasets, it is very likely that their effect on the modelling is negligible. Whereas we recommend further exploring this hypothesis, we do not consider the handling of these two kinds of outliers a top priority.

The main challenge in outlier handling is the identification of two other kinds of outliers as follows.

- Falsely detected PS and temporally correlated unwrapping errors in areas with low density of PS: Due to the poor point density, we cannot exploit spatial correlation properties for validation and all PS in that area have to be considered potential outliers. By disregarding them, we may lose valuable information. Thus, it is essential to perform sound consistency checks in the InSAR processing and to quantitatively assess the reliability of PS. Spatially isolated PS that are not deemed sufficiently reliable or verifiable should be flagged and interpreted with care.
- Both spatially and temporally correlated unwrapping errors: These are the most critical kind of outliers or errors in the InSAR datasets and cannot be easily detected or identified. However, the presence of this kind of outliers in InSAR data from the Netherlands is very unlikely. First of all, most of the time series InSAR processing routines apply intensive consistency checks in the unwrapping steps in order to prevent this kind of errors. Second, in time series InSAR methodologies, most of the points are unwrapped independently from their neighboring points. So the likelihood that all points in one particular area have the same unwrapping error is very small. Third, the chance of unwrapping errors is much lower for the InSAR stacks with a large number of images ( $>50$ ) like those that are available from the northern

Netherlands. Fourth, the chance of unwrapping errors is even lower in areas with negligible topography variations as in the northern Netherlands. Based on all these arguments, we conclude that spatio-temporally correlated unwrapping errors are not likely for the application of geomechanical modelling based on InSAR data, and we do not propose any kind of testing for this kind of unlikely scenarios. However we do recommend to raise awareness for this kind of errors in the post-analysis of the geomechanical modelling.

## 4.4. Outlier handling: summary

The main focus of outlier handling is on levelling, in which unavoidable human errors regularly cause huge discrepancies. We proposed a pragmatic approach for outlier removal in levelling datasets. Our proposal can be summarized as:

1. First, apply testing on epoch-wise (single-difference) adjusted observations using a standard geodetic testing methodology
2. Then, apply testing to individual time series of double differences with following considerations:
  - Base the test on time series residuals with respect to a first order geomechanical model.
  - Use the stochastic model proposed in chapter 2.
  - Use an extremely relaxed test sensitivity to identify only most obvious outliers.
  - Consider three kinds of alternative hypotheses as identification error, disturbance, and abnormal behavior
  - Use the B-method of testing to decide for the most likely hypothesis
  - Remove the detected identification errors, split the time series in case of disturbances, and flag the detected abnormal time series
  - Iterate with geomechanical modelling.

For InSAR, outliers are generally a minor issue due to the high spatio-temporal sampling. Nevertheless, thorough reliability checks are essential here.

For both levelling and InSAR, a sensitivity analysis in operational modelling can help quantifying the actual impact of outliers.



## 5. InSAR data reduction

### 5.1. Introduction

Handling large volumes of InSAR data is an important aspect of the geomechanical modelling workflow. An InSAR dataset consists of deformation time series for a large number of radar targets. Contrary to levelling datasets, which typically cover hundreds of benchmarks and up to 10-30 epochs, InSAR datasets often comprise tens of thousands of points with 40-100 epochs and increasing tendency. This has two consequences:

1. Computationally expensive geomechanical modelling,
2. Too large full covariance matrix to work with in practice (Note that we need the matrix inverse in the modelling).

Hence, a proper data reduction technique is required for InSAR data. For data reduction, some assumptions about spatio-temporal variability of the deformation signal are needed. In developing a proper methodology for data reduction, the clarity, simplicity, plausibility, and therefore falsifiability of the used assumption should be considered. In this chapter, we propose a data reduction approach for InSAR data in both the time and space domains with clear and plausible assumptions.

In general, there are two different lines of thought towards data reduction:

1. Methodologies that transform the data to another domain, for example principle component analysis or FFT/wavelet-based data techniques. These are popular in image processing. A good example of such an approach applied to InSAR can be found in (Hetland, et al. 2012).
2. Methodologies that keep the data in the original domain. A popular example of this kind is the quadtree decomposition.

The disadvantage of the first kind of methodologies is that it creates extra complexity in geomechanical modelling and interpretation. In other words, the geomechanical modelling should then be applied in another domain, e.g., wavelet domain, which makes the interpretation complicated. Furthermore, propagation of the full covariance matrices to another domain is challenging and error-prone, especially for nonlinear transformations. As a consequence, our proposal is based

on the second kind of methodologies that keep the deformation data in the deformation domain.

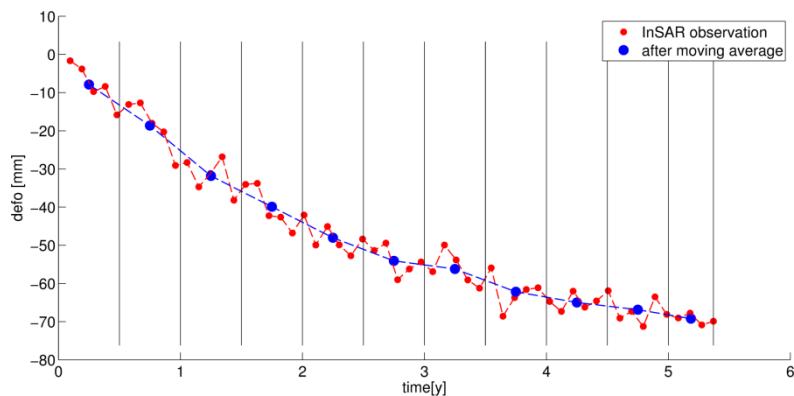
## 5.1. InSAR data reduction in time

For data reduction in time we propose to use “binned averaging” with a fixed time interval (e. g., 6 or 12 months), depending on the characteristics of the subsidence signal. The method consists in subdividing every double difference time series in intervals (“bins”), in which all contained values are averaged. The underlying assumption of this simple approach is (in case of regularly sampled time series) that the deformation rate is constant during the bin time interval. This clear assumption allows the user to simply choose the time interval in which he expects a constant rate.

We should note that the main cause of the big data volume in InSAR data is the large number of persistent scatterers, and the data reduction in the time domain is of secondary importance compared to data reduction in space. So if plausible information about the rate variability is not available, it is recommended not to apply any temporal data reduction. Furthermore, if the temporal evolution of subsidence is of particular interest, the time interval should be chosen conservatively small.

The binned data averaging can be formulated as a linear transformation, allowing the application of linear error propagation to derive the InSAR covariance matrix after data reduction.

Figure 5.1 demonstrates the application of the binned averaging (with six month intervals) on a synthetic InSAR time series.



*Figure 5.1: demonstration of binned averaging with 6 month intervals applied on a synthetic InSAR time series. Blue and red dots are InSAR observations before and after data reduction.*

We should note that for temporal data reduction, it is possible to apply more sophisticated (but also more complicated) alternative approaches. An example is to predict InSAR deformation at temporal sampling points using best linear unbiased prediction (BLUP, also known as least squares collocation), taking the full covariances into account with the same assumption of a constant rate in the time intervals. However, we think such an approach is unnecessarily complex for the application at hand.

## 5.2. InSAR data reduction in space

For spatial data reduction, we consider two potential methodologies: quadtree decomposition and hierarchical K-means clustering. Both of techniques are based on recursive partitioning of point locations into clusters. In the following, we review the two techniques.

### 5.2.1 Quadtree decomposition

Quadtree decomposition is a common compression algorithm in image/signal processing to subdivide an image or a data cloud into sub-blocks. It has been already applied on InSAR data in different studies such as (Jonsson, et al. 2002, Ketelaar 2009). In this procedure, the scene covered by PS points is divided into four quadrants and the mean deformation of each quadrant is calculated. If the variation of deformation in each quadrant (based on a defined criterion) exceeds a given threshold, the quadrant is divided into four new quadrants and again the variation is evaluated. The process continues recursively until convergence. Data reduced by this approach then represent the statistically significant portion of the original data information but with far fewer sampling points.

An important parameter in this quadtree procedure is the criterion to evaluate the significance of signal variation in each quadrant. For application on time series InSAR datasets, we propose the following parameter as a measure of signal variability within a quadrant:

$$T_{\text{quadtree}} = \text{Max} ([\sigma_{q_1} \sigma_{q_2} \dots \sigma_{q_N}]),$$

where  $N$  is the number of SAR images, and  $\sigma_{q_i}$  is the standard deviation of the deformation measurements within the quadrant for the  $i$ th image. An alternative to using standard deviations is to use the range of the observed deformations (i.e., the difference between the maximum and minimum deformation) in each cluster. In that case  $T_{\text{quadtree}}$  is defined as:

$$T_{\text{quadtree}} = \text{Max} ([r_{q_1} r_{q_2} \dots r_{q_N}]),$$

with  $r_{q_i}$  is the range of the observed deformations in the  $i$ th image. We should note that using the deformation range as a measure of consistency is better than the standard deviation, as it is more robust to outliers. This is an advantage that prevents outliers to contribute in the averaging. In addition to  $T_{\text{quadtree}}$ , an extra criterion on the final blocks size can also be added in the procedure. For example the algorithm can be iterated till all the blocks have  $T_{\text{quadtree}}$  smaller than a given threshold and also all the blocks have a size smaller than a given size threshold. In this way, we avoid to have very large blocks. The thresholds for  $T_{\text{quadtree}}$  and block-size can be chosen by the user based on expected spatial variation of deformation.

As a demonstration, Figure 5.2 shows a quadtree decomposition applied to InSAR deformation rates over the Veendam salt mining area (The Netherlands). Note that one disadvantage of quadtree decomposition is that the final location of averaged PS is always the center of a grid cell. Hence, the spatial distribution of the results is defined by the grid geometry rather than the location of PS.

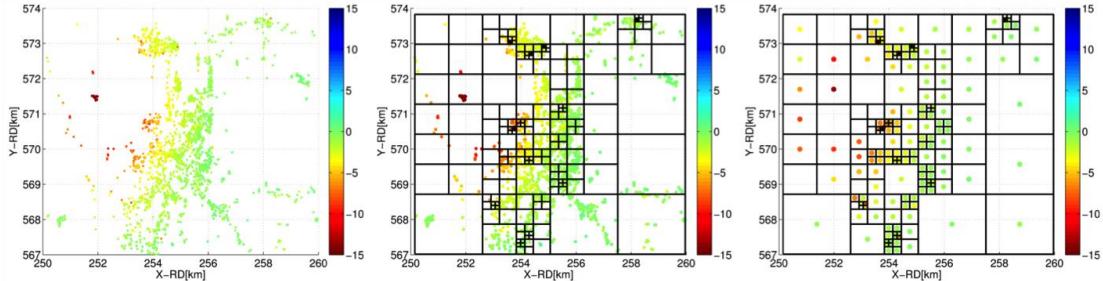


Figure 5.2: Quadtree decomposition procedure applied to InSAR deformation rates over the Veendam area. Left) deformation rates, Middle) quadtree blocks, Right) deformation rates after data reduction.

### 5.2.2 Hierarchical K-means clustering

This approach has been developed based on the concept of “K-means clustering”. We first review first the generic methodology of k-means clustering, and then we describe the specific recursive implementation of the technique, that is “hierarchical” k-means clustering.

**K-means clustering:** Let’s assume a dataset of  $m$  spatial observations  $d_i$  with coordinates  $(x_i, y_i)$ , where  $i = 1 \dots m$ . This technique partitions the  $m$  observations into  $k$  clusters in a way that each observation belongs to the cluster with the nearest mean coordinates. In other words the  $m$  observations are clustered into sets  $S_j$  ( $j = 1 \dots k$ ) in a way to minimise the following objective function:

$$\min_s \sum_{j=1}^k \sum_{d \in S_j} \sqrt{(x_i - \bar{x}_j)^2 + (y_i - \bar{y}_j)^2}$$

where  $\bar{x}_j$  and  $\bar{y}_j$  are the mean coordinates of the points in the  $j$ th cluster. To start the algorithm, the initial positions of the  $k$  clusters are required.

The k-means clustering algorithm can be summarized as:

1. Input:  $m$  spatial observations  $d_i$  with coordinates  $(x_i, y_i)$  where  $i = 1 \dots m$ , and the number of clusters  $k$
2. Initialize:  $k$  cluster center coordinates  $(x_0, y_0)$ , where  $j = 1 \dots k$
3. Iterate:
  - Assign each sample  $d_i$  to its closest cluster center
  - Recompute the new cluster centers as the mean coordinates of the points in each cluster ( $\bar{x}_i$  and  $\bar{y}_i$ )
  - Repeat while not converged
4. Convergence criteria: cluster centers do not change.

#### Hierarchical k-means clustering for InSAR:

This concept is similar to the quadtree decomposition but instead of iteratively dividing the data into quadrants, the data are iteratively portioned into two clusters based on k-means clustering approach with  $k=2$ . The iteration continues until convergence.

In order to apply this approach to InSAR data, we introduce two new convergence criteria:

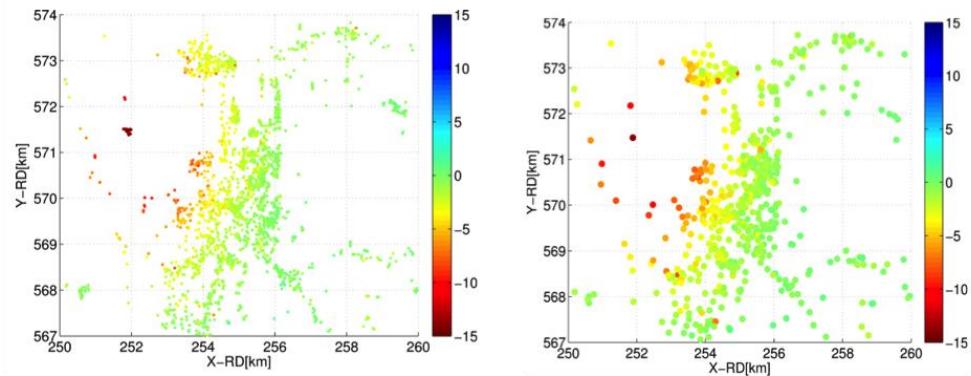
1. The maximum distance between points in all clusters should be smaller than a given threshold. By using a threshold on this criterion, we avoid averaging points that are too far from each other.

2. The variation of deformation in all clusters is beneath a given threshold. To evaluate the variation, the same parameter as in the quadtree decomposition can be used ( $T_{\text{quadtree}}$ ).

The algorithm can be summarized as:

1. The dataset of all PS points is divided first into two clusters using k-means clustering. The initial values of the cluster centers are selected randomly.
2. Then the mean deformation of each cluster is calculated.
3. If the variation of deformation in each cluster based on  $T_{\text{quadtree}}$  exceeds a given threshold or the maximum distance between points in each cluster exceeds a given threshold, the cluster is divided into two new clusters.
4. Steps 2-4 are iterated till all the clusters are smaller than the maximum-size threshold and all the clusters pass the consistency test based on  $T_{\text{quadtree}}$ .

Figure 5.3 shows the results of the hierarchical k-means clustering applied to InSAR deformation rates over the Veendam mining area. The key advantage with respect to quadtree decomposition is that the final locations of the points in the reduced dataset are defined based on the location of existing PS rather than a grid geometry.



*Figure 5.2: hierarchical k-means clustering applied to InSAR deformation rates over the Veendam area. Left) deformation rates before clustering (~3000 points), Right) deformation rates after clustering (~400 points).*

### 5.3. Summary/Conclusion

In this section, we proposed simple methodologies for InSAR data reduction in time and space.

We proposed simple binned averaging for data reduction in the time domain based on the assumption of a constant (deep sourced) deformation rate within the bin time intervals.

For data reduction in the space domain, we considered both quadtree decomposition and hierarchical k-means clustering. Based on the advantages of the latter, we propose to use hierarchical k-means clustering. The main assumption for this data reduction is that the deformation is constant within each cluster. The validity of this assumption depends on the size of the cluster. The user can steer the size of the clusters by choosing a reasonable threshold for the cluster size.

Both proposed approaches for temporal and spatial data reduction can be formulated as linear transformation. This is beneficial as it allows for rigorous linear error propagation.



## 6. Data integration

Geodetic subsidence datasets include data from levelling, InSAR, and GPS. Due to their different properties regarding precision, spatio-temporal sampling rate, continuity and sensitivity to horizontal and vertical deformation components, the measurements from these three techniques have a complementary role in subsidence modelling. For example the high temporal sampling of GPS and InSAR data together with their sensitivity to both horizontal and vertical deformation can complement levelling measurements that have a much higher precision but have a sensitivity only to vertical displacements and have a lower temporal sampling. Using all the available geodetic datasets better constrains the modelling and consequently has a potential to improve the subsidence prediction. How to properly integrate different datasets together with their stochastic models into the modelling process is subject of this chapter.

Conceptually, it is possible to combine the datasets in the observation space prior to geomechanical modelling. As the temporal and spatial sampling of different kind of geodetic data is not the same, this approach obviously requires some sort of interpolation (collocation) in order to evaluate the deformation/subsidence at certain locations and times. Then the interpolated deformations can be directly used for the modelling. The disadvantage of this approach is that interpolation or collocation methods need a-priori knowledge about the spatio-temporal variability of the signal of interest. Wrong assumptions and imperfect a-priori knowledge can result in bias/error in the interpolated deformations, which may be propagated to the biased subsidence modelling. To avoid interpolation artifacts, we propose not to combine different datasets prior to the modelling. Thus, techniques are introduced separately into the modelling, and implicitly combined in the inversion/estimation of the geomechanical model parameters. In this manner, there is no need for interpolation or collocation.

Here, we propose a conceptual mathematical framework for joint inversion of geodetic data. Let's assume that  $y_L$ ,  $y_G$ , and  $y_I$  are the vectors of double differences from levelling, GPS, and InSAR respectively. Each of these sets of observations may have a different datum. The three vectors can be merged to one observation vector as:

$$y = \begin{bmatrix} y_L \\ y_G \\ y_I \end{bmatrix}$$

Then, the dispersion of  $y$  is described by its block covariance matrix as:

$$Q_y = \begin{bmatrix} Q_{y_L} & Q_{y_L y_G} & Q_{y_L y_I} \\ Q_{y_G y_L} & Q_{y_G} & Q_{y_G y_I} \\ Q_{y_I y_L} & Q_{y_I y_G} & Q_{y_I} \end{bmatrix}.$$

The diagonal elements of this block matrix are covariance matrices of levelling, GPS, and InSAR respectively, and the off-diagonal elements contain the cross-covariances between different techniques. Whereas measurement noise components of different techniques are not correlated, idealisation noise components can very well be correlated. So the off-diagonal blocks of  $Q_y$  describe only the correlation due to the idealisation noise.

Having an observation vector  $y$  and its covariance matrix  $Q_y$ , the objective function of geomechanical modelling can be written in general form as:

$$\hat{m} = \operatorname{argmin}_m \|y - G(m)\|_{Q_y^{-1}}^2,$$

where  $m$  are the model parameters, and  $G(m)$  is the function that explains the relation between model parameters and observations.  $\|\cdot\|_{Q_y^{-1}}^2$  is the L2-norm operator as  $\|(.)\|_{Q_y^{-1}}^2 = (.)^T Q_y^{-1} (.)$ .

Let  $F_L(m)$ ,  $F_G(m)$ , and  $F_I(m)$  be the geomechanical functional models which predict the subsidence at locations of levelling benchmarks, GPS stations and InSAR persistent scatterers, respectively. Then,  $G(m)$  can be written as:

$$G(m) = \begin{bmatrix} D_L F_L(m) \\ D_G F_G(m) \\ D_I F_I(m) \end{bmatrix},$$

where  $D_L$ ,  $D_G$ , and  $D_I$  are matrices which transform the geomechanical model prediction to the spatio-temporal structure of levelling, GPS, and InSAR data respectively. For example in case of LOS InSAR observations with multiple reference points and multiple reference images, the matrix  $D_I$  accounts for the projection of the vertical and horizontal subsidence components (i.e. output of  $F_I(m)$ ) on the LOS direction as well as for multiple references by temporal or spatial differentiation. If the datum (or spatio-temporal reference) of the geomechanical model is different than the datum(s) of geodetic observations, the matrices  $D_L$ ,  $D_G$ , and  $D_I$  can transform the model predictions to the datum(s) of the observations.

#### **Discussion on using inter-techniques cross-correlations in the modelling:**

As we discussed before, observations from different techniques can be correlated due to their idealisation noise. Taking into account inter-technique correlations in the proposed mathematical framework would imply two difficulties as follows.

- It increases the computational complexity due to propagation of cross-covariances through InSAR data reduction. Consider that if data reduction is applied on InSAR data, to do the error propagation through this transformation, not only the InSAR covariance matrix should be propagated but also the cross-covariance matrices between InSAR and the other two techniques also should be propagated. Although this is possible, it increases the computational complexity of the error propagation significantly.
- It increases the complexity due to lacking modularisability of the techniques in geomechanical inversion. If we do **not** consider all inter-techniques cross-

correlations, the inversion objective function can be simply modularized (i. e., partitioned) as:

$$\begin{aligned} \text{argmin}_m \|y - G(m)\|_{Q_y^{-1}}^2 = \\ = \text{argmin}_m \left( \|y_L - D_L F_L(m)\|_{Q_{y_L}^{-1}}^2 + \|y_G - D_G F_G(m)\|_{Q_{y_G}^{-1}}^2 + \|y_I - D_I F_I(m)\|_{Q_{y_I}^{-1}}^2 \right). \end{aligned}$$

This modularisability would allow that the L2-norm components can be evaluated separately for each technique.

Due to the aforementioned complications, and also due to difficulties in assessing correctly the cross-covariance matrices, we propose to neglect inter-technique correlations in the geomechanical inversion.

**Summary:** In this chapter, we proposed the conceptual mathematical framework for joint geomechanical inversion of all the geodetic techniques together. This approach does not need any collocation or interpolation of different data sources before the inversion.



## A.1. Appendix 1- Demo: Simulation of deformation for InSAR observations

In this appendix we demonstrate the simulation of the deformation signal for a set of PS over an area of 13x13 km. We assume 25 radar acquisitions with a time interval of 96 days. The parameters used in the simulation are reported in Table 3.3.

The simulated data are visualized here in the order of:

- Figure A1.1: vertical deformation
- Figure A1.2: magnitude of horizontal deformation
- Figure A1.3: magnitude of the east component of horizontal deformation
- Figure A1.4: magnitude of the north component of horizontal deformation
- Figure A1.5: LOS deformation
- Figure A1.6: Converted-to-vertical deformation assuming zero horizontal deformation
- Figure A1.7: Difference between the true vertical signal (Figure A1.1) and the converted one (Figure A1.6). This is the bias due to neglect of horizontal deformation
- Figure A1.8: Double difference LOS deformation with respect to the reference image and the reference point.

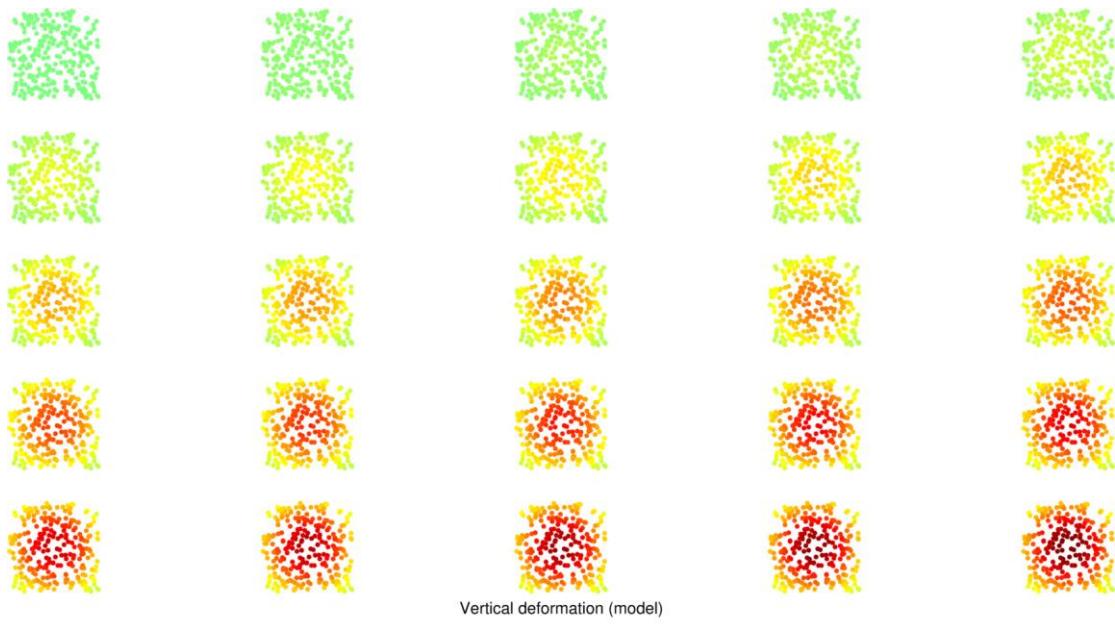


Figure A1.1: vertical deformation. Color-scale is in [mm].

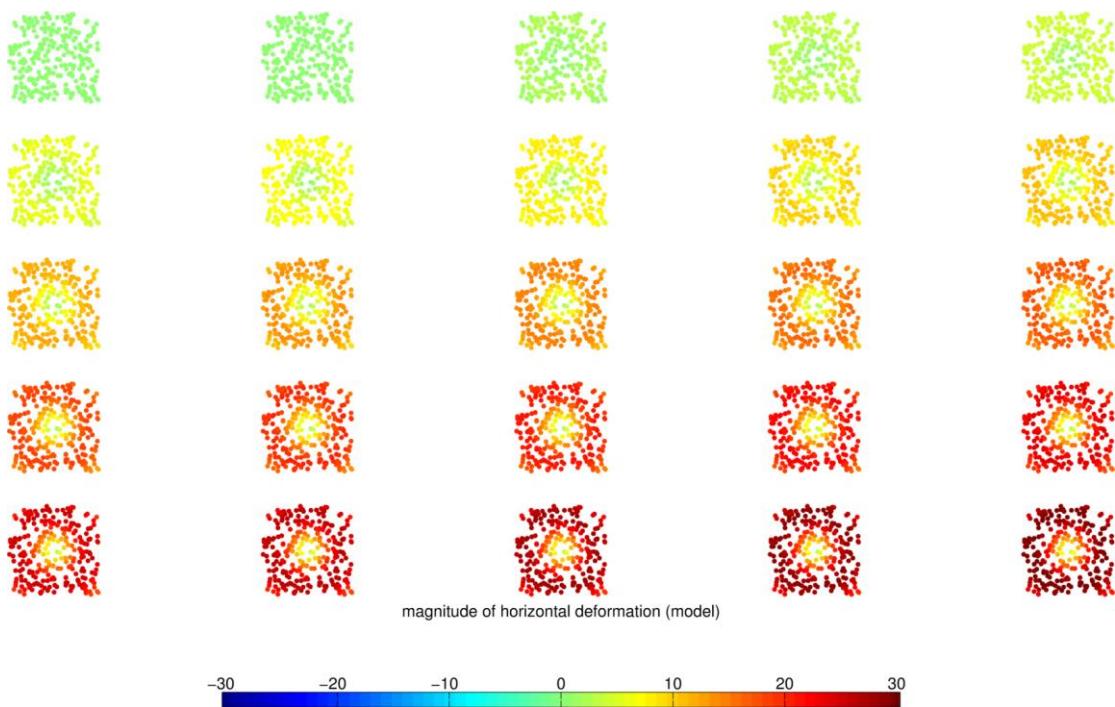


Figure A1.2: magnitude of horizontal deformation Color-scale is in [mm].

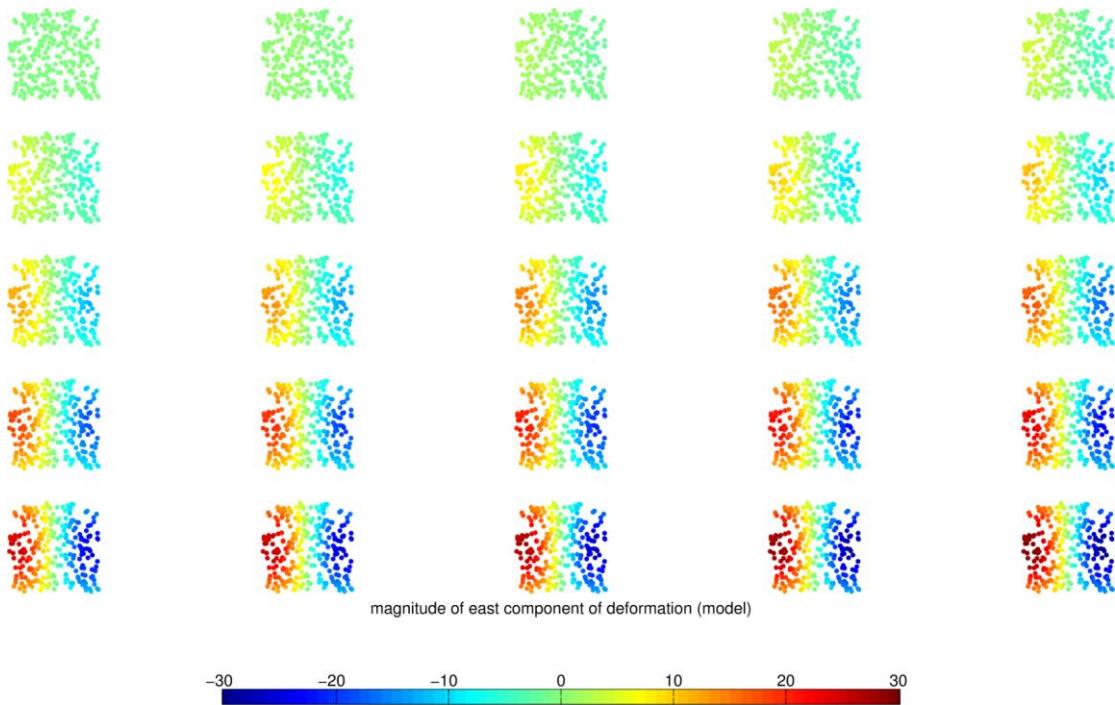


Figure A1.3: magnitude of the east component of horizontal deformation Color-scale is in [mm].

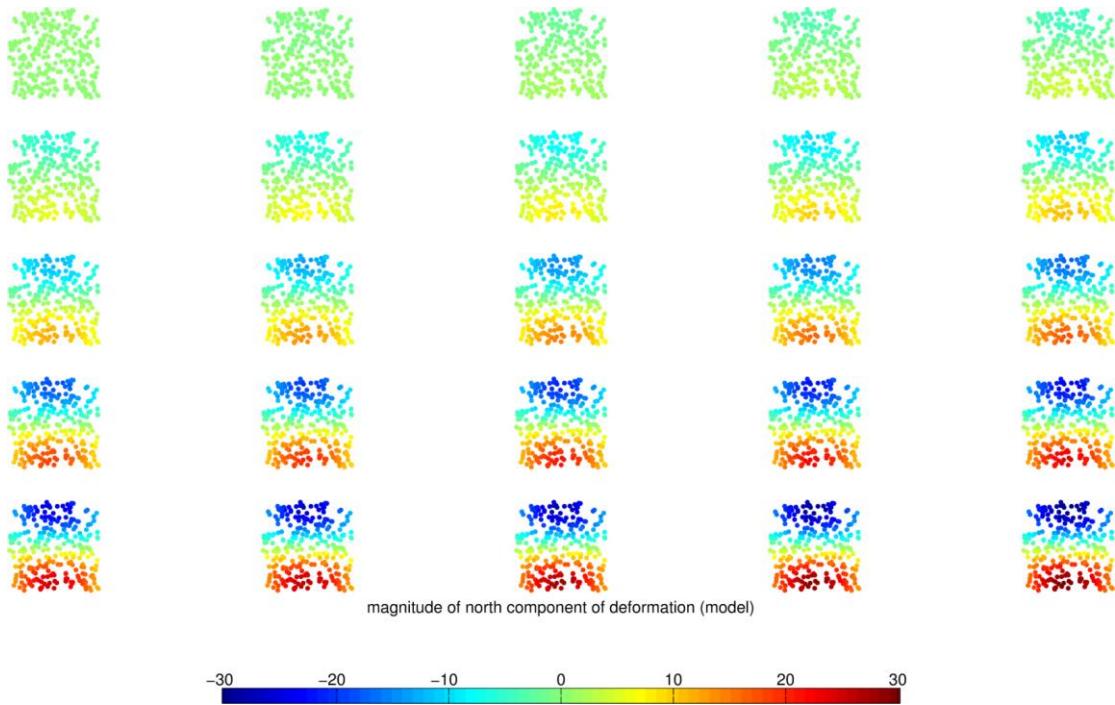


Figure A1.4: magnitude of the north component of horizontal deformation Color-scale is in [mm].

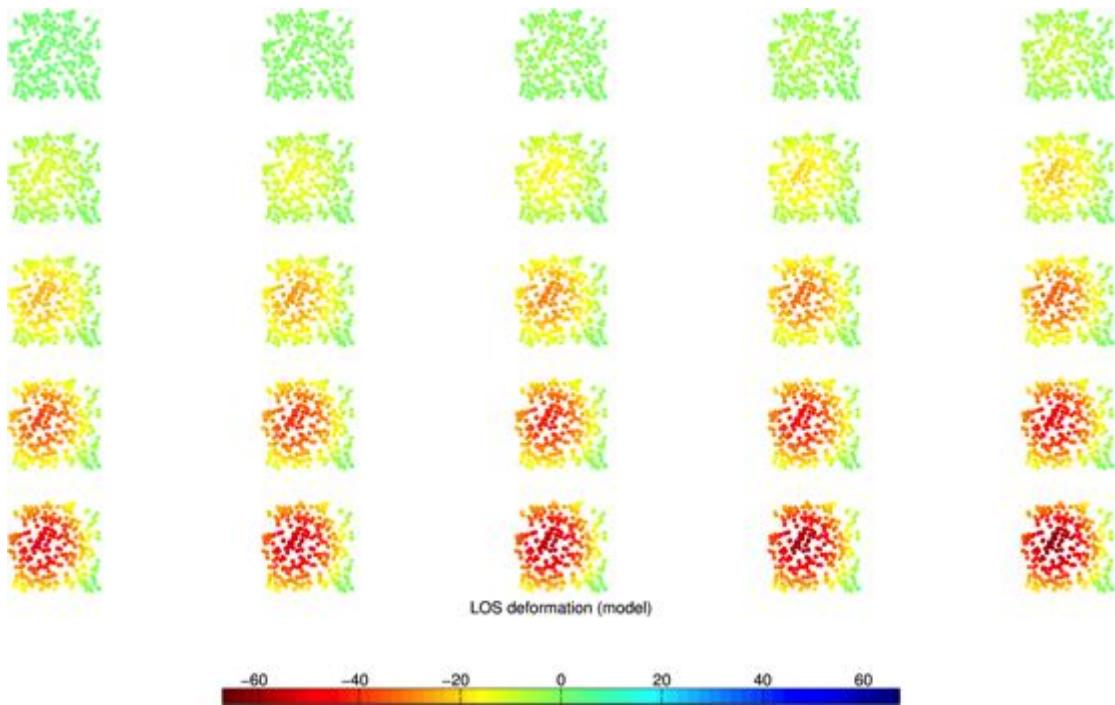


Figure A1.5: LOS deformation. Color-scale is in [mm].

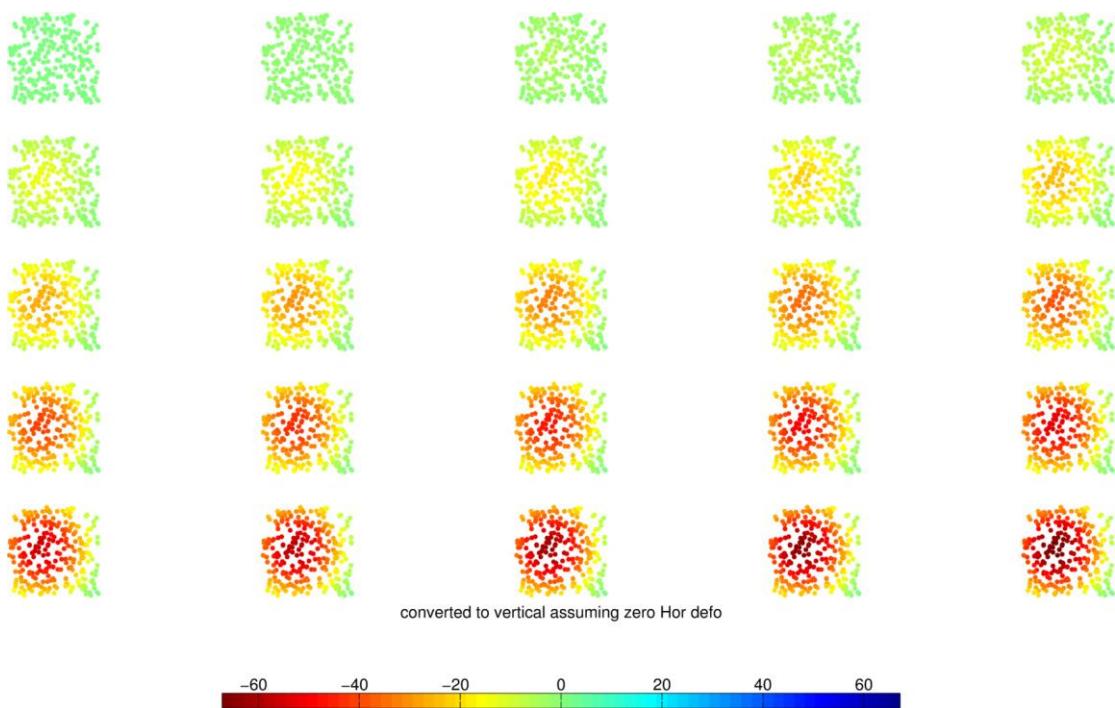


Figure A1.6: Converted-to-vertical deformation assuming zero horizontal deformation Color-scale is in [mm].

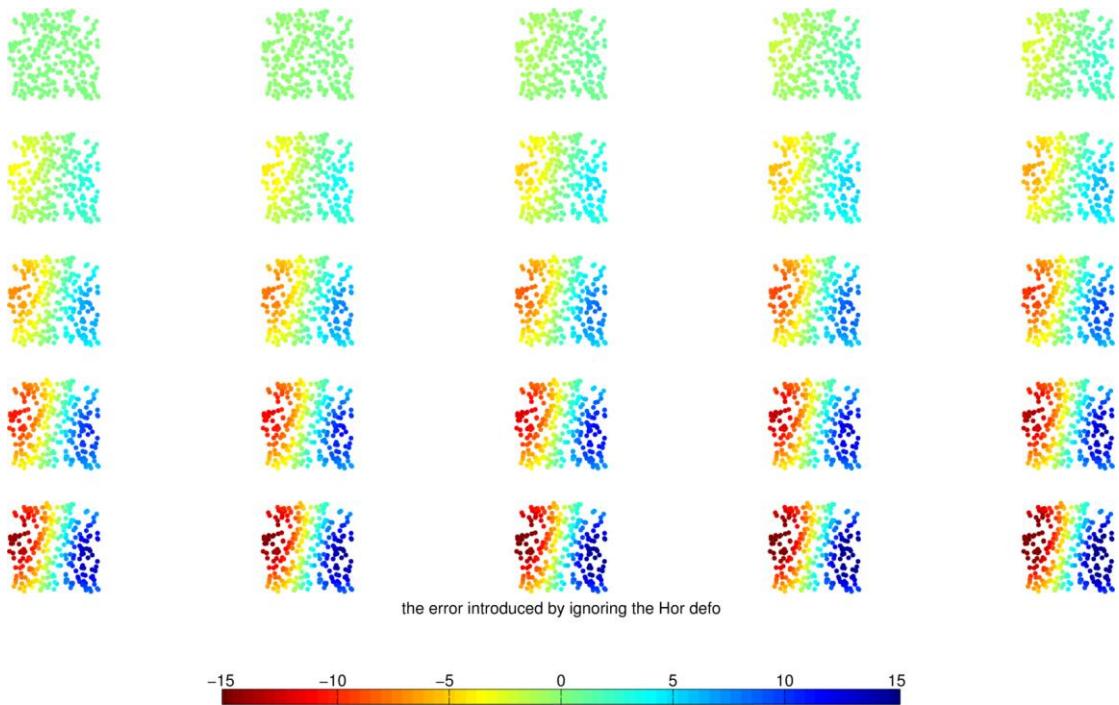


Figure A1.7: Difference between the true vertical signal (Figure A1.1) and the converted one (Figure A1.6). This is the bias due to neglect of horizontal deformation Color-scale is in [mm].

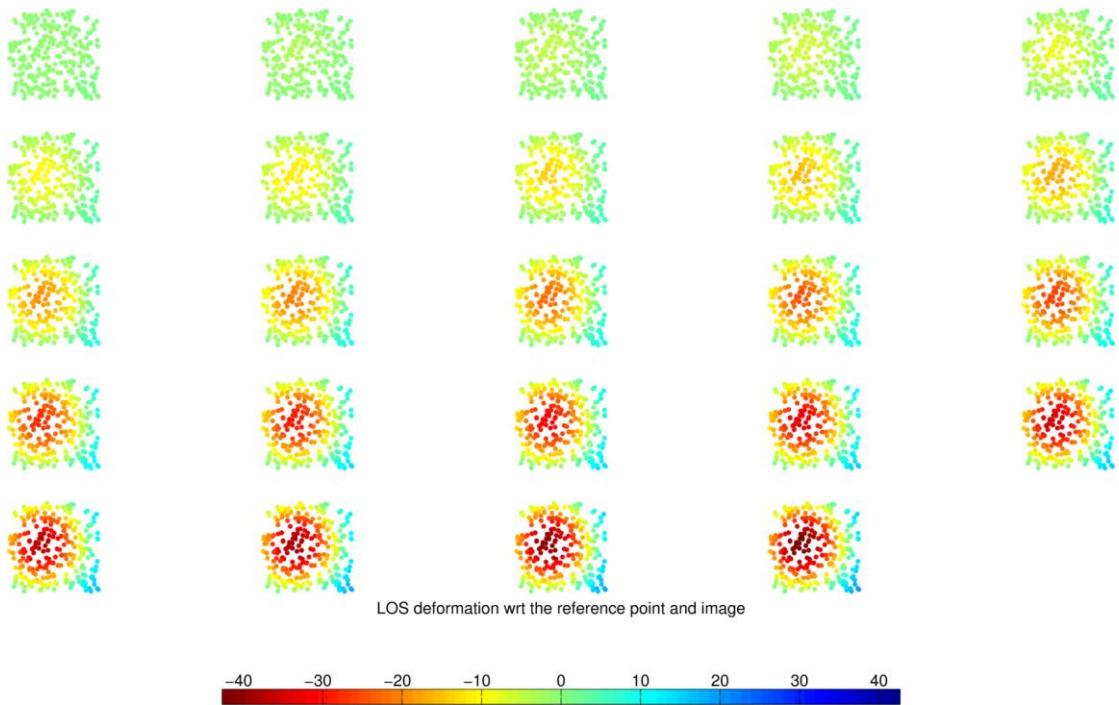


Figure A1.8: Double difference LOS deformation with respect to the reference image and the reference point. Color-scale is in [mm].



## **A.2. Appendix 2- Aspects of the current NAM geodetic data processing workflow**

In this study, we provided recommendations regarding five different aspects of exploitation of geodetic data in subsidence modelling. These aspects are: stochastic modelling, optimal output level, outlier handling, InSAR data volume reduction and data integration. In order to facilitate the appreciation of innovations to the current workflow, this appendix summarizes the current status of the five studied aspects in the geodetic data processing workflow as it is currently implemented at NAM.

### **Stochastic model used in the modelling:**

- Idealisation noise is not taken into account.
- For levelling data, only standard deviations of spatial single differences, propagated to spatio-temporal double differences, are taken into account as a weighting. Any covariances are neglected.
- For InSAR data, no formal quantification of uncertainties is provided at all. InSAR observations are spatially resampled to the positions of levelling benchmarks. Due to variable PS density in the vicinity of the levelling benchmarks, the actual uncertainty is poorly reflected by the implicitly assumed model of uniform weights.

### **Optimal output level:**

- Levelling: Observations are adjusted epoch-wise as a free network with one reference benchmark under consideration of measurement noise (as reported in Section 2.1). The adjusted height differences are propagated to the final output level, which is double differences with respect to a common reference benchmark but multiple reference epochs (i.e., output level 3 in Section 3.2.1). In the modelling, these double differences are interpreted as temporal single differences with a pseudo-absolute reference in space by assuming stability of the reference benchmark.
- InSAR: The output level used in the modelling is single-master vertical deformation time series per PS, excluding atmospheric effect and topography, and converted to vertical assuming zero horizontal deformation (w.r.t. one reference point and but multiple reference epochs). This is equivalent to output level 3 in Section 3.3.1.
- GPS: The output level is temporal single differences with a pseudo-absolute reference in space, manifested by a network of assumedly stable reference stations.

### **Outlier handling:**

- For leveling, only gross errors are detected and removed during the epoch-wise network adjustment based on the closing loop conditions and using standard geodetic testing methodology (i.e., overall model test and w-test). Additionally, spatio-temporal deformation analysis is used to formally identify a class of “potentially unstable” benchmarks. The assessment makes use of a geomechanical model prediction. A “potentially unstable” benchmark has an apparent displacement behaviour that is significantly different from that of surrounding benchmarks and can be explained by either an actual deviation of the displacement or by human errors such as benchmark misidentifications. The modeller decides heuristically if an individual “potentially unstable” benchmark can be trusted or not.
- For InSAR, the PS that show a different deformation pattern with respect to the surrounding PS are detected and removed.

### **InSAR data volume:**

InSAR data are resampled to the locations of levelling benchmarks to create a combined dataset with levelling:

- All PS within a radius of 500 m about a benchmark are attributed to that benchmark.
- InSAR time series per benchmark are generated as the (spatial) median of all individual time series.
- Yearly averaging of InSAR time series is used for data reduction in time.

### **Data integration:**

In areas where only levelling data are available, geomechanical model calibration is based on levelling data only.

Where InSAR data are available, geomechanical model calibration is based on a combined dataset of leveling and InSAR data, sampled at the benchmark locations:

- Time series from different InSAR missions are aligned by offset estimation in overlapping periods.
- Combined InSAR time series are aligned with levelling time series by offset estimation in overlapping periods.

For gas fields close to the Wadden Sea, geomechanical modelling is complementarily based on GPS campaign measurements that are provided in a separate dataset. Note that due to the observation design, which is subject to practical constraints, only the vertical component can be used for geomechanical model calibration.

## Bibliography

Baarda, W. *A testing procedure for use in geodetic networks*. Edited by Netherlands Geodetic Commission. Vol. Vol. 5 . Delft: Publications on Geodesy, 1968.

Cressie N., Hawkins D.H. *Robust estimation of the variogram: I*. Vol. 12. 2 vols. Mathematical Geology, 1980.

Emardson, T. R., Simons, M., Webb, F. H. "Neutral atmospheric delay in interferometric synthetic aperture radar applications: Statistical description and mitigation." *Journal of Geophysical Research: Solid Earth*, 2003: 2156-2202.

Ferretti, A., C. Prati, and F. Rocca. "Permanent scatterers in SAR interferometry." *IEEE Transactions on Geoscience and Remote Sensing* 39, no. 1 (2001): 8,20.

Ferretti, A., C. Prati, and F. Rocca. "Permanent scatterers in SAR interferometry." *IEEE Transactions on Geoscience and Remote Sensing* 39, no. 1 (2001): 8–20.

Ferretti, A., et al. "Submillimeter Accuracy of InSAR Time Series: Experimental Validation." *IEEE Transactions on Geoscience and Remote Sensing* 45, no. 5 (2007): 1142,1153.

Ferretti, A., Savio, G., Barzaghi, R., Borghi, A., Musazzi, S., Novali, F., Prati, C.,

Rocca, F. "Submillimeter Accuracy of InSAR Time Series: Experimental Validation." *IEEE Transactions on Geoscience and Remote Sensing* 45, no. 5 (2007): 1142,1153.

Fjaer, Erling, Holt, Rune, Horsrud, Per, Raaen, Arne M, Risnes, Rasmus. *Petroleum Related Rock Mechanics*. Elsevier Science Publisher, 1992.

Geertsma, J. "A basic theory of subsidence due to reservoir compaction: The homogeneous case." Verhandelingen Kon. Ned. Geol. Mijnbouwk. Gen., 1973.

Genton M. *Highly Robust Variogram Estimation*. Vol. 30. 2 vols. Mathematical Geology, 1998.

Grebenitcharsky, R. and R. Hanssen. "Spatial modeling of atmospheric delays for InSAR and GPS sensors." *5th Annual Meeting*. Utrecht,: European Meteorological Society, 2005.

Handcock, M. S., J. R. Wallis. "An approach to statistical spatialtemporal modeling of meteorological fields." *Journal of the American Statistical Association* 89 (1994): 368–390.

Hanssen, R.F.. *Atmospheric Heterogeneties in ERS Tandem SAR Interferometry*. Delft University Press, 1998.

Hanssen, R.F. *Atmospheric Heterogeneties in ERS Tandem SAR Interferometry*. Delft University Press, 1998.

Hanssen, R.F.,. *Radar Interferometry: Data Interpretation and Error Analysis*. Springer, 2001.

Hetland, E. A., P. Musé, M. Simons, Y. N. Lin, Agram P. S., and C. J. DiCaprio. "Multiscale InSAR Time Series (MInTS) analysis of surface deformation." *Journal of Geophysical Research: Solid Earth* 117 (February 2012).

Houtenbos, A.P.E.M. *Subsidence Residual Modeling, SuRe user manual*. Houtenbos Geodetic Consultancy, 2004.

Houtenbos, A.P.E.M., and F. Kenselaar. *Peilmerk hoogte variaties – stochastische analyse van peilmerkbeweging in Nederland*. TUDelft, 2001.

Jonsson, S., H. Zebker, P. Segall, and F. Amelung. “Fault Slip Distribution of the 1999 Mw7.1 Hector Mine, California, Earthquake, Estimated from Satellite Radar and GPS Measurements.” *Bulletin of the Seismological Society of America* 92, no. 4 (May 2002).

Ketelaar, G. *Satellite Radar Interferometry: Subsidence Monitoring Techniques*. Springer, 2009.

Kolmogorov, A. N. “Dissipation of energy in locally isotropic turbulence.” *Doklady Akademii Nauk SSSR* 32 (1941): 16.

Liu, S. *Satellite radar interferometry: estimation of atmospheric delay*. Delft: Delft University of Technology, 2012.

Marinkovic, P., Ketelaar, G., van Leijen, F., and Hanssen, R. “InSAR quality control: Analysis of five years of corner reflector time series.” *FRINGE07*. Frascati, Italy, 2008.

Onn, F. AU , Zebker, H. A. “Correction for interferometric synthetic aperture radar atmospheric phase artifacts using time series of zenith wet delay observations from a GPS network.” *Journal of Geophysical Research: Solid Earth*, 2006: 2156-2202.

Samiei-Esfahany, S., Hanssen, R.F., Thienen-Visser, K., Muntendam-Bos, A. “On The Effect of Horizontal Deformation on InSAR Subsidence Estimates.” *Fringe 2009 workshop proceedings*. Frascati: ESA. 2009.

Shanker Agram, P., *Persistent Scatterer Interferometry in Natural Terrain*. Stanford University, 2010.

Stein, M. L. “Space-time covariance functions.” *Journal of the American Statistical Association* 100 (2005): 310–321.

Tatarskii, V. I. *The Effects of the Turbulent Atmosphere on Wave Propagation*. Jerusalem:: Israel Program for Scientific Translations, 1971.

Teunissen, P. J. G. *Testing theory; an introduction*. Delft: Delft University Press, 2000.

van Leijen, F. *Persistent Scatterer Interferometry based on geodetic estimation theory*. Delft: Delft University of Technology, 2014.

Wackernagel, H. *Multivariate Geostatistics*. Springer Science & Business Media,, 2003.

Yaglom, A.M. *An Introduction to the Theory of Stationary Random Functions*. Prentice-Hall, 1962.