# QUInSAR: Temporal Parameter and Ambiguity Estimation Using Recursive Least-Squares

## A Methodology for Persistent Scatterer Interferometry

Q.E. Verburg

Technische Universiteit Delft

TU Delft
Delft University of Technology

Challenge the future

# QUInSAR: Temporal Parameter and Ambiguity Estimation Using Recursive Least-Squares

## A Methodology for Persistent Scatterer Interferometry

by

**Q.E. Verburg**

in partial fulfilment of the requirements for the degree of

**Master of Science**
in Civil Engineering

at the Delft University of Technology,
to be defended publicly on Thursday August 31, 2017 at 14:00.

| | | |
|---|---|---|
| Student number: | 1354507 | |
| Supervisor: | Prof. dr. ir. R. F. Hanssen, | TU Delft |
| Thesis committee: | dr. ir. F. van den Heuvel, | SkyGeo |
| | dr. ir. F. J. van Leijen, | TU Delft |
| | dr. ir. F. C. Vossepoel, | TU Delft |

*This thesis is confidential and cannot be made public until publication.*

# Abstract

Parameter and ambiguity estimation in the temporal domain, for arcs of differential phase observations between two persistent scatterers (PS), is a critical part in Persistent Scatterer Interferometry (PSI). Deformation models, used as constraint in the parameter estimation, often do not capture the full extent of the deformation behaviour. This results in a poor separation of signal and noise, and rejection of arcs that do not behave conform the functional model. Previous work assumed that deformation behaviour is stationary and that a full time series can be described with a single set of deformation parameters. In order to develop a more broadly applicable deformation model, this study applies a temporal smoothness constraint during parameter estimation, by assuming that deformation rates are affected by a temporally correlated zero-mean random acceleration. This constraint is implemented using recursive least-squares, similar to Kalman filtering, which also enables efficient updating of arcs when new acquisitions are available.

Various kind of deformation types are simulated to create phase observations based on real *TerraSAR-X*, *Radarsat2* and *ERS* stacks of interferograms. This simulated data is processed using the new recursive estimator and results are compared to that of a batch estimator using a steady-state assumption, to analyse the impact of adding a priori information about the smoothness of the physical signal. Furthermore, a case study on real data is performed on an area where non-linear subsidence has occurred, due to soil remediation. This study presents a mathematical framework for incorporating a priori knowledge about the smoothness of the deformation signal as constraint for parameter and ambiguity estimation. Especially non-linear deformation is better estimated using this method, resulting in a higher success-rate, better separation of signal and noise, and more PS passing quality thresholds. The framework moreover enables efficient updating of existing datasets when new acquisition are available.

# Preface

This thesis is written as result of the master program *Geoscience and Remote Sensing (GRS)*. It concludes the final research project needed to obtain the degree of *Master of Science*. I had the pleasure to conduct this study in the radar group of *Delft University of Technology*, under direct supervision of Ramon Hanssen. During this study I focussed on two topics that are of interest for the Persistent Scatterer Interferometry (PSI) community:

> *resolving various types of non-linear deformation patterns*

and

> *efficient updating of datasets when new satellite acquisitions become available.*

Working on these topics was very satisfying and I hope that my contributions can be used in the improvements of this impressive technique. Before diving into depths, I want to thank my committee for all the supervision, *SkyGeo* for providing me data and resources, and all my fellow students and co-workers for making my time in this master program so memorable. Finally, I want to thank Inge for the limitless support during this master program. I am sure that after all the proof reading she could pass a variety of GRS courses.

*Q.E. Verburg*
*Delft, 2017*

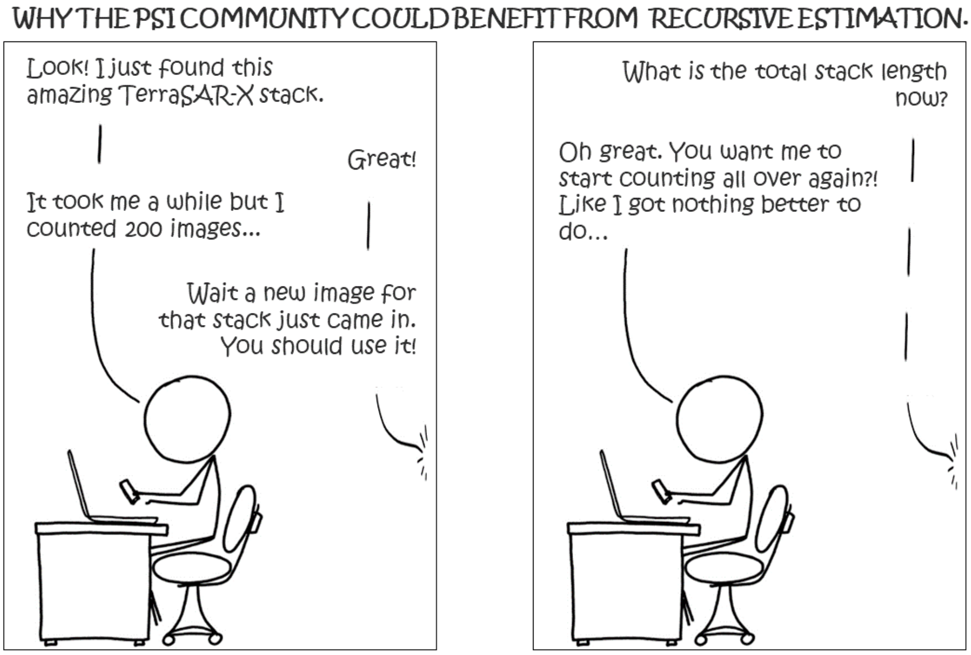


Figure 1: The current PSI approach.

# Contents

# Delft-Radar Notation

## List of Operators

| | |
|---|---|
| $\{\underline{.}\}$ | Stochastic variable |
| $\{\hat{.}\}$ | Estimator |
| $\{.\}^\top$ | Transpose |
| $\{.\}^{-1}$ | Inverse |
| $\{.\}^\star$ | Complex conjugate |
| $\{\hat{.}\}_t$ | Epoch $t$ |
| $\{\hat{.}\}_T$ | Most recent epoch $T$ |
| $\{\hat{.}\}_{t\|t-1}$ | A priori estimator at time $t$ |
| $\{\hat{.}\}_{t\|t}$ | A posteriori estimator at time $t$ |
| $\\| . \\|_Q^2$ | Weighted norm, $(.)^\top Q^{-1}(.)$ |
| $E\{.\}$ | Expectation |
| $D\{.\}$ | Dispersion |
| $\det(.)$ | Determinant |
| $\mathcal{W}\{.\}$ | Wrapping operator |
| $\delta_{t_1 t_2}$ | Kronecker delta $\begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$ |

## List of Symbols

| | |
|---|---|
| $A$ | Amplitude; Design matrix |
| $B^\perp$ | Perpendicular baseline |
| $\psi_t^i$ | SLC phase of pixel $i$ on epoch $t$ |
| $\phi_{0,t}^{i,j}$ | Absolute double-difference phase between epoch $t$ and $0$, of arc $i,j$ |
| $\phi_t$ | Short for $\phi_{0,t}^{i,j}$ |
| $\varphi_{0,t}^{i,j}$ | Wrapped double-difference phase observation between epoch $t$ and $0$, of arc $i,j$ |
| $\varphi_t$ | Short for $\varphi_{0,t}^{i,j}$ |
| $x$ | Vector of unknown variables |
| $\hat{\underline{x}}$ | Estimated unknowns |

| | |
|---|---|
| $\underline{\hat{x}}_{t\|t-1}$ | A priori estimator for epoch $t$, given epoch $t-1$ |
| $\underline{\hat{x}}_{t\|t}$ | A posteriori estimator for epoch $t$ |
| $Q$ | Variance-Covariance Matrix (VCM) |
| $Q_y$ | VCM of observation vector $\underline{y}$ |
| $Q_\phi$ | VCM of the absolute double-difference phases $\phi$ |
| $I_n$ | Identity matrix of size $n \times n$ |
| $\lambda$ | Radar wavelength |
| $\gamma$ | Coherence |
| $\hat{\gamma}$ | Estimate of coherence / ensemble coherence |

# 1

# Introduction

## 1.1 Motivation

Persistent Scatterer Interferometry (PSI) is becoming a more frequently used geodetic technique for deformation measurements. However, it remains difficult to extract various types of non-linear deformation patterns with PSI. The wish of near real-time monitoring and increasing sizes of datasets, moreover provides the need for efficiently updating of datasets when new satellite acquisitions become available. In this study, the deformation model parameterisation is changed to that of a temporal smoothness constraint. The recursive form in which this constraint is implemented also allows for efficient updating of time series.

## 1.2 Background

Since 1978, satellites have been in orbit that are equipped with synthetic aperture radar (SAR). This type of satellites emits electro-magnetic pulses that are reflected by ground elements. The wavelength of the electro magnetic pulses enables the satellite to look through clouds, making it able to monitor the earth surface at a constant rate in contrary to optical satellites. Due to expanding infrastructure and the need to monitor surface deformations in a timely fashion (e.g. preventing dike failures), SAR has become an increasingly popular remote sensing technique. Using the complex phase data of multiple SAR acquisitions of the same area, surface deformations can be measured [Gabriel et al., 1989, Zebker and Goldstein, 1986]. Generally, man-made structures like roads, railways and buildings have a high backscatter, making them highly visible in SAR imagery [Usai, 1997, Usai and Hanssen, 1997]. Furthermore, natural objects like hard surfaces and rock-covered mountains can backscatter coherently, making them clearly visible in SAR.

A lot of progress has been made at the beginning of this century [Ferretti et al., 2000a, 2001, Hanssen, 2001] resulting in the group of techniques referred to as persistent scatterer interferometry (PSI). PSI is based on the backscatter of a coherent radar target, which can be one of earlier mentioned natural or man-made objects. A target with a low geometrical and temporal decorrelation and a high phase coherence is in general a good candidate for PSI. The relative deformation of this point can be found by comparing it with another radar target in the same image at two successive times.

Despite the progress made so far, the ambiguous nature of phase measurements remains. SAR only measures the fractional wave cycle, while the amount of integer cycles between target and satellite remain unknown. This means that movement of a radar target, geometry changes, and propagation delay can cause phase differences to another target to skip phase cycles. An analogy to clarify this phenomenon in a simple way, would be that of a clock that shows the minutes, but not the hours. One would still be able to keep track of time by looking at the clock regularly, but uncertainties become larger when only looking at the clock once or less per hour, as is visualized in figure 1.1.

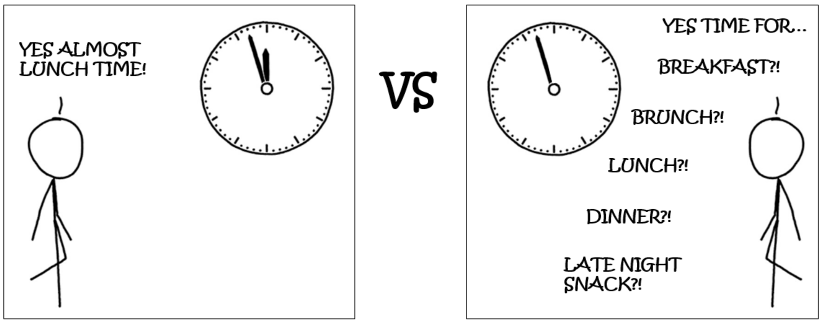


Figure 1.1: Ambiguous phase measurement analogy. When information is available about the hours and minutes, a pretty good guess can be made if it is lunchtime already. However, when the clock only shows the minute dial, we would need to check the clock very frequently to get an indication of how many hours have passed.

In terms of deformation, the same happens in PSI. Given a wavelength of $\lambda = 31$ mm, only a maximum deformation of $D \approx 8$ mm can be measured. A deformation of $D = 16$ mm would appear as $D \approx 0$ mm. This does not necessarily pose a problem when the deformation signal is a slow temporal process. However, when measured sparsely, e.g. yearly, more uncertainty enters the deformation estimation, resulting in an unresolvable signal.

A common technique to resolve the ambiguities of these phase measurements is to assume a certain deformation model for the movement of radar targets. A common made assumption is that targets subside with a constant velocity (steady-state deformation) [Ferretti et al., 2001, Wright et al., 2001, Zebker et al., 1997], but also more complex models like quadratic, seasonal and exponential deformation [Ferretti et al., 2000b, Kampes and Hanssen, 2004, Samiei-Esfahany, 2008, van Leijen, 2014, Van Leijen and Hanssen, 2007] can be used.

Although the steady-state model is valid in many situations, many interesting deformation cases are unresolvable. Especially when observing a longer period of time, the currently used deformation models can no longer accurately describe the occurring patterns. External influences like major construction work or fluid injection and extraction [van der Kooij, 1997] can induce sudden changes in deformation patterns [Ketelaar, 2008] which would, with the currently used models, need very specific knowledge about the resulting deformation shape and times of external interferences. A paradox comes into play here: The model needs to be known in order to estimate the unknown parameters, however, to know the model, the unknown parameters need to be estimated. This implies the need for incorporating a priori knowledge, which unfortunately is not always available. Improvements of these commonly used unwrapping models, in the form of a more broadly applicable deformation model, is hence desirable. Better deformation estimations and fewer errors in estimated ambiguities, will result in more successfully identified Persistent Scatterers.

*In this study the idea is explored to unwrap based on a temporal smoothness constraint.* The functional model is based on a steady-state deformation model which is not stationary over the full time-series, but changes gradually over time. For this specific type of dynamic model Teunissen [2007] has derived a solution based on recursive least squares, similar to Kalman filtering [Kalman, 1960], which is the basis for applying the smoothness constraint and resolving the ambiguities. Moreover, the recursive nature of this solution enables efficient updating when new observations become available. Instead of reprocessing the complete time series in a batch adjustment, only the latest observation is used for adjusting the unknown parameters and, implicitly, resolving the unknown ambiguity.

## 1.3 Research question

This research uses existing theory about recursive estimation and extents it to a suitable methodology for phase unwrapping in PSI. To guide the scope of this study a broad research question is formulated:

> *Can a recursive estimator be used for phase unwrapping in PSI by applying a smoothness constraint in the temporal domain? What are its advantages and challenges?*

To help find an answer, the research question is split up into various sub-questions which will be answered subsequently as the study progresses:

1. What are the shortcomings of currently used unwrapping algorithms in PSI?
2. How can the temporal smoothness constraint be modelled and how can contextual knowledge be incorporated in the parameters of this model?
3. How does the recursive estimator performs in comparison with the currently used algorithms?
4. Does the recursive estimator performs better on deformation types where current algorithms have issues?

## 1.4 Goals

To set certain boundaries to the scope of the study a set of goals and limitations are defined:

1. Develop an recursive algorithm that can be used in the PSI processing chain to unwrap single-arcs of double-difference phase observations.
   (a) Unwrap steady-state deformation.
   (b) Unwrap non linear deformation like breakpoint, quadratic and other types of temporally smooth deformation.
   (c) Estimate a thermal coefficient that describes the relation between temperature and thermal dilations.
   (d) Refine point height (Residual DEM height) during the recursive estimation.
2. Implement the algorithm in the SkyGeo processing chain.
3. Determine full variance-covariance matrix for the estimated parameters.
4. Demonstrate the recursive method on real data and compare the results with currently used algorithms.
5. Propose a method to use the recursive estimator to update datasets when new acquisitions come in.

## 1.5 Limitations

This study focusses on the unwrapping of single-arc double-difference phase observations, obtained from differencing the interferometric phases, of two pixels from a coregistered single-master stack. To narrow down the scope, this study will not focus on network creation and testing, nor will it concern with estimation of orbital and atmospheric signals.

## 1.6 Outline

A short introduction to radar remote sensing and PSI is given in chapter 2. Thereafter, a short recap of adjustment theory is provided in chapter 3, followed by the recursive least squares theory n which this study is based. The practical implementation of this theory for unwrapping is discussed in chapter 4. Chapter 5 and 6 are devoted to testing the proposed methods on simulated and real data. The same data is also processed using conventional phase unwrapping methods for comparison. The results are summarised and discussed in chapter 7.

# 2

# Phase unwrapping in PSI

*This introductory chapter will put phase unwrapping in the context of radar remote sensing. First a short introduction to Radar remote sensing is given, thereafter, it is explained how these radar observations can be used for observing a certain area using a repeat-pass configuration. To accomplish this, first the concept of interferometric radar is introduced, to follow with an introduction to Persistent Scatterer Interferometry (PSI).*

## 2.1 Introduction to radar

Imaging radars, employed in some earth observation satellite missions are, among others, used for earth surface measurements [Gabriel et al., 1989]. They produce imagery that differ from optical images in multiple aspects:

- Radar operates in a much longer wavelength than optical sensors, which operate in the visible light bandwidth, with wavelengths in the order of $10000$ times larger. By using this longer wavelength, radar is insensitive for optical obstructions like clouds, enabling radar to measure the earth surface continuously.
- Whereas optical sensors rely on reflected sunlight, radar is self-illuminating by emitting electromagnetic pulses that are reflected by the earth's surface. This ability makes it independent from external illumination, and adds the capability to compare the emitted signal with the received signal.
- Optical sensors can map pixels to a ray, such that it can derive from which direction a signal comes from. Space-born radar sensors measure in a different way. They measure the two-way *range* between sensor and earth surface, while travelling along the *azimuth* direction of the satellite orbit. Hence the reference system for these imagery is range and azimuth.

From the complex observation of a radar pixel, two properties can be derived; amplitude and phase. The amplitude is a measurement of the level of reflection back to the satellite. The phase indicates the fraction of a wave cycle measured by the radar instrument. If an integer amount of wave cycles fits in the 2-way distance between earth and the satellite, the measured phase is $0$ radians. When this distance is half a wavelength longer, the measured phase is $\pi$ radians. The exact amount of phase cycles is not measured, and can only be computed when the exact distance between satellite and earth surface is known. The measurement of pixel $i$ is stored as a complex value:

$$y^i = \mid A^i \mid e^{j\psi^i}, \tag{2.1}$$

where the real part $A^i$ is the amplitude of the signal, and $\psi^i$ is the phase with $\psi^i \in [-\pi, \pi)$.

## 2.2 The phase measurement

In this study the interferometric phase observations of multiple radar images are used. This manifests the difference in phase of one pixel between two epochs, which is calculated by multiplying the first pixel times the complex conjugate of the second pixel:

$$\begin{aligned} z &= y_{t_1} y^*_{t_2}, \\ &=\mid A_{t_1} A_{t_2} \mid e^{j(\psi_{t_1}-\psi_{t_2})}, \\ &=\mid z \mid e^{j\varphi_{t_1,t_2}}, \end{aligned} \tag{2.2}$$

where $z$ indicates the complex interferometric observation and $\varphi_{t_1,t_2}$ indicates the relative interferometric phase between epoch $t_1$ and $t_2$. The absolute phase difference is given by [Bamler and Hartl, 1998]:

$$\begin{aligned} \phi_{t_1,t_2} &= \frac{-2\pi 2R_1}{\lambda} - \frac{-2\pi 2R_2}{\lambda}, \\ &= -\frac{4\pi \Delta R}{\lambda}, \end{aligned} \tag{2.3}$$

where $2R$ is the 2-way distance between the satellite and earth, $\lambda$ is the wavelength of the satellite and $2\pi$ is the amount of radians in a single wave-cycle. There are many contributors to this phase measurement that change its value [Hanssen, 2001]:

$$\phi = \phi_{\text{ref}} + \phi_{\Delta H} + \phi_{\text{defo}} + \phi_{\text{atmo}} + \phi_{\text{orb}} + \phi_{\text{scat}} + \phi_{\text{noise}}. \tag{2.4}$$

**Reference surface $\phi_{\text{ref}}$.** A trivial contributor to the phase measurement is the reference surface, $\phi_{\text{ref}}$. Assuming that a reference surface is available together with the satellite position during acquisition, it can be computed what the phase value would be for a reflection on the reference surface. This is done by computing the range $R$ from the satellite to a reference surface and multiplying it by the *double wavenumber*, which follows from Eq. (2.3):

$$\phi_{\text{ref}} = -\frac{4\pi}{\lambda} \Delta R. \tag{2.5}$$

**Residual DEM height $\phi_{\Delta H}$.** In practice the used reference surface, a digital elevation model (DEM), is not enough to describe the position of a reflecting object, and hence a height difference is present between the reflecting object and the DEM; the *residual DEM height* $\Delta H$. This residual DEM height causes a phase contribution which is different for every interferometric combination, and can be expressed in terms of the perpendicular baseline of the interferometric combination [Zebker and Goldstein, 1986]:

$$\begin{aligned} \phi_{\Delta H} &= \frac{-4\pi}{\lambda} \frac{B_\perp}{R \sin(\theta_{\text{inc}})} \Delta H, \\ &= H2PH \times \Delta H. \end{aligned} \tag{2.6}$$

The slant range from the master orbit pass and the reflecting object is indicated as $R$, $\theta_{\text{inc}}$ is the iteratively updated incidence angle of the radar pulse, $B_\perp$ is the perpendicular baseline of the slave satellite pass to the location of the master pass and the location of the reflecting object, $\Delta H$ is the height of the backscattering object above the reference surface and $H2PH$ is the *Height-to-Phase* conversion factor that shortens the notation.

**Deformation $\phi_{\text{defo}}$** The earth's surface is not constant over time. The subsurface is a complex system that can deform as result of multiple factors like external loadings, injection or extraction of fluids or changing ground water levels. These factors can cause deformation of the backscattering object, which may be also the parameter of interest in PSI research. The deformation phase contribution between two acquisitions is given by:

$$\phi_{t_1,t_2,\text{defo}} = -\frac{4\pi}{\lambda} D_{t_1,t_2}, \tag{2.7}$$

where $D_{t_1,t_2}$ is the *Line-of-Sight (LOS)* deformation between epoch $t_1$ and $t_2$, seen from the satellite.

**Orbit trend $\phi_{\text{orb}}$.** The orbit trend is caused by the imprecision of the orbit estimation. If the provided orbit is incorrect, the reference phase will be incorrect as well. The phase contribution is referred to as the *orbital phase screen (OPS)* [van Leijen, 2014], since the resulting phase, $\phi_{\text{orb}}$, is a low frequency spatially correlated signal.

**Atmospheric delay $\phi_{\text{atmo}}$.** The error signal caused by the electro magnetic pulse travelling through the atmosphere, is called the atmospheric delay. This contribution is also referred to as the *atmospheric phase screen (APS)*. It is estimated by using the full stack of interferograms and the assumption that the atmospheric signal is uncorrelated in time [Hanssen, 2001].

**Scattering phase & noise $\phi_{\text{scat}}$,$\phi_{\text{noise}}$.** These two contributions describe the noise in the phase data. The scattering phase noise, $\phi_{\text{scat}}$, is caused by temporal, geometrical, and volume decorrelation. The other noise term, $\phi_{\text{noise}}$, consists mainly out of thermal instrument noise and processing errors.

### 2.2.1 Double-Phase Differences

In PSI information is extracted from scatterers by looking at their difference to other scatterers. Since one scatterer already embeds interferometric information of multiple acquisitions (i.e. the difference of one acquisition to the master acquisition), the differences between different pixels are called *Double-Difference (DD)* phase observations; phase differences in time and space. The connection between the two pixels is also referred to as *arc*. The functional model for this is [van Leijen, 2014]:

$$\begin{aligned} \underline{\varphi}^{i,j}_{0,s} &= \underline{\varphi}^{i}_{0,s} - \underline{\varphi}^{j}_{0,s} , \\ &= -2\pi a^{i,j}_{0,s} + \phi^{i,j}_{0,s,\Delta H} + \phi^{i,j}_{0,s,\text{defo}} + \phi^{i,j}_{0,s,\text{OPS/APS}} + \underline{\phi}^{i,j}_{0,s,\text{noise}} , \end{aligned} \tag{2.8}$$

where $i$ and $j$ are two different pixels, $0,s$ is the difference in time of slave acquisition $s$ to master acquisition $0$ and $a^{i,j}_{0,s}$ is the integer valued phase ambiguity. As can be seen, the notation $\underline{\varphi}$ is now used instead of $\underline{\phi}$. The difference between these two measures is that $\underline{\varphi}$ is the wrapped version of the absolute phase difference $\underline{\phi}$, meaning that the observation is now in the $[-\pi,\pi)$ interval. The difference between $\underline{\varphi}$ and $\underline{\phi}$ is $-2\pi a$.

One of the advantages of using DD phase observations is that the contribution of $\phi_{\text{OPS/APS}}$ is generally very low when 2 nearby pixels (e.g. arc distance 200 m) are chosen, because both signals are spatially smooth and hence do not change significantly over this short distance.

## 2.3 PSI processing steps

This section discusses the various steps of a PSI processing chain [van Leijen, 2014], in order to understand where and how unwrapping plays a role in this chain.

### 2.3.1 Creating a stack of images

Radar image acquisitions are never taken from the exact same position, resulting in a slightly different acquisition geometry per satellite pass. To be able to use multiple acquisitions for PSI, all slave acquisitions must be *coregistered* to the master acquisition. All slave images get resampled in this process, such that a stack of pixels, covering the same geometry, is created right on top of each other. This process is illustrated in figure 2.1. The result is a *stack* of differential interferograms, often referred to as simply *stack* [Kampes and Usai, 1999].

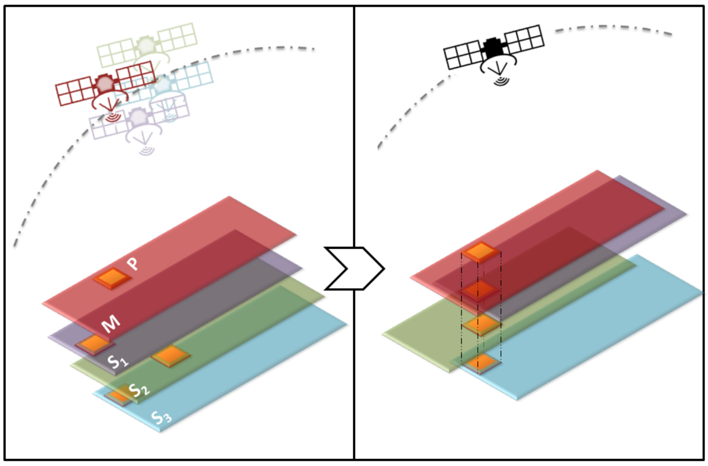


Figure 2.1: The process of *coregistration*. Due to differences in acquisition geometry (i.e. a slightly different acquisition position), the resulting images cover a slightly different area. In other words, a scatterer located in pixel $i$ of the master acquisition, can be located in pixel j of a slave acquisition. The process of *coregistration* shifts the slave acquisitions such that the scattering objects are aligned, which is crucial for PSI. Thereafter, the slave acquisitions get resampled to the same pixel locations as the master acquisition, creating a *stack* of coregistered radar images.

### 2.3.2 PS selection

After obtaining the stack of images for a certain area, the next step is to start PSI processing of a set of low-noise persistent scatterers to create an initial network. For this, an estimation should be made regarding which radar pixels could be potential persistent scatterer candidates for this first order network. The amplitude is used to approximate the phase dispersion of a scatterer. Ferretti et al. [2001] showed that normalised amplitude dispersion is a good indicator of phase dispersion for points with high signal to noise ratio. The normalised amplitude dispersion is given by:

$$D_a = \frac{\sigma_a}{\mu_a} \approx \sigma_\phi, \tag{2.9}$$

where $D_a$ is the normalized amplitude dispersion, $\sigma_a$ is the standard deviation of amplitude values, $\mu_a$ is the mean amplitude value and $\sigma_\phi$ is the dispersion of phase values. A threshold is chosen such that only pixels with a low normalised amplitude dispersion cover the area of interest. These pixels are called *1st order Persistent Scatterer candidates (PSc1)* [van Leijen, 2014].

### 2.3.3 Network construction

The threshold in the previous step is chosen in such a way that PSc1's are equally distributed over the whole area of interest. This is done to create a first order network over this area of interest that can aid in resolving more PS. A network is a collection of PS that are connected via arcs to form a redundant system. This network is often formed by performing a *Delauney triangulation* [Delaunay, 1934]. The DD phase observations are computed for all arcs and form the input for the next step.

### 2.3.4 Parameter estimation

With the arcs of the network as input, we arrive at the step where this study is centred around; *parameter estimation*, also named *unwrapping*. The term parameter estimation is used in this study because the phase ambiguities are implicitly resolved by estimating the deformation over an arc. The unknown parameters are the unknown integer ambiguities $a_{0,s}$, the height difference between the two PSc1 $\Delta H$, and the deformation for every interferometric pair $D_{0,s}$. With a stack of $T$ images, this leads to $(T-1)+1+(T-1)=2T-1$ unknown variables. Some assumptions have to be made to constrain this estimation problem. A popular assumption is a steady-state model for the deformation, described by one parameters: deformation rate. This changes the deformation to:

$$D_{0,s}^{i,j} = v^{i,j}(t_s - t_0) \tag{2.10}$$

where $i, j$ indicates the arc between pixel $i$ and $j$, and $t_s, t_0$ indicate the time of the slave and master acquisition, respectively. The amount of unknown variables is now $(T-1)+2=T+1$. Although an rank defect remains, the ambiguities can be resolved by constraining the height and deformation parameters, using pseudo-observations.

The pseudo observation for the residual DEM height depends on the quality of the DEM and the area of interest. When an urban area is being processed, the DEM accuracy is most likely not describing the building heights very well. In order to resolve large height differences over arcs, the pseudo-observation should be of lower precision. The dispersion of the pseudo-observation can for example be loosened in the form of $\sigma_{\Delta H}$, indicating the standard deviation of possible $\Delta H$. With a $\sigma_{\Delta H}=30$ m, the 2-sigma confidence interval covers a $\Delta H$ of 60 m. The same can be done for creating a pseudo-observation for the deformation rate. When it is known a priori that the range of deformation rates within a certain area is $\pm 20 mm/y$, the dispersion of the pseudo-observation could be set on $\sigma_v = 10$ mm/y.

Due to the ambiguous nature of the phase measurement, there is a maximum resolvable deformation rate. Assuming that deformation is the only factor influencing two subsequent observations, the maximum rate of deformation is given by:

$$v = \frac{\frac{\lambda}{4}}{\Delta t}, \tag{2.11}$$

where $\lambda$ is the wavelength and $\Delta t$ is the nominal repeat cycle of the satellite. In the case of *TerraSAR-X* with $\lambda = 31$ mm and $\Delta t = 11$ days, this results in a LOS deformation rate of 257 mm/y. However, this is just a theoretical value assuming no noise. The addition of noise will lower this value, because the maximum resolvable difference between two consecutive observations of $\frac{\lambda}{4}$ is now lowered by observation phase noise, as illustrated in figure 2.2.

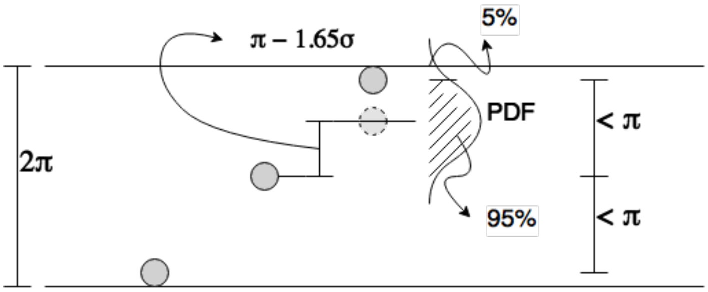


Figure 2.2: Derivation for maximum resolvable deformation rate, given that two consecutive phase observations are only influenced by deformation and observation phase noise. The maximum resolvable deformation between the two phases is $\pi$, minus the phase noise. Assuming a resolvability success rate of 95 %, and ignoring the lower tail of the probability density function (PDF), the maximum resolvable deformation is given by $\pi - \sqrt{2}(1.65\sigma_{\text{noise [rad]}}) = \frac{\lambda}{4} - 2(1.65\sigma_{\text{noise [m]}})$. The factor $\sqrt{2}$ is to include that both observations are stochastic.

When the constraint is set that 95 % of observation noise instances should not cause violation of the $\frac{\lambda}{4}$ restriction, the maximum resolvable linear rates can be computed by[1]:

$$v = \frac{\frac{\lambda}{4} - \sqrt{2}\,(1.65\sigma_{\text{noise}})}{\Delta t}. \tag{2.12}$$

The result of the maximal resolvable deformation rates, for different satellites and noise levels, is displayed in figure 2.3. Note that deformation is measured in the line-of-sight (LOS) of the satellite, therefore the deformation rates are also in the LOS direction. Additionally, note that these values are deformation rates over an arc, thus the difference in deformation rate between two PS.

---

[1] assuming that the effect of the lower tail of observation phase noise in figure 2.2 can be neglected.

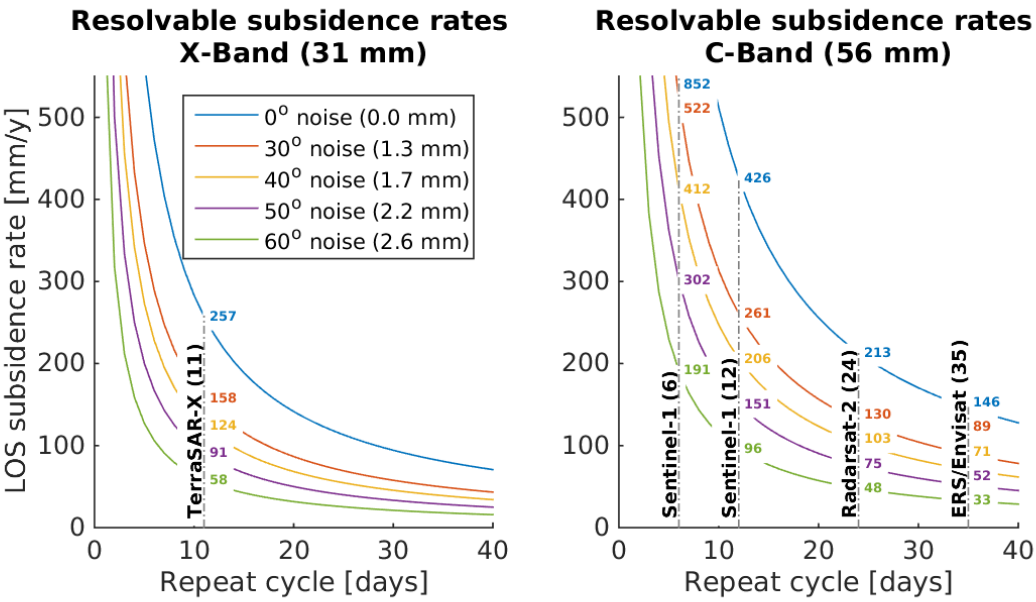


Figure 2.3: Maximum resolvable deformation rate as a function of nominal repeat cycle and observation phase noise, assuming only two consecutive images. It can be clearly observed that the maximum resolvable deformation rate decreases when noise and repeat cycle time increase. The values for some satellite mission are plotted in the graph, with the nominal repeat cycle of the satellite in parentheses behind the satellite name.

### Integer least squares

One way of resolving the ambiguity resolution is using Integer Least-Squares (ILS) [Teunissen, 1993]. In this technique the system of equations:

$$E\{\underline{y}\} = Aa + Bb, \tag{2.13}$$

is solved where $a$ is the vector of unknown integer ambiguities and $b$ is the vector of parameters to be estimated. First, the float solution of vector $a$ is obtained using ordinary Least-Squares. Thereafter, an estimator is chosen to approximate the integer ambiguities. The simplest way is to round the float ambiguities to the nearest integer. The best method to do this is to use Integer Least Squares (ILS) [Teunissen, 1993] which presents an optimal estimation of the ambiguity resolution [Teunissen, 1999], but does this at the cost of computational efficiency [van Leijen, 2014] compared to the Integer Bootstrapping method [Teunissen and Kleusberg, 1996].

### Periodogram

A metric for quality is to estimate the temporal ensemble coherence [Ferretti et al., 2001, van Leijen, 2014]:

$$\hat{\underline{\gamma}} = \left| \frac{1}{T-1} \sum_{t=1}^{T-1} e^{j(\underline{\varphi}_{0,t} - \hat{\underline{\varphi}}_{0,t})} \right|, \tag{2.14}$$

where $\hat{\underline{\gamma}}$ is the estimated temporal ensemble coherence, $t$ is an acquisition, $T$ is the total amount of acquisitions, $\underline{\varphi}_{0t}$ is the phase observation of interferometric pair $0,t$ and $\hat{\underline{\varphi}}_{0,t}$ is the estimated phase observation of interferometric pair $0,t$. This quality measure can be used as cost function for unwrapping using a periodogram [Counselman and Gourevitch, 1981, Ferretti et al., 2001]. This method searches the solution space for the parameters that maximize the temporal ensemble coherence as described in van Leijen [2014] as the *Ambiguity Function*:

$$\arg\max_{b} |\hat{\gamma}| = \left| \frac{1}{T} \sum_{t=1}^{T} e^{j(\underline{\varphi}_{0,t} - \hat{\underline{\varphi}}_{0,t}(b))} \right|, \tag{2.15}$$

where $b$ is the vector of model parameters and $\underline{\hat{\varphi}}_{0,t}(b)$ is the modelled deformation phase of epoch $t$ using parameters $b$. This method can be used to iteratively converge to a maximum by narrowing the solution space stepwise. The solution space of this method gains an extra dimension for every parameter to be estimated and is therefore computationally expensive for higher order deformation models. For this reason, the periodogram is mostly used for estimating the residual height in combination with a constant velocity model, by modelling the phase as:

$$\begin{aligned}\varphi_{\text{modeled}}(B_\perp,t) &= \varphi_{\Delta h}(B_\perp)+\varphi_v(t),\\ &= -\frac{4\pi}{\lambda}\frac{B_\perp}{R\sin\theta_{\text{inc}}}\Delta H+\frac{-4\pi}{\lambda}t.\end{aligned} \tag{2.16}$$

## 2.3.5 Network parameter estimation

For every arc the deformation parameters, residual height and absolute phase observations are now known. These values remain relative in space and only say something about the difference between two PSc. Using the network of arcs created earlier, the parameters and ambiguities can be spatially estimated using path integration to a single reference point, such that PS parameters can be compared to each other now that they use the same reference. However, residues will inevitably be found due to errors in some arcs, caused by wrongly unwrapped ambiguities. Kampes [2006] proposes the testing of the network based on the estimated parameters, while van Leijen [2014] proposes testing of the ambiguities per interferogram. The latter incorporates the B-method of testing [Baarda, 1968], and allows for testing of different deformation models.

## 2.3.6 OPS/APS estimation

With the unwrapped and tested 1st order network, an estimation can be made concerning the orbital and atmospheric phase screen (OPS/APS) [Ferretti et al., 2000a], both assumed to be spatially smooth signals. Recap that the contributions to $\phi$ of a PS are:

$$\underline{\phi}=\underbrace{\phi_{\Delta H}+\phi_{\text{defo}}}_{\text{estimated}}+\phi_{\text{atmo}}+\phi_{\text{orb}}+\underbrace{\underline{\phi}_{\text{scatt}}+\underline{\phi}_{\text{noise}}}_{\text{low}}. \tag{2.17}$$

This shows that the phase residuals of a PS, when corrected for height and deformation, are mainly caused by $\phi_{\text{orb}}$ and $\phi_{\text{atmo}}$. The atmospheric and orbit error signal of a PS can therefore be found by high-pass filtering a PS, or simply by taking the residuals to the adjusted observations of the PS. Per acquisition, these values of all PS are used to estimate an orbital and atmospheric phase screen. The OPS is estimated as a plane and removed from the data. Then the APS can be estimated using interpolation.

## 2.3.7 Densification

Using the 1st order network, together with the OPS and APS estimation, more PSc can be unwrapped. Up until recently this was often done by solely redoing the initial point selection of step 2.3.2 while using a lower threshold [van Leijen, 2014]. However, the increased computation power makes it feasible to unwrap all radar pixels. This is of great interest, because not all pixels with usable signal-to-noise ratios have a low amplitude dispersion. Unwrapping all pixels increases the amount of PS, as shown in the *area of interest aproach* in [van Leijen, 2014] where for a certain small AOI all pixels are processed for potential scatterers, to obtain more PS.

The densification is is done by correcting the new PSc by using the OPS and APS, and by creating arcs to the 3 nearest network PS. The amount of 3 arcs enables testing of the unwrapping result [van

Leijen, 2014]. PSc with a low quality metric value will be rejected and will hence not be used for the final result.

This study focusses on the parameter estimation of arcs in PSI. It is used in the network creation and densification steps of PSI processing.

# 3

# Batch and recursive estimation theory

*A recap of adjustment theory is provided in this chapter. First, simple least squares will be explained together with the weighted version of this algorithm. Second, it will be shown what specific set of weights will lead to the Best Linear Unbiased Estimator (BLUE) and will the relation between Maximum Likelihood Estimation (MLE) and the BLUE moreover be explained. It will then be shown how these algorithms can be used in a recursive way. First this will be done for fitting a static model, meaning that a reduced set of parameters define the complete deformation time series of a PS. Then it is shown how this can be rewritten for a time-varying case, to allow for a methodology for estimating a dynamic process, resulting in displacement estimates per epoch.*

## 3.1 Batch estimation

### 3.1.1 (Weighted) Least Squares

One of the most common methods for parameter estimation is Least Squares [Gauss, 1809]. The method is built on the method of estimating parameters $x$ that minimise the sum of squared errors between the observations $\underline{y}$ and the adjusted observations $\underline{\hat{y}}$. Given the functional model [Gauss, 1809, Koch, 1999, Markov, 1900]:

$$E\left\{\underline{y}\right\} = Ax, \tag{3.1}$$

where $\underline{y}$ is an $m \times 1$ vector of $m$ stochastic observables, $A$ is an $m \times n$ model matrix that transforms the $n \times 1$ vector $x$ of $n$ unknowns into $y$. The estimator $\underline{\hat{x}}$ is found by [Koch, 1999]:

$$\begin{aligned} \underline{\hat{x}}_{LSQ} &= \arg\min_{x}(\| \underline{y} - Ax \|^2), \\ \underline{\hat{x}}_{LSQ} &= (A^\intercal A)^{-1} A^\intercal \underline{y}, \end{aligned} \tag{3.2}$$

Different weights can be given to each observation, provided in a weight matrix $W$. This can be useful when it is known that some observations have an higher precision than others [Koch, 1999]:

$$\begin{aligned} \underline{\hat{x}}_{WLSQ} &= \arg\min_{x}(\| \underline{y} - Ax \|_W^2), \\ \underline{\hat{x}}_{WLSQ} &= (A^\intercal W A)^{-1} W A^\intercal \underline{y}. \end{aligned} \tag{3.3}$$

Observe that Eq. (3.3) equals (3.2) in the case that the weight matrix $W$ equals $c \times I$, meaning any identity matrix times a scalar, resulting in a diagonal matrix with equal entries for the weights.

### 3.1.2 Maximum Likelihood Estimation and BLUE

MLE is an alternative method to estimate the parameters $x$ by maximising the conditional likelihood of observation vector $\underline{y}$. Given measurements with Gaussian-distributed errors, we can express the functional model as [Koch, 1999]:

$$\begin{aligned} E\left\{\underline{y}\right\} &= Ax, \\ D\left\{\underline{y}\right\} &= Q_{yy} \text{ with } \underline{y} \sim \mathcal{N}(Ax, Q_{yy}). \end{aligned} \tag{3.4}$$

The likelihood $f_y(\underline{y} \mid x)$ is given by:

$$f_y(\underline{y} \mid x) = \frac{1}{\sqrt{\det(2\pi Q_{yy})}} \exp\left(-\frac{1}{2} \parallel \underline{y} - Ax \parallel^2_{Q_{yy}^{-1}}\right). \tag{3.5}$$

Since maximising the likelihood is the same as maximising the log likelihood, the problem can be reduced to [Koch, 1999]:

$$\begin{aligned} \underline{\hat{x}}_{MLE} &= \arg\max_x(\ln f_y(\underline{y} \mid x)), \\ &= \arg\min_x(\parallel \underline{y} - Ax \parallel^2_{Q_{yy}^{-1}}), \end{aligned} \tag{3.6}$$

which has an analytical solution for linear problems [Koch, 1999]:

$$\begin{aligned} \underline{\hat{x}}_{MLE} &= (A^{\intercal} Q_{yy}^{-1} A)^{-1} A^{\intercal} Q_{yy}^{-1} \underline{y}, \\ Q_{\hat{x}_{\text{MLE}}} &= (A^{\intercal} Q_{yy}^{-1} A)^{-1}. \end{aligned} \tag{3.7}$$

Observe that this analytical solution is the same as Eq. (3.3) when $Q_{yy}^{-1}$ is used as weight matrix $W$ and $\underline{y}$ is Gaussian distributed. This is also known as the Best Linear Unbiased Estimator (BLUE).

## 3.2 Recursive Least-Squares

When a continuous data stream is being estimated, measurements can come in on a fixed time interval. In modern applications, the number of measurements can easily reach millions. For this purpose, it would be beneficial if the estimator does not have to be recalculated for every incoming measurement. This would not only save a lot of processing time, but also reduce the number of measurements that have to be stored. This section is a summary of Teunissen [2007].

### 3.2.1 Static case

This section will show how the estimation of a linear system can be done recursively. Static refers to the fact that a single set of parameters describes the complete model for every epoch. This is contrary to estimating a dynamic process like tracking the position of a moving car, discussed later in this chapter, which often cannot be described by a single set of parameters describing the period of interest. Given a partitioned model for estimating a static model [Teunissen, 2007]:

$$E\{\begin{bmatrix} \underline{y}_0 \\ \underline{y}_1 \\ \vdots \\ \underline{y}_T \end{bmatrix}\} = \begin{bmatrix} A_0 \\ A_1 \\ \vdots \\ A_T \end{bmatrix} x \quad ; \quad D\{\begin{bmatrix} \underline{y}_0 \\ \underline{y}_1 \\ \vdots \\ \underline{y}_T \end{bmatrix}\} = \begin{bmatrix} Q_0 & & & 0 \\ & Q_1 & & \\ & & \ddots & \\ 0 & & & Q_T \end{bmatrix}, \tag{3.8}$$

where $\underline{y}_{0,1,\ldots,T}$ are assumed to be uncorrelated vectors with corresponding variance-covariance matrices $Q_{1,2,\ldots,T}$. The solution of Eq. (3.8) is given by [Teunissen, 2007]:

$$\begin{aligned} \underline{\hat{x}}_{(T)} &= (\sum_{t=0}^{T} A_t^{\intercal} Q_t^{-1} A_t)^{-1} (\sum_{t=0}^{T} A_t^{\intercal} Q_t^{-1} \underline{y}_t), \\ Q_{\hat{x}_{(T)}} &= (\sum_{t=0}^{T} A_t^{\intercal} Q_t^{-1} A_t)^{-1}. \end{aligned} \tag{3.9}$$

Writing this in a recursive form leads to [Teunissen, 2007]:

$$\boxed{\begin{array}{l} \begin{cases} \underline{\hat{x}}_{(t)} &= (Q^{-1}_{\hat{x}_{(t-1)}} + A_t^{\intercal} Q_t^{-1} A_t)^{-1} (Q^{-1}_{\hat{x}_{(t-1)}} \underline{\hat{x}}_{(t-1)} + A_t^{\intercal} Q_t^{-1} \underline{y}_t) \\ Q_{\hat{x}_{(t)}} &= (Q^{-1}_{\hat{x}_{(t-1)}} + A_t^{\intercal} Q_t^{-1} A_t)^{-1} \end{cases} \\ \text{Given the initial state:} \\ \begin{cases} \underline{\hat{x}}_{(0)} &= (A_0^{\intercal} Q_0^{-1} A_0)^{-1} A_0^{\intercal} Q_0^{-1} \underline{y}_0 \\ Q_{\hat{x}_{(0)}} &= (A_0^{\intercal} Q_0^{-1} A_0)^{-1} \end{cases} \end{array}}\ , \tag{3.10}$$

from which the recursive definition can also be written as:

$$\begin{aligned} Q_{\hat{x}_{(t)}} &= (Q^{-1}_{\hat{x}_{(t-1)}} + A_t^{\intercal} Q_t^{-1} A_t)^{-1}. \\ \underline{\hat{x}}_{(t)} &= \underline{\hat{x}}_{(t-1)} + Q_{\hat{x}_{(t)}} A_t^{\intercal} Q_t^{-1} (\underline{y}_t - A_t \underline{\hat{x}}_{(t-1)}), \end{aligned} \tag{3.11}$$

This notation is often used in Maximum A Posteriori (MAP) estimation [e.g. Kalman, 1960], discussed in section 3.4 (p. 23) for its similarities to recursive least-squares, where $\underline{\hat{x}}_{(t)}$ would indicate the prior estimator and $Q_{\hat{x}_{(t-1)}}$ its covariance matrix.

### 3.2.2 Time varying case

The previous section describes the static case, for example estimating body temperature based on a series of measurements. Assuming that a time varying parameter has to be estimated (e.g. position of a moving car), Eq. (3.8) gets extended to [Teunissen, 2007]:

$$E\{\begin{bmatrix} \underline{y}_0 \\ \underline{y}_1 \\ \vdots \\ \underline{y}_T \end{bmatrix}\} = \begin{bmatrix} A_0 & & & 0 \\ & A_1 & & \\ & & \ddots & \\ 0 & & & A_T \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_T \end{bmatrix} \quad ; \quad D\{\begin{bmatrix} \underline{y}_0 \\ \underline{y}_1 \\ \vdots \\ \underline{y}_T \end{bmatrix}\} = \begin{bmatrix} Q_0 & & & 0 \\ & Q_1 & & \\ & & \ddots & \\ 0 & & & Q_T \end{bmatrix}. \tag{3.12}$$

This amount of unknown parameters is impractical for estimation due to rank deficit. However, the fact that there may be a linear relationship between the values of the vector $x_{1:T}$ can be used:

$$x_t = \Phi_{t,0} x_0, \tag{3.13}$$

where $\Phi_{t,0}$ is the $n \times n$ transition matrix that transforms $x_0$ into $x_t$.

#### Transition matrix

Suppose a deformation event can be described by:

$$u(t) = u(t_0) + \int_{t_0}^{t} \dot{u}(\tau) d\tau, \tag{3.14}$$

where $u(t)$ describes the 1-D deformation of the surface and $\dot{u}(t)$ is the continuous time derivative, indicating the deformation rate. In the deterministic case, the first time derivative is zero and Eq. (3.14) simplifies to:

$$\begin{aligned} u(t) &= \Phi_{t,t_0} \cdot u(t_0), \\ \Phi_{t,t_0} &= I, \end{aligned} \tag{3.15}$$

showing that in the deterministic case Eq. (3.12) is the same as Eq. (3.8). Now assume that subsidence is occurring with a continuous time derivative:

$$\dot{u}(t) = \dot{u}(t_0) + \int_{t_0}^{t} \ddot{u}(\tau) d\tau. \tag{3.16}$$

Substituting Eq. (3.16) in (3.14) and using integration by parts, results in:

$$\begin{aligned} u(t) &= u(t_0) + \int_{t_0}^{t} \left( \dot{u}(t_0) + \int_{t_0}^{\tau} \ddot{u}(\rho) d\rho \right) d\tau, \\ &= u(t_0) + \dot{u}(t_0)(t - t_0) + \int_{t_0}^{t} (t_0 - \tau) \ddot{u}(\tau) d\tau, \end{aligned} \tag{3.17}$$

showing that $u(t)$ is known when $u(t_0), \dot{u}(t_0)$ and $\ddot{u}(t)$ are known. Writing Eq. (3.14) and (3.17) together in vector form:

$$\begin{bmatrix} u(t) \\ \dot{u}(t) \end{bmatrix} = \begin{bmatrix} 1 & (t - t_0) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u(t_0) \\ \dot{u}(t_0) \end{bmatrix} + \int_{t_0}^{t} \begin{bmatrix} (t - \tau) \\ 1 \end{bmatrix} \ddot{u}(\tau) d\tau, \tag{3.18}$$

results in a direct expression for $u(t)$ and $\dot{u}(t)$. Assuming a constant velocity model, thus $\ddot{u} = 0$, a clear linear relation between $x(t)$ and $x(t_0)$ is visible:

$$\underbrace{\begin{bmatrix} u(t) \\ \dot{u}(t) \end{bmatrix}}_{x_t} = \Phi_{t,t_0} \underbrace{\begin{bmatrix} u(t_0) \\ \dot{u}(t_0) \end{bmatrix}}_{x_0}, \tag{3.19}$$

so $\Phi_{t,t_0}$ in the case of constant velocity is given by:

$$\boxed{\Phi_{t,t_0} = \begin{bmatrix} 1 & (t - t_0) \\ 0 & 1 \end{bmatrix}}. \tag{3.20}$$

Similarly, it can be shown that for a constant acceleration model the following transition matrix can be obtained:

$$\boxed{\begin{aligned} \begin{bmatrix} u(t) \\ \dot{u}(t) \\ \ddot{u}(t) \end{bmatrix} &= \Phi_{t,t_0} \begin{bmatrix} u(t_0) \\ \dot{u}(t_0) \\ \ddot{u}(t_0) \end{bmatrix} \\ \Phi_{t,t_0} &= \begin{bmatrix} 1 & (t - t_0) & \frac{1}{2}(t - t_0)^2 \\ 0 & 1 & (t - t_0) \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}}. \tag{3.21}$$

As shown in Teunissen [2007, eq. 3.23] there are three important properties of transition models:

$$\boxed{\begin{array}{rl} \text{Initial value:} & \Phi_{t_0,t_0} = I \\ \text{Transition property:} & \Phi_{t,t_0} = \Phi_{t,t_1} \Phi_{t_1,t_0} \;\; \forall \;\; t, t_0, t_1 \\ \text{Inverse property:} & \Phi_{t,t_0}^{-1} = \Phi_{t_0,t} \;\; \forall \;\; t, t_0 \end{array}}. \tag{3.22}$$

In Eq. (3.13) and Eq. (3.19) it is shown that $x_t = \Phi_{t,0} x_0$. The transition and inverse properties of the transition model mentioned in (3.22) can be used to create an expression for an arbitrary time $x_i$ based on $x_t$:

$$x_i = \Phi_{i,t} x_t. \tag{3.23}$$

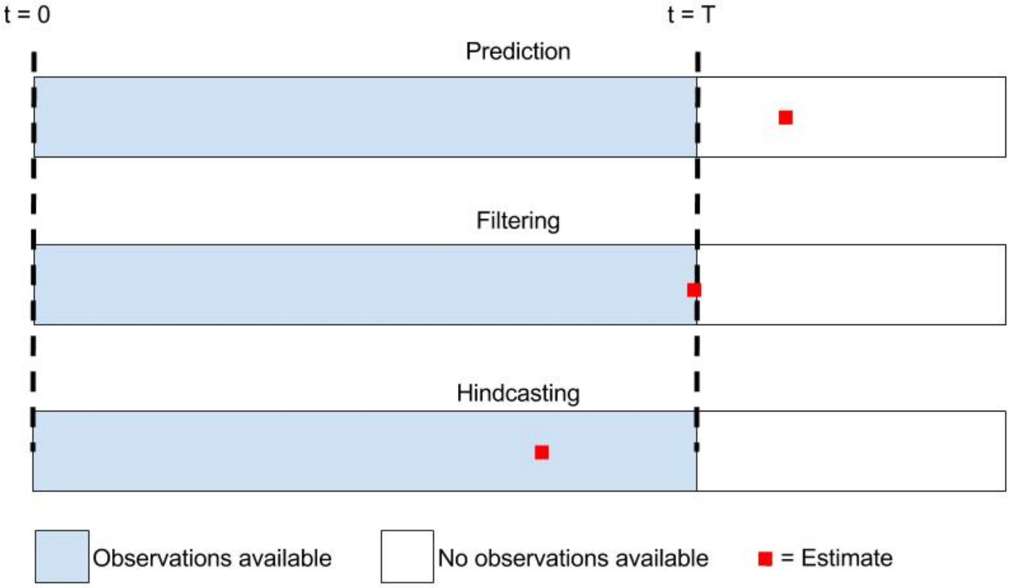


Figure 3.1: The difference between predicting, filtering, and hindcasting. Predicting is estimating a future epoch of which no observations are available yet, filtering is estimating the most recent observation and hindcasting is updating past filter estimates using more recent observations.

Substituting this in Eq. (3.12) results in the new expression for the time-varying case [Teunissen, 2007]:

$$E\{\begin{bmatrix}\underline{y}_0\\ \underline{y}_1\\ \vdots\\ \underline{y}_T\end{bmatrix}\} = \begin{bmatrix}A_0\Phi_{0,t}\\ A_1\Phi_{1,t}\\ \vdots\\ A_T\Phi_{T,t}\end{bmatrix} x_t \qquad ; \qquad D\{\begin{bmatrix}\underline{y}_0\\ \underline{y}_1\\ \vdots\\ \underline{y}_T\end{bmatrix}\} = \begin{bmatrix}Q_0 & & & 0\\ & Q_1 & & \\ & & \ddots & \\ 0 & & & Q_T\end{bmatrix}, \tag{3.24}$$

from which we denote the estimator as $\hat{\underline{x}}_{t|T}$, the estimator at time $t$, given the time series $1:T$.

Three cases can be defined which will now be shortly discussed, and are clarified in figure 3.1.

**1. Predicting:** $t > T$ When a time $t$ later than the most recent epoch $T$, is being estimated, this is called predicting. The BLUE of this prediction, given $\hat{\underline{x}}_{T|T}$ is known, is given by [Teunissen, 2007]:

$$\begin{aligned}\hat{\underline{x}}_{t|T} &= \Phi_{t,T}\hat{\underline{x}}_{T|T} \; ; \qquad t > T,\\ Q_{\hat{x}_{t|T}} &= \Phi_{t,T} Q_{\hat{x}_{T|T}} \Phi_{t,T}^{T}.\end{aligned} \tag{3.25}$$

**2. Filtering:** $t = T$ Filtering is performed when estimating the most current time step. It is needed for both predicting and hindcasting to get the BLUE. The estimator $\hat{x}_{T|T}$ is given by the Least Squares solution of (3.24) [Teunissen, 2007]:

$$\begin{aligned}\hat{\underline{x}}_{T|T} &= (\sum_{t=0}^{T} \Phi_{t,T}^{T} A_t^{T} Q_t^{-1} A_t \Phi_{t,T})^{-1} (\sum_{t=0}^{T} \Phi_{t,T}^{T} A_t^{T} Q_t^{-1} \underline{y}_t),\\ Q_{\hat{x}_{T|T}} &= (\sum_{t=0}^{T} \Phi_{t,T}^{T} A_t^{T} Q_t^{-1} A_t \Phi_{t,T})^{-1}.\end{aligned} \tag{3.26}$$

**3. Hindcasting:** $t < T$ We refer to the situation hindcasting when parameters of ”past” time instances are updated based on the most current estimator. This optimal estimation of past epochs is not to be confused with *smoothing* [Teunissen, 2007] which is a more general term for conditioning past estimates to some past and future observations. Section 3.3 will go into more detail about hindcasting. The estimator looks similar to that of the predicting estimator but for the case of estimating stochastic processes, this will no longer be the case.

$$\begin{aligned} \underline{\hat{x}}_{t|T} &= \Phi_{t,T}\underline{\hat{x}}_{T|T}, \qquad t<T\;; \\ Q_{\hat{x}_{t|T}} &= \Phi_{t,T}Q_{\hat{x}_{T|T}}\Phi_{t,T}^{\mathsf{T}}. \end{aligned} \tag{3.27}$$

### Recursive equations

We now want a set of recursive equations that express the estimator $\underline{\hat{x}}_{t,t}$ in terms of the previous estimator $\underline{\hat{x}}_{t-1,t-1}$ and the latest measurement $\underline{y}_t$, to obtain a solution for the time-varying functional model of Eq. (3.12). This is done by solving the model [Teunissen, 2007]:

$$E\{\begin{bmatrix} \underline{\hat{x}}_{t-1|t-1} \\ \underline{y}_t \end{bmatrix}\} = \begin{bmatrix} \Phi_{t-1,t} \\ A_t \end{bmatrix} x_t \;\;;\;\; D\{\begin{bmatrix} \underline{\hat{x}}_{t-1|t-1} \\ \underline{y}_t \end{bmatrix}\} = \begin{bmatrix} Q_{\hat{x}_{t-1|t-1}} & 0 \\ 0 & Q_{y_t} \end{bmatrix}. \tag{3.28}$$

The model is solved in two parts. First, the part of the model is solved that is referred to as the *time-update*, or *prediction*:

$$E\{\underline{\hat{x}}_{t-1|t-1}\} = \Phi_{t-1,t}x_t \;\;;\;\; D\{\underline{\hat{x}}_{t-1|t-1}\} = Q_{\hat{x}_{t-1|t-1}}. \tag{3.29}$$

The solution of this model is given directly by inverting the transition matrix (see property: (3.22)) and result in the *time-update equations* [Teunissen, 2007]:

$$\boxed{\begin{aligned} \underline{\hat{x}}_{t|t-1} &= \Phi_{t,t-1}\underline{\hat{x}}_{t-1|t-1} \\ Q_{\hat{x}_{t|t-1}} &= \Phi_{t,t-1}Q_{\hat{x}_{t-1|t-1}}\Phi_{t,t-1}^{\mathsf{T}} \end{aligned}}\,. \tag{3.30}$$

It can be seen that this set of equations is similar to the prediction equation in (3.25). The solution of the *time-update equations* can now be used to solve (3.28) by solving:

$$E\{\begin{bmatrix} \underline{\hat{x}}_{t|t-1} \\ \underline{y}_t \end{bmatrix}\} = \begin{bmatrix} I \\ A_t \end{bmatrix} x_t \;\;;\;\; D\{\begin{bmatrix} \underline{\hat{x}}_{t|t-1} \\ \underline{y}_t \end{bmatrix}\} = \begin{bmatrix} Q_{\hat{x}_{t|t-1}} & 0 \\ 0 & Q_{y_t} \end{bmatrix}, \tag{3.31}$$

of which the solution is given in [Teunissen, 2007]:

$$\boxed{\begin{aligned} Q_{\hat{x}_{t|t}} &= (Q_{\hat{x}_{t|t-1}}^{-1} + A_t^{\mathsf{T}}Q_{y_t}^{-1}A_t)^{-1} \\ \underline{\hat{x}}_{t|t} &= \underline{\hat{x}}_{t|t-1} + Q_{\hat{x}_{t|t}}A_t^{\mathsf{T}}Q_{y_t}^{-1}(\underline{y}_t - A_t\underline{\hat{x}}_{t|t-1}) \end{aligned}}\,. \tag{3.32}$$

Because the *time-update* is here combined with the new measurement these equations are called the *measurement-update* which are similar to the previously described *filtering* step. An example result of the time-varying recursive estimation is shown in figure 3.2.

### 3.2.3 Dynamic case

In the previous section of the *time-varying* case, it is assumed that the estimated system is static, meaning that the time-varying vector $x$ is determined completely using a single time instance $x_t$. This is a strict assumption that will not hold for many real life applications. Deformation is the result of a complex combination of factors such as surface loads, subsurface injection, construction dynamics or the composition of the soil. Treating deformation as a stationary static system, is hence a stringent assumption which can cause a strong reduction of identified scatterers.

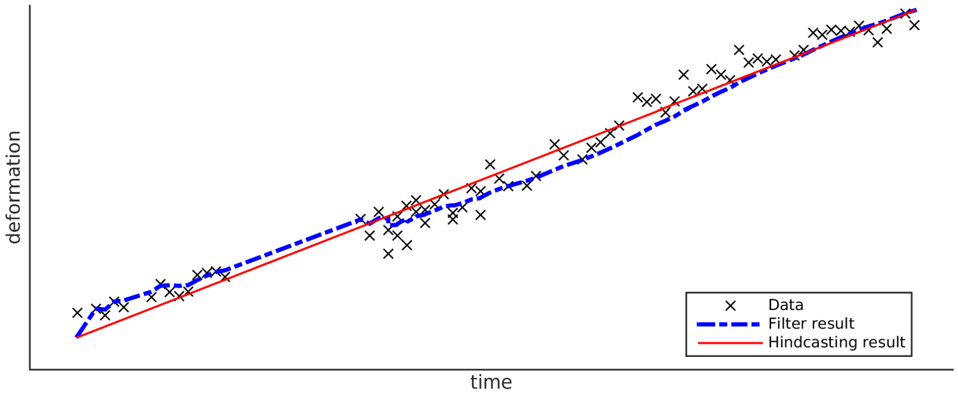


Figure 3.2: Filtering with the time-varying method. The adjusted displacement results of the filter steps are shown in blue. Keep in mind that every adjusted displacement value is only computed using all the observations on the left of that value, i.e. only historic observations are used for adjusting a certain epoch. Due to this, the blue line behaves more smooth towards the end of the time-series, where more data is used for the adjustment. The red line indicates the best estimate (the last epoch), after hindcasting. This is done as described in Eq. (3.27), and is hence a linear extrapolation of the estimate at time $T$, and the same result as after doing batch least-squares adjustment. This same time-series is used later in this chapter to illustrate the result of *dynamic* adjustment and the corresponding hindcasting.

In this section, the functional model will be changed, in such a way that it loses it static properties. The assumption is made that the signal of interest is a temporal smooth signal. On small time scales the behaviour approximates steady-state behaviour, while on longer time scales the steady-state behaviour changes under the influence of a (random) acceleration. These changes are modelled in a difference vector $d$.

Assume that the surface deformation dynamics are described by [Teunissen, 2007]:

$$\dot{x}(t) = F(t)x(t) + G(t)z(t), \tag{3.33}$$

where $F(t)$ and $G(t)$ are system matrices, $x(t)$ are the steady state parameters, and $z(t)$ is an external input function, in this case describing the changes due to acceleration to the steady-state parameters. Now the state transition Eq. (3.23) can be updated to [Teunissen, 2007]:

$$x_t = \Phi_{t,t_0} x_{t_0} + d_{t,t_0}, \tag{3.34}$$

where $d_{t,t_0}$ is the *difference vector*, incorporating the changes due to acceleration to the steady-state parameters. Difference vector $d_{t,t_0}$ is of the same size and units as $x_t$ and is given by the integral [Teunissen, 2007]:

$$d_{t,t_0} = \int_{t_0}^{t} \Phi(t,\tau)G(\tau)z(\tau)d\tau. \tag{3.35}$$

This changes the partial model to:

$$E\{\begin{bmatrix} \underline{y}_0 \\ \underline{y}_1 \\ \vdots \\ \underline{y}_T \end{bmatrix}\} = \begin{bmatrix} A_0\Phi_{0,t} \\ A_1\Phi_{1,t} \\ \vdots \\ A_T\Phi_{T,t} \end{bmatrix} x_t + \begin{bmatrix} A_0 d_{0,t} \\ A_1 d_{1,t} \\ \vdots \\ A_T d_{T,t} \end{bmatrix} \quad ; \quad D\{\begin{bmatrix} \underline{y}_0 \\ \underline{y}_1 \\ \vdots \\ \underline{y}_T \end{bmatrix}\} = \begin{bmatrix} Q_0 & & 0 \\ & Q_1 & \\ & & \ddots & \\ 0 & & & Q_T \end{bmatrix}. \tag{3.36}$$

The values for the difference vector in Eq. (3.36) are not known deterministically and can not be estimated as unknowns due to the lack of redundancy in observations. Fortunately, they can be for some applications added as *pseudo-observations*, when assuming that the acceleration is a zero-mean random process [Teunissen, 2007]:

$$E\left\{\begin{bmatrix}\underline{y}_0\\ \underline{d}_1\\ \underline{y}_1\\ \vdots\\ \underline{d}_T\\ \underline{y}_T\end{bmatrix}\right\}=\begin{bmatrix}A_0 & & & & 0\\ -\Phi_{1,0} & I & & & \\ & A_1 & & & \\ & & \ddots & & \\ & & & -\Phi_{T,T-1} & I\\ 0 & & & & A_T\end{bmatrix}\begin{bmatrix}x_0\\ x_1\\ \vdots\\ x_{T-1}\\ x_T\end{bmatrix}\quad;\quad D\left\{\begin{bmatrix}\underline{y}_0\\ \underline{d}_1\\ \underline{y}_1\\ \vdots\\ \underline{d}_T\\ \underline{y}_T\end{bmatrix}\right\}=\begin{bmatrix}Q_{y_0} & & & & & 0\\ & Q_{d_1} & & & & \\ & & Q_{y_1} & & & \\ & & & \ddots & & \\ & & & & Q_{d_T} & \\ 0 & & & & & Q_{y_T}\end{bmatrix},\tag{3.37}$$

where $Q_{d_t}$ is the VCM for the difference vector $\underline{d}$. From this equation it can be derived that the estimators $\underline{\hat{x}}_{t-1|t}$ and $\underline{\hat{x}}_{t|t}$ of $x_{t-1}$ and $x_t$ depend solely on $\underline{\hat{x}}_{t-1|t-1}$, $\underline{d}_T$ and $\underline{y}_T$. Therefore, $\underline{\hat{x}}_{t-1|t}$ and $\underline{\hat{x}}_{t|t}$ can be computed by solving [Teunissen, 2007]:

$$E\left\{\begin{bmatrix}\underline{\hat{x}}_{t-1|t-1}\\ \underline{d}_t\\ \hdashline \underline{y}_t\end{bmatrix}\right\}=\begin{bmatrix}I & 0\\ -\Phi_{t,t-1} & I\\ \hdashline 0 & A_t\end{bmatrix}\begin{bmatrix}x_{t-1}\\ x_t\end{bmatrix}\quad;\quad D\left\{\begin{bmatrix}\underline{\hat{x}}_{t-1|t-1}\\ \underline{d}_t\\ \hdashline \underline{y}_t\end{bmatrix}\right\}=\left[\begin{array}{cc:c}Q_{\hat{x}_{t-1|t-1}} & 0 & 0\\ 0 & Q_{d_t} & 0\\ \hdashline 0 & 0 & Q_{y_t}\end{array}\right],\tag{3.38}$$

which will be done in two steps. First, the top block of 3.38 is estimated:

$$E\left\{\begin{bmatrix}\underline{\hat{x}}_{t-1|t-1}\\ \underline{d}_t\end{bmatrix}\right\}=\begin{bmatrix}I & 0\\ -\Phi_{t,t-1} & I\end{bmatrix}\begin{bmatrix}x_{t-1}\\ x_t\end{bmatrix}\quad;\quad D\left\{\begin{bmatrix}\underline{\hat{x}}_{t-1|t-1}\\ \underline{d}_t\end{bmatrix}\right\}=\begin{bmatrix}Q_{\hat{x}_{t-1|t-1}} & 0\\ 0 & Q_{d_t}\end{bmatrix}.\tag{3.39}$$

This can be done directly by inverting $\left[\begin{smallmatrix}I & 0\\ -\Phi_{t,t-1} & I\end{smallmatrix}\right]$, because it is full rank:

$$\begin{bmatrix}\underline{\hat{x}}_{t-1|t-1}\\ \underline{\hat{x}}_{t|t-1}\end{bmatrix}=\begin{bmatrix}I & 0\\ \Phi_{t,t-1} & I\end{bmatrix}\begin{bmatrix}\underline{\hat{x}}_{t-1|t-1}\\ \underline{d}_t\end{bmatrix}.\tag{3.40}$$

The corresponding covariance matrices can be found by error propagation:

$$\begin{bmatrix}Q_{\hat{x}_{t-1|t-1}} & Q_{\hat{x}_{t-1|t-1}\hat{x}_{t|t-1}}\\ Q_{\hat{x}_{t|t-1}\hat{x}_{t-1|t-1}} & Q_{\hat{x}_{t|t-1}}\end{bmatrix}=\begin{bmatrix}Q_{\hat{x}_{t-1|t-1}} & Q_{\hat{x}_{t-1|t-1}}\Phi^{\mathrm{T}}_{t,t-1}\\ \Phi_{t,t-1}Q_{\hat{x}_{t-1|t-1}} & \left(\Phi_{t,t-1}Q_{\hat{x}_{t-1|t-1}}\Phi^{\mathrm{T}}_{t,t-1}+Q_{d_t}\right)\end{bmatrix}.\tag{3.41}$$

The resulting recursive equations below are called the *time-update* equations. They show that the predicted estimator $\underline{\hat{x}}_{t|t-1}$ is only dependent on the filtered estimator $\underline{\hat{x}}_{t-1|t-1}$ and the difference vector $d_t$ [Teunissen, 2007]:

$$\boxed{\begin{array}{c}\textbf{Time-update equations}\\ \begin{aligned}\underline{\hat{x}}_{t|t-1}&=\Phi_{t,t-1}\underline{\hat{x}}_{t-1|t-1}+\underline{d}_t\\ Q_{\hat{x}_{t|t-1}}&=\Phi_{t,t-1}Q_{\hat{x}_{t-1|t-1}}\Phi^{\mathrm{T}}_{t,t-1}+Q_{d_t}\end{aligned}\end{array}}\;.\tag{3.42}$$

These equations are the result of solving Eq. (3.39). Using these *time-update* equations, Eq. (3.38) can be solved by solving:

$$E\left\{\begin{bmatrix}\underline{\hat{x}}_{t-1|t-1}\\ \hdashline \underline{\hat{x}}_{t|t-1}\\ \underline{y}_t\end{bmatrix}\right\}=\left[\begin{array}{c:c}I & 0\\ \hdashline 0 & I\\ 0 & A_t\end{array}\right]\begin{bmatrix}x_{t-1}\\ x_t\end{bmatrix}\quad;\quad D\left\{\begin{bmatrix}\underline{\hat{x}}_{t-1|t-1}\\ \hdashline \underline{\hat{x}}_{t|t-1}\\ \underline{y}_t\end{bmatrix}\right\}=\left[\begin{array}{c:cc}Q_{\hat{x}_{t-1|t-1}} & Q_{\hat{x}_{t-1|t-1}\hat{x}_{t|t-1}} & 0\\ \hdashline Q_{\hat{x}_{t|t-1}\hat{x}_{t-1|t-1}} & Q_{\hat{x}_{t|t-1}} & 0\\ 0 & 0 & Q_{y_t}\end{array}\right].\tag{3.43}$$

It can be seen that $x_t$ only depends on the lower right block of the model matrix. Given this fact, the system can be solved by reducing the functional and stochastic model to:

$$E\left\{\begin{bmatrix}\underline{\hat{x}}_{t|t-1}\\ \underline{y}_t\end{bmatrix}\right\}=\begin{bmatrix}I\\ A_t\end{bmatrix}x_t\quad;\quad D\left\{\begin{bmatrix}\underline{\hat{x}}_{t|t-1}\\ \underline{y}_t\end{bmatrix}\right\}=\begin{bmatrix}Q_{\hat{x}_{t|t-1}} & 0\\ 0 & Q_{y_t}\end{bmatrix},\tag{3.44}$$

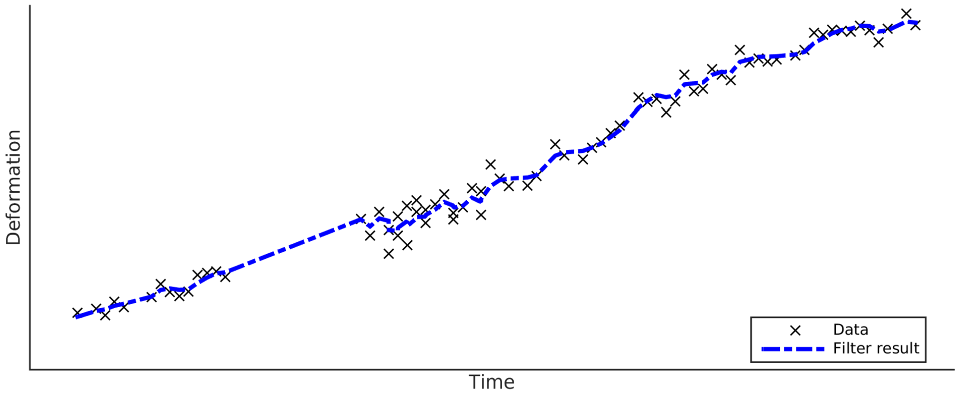


Figure 3.3: The same data as used in figure 3.2 is processed, but now while considering the signal of interest to be a dynamic process. The blue line follows the data more closely than previously shown in figure 3.2, but appears less smooth than the signal estimate in figure 3.2.

of which the solution is previously (3.32) given by:

$$\boxed{\begin{aligned} &\textbf{Measurement-update equations}\\ \underline{\hat{x}}_{t|t} &= \underline{\hat{x}}_{t|t-1} + (Q_{\hat{x}_{t|t-1}}^{-1} + A_t^\intercal Q_{y_t}^{-1} A_t)^{-1} A_t^\intercal Q_{y_t}^{-1} (\underline{y}_t - A_t \underline{\hat{x}}_{t|t-1})\\ Q_{\hat{x}_{t|t}} &= (Q_{\hat{x}_{t|t-1}}^{-1} + A_t^\intercal Q_{y_t}^{-1} A_t)^{-1} \end{aligned}} \;. \tag{3.45}$$

These equations are the *measurement-update* equations of the stochastic surface deformation model. The final result of filtering the dataset used in figure 3.2 can be seen in figure 3.3. The filtered estimate seems to follow the data more closely, but as a result of only using past epochs, the estimate is more rough than the signal is expected to be.

## 3.3 Hindcasting

The filter algorithms that are discussed in this chapter until now, are optimal estimators for the most recent observation. Those estimators only describe their current epoch, and as soon as new observations are available, the estimator is no longer based on all observations and is hence no longer the best possible estimator.

To obtain the most optimal estimation of the parameters at a certain epoch, it is important to *update* past estimators with newer observations. This process is often referred to as smoothing.

Smoothing is estimating a system state at a given time, conditioned to several past, present, and future observations [e.g. Cosme et al., 2012]. This means that e.g. moving average and low-pass filters both categorise as smoothing methods. Hence we introduce a new term to refer to optimal updating of past estimators: *hindcasting*; a specific type of smoothing that returns the maximum a posteriori (MAP) estimator [e.g. Rauch et al., 1965] for all or some specific epochs. MAP estimation can be used to find the value of an unknown parameter, based on prior estimate and distribution, and empirical data with a specific distribution. It differs from Maximum Likelihood estimation by incorporating a prior distribution of the signal. The term hindcasting is used in meteorology and oceanography [e.g. Cardone et al., 1975, Leetmaa and Ji, 1989] to indicate the testing of the performance of prediction models, and is from now on used in the slightly different context of optimally updating parameter estimations using all available observations. The difference in formulation between the MAP estimate of the filter step, and the hindcasted MAP estimate is:

$$\begin{aligned} \underline{\hat{x}}_{\text{MAP,Filter}} &= \underline{\hat{x}}_{T|T} = \arg\max_x f(x_T \mid \underline{y}_{1:T}),\\ \underline{\hat{x}}_{\text{MAP,Hindcasted}} &= \underline{\hat{x}}_{t|T} = \arg\max_x f(x_t \mid \underline{y}_{1:T}) \quad \text{for} \quad t_0 \leq t \leq T. \end{aligned} \tag{3.46}$$

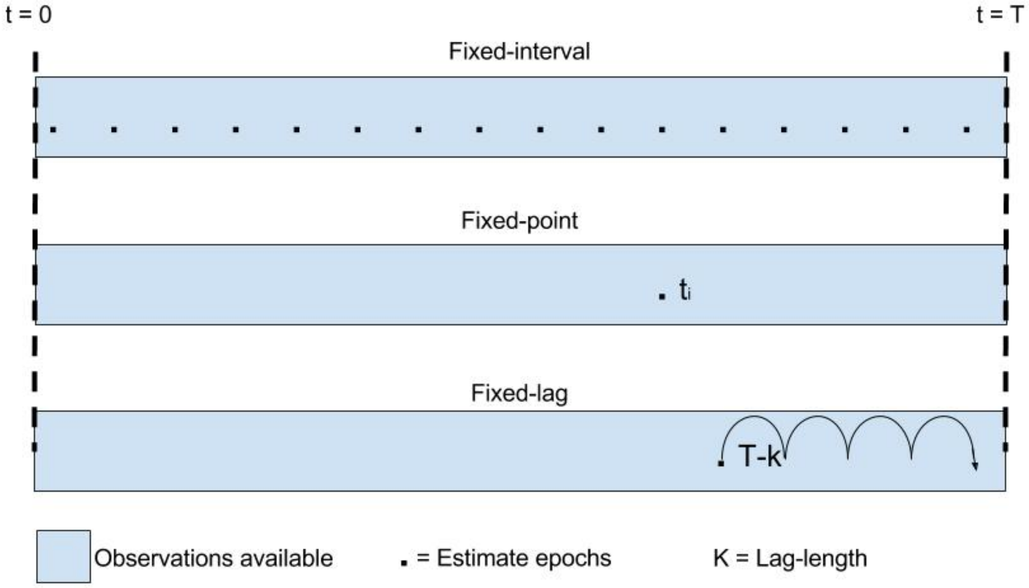


Figure 3.4: Three types of hindcasting. Fixed-interval hindcasting is used to obtain estimates for *all* observed epochs (e.g. what was the missile trajectory given the measurements of the whole flight). Fixed-point hindcasting is used to get the estimate on one single epoch (e.g. what was the temperature on time $t$ given all surrounding observations). Fixed-lag hindcasting gives an estimate of the $k$ epochs before the most recent one and is used for numerically efficient correction of only the latest number of epochs (e.g. what is the current driving speed given the GPS measurements of the last few seconds).

The MAP *filter* estimate is the estimate of the *most recent* observation, at time $T$, based on all measurements, while the MAP *hindcasted* estimate is the estimate of an arbitrary epoch based on all the available observations. It can be seen from this formulation, by substituting $t$ with $T$, that the estimate of the most recent observation is the same for the filter and hindcasting output. This is the reason that hindcasting algorithms are often implemented in a backwards recursion [Rauch et al., 1965], after the forward recursive filter pass that is performed by using Eq. (3.42) and (3.45).

Three types of hindcasting are highlighted [Cosme et al., 2012]:

- **Fixed-interval:** This type gives estimates for all epochs $t$ in the interval $[t_0, T]$. This type is mostly used when an optimal estimate of a dynamic process is needed, incorporating as much data as possible.

- **Fixed-point:** This type calculates the estimate for one specific epoch $t$ in the $[t_0, T]$ interval, every time a new measurement comes in.

- **Fixed-lag:** This type gives an estimate $\hat{\underline{x}}_{T-k|T}$ of a limited fixed amount of epochs, k, in the past. The application of this type is to get a near real time estimate when the filter output is not precise enough, but *fixed-interval* estimation is numerically too expensive.

Figure 3.4 visualises three different types of hindcasting. It shows that *fixed-interval* is most suitable for deformation estimation, since it returns estimates for *all* epochs based on *all* observations. However, this is also the most expensive one in computation time, since all epochs need to be re-estimated once a new observation comes in. *Fixed-lag* smoothing could be of interest for longer time-series, where hindcasting a few years in the past makes no significant change anymore to these estimates.

In the following, we will consider *fixed-interval* hindcasting as the used hindcasting method for this study.

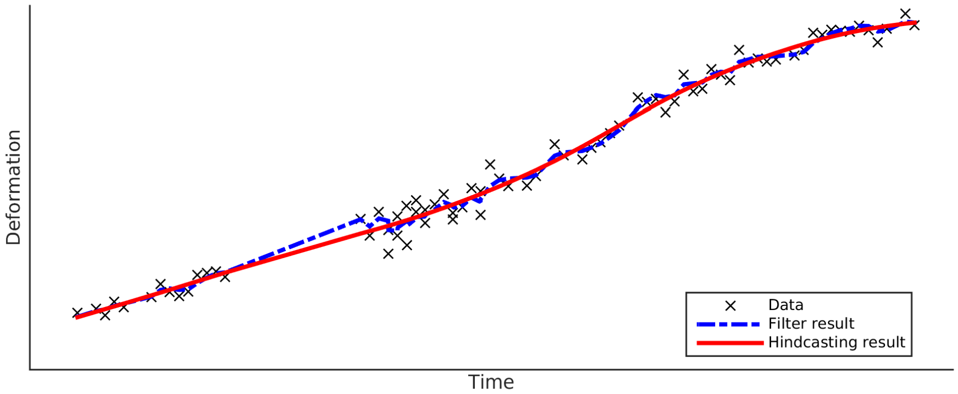


Figure 3.5: Example of results after forward filtering, the subsequent combination of the *time-update* and *measurement-update* equations, and backward hindcasting, the time-reversed updating of previous estimates which were based on limited time series. A certain a priori assumption is made about the smoothness of the physical signal in the VCM of the *difference vector* $Q_d$. The filtered estimate is still more fluctuant than would be expected from the a priori estimated stochasticity of the dynamics—$Q_d$. The hindcasting estimate, however, shows a fit to the data corresponding with the expected signal behaviour. This visualisation is included to show that the used functional model is now based on a smoothness constraint, as described in section 3.2.3 (p. 18).

### 3.3.1 Fixed-interval hindcasting

There are multiple methods to achieve fixed-interval hindcasting. One of the most popular algorithms is that of Rauch, Striebel, and Tung [1965] who presented an efficient algorithm implemented in a backwards recursive pass. The hindcasting is done purely in the parameter space and starts at the most recent filter estimate of epoch $T$. As shown in Eq. (3.46), this filter estimate is equal to the hindcasting estimate and is hence the initialisation of the hindcasting algorithm. From that point onwards, the epochs are processed recursively in a time-reversed order back to the first epoch $t_0$ [Rauch et al., 1965]:

$$\boxed{\begin{array}{l} \text{Iterate}\quad t \rightarrow T:-1:t_0 \\ \qquad\qquad \begin{aligned} G_t &= Q_{\hat{x}_{t|t}} \Phi^{\mathrm{T}}_{t+1,t} Q^{-1}_{\hat{x}_{t+1|t}} \\ \underline{\hat{x}}_{t|T} &= \underline{\hat{x}}_{t|t} + G_t \left( \underline{\hat{x}}_{t+1|T} - \underline{\hat{x}}_{t+1|t} \right) \\ Q_{\hat{x}_{t|T}} &= Q_{\hat{x}_{t|t}} + G_t \left( Q_{\hat{x}_{t+1|T}} - Q_{\hat{x}_{t+1|t}} \right) G_t^{\mathrm{T}} \end{aligned} \end{array}} \ . \tag{3.47}$$

Here $G_t$ is the hindcasting *gain* and $\underline{\hat{x}}_{t|T}$ indicates the estimate of epoch $t$ after hindcasting. The parameter hindcasting is done purely based on filter estimates and VCM's, as shown in Eq. (3.47). This also means that the smoothness of the hindcasting result depends partly on the stochasticity of the difference vector—$Q_{d_t}$, which is implicitly in $Q_{\hat{x}_{t|t}}$ and $Q_{\hat{x}_{t+1|t}}$. An example of the difference between the filter and the hindcasting estimate is displayed in figure 3.5. The hindcasting estimate, indicated by the red line, is based on all available observations, where the filter estimates, indicated by the blue line, for each epoch are only based on previous epochs.

This study uses recursive least-squares, and models deformation as a dynamic process [Teunissen, 2007]. In signal processing another estimator is commonly used, of which the similarities will be discussed in the next section.

## 3.4 Similarities to Kalman filtering

The groundwork for the theory used in this study is based on adjustment theory commonly used in mathematical geodesy. In signal processing one might note the resemblance to the dynamic case in section 3.2.3 with the theory of *Kalman filtering* [Kalman, 1960]. He provides a commonly used filter method in signal processing, because it returns a maximum a posteriori filter estimate when observation noise follows a known Gaussian distribution, and given that the stochasticity of the process noise is known.

It can be shown that *Kalman filtering* is the same as recursive least-squares when describing the signal as a dynamic process (section 3.2.3). A common notation for the Kalman equations is [Bishop and Welch, 2001][1]:

$$\boxed{\begin{aligned} &\textbf{Time-update} \\ \hat{x}_k^- &= A_k \hat{x}_{k-1} \\ P_k^- &= A_k P_{k-1} A_k^\intercal + Q_k \\ &\textbf{Measurement-update} \\ K_k &= P_k^- H_k^\intercal \left( H_k P_k^- H_k^\intercal + R \right)^{-1} \\ \hat{x}_k &= \hat{x}_k^- + K_k \left( z_k - H_k \hat{x}_k^- \right) \\ P_k &= \left( I - K_k H_k \right) P_k^- \end{aligned}} \,, \tag{3.48}$$

where a direct link can be made from Kalman filtering notation, to the notation used in this thesis:

$$\begin{aligned} \hat{x}_k &= \underline{\hat{x}}_{t|t}, \\ \hat{x}_k^- &= \underline{\hat{x}}_{t|t-1}, \\ A_k &= \Phi_{t,t-1}, \\ P_k^- &= Q_{x_{t|t-1}}, \\ P_k &= Q_{x_{t|t}}, \\ H_k &= A_t, \\ R &= Q_{y_t}, \\ z_k &= y_t, \end{aligned} \tag{3.49}$$

resulting in a formulation that already looks similar to Eq. (3.42) and Eq. (3.45) used in the stochastic case:

$$\boxed{\begin{aligned} &\textbf{Time-update} \\ \underline{\hat{x}}_{t|t-1} &= \Phi_{t,t-1} \underline{\hat{x}}_{t-1|t-1} \\ Q_{x_{t|t-1}} &= \Phi_{t,t-1} Q_{x_{t-1|t-1}} \Phi_{t,t-1}^\intercal + Q_{d_t} \\ &\textbf{Measurement-update} \\ K_t &= Q_{x_{t|t-1}} A^\intercal \left( A_t Q_{x_{t|t-1}} A_t^\intercal + Q_{y_t} \right)^{-1} \\ \underline{\hat{x}}_{t|t} &= \underline{\hat{x}}_{t|t-1} + K_t \left( y_t - A \underline{\hat{x}}_{t|t-1} \right) \\ Q_{x_{t|t}} &= \left( I - K_t A \right) Q_{x_{t|t-1}} \end{aligned}} \,. \tag{3.50}$$

The equality of the solution of the stochastic case and that of Kalman filtering can be shown when writing the functional model of the *measurement-update* in Eq. (3.44) in terms of condition equations [Teunissen, 2000, Ch. 3] [Teunissen, 2007, Eq. (3.49)]:

$$\left[ A_t \;\; I \right] E\{\begin{bmatrix} \underline{\hat{x}}_{t|t-1} \\ \underline{y}_t \end{bmatrix}\} = 0 \;\; ; \;\; D\{\begin{bmatrix} \underline{\hat{x}}_{t|t-1} \\ \underline{y}_t \end{bmatrix}\} = \begin{bmatrix} Q_{\hat{x}_{t|t-1}} & 0 \\ 0 & Q_{y_t} \end{bmatrix}, \tag{3.51}$$

---

[1] Note that the control-input matrix and vector of the Kalman filter are omitted from the formulation, since they are irrelevant for this application.

which when solved, gives for the recursive measurement-update equations [Teunissen, 2007, Eq. (6.34)]:

$$\boxed{\begin{aligned} &\textbf{Measurement-update} \\ Q_{v_t} &= Q_{y_t} + A_t Q_{x_{t|t-1}} A_t^{\mathsf{T}} \\ \underline{\hat{x}}_{t|t} &= \underline{\hat{x}}_{t|t-1} + Q_{x_{t|t-1}} A_t^{\mathsf{T}} Q_{v_t}^{-1} \left( y_t - A_t \underline{\hat{x}}_{t|t-1} \right) \\ Q_{x_{t|t}} &= \left( Q_{x_{t|t-1}} - Q_{x_{t|t-1}} A_t^{\mathsf{T}} Q_{v_t}^{-1} A_t Q_{x_{t|t-1}} \right) \end{aligned}} . \tag{3.52}$$

This result can, with simple substitution, be changed to the formulation of (3.50) where the *Kalman Gain*, is given by:

$$K_t = Q_{x_{t|t-1}} A_t^{\mathsf{T}} Q_{v_t}^{-1}, \tag{3.53}$$

and $Q_{v_t}$ is considered to be the variance matrix of the predicted residual:

$$Q_{v_t} = y_t - A_t \underline{\hat{x}}_{t|t-1}. \tag{3.54}$$

Therefore it is shown that recursive least-squares for describing a dynamic process is the same as Kalman filtering.

# 4

# Recursive Least-Squares in Phase Unwrapping

*Recursive least squares can have multiple applications for InSAR. First of all, it can be used for model fitting when new observations come in. This enables efficient updating of model parameters without re-using whole time series in the calculation and moreover, without changing past estimates. However, the scope of this study is phase unwrapping of InSAR time series. Here we show that the set of recursive equations for the static case (Eq. 3.10) can be used for real-time phase unwrapping. This means that the previously adjusted parameters are used for unwrapping the newest observation. First the application for the* static case *is shown, thereafter the theory is extended to be used in the* dynamic case *for parameter and ambiguity estimation. The main goal is to limit the amount of unmodeled deformation and reduce the amount of type-1 errors*[1]*, by using the dynamic functional model in PSI processing.*

## 4.1 Deformation as a static process

Chapter 3 discussed how the recursive equations are derived for the static and dynamic case. This section will explain how the ambiguities of the double-difference phase observations can be resolved using the recursive least squares equations with a static functional model. The term static indicates here that a parametric model is used, for instance a steady-state model, to constrain the ambiguity estimation. The notation in the equations of this chapter is adjusted to comply with Delft PSI literature [e.g. Hanssen, 2004, Samiei-Esfahany, 2017, van Leijen, 2014]:

$\underline{\varphi}_t = \underline{\varphi}^{i,j}_{0,t}$, Wrapped double-difference phase observation of arc $i,j$, on epoch $t$ compared to reference epoch $0$.

$\underline{\phi}_t = \underline{\phi}^{i,j}_{0,t}$, Absolute double-difference phase of arc $i,j$, on epoch $t$ compared to reference epoch $0$.

---

[1] Rejection of scatterers that should be accepted, also referred to as: *false negatives*.

### 4.1.1 Functional and stochastic model

Chapter 3 showed that the functional model can be written in a partitioned way as long as the stochastic model is in the form of a block diagonal matrix:

$$E\{\underbrace{\begin{bmatrix}\underline{\phi}_0\\ \vdots\\ \underline{\phi}_t\\ \vdots\\ \underline{\phi}_T\end{bmatrix}}_{y}\} = \underbrace{\begin{bmatrix}A_0\\ \vdots\\ A_t\\ \vdots\\ A_T\end{bmatrix}}_{A} x \qquad ; \qquad D\{\underbrace{\begin{bmatrix}\underline{\phi}_0\\ \vdots\\ \underline{\phi}_t\\ \vdots\\ \underline{\phi}_T\end{bmatrix}}_{y}\} = \underbrace{\begin{bmatrix}Q_{\phi_0} & & & 0\\ & \ddots & & \\ & & Q_{\phi_t} & \\ & & & \ddots & \\ 0 & & & & Q_{\phi_T}\end{bmatrix}}_{Q_y}, \tag{4.1}$$

where:

- $\underline{\phi}_t$ is short for $\underline{\phi}^{i,j}_{0,t}$, the absolute double phase difference between epoch $0$ and $t$ of arc $i,j$,
- $x$ is the vector of unknown parameters,
- $A_t$ is the $1 \times n$ matrix that transforms the parameters $x$ into the expectation of the double phase difference at epoch $t$: $\underline{\phi}_t$, where $n$ is the number of adjusted parameters.

Eq. (4.1) can be solved recursively based on the estimator of the previous iteration for which the functional and stochastic model reduces to [Teunissen, 2007]:

$$E\{\begin{bmatrix}\underline{\hat{x}}_{T-1}\\ \underline{\phi}_T\end{bmatrix}\} = \begin{bmatrix}I\\ A_T\end{bmatrix} x \qquad ; \qquad D\{\begin{bmatrix}\underline{\hat{x}}_{T-1}\\ \underline{\phi}_T\end{bmatrix}\} = \begin{bmatrix}Q_{\underline{\hat{x}}_{T-1}} & 0\\ 0 & Q_{\underline{\phi}_T}\end{bmatrix}, \tag{4.2}$$

where $\underline{\hat{x}}_{T-1}$ is the $n \times 1$ estimator of the previous iteration of the $n$ adjusted parameters.

Now assume that during parameter and ambiguity estimation, given a steady-state deformation model, the following parameters may need to be estimated:

1. **Velocity parameter** $v$ $[m/y]$. The parameter modelling the steady-state deformation rate.
2. **Height difference** $\Delta H$ $[m]$. This is the height of the scatterer with respect to the reference scatterer, as is discussed in chapter 2.
3. **Temperature dependence** $\alpha$ $[\mathcal{K}^{-1}]$. It is also possible to estimate the temperature dependence $\alpha$ of a scatterer. This is done by assuming that a scatterer is based on an object which reacts linearly on temperature changes, such that the dilation in line-of-sight of the object is given by [Chang, 2015]:

$$\Delta L_t = \alpha \cdot \Delta K_t. \tag{4.3}$$

Here are $\Delta L_t$ and $\Delta K_t$ respectively the change in dilation and temperature at time $t$, compared to the reference epoch, and $\alpha$ is the thermal coefficient of the object. When temperature information is not available or the seasonal trend is not a linear function of temperature, a linearised sinusoidal model can be used, adding two instead of one extra parameter [Kampes, 2005, Van Leijen and Hanssen, 2008].

The vector $x$ is now formed by:

$$x = \begin{bmatrix}v\\ \Delta H\\ \alpha\end{bmatrix} \quad . \tag{4.4}$$

Matrix $A_t$ transforms the parameters $x$ into the expectation of the absolute double difference phase observations $\underline{\phi}_t$. The line-of-sight deformation is estimated in meters and is thus multiplied by the double wavenumber $\frac{-4\pi}{\lambda}$ times $t$, to obtain deformation in radians. The height difference $\Delta H$ is multiplied by the height-to-phase conversion factor (Eq. 2.6), and the temperature coefficient is multiplied by the temperature times the double wavenumber, to obtain the temperature dilation in radians. Matrix $A_t$ is thus given by:

$$A_t = \begin{bmatrix} \frac{-4\pi}{\lambda} t & \frac{-4\pi}{\lambda} \frac{B_t^{\perp}}{R \sin \theta_{inc}} & \frac{-4\pi}{\lambda} \Delta K_t \end{bmatrix} . \tag{4.5}$$

Inserting these expressions in Eq. (4.2) results in the recursive functional model:

$$\begin{gathered} E\{\begin{bmatrix} \underline{\hat{v}}_{t-1} \\ \underline{\hat{\Delta H}}_{t-1} \\ \underline{\hat{\alpha}}_{t-1} \\ \underline{\varphi}_t \end{bmatrix}\} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{-4\pi}{\lambda} t & \frac{-4\pi}{\lambda} \frac{B_t^{\perp}}{R \sin \theta_{inc}} & \frac{-4\pi}{\lambda} \Delta K_t \end{bmatrix} \begin{bmatrix} v_t \\ b \\ \Delta H_t \\ \alpha_t \end{bmatrix} , \\ D\{\begin{bmatrix} \underline{\hat{v}}_{t-1} \\ \underline{\hat{\Delta H}}_{t-1} \\ \underline{\hat{\alpha}}_{t-1} \\ \underline{\varphi}_t \end{bmatrix}\} = \begin{bmatrix} \sigma_v^2 & & & sym \\ \sigma_{v,\Delta H} & \sigma_{\Delta H}^2 & & \\ \sigma_{v,\alpha^T} & \sigma_{\Delta H,\alpha} & \sigma_\alpha^2 & \\ 0 & 0 & 0 & \sigma_{\varphi_t}^2 \end{bmatrix}, \end{gathered} \tag{4.6}$$

where ${}_{t-1}$ denotes the estimator of the previous time step and $\sigma_{v,\Delta H}$ is the covariance between those two estimators at the specified epoch $t-1$.

Equations (4.1) to (4.6) use the absolute double-difference phase $\underline{\phi}_t$. Unfortunately, the available double-difference phase observations, $\underline{\varphi}_t$, are wrapped due to the complex nature of the radar observation. In the next section it is discussed how to treat with the unknown integer amount of phase cycles, that wrapped double-difference phase observations have to shift to equal the absolute double-difference phase observation $\underline{\phi}_t$.

### 4.1.2 Ambiguity resolution

As previously mentioned in chapter 1, the aim of this study is to apply a *temporal smoothness constraint* to aid the ambiguity resolution. To accomplish this, the assumption is made that the expectation of the absolute double-difference phase observation of the next epoch:

$$\underline{\hat{\phi}}_{t|t-1} = A_t \underline{\hat{x}}_{t-1}, \tag{4.7}$$

is within half a wave cycle of the actual observation. In mathematical form:

$$\left| \underline{\phi}_t - \underline{\hat{\phi}}_{t|t-1} \right| < \pi, \tag{4.8}$$

where $\underline{\phi}_t$ is the absolute double-difference phase at epoch $t$ and $\underline{\hat{\phi}}_{t|t-1}$ is the estimator of the absolute double-difference phase observation of the same arc and epoch, based on epoch $t-1$. The static recursive equations are straight forward when the absolute double-difference phases, $\phi_{0:T}$, are known, see Eq. (3.11):

$$\boxed{\begin{aligned} \underline{\hat{x}}_t &= \underline{\hat{x}}_{t-1} + \left(Q_{\hat{x}_{t-1}}^{-1} + A_t^{\intercal} Q_{\phi_t}^{-1} A_t\right)^{-1} A_t^{\intercal} Q_{\phi_t}^{-1} \left(\underline{\phi}_t - A_t \underline{\hat{x}}_{t-1}\right) \\ Q_{\hat{x}_t} &= \left(Q_{\hat{x}_{t-1}}^{-1} + A_t^{\intercal} Q_{\phi_t}^{-1} A_t\right)^{-1} \end{aligned}} . \tag{4.9}$$

Unfortunately, the absolute phases, $\phi_t$, are not known, but the wrapped observations, $\varphi_t$, can also be used as long as the condition of Eq. (4.8) holds. This can be shown by using the fact that:

$$\underline{\phi}_t = \underline{\varphi}_t + a 2\pi, \quad (a \in \mathbb{Z}), \tag{4.10}$$

where $a$ is the unknown amount of integer phase cycles. The wrapping operator $\mathcal{W}\{...\}$ is introduced which is short for [Hanssen, 2001]:

$$\mathcal{W}\{\underline{\phi}\} = \quad \mathrm{mod}_{2\pi}(\underline{\phi} + \pi) - \pi, \tag{4.11}$$

where $\mathrm{mod}_{2\pi}$ is the modulo $2\pi$ operator; in this case transforming $\underline{\phi}$ to the $[-\pi, \pi)$ interval. It can be shown that the wrapped observations can be used when Eq. (4.8) holds:

$$\begin{aligned}
\mathcal{W}\{\underline{\varphi}_t - A_t\underline{\hat{x}}_{t-1}\} &= \mathcal{W}\{\underline{\phi}_t - A_t\underline{\hat{x}}_{t-1} - a2\pi\}, \\
&= \mathrm{mod}_{2\pi}\left(\underbrace{\underline{\phi}_t - A_t\underline{\hat{x}}_{t-1}}_{[-\pi,\pi](4.8)} - a2\pi + \pi\right) - \pi, \\
&= \mathrm{mod}_{2\pi}\left(\underbrace{\underline{\phi}_t - A_t\underline{\hat{x}}_{t-1} + \pi}_{[0,2\pi]} - a2\pi\right) - \pi, \\
&= \cancel{\mathrm{mod}_{2\pi}}\left(\underline{\phi}_t - A_t\underline{\hat{x}}_{t-1} \cancel{+\pi} \cancel{-a2\pi}\right) \cancel{-\pi}, \\
&= \underline{\phi}_t - A_t\underline{\hat{x}}_{t-1}.
\end{aligned} \tag{4.12}$$

Using this expression, the static recursive equations (3.11) can be written as:

$$\boxed{\begin{aligned}
\underline{\hat{x}}_t &= \underline{\hat{x}}_{t-1} + \left(Q_{\hat{x}_{t-1}}^{-1} + A_t^{\intercal} Q_{\phi_t}^{-1} A_t\right)^{-1} A_t^{\intercal} Q_{\phi_t}^{-1} \mathcal{W}\{\underline{\varphi}_t - A_t\underline{\hat{x}}_{t-1}\} \\
Q_{\hat{x}_t} &= \left(Q_{\hat{x}_{t-1}}^{-1} + A_t^{\intercal} Q_{\phi_t}^{-1} A_t\right)^{-1}
\end{aligned}}\,. \tag{4.13}$$

### The unwrapping constraint

The assumption made in Eq. (4.8) puts a constraint on the magnitude of acceptable observation phase noise, $Q_{\phi_t}$, and on the precision of the previous estimator projected onto the current epoch, $Q_{\hat{\phi}_{t|t-1}}$. The stochastic properties of the observation phase noise, for points with high SNR, are in PSI methods often modelled as additive zero-mean and Gaussian distributed [Hanssen, 2004]. Eq. (4.8) can be rewritten to:

$$\begin{aligned}
\left|\underline{\phi}_t - \underline{\hat{\phi}}_{t|t-1}\right| &< \pi, \\
\left|\Delta\underline{\phi}_t\right| &< \pi,
\end{aligned} \tag{4.14}$$

where $\underline{\phi}_t$ is the latest double-difference phase observation, and $\underline{\hat{\phi}}_{t|t-1}$ is the a priori adjustment of the same phase observation, based on the parameter estimator of the previous epoch:

$$\underline{\hat{\phi}}_{t|t-1} = A_t\underline{\hat{x}}_{t-1}. \tag{4.15}$$

The variance-covariance matrix (VCM) of this matrix is given by:

$$Q_{\hat{\phi}_{t|t-1}} = A_t Q_{\hat{x}_{t-1}} A_t^{\intercal}. \tag{4.16}$$

The VCM of the residual double-difference phase observation $\Delta\underline{\phi}_t$ is now given by:

$$Q_{\Delta\phi_t} = Q_{\phi_t} + A_t Q_{\hat{x}_{t-1}} A_t^{\intercal}. \tag{4.17}$$

Figure 4.1 illustrates how the standard deviation of the residual phase observation can be used to determine the percentage of observations that comply with the unwrapping constraint in Eq. (4.14)—the *success-rate*. The probability that:

$$P = P(\Delta\phi_t < -\pi) + P(\Delta\phi_t > \pi), \tag{4.18}$$

given the normal distribution:

$$\mathcal{N}(0, Q_{\Delta\phi_t}), \tag{4.19}$$

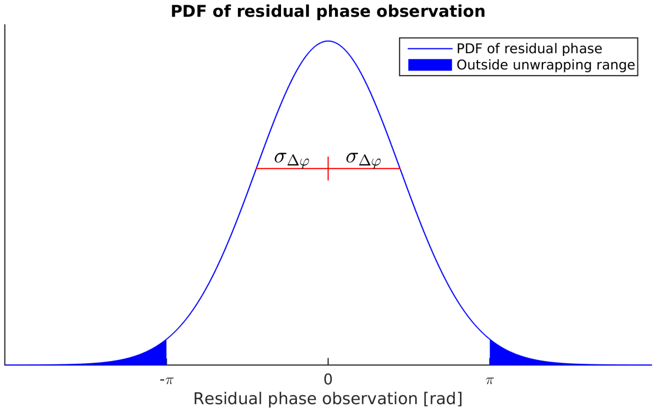


Figure 4.1: Theoretical percentage of observations that comply with the unwrapping constraint of Eq. (4.14). The areas indicated in blue correspond to the percentage of observations that can not be unwrapped to the correct cycle, due to the combination of observation phase noise, $Q_{\phi}$, and the precision of the a priori estimator of the phase observation, $Q_{\hat{\phi}_{t|t-1}}$.

equals the probability that an observation is unwrapped incorrectly. This distribution is zero-mean due to $\underset{-t|t-1}{\hat{\phi}}$ from Eq. (4.15) being the unbiased a-priori estimator of epoch $t$ [Teunissen, 2007]. This probability of incorrect unwrapping is illustrated in blue in figure 4.1. Figure 4.2 shows the unwrapping success-rate as a function of unavailable acquisitions. For TerraSAR-X this means that when the next epoch is the typically $11$ days later, the probability is listed above zero on the x-axis. When the next acquisition is $22$ days later, this indicates that one acquisition was unavailable. The probabilities are computed for 3 noise levels, ranging from $50°$ (X-Band: $2.2mm$, C-Band: $3.9mm$) to $70°$ (X-Band: $3.0mm$, C-Band: $5.4mm$). This is Gaussian distributed additive noise to the double-difference phase observations. The decision of these noise levels will be motivated in chapter 5.3.3 (p. 56). From figure 4.2 it can be deduced that when it is required to have a confidence interval of 2-sigma ($\approx 95\%$), observations can be unwrapped to the correct phase cycle, if the phase noise is $60°$ or lower.

As discussed in the previous paragraph, not all observations will be lying in the required interval and hence some observations might be unresolvable and unwrapped to the wrong phase cycle. Therefore, it is necessary to discuss the effect of those *outliers* on the overall result.

First, when looking at the *gain* part of Eq. (4.13), which determines the contribution of the new observation to the adjustment:

$$\text{Gain} = \left(Q_{\hat{x}_{t-1}}^{-1} + A_t^{\intercal} Q_{\phi_t}^{-1} A_t\right)^{-1} A_t^{\intercal} Q_{\phi_t}^{-1}, \tag{4.20}$$

it can be seen that the gain converges to very low values when the relative contribution of the observation variance to the VCM of adjusted parameters is very small:

$$A_t^{\intercal} Q_{\phi_t} A_t \gg Q_{\hat{x}_{t-1}}. \tag{4.21}$$

From this it can be deduced that the gain is low as long as a large enough amount of observations is already used for computing $Q_{\hat{x}_0}$; the initialisation VCM of the recursive equations (Eq. 3.10). Every iteration decreases the values of $Q_{\hat{x}_t}$, hence when assuming that $Q_{\phi_t}$ does not change significantly over time, Eq. (4.21) remains true. This process is illustrated in figure 4.3, showing that the effect of unresolvable observations on the estimator is small. The gain is a vector of the same size as parameter vector $x$, so for illustration purposes only the values corresponding to deformation are shown.

The second argument for the effect of occasionally unresolvable observations being minimal, is that the distribution of $Q_{\Delta\phi_t}$ is zero-mean. Therefore, the instances of unresolvable observations happen equally on both sides of the correct phase cycle—not biasing the estimator. This effect also means that the apparent observation variance is slightly lower than what is described in $Q_{\phi_t}$. This results in a computed $Q_{\hat{x}_t}$ which is slightly overestimated, nevertheless, given the minor magnitude and the low occurrence of this effect (see figure 4.2), this effect is acknowledged but will be neglected in the calculations.

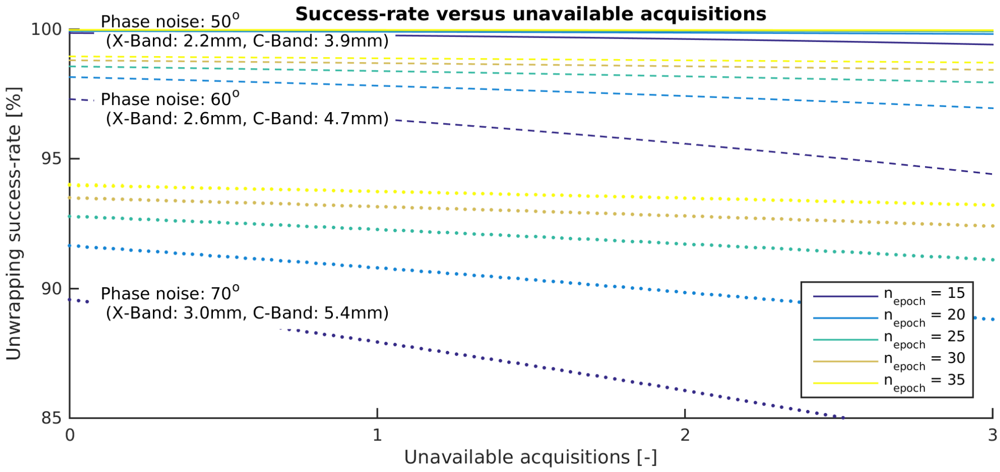


Figure 4.2: Computation of the theoretical probability that an observation does not comply with the unwrapping constraint of Eq. (4.14), given different levels of observation phase noise. Because the value of $Q_{\hat{x}_{t-1}}$ depends on the amount of epochs already processed and decreases over time, different lines are drawn given that a certain amount of observations are already processed. As expected, the probabilities of incorrect unwrapping decrease when more epochs are already processed. On the x-axis the effect is shown when one or more observations are not acquired. This shows that the chance of an observation being unwrapped to the wrong phase cycle increases when $\Delta t$ increases. The displayed behaviour is the same for different kind of satellites, because the noise level in degrees is independent on the wavelength and the repeat-cycle $\Delta t$ moreover scales $Q_{\hat{x}_{t-1}}$ equally.

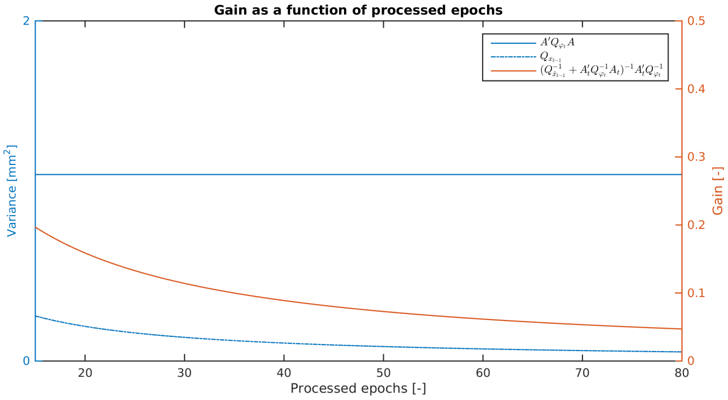


Figure 4.3: Decaying of the gain as a function of time (red line). The gain is a vector of the same size as parameter vector $x$, determining how much the parameter adjustment is influenced by the new phase residual $\Delta\underline{\hat{\phi}}_t$. A large gains means that the new observation contributes greatly to the parameter adjustment, a small gain means the contrary. Only one of the gain values is shown to indicate the principle that the deformation adjustment, and with that the phase unwrapping, is barely influenced by an occasional unresolvable residual phase observation. The blue lines indicate how the different parts of the gain function compare to each other, to motivate the claim made in Eq. 4.21. Only the variance value related to deformation is plotted, as is similarly done for the plotted gain.

## 4.2 Deformation as a dynamic process

As discussed in section 3.2.3, treating surface deformation in a parametric way is a stringent assumption that puts a large constraint on the final unwrapping result. Especially on longer time series it does become more difficult to define a functional model, able to describe the sometimes complex and time-variable behaviour of surface deformation. Usually a steady-state (constant velocity) deformation model is applied in temporal unwrapping algorithms as they provide maximum redundancy, but algorithms also allow for more complex deformation models like seasonal, temperature, breakpoint, jump and higher order polynomial models [Chang and Hanssen, 2016, Samiei-Esfahany, 2008]. While these models can capture some of the dynamics in deformation time series, there are still a few disadvantages. First, they are only applied when there is a priori information indicating the need for such models. Second, using more models for unwrapping will increase the computation time, depending on the used method, due to the increased amount of unknowns to be estimated. Third, the parameterisation of a deformation process may not be stationary, that is, it may change over time. Particularly considering human interaction with the subsurface this will be a problem. Another disadvantage is that some models, like breakpoint and jump models, need a priori information about the date that a change happened, or when this information is unavailable need to test the whole spectrum of possible dates. Multiple approaches to adaptively select and test more complex deformation models during unwrapping are proposed in van Leijen [2014], Van Leijen and Hanssen [2007], while the testing of a large library of deformation models in an a posteriori step is performed in Chang and Hanssen [2016]. However, these approaches are still relying on predefined deformation patterns.

All these problems may be mitigated when deformation is treated in a non-parametric way, using a priori information on the expected smoothness of the deformation process. For this study it is assumed that the behaviour of deformation on short time scales, e.g. months, approaches that of a steady-state system. On longer time scales however, the steady-state can be influenced by an acceleration term, required to change the subsidence rate over time. This temporal smoothness constraint can be applied by combining the theory from section 3.2.3, making some assumptions about the expected temporal smoothness of the signal.

### 4.2.1 Functional and stochastic model

The mathematical framework is based on the Gauss-Markov model and the partitioned functional model, that were given in Eq. (3.36), showing that any difference to the modelled deformation is given by the *difference vector* $d_t$ [Teunissen, 2007]:

$$E\{\begin{bmatrix}\underline{\phi}_0\\ \vdots\\ \underline{\phi}_t\\ \vdots\\ \underline{\phi}_T\end{bmatrix}\} = \begin{bmatrix}A_0 & & & & \\ & \ddots & & & \\ & & A_t & & \\ & & & \ddots & \\ & & & & A_T\end{bmatrix}\begin{bmatrix}x_0\\ \vdots\\ x_t\\ \vdots\\ x_T\end{bmatrix} + \begin{bmatrix}A_0 d_{0,T}\\ \vdots\\ A_t d_{t,T}\\ \vdots\\ A_T d_{T,T}\end{bmatrix} \;\; ; \;\; D\{\begin{bmatrix}\underline{\phi}_0\\ \vdots\\ \underline{\phi}_t\\ \vdots\\ \underline{\phi}_T\end{bmatrix}\} = \begin{bmatrix}Q_{\phi_0} & & & & 0\\ & \ddots & & & \\ & & Q_{\phi_t} & & \\ & & & \ddots & \\ 0 & & & & Q_{\phi_T}\end{bmatrix}, \tag{4.22}$$

with:

$$\begin{bmatrix}x_0\\ \vdots\\ x_t\\ \vdots\\ x_T\end{bmatrix} = \begin{bmatrix}\Phi_{0,T}\\ \vdots\\ \Phi_{t,T}\\ \vdots\\ \Phi_{T,T}\end{bmatrix} x_T, \tag{4.23}$$

where:

- $\Phi_{t,T}$ is the transition matrix from Eq. (3.20), transforming parameter vector $x_T$ to its expected value $x_t$ at epoch $t$.
- $A_t$ is the system matrix that transforms the unknown parameters to the expectation of the absolute double-difference phase $\underline{\phi}_t$.

- $d_{t,T}$ is the *difference vector* that describes how the unknown parameters of epoch $t$ differ from epoch T. By multiplying this *difference vector* times $A_t$, the contribution to the expectation of the absolute double-difference phase of epoch $t$ is computed.

Teunissen [2007] showed that when the *difference vector*, $d_t$, is treated as a pseudo-observation, Eq. (4.22) can be rewritten to:

$$E\{\begin{bmatrix} \underline{\phi}_0 \\ \underline{d}_1 \\ \underline{\phi}_1 \\ \vdots \\ \underline{d}_T \\ \underline{\phi}_T \end{bmatrix}\} = \begin{bmatrix} A_0 & & & & 0 \\ -\Phi_{1,0} & I & & & \\ & A_1 & & & \\ & & \ddots & & \\ & & & -\Phi_{T,T-1} & I \\ 0 & & & & A_T \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{T-1} \\ x_T \end{bmatrix};$$
$$D\{\begin{bmatrix} \underline{\phi}_0 \\ \underline{d}_1 \\ \underline{\phi}_1 \\ \vdots \\ \underline{d}_T \\ \underline{\phi}_T \end{bmatrix}\} = \begin{bmatrix} Q_{\phi_0} & & & & & 0 \\ & Q_{d_1} & & & & \\ & & Q_{\phi_1} & & & \\ & & & \ddots & & \\ & & & & Q_{d_T} & \\ 0 & & & & & Q_{\phi_T} \end{bmatrix}, \tag{4.24}$$

where:

- $x_t$ are the unknown parameters of epoch $t$. Compared to the previous section, 4.1, the parameterisation of the vector of unknown parameters has changed to allow for changing parameters over time. This is necessary because the time-varying pseudo-observation $\underline{\mathrm{d}}$ is included in the parameterisation. The new parameterisation allows for a changing deformation rate over time, hence enabling deformation modelling based on expected smoothness. The functional model of a single absolute double-difference phase is now given by:

  $$\underline{\phi}_t = A_t x_t + \epsilon, \tag{4.25}$$

  where $\epsilon$ indicates the observation noise. Considering the steady-state assumption for short time spans, the parameter vector now consists of:

  $$x_t = \begin{bmatrix} D_t \\ v_t \\ \Delta H_t \\ \alpha_t \end{bmatrix}, \tag{4.26}$$

  where $D_t$ is the deformation compared to the reference epoch, $v_t$ is the deformation rate of epoch $t$, $\Delta H$ is the height difference with respect to the reference scatterer, and $\alpha_t$ is the temperature correlation coefficient. While including the deformation rate $v_t$ in this vector might seem redundant since $D_t$ already describes the deformation signal, it is of importance in the *time-update* (see Eq. 3.42) step to provide an a priori estimate for the next epoch.

- $A_t$, as described in Eq. (4.5), also changes due to the new formulation of $x_t$:

  $$A_t = \begin{bmatrix} \frac{-4\pi}{\lambda} & 0 & \frac{-4\pi}{\lambda} \frac{B_t^{\perp}}{R \sin \theta_{inc}} & \frac{-4\pi}{\lambda} \Delta K_t \end{bmatrix}, \tag{4.27}$$

  where the first entry, the double wavenumber, transforms the deformation $D_t$ into radians. The second entry is $0$, because the *time-update* is made in the transition matrix and $v_t$ is not necessary in the expectation of $\underline{\phi_t}$ since all deformation is captured in $D_t$. The third entry is the H2PH conversion factor and the last entry is the double wavenumber multiplied by the relative temperature change between the current and reference epoch.

- $\Phi_{t,t-1}$ is the transition matrix from Eq. (3.20), transforming parameter vector $x_t$ to $x_{t-1}$. When extended for the two extra unknown parameters, $\Delta H$ and $a_t$, of which the expectation does not change over time, $\Phi_{t,t-1}$ becomes:

$$\Phi_{t_1,t_2} = \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.28}$$

- $\underline{d}_t$ is short for $\underline{d}_{t-1,t}$, indicating the *difference vector* of epoch $t$, compared to epoch $t-1$. The *difference vector* is added as a pseudo-observation in the functional model of Eq. (4.24) with a zero-mean expectation [Teunissen, 2007].

- $Q_{d_t}$ is the VCM that describes the stochasticity of the pseudo-observation $d_{t-1,t}$. This implicitly contains the expected smoothness of the deformation signal and should be obtained using a priori knowledge, for instance obtained with expert elicitation. This topic will be discussed in more detail in section 4.2.3.

The recursive equations needed to solve Eq. (4.24) are very similar to Eq. (4.13) [Teunissen, 2007]:

$$\boxed{\begin{aligned} &\qquad\text{Time-update equations} \\ \underline{\hat{x}}_{t|t-1} &= \Phi_{t,t-1}\underline{\hat{x}}_{t-1|t-1} + \underline{d}_t \\ Q_{\hat{x}_{t|t-1}} &= \Phi_{t,t-1}Q_{\hat{x}_{t-1|t-1}}\Phi^{\intercal}_{t,t-1} + Q_{d_t} \\ &\qquad\text{Measurement-update equations} \\ \hat{x}_{t|t} &= \hat{x}_{t|t-1} + (Q^{-1}_{\hat{x}_{t|t-1}} + A^{\intercal}_t Q^{-1}_{\phi_t} A_t)^{-1} A^{\intercal}_t Q^{-1}_{\phi_t}\ \mathcal{W}\{\underline{\varphi}_t - A_t\hat{x}_{t|t-1}\} \\ Q_{\hat{x}_{t|t}} &= (Q^{-1}_{\hat{x}_{t|t-1}} + A^{\intercal}_t Q^{-1}_{\phi_t} A_t)^{-1} \end{aligned}}. \tag{4.29}$$

The difference between Eq. (4.13) and (4.29) is in the time-update part of the equations, where the difference vector, $d_t$, as well as its stochastic properties, $Q_{d_t}$, are both taken into account. Without those two terms the same solution would be yielded as in the static case in Eq. (4.13). Similarly as in the static case, the wrapping operator $\mathcal{W}$ is used in the *measurement-update*. It will be shown in section 4.2.4 (p. 42) under which conditions usage of the wrapping operator is allowed.

### 4.2.2 Stochastics of the observation vector

The variance-covariance matrix, $Q_\phi$, describes the statistical properties of the double difference observations $\underline{\varphi}^{0,t}_{i,j}$. This matrix can be derived from the original vector of SLC observations as done by Hanssen [2004]. The vector of wrapped SLC phase observations for two points, $i$ and $j$, is given by:

$$\begin{aligned} \underline{\psi} &= \begin{bmatrix} \underline{\psi}^i \\ \underline{\psi}^j \end{bmatrix} = \begin{bmatrix} \underline{\psi}^i_0 & \cdots & \underline{\psi}^i_T & \underline{\psi}^j_0 & \cdots & \underline{\psi}^j_T \end{bmatrix}^{\intercal}, \\ D\{\underline{\psi}\} &= Q_\psi = Q_{\text{noise}} + Q_{\text{atmo}} + Q_{\text{defo}}. \end{aligned} \tag{4.30}$$

The expression of $Q_\psi$ in $Q_{\text{noise}} + Q_{\text{atmo}} + Q_{\text{defo}}$ is an addition by van Leijen et al. [2006] where:

- $Q_{\text{noise}}$ is a block-diagonal matrix, describing temporal decorrelation and thermal noise.

- $Q_{\text{atmo}}$ is a block-diagonal matrix, describing the contribution of atmospheric and orbit errors. These atmospheric errors are correlated in space but not in time [Hanssen, 2001].

- $Q_{\text{defo}}$ is describes the unmodeled deformation, correlated in time and space.

The full covariance matrix for the vector of wrapped SLC phase observations $\underline{\psi}$ is given by [Hanssen, 2004]:

$$D\left\{\begin{bmatrix} \underline{\psi}^i_0 \\ \vdots \\ \underline{\psi}^i_T \\ \underline{\psi}^j_0 \\ \vdots \\ \underline{\psi}^j_T \end{bmatrix}\right\} = \begin{bmatrix} \sigma^2_{\psi^i_0} & & & & & \\ \vdots & \ddots & & & sym & \\ \sigma_{\psi^i_0,\psi^i_T} & \cdots & \sigma^2_{\psi^i_T} & & & \\ \sigma_{\psi^i_0,\psi^j_0} & \cdots & \sigma_{\psi^i_T,\psi^j_0} & \sigma^2_{\psi^j_0} & & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \\ \sigma_{\psi^i_0,\psi^j_T} & \cdots & \sigma_{\psi^i_T,\psi^j_T} & \sigma_{\psi^j_0,\psi^j_T} & \cdots & \sigma^2_{\psi^j_T} \end{bmatrix}, \tag{4.31}$$

where:

- $\sigma^2_{\psi^i_t}$ is the variance of one phase observation, consisting of: scattering noise, thermal noise, atmospheric signal, orbital signal and unmodeled deformation.
- $\sigma_{\psi^i_{t_1},\psi^i_{t_2}}$ is the covariance of one point between epoch $t_1$ and $t_2$. Its main contributor is the coherence change, but also unmodeled deformation contributes to this factor.
- $\sigma_{\psi^i_{t_1},\psi^j_{t_2}}$ is the covariance between two points at two different epochs. The only effector of this value is the spatio-temporal correlation of unmodeled deformation.
- $\sigma_{\psi^i_t,\psi^j_t}$ is the covariance between two points at the same epoch and is here dominated by the spatial correlated contributions of the atmospheric signal, orbital signal and unmodeled deformation.

The variances and covariances of the double difference phase observations can now be obtained by error propagation [Kampes, 2006, van Leijen et al., 2006], with:

$$\underline{\varphi}^{i,j}_{0,t} = (\underline{\psi}^i_t - \underline{\psi}^i_0) - (\underline{\psi}^j_t - \underline{\psi}^j_0), \tag{4.32}$$

the variances are given by:

$$\begin{aligned}\sigma^2_{\varphi^{i,j}_{0,t}} = &\sum_{p=i,j}\sum_{q=0,t}\sigma^2_{\psi^p_q} \\ &- 2\Big(\underbrace{\sigma_{\psi^i_0,\psi^j_0}}_{\substack{C_{\text{atmo},\psi_0}(l^{ij}) \\ + \\ C_{\text{defo},\psi_0}(l^{ij})}} + \underbrace{\sigma_{\psi^i_t,\psi^j_t}}_{\substack{C_{\text{atmo},\psi_t}(l^{ij}) \\ + \\ C_{\text{defo},\psi_t}(l^{ij})}} + \underbrace{\sigma_{\psi^i_0,\psi^i_t}}_{C_{\text{defo},\psi^i}(\Delta t_{0,t})} + \underbrace{\sigma_{\psi^j_0,\psi^j_t}}_{C_{\text{defo},\psi^j}(\Delta t_{0,t})}\Big) + 4\underbrace{\sigma_{\psi^i_0,\psi^j_t}}_{C_{\text{defo}}(l^{ij},\Delta t_{0,t})},\end{aligned} \tag{4.33}$$

where:

$$\sigma^2_{\psi^p_q} = \sigma^2_{\text{scat},\psi^p} + \sigma^2_{\text{atmo},\psi_q} + \sigma^2_{\text{defo},\psi^p_q}. \tag{4.34}$$

The underbraces indicate the contributions of all components, in which the capital $C$ stands for a covariance function. It can be seen that unmodeled deformation has a large influence on the variance of the DD observation. In this study we assume that all of the unmodeled deformation is captured in the difference vector $\underline{d}$ of the functional model (Eq. 4.22) and that its statistics are accounted for in the covariance matrix $Q_d$. The expression for the variance can therefore be simplified to:

$$\sigma^2_{\varphi^{i,j}_{0,t}} = \sum_{p=i,j}\sum_{q=0,t}\left(\sigma^2_{\text{scat},\psi^p} + \sigma^2_{\text{atmo},\psi_q}\right) - 2\Big(\underbrace{\sigma_{\psi^i_0,\psi^j_0}}_{C_{\text{atmo},\psi_0}(l^{ij})} + \underbrace{\sigma_{\psi^i_t,\psi^j_t}}_{C_{\text{atmo},\psi_t}(l^{ij})}\Big), \tag{4.35}$$

where:

- $\sigma^2_{\text{scat},\psi^p}$ is the variance of the scattering phase noise for the whole time series of point $p$. This term can be approximated by using the normalized amplitude dispersion $D_{\text{amp},\psi^p}$ of that point [Ferretti et al., 2001].

- $\sigma^2_{\text{atmo},\psi_t}$ is the variance of the atmospheric signal at epoch $t$, and is commonly approximated during the APS estimation via the nugget of the experimental variogram.
- $C_{\text{atmo},\psi_t}(l^{ij})$ is the covariance at epoch $t$ as function of the arc-length, $l^{ij}$, given by the computed experimental variogram of the APS estimation.

### Practical simplifications

In Ketelaar [2009], the covariances of $Q_\phi$ are shown to be dominated by unmodeled deformation. In this study it is assumed that all deformation is *modelled*, making $Q_\phi$ in Eq. (4.29) a convenient diagonal matrix. The diagonal terms of this matrix are variances of the atmospheric signal and the temporal decorrelation. In the first parameter adjustment, the experimental variograms cannot yet be computed because they rely on the results of this first parameter adjustment. As solution, the arcs of the network are chosen to be sufficiently short, such that the influence of atmosphere over these arcs is assumed to be small and zero-mean distributed. After APS estimation it is assumed that most atmospheric signal is removed and only the temporal decorrelation term in Eq. (4.35) is left. For both situations the temporal decorrelation is the dominant factor, which, with the method of normalized amplitude dispersion [Ferretti et al., 2001] can be approximated with a variance factor:

$$\begin{aligned} Q_\phi &= \left(2\sigma^2_{\text{scat},\psi^i} + 2\sigma^2_{\text{scat},\psi^j}\right) I_{T\times T}, \\ &= \sigma^2_{i,j} I_{T\times T}, \end{aligned} \tag{4.36}$$

where $I_{T\times T}$ is the identity matrix of size equal to the amount of observations, $\sigma^2_{\text{scat},\psi^i}$ is the variance contribution of temporal decorrelation approximated by normalized amplitude dispersion, and $\sigma^2_{i,j}$ is the total a priori variance factor for arc $i,j$. The downside of determining $Q_\phi$ with this method is that the normalized amplitude dispersion is only a precise and accurate estimator of scattering noise for points with a rather low amplitude dispersion, i.e. $< 0.25$ [Ferretti et al., 2001]. Moreover, the implementation of this parameter and ambiguity estimator would benefit from a fixed variance factor for all evaluated arcs, to improve computational efficiency by vectorisation. In chapter 5 it will hence be investigated if a fixed a priori variance factor has a large influence on the parameter and ambiguity estimation. If this is not the case, this method could be used to provide an a priori variance factor to approximate $Q_\phi$. After unwrapping an a posteriori variance factor $\hat{\sigma}^2_{i,j}$ can be calculated via [Kampes, 2005]:

$$\hat{\sigma}^2 = \frac{\underline{\hat{e}}^\intercal Q_\phi^{-1} \underline{\hat{e}}}{r}, \tag{4.37}$$

with:

$$\underline{\hat{e}} = \underline{\hat{\phi}} - \underline{\phi}, \tag{4.38}$$

where $\underline{\hat{e}}$ is the vector of observation residuals, $\underline{\hat{\phi}}$ is the vector of adjusted absolute double-difference phase observations, and $\underline{\phi}$ is the vector of unwrapped double-difference phase observations. $r$ stands for the redundancy of the functional model—the number of observations minus amount of adjusted parameters. This a posteriori variance factor is an indication of how well $Q_\phi$ describes the dispersion of the observations, assuming that the functional model and stochasticity of the difference vector, $Q_d$, are correct.

For the *static case* an a posteriori correction to the VCM of adjusted parameters can be made, given that the a priori VCM of adjusted parameters is given by:

$$Q_{\hat{x}} = (A^\intercal Q_\phi^{-1} A)^{-1}, \tag{4.39}$$

where $Q_\phi$ is estimated a priori by:

$$Q_\phi = \sigma^2_{i,j} I. \tag{4.40}$$

Using the a posteriori variance factor of Eq. (4.37), $Q_\phi$ can be recalculated using the a posteriori corrected $Q_\phi$:

$$Q_{\hat{\phi}} = \hat{\sigma}^2 Q_\phi. \tag{4.41}$$

Equations (4.39) and (4.41) are combined to create an expression for the a posteriori VCM of adjusted parameters:

$$\begin{aligned} Q_{\hat{x},\text{a-posteriori}} &= (A^{\mathsf{T}} Q_{\hat{\phi}}^{-1} A)^{-1}, \\ &= \hat{\sigma}^2 (A^{\mathsf{T}} Q_{\phi} A)^{-1}, \\ &= \hat{\sigma}^2 Q_{\hat{x}}. \end{aligned} \tag{4.42}$$

The same can be done for the *dynamic case*, however, the above-mentioned direct method would lead to an too optimistically $Q_{\hat{x}}$, due to the addition of $Q_d$ in the *measurement-update* of Eq. (4.29). The iteration from Eq. (4.29) for computing $Q_{\hat{x}_{t|t-1}}$ has to be rerun with an a posteriori $Q_{\hat{\phi}}$ from Eq. (4.41) to obtain the correct $Q_{\hat{x}}$, adding computational effort.

### 4.2.3 Stochastics of the difference vector

The *difference vector*, $\underline{d}$, is added in the functional model of the dynamic case in Eq. (4.24) as a pseudo-observation. This pseudo-observation needs to be stochastically constrained, using $Q_d$, in order to be used. A popular way to do this for applications with steady-state behaviour on short time-scales, is to assume that the steady-state is influenced by a random zero-mean acceleration term over time [Bar-Shalom et al., 2004, Genovese, 2001, Teunissen, 2007]. When $Q_d$ is set to be infinitely small, or zero, it means that the dispersion of the pseudo-observation $\underline{d}$ equals zero. In other words, $\underline{d}$ is deterministic with value $0$, because of the zero-mean assumption. In this case the solution becomes equal to that of the recursive equations (4.13) of the static deformation case in section 4.1.

On the other hand, when $Q_d$ is set infinitely large, $Q_{\hat{x}}$ will also become infinitely large because the time-update equation in Eq. (4.29) adds $Q_d$ completely to $Q_{\hat{x}_{t|t-1}}$. As discussed before in chapter 4.1.2, the gain of the *measurement-update* equation will be close to $0$ as long as the contribution of observation variance to the VCM of adjusted parameters is small:

$$A_t^{\mathsf{T}} Q_{\phi_t}^{-1} A_t \ll Q_{\hat{x}_{t|t-1}}^{-1}. \tag{4.43}$$

However, the opposite is the case when $Q_d$ is set to infinite, meaning that the *gain* is close to $1$ and thus ignoring the a priori estimate from the time-update equations: $\underline{\hat{x}}_{t|t-1}$. This means that the next deformation estimate is just the next phase observation, unwrapped to the nearest phase-cycle, completely ignoring the a priori estimate from the *time-update* step, and with that ignoring the steady-state assumption. It follow from this that $Q_{d_t}$ can be regarded as a smoothness parameter, indicating how temporally smooth the deviation from the steady-state deformation model may be.

As stated in the above paragraph, it is necessary to describe the stochasticity of the difference vector $\underline{d}$. Teunissen [2007] describes two methods for doing this. Both methods use the fact that changes in the steady-state model between two epochs are caused by an acceleration term. Hence. instead of the steady-state propagation from epoch $t_1$ to $t_2$:

$$E\{\underline{v}_2\} = E\{\underline{v}_1\}, \tag{4.44}$$

the expectation of $\underline{v}_2$ changes to:

$$E\{\underline{v}_2\} = E\{\underline{v}_1\} + E\{\underline{a}_{1,2}\}(t_2 - t_1). \tag{4.45}$$

The first way of describing pseudo-observation $\underline{d}$ regards the acceleration term in the difference vector, as a Gaussian distributed zero-mean random acceleration. The zero-mean assumption is convenient, because it turns the expectation of the difference vector $E\{\underline{d}\}$ to $0$, simplifying the *time-update* step.

This study focus on deformation signals that show a slow temporal-smooth signal, which on short time scales (e.g. months), approximates stead-state behaviour. On longer time scales however, the steady-state behaviour can change under influence of external forces (e.g. loading changes, subsurface mechanics etc.), which induces acceleration per Newton's second law. This acceleration is a temporal-slow process. Describing this acceleration with a zero-mean Gaussian distribution would hence not be accurate.

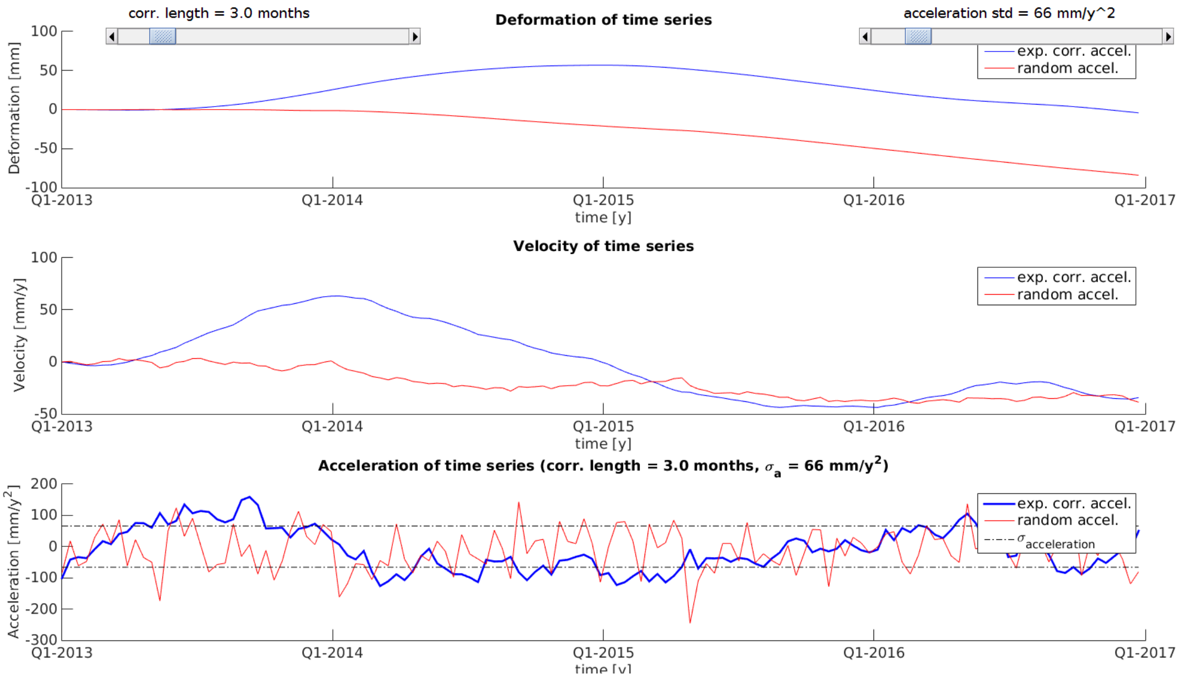


Figure 4.4: Two time series, displayed in red and blue, are simulated by adding random acceleration to a not depicted stable (0mm/y) time series. The used acceleration can be seen in the lowest sub-plot. The resulting velocity is shown in the middle and the final deformation pattern in the first plot. The time series displayed in red is generated by uncorrelated normally distributed acceleration and the time series displayed in blue is generated by temporally correlated acceleration using an exponential function. The correlation of the acceleration can be clearly recognized in the third subplot where the blue line deviates more smoothly over time between positive and negative values, while showing a similar standard deviation as the red line. The used acceleration standard deviation is shown as the black dashed-dotted line.

Hence, a second way of describing $\underline{d}$ is proposed [Teunissen, 2007]. A temporal correlation to the acceleration process is introduced, such that the acceleration will only change slowly over time. The diagonal form of the partitioned stochastic model in Eq. (4.24) makes it difficult to implement time correlation, however, a solution is available for exponential correlation [Teunissen, 2007].

### Correlated versus uncorrelated acceleration

The difference between the two models of acceleration is illustrated in figure 4.4. The model indicated by the red line is accomplished by a adding a Gaussian distributed zero-mean acceleration to a stable (0mm/y deformation) point. The model indicated by the blue line is similarly generated with the exception that the acceleration is exponentially correlated in time, as evident from the third sub-plot. In the plot the values of the uncorrelated time series acceleration can vary tremendously between two epochs, while the correlated time series show a more smooth transition from a positive to a negative acceleration. This smooth behaviour is more realistic for the slow dynamic subsurface processes that are of interest in this study. The smooth acceleration also causes the deformation velocity to change gradually over time which can be frequently observed in longer time series, where the deformation pattern is more likely to be no longer stationary.

### Exponential correlated acceleration in time [Teunissen, 2007]

This section will explain how the stochastics of the *difference vector* should be modelled to achieve exponential correlated acceleration in time for the recursive mathematical functional model in Eq. (4.24). Given exponentially correlated acceleration with auto-covariance function [Teunissen, 2007]:

$$\sigma_{aa}(\Delta t) = \sigma^2_{\text{accel}} e^{-\frac{\Delta t}{L_{\text{corr}}}};\ L_{\text{corr}} > 0, \tag{4.46}$$

which is plotted in figure 4.5. The variables of this function are:

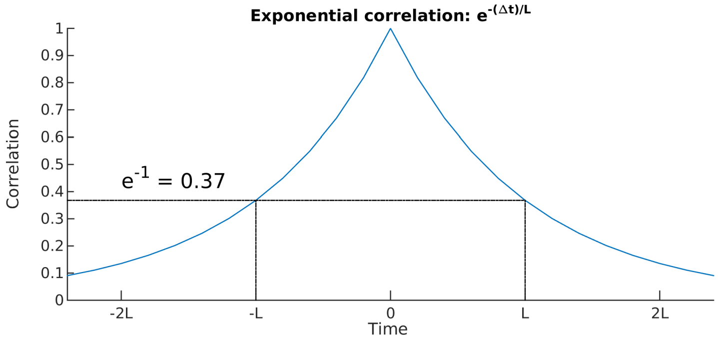


Figure 4.5: Exponential correlation as a function of $\Delta t$. It can be seen from this figure that if for instance the correlation length $L_{\text{corr}} = 5$ months, the correlation of two values with a $\Delta t = 5$ months is $e^{-1} \approx 0.37$. After two times the correlation length, the correlation is only $e^{-2} \approx 0.14$.

- **Auto-covariance** $\sigma_{aa}$ of the acceleration between two epochs. This value describes the correlation of acceleration as a function of time difference $\Delta t$.

- **Time separation** $\Delta t$ between two epochs.

- **Variance** $\sigma^2_{\textbf{accel}}$ of the acceleration of the system, to be estimated a priori using e.g. expert elicitation, which will be discussed in more detail on page 42. This term describes the magnitude of the expected acceleration. In an area where no known effects are expected to change the steady-state behaviour, this value will be very small, e.g. $\sigma = 0.1mm/y^2$. In an area with variable sub-surface dynamics, like during variations in gas extraction, the expected magnitude of acceleration will be larger, e.g. $\sigma = 10mm/y^2$.

- **Correlation length** $L_{\textbf{corr}}$ of the system acceleration. This parameter needs to be estimated a priori which will be discussed in more detail on page 42. The influence of the correlation length on the actual correlation is illustrated in figure 4.5. After passing a time period, equal to the correlation length, the correlation between the two acceleration values is $e^{-1} = 0.37$. After passing a time period twice this length, the correlation is reduced to $e^{-2} = 0.14$. In practice this means that with a correlation length of for instance $L_{\text{corr}} = 5$ months, the correlation would be significantly reduced after $5$ months, and almost negligible after $10$ months. The value $L_{\text{corr}} = 5$ months would result in a correlation between two typical TerraSAR-X observations ($\Delta t = 11$days) of $0.93$.

The addition of correlated acceleration results in the need to add this parameter to the vector of unknowns and the difference vector $\underline{d}$ [Teunissen, 2007]:

$$\underline{d}_t = \begin{bmatrix} \underline{D}_t \\ \underline{v}_t \\ \underline{a}_t \end{bmatrix}, \tag{4.47}$$

where $\underline{D}_t$, $\underline{v}_t$ and $\underline{a}_t$ indicate the deformation, deformation rate, and deformation acceleration difference compared to the previous epoch $t-1$. When using auto-covariance function (4.46), $Q_{d_t}$ is given by [Teunissen, 2007, Eq. 6.94]:

$$Q_{d_t} = \begin{bmatrix} \sigma^2_{D_t} & & symm. \\ \sigma_{D_t,v_t} & \sigma^2_{v_t} & \\ \sigma_{D_t,a_t} & \sigma_{v_t,a_t} & \sigma^2_{a_t} \end{bmatrix},$$

$$= \sigma^2_{\text{accel}} \begin{bmatrix} q_{11} & & symm. \\ q_{21} & q_{22} & \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \begin{cases} q_{11} = 2L^3 \left[\Delta t - \dfrac{\Delta t^2}{L} + \dfrac{\Delta t^3}{3L^2} - 2e^{-\frac{\Delta t}{L}}\Delta t + \dfrac{L}{2}(1 - e^{-2\frac{\Delta t}{L}})\right] \\ q_{21} = 2L^2 \left[-\Delta t + \dfrac{1}{2L}\Delta t^2 + e^{-\frac{\Delta t}{L}}\Delta t - Le^{-\frac{\Delta t}{L}} + \dfrac{L}{2}(1 + e^{-2\frac{\Delta t}{L}})\right] \\ q_{31} = 2L \left[-e^{-\frac{\Delta t}{L}}\Delta t + \dfrac{L}{2}(1 - e^{-2\frac{\Delta t}{L}})\right] \\ q_{22} = 2L \left[\Delta t - \dfrac{3L}{2} + 2Le^{-\frac{\Delta t}{L}} - \dfrac{L}{2}e^{-2\frac{\Delta t}{L}}\right] \\ q_{32} = 2L \left[-e^{-\frac{\Delta t}{L}} + \dfrac{1}{2}(1 + e^{-2\frac{\Delta t}{L}})\right] \\ q_{33} = \left[1 - e^{-2\frac{\Delta t}{L}}\right] \end{cases},$$

(4.48)

where $\Delta t$ indicates the absolute time difference between epoch $t$ and $t-1$. Because of the addition of this exponentially correlated acceleration, the transition matrix also needs to be adjusted. [Teunissen, 2007, Eq. 6.93] shows that the transition matrix becomes:

$$\Phi_{t,t-1} = \begin{bmatrix} 1 & (t_t - t_{t-1}) & L^2\left(-1 + \frac{(t_t - t_{t-1})}{L} + e^{-\frac{(t_t - t_{t-1})}{L}}\right) \\ 0 & 1 & L\left(1 - e^{-\frac{(t_t - t_{t-1})}{L}}\right) \\ 0 & 0 & e^{-\frac{(t_t - t_{t-1})}{L}} \end{bmatrix}, \tag{4.49}$$

which is different from the transition matrix for a constant acceleration model, as shown in Eq. (3.21). The difference is that the model of Eq. (4.49) assumes a zero-mean acceleration, while the model of Eq. (3.21) assumes a constant acceleration. The zero-mean acceleration assumption can most clearly be seen in the lower right entry of the transition model. The constant acceleration transition model has $1$ as entry there, meaning that in the next time step the acceleration is expected to be the same. Meanwhile, the transition model of the exponentially correlated acceleration has $e^{-\frac{(t_t - t_{t-1})}{L}}$ as entry, which is the expression for the correlation. This value converges to zero for time steps $\Delta t >> L$, effectively decorrelating the acceleration value and setting it to the expectation value—zero.

The entries for $\underline{d}_t$, $Q_{d_t}$ and $\Phi_{t,t-1}$ have now been described for the deformation related parameters. However, during parameter and ambiguity estimation the height difference $\Delta H$ and the temperature correlation $\alpha$ may also be estimated, expanding the vector of unknown parameters to:

$$x_t = \begin{bmatrix} D_t \\ v_t \\ a_t \\ \Delta H_t \\ \alpha_t \end{bmatrix} \tag{4.50}$$

of which the last two, $\Delta H$ and $\alpha$ are assumed to be constant over time. For the equations to be consistent these parameters should be added to $\Phi_{t,t-1}$, $\underline{d}$ and $Q_d$ which results in:

$$\Phi_{t,t-1} = \begin{bmatrix} 1 & (t_t - t_{t-1}) & L^2\left(-1 + \frac{(t_t - t_{t-1})}{L} + e^{-\frac{(t_t - t_{t-1})}{L}}\right) & 0 & 0 \\ 0 & 1 & L\left(1 - e^{-\frac{(t_t - t_{t-1})}{L}}\right) & 0 & 0 \\ 0 & 0 & e^{-\frac{(t_t - t_{t-1})}{L}} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4.51}$$

and

$$\underline{d}_t = \begin{bmatrix} \Delta \underline{D}_t \\ \Delta \underline{v}_t \\ \Delta \underline{a}_t \\ \Delta(\Delta \underline{H}_t) \\ \Delta \underline{\alpha}_t \end{bmatrix}; \; Q_{d_t} = \sigma^2_{\text{accel}} \begin{bmatrix} q_{11} & & & & symm. \\ q_{21} & q_{22} & & & \\ q_{31} & q_{32} & q_{33} & & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{4.52}$$

It can be seen from the definition of $Q_{d_t}$ that the residual DEM height $\underline{\Delta H}$ and the temperature coefficient $\underline{\alpha}$ are assumed to be constant over time, since their variance for the difference vector is set to $0$. These matrices are then used in the time-update equations (4.29), where the sample values for $\underline{d}_t$ are set to zero. The initial value for acceleration $\underline{a}_0$ to initialise the recursive equations is set to zero as well, with variance $\sigma^2_{\text{accel}}$, Eq. (4.46). The other values in the state vector $\underline{x}_0$ result from adjusting a short time series, e.g. 25 observations, with integer least squares or ensemble coherence maximisation (both described in section 2.3.4). In the latter case the initial state covariance matrix $Q_{x_0}$ can be approximated by applying error propagation based on $Q_{\phi_0}$.

### Defining acceleration parameters

The assumption of exponentially correlated acceleration introduces two new parameters: the *acceleration variance*, $\sigma^2_a$, and *correlation length*, $L_{\text{corr}}$. Those parameters describe the expected deformation signal and should either be defined a priori by expert elicitation, or be determined based on experience. These are very important parameters that ultimately define the expected smoothness of time series, and consequently, the unwrapping result.

The correlation length can be determined based on a graph as depicted in figure 4.5. Remember from Eq. (4.49) that the third column of the transition matrix determines how the acceleration influences the expectation of the parameters of the next epoch. When a shorter correlation length is chosen, it is expected that the acceleration returns to zero more quickly. This means that when a sudden change of the deformation rate is expected, the correlation length should also be chosen accordingly.

If it is observed in situ for instance that after the maintenance to a building that lasted two months, the steady-state behaviour of the object changed with an increase in subsidence rate, the correlation length should be chosen to be about the same length. This causes the acceleration before and after the maintenance to be only slightly ($37\%$) correlated. On the other side of the scenario spectrum could for instance be a long-term process like the reduction of ground water extraction in Delft over a matter of years [Delft, 2017]. A suitable correlation length may be in the order of magnitude of one year, but consulting domain experts about the expected dynamics of the subsidence could improve the estimate, by incorporating for instance results of soil simulations. A sensitivity study on the influence of this parameter is performed in chapter 5.

It is difficult to get an accurate value for the acceleration variance, because it is a harder to interpret than for instance deformation and velocities. Therefore, it is proposed to express the acceleration variance in the expected deformation rate behaviour. A function can be plotted, using entry $q_{22}$ from Eq. (4.48), indicating the added uncertainty in the subsidence rate adjustment $\Delta \underline{v}_t$ in difference vector $\underline{d}_t$. When the correlation length is already determined, multiple lines can be chosen from as options, as is shown in figure 4.6. The $\Delta v$ values on the y-axis are $2 \times \sigma_{\Delta v}$ to obtain the 2-sigma interval of expected change of subsidence rate.

## 4.2.4 Wrapping operator in the dynamic case

Similarly as in section 4.1, the set of equations given in Eq. (4.29) uses the wrapping operator in the measurement-update. In the previous section about deformation as a static process, an expression for the uncertainty of the residual phase observation was already made in Eq. (4.14):

$$\Delta \underline{\varphi}_t = \underline{\varphi}_t - A_t \underline{\hat{x}}_{t|t-1}.$$

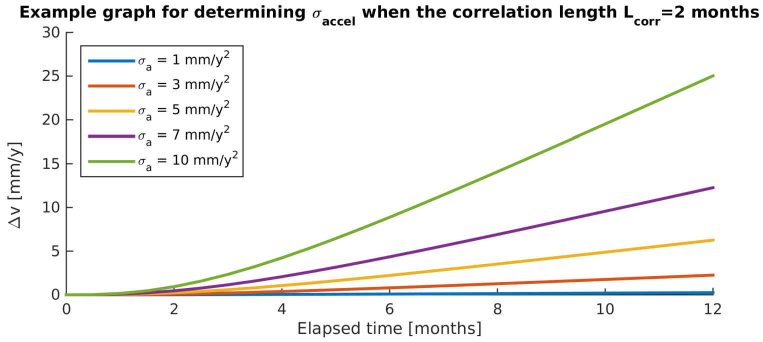


Figure 4.6: Multiple lines showing the range of possible $\Delta v$, for different values of $\sigma_{\text{accel}}$. In this case a correlation length of $L_{\text{corr}} = 2$ months is assumed. The value $\Delta v$ is estimated by taking the 2-sigma interval of $\sigma^2_{\Delta v}$.

Note that the notation changed slightly here due to the prior notation, $t \mid t-1$, of the estimator after the time-update equation. At first glance it might seem that the VCM of the residual phase does not change from the equation given in the previous chapter, because it still holds that:

$$Q_{\Delta\phi_t} = Q_{\phi_t} + Q_{\hat{x}_{t|t-1}}. \tag{4.53}$$

However, the difference is that the formulation of $Q_{\hat{x}_{t|t-1}}$ changed in Eq. (4.29) to:

$$Q_{\hat{x}_{t|t-1}} = \Phi_{t,t-1} Q_{\hat{x}_{t-1|t-1}} \Phi^{\mathrm{T}}_{t,t-1} + Q_{d_t}, \tag{4.54}$$

thus now incorporating the stochastic properties of the difference vector $\underline{d_t}$. As seen in Eq. (4.48), the values of $Q_{d_t}$ are now based on correlation length and acceleration variance, which describe the physical subsidence process. This means that the influence of $Q_{d_t}$ is now different per satellite characteristic because $Q_{d_t}$ is dependent on the repeat cycle time. Figure 4.7 and 4.8 show the theoretical success rates based on different values of $\sigma_{\text{accel}}$ given as a function of unavailable observations. Because the influence of $\sigma_{\text{accel}}$ is very low the x-axis shows values for gaps in the stack of images up to half a year. The y-axis is limited to $5\%$, because higher values lie outside the 2-sigma confidence interval for resolvable observations. For figure 4.7 it is assumed that $Q_{\hat{x}_{t-1|t-1}}$ is based on $25$ observations, while for figure 4.8 this is $50$ observations. At first glance the results may look quite counter-intuitive, since a larger $\Delta t$ means a better resolvability. However, the better resolvability is due to the fact that, although the same number of observations is used, the period that those observations span increases with $\Delta t$. Implicitly also resulting in a lower $Q_{\hat{x}_{t-1|t-1}}$ for higher repeat cycles, which was previously shown in figure 4.2 to be an important factor for deformation as a static process. It must be noted that this is just a theoretical probability, which assumes that the functional model accurately describes the physical process. Nevertheless, it gives insights into the behaviour around the important unwrapping constraint in Eq. (4.8).

## 4.3 Processing scheme

In this section, all the necessary steps for unwrapping arcs are listed for the recursive estimator. One arc consists out of the double-difference phase observations for two different pixels. In the implementation of the recursive estimator, multiple arcs can be processed at the same time due to vectorisation, as long as they share the same input parameters for the algorithm. The input parameters are listed below. Most of them are commonly used in current approaches, with the exception of the acceleration variance and correlation length:

- **Double-difference wrapped phase observations ($\varphi_{1:T}$):** Depending on the processing step, these observations can already be corrected with estimated orbital and atmospheric phase screens .

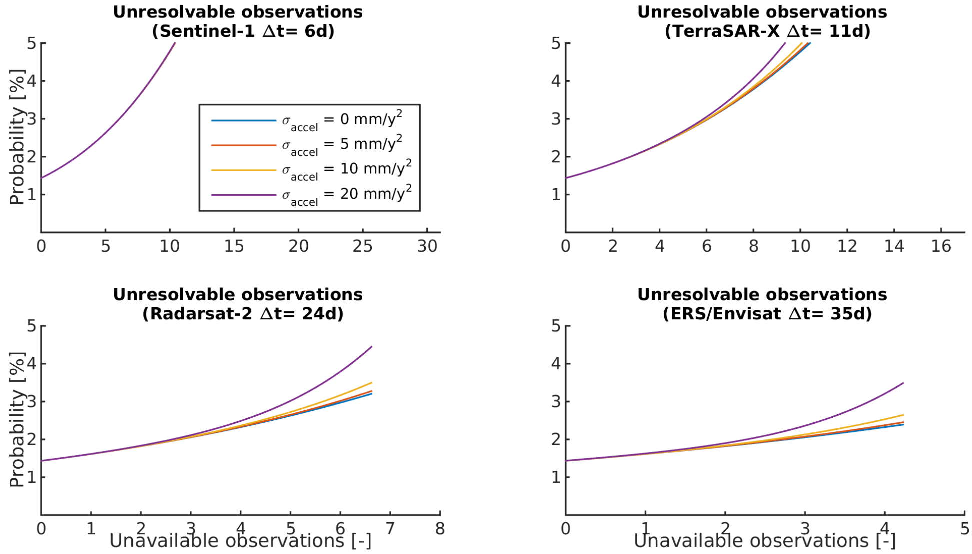


Figure 4.7: Theoretical estimate for the resolvability of double-difference phase observations with recursive least squares, given exponentially correlated acceleration. The graphs assume that $25$ epochs are already processed for the determination of $Q_{\hat{x}_{t-1|t-1}}$. Phase noise is based on $\sigma_{\varphi} = 60°$ (X-Band: $2.6mm$, C-Band: $4.6mm$), since it shown to be a good approximation of maximum allowable phase noise in figure 4.2. The correlation length $L_{\text{corr}}$ is fixed on 3 months, because this value does not change the graphs significantly. The x-axis spans half a year for all graphs and shows different values due to the different repeat cycles of the satellites. It can be seen that $\sigma_{\text{accel}}$ only starts to play a role after approximately a data gap of 3 months. Therefore, the conclusions of section 4.1 concerning the ambiguity resolution in the case of deformation as a static process, still hold. The unresolvability of observations is shown to be approximately $1.4\%$ for normal repeat cycles. It must be noted that this is just a theoretical probability, which assumes that the functional model accurately describes the physical process. Nevertheless, it gives insights into the behaviour around the important unwrapping constraint in Eq. (4.8).

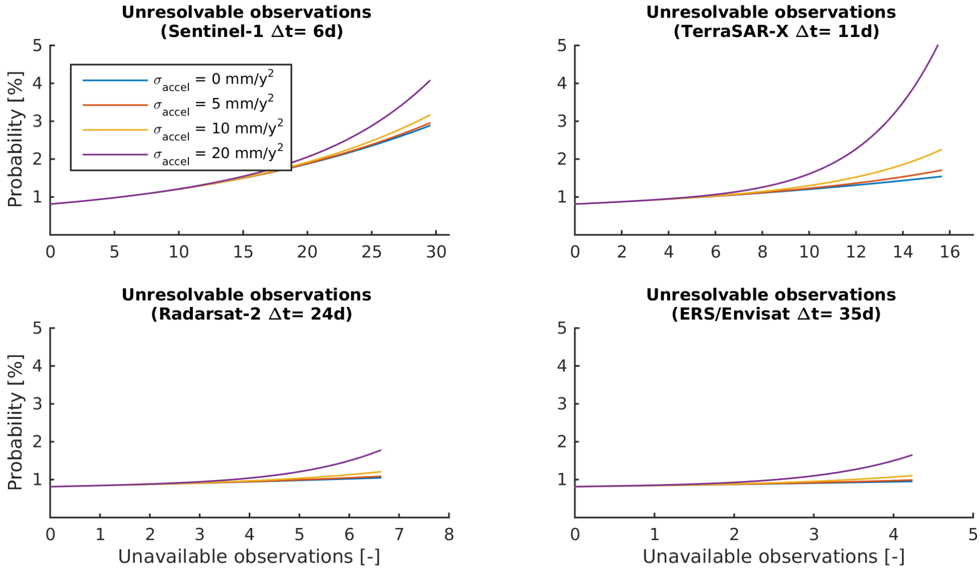


Figure 4.8: For this figure the same theory applies as described in figure 4.7, with the exception of $Q_{\hat{x}_{t-1|t-1}}$ now being computed based on $50$ processed epochs instead of $25$. It can be seen that the influence of including more epochs is significant, because all parameters result in a unresolvability of less than $5\%$, and for a typical repeat cycle less than $1\%$. It must be noted that this is again just a theoretical probability, which assumes that the functional model accurately describes the physical process. Nevertheless, it gives insights into the behaviour around the important unwrapping constraint in Eq. (4.8).

- **Phase observation epochs ($t$):** This seems trivial but all the epochs of the double-difference phase observations should be the same for all arcs being processed. This is to ensure that the stochastic properties of the computed estimates are correct and only need to be computed once.

- **Phase observation variance ($Q_\varphi$):** An estimate of the phase observation variance should be available for every epoch, valid for all arcs being processed. This can be approximated based on amplitude dispersion, but as discussed in section 4.2.2, it might be more convenient to model $Q_\varphi$ with an a priori variance factor: $\sigma_\varphi^2 \times I_{T\times T}$.

- **Initialisation parameters:** The algorithm needs an initial set of parameters per arc to initialise the recursive updating. The initialisation is done on a smaller subset of observations (e.g. the first 25 observations) with one of the unwrapping methods described in chapter 2.
  - **Initialisation length ($n_{\text{init}}$):** The initialisation length determines how many phase observations are used to get an initial approximation for estimator $\underline{\hat{x}}_0$. A steady-state deformation model is always used in this step under the assumption that the behaviour of the subsidence is steady-state for shorter time spans. Depending on the repeat cycle of the used sensor, this spans a certain length in time. For initialisation a trade-off has to be made between including as many observations as possible to increase the success rate, and keeping a short initialisation time span to increase the correctness of the steady-state assumption.
  - **Deformation rate regularisation ($\sigma_v$):** The expectation of subsidence rates is set through this variance term. It should be based upon a priori information or expert elicitation. The maximum values can be derived from the maximum resolvable rates per sensor and expected osbervation phase noise (see section 2.3.4).
  - **Residual height regularisation ($\sigma_{\Delta H}$):** Inaccuracies in the used DEM cause the need to estimate residual heights. The magnitude of the found residual heights depends on the used DEM, type of land coverage, world region and objects of interest. Often a variance value based on experience is used to estimate this term.
  - **Temperature correlation regularisation ($\sigma_\alpha$):** This parameter regularises the temperature correlation during initialisation. This parameter can be tuned based on objects of interest, type of stack, and experience. The temperature correlation can be estimated in the initialisation, but can also be assumed zero with this regularisation parameter as initial variance.

- **Correlation length ($L_{\text{corr}}$):** The correlation length determines the degree of correlation of acceleration in time. When acceleration occurs, the correlation length determines the shape of expected acceleration reduction.

- **Acceleration variance ($\sigma_{\text{accel}}$):** The magnitude of acceleration is determined by this variance factor. When the expected deformation is steady-state, this term should be zero. When a lot of dynamics are expected this term should be scaled accordingly.

- **Temperature observations (optional):**($K_{1:t}$): A strong temperature dependence in the scatterers conflicts with the steady-state behaviour, because this term is momentarily not in the functional model. Therefore, temperature observations corresponding to the observation epochs can be added to model the dilation of objects as a results of temperature changes.

The proposed algorithm can be best schematised in three essential steps:

1. **Initialisation:** The phase data of arcs to be processed simultaneously are fed into the initialisation algorithm. Initialisation is done based on the first $n_{init}$ observations of each arc and results in an estimate per arc for: $t_0$ offset, subsidence rate and residual height. Two possible methods are described to obtain the initial parameters:
   (a) **Ensemble coherence maximisation:** This method is also referred to as *the ambiguity function* [Counselman and Gourevitch, 1981] and was discussed in more detail in chapter 2. This method results in an unwrapped time series for the first $n_{init}$ observations and estimates of the parameters. However, it does not return a full covariance matrix. This could be solved

   by applying Least Squares error propagation on the unwrapped time series, although this does not take into account the uncertainty of the integer ambiguities. Nevertheless, this method is often preferred on account of its computational efficiency.

   (b) **Integer Least Squares:** This method does result in a full variance-covariance matrix and optimal estimate, it is however up to a factor $100$ slower compared to ensemble coherence maximisation when noise and amount of observations increases [van Leijen, 2014].

2. **Recursive Least Squares (Iterate: $t_0 \rightarrow t_T$):** This part updates the previous estimator with the latest observation and simultaneously unwraps that observation to the correct phase cycle, given that the assumptions made about the signal are correct. The recursive Least Squares algorithm is initiated with the parameters of step 1 and thereafter re-estimated for every new observation. The result is a parameter estimation for every epoch $\underline{\hat{x}}_{t|t}$, based on all previous observations; $\underline{\hat{x}}_{t|1:t-1}$. This part consists out of 2 steps (Eq. 4.29):

   (a) **Time-update:** This part creates the prior estimation for the measurement-update. This prior is created by "transitioning" the parameters of the previous time step to the current time step with the necessary error propagation.

   (b) **Measurement-update:** In the measurement-update step the prior is updated with the new observation to obtain a new estimate for the latest epoch. The ambiguity assumption in Eq. (4.8) allows for the use of the wrapping operator, which enables the comparison of the unwrapped prior with the wrapped observation. The results of this step become the input for the time-update of the next epoch.

3. **Hindcasting (Iterate: $t_T \rightarrow t_0$):** In this step the estimate $\underline{\hat{x}}_{t|1:t}$ is transformed into $\underline{\hat{x}}_{t|1:T}$, $T$ being the latest observation, meaning that the estimates are updated using all available observations. This thesis uses the Rauch-Tung-Striebel fixed-interval hindcasting algorithm [Rauch, Striebel, and Tung, 1965] which is a popular and efficient fixed-interval hindcasting implementation.

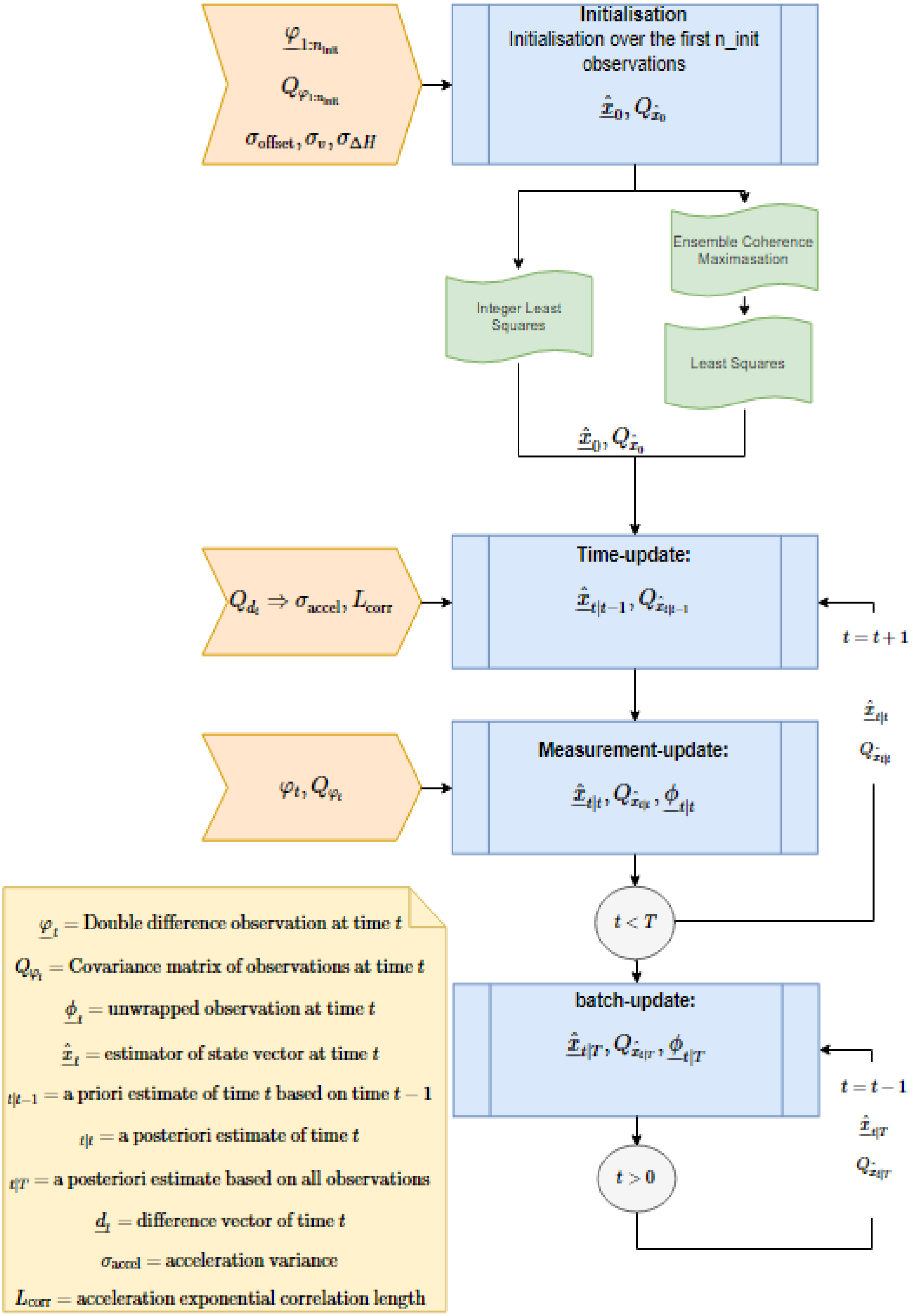


Figure 4.9: Processing scheme of the recursive parameter estimation algorithm. The estimation can be done for multiple arcs at the same time, where above process is done in an vectorised implementation. The requirement for this approach is that all arcs use the same input parameters (e.g. observation phase noise $Q_{\varphi}$.)

# 5

# Application on Simulated Data

*To assess the performance of the recursive method, data is simulated based on the characteristics of three different sensors: TerraSAR-X, Radarsat-2 and ERS. Based on those characteristics, arcs of double-difference phase observations are simulated for different kind of deformation types and phase noise levels. These arcs are processed multiple times using the recursive parameter estimation method, using different input parameters for every run. The results are finally quantified to assess their performance and sensitivity to different input parameters.*

## 5.1 Stack characteristics

To test the performance of the recursive parameter estimation method, data is simulated based on different satellite sensor characteristics. The main reason for this approach is that different radar satellites have different repeat cycle times, operate with different wavelengths, and have different perpendicular baseline configurations. For that matter, three actual stacks have been chosen that cover the city of Delft for over more than 5 years. For each stack an already processed dataset of Persistent Scatterers is selected and trimmed to only have $1000$ PS. The actual perpendicular baselines of each stack are used to simulate heights for every PS. Different type of deformation signals are simulated and sampled on the acquisition dates of every stack. The satellite wavelength is used to compute the phase contribution of the deformation signal. The resulting phase of every PS is stored as complex value retain the wrapped nature of the simulated phase observation.

### 5.1.1 TerraSAR-X

*TerraSAR-X* operates in X-band with a wavelength of $\lambda = 31mm$. This means that a line-of-sight (LOS) deformation of $\pi$ radians equals $15.5mm$. *TerraSAR-X* is often used when high resolution is needed because the default mode, *stripmap*, has a ground resolution of $3m \times 3m$, which can be increased with *spotlight* to $1m \times 1m$ [Werninghaus et al., 2003]. It has been operational since 2008 and is still actively obtaining imagery. Figure 5.1 shows the observation dates for the *descending* stack covering Delft, on track $48$.

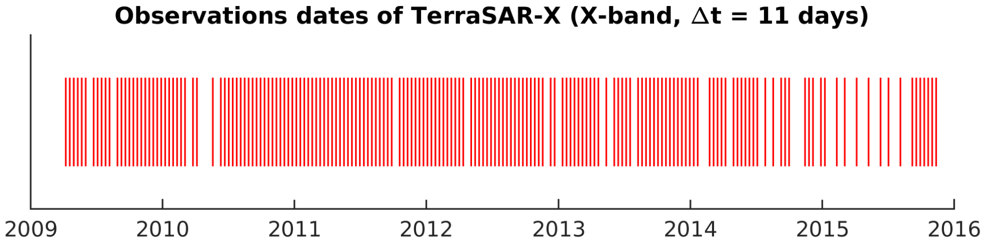


Figure 5.1: Observation dates of the *descending* TerraSAR-X stack of track $48$. The default repeat cycle time is $11$ days, but regularly some acquisitions are skipped. In 2015 the acquisition intervals even increase to once a month for a longer period of time, and is kept in the simulation to have a more realistic stack. The total amount of acquisitions is $182$.

### 5.1.2 Radarsat-2

The SAR satellite *Radarsat-2* is owned by the *Canadian Space Agency* and has been operational since 2008. It operates in C-band with a wavelength of $56mm$ and is still actively acquiring images. The used stack covers Delft and is recorded in the *Descending* track, with track number $102$. Standard ground resolution is approximately $5m \times 20m$ [Morena et al., 2004]. Figure 5.2 shows the observation dates of the stack used for simulation.

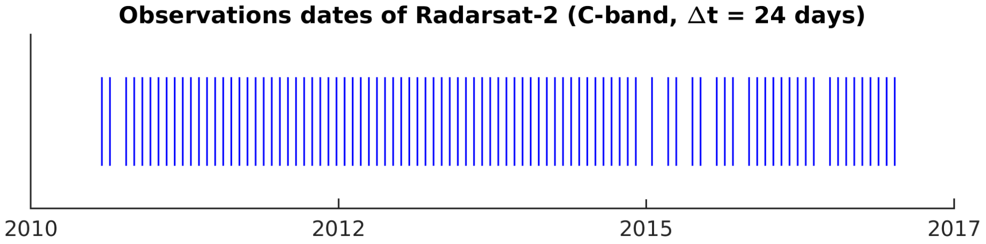


Figure 5.2: Observation dates for the Radarsat-2 stack of descending track $102$. The nominal repeat cycle is $24$ days, but some acquisitions are missing, especially during 2015. The total amount of acquisitions is $92$.

### 5.1.3 ERS

The *European Remote sensing Satellite* (ERS) was the first earth observation satellite of the *European Space Agency* (ESA). The first one, ERS-1, was launched in 1991, the second one, ERS-2, was launched in 1995. They both have the same SAR sensor, C-band, operating on a wavelength of $56mm$ [Attema, 1991]. For the greater part of their operations they shared the same nominal repeat cycle of 35 days, flying in a tandem orbit with an one day interval. ERS-1 was operational until 2000 and ERS-2 up until to 2011, while the data usable for inSAR only spans up to 2001 due to a gyroscope malfunction. The used stack covers Delft during descending track number $423$. The original stack does have data available before 1995, but the large observation gap of more than a year has motivated the decision to remove that data, to avoid unwrapping errors caused by a lack of data. The ground resolution is $4m \times 20m$.

### 5.1.4 Stack comparison

Figure 5.4 shows the three different stacks used for simulation within one figure, with on the x-axis the duration of the stacks. The improvements in repeat cycle times make near real-time monitoring of subsidence more viable.

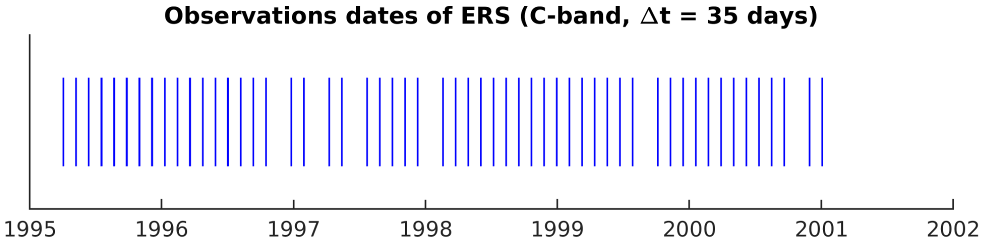


Figure 5.3: Observation dates of the ERS descending track number $423$, covering Delft. The nominal repeat cycle time is $35$ days and this stack only misses a few of those observations. The data before 1995 has been removed due to a very large data gap of more than a year, making it a challenge to resolve the ambiguities correctly due to unknown possible subsidence behaviour over this long time. The total amount of acquisitions is $62$.

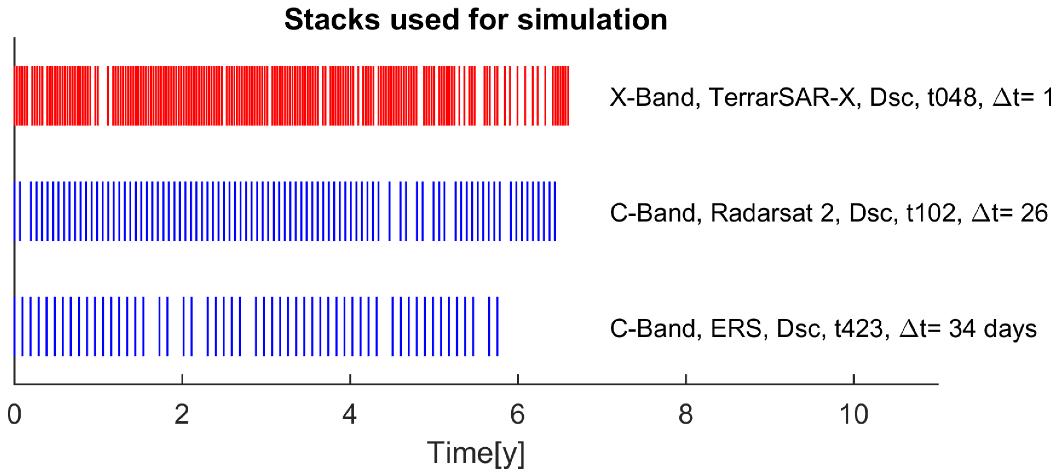


Figure 5.4: Comparison of the three stacks used for simulation to clearly visualize the differences in nominal repeat cycle times. The $\Delta t$ listed on the right are the average repeat cycle times for that specific stack.

## 5.2 Deformation characteristics

One of the advantages of the recursive parameter estimation is the ability to capture different types of non steady-state deformation behaviour, without changing the functional model. There are two *input-parameters* in this functional model that describe the expected smoothness of the deformation: correlation length $L_{\text{corr}}$ and acceleration variance $\sigma^2_{\text{accel}}$. Later in this chapter it is explored how sensitive those parameters are and if a rough approximation of these values yields the same results as the exact values. To achieve this, sets of different deformation types are modelled. Of each type, $1000$ realisations are made to test a wide range of deformation parameters. The different types of deformation are discussed in the following sections.

### 5.2.1 Steady-state

One of the most common assumptions for deformation estimation in PSI is steady-state deformation, shown in figure 5.5. This implies that the deformation rate is constant during the time series. In practice there may be deviations, but with the steady-state assumption small deviations from that model, like a small temperature correlation, are still resolvable in unwrapping. For this type of deformation the range of values used in the simulation is $\underline{v} = [0, 20]$ mm/y [SkyGeo, 2016].

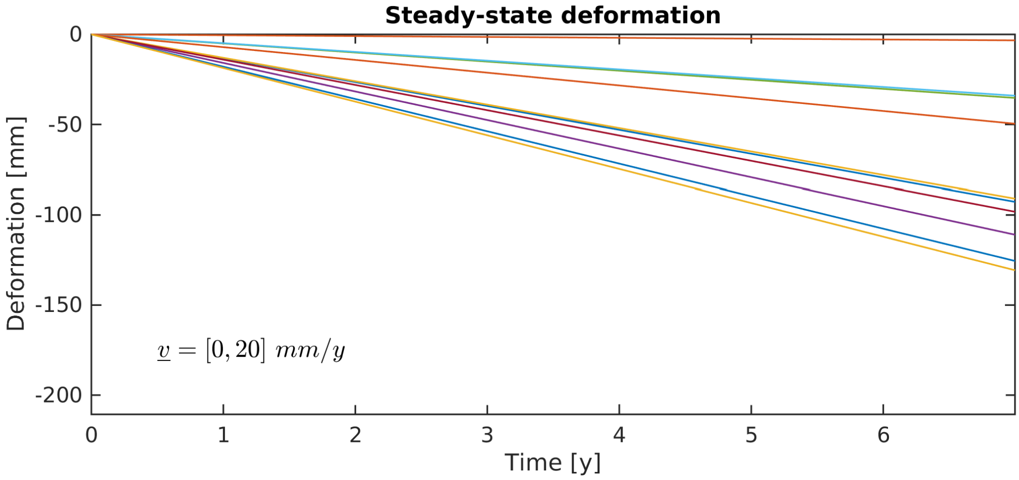


Figure 5.5: Ten realisations of the steady-state deformation. The deformation rates are in the range of 0 mm/y to 20 mm/y.

### 5.2.2 Steady-state with a constant acceleration term

The next deformation type is an extension to the steady-state model, adding an additional constant acceleration term. The equation used for this is:

$$\underline{D} = \underline{v}(t - t_0) + \underline{a}(t - t_0)^2, \tag{5.1}$$

and the results can be seen in figure 5.6. The main motivation for this deformation type is that this specific type of model is often used for unwrapping by the graduation company, whenever the results of the default steady-state model indicate that many arcs are unwrapped incorrectly. The steady-state rates vary from $\underline{v} = [0, 20]$ mm/y and the acceleration terms are randomly chosen from the interval $\underline{a} = [-1, 1]$ mm/y$^2$ [SkyGeo, 2016].

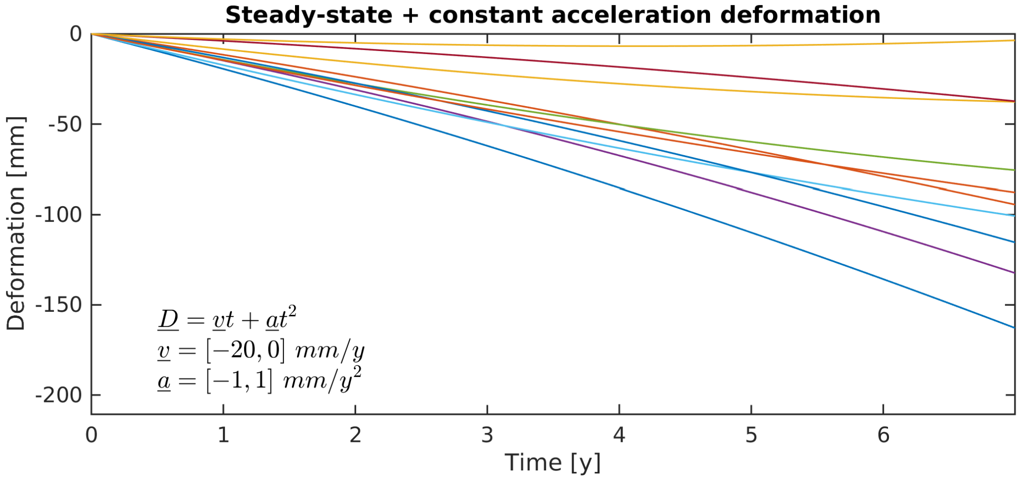


Figure 5.6: Ten realisations of the extended steady-state deformation model, which has an additional constant acceleration term. The steady-state rates are chosen from the interval $\underline{v} = [0, 20]$ mm/y and the acceleration term is chosen from $\underline{a} = [-1, 1]$ mm/y$^2$.

### 5.2.3 Dynamic

The second deformation type is called *dynamic deformation*. It is generated under the assumption of the exponentially correlated acceleration (see section 4.2.3) and works in three steps:

1. Per realization a random steady-state deformation parameter is initiated in the range of $\underline{v}_0 = [0, 20]$ mm/y.
2. Time series of exponentially correlated acceleration, $a_{0:T}$, are generated based on a correlation length $L_{\text{corr}}$ =5 months and an acceleration standard deviation of $\sigma_{\text{accel}}$ =5 mm/y$^2$, 10 mm/y$^2$ and 20 mm/y$^2$. These values have been chosen based on visual inspection of the resulting time-series, agreeing that these values result in different orders of non-linear behaviour as sometimes experienced in real projects [SkyGeo, 2016].
3. Deformation values are generated iteratively based on the following recursion:

$$\begin{aligned} D_t &= D_{t-1} + v_{t-1}\Delta t + a_{t-1}\frac{1}{2}(\Delta t)^2, \\ v_t &= v_{t-1} + a_{t-1}\Delta t, \end{aligned} \tag{5.2}$$

   where $D_0$ =0 mm, $v_0$ is the in step 1 determined deformation rate and $a_{0:T}$ is the realization of the exponentially correlated acceleration.

This deformation model is thus generated under the same assumption as used in the functional model of the recursive parameter estimation (section 4.2.3). Some sample realisations using the three different values of $\sigma_{\text{accel}}$ are shown in figure 5.7.

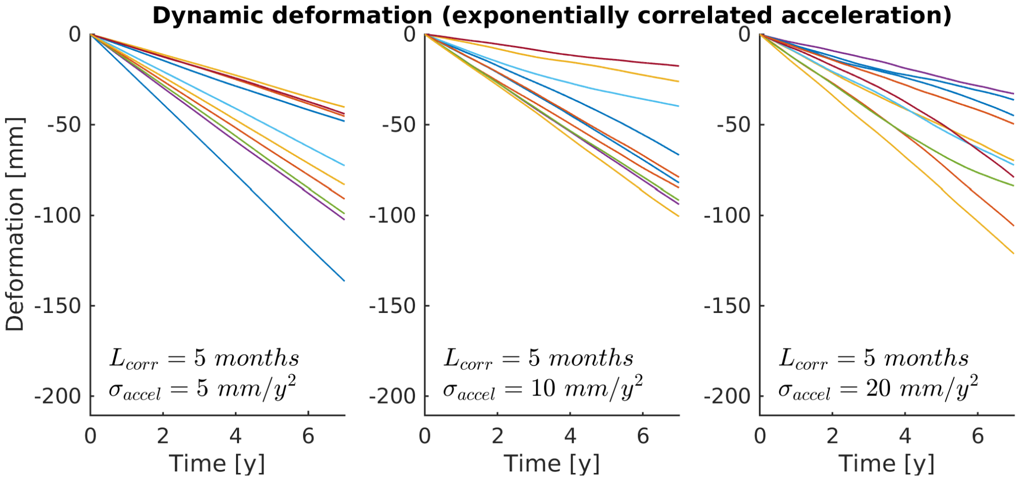


Figure 5.7: Realisations of the dynamic deformation type, for three different values of acceleration variance (5 mm/y$^2$, 10 mm/y$^2$ and 20 mm/y$^2$). It can be seen that the deviations from the steady-state model increase when $\sigma_{\text{accel}}$ increases.

### 5.2.4 Exponentially decaying

Due to consolidation of the soil, exponentially decaying subsidence often occurs in areas where new objects are build. As a result of loading, e.g. newly build houses, the soil tends to subside rapidly over a relatively short time span. This rapid deformation gradually slow downs until the remaining deformation is approximately steady-state. The deformation is approximated by using:

$$\underline{D} = \underline{b}_{\text{exp}} e^{\ln(1-\rho)\frac{(t-t_0)}{t_\rho}}, \tag{5.3}$$

where $\rho = .99$ and $t_{\rho=99\%} = 700$ days, based on settlement rod data from an high way construction work [SkyGeo, 2014]. The magnitude of exponential decaying deformation is indicated as $\underline{b}_{\text{exp}}$. $t_{99\%}$ Is the time when 99% of the exponentially decaying deformation has taken place. It can be seen from figure 5.9 that the initial deformation rate $\underline{v}_0$ is very high and it is therefore important to determine for which magnitude $\underline{b}_{\text{exp}}$ the deformation is still resolvable. The resolvability of the deformation is determined similarly as previously depicted in figure 2.2, where it was shown which linear deformation rates are resolvable for different sensors and noise levels. For this case it is desirable that the difference between the first two observations is less than (figure 2.2):

$$\frac{4\pi}{\lambda}\underline{b}_{\text{exp}}\left(e^{\ln(1-\rho)\frac{\Delta t}{t_\rho}}-1\right)<\pi-r_\alpha\sigma_{\text{phase noise}}, \tag{5.4}$$

where $\Delta t$ indicates the nominal repeat cycle of the satellite, $r_\alpha$ is the significance level factor of the one-sided Gaussian distribution, and $\sigma$ is the standard deviation of the double-difference phase observation. For a significance level of $\alpha = .05$, $r_\alpha = 1.65$. When Eq. (5.4) is expressed in terms of $b_{\text{exp}}$:

$$b_{\text{exp}} = \left(\pi - r_\alpha\sigma_{\text{phase noise}}\right)\frac{\lambda}{4\pi}\left(e^{\ln(1-\rho)\frac{\Delta t}{t_\rho}}-1\right)^{-1}. \tag{5.5}$$

The expression of $b_{\text{exp}}$ can be plotted as a function of repeat cycle and observation phase noise. The result from this can be seen in figure 5.8, where theoretical resolvable values for $b_{\text{exp}}$ are shown as function of repeat cycle and phase noise. Figure 5.8 can in combination with expert elicitation, provide an approximation on when to start monitoring consolidation. Based on figure 5.8 it is decided to model the consolidation with values of $b_{\text{exp}} = [50, 100]$ mm to be within the theoretical resolvable range of TerraSAR-X. The reason why these values are chosen in the range of TerraSAR-X, is because of the fast repeat cycle of this satellite, making it a popular choice for near real-time monitoring. Also, lower values of $b_{\text{exp}}$ would make the deformation lie within only one or two phase cycles for C-band, and would therefore be relatively easy for the ambiguity estimation. Figure 5.9 show ten realisations of the consolidating soil.

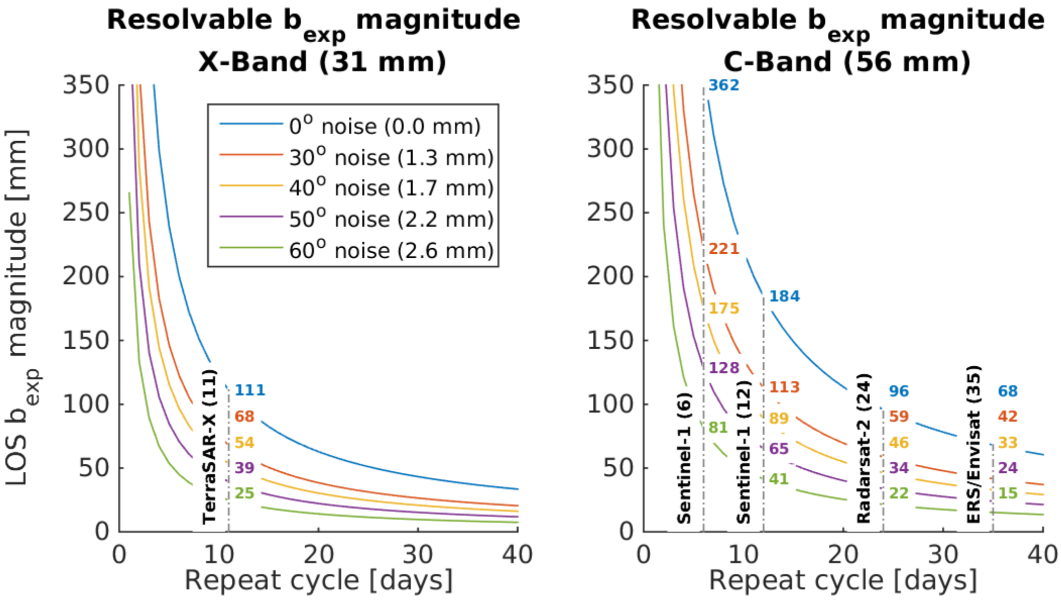


Figure 5.8: Indication of the resolvable consolidation of soil when approximated by Eq. (5.3). The significance level is set to be $\alpha = .05$, $\rho = .99$, and $t_{\rho=99\%} = 700$ days [SkyGeo, 2016]. It can be seen that for long repeat cycles, like Radarsat-2 and ERS, the resolvable consolidation is low. The resulting deformation phases would lie within one or two phase cycles, resulting in the ambiguity resolution that can even be solved by a steady-state model. For simulation purposes the parameters are therefore chosen in the range of $b_{\text{exp}} = [50, 100]$ mm, making them theoretically resolvable.

### 5.2.5 Single and double breakpoint

The last type of simulated deformation is the breakpoint model, used in e.g. Samiei-Esfahany [2008]. The breakpoint model is a steady-state model that changes on a certain moment in time to a new

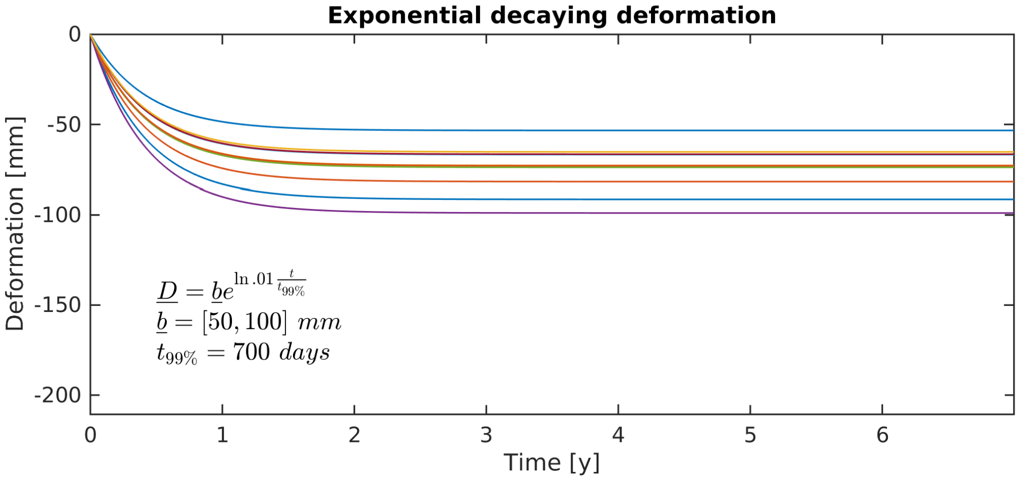


Figure 5.9: realisations of the approximated consolidation of soil. It can be seen that the total deformation ranges from $[50, 100]$ mm.

steady-state model. The model is continuous on the breakpoint and are for the simulations of this thesis modelled as single and double breakpoint. The steady-state parts of the deformation are at least $20$ observations long, such that during initialisation the steady-state assumption holds. This results in different periods of time that steady-state deformation is occurring, depending on the nominal repeat cycle of the satellite. The location of the breakpoint in time is random per realisation and unknown a priori for processing. The initial deformation rate is modelled as $\underline{v} = [0, 20]$ mm/y and the size of $\Delta\underline{v}$ is modelled to be in the range of $\Delta\underline{v} = \pm[5, 10]$ mm/y to show a significant break in the time series. Ten realisations of the single and double breakpoint models are shown in figure 5.10.

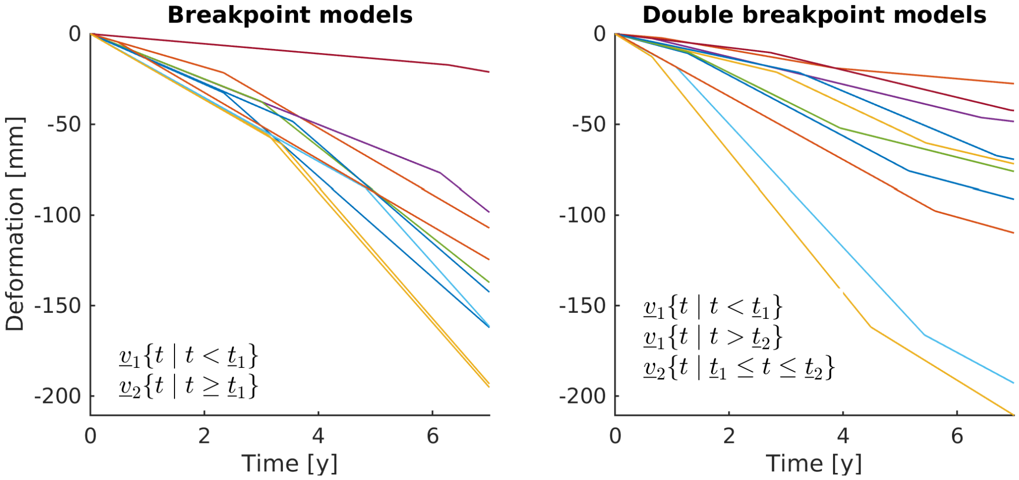


Figure 5.10: realisations of the single and double breakpoint models. The location in time of the breakpoints are random and have the constraint that at least 20 observations are available before and after the breakpoint, subsiding in steady-state behaviour. The initial deformation rate is $\underline{v} = [0, 20]$ mm/y and the rate of change on a breakpoint is modelled as $\Delta\underline{v} = \pm[5, 10]$ mm/y.

## 5.3 Simulated phase contributions

Besides deformation there are more phase contributions in the signal. They were previously discussed in section 2.2 and in this section it will be described which contributions will be modelled in the simulation.

### 5.3.1 Atmospheric and orbital phase screens

The atmospheric and orbital phase contributions are considered to be correlated in space, but not significantly in time [Hanssen, 2001]. Because of this, these contributions can be estimated when a significant amount of images is available and as long as the spatio-temporal characteristics of the signal do not show equivalence with the signal of interest. In this thesis we consider unwrapping of double-difference phase observations obtained over arcs, generated between persistent scatterer candidates (PSc). Two cases can be distinguished:

1. **Initial network:** The initial network is generated before orbital phase screen (OPS) and atmospheric phase screen (APS) estimation. The PSc's chosen for this network are expected to have a low observation phase noise, and because the arcs are only made over short distances ($< 800$ m), the influences of the OPS and APS are assumed to be insignificant. This is due to the maximum arc distance being set based on the spatial correlation length of the OPS and APS signal.
2. **Densification:** Since in the initial network arcs with low phase noise were used, an estimation of the APS and OPS could be made. The estimated phase contributions of those phase screens can now be subtracted from all PSc during densification, reducing the influence on arcs that have larger observation phase noise.

For both cases it is thus assumed that the influence of the OPS and APS on the double-difference phase observations is negligible. No specific phase screen is therefore simulated, because it is assumed negligible or to be captured by the observation phase noise contribution.

### 5.3.2 Residual DEM heights

It cannot be assumed that a perfect DEM or digital surface model (DSM) is available in processing and therefore residual DEM/DSM heights will always influence the double-difference phase observations. It is essential to estimate this $\Delta H_{\text{dem}}$, to be able to resolve the deformation signal. Residual DEM heights are therefore simulated for the arcs using Eq. (2.6) and the actual perpendicular baselines, $B_{\perp}$, of the stacks described in section 5.1. The range of residual heights are set to $\Delta H_{\text{dem}} = [-30, 30]$ m.

### 5.3.3 Observation phase noise

The contributions to the double-difference phase observations, other than deformation and residual DEM heights, are modelled as zero-mean Gaussian noise [Hanssen, 2004]. The important question then remains how to choose a realistic variance value for this noise. For this purpose some PSI datasets are investigated. Datasets of TerraSAR-X, Radarsat-2, ERS and Envisat are chosen to get representative values encountered in practice. These datasets have the estimated *temporal ensemble coherence* stored per processed arc, which is described by Ferretti et al. [2001] as the absolute value of the complex temporal *ensemble coherence*:

$$\underline{\hat{\gamma}} = \left| \frac{1}{T} \sum_{t=1}^{T} \exp(j(\underline{\varphi}_t - A_t \underline{\hat{x}}_t)) \right|, \tag{5.6}$$

where $\underline{\hat{\gamma}}$ is the estimator of the temporal ensemble coherence, $j$ is the imaginary unit and $A_t \underline{\hat{x}}_t$ is the modelled phase observation of time $t$. It is essentially a normalized measure of fit to the estimated phase observations and, assuming that the temporal correlated deformation term is estimated correctly, a measure of the temporal phase dispersion. Since the estimated *temporal ensemble coherence* is the only stored characteristic in the used datasets on arc-level, it's desirable to approximate the observation phase noise per arc based on these values.

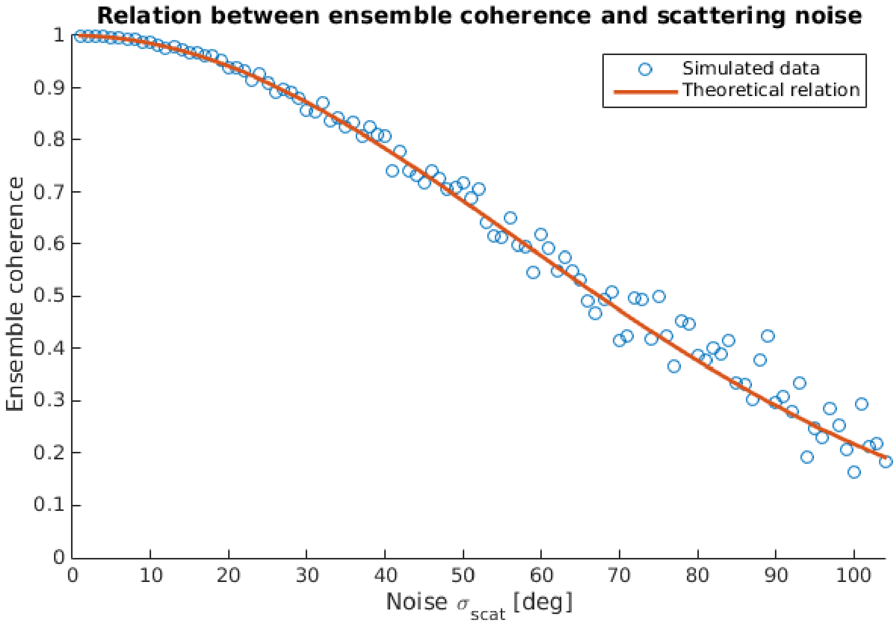


Figure 5.11: Relation between $\underline{\hat{\gamma}}$ and $\sigma$. The values are generated by simulating $1000$ time series of $200$ observations. The time series are generated with increasing values of $\sigma$ and their $\underline{\hat{\gamma}}$ estimated by using Eq. (5.6). The dispersion of the simulated values around the theoretical relation increases with the noise level.

The relation between temporal ensemble coherence and observation phase noise is given by [e.g. Pewsey et al., 2013]:

$$R = \left| e^{-\frac{\sigma^2}{2}} \right|, \tag{5.7}$$

where $R$ is in literature referred to as the *mean resultant vector* which equals the expectation of the temporal ensemble coherence, $E\{\underline{\hat{\gamma}}\}$. Eq. 5.6 can be used to show the relation between $\underline{\hat{\gamma}}$ and noise level $\sigma$ as is shown in figure 5.11. The relation in Eq. 5.6 can be used to get realistic phase noise variances for the simulations in this chapter, based on real processed datasets. The results are shown in figure 5.12. The datasets do not have PS with temporal ensemble coherence values lower than $0.5$. That threshold is set to distinguish points with significant signal from points with mostly noise. The values found in these datasets are most likely negatively biased (showing more phase noise), due to unmodeled deformation being interpreted as noise. Based on the histograms in figure 5.12, observation phase noise levels of $40°$ (X-Band: $1.7mm$, C-Band: $3.1mm$), $50°$ (X-Band: $2.2mm$, C-Band: $3.9mm$) and $60°$ (X-Band: $2.6mm$, C-Band: $4.7mm$) are chosen, to cover the whole spectrum of phase noise levels encountered in current projects.

## 5.4 Simulation input parameters

In this section the different kinds of simulation input parameters are discussed. Every instance of every deformation type will be unwrapped by the recursive parameter estimator. As a reference, the data will also be unwrapped by using a periodogram (section 2.3.4). The periodogram is applied with (i) solely a steady-state model, and (ii) a combination of steady-state with a constant acceleration component. Different input parameters will be varied independently, to investigate the sensitivity of the deformation types to these parameters.

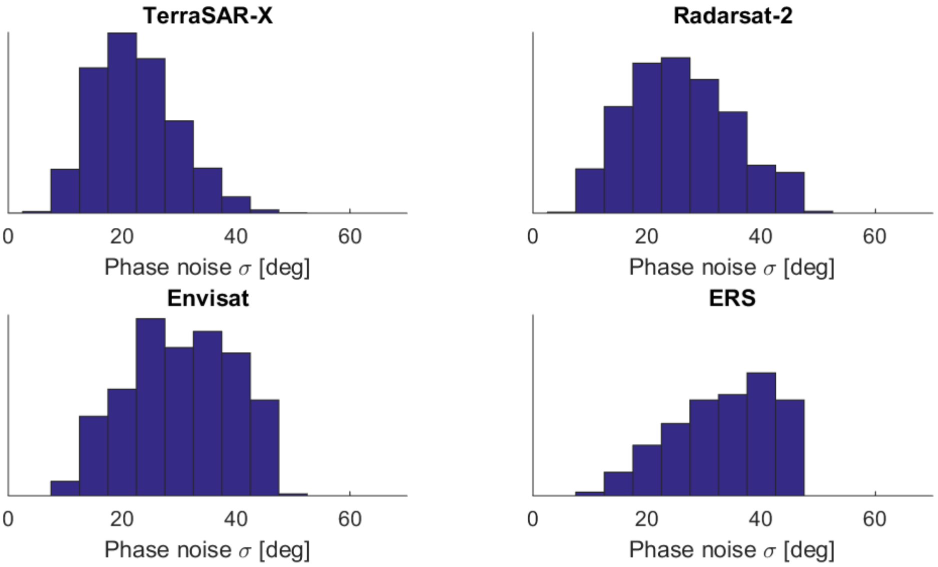


Figure 5.12: Estimated noise levels of arcs in four different data sets. These histograms are the basis for the decision of the observation phase noise levels of $40^\circ$ (X-Band: $1.7mm$, C-Band: $3.1mm$), $50^\circ$ (X-Band: $2.2mm$, C-Band: $3.9mm$) and $60^\circ$ (X-Band: $2.6mm$, C-Band: $4.7mm$).

### 5.4.1 Initialisation length

First the influence of the initialisation length is investigated. The initialisation length $n_{\text{init}}$, is the amount of observations from the start of the time series that is used for an initial parameter estimation. The estimated ambiguities of the initialisation interval is not used, only the parameters needed to initiate the recursive parameter estimation are used. Initialisation is done based on a steady-state assumption and estimates:

$$\underline{v}_0, \Delta \underline{H}, \tag{5.8}$$

where $\underline{v}_0$ is the deformation rate at epoch $0$ and $\Delta \underline{H}$ is the height difference over the arc. The time subscript is omitted here because this is not a time-varying parameter, however, the parameter is updated on every observation. While it may be tempting to choose the initialisation length as long as possible, it must be considered that this will challenge the steady-state assumption. The longer the initialisation length, the less deviation from the steady-state model is allowed. This is difficult to quantify, but a period of time should be chosen during which the steady state behaviour is very probable, and enough observations are available for redundancy of the estimation. Figure 5.13 shows the relation between amount of observations and time, based on the nominal repeat cycle. This figure can help to determine the amount of observations for initialisation. The lower limit is set to $n_{\text{init}} = 15$ observations to assure a certain level of redundancy, which is a value based on experience [SkyGeo, 2016] for which a stack usually seems to generate reasonable results. To investigate the effect on the success rate, the initialisation will be done with increasing values of $n_{\text{init}}$ of $[15, 20, 25, 30, 35]$.

### 5.4.2 A priori acceleration correlation length

For the exponential correlated acceleration assumption, a value is needed that describes the smoothness of the deformation, i.e., the correlation length $L_{\text{corr}}$. It was already discussed in chapter 4 that physical knowledge about the signal of interest, for instance obtained with expert elicitation, should guide the decision of an a priori value for $L_{\text{corr}}$. The correlation length can be interpreted as the amount of time that an event can result in a certain acceleration. This can range from a very short period, e.g. a month, caused by an instantaneous event, up to a long period of even a year due to e.g. permanent

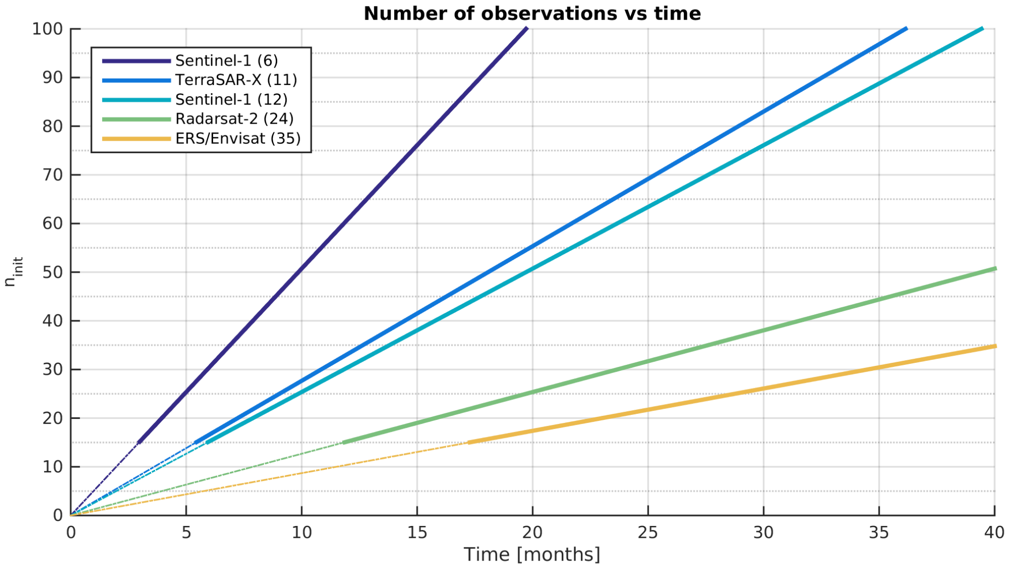


Figure 5.13: Relation between number of observations and the time span based on nominal repeat cycles (indicated in the parenthesis in the legend). This figure can be used for determining the initialisation length. When the subsidence is expected to be approximately steady-state in a period of 1 year, it can be determined that the ideal initialisation length for TerraSAR-X is 33 observations. The used initialisation length is then $\min([n_{\text{stack}}, 33])$, where $n_{\text{stack}}$ is the amount of images in the stack. As the stack grows over time, to over more than $33$ images, observations will be left out of initialisation.

ground water extraction. Because of this broad range the simulated input parameters for correlation length $L_{\text{corr}}$ are chosen as $[1, 3, 5, 7]$ months.

### 5.4.3 A priori acceleration variance

The acceleration variance determines the magnitude of the expected correlated acceleration. Chapter 4 provided the idea to determine this variance based on the expected steady-state subsidence rate change over time, see figure 4.6. For this analysis, the same acceleration variance values are used, as in the simulation of the *dynamic* subsidence data type: $\sigma_{\text{accel}} = [0.1, 5, 10, 20] mm/y^2$. The value of $0.1 mm/y^2$ is added to show that for very small values of acceleration, the functional model of the recursive parameter estimation method approximates a steady-state model.

### 5.4.4 A priori phase noise variance factor

The observation phase noise is described in $Q_\phi$, see Eq. (4.24), and indicates the dispersion of the phase observations. The phase noise in the simulations is modelled as a zero-mean Gaussian variance factor, such that $Q_\phi = \sigma^2_{\text{phase noise}} \times I_T$. Since a quantitative prediction of the observation phase noise is difficult, due to the many varying factors influencing this noise, the variance factor is also used as an a priori estimation of the phase noise. Based on the analysis of phase noise in real data in figure 5.12, and the consequences of the wrapping operator on the amount of acceptable phase noise in figure 4.2, the maximum variance factor is set to $\sigma_{\text{phase noise}} = 70°$. To investigate the influence of the a priori phase noise factor on the unwrapping result, values in the range of $\sigma_{\text{phase noise}} = [30°.40°.50°.60°.70°][deg]$ are used in unwrapping the simulated deformation.

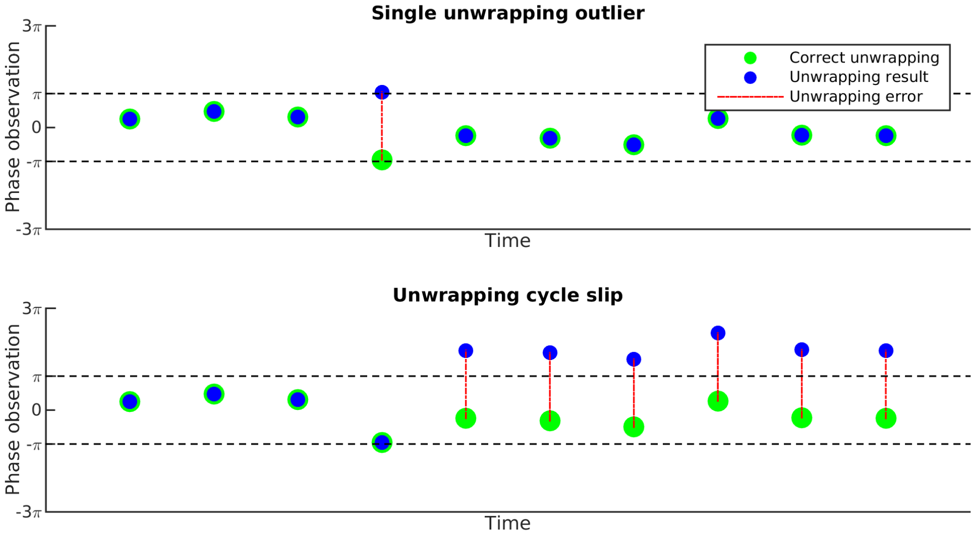


Figure 5.14: Example showing the difference between a single unwrapping outlier and a *cycle-slip*. It can be seen that where the cycle-slip has a huge impact on the time series solution, the single unwrapping outlier time series still contains most of the information content and is therefore allowed to occur. This decision will increase the overall success-rates.

## 5.5 Quantification

Observations are simulated for three satellites, with eight deformation types and three degrees of noise. These observations are in total unwrapped $20$ times, using different input parameters. This results in a total of $3 \times 8 \times 3 \times 20 = 1440$ data sets, each containing $1000$ unwrapped time series. To extract useful statistics from those $1.4$ million unwrapped arcs, three quality metrics are computed.

### 5.5.1 Success rate

The goal of unwrapping is to determine the amount of integer phase cycles that a wrapped observation has to be shifted to obtain the original unwrapped phase observation. With real data it is impossible to know if an observation is unwrapped correctly due to the ambiguous nature of the phase measurement. Fortunately, all observations in this chapter are simulated and thus the original unwrapped phase observations can be checked against the unwrapping result. From this the success rate can be determined—the percentage of arcs in one data set that are correctly unwrapped. An arc is considered to be unwrapped successful when all ambiguities are correctly estimated, with the exception of one or more single outliers that are caused by large phase noise or other uncertainties. The effect of those outliers on the deformation result are considered to be minimal. The overall success-rate will be higher due to this. An example of a single outlier is given in figure 5.14.

### 5.5.2 Initialisation errors

Since the initialisation is paramount for the success of the recursive parameter estimation, it is done by using already existing unwrapping methods. It is therefore interesting to distinguish incorrect unwrapping caused by initialisation from incorrect unwrapping caused by the recursive algorithm. However, this is not trivial, because initialisation solely provides estimates of the parameters and the final unwrapping is completely done by the recursive algorithm. To approximate if the initialisation parameters are the cause of an incorrect unwrapped time series, it is assumed that a correct set of initialisation parameters should at least be able to unwrap the first $10$ observation correctly. If the first $10$ observations are unwrapped correctly, mistakes are thereafter assumed to be caused by the unwrapping algorithm.

### 5.5.3 Deformation estimation

The final quality metric to introduce is the standard deviation of the displacement estimation. Since displacement and derived products like subsidence rates are the main use of PSI, it is of interest to know the degree of fit of the modelled displacement to the actual simulated displacement. For this the standard deviation between simulated and adjusted displacement is calculated by:

$$\sigma_{\text{defo}} = \sqrt{\frac{\sum_{t=1}^{T} \phi_{t,\text{defo}} - \underline{\hat{\phi}}_{t,\text{defo}}}{T-1}}, \tag{5.9}$$

where T is the amount of observations, $\phi_{t,\text{defo}}$ is the simulated displacement phase at time $t$ and $\underline{\hat{\phi}}_{t,\text{defo}}$ is the adjusted displacement phase at time $t$.

## 5.6 Analysis of results

In this section the results of the simulations are summarised. The full set of results can be found in appendix A, since not everything can be discussed in detail. First an overall comparison is made between the recursive estimator, where extra a priori information is added in the functional model, and the periodogram where only the (i) steady-state assumption and (ii) steady-state plus constant acceleration assumption is made. This is done in order to evaluate the improvements that this kind of a priori information can cause. Next, the sensitivity of the results to the different input parameters is investigated. This is of importance, because in the first comparison the input parameters are chosen in such a way that they are assumed to be optimal for the simulated data. In practice the exact behaviour of subsidence can only be approximated, so it is of importance to know how precise this should be done.

### 5.6.1 The value of adding a priori information

The recursive estimator is in the figures referred to as *QUInSAR*. Figure 5.15 shows the performance of the recursive estimator versus the periodogram for TerraSAR-X data with a noise level of $40°$. Section 5.3.3 explained that this level of observation phase noise in the data is already on the high end of the spectrum of expected levels. For this level of phase noise the recursive estimator performs $100\%$ on all deformation types, this in comparison to the very poor performance of the periodogram with steady-state assumption, highlighted in red. The lower part of figure 5.15 displays the standard deviation of the adjusted displacement signal to the simulated signal, $\sigma_{\text{defo}}$, of correctly unwrapped arcs. The recursive estimator consistently has a high precision and is only outperformed by the periodogram on the steady-state deformation and the steady-state deformation including constant acceleration, since their functional models describes these deformation type exactly.

Figure 5.16 displays the success rates of the same deformation signals, but now being simulated using the characteristics of Radarsat-2 and ERS. The same remarks as for figure 5.15 hold, except that the performance of the periodogram improved. This is caused by the fact that the deformation is 'observed' with *C-band* which has almost double the wavelength as *X-band*, i.e. the same deformation signal spans fewer wave cycles and thus is more likely to unwrap correctly with steady-state models.

Table 5.1: Relation between time and $n_{\text{init}}$. Time values are shown in years. Whereas 15 observations corresponds to half a year of TerraSAR-X data, for ERS this corresponds to almost triple that amount; 1.4 years.

| $n_{\text{init}}[-]$ | 15 | 20 | 25 | 30 | 35 |
|---|---|---|---|---|---|
| TerraSAR-X ($\Delta t = 11$ days)[y] | 0.5 | 0.6 | 0.8 | 0.9 | 1.1 |
| Radarsat-2 ($\Delta t = 24$ days)[y] | 1.0 | 1.3 | 1.6 | 2.0 | 2.3 |
| ERS ($\Delta t = 35$ days)[y] | 1.4 | 1.9 | 2.4 | 2.9 | 3.4 |

*QUInSAR* outperforms the periodogram on all deformation types, with the exception of the steady-state case with and without an acceleration term. This is because these deformation types are made by the same functional model as used by the periodograms. Nevertheless, the results of the recursive estimator are only marginally behind that of the periodogram, as can be seen in Figure 5.15 and 5.16.

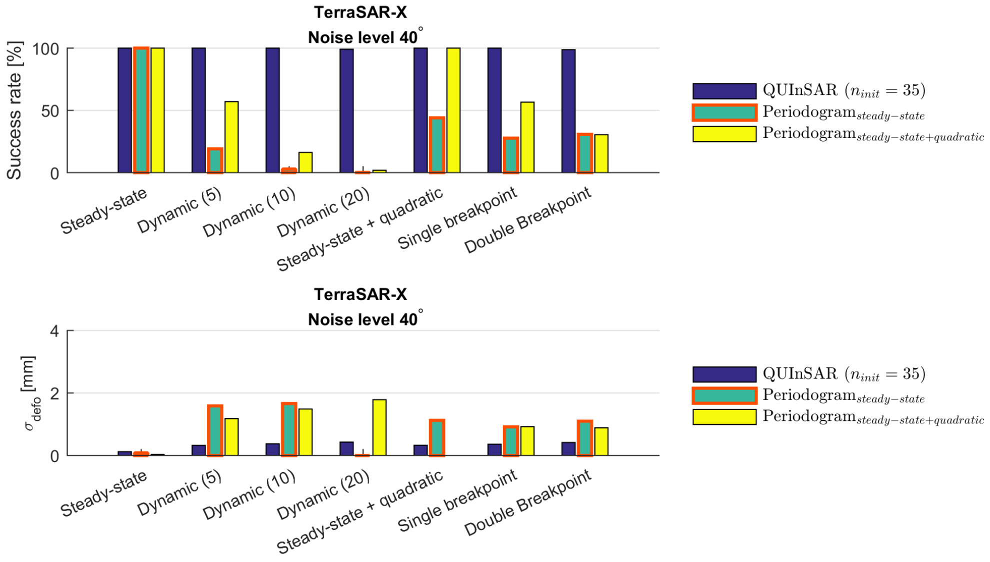


Figure 5.15: Results of recursive ambiguity estimator and periodograms for data simulated with TerraSAR-X characteristics. The top graph shows the success rates of the three methods per deformation type. *Dynamic(5)* indicates that the dynamic deformation model is used with an acceleration variance of 5 mm/y$^2$. The lower graph shows the standard deviation of the estimated displacement, $\sigma_{\text{defo}}$. The recursive estimator is initiated with 35 observations, correlation length $L_{\text{corr}}$ = 5 months, and acceleration variance $\sigma_{\text{accel}}$ = 10 mm/y$^2$. For the dynamic deformation types the same $\sigma_{\text{accel}}$ as used for creation of that signal is used; 5 mm/y$^2$, 10 mm/y$^2$ and 20 mm/y$^2$. The commonly used steady-state periodogram estimator is highlighted in red. It is clear that this estimator only performs well on the steady-state deformation type, while the recursive estimator performs successfully on all types. The $\sigma_{\text{defo}}$ graph shows that the displacement estimation is also more precise for the recursive estimator. This is not only true for the *dynamic* deformation type which is generated using a similar functional model as the estimator is using, but also for other types like the *breakpoint* models.

### 5.6.2 Initialisation length

The initialisation length determines how many observations are used for estimating the initialisation parameters. Using a fixed amount of observations results in different periods of time for different nominal repeat cycle times. For evaluation purposes the range of used initialisation lengths is $n_{\text{init}} = [15, 20, 25, 30, 35]$. Corresponding time values can be found in table 5.1.

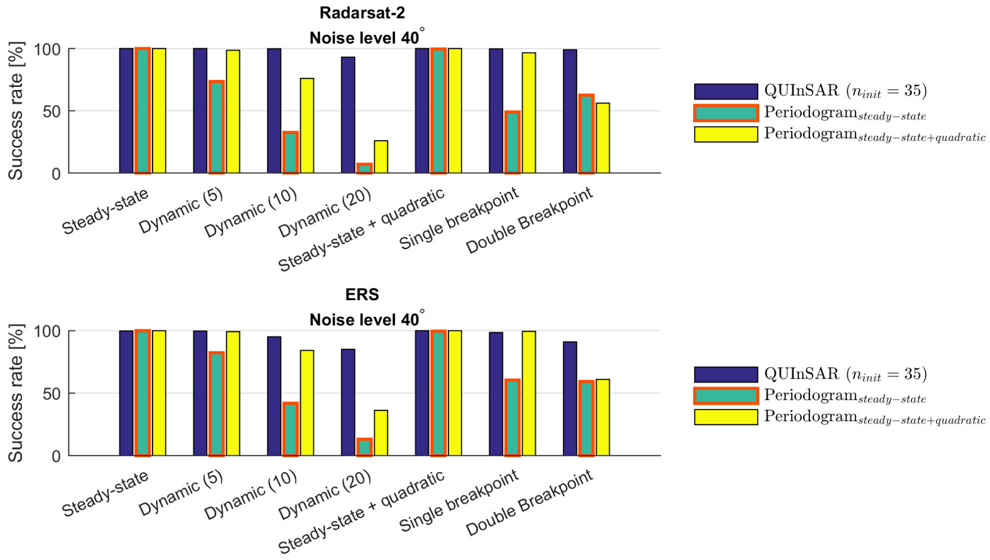


Figure 5.16: Success rates for the *Radarsat-2* and *ERS* datasets. The same remarks hold as for figure 5.15. The results of ERS are slightly less than *Radarsat-2*, most likely due to the increased repeat cycle time, causing more dynamics to have occurred per observation. It is interesting to note that the periodogram with an added constant acceleration term also has an high performance, with exception of deformation cases where the dynamics are more likely to have both an accelerating and decelerating part (e.g. double breakpoint model, modelled as an increase in deformation rate on the first break, followed by a decrease to the first deformation rate in the second breakpoint).

### Steady-state success rates

First the results of the steady-state deformation are inspected. Figure 5.17 displays the success rates of the different initialisation lengths together with the periodogram results. For low noise levels ($\leq 40°$) no significant difference is observed in the performance of different initialisation lengths. For higher noise levels however it is found that success rate increases with larger values of $n_{\text{init}}$, with a value of $n_{\text{init}} = 35$, highlighted in red, performing on par with the periodogram. Figure 5.18 reveals that unwrapping errors for high noise (60°) are mainly caused by an erroneous initialisation.

### Non-linear success rates

Only one of the three tested deformation types is steady-state deformation. The other types follow more non-linear behaviour, which due to the used steady-state initialisation assumption could result in additional errors. Figure 5.19 illustrates this effect. Where in figure 5.17 the success rate increased with growing $n_{\text{init}}$ values, in figure 5.19 there seems to be an optimum. The shown values are for the *Dynamic* ($\sigma_{\text{accel}} = 20mm/y^2$) very non-linear deformation type. The lower part of the figure shows that the amount of initialisation errors increases again after a certain value, indicated by the increasing heights of the bars coloured in red. For TerraSAR-X this occurs after $0.8$ years, for Radarsat-2 after $1.3$ years and for ERS after $2.4$ years. The timespan for TerraSAR-X is shorter because of the use of X-band. This increases the displacement steps between observations when expressed in wave-cycles compared to C-band. It was expected for Radarsat-2 and ERS that the same length in time would be found. However, ERS has its optimum on twice the length as Radarsat-2. A possible explanation can be found in figure 5.17 where it was shown that ERS shows a stronger decrease of initialisation errors than Radarsat-2 when $n_{\text{init}}$ increases. This effect could be superimposing the effect of decreasing success rate for non-linear deformation, hence shifting the optimum value for $n_{\text{init}}$.

Figure 5.20 shows the success rates and initialisation errors of the exponentially decaying deformation case, for TerraSAR-X. The success rates are the lowest of all in this simulation. The lower part of the figure shows that the vast majority of errors are due to initialisation. This is not surprising since initialisation is performed on the most non-linear part of the time-series, meaning that the steady-state assumption used during initialisation does not hold. If there is reason to expect this kind of deformation in real cases, the initialisation method should be adjusted in order to increase the success-rate.

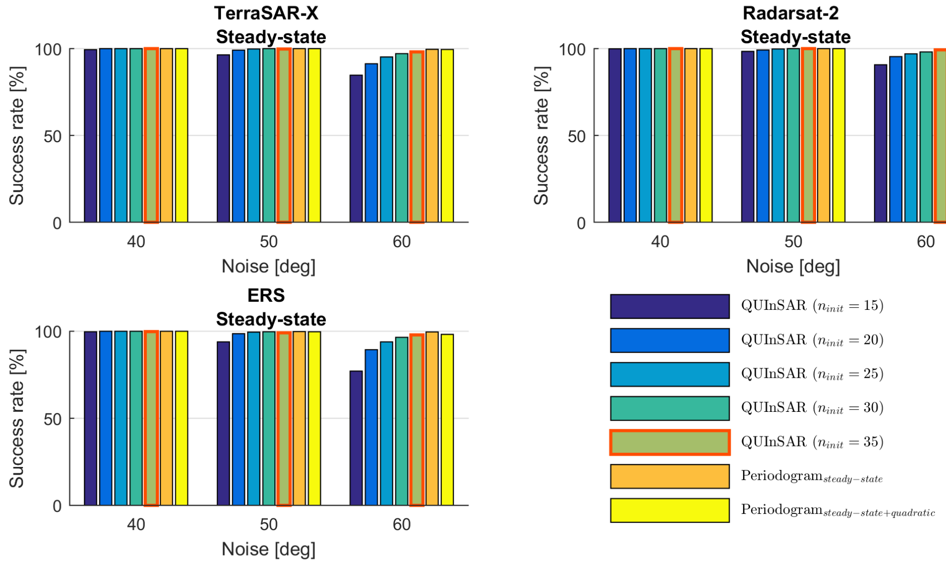


Figure 5.17: Success rates of unwrapping steady-state deformation with different initialisation lengths. The success rates of the recursive estimator are shown together with the periodogram results, which, as a batch estimator, uses all observations at once. For low noise levels (≤ 40°) the results are equally well, but for higher noise levels (50°,60°) the need for longer initialisation length is visible, in the increasing success rates when $n_{\text{init}}$ increases. For $n_{\text{init}} = 35$ the recursive estimator and the periodogram show similar results.

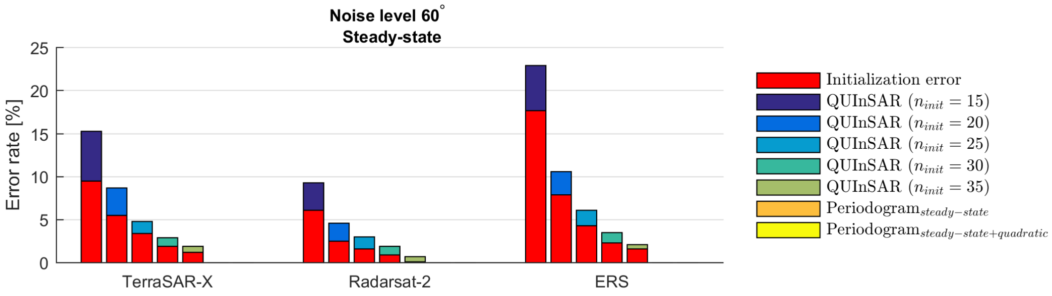


Figure 5.18: Initialisation errors of the recursive estimator for steady-state deformation and $60°$ phase noise. The larger part of errors are caused by errors in initialisation.

### 5.6.3 A priori acceleration variance

The input parameter defining the magnitude of non-linear deformation is the acceleration variance factor $\sigma^2_{\text{accel}}$. An a priori estimate of this factor is needed to describe the amount of expected diversion from the steady-state model. In this section the influence of and sensitivity to this parameter is discussed. Data is processed with values of $\sigma_{\text{accel}} = 0\,\text{mm/y}^2$, $5\,\text{mm/y}^2$, $10\,\text{mm/y}^2$ and $20\,\text{mm/y}^2$. These values are chosen, because they are the optimal values for respectively the *steady-state* and the three realisations of the *dynamic* model.

#### Results of $\sigma_{\text{accel}}$ on *dynamic* deformation models

Figure 5.21 presents the success rates of the dynamic deformation models for different values of $\sigma_{\text{accel}}$. Per graph, the dataset using the optimal value of $\sigma_{\text{accel}}$ is highlighted in red. As expected these values result in the highest success rates. Results of $\sigma_{\text{accel}} = 0\,\text{mm/y}^2$ are similar to the periodogram with steady-state model, showing that the functional model of the recursive estimator indeed reduces to steady-state for this value. An interesting observation is that the parameter $\sigma_{\text{accel}}$ is not very sensitive, since a high value of e.g. $\sigma_{\text{accel}} = 20\,\text{mm/y}^2$ also performs very well on the *Dynamic*($5\,\text{mm/y}^2$) and *Dynamic*($10\,\text{mm/y}^2$) data. This observation is not applicable vice versa.

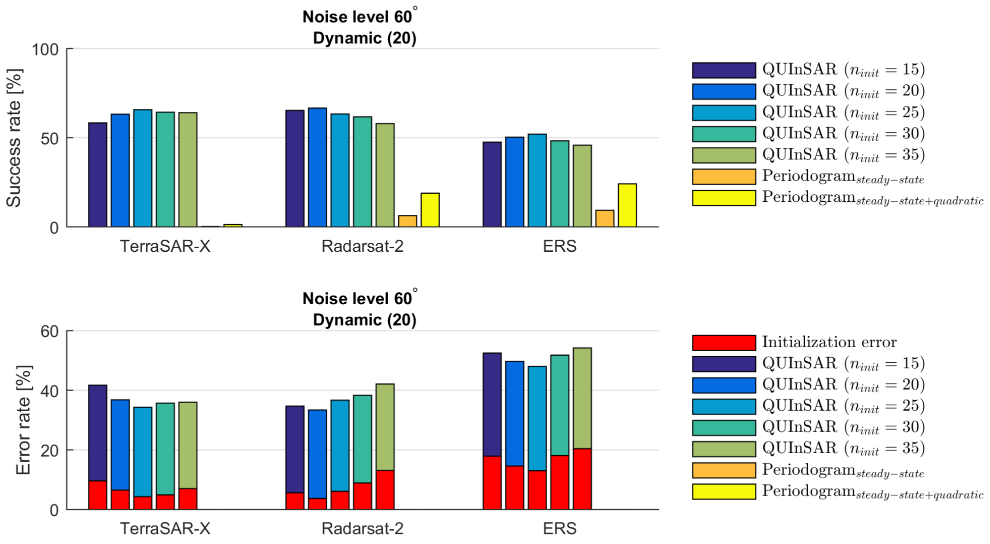


Figure 5.19: Success rates and initialisation errors for the most non-linear deformation case; *Dynamic*($\sigma_{\text{accel}} = 20mm/y^2$). The optimum values for $n_{\text{init}}$ are different per satellite. This is caused by the differences in wavelength and repeat cycle times.

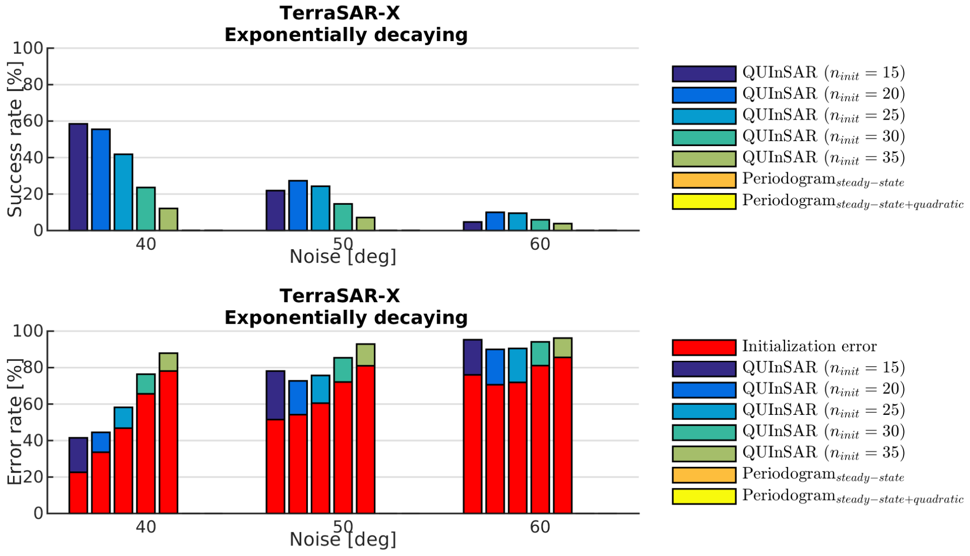


Figure 5.20: Success rates and initialisation errors for the exponentially decaying deformation case. The majority of errors are caused by initialisation. The high non-linear behaviour of this deformation type in the first year causes the initialisation, using a steady-state assumption, to fail.

The displacement standard deviations are displayed in figure 5.22. Here it is once more shown that choosing the optimal $\sigma_{\text{accel}}$ value yields the highest precision, while not being crucial for a good displacement estimation, since using other $\sigma_{\text{accel}}$ values also yield low standard deviations for the adjusted displacement. This low sensitivity of $\sigma_{\text{accel}}$ results in the fact that the acceleration can be estimated to be relatively high, while still yielding good results for steady-state deformation.

### Results of $\sigma_{\text{accel}}$ on other deformation models

The success rates of the other deformation models are shown in figure 5.23. Those models do not have a specific optimal value of $\sigma_{\text{accel}}$, since this value was not used in the creation of the deformation. Despite that, some experimentally obtained optimal values can be derived from this figure. The results

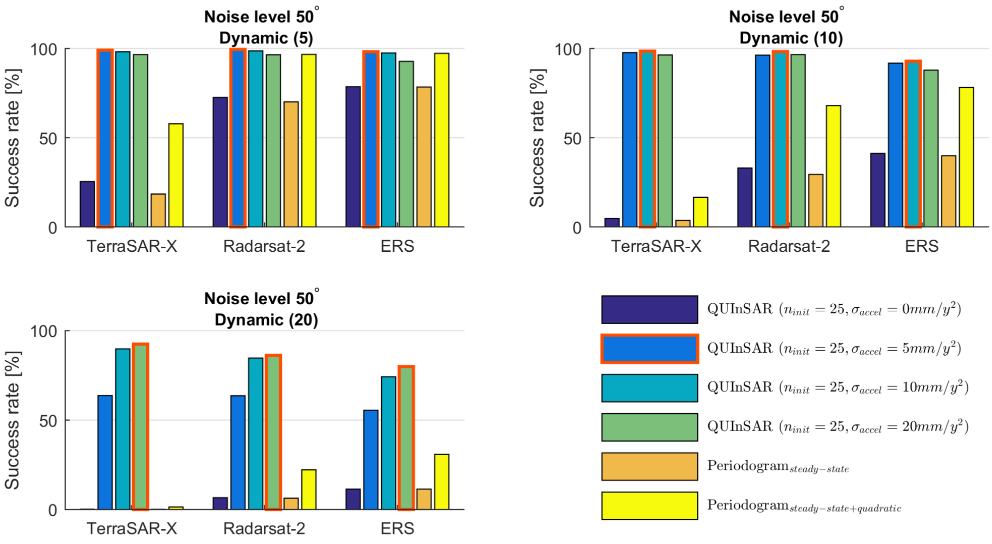


Figure 5.21: Success rates of the *dynamic* deformation types for values of $\sigma_{\text{accel}} = 0\,\text{mm/y}^2$, $5\,\text{mm/y}^2$, $10\,\text{mm/y}^2$ and $20\,\text{mm/y}^2$. The optimal values for the different deformation types are highlighted in red. Using the optimal value for the specific data type increases the success rate, but does not increase it significantly. Overestimating $\sigma_{\text{accel}}$ only lowers the success-rate by a few percent, while underestimating $\sigma_{\text{accel}}$ can drastically lower the success-rate.

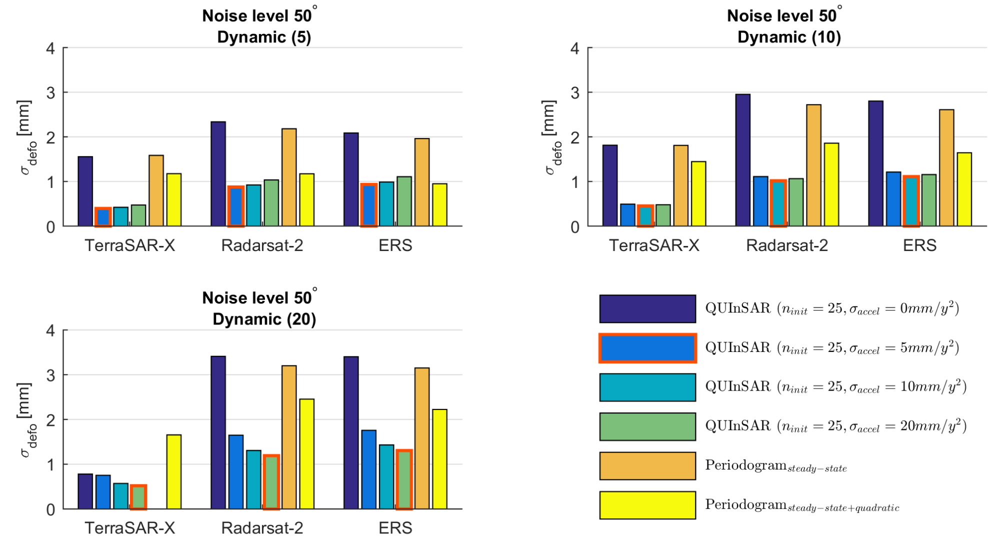


Figure 5.22: Standard deviations of the modelled displacement. The optimal values highlighted in red yield the best results. Overestimating $\sigma_{\text{accel}}$ changes the results only slightly.

for the highest phase noise level of 60° are displayed because the effect is most visible there. The highlighted values of $\sigma_{\text{accel}} = 10\,\text{mm/y}^2$ in the top graph show the best result on average. In the bottom graph however, the best results are obtained with $\sigma_{\text{accel}} = 5\,\text{mm/y}^2$, but the differences are very minimal.

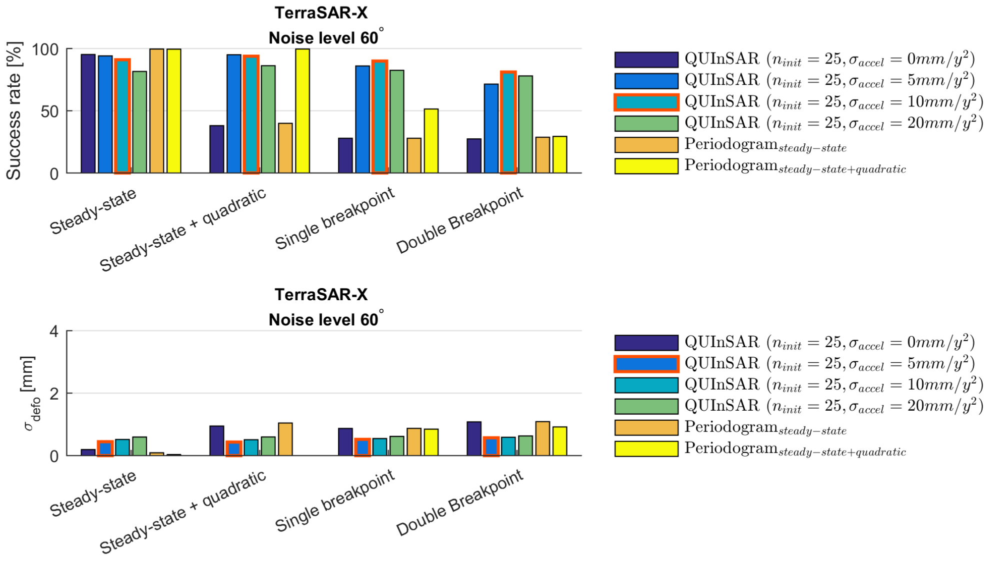


Figure 5.23: Results for TerraSAR-X on the *non-dynamic* deformation types, for a noise level of 60°. The top graph shows the success rates for values of $\sigma_{\text{accel}}$ =0 mm/y$^2$, 5 mm/y$^2$, 10 mm/y$^2$ and 20 mm/y$^2$. The bottom graph displays the quality of the estimated displacement. Depending on the expected deformation type, a value of 5 mm/y$^2$ or 10 mm/y$^2$ yields the best results.

### 5.6.4 A priori acceleration correlation length

Figure 5.24 shows the success rates of input values of $L_{\text{corr}}$ =1 months, 3 months, 5 months and 7 months on TerraSAR-X data with a phase noise level of 60°. The theoretical optimal values per deformation type have been highlighted. Only a small decrease in successes is visible for $L_{\text{corr}} = 1$ month, where a much lower value is indicated than the used 5 months for generating the data. The effect on the deformation estimation $\sigma_{\text{defo}}$ in the lower part of the figure is negligible. Similar results are found for the other deformation types and are therefore not displayed here, but can be found in appendix A. The correlation length $L_{\text{corr}}$ has arisen to have the least influence on the success-rate and displacement precision.

### 5.6.5 A priori phase noise variance factor

For computing the variance covariance matrix of observations, $Q_\phi$, a fixed phase noise variance factor is used; $\sigma_\phi^2$. The used input parameters for the recursive estimation are $\sigma_\phi$ = 30°, 40°, 50°, 60° and 70°.

#### Effect of correct $Q_\phi$ on success rate

Figure 5.25 shows the success rates for the *double breakpoint* deformation type. In red the results are highlighted that are obtained by using the simulated noise level as variance factor $\sigma_\phi^2$. The highest success rates are mostly found close to this optimal $\sigma_\phi^2$ value. Figure 5.26 shows the precision of estimated deformation for the different phase noise variance factors. The optimal values are highlighted in red and are found there where the best $\sigma_{\text{defo}}$ values are used.

#### Consequences of using a fixed variance factor

In section 4.2.2 it was argued that using a fixed variance factor is a practical simplification, not expected to have big consequences on the success rate of unwrapping. This statement can now be validated

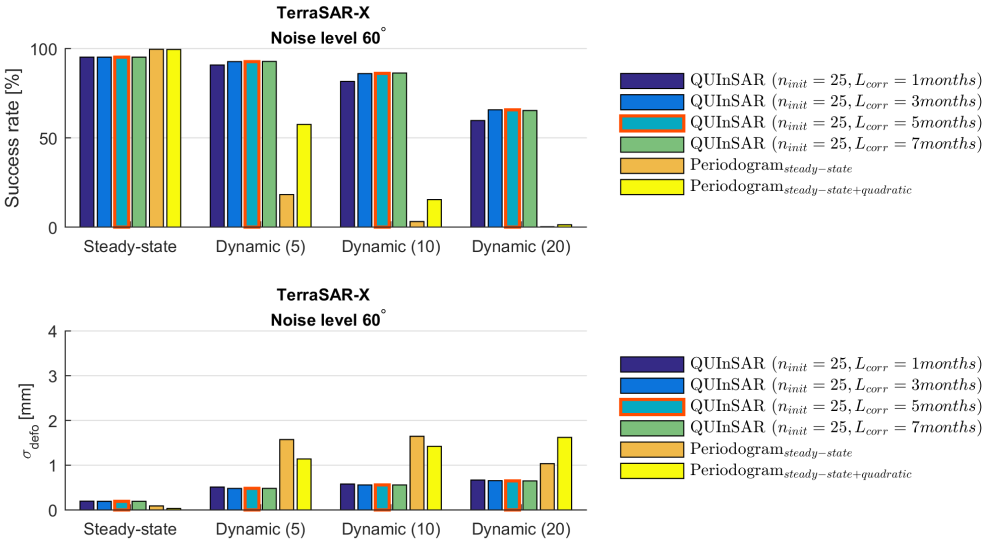


Figure 5.24: Success rates and deformation estimation precisions for the steady-state and dynamic deformation types for different values of $L_{\text{corr}}$. The dynamic data is generated with a value of $L_{\text{corr}} = 5$ months, but this is barely visible in both the success rates and $\sigma_{\text{defo}}$ values.

by looking at figure 5.27 where the success rates and deformation estimation precisions are shown for the *dynamic*(10 mm/y$^2$*)* deformation type. Using the a priori variance factor $\sigma_\varphi = 60°$, highlighted in red, has a high performance on all simulated phase noise levels, while having no noticeable effect on the precision of the estimated deformation.

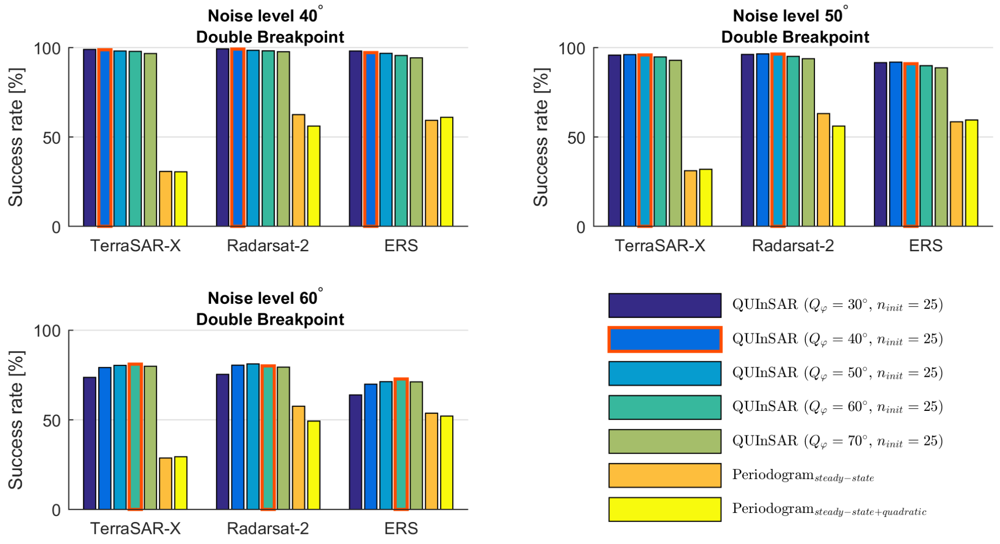


Figure 5.25: Success rates for the double breakpoint deformation type, given different satellites and noise levels. The data is processed using different values as a priori variance factor $\sigma_\phi$ = 30°, 40°, 50°, 60° and 70°. The best success rates are found when the correct variance factor value is used, but the differences are small.

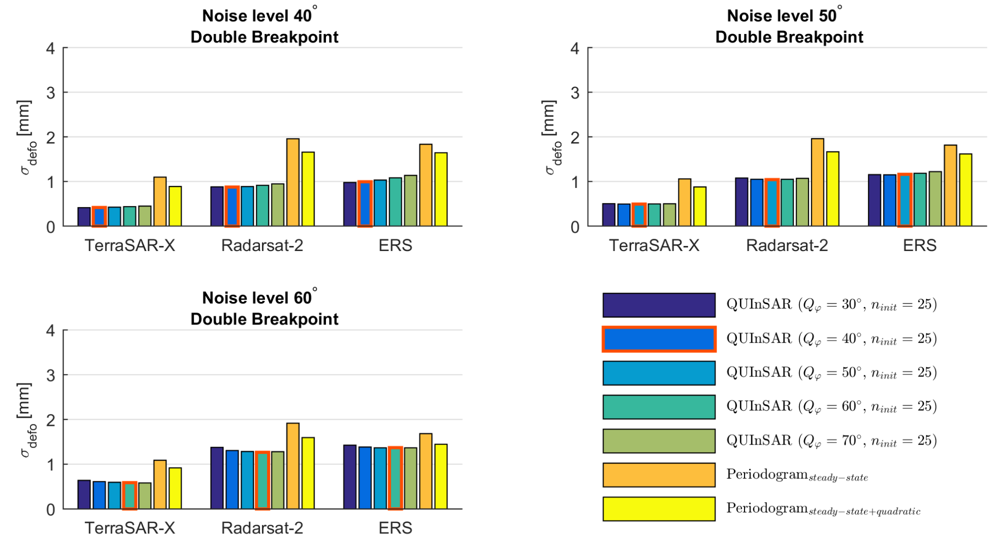


Figure 5.26: Precision of the estimated deformation when using different values as a priori variance factor. The results show that the precision is not strongly influenced by the choice of $\sigma_\phi$.

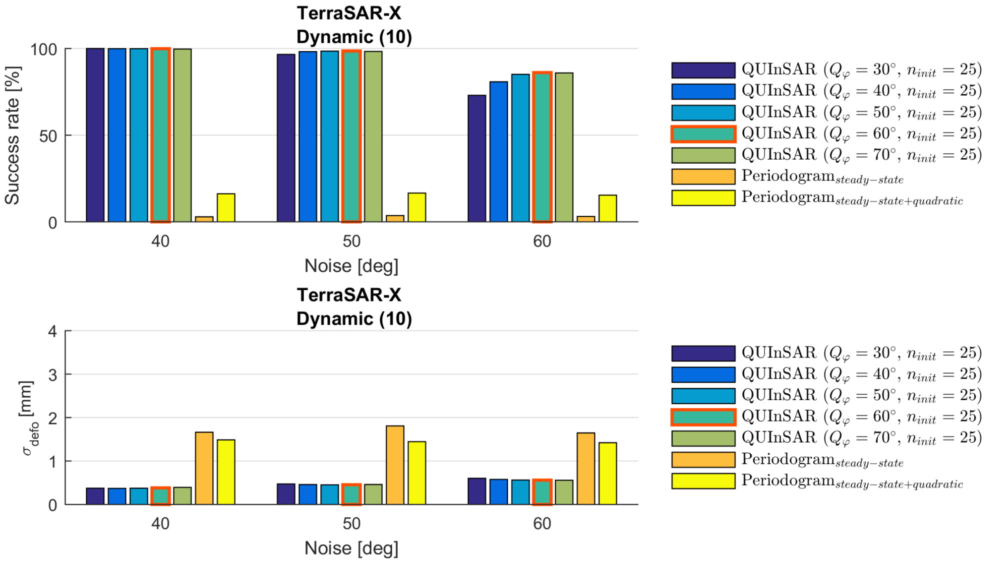


Figure 5.27: Success rates and estimated deformation precisions for the *dynamic(*10 mm/y$^2$*)* deformation type for TerraSAR-X data. High $\sigma_\phi$ also performs well on data with lower simulated phase noise levels, showing that the decision for using an a priori variance factor over a theoretically derived $Q_\phi$ is an efficient simplification.

# 6

# Application on Real Data

*Chapter 5 showed that the recursive estimator has high success-rates on different simulated deformation circumstances. A small case study on real data will be presented in this chapter as a proof of concept. The case study is centered around a building block located in The Hague that subsided significantly during a small time window. The increase of subsidence is correlated with ground water extraction in the area due to soil remediation.*

## 6.1 Description

The case study in The Hague is located around the *Anna Paulownaplein*. In this area the soil is polluted by a chemical cleaning corporation that was located at this square from 1960 until 1987. In October 2010, a soil redemption project started that extracted and cleansed the groundwater ["Municipality of The Hague", 2016 (accessed August 22, 2017]. The area processed for this case study is illustrated in figure 6.1. It consists out of the main square and surrounding building blocks.

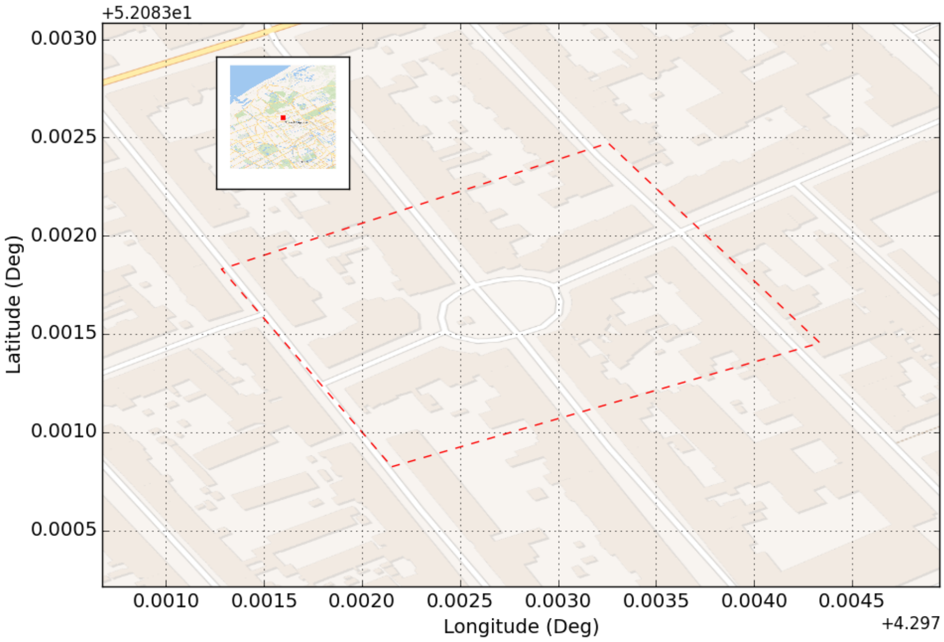


Figure 6.1: Extent of the area processed for the *Anna Paulownaplein* in the city of The Hague [Ku, 2017].

## 6.2 Data & input

### 6.2.1 Stack

The area of interest (AOI) falls within a *TerraSAR-X* descending stack of track 048. This stack has 235 acquisitions, ranging from *08-Apr-2009* to *28-Jun-2017*. The overview of acquisition dates is displayed in figure 6.2 and the extent of the used *TerraSAR-X* stack is provided in figure 6.3.

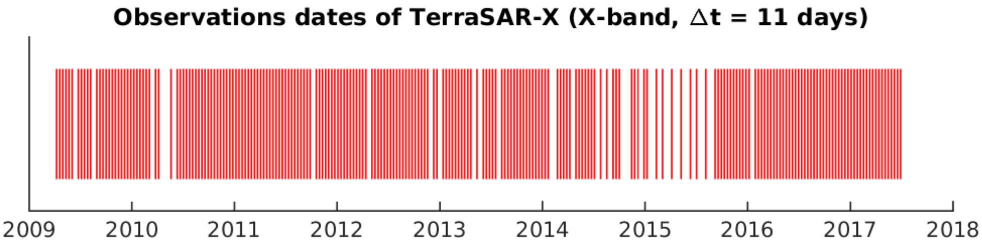


Figure 6.2: Observation epochs of the used *TerraSAR-X* descending stack of track 048 with 235 acquisitions (see figure 6.3).

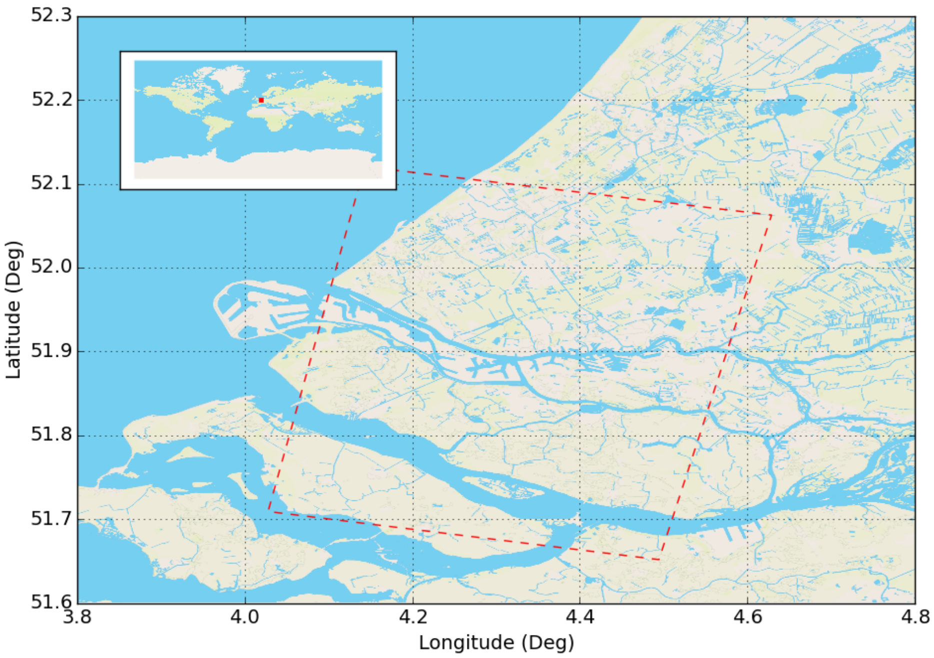


Figure 6.3: Extent of the *TerraSAR-X* descending stack of track 048 [Ku, 2017]. The stack covers The Hague, Delft and Rotterdam in The Netherlands.

### 6.2.2 Processing scheme

A simple processing scheme is used for the proof of concept. Since the selected area is only a few hundred meters in diameter, the effect of atmosphere and orbital errors are assumed to be negligible. For all pixels within the AOI, arcs are created to the same pixel, creating a star network. This pixel is chosen based on lowest normalized amplitude dispersion [Ferretti et al., 2001]. The recursive estimator processed all arcs, using the same input parameters for every arc. Thereafter, it is determined which arcs have a temporal ensemble coherence value higher than a certain threshold. The arcs that pass the threshold are unwrapped once more, now while using the periodogram with a steady-state assumption. This scheme is used to show single-arc results of real data, processed with two different methods. The single-arc part is important because it excludes influences in the final result caused by networks that are often used as base for unwrapping.

In the following sections, the methodology and results of both methods are given. Finally the two datasets will be compared in the analysis.

### 6.2.3 Input parameters

Based on the parameterisation described in chapter 5, and knowledge about the soil remediation, the choices for the input parameters are described below.

#### Initialisation length

The case study data is expected to have non-linear deformation behaviour. The simulations revealed that setting the initialisation length too long would cause these kind of deformation behaviour to be wrongly unwrapped, while setting it too short would result in wrong initialisation due to the observation phase noise. The starting date of the soil remediation is known, and it is assumed that during the period before the deformation behaviour was steady-state. Therefore, the initialisation length is set to 35 observations.

#### A priori acceleration correlation length

As described in section 5.6.4, the correlation length $L_{\text{corr}}$ is an insensitive parameter. For breakpoint models, the deformation type expected in this case study, it was shown that there is only a small decrease in success rate when $L_{\text{corr}}$ is set to 1 month, while the success rate stayed constant for the remaining values of $L_{\text{corr}}$. For the case-study the correlation length is set to 5 months, due to lack of further knowledge about the physical signal.

#### A priori acceleration variance

The sensitivity of the expected acceleration variance, $\sigma^2_{\text{accel}}$, is higher than that of the correlation length. No knowledge about the expected acceleration variance is available for this case study, so the a priori estimate will be based on results of the simulations. In section 5.6.3 it was shown that a value of $\sigma_{\text{accel}} = 10\,\text{mm/y}^2$ performed best for multiple non-linear deformation types, but has as trade off that some high-noise steady-state deformation might be missed (see figure A.5). Ideally, this decision should be made based on external knowledge about the deformation patterns that occurred.

#### A priori observation phase noise variance factor

Observation phase noise is described by an a priori variance factor. The simulations of section 5.6.5 showed that overestimating the observation phase noise only has a minimal effect on arcs with lower phase noise. In figure 6.3 it was shown that arcs with observation phase noise of $\leq 60°$ had a success-rate of $\geq 95\,\%$. Therefore the a priori observation phase noise variance factor is set to the theoretical maximum allowable $\sigma_{\phi} = 60°(2.6\,\text{mm})$.

## 6.3 Results

The area of the *Anna Paulownaplein* is processed, using the above described processing scheme and input parameters. The result is a data set of 4685 arcs that passed the temporal ensemble coherence threshold of 0.6 after being unwrapped with the recursive estimator. All 6224 arcs are then also processed using the steady-state periodogram and the quality values are similarly recalculated for this data set. To aid the analysis of both performances, figures with unwrapped time series are indicated. These figures display data from an arc, first unwrapped with the recursive estimator, referred to as *QUInSAR*, and thereafter unwrapped with the periodogram. The computed *Root Mean Square Error* (RMSE) and arc quality value is moreover listed in the titles to indicate the quality of fit, together with the estimated $\Delta H$ to the reference point. The arc quality represents the estimated ensemble phase coherence, see Eq. (2.14), and can be interpreted as a goodness-of-fit value.

### 6.3.1 Steady-state behaviour

To start with the analysis, it is first investigated how arcs showing approximately steady-state deformation are unwrapped. It is assumed that the arc temporal ensemble coherence, unwrapped with the periodogram under the steady-state assumption, is an indicator of linearity of deformation over an arc. Hence, in order to select arcs that have steady-state behaviour, only those arcs from the periodogram dataset are chosen with a temporal ensemble coherence greater than $0.85$. The RMSE and temporal ensemble coherence values of these arcs are plotted against the quality values of the same arcs unwrapped with the recursive estimator in figure 6.4. A blue dot under the red line indicates an arc that received a higher quality value when unwrapped with the recursive estimator. The majority of points lies under the red line, which is as expected, since the recursive estimator should model most of the deformation that remain unmodelled by the periodogram, and thus yields slightly higher quality values. An example of an arc having merely the same quality value when unwrapped with both methods, can be seen in figure 6.5. An arc with higher quality after unwrapping with the periodogram can be seen in figure 6.6. The reason for the latter receiving a lower quality value, is probably due to an initialisation error. The upper graph of figure 6.6, showing the recursive estimator, seems to start with a strong uplift that eventually causes two cycle-slips before the deformation pattern resembles the deformation pattern of the bottom graph.

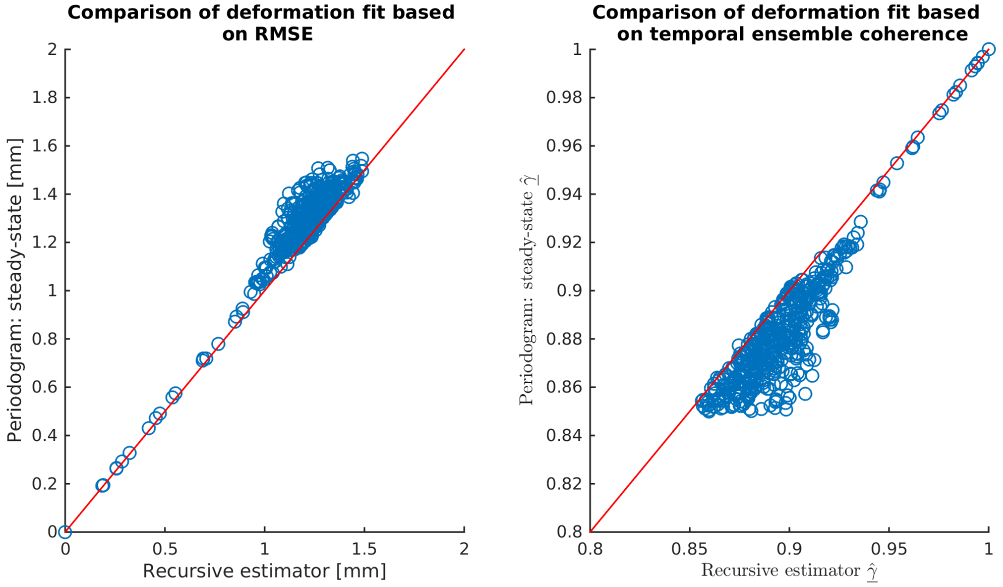


Figure 6.4: RMSE and temporal ensemble coherence of the recursive estimator versus the periodogram. The red line indicates where both metrics are equal. Left: Arcs that lie above the red line have received a lower RMSE value with the recursive estimator than with the periodogram. Right: Arcs that lie under the red line have received a higher temporal ensemble coherence value with the recursive estimator. The better quality metric value results from the lower amount of unmodelled deformation with this recursive estimator, resulting in less residuals and thus a better quality metric value.

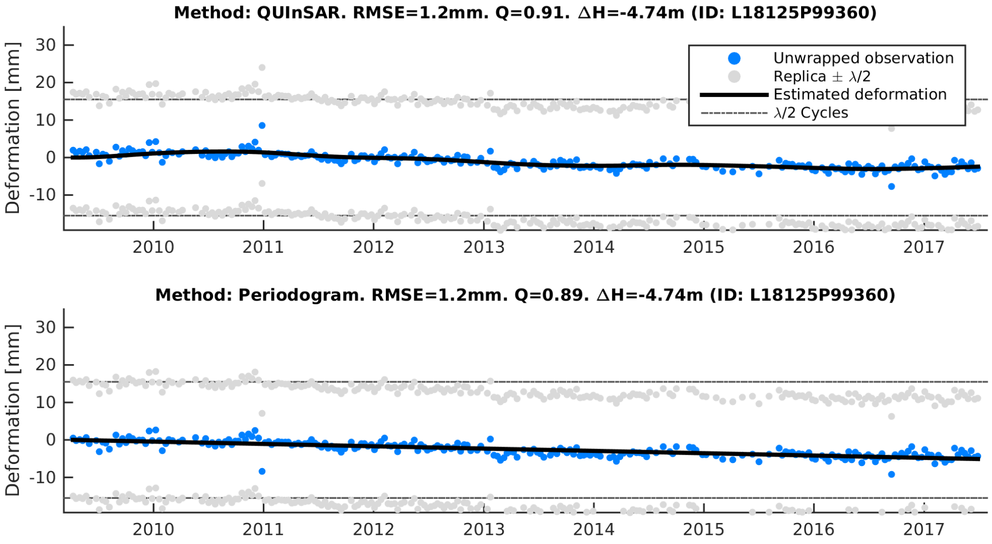


Figure 6.5: Arc with a low amount of observation phase noise and deformation behaviour that approaches steady-state. The solution provided by both methods is practically the same.

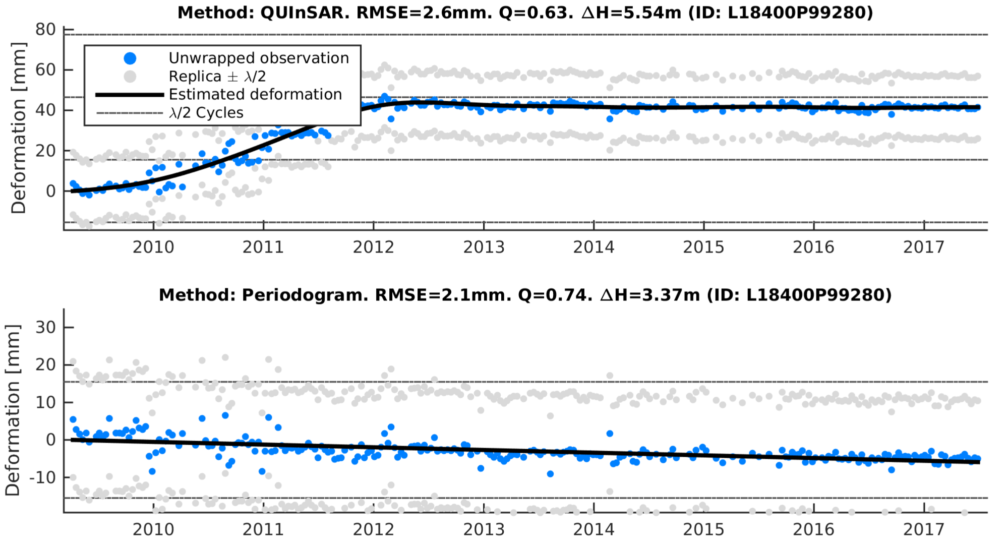


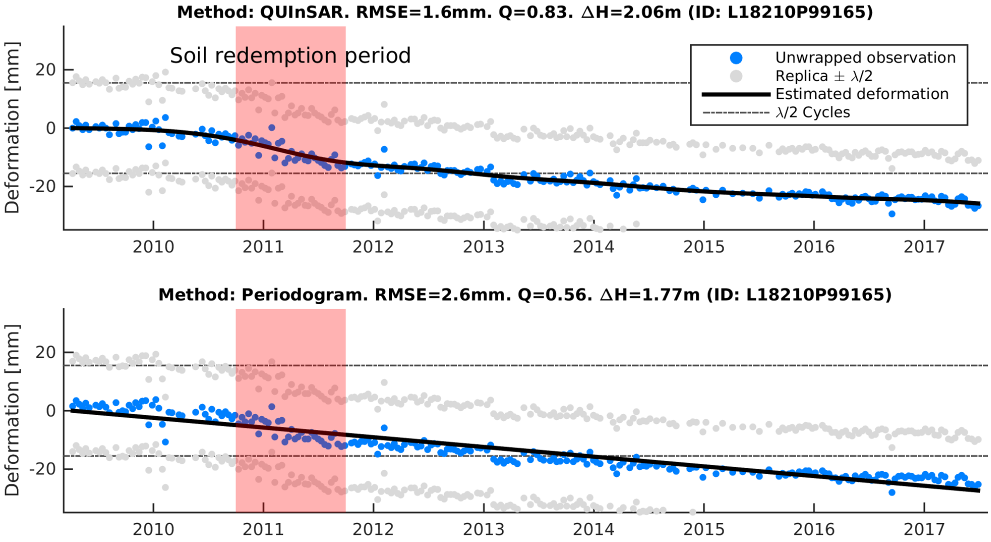
Figure 6.6: Arcs with a large difference in quality metric values between the two used methods. This difference is caused by an initialisation error in the recursive estimator, which estimated an uplift when almost no deformation was happening. After two years, *QUInSAR* recovered from the wrongly estimated uplift, and the solution becomes similar to that of the periodogram.

### 6.3.2 Examples of non-linear deformation

Within the area of interest resides a building block that shows non-linear deformation due to soil remediation [Hopman, 2016]. Two arcs are shown in figures 6.7 and 6.8 that were created between the reference point and this building block. It can be seen that in the time frame of the soil remediation, marked in red, the deformation rate increased temporarily. The periodogram of figure 6.7 was able to capture this completely whereas the periodogram in figure 6.8 resolved a steady-state deformation by causing a cycle-slip, resulting in missing about $15mm$ of deformation. Moreover, this reduced the quality value by $0.1$, because the unmodelled deformation was interpreted as phase noise.

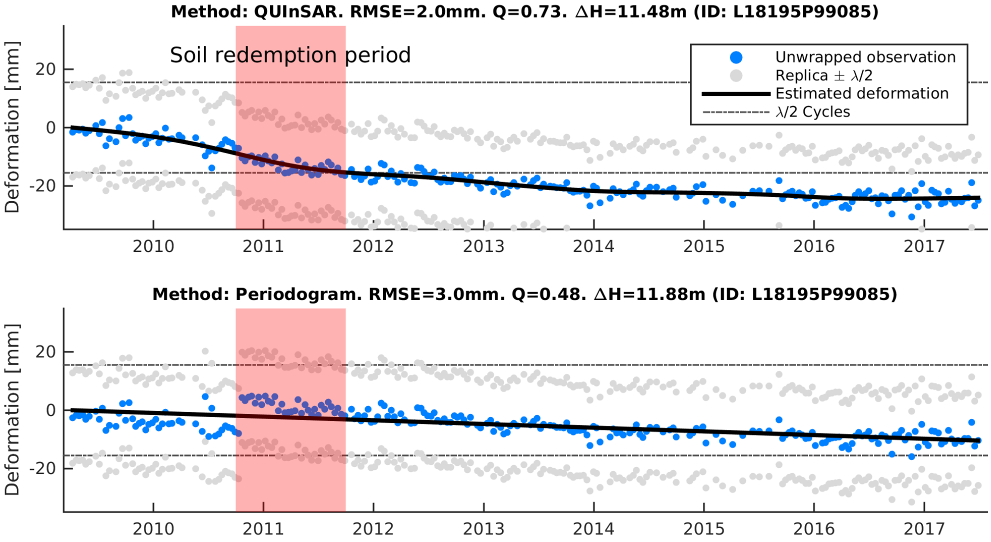
### 6.3.3 Unmodelled deformation

It is shown in figure 6.8 that the recursive estimator can increase the quality value by modelling deformation that a steady-state model would miss. This can prevent arcs from being falsely rejected in the network formation or during the final quality threshold. The term *falsely* relates to the fact that the arc might have a higher signal to noise ratio than is assumed based on the steady-state assumption. Figure 6.9 shows a RMSE and temporal ensemble coherence comparison for the whole dataset. The arcs plotted below the red line increased in temporal ensemble coherence by using the recursive estimator while the arcs above the red line decreases in temporal ensemble coherence by using the periodogram. Two extreme cases, plotted in red, are investigated because of their unexpected extreme values. First the arc represented by the top-left red dot in the RMSE plot is analysed (figure 6.10). This arc gained $0.34$ in temporal ensemble coherence when unwrapped with the recursive estimator compared to the periodogram result. The result of the periodogram got a value of $0.53$; lying outside the most used quality threshold. The low temporal ensemble coherence value is caused by a temporary subsidence around 2011 as illustrated in figure 6.10, which distorts the steady-state. The arc represented by the upper red dot in figure 6.9 is displayed in figure 6.11. Figure 6.11 clearly indicates that the reduction of temporal ensemble coherence for the recursive estimator is caused by an initialisation error from which it eventually recovers after having processed two years of observations. Although the estimation for the deformation rate after these two years is similar to the steady-state results, the deformation estimate is off by two cycles ($31mm$).


Method: QUInSAR. RMSE=1.6mm. Q=0.83. ΔH=2.06m (ID: L18210P99165)

Method: Periodogram. RMSE=2.6mm. Q=0.56. ΔH=1.77m (ID: L18210P99165)

Figure 6.7: Arc with double breakpoint behaviour due to soil remediation. The arc is similarly unwrapped with both the recursive estimator and the periodogram. Since the periodogram modelled the deformation as steady-state, the quality value of that arc is lower.


Method: QUInSAR. RMSE=2.0mm. Q=0.73. ΔH=11.48m (ID: L18195P99085)

Method: Periodogram. RMSE=3.0mm. Q=0.48. ΔH=11.88m (ID: L18195P99085)

Figure 6.8: Another arc with double breakpoint behaviour. This time the periodogram was not able to resolve the deformation pattern. One consequence of this is that the arc quality value of this point dropped, increasing the chance that this arc would be rejected in standard processing procedures. Because the arc was not rejected due to its quality value, the estimated deformation of this arc is about $15mm$ off due to the cycle slip that happened during the period of increased subsidence rate.

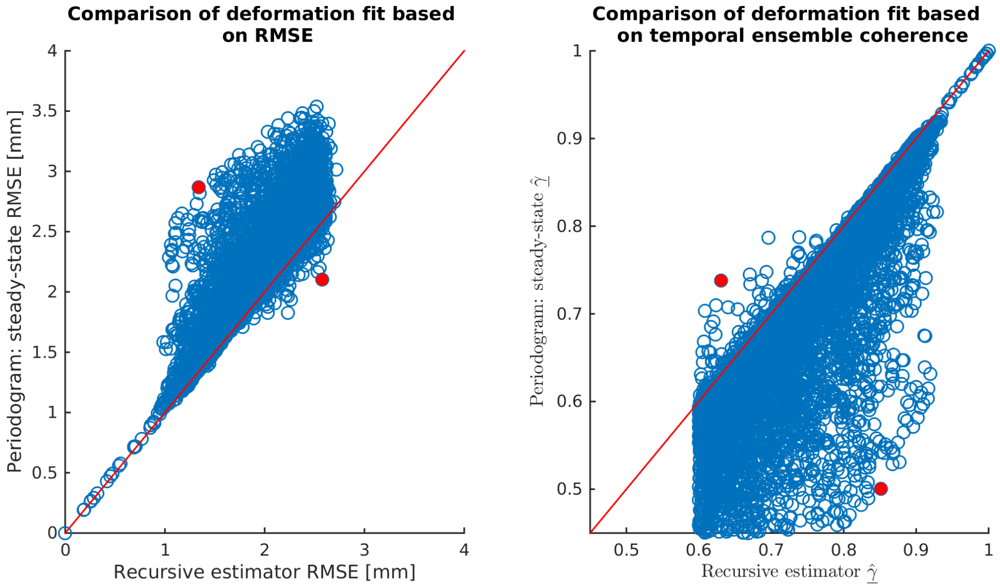


Figure 6.9: RMSE and temporal ensemble coherence of the recursive estimator versus the periodogram. The red line indicates where both metrics are equal. Left: Arcs that lie above the red line have received a lower RMSE value with the recursive estimator than with the periodogram. Right: Arcs that lie under the red line have received a higher temporal ensemble coherence value with the recursive estimator. The better quality metric value results from the lower amount of unmodelled deformation with this recursive estimator, resulting in less residuals and thus a better quality metric value. The red dots indicate the arcs with the biggest difference in quality metric between the two used methods.

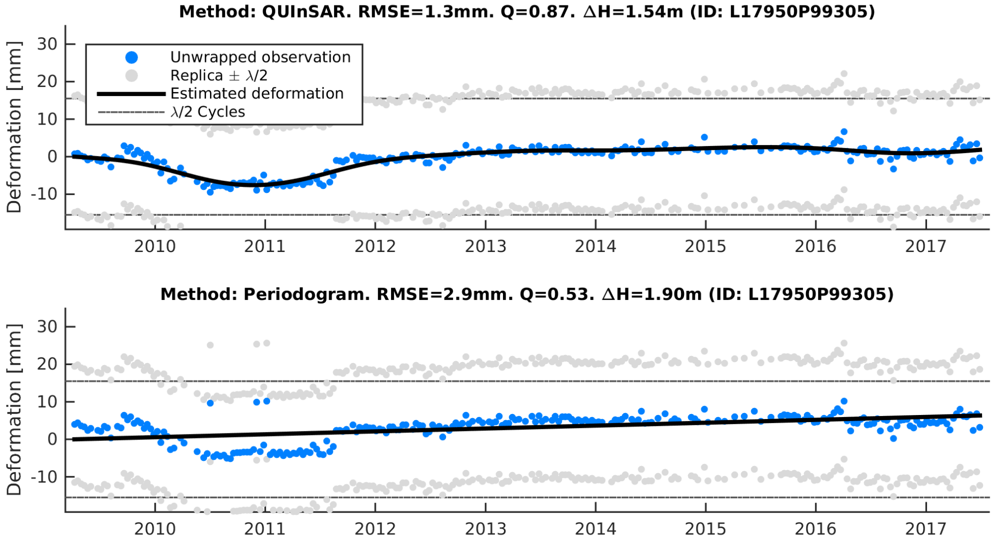
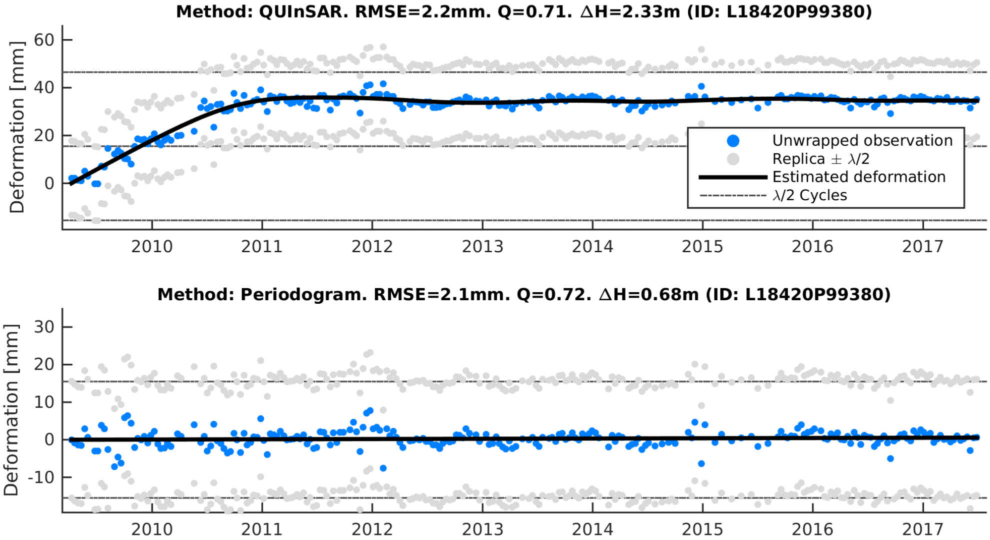


Figure 6.10: Arc that experiences some non-linear deformation in the form of temporary subsidence around 2011. The recursive estimator adapts to the local change of steady-state behaviour whereas the periodogram does not model this deformation. This results in a quality value that is 0.35 lower than that of the recursive estimator. This would mean in practice that the arc is rejected for further processing.



**Method: QUInSAR. RMSE=2.2mm. Q=0.71. ΔH=2.33m (ID: L18420P99380)**

**Method: Periodogram. RMSE=2.1mm. Q=0.72. ΔH=0.68m (ID: L18420P99380)**

Figure 6.11: Arc where the recursive estimator performed significantly less than the periodogram. Close inspection reveals an initialisation error for the recursive estimator, resulting in an estimated uplift of two cycles ($31mm$). Thereafter, the recursive estimator recovers by resolving the steady-state rate, nevertheless, the deformation estimation remains biased.

# 7

# Discussion & Conclusions

Persistent Scatterer Interferometry (PSI) is becoming a more frequently used geodetic technique for deformation measurements. However, this technique experiences difficulties with the extraction of various types of non-linear deformation patterns. The wish of near real-time monitoring and increasing sizes of datasets, moreover provides the need for efficiently updating of datasets when new satellite acquisitions become available.

This study demonstrates that recursive least-squares can be used for phase unwrapping in PSI. It moreover allows for applying a smoothness constraint in the temporal domain to help resolve different kinds of deformation types. The designed mathematical framework can therefore both aid in resolving temporal smooth deformation, and updating datasets with new acquisitions.

## Extraction of temporal smooth deformation patterns

By changing the deformation model parameterisation to that of a temporal smoothness constraint, a functional model emerges , similar to Kalman filtering, that describes deformation as a dynamic process. More types of deformation patterns can be detected via this model compared to steady-state models by adding more information about the expected temporal smoothness of the signal. The amount of unmodelled deformation after adjustment is reduced by this model, aiding a better separation of signal and noise in the time series. This results in a better quality value describing the signal to noise ratio, improving the quality subsets made during processing.

The proposed method depends on an initialisation per arc on a subset of the time series. The amount of observations used in the initialisation should be chosen based on the expected steady-state behaviour of the deformation, while maintaining the redundancy of the estimation by including a minimal amount of observations. The simulations in this study show an optimum of about $25$ observations. Using more observations has no significant result on the success-rate for steady-state deformation, while even decreasing success-rates for non-linear deformation types, due to the steady-state assumption used during initialisation. Using a fixed amount of observations results in different lengths of time depending on the nominal repeat cycle of the used satellite. The resolvability of a signal therefore strongly depends on the dynamic behaviour during the used initialisation period.

The unwrapping methods of *Integer Least Squares (ILS)* and the periodogram are discussed in this research. Both can be used for initialisation. While ILS has the advantage of providing a full variance-covariance matrix (VCM) of estimated parameters by taking into account the estimated integer ambiguities, the periodogram has as advantage to be an order of magnitude faster, though at cost of a slightly lower success-rate in the simulation. The VCM of the estimated parameters using the periodogram is approximated by applying Least Squares error propagation. An unexpected result was found when the recursive estimator was applied on real data. Despite the fact that the periodogram initialisation performed slightly worse on simulated data compared to ILS, the periodogram resulted in approximately $30\%$ more points passing the quality threshold than when using ILS.

The proposed method requires the VCM of the double-difference phase observations to be diagonal. It is shown that this simplification does not pose a problem when describing the deformation as a dynamic process. Furthermore it is shown in the simulations that using a high a priori *observation phase noise variance factor* does not significantly influence the success-rate of arcs with a lower observation phase noise than assumed a priori. However, arcs with a higher observation phase noise than assumed a priori have a lower success-rate. It is therefore concluded that an overestimated a priori *observation phase noise variance factor* can be used, that is tuned to the theoretical maximum allowed observation phase noise. The VCM of the adjusted parameters can afterwards be adjusted with an *a posteriori variance factor*, to give a better precision indication to the adjusted observations.

The difference vector, describing the deviations from the steady-state model, is in this study the key to modelling the usually unmodelled deformation when a steady-state assumption is used. It is shown that temporally smooth deformation signals can be modelled by being the result of a temporally correlated acceleration, of which the correlation is described by an exponential function. The simulations indicate that the a priori *acceleration correlation length* does not significantly affect the unwrapping success rate. In contrast, the a priori *acceleration variance* does influence the success rate significantly. Underestimating the acceleration variance causes non-linear deformation types to be unwrapped incorrectly, while overestimating only has a minimal negative effect on the success rates. This motivates the conclusion that even when the majority of the deformation is expected to be steady-state, the a priori acceleration variance factor should be tuned to the small amount of other deformation signals. Tuning the variance factor in this way avoids a bias towards steady-state deformation while it does not significantly alter the success rates for unwrapping steady-state deformation signals.

Conclusively, the recursive estimator does not only have a higher success-rate than the steady-state periodogram during simulations, but the degree of fit to the simulated deformation is also better for the recursive method. This is a trivial result since the functional method of the proposed recursive estimator incorporates more knowledge about the deformation signals. The recursive estimator, however, estimates deformation patterns that are difficult to approximate with current reduced parametric deformation models. This phenomenon is also evident in real data, as arcs with very low root mean square error (rmse) can be identified after recursive processing, while those same arcs would have been rejected when being processed by the steady-state periodogram. Even arcs with 'steady-state' deformation are better adjusted by the recursive method than by the steady-state periodogram method, since exact steady-state deformation does rarely occur in real life. Contrary to the periodogram, the recursive estimator is capable to adapt to these small deviations from steady-state deformation rate.

Nevertheless, despite promising results, some remarks have to be made about the limitations of this methodology. The unwrapping success-rate depends greatly on the adjusted parameters of an arc, computed during the initialisation. Whereas batch methods use the full time series to estimate the ambiguities, the recursive estimator only uses a subset. While the reason for this selection is well-argued, it remains a weak point of the methodology. The simulation results showed that a significant part of the wrongly unwrapped time series were caused by initialisation errors, pointing out that the current approach/methodology leaves room for improvement.

## Efficiently updating of datasets

The recursive form in which the temporal smoothness constraint is implemented in the recursive estimator, allows efficiently updating of time series as it adjusts the newest observation of an arc, based on the previously adjusted parameters. In this adjustment the unknown phase ambiguity is implicitly estimated. By only executing one extra filter step, previous filter steps and initialisation can be skipped, which improves the computational efficiency.

Performing the hindcasting step after every new acquisition may be unwanted, because this procedure changes historic adjusted observations. An optimal procedure in terms of user expectation and processing efficiency would be to update time series in *real-time* (i.e. when a new acquisition is acquired) by only doing the filter step of the recursive method. This is a highly efficient step, even for datasets containing millions of points. A full reprocessing that includes orbital an atmospheric phase screen estimation, point selection, and *hindcasting* can then be done periodically, e.g. every 3 months.

This approach reduces computational load, while increasing the possibilities of *real-time* monitoring.

Computational load of hindcasting could be even further reduced, when replacing the *fixed-interval* hindcasting method with a *fixed-lag* hindcasting method. A *fixed-lag* hindcasting only adjusts the $n$, latest observations, where $n$ is the lag-length. The *lag-length* can be based on the chosen acceleration correlation length, because both parameters assume a certain independence over larger time-spans. If there is need for further reducing the numerical complexity of the algorithm, it could be researched if changing the recursive formulation to *condition equations* [Teunissen, 2007] yields any further computational improvements due to a reduction of matrix inversions.

## Future research

Recursive estimation can be a powerful addition to PSI. The possibility to efficiently update data sets when new acquisitions are available, while at the same time resolving deformation patterns that are frequently missed with other methods, is of great value for this technique.

Furthermore, improvements can be realised with regularisation. Regularisation is used in initialisation, where it constrains the solution space. However, during the recursive estimation no regularisation is currently applied, with the exception of the zero-mean acceleration assumption. Due to the lack of regularisation, the adjustment sometimes diverges to very unlikely values of the unknown parameters.

Also, the decision for the a priori acceleration input parameters should be researched in more detail. It is of importance to let knowledge of the physical signal guide the decision for the a priori correlation length and acceleration variance. This knowledge could be obtained by expert elicitation, but a translation step has to be made from this knowledge to actual parameters. Often, large areas are processed using the same input parameters. Expert elicitation could help tune the input parameters to specific locations to have the most optimal representation of the physical signal.

In addition, more research is needed in determining the quality of unwrapping. The functional model of this recursive methodology has more degrees of freedom, which inevitably results in the risk of cycle errors due to observation phase noise. This is especially important in the processing of real data, where the actual observation phase noise does necessarily behave like Gaussian distributed noise, as described in the stochastic model.

Last but not least, the implementation of this unwrapping methodology in current processing schemes should be researched extensively. Especially the updating of the first order network should be handled with care, to assure that this network is thoroughly tested, preventing errors propagating over the dataset.

# A

# Full simulation results

*In this chapter the full simulation results figures are shown. They are ordered per satellite and per satellite the order is: initialization length, a priori acceleration correlation length, a priori acceleration variance and observation phase noise variance factor. Each input parameter has two figures: Success rates and, initialization errors or estimated deformation standard deviation. A list of result figures is given below for easy reference.*

## List of Figures in Appendix A

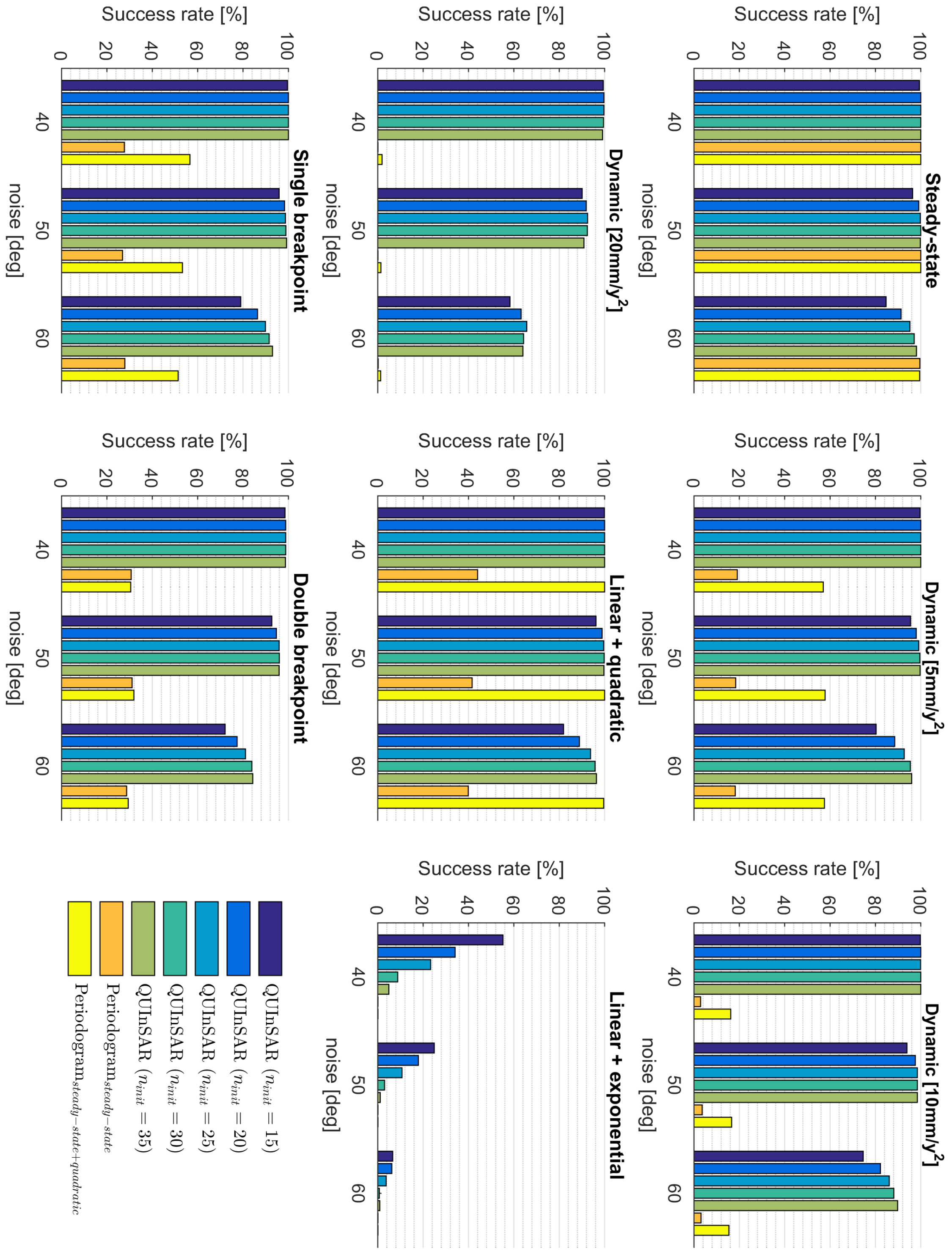


Figure A.1: TerraSAR-X success rates with varying initialization length.

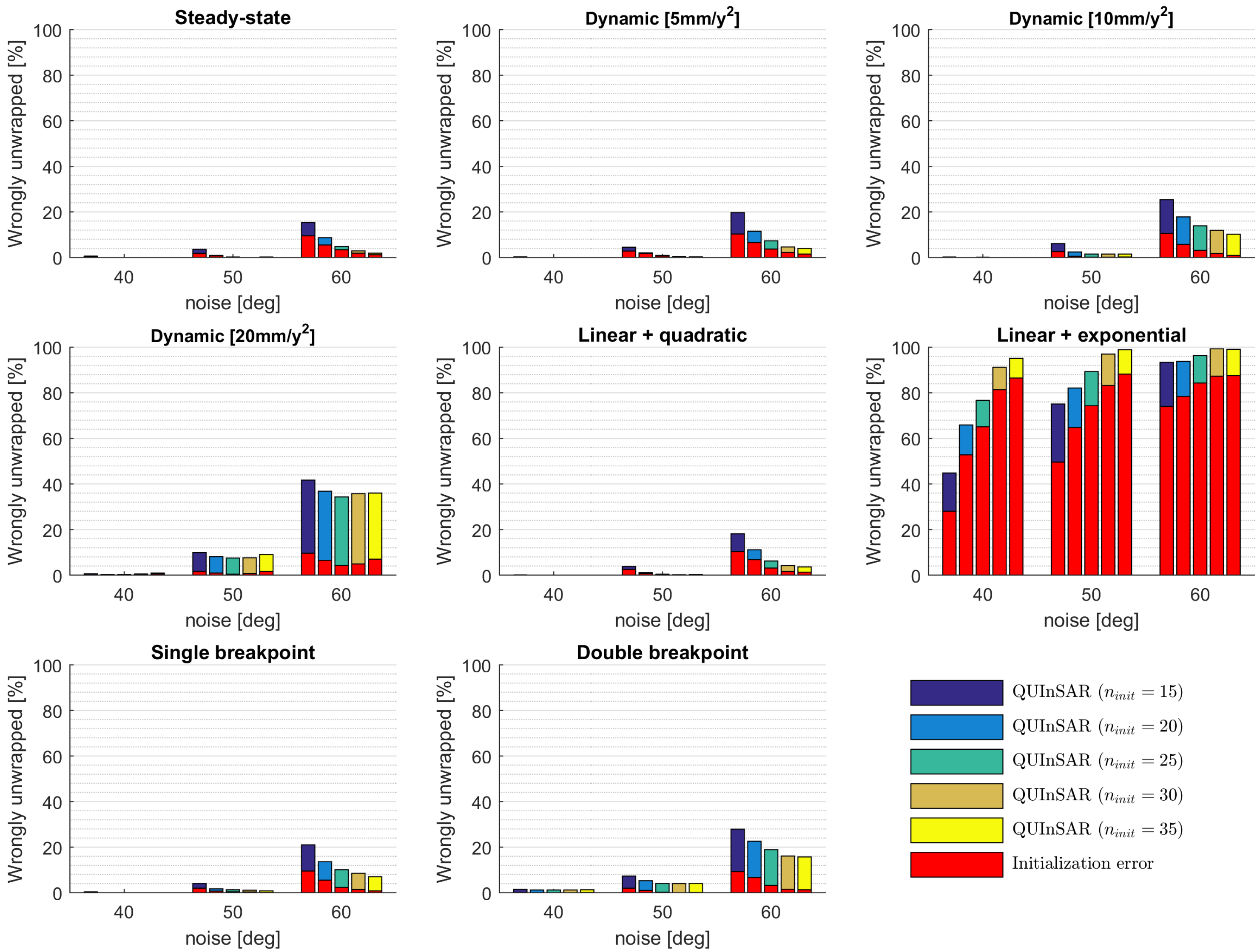


Figure A.2: TerraSAR-X initialization errors with varying initialization length.

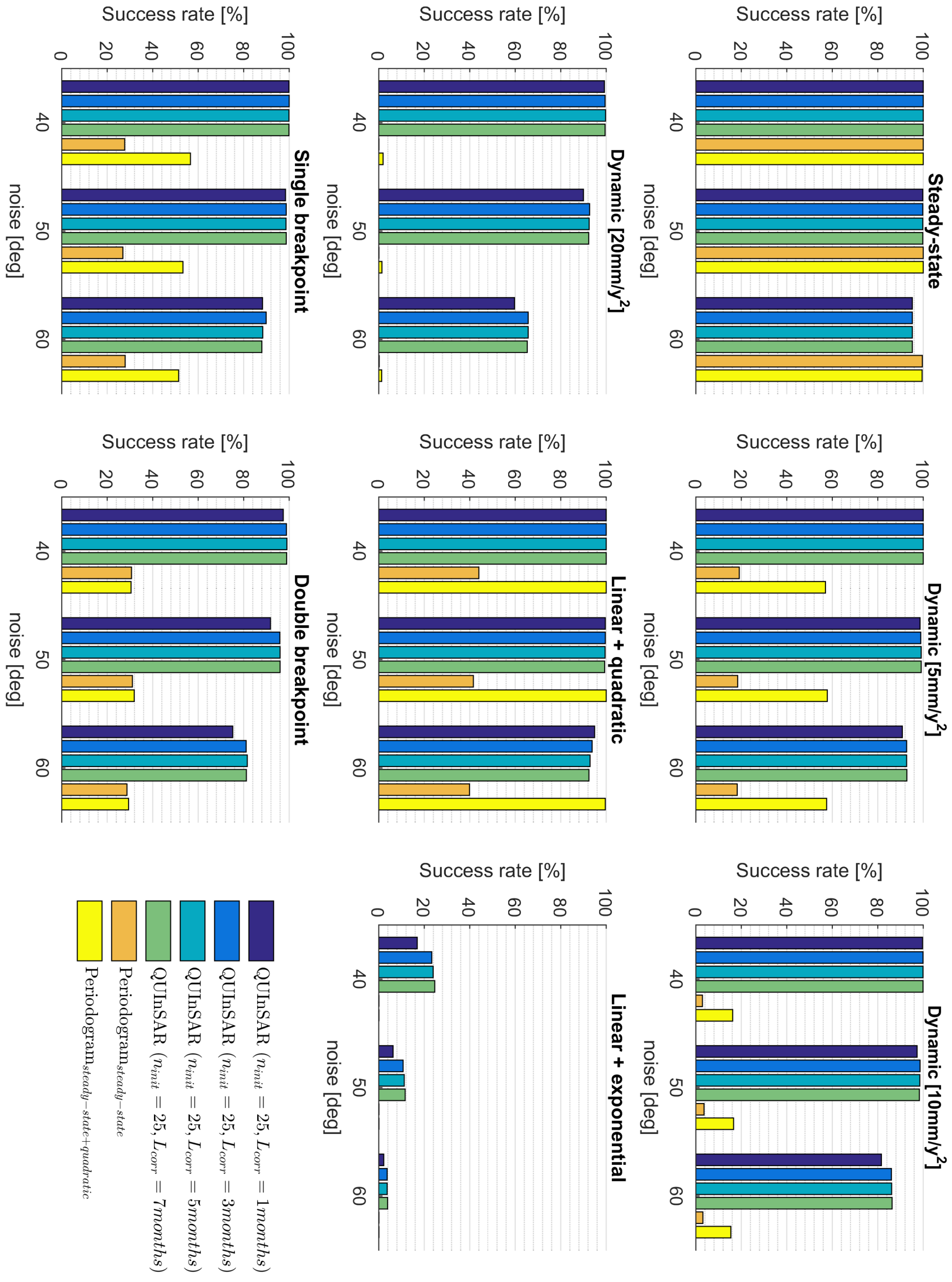


Figure A.3: TerraSAR-X success rates with varying a priori acceleration correlation length.

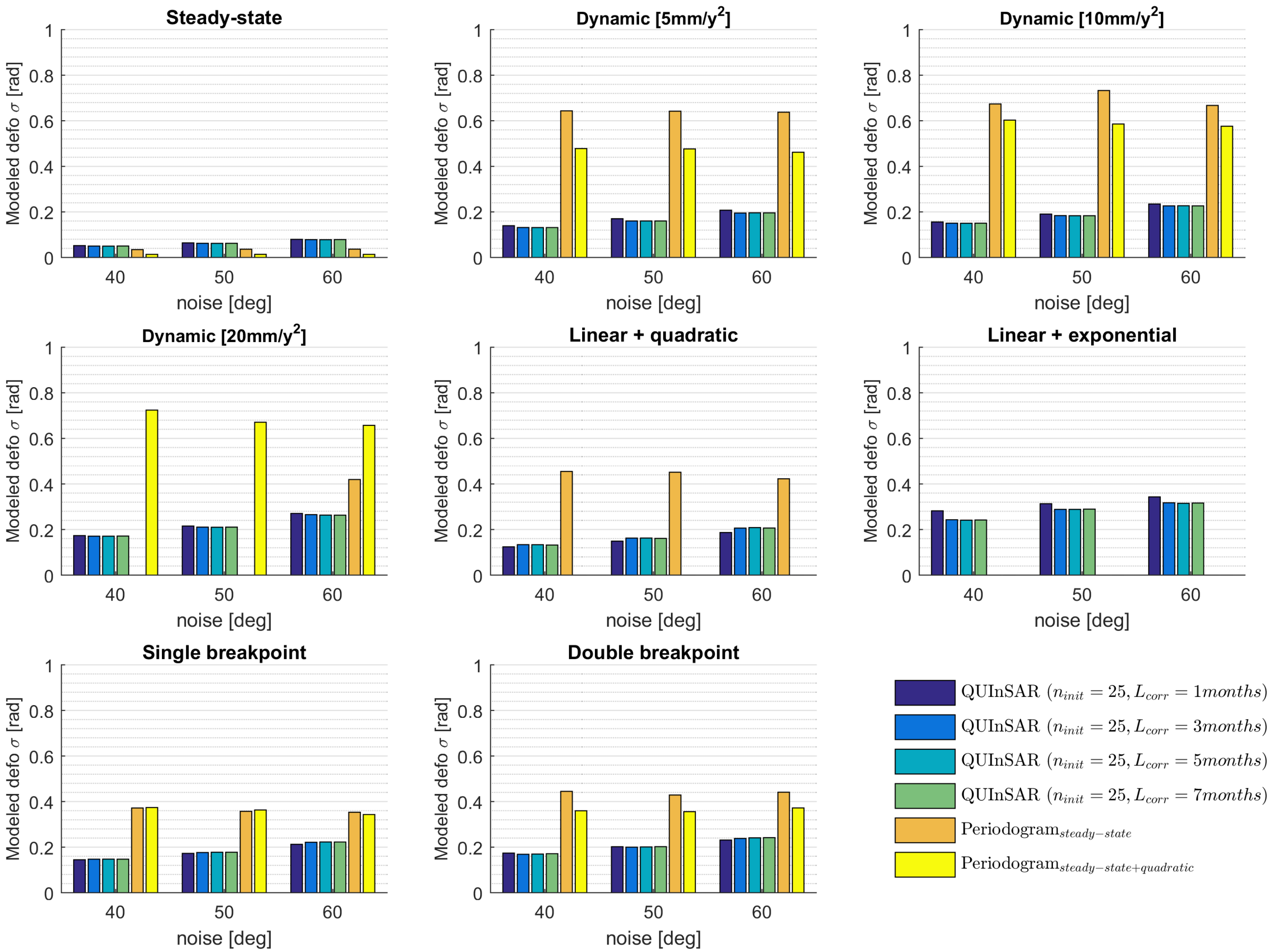


Figure A.4: TerraSAR-X estimated deformation standard deviation with varying a priori acceleration correlation length.

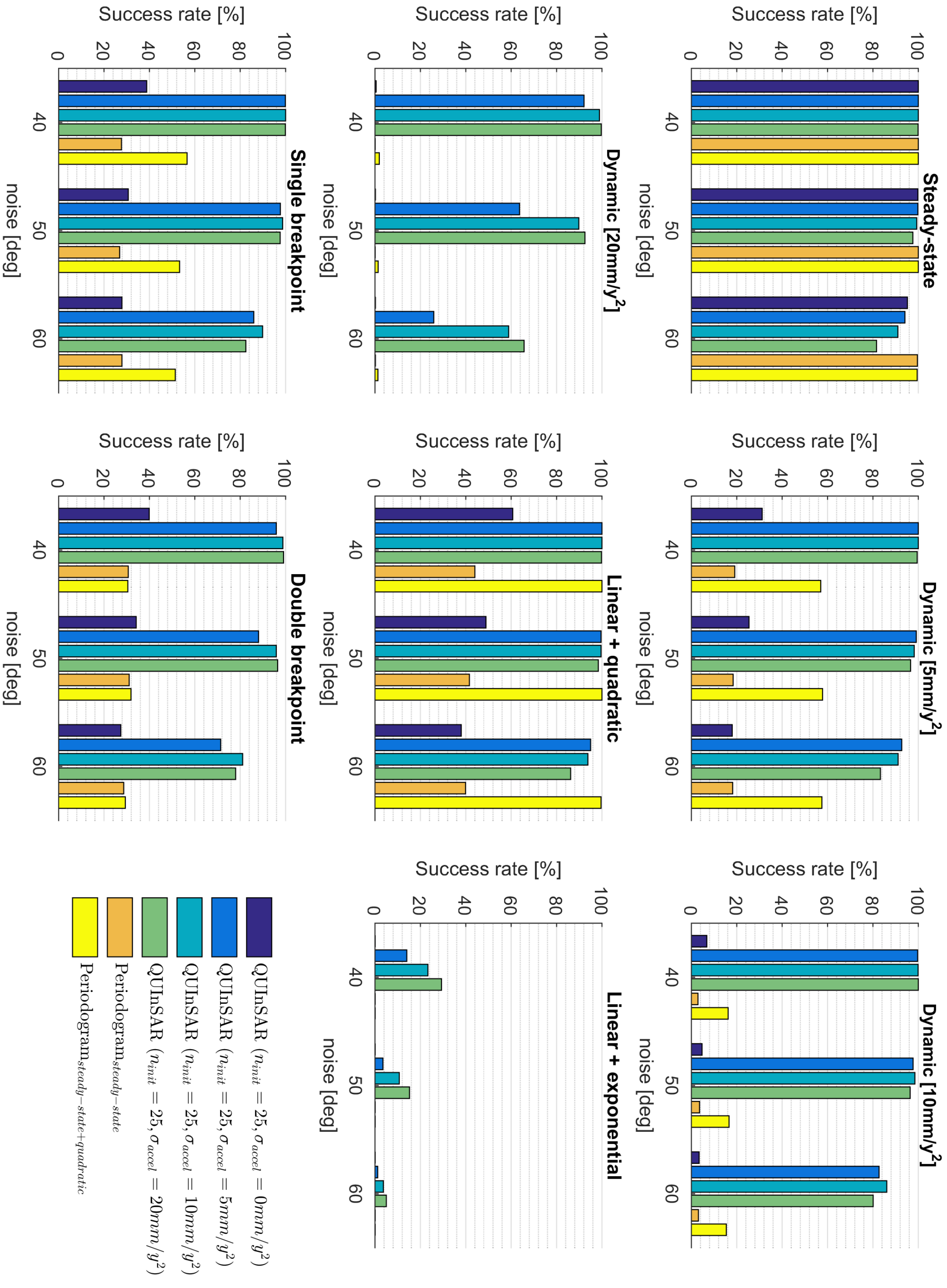


Figure A.5: TerraSAR-X success rates with varying a priori acceleration variance.

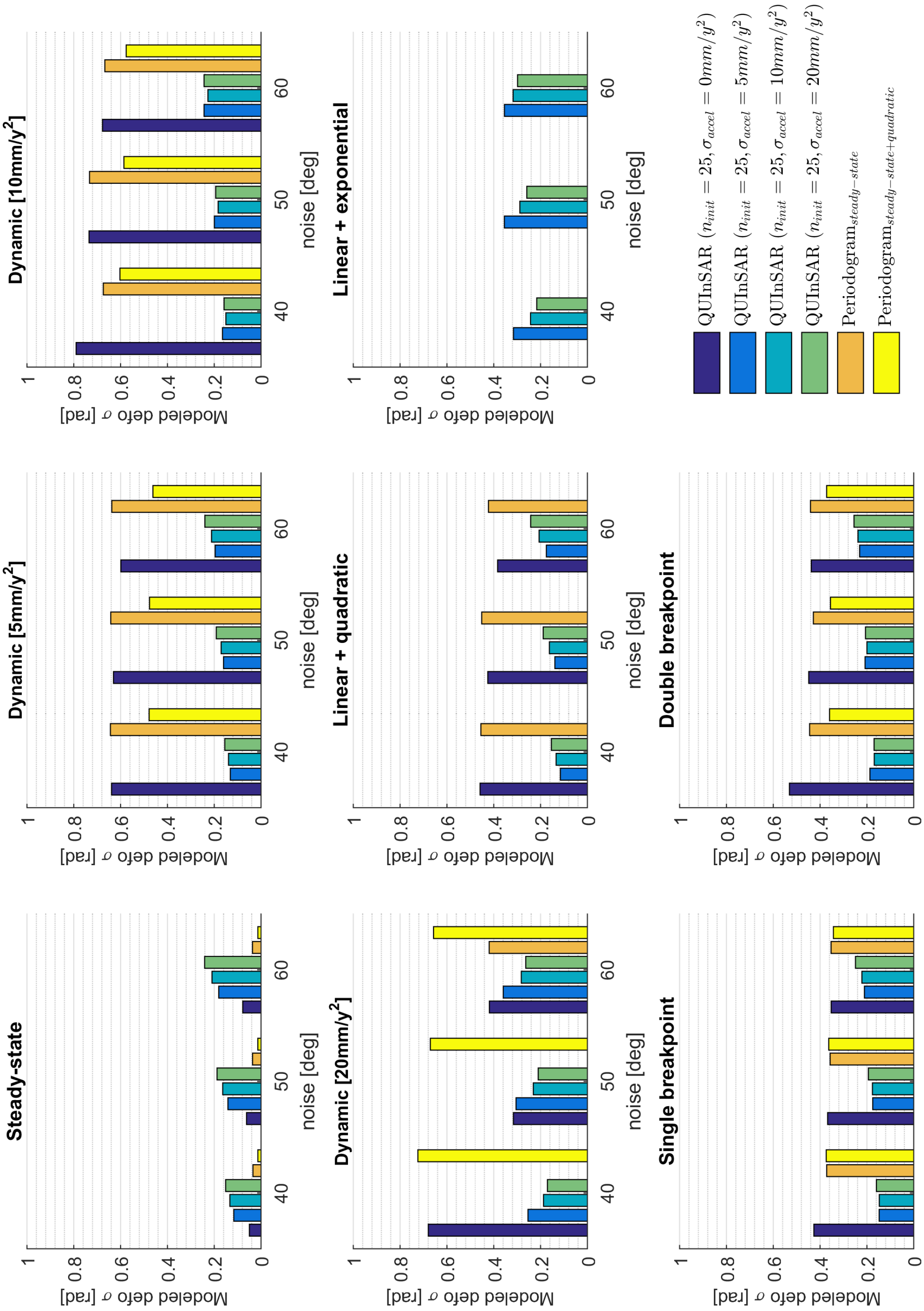


Figure A.6: TerraSAR-X estimated deformation standard deviation with varying a priori acceleration variance.

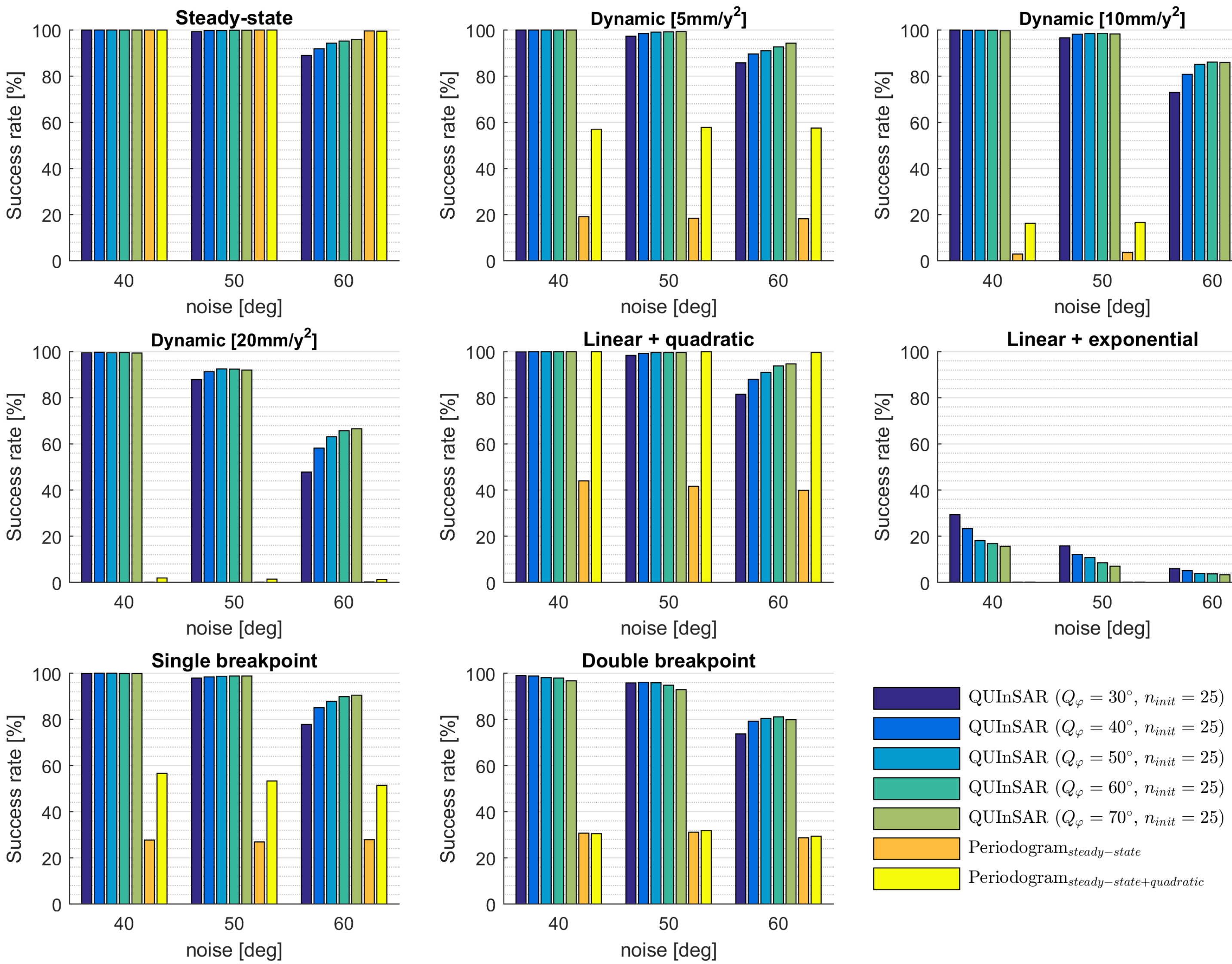


Figure A.7: TerraSAR-X success rates with varying observation phase noise factors.

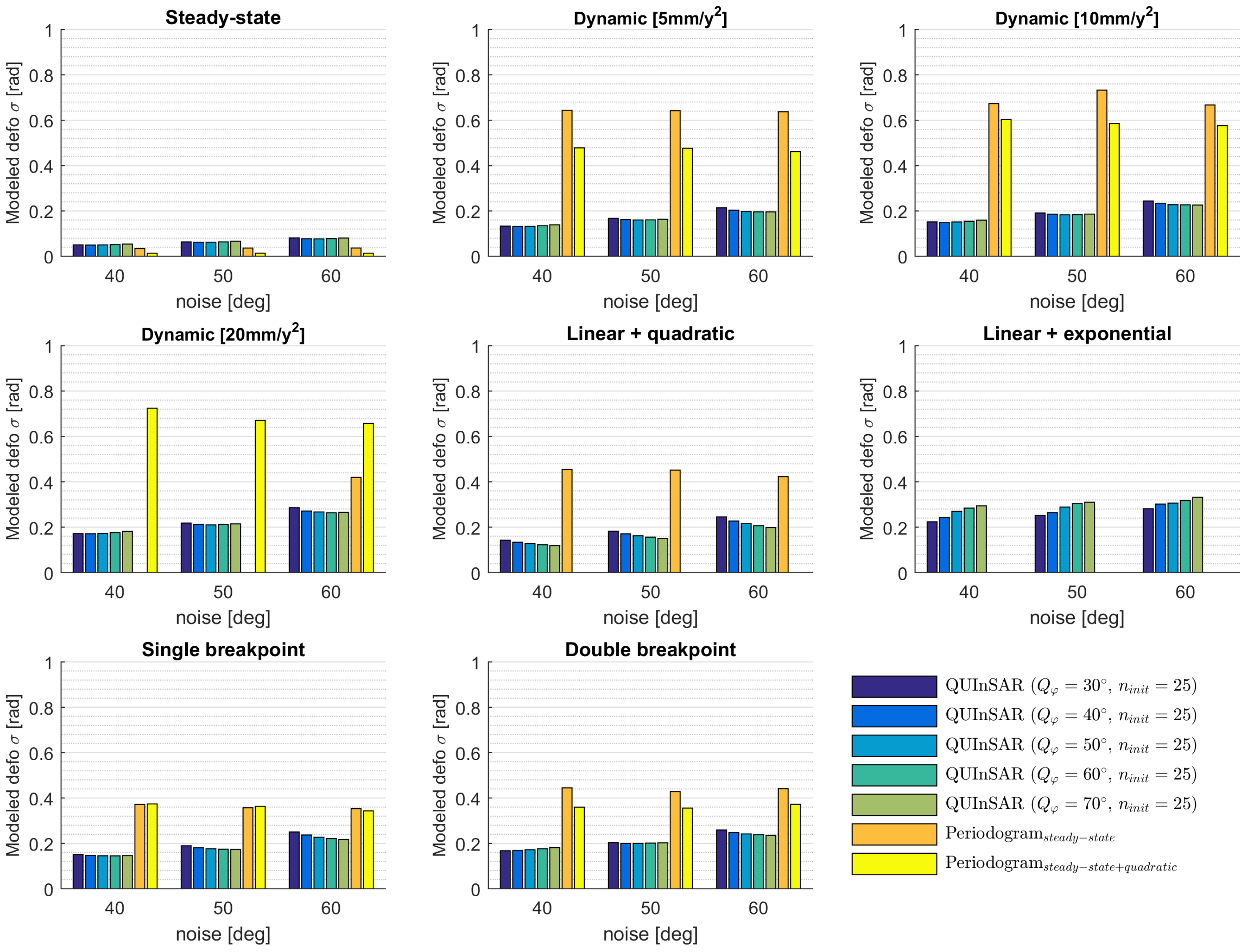


Figure A.8: TerraSAR-X estimated deformation standard deviation with varying observation phase noise factors.

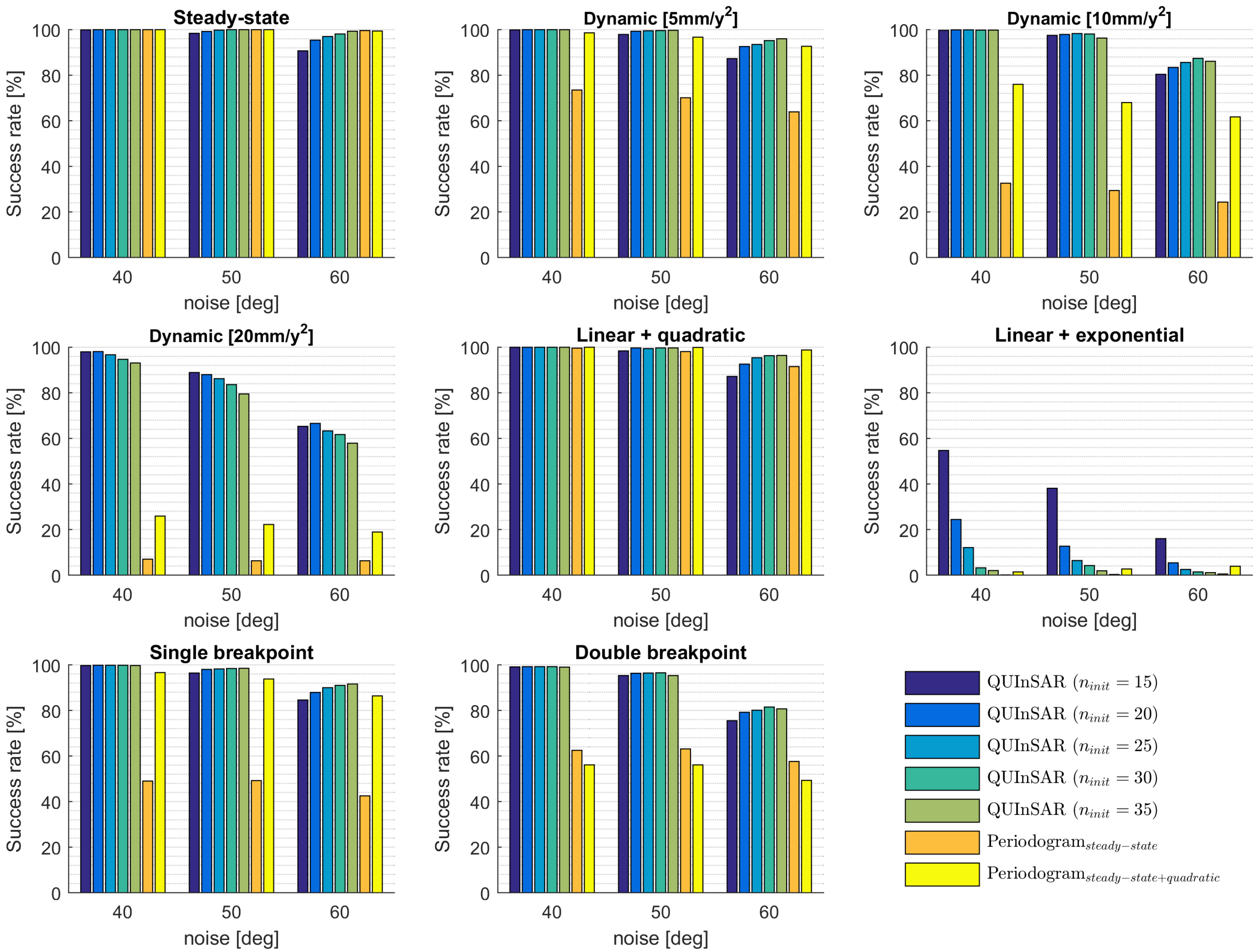


Figure A.9: Radarsat-2 success rates with varying initialization length.

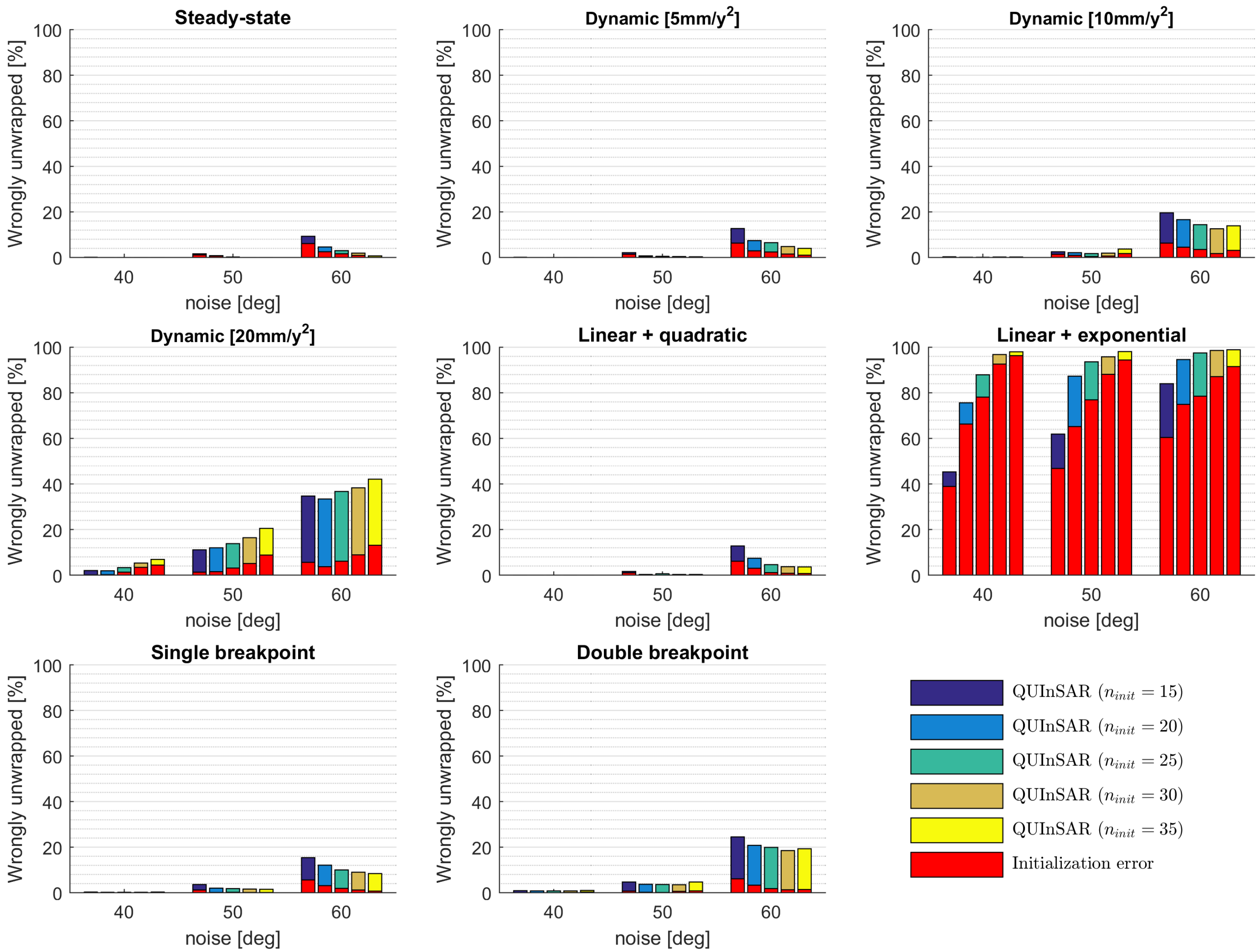


Figure A.10: Radarsat-2 initialization errors with varying initialization length.

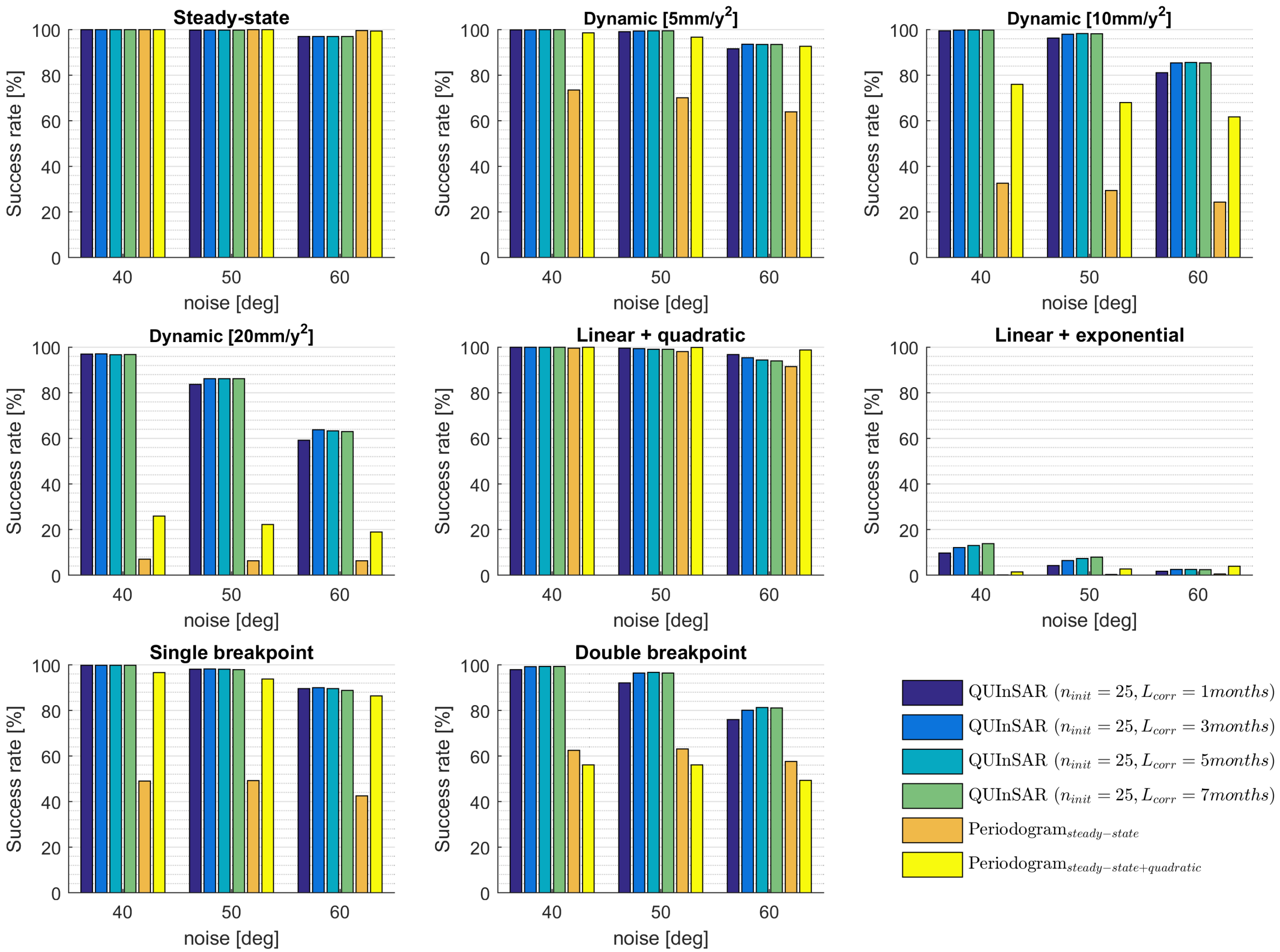


Figure A.11: Radarsat-2 success rates with varying a priori acceleration correlation length.

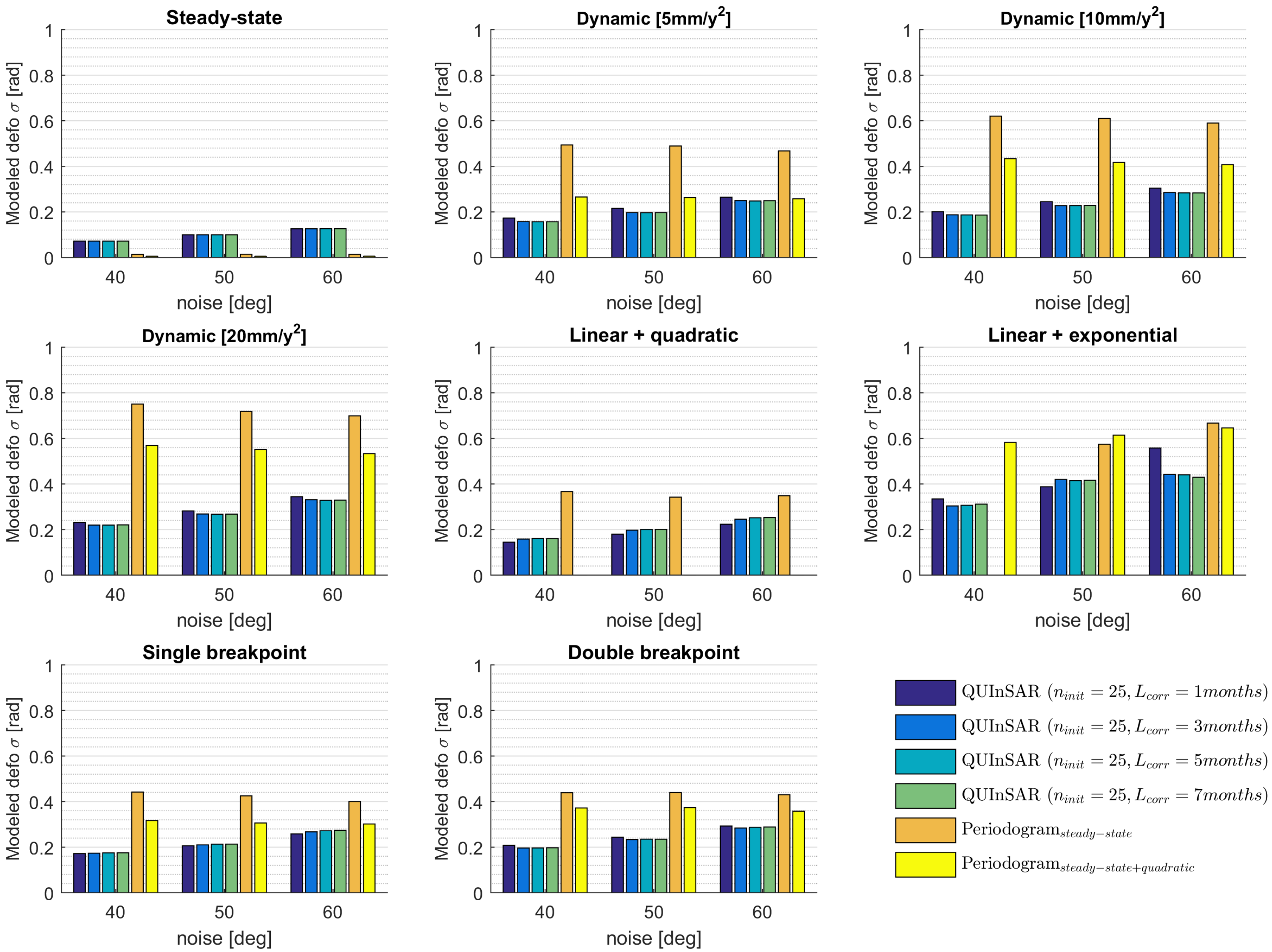


Figure A.12: Radarsat-2 estimated deformation standard deviation with varying a priori acceleration correlation length.

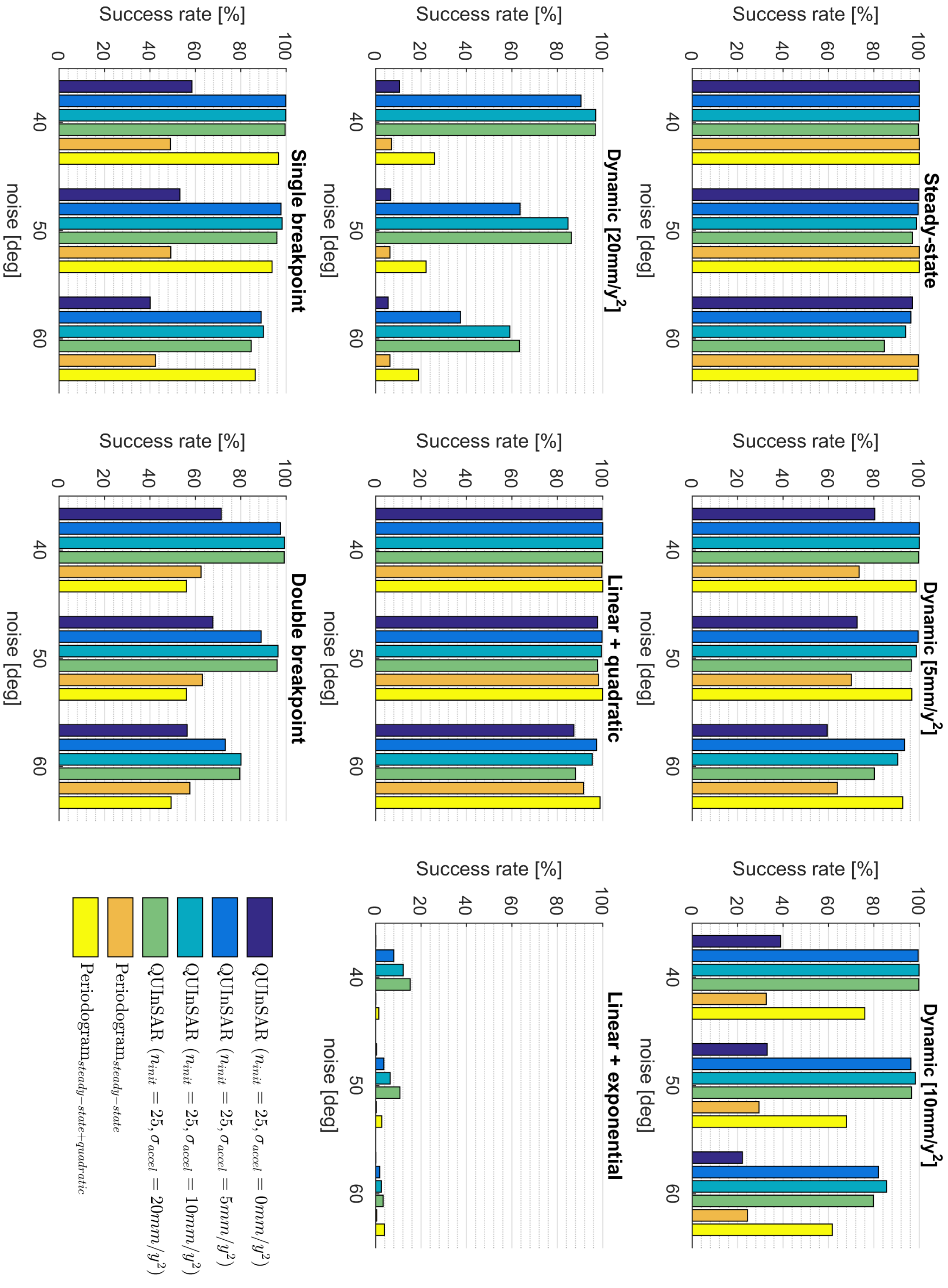


Figure A.13: Radarsat-2 success rates with varying a priori acceleration variance.

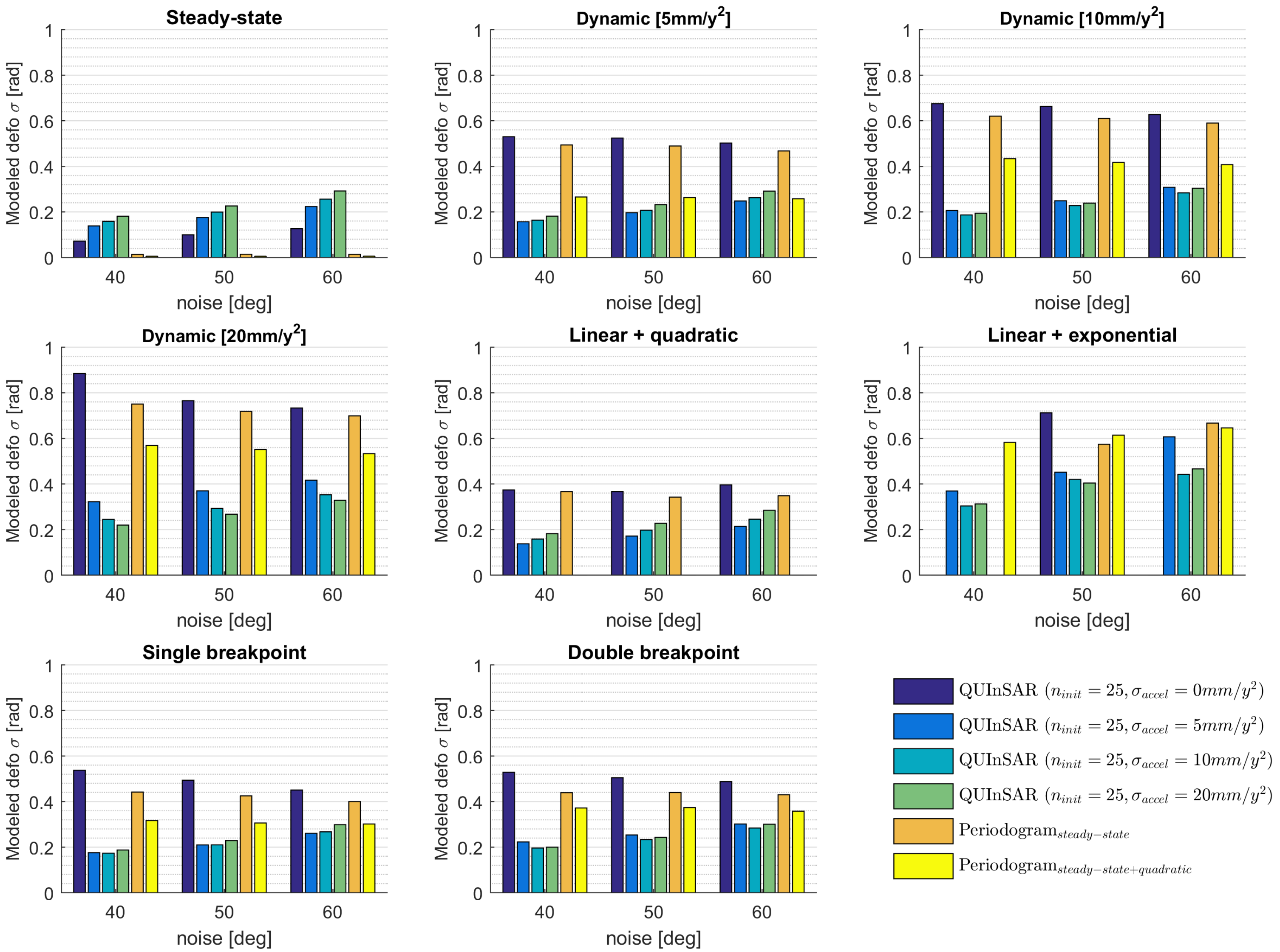


Figure A.14: Radarsat-2 estimated deformation standard deviation with varying a priori acceleration variance.

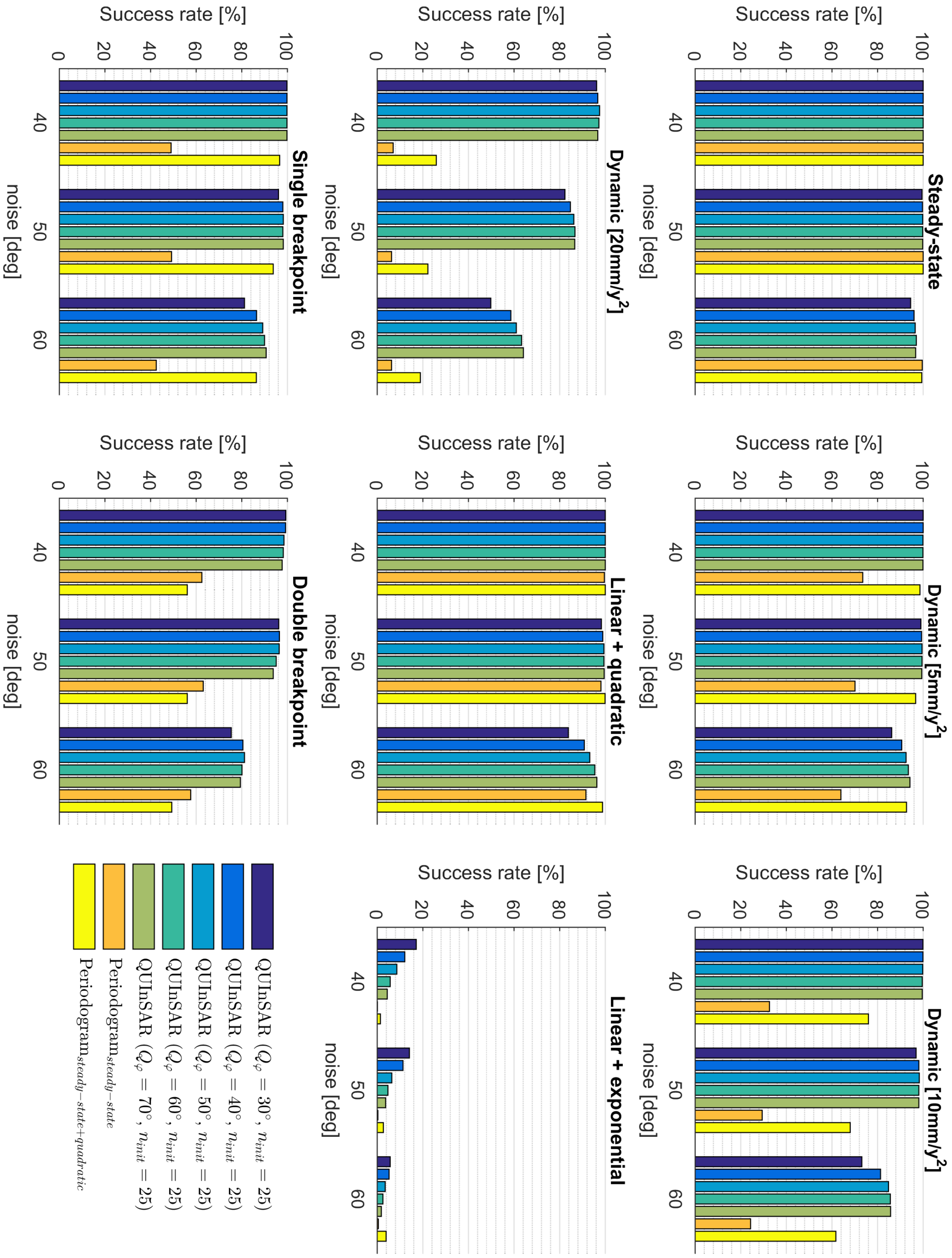


Figure A.15: Radarsat-2 success rates with varying observation phase noise factors.

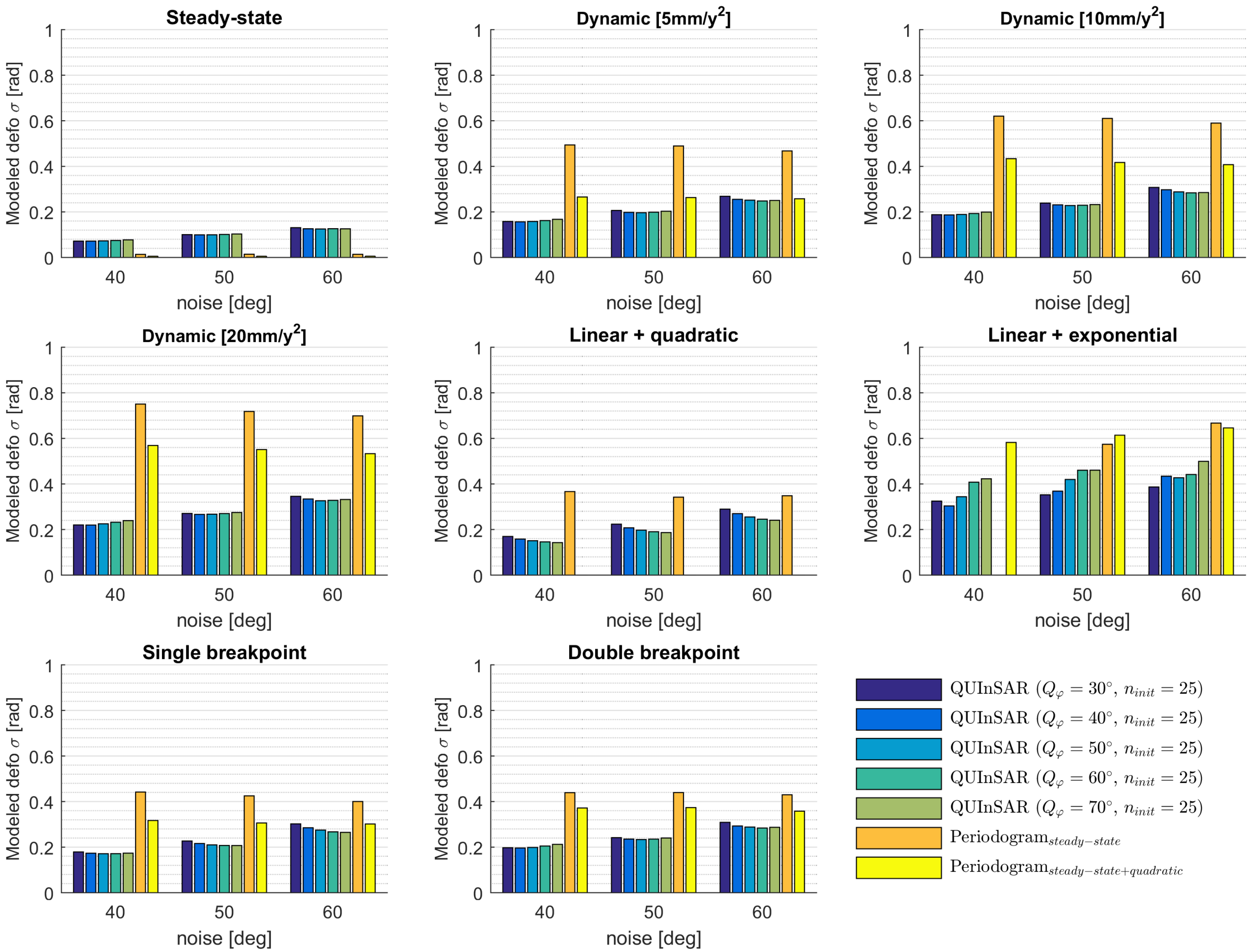


Figure A.16: Radarsat-2 estimated deformation standard deviation with varying observation phase noise factors.

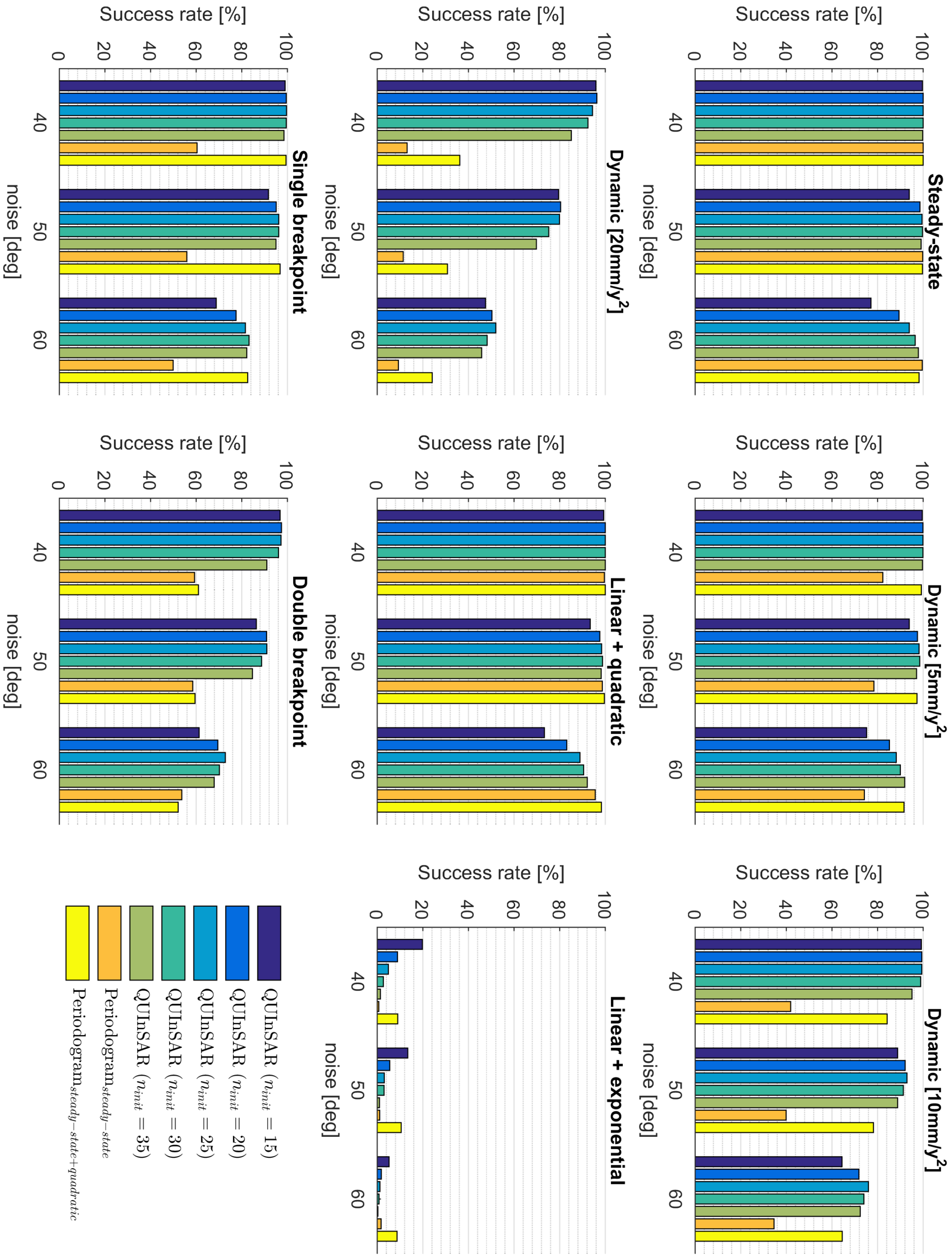


Figure A.17: ERS success rates with varying initialization length.

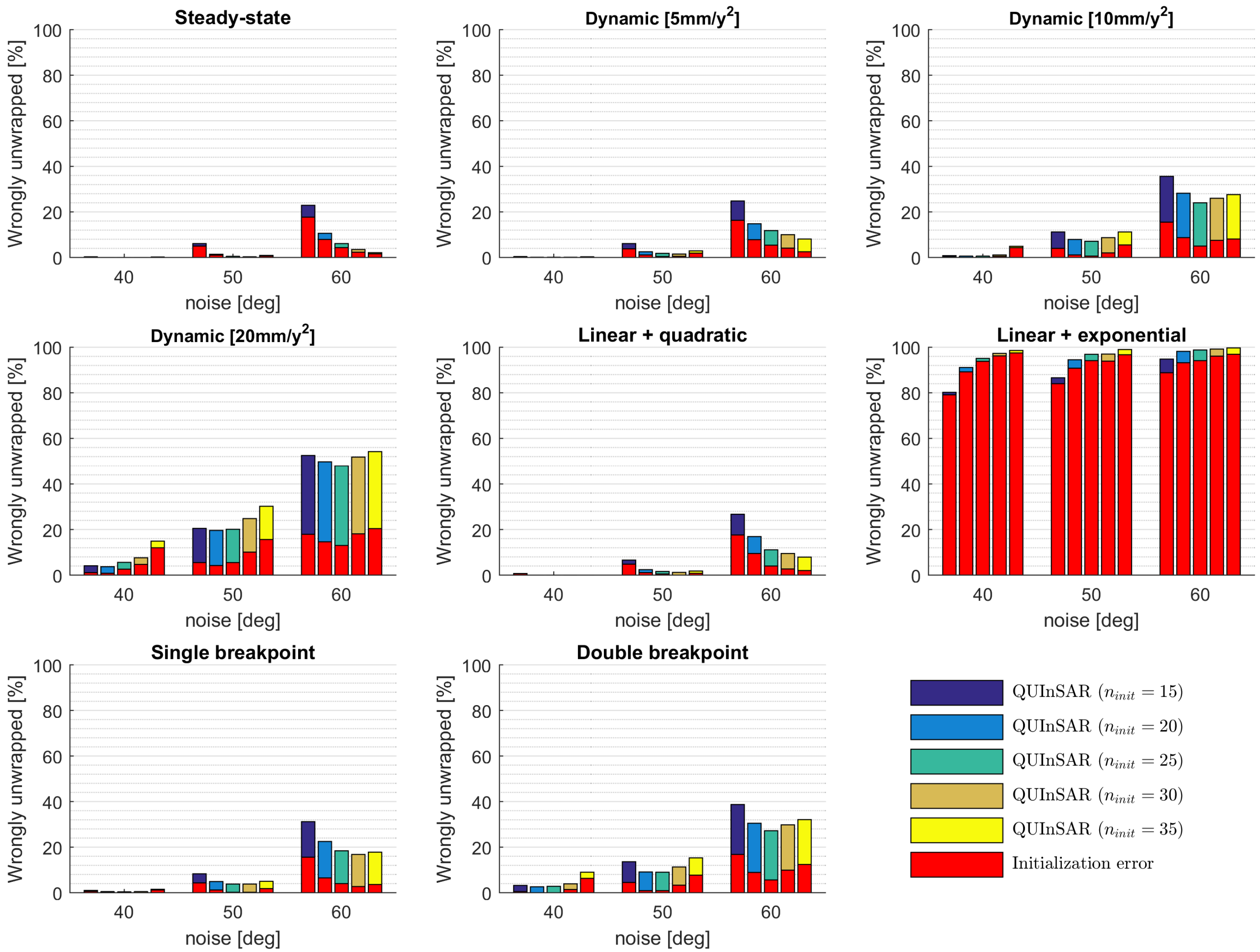


Figure A.18: ERS initialization errors with varying initialization length.

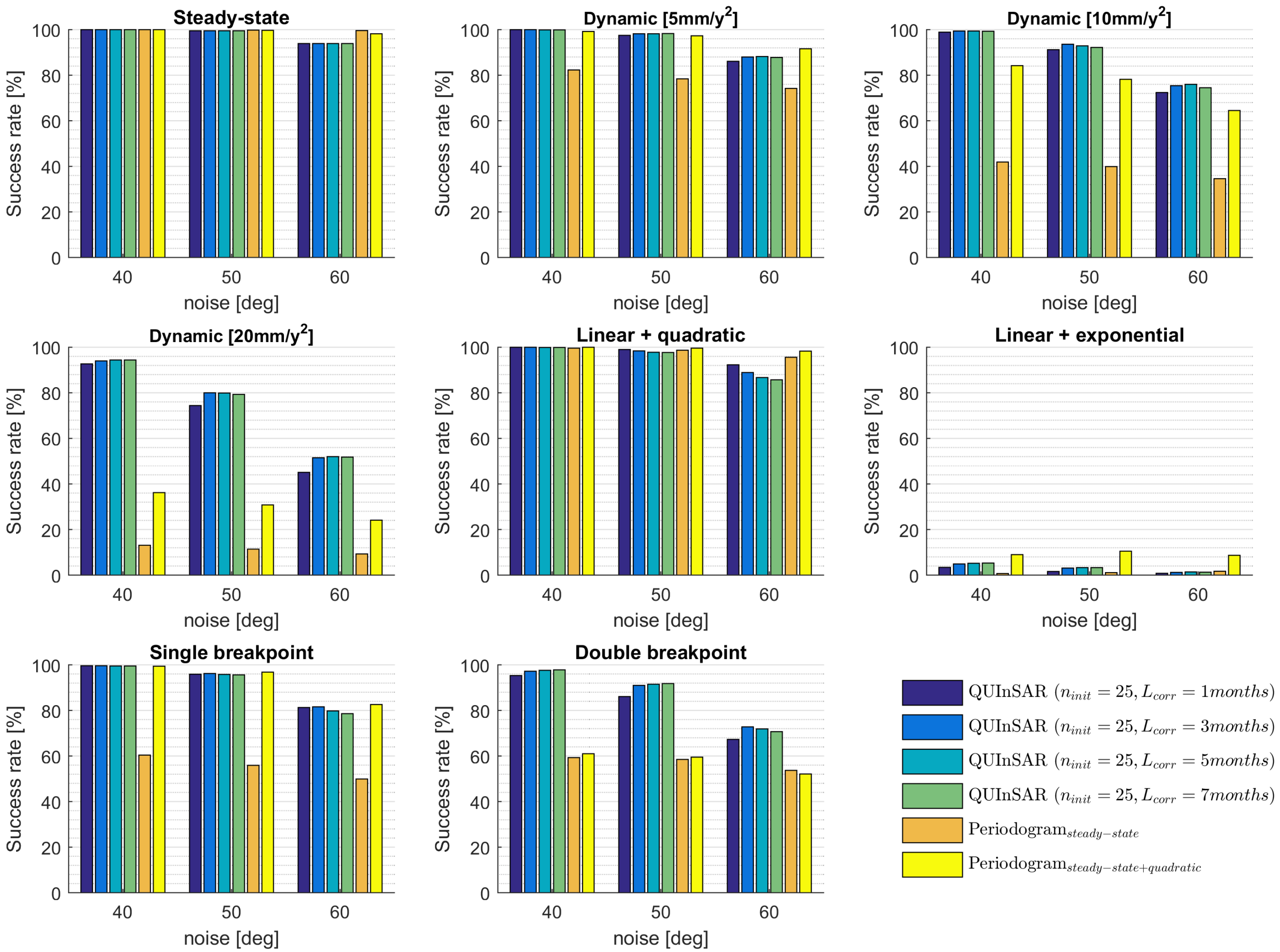


Figure A.19: ERS success rates with varying a priori acceleration correlation length.

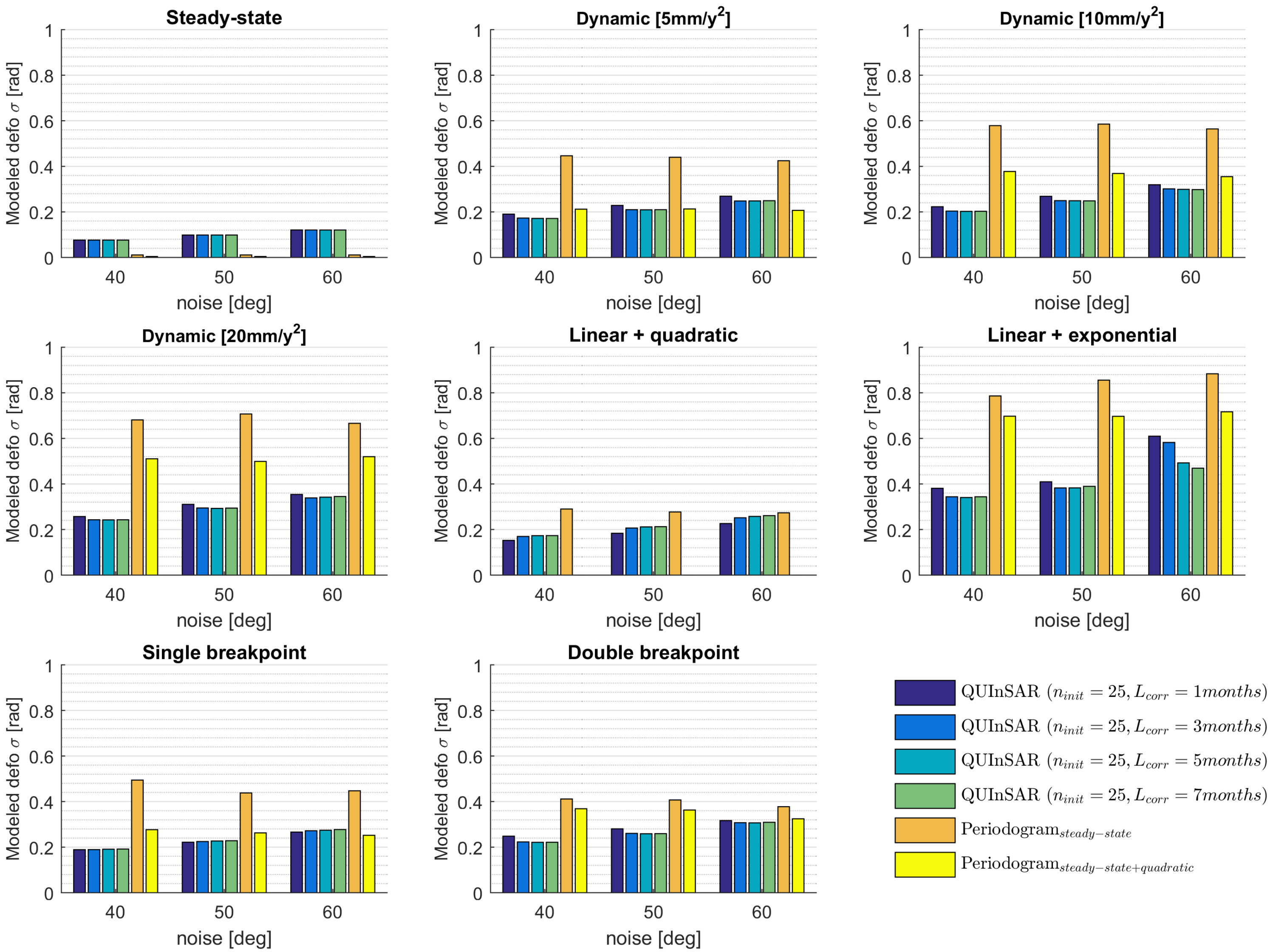


Figure A.20: ERS estimated deformation standard deviation with varying a priori acceleration correlation length.

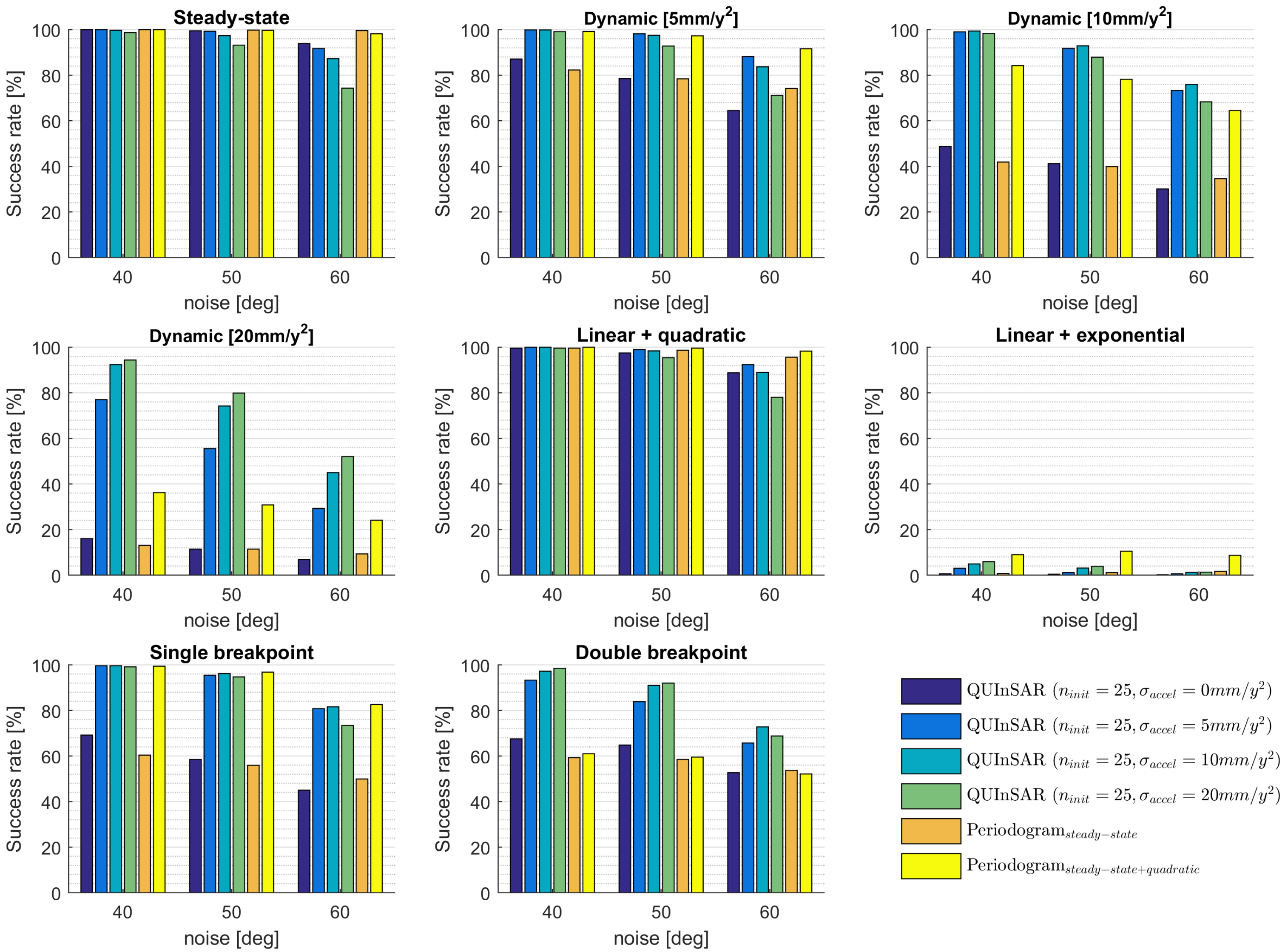


Figure A.21: ERS success rates with varying a priori acceleration variance.

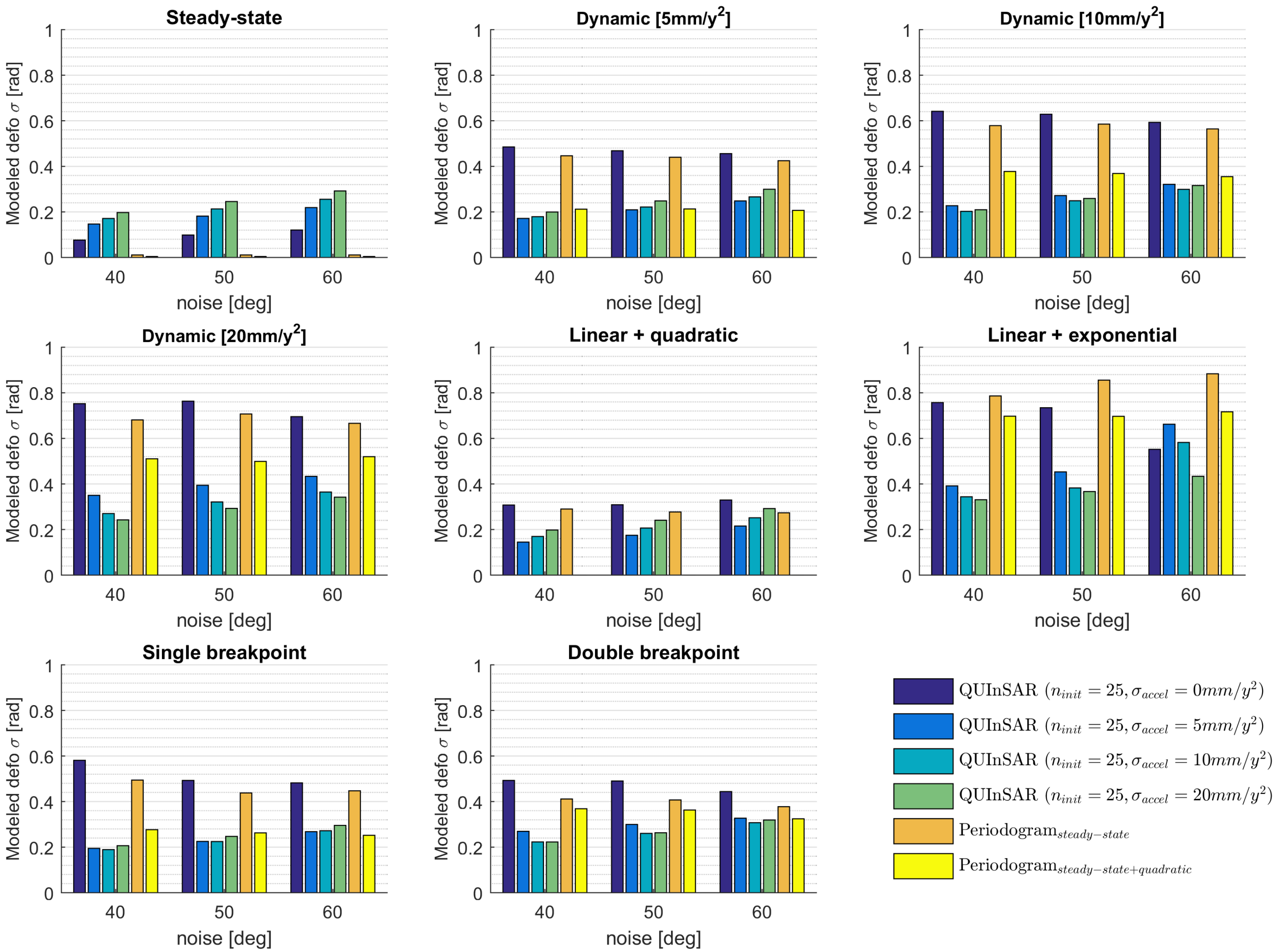


Figure A.22: ERS estimated deformation standard deviation with varying a priori acceleration variance.

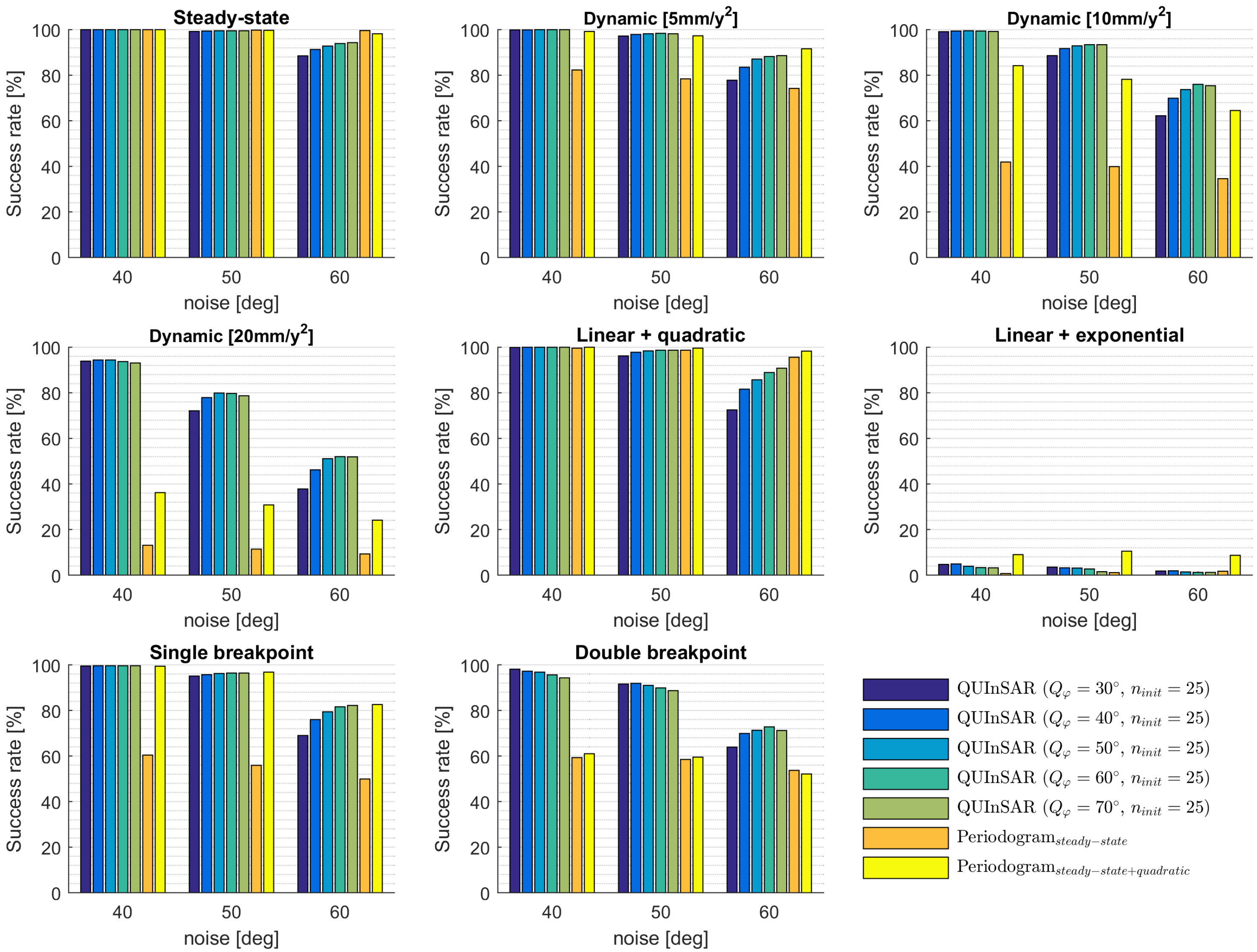


Figure A.23: ERS success rates with varying observation phase noise factors.

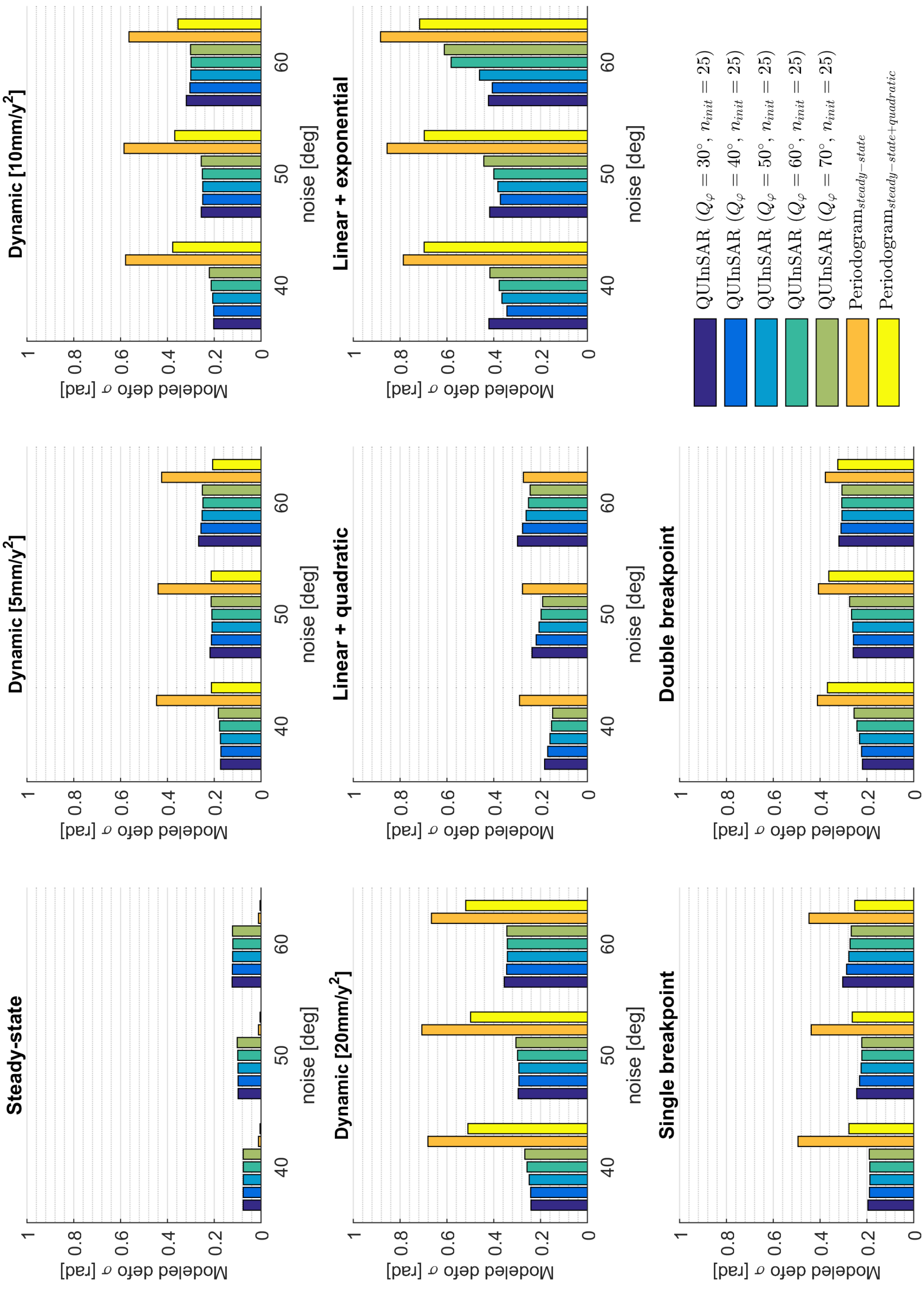


Figure A.24: ERS estimated deformation standard deviation with varying observation phase noise factors.

# Bibliography

E. P. W. Attema. The active microwave instrument on-board the ERS-1 satellite. *Proceedings of the IEEE*, 79(6):791–799, June 1991.

W. Baarda. *A testing procedure for use in geodetic networks*, volume 5 of *Publications on Geodesy*. Netherlands Geodetic Commission, Delft, 2 edition, 1968. ISBN 9789061322092.

R. Bamler and P. Hartl. Synthetic aperture radar interferometry. *Inverse Problems*, 14(98):54, 1998.

Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc, New York, 2004. ISBN 9780471465218.

G. Bishop and G. Welch. An introduction to the kalman filter. *Proc of SIGGRAPH, Course*, 8:41, 2001. URL http://old.shahed.ac.ir/references/kalman_filter_notes.pdf.

V. Cardone, W. Pierson, E. Ward, et al. Hindcasting the directional spectra of hurricane generated waves. In *Offshore Technology Conference*. Offshore Technology Conference, Houston, Texas, 5-8 May, 1975.

L. Chang. *Monitoring civil infrastructure using satellite radar interferometry*. PhD thesis, Delft, University of Technology, 2015. URL http://doi.org/10.4233/uuid:f4c6a3a2-73a8-4250-a34f-bc67d1e34516.

L. Chang and R. F. Hanssen. A Probabilistic Approach for InSAR Time-Series Postprocessing. *IEEE Transactions on Geoscience and Remote Sensing*, 54(1):421–430, 1 2016. ISSN 0196-2892. doi: 10.1109/TGRS.2015.2459037. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7208887.

E. Cosme, J. Verron, P. Brasseur, J. Blum, and D. Auroux. Smoothing Problems in a Bayesian Framework and Their Linear Gaussian Solutions. *Monthly Weather Review*, 140(2):683–695, 2012. ISSN 0027-0644. doi: 10.1175/MWR-D-10-05025.1.

C. C. Counselman and S. A. Gourevitch. Miniature Interferometer Terminals for Earth Surveying: Ambiguity And Multipath with Global Positioning System. *IEEE Transactions on Geoscience and Remote Sensing*, GE-19(4):244–252, 1981. ISSN 0196-2892. doi: 10.1109/TGRS.1981.350379. URL http://ieeexplore.ieee.org/document/4157258/.

B. Delaunay. Sur la sphere vide. *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk*, 7(793-800):2, 1934.

Delft. GRONDWATERONTTREKKING, 2017. URL https://www.delft.nl/Inwoners/Actueel/Dossiers/Grondwateronttrekking.

A. Ferretti, C. Prati, and F. Rocca. Analysis of Permanent Scatterers in SAR interferometry. In *International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No.00CH37120)*, volume 2, pages 761–763. IEEE Transactions on Geoscience and Remote Sensing, 2000a. ISBN 0-7803-6359-0. doi: 10.1109/IGARSS.2000.861695. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=861695http://ieeexplore.ieee.org/document/861695/.

A. Ferretti, C. Prati, and F. Rocca. Nonlinear Subsidence Rate Estimation Using Permanent Scatterers in Differential SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5):2202–2212, 2000b. doi: 10.1109/36.868878.

A. Ferretti, C. Prati, and F. Rocca. Permanent Scatters in SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 39(1):8–20, 2001. ISSN 01962892. doi: 10.1109/36.898661.

A. K. Gabriel, R. M. Goldstein, and H. A. Zebker. Mapping small elevation changes over large areas: Differential radar interferometry. *Journal of Geophysical Research*, 94(B7):9183, 1989. ISSN 0148-0227. doi: 10.1029/JB094iB07p09183. URL http://doi.wiley.com/10.1029/JB094iB07p09183.

C. F. Gauss. *Theoria motus corporum coelestium in sectionibus conicis solem ambientium auctore Carolo Friderico Gauss*. sumtibus Frid. Perthes et IH Besser, 1809.

A. F. Genovese. The interacting multiple model algorithm for accurate state estimation of maneuvering targets. *Johns Hopkins APL Technical Digest (Applied Physics Laboratory)*, 22(4):614–623, 2001. ISSN 02705214.

R. F. Hanssen. *RADAR INTERFEROMETRY Data Interpretation and Error Analysis*. PhD thesis, Delft, University of Technology, 2001.

R. F. Hanssen. Stochastic modeling of time series radar interferometry. In *International Geoscience and Remote Sensing Symposium, Anchorage, Alaska, 20–24 September 2004, pages cdrom*, page 4, 2004. ISBN 0-7803-8742-2. doi: 10.1109/IGARSS.2004.1369832.

L. Hopman. Beschikking wijziging op het saneringsplan. Letter on municipality website, 2016. URL https://www.denhaag.nl/nl/in-de-stad/natuur-en-milieu/water-en-bodem/sanering-in-uitvoering-anna-paulownaplein.htm.

R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35, 8 1960. ISSN 00219223. doi: 10.1115/1.3662552. URL http://fluidsengineering.asmedigitalcollection.asme.org/article.aspx?articleid=1430402.

B. Kampes and S. Usai. Doris: the Delft Object-oriented Radar Interferometric Software. In *2nd International Symposium on Operationalization of Remote Sensing, Enschede, The Netherlands, 16–20 August*, 1999. URL http://doris.tudelft.nl/Literature/kampes99.pdf.

B. M. Kampes. *Displacement Parameter Estimation using Permanent Scatterer Interferometry*. PhD thesis, Delft University of Technology, 2005. URL http://resolver.tudelft.nl/uuid:9f11f2ef-2db4-4583-a763-aee88e06ee3b.

B. M. Kampes. *Radar interferometry: Persistent Scatterer Technique*. Springer Netherlands, 1 edition, 2006. ISBN 978-1-4020-4576-9. doi: 10.1007/978-1-4020-4723-7. URL http://link.springer.com/10.1007/978-1-4020-4723-7.

B. M. Kampes and R. F. Hanssen. Ambiguity resolution for permanent scatterer interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 42(11):2446–2453, 2004. ISSN 01962892. doi: 10.1109/TGRS.2004.835222.

G. Ketelaar. *Satellite Radar Interferometry: Subsidence Monitoring Techniques*, volume 14 of *Remote Sensing and Digital Image Processing*. Springer Netherlands, Dordrecht, 2009. ISBN 978-1-4020-9427-9. doi: 10.1007/978-1-4020-9428-6. URL http://link.springer.com/10.1007/978-1-4020-9428-6_2http://link.springer.com/10.1007/978-1-4020-9428-6.

V. Ketelaar. *Monitoring surface deformation induced by hydrocarbon production using satellite radar interferometry*. PhD thesis, Delft University of Technology, Delft, 2008.

K.-R. Koch. *Parameter Estimation and Hypothesis Testing in Linear Models*. Springer-Verlag, New York, 2 edition, 1999. ISBN 978-3-662-03976-2.

O. Ku. Python code for generation of aoi figures, 2017.

A. Leetmaa and M. Ji. Operational hindcasting of the tropical pacific. *Dynamics of atmospheres and oceans*, 13(3-4):465–490, 1989. doi: 10.1016/0377-0265(89)90050-X.

A. A. Markov. *Wahrscheinlichkeitsrechnung*. Tuebner, Leipzig, 1900.

L. C. Morena, K. V. James, and J. Beck. An introduction to the RADARSAT-2 mission. *Canadian Journal of Remote Sensing*, 30(3):221–234, 2004. ISSN 07038992. doi: 10.5589/m04-004.

"Municipality of The Hague". Anna paulownaplein: bodemsanering in uitvoering, 2016 (accessed August 22, 2017). URL https://www.denhaag.nl/nl/in-de-stad/natuur-en-milieu/water-en-bodem/sanering-in-uitvoering-anna-paulownaplein.htm.

A. Pewsey, M. Neuhäuser, and G. D. Ruxton. *Circular statistics in R*. Oxford University Press, 2013. ISBN 9780199671137.

H. E. Rauch, C. T. Striebel, and F. Tung. Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, 3(8):1445–1450, 8 1965. ISSN 0001-1452. doi: 10.2514/3.3166. URL http://arc.aiaa.org/doi/10.2514/3.3166.

S. Samiei-Esfahany. *Improving Persistent Scatterer Interferometry Results for Deformation Monitoring (Case study on the Gardanne mining site)*. PhD thesis, Delft, University of Technology, 2008.

S. Samiei-Esfahany. *On exploitation of distributed scatterers in synthetic aperture radar interferometry*. PhD thesis, Delft University of Technology, 2017.

SkyGeo. Internal SkyGeo memo with settlement rod data of highway construction work, 2014.

SkyGeo. Personal communication with InSAR engineers about valid ranges of deformation parameters, 2016.

P. Teunissen. *Adjustment theory; an introduction*. VSSD, Delft, Delft, 1 edition, 2000. ISBN 90-407-1974-8.

P. J. G. Teunissen. Least-Squares Estimation of the Integer GPS Ambiguities. In *Invited Lecture for Section IV "Theory and Methodology" at the General Meeting of the International Association of Geodesy, Beijing, China, August 1993*, page 16, 1993. URL http://pages.citg.tudelft.nl/fileadmin/Faculteit/CiTG/Over_de_faculteit/Afdelingen/Afdeling_Geoscience_and_Remote_Sensing/pubs/PT_BEIJING93.PDF.

P. J. G. Teunissen. An optimality property of the integer least-squares estimator. *Journal of Geodesy*, 73(11):587–593, 1999. ISSN 09497714. doi: 10.1007/s001900050269.

P. J. G. Teunissen. *Dynamic Data Processing*. VSSD, Delft, 1 edition, 2007. ISBN 978-90-407-1976-9. URL http://www.delftacademicpress.nl/a031.php.

P. J. G. Teunissen and A. Kleusberg. *GPS for Geodesy*, volume 60 of *Lecture Notes in Earth Sciences*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1996. ISBN 978-3-540-60785-4. doi: 10.1007/BFb0117676. URL http://link.springer.com/10.1007/BFb0117676.

S. Usai. The use of man-made features for long time scale INSAR. In *In International Geoscience and Remote Sensing Symposium, Singapore, 3–8 Augustus 1997*, volume 4, pages 1542–1544, 1997. ISBN 0780338367. doi: 10.1109/IGARSS.1997.608936. URL http://doris.tudelft.nl/Literature/usai97a.pdf.

S. Usai and R. Hanssen. Long time scale INSAR by means of high coherence features. In *Third ERS Symposium - Space at the Service of our Environment, Florence, Italy, 17-21 March, 1997*, pages 225–228, 1997. ISBN 0379-6566.

M. van der Kooij. Land Subsidence Measurements at the Belridge Oil Fields from ERS Insar Data. In *Third ERS Symposium on Space at the Service of our Environment, Florence, Italy, 14-21 March*, 1997.

F. J. van Leijen. *Persistent Scatterer Interferometry based on geodetic estimation theory*. PhD thesis, Delft, University of Technology, 2014. URL http://doi.org/10.4233/uuid:5dba48d7-ee26-4449-b674-caa8df93e71e.

F. J. Van Leijen and R. F. Hanssen. Persistent scatterer density improvement using adaptive deformation models. In *International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–27 July 2007*, number 2, pages 2102–2105, 2007. ISBN 1424412129. doi: 10.1109/IGARSS.2007.4423248.

F. J. Van Leijen and R. F. Hanssen. Ground water management and its consequences in Delft, The Netherlands as observed by persistent scatterer interferometry. In *Fifth International Workshop on ERS/Envisat SAR Interferometry, 'FRINGE07', Frascati, Italy, 26 Nov-30 Nov 2007*, page 6, 2008.

F. J. van Leijen, R. F. Hanssen, P. S. Marinkovic, and B. M. Kampes. SPATIO-TEMPORAL PHASE UNWRAPPING USING INTEGER LEAST-SQUARES. In *Fourth International Workshop on ERS/Envisat SAR Interferometry, 'FRINGE05', Frascati, Italy, 28 Nov-2 Dec 2005*, page 6, 2006.

R. Werninghaus, W. Balzer, S. Buckreuss, J. Mittermayer, and P. Mühlbauer. The TerraSAR-X Mission. In *Proceedings of SPIE*, pages 9–16, 2003.

T. J. Wright, B. Parsons, and E. Fielding. Measurement of interseismic strain accumulation across the {N}orth {A}natolian {F}ault by satellite radar interferometry. *Geophys. Res. Lett.*, 28(10):2117–2120, 2001. doi: 10.1029/2000GL012850.

H. A. Zebker and R. M. Goldstein. Topographic mapping from interferometric synthetic aperture radar observations. *Journal of Geophysical Research*, 91(B5):4993, 1986. ISSN 0148-0227. doi: 10.1029/JB091iB05p04993. URL http://adsabs.harvard.edu/abs/1986JGR....91.4993Z%5Cnhttp://doi.wiley.com/10.1029/JB091iB05p04993.

H. A. Zebker, P. A. Rosen, and S. Hensley. Atmospheric effects in interferometric synthetic aperture radar surface deformation and topographic maps. *Journal of Geophysical Research*, 102(10):7547–7563, 1997. ISSN 0148-0227. doi: 10.1029/96JB03804.