TIME-SERIES INSAR: AN INTEGER LEAST-SQUARES APPROACH FOR DISTRIBUTED SCATTERERS

Sami Samiei-Esfahany, Ramon F. Hanssen

Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands.
Email: s.samieiesfahany@tudelft.nl

ABSTRACT

The objective of this research is to extend the geodetic mathematical model which was developed for persistent scatterers to a model which can exploit distributed scatterers (DS). The main focus is on the integer least-squares framework, and the main challenge is to include the decorrelation effect in the mathematical model. In order to adapt the integer least-squares mathematical model for DS we altered the model from a single master to a multi-master configuration and introduced the decorrelation effect stochastically. This effect is described in our model by a full covariance matrix. We propose to derive this covariance matrix by numerical integration of the (joint) probability distribution function (PDF) of interferometric phases. This PDF is a function of coherence values and can be directly computed from radar data. We show that the use of this model can improve the performance of temporal phase unwrapping of distributed scatterers.

Key words: Distributed Scatterer ; Time-Series InSAR; Integer Least-Squares; Decorrelation.

1. INTRODUCTION

During the last decade, different time-series InSAR methodologies have been developed capable of detecting and monitoring various ground deformation mechanisms. A first-generation time-series InSAR technique, called Persistent Scatterer Interferometry (PSI) [4, 5, 11, 9], focuses on targets with persistent phase behavior in time. These targets, named persistent scatterers (PS), are coherent in the entire data stack and are minimally affected by temporal and geometric decorrelation. However, the availability of persistent scatterers is not guaranteed. For instance, in rural and agricultural areas, a low density of PS causes low spatial sampling of deformation measurements. This limitation raises the question: is it possible to extract useful information from resolution cells with distributed scattering mechanism? During the last few years, due to processing more datasets and especially due to availability of data with shorter satellite revisit times, it has become evident that there are distributed scatterer (DS) resolution cells which contain coherent information not in the whole data stack but only in interferograms with short temporal baseline. That is, there are some areas which preserve coherence during some time interval, and then gradually lose their coherence. Consequently, the second generation of time-series InSAR methodologies based on small baselines processing [2, 13, 8] and recently the SqueeSAR [3] method were developed to extract information also from distributed scatterers. The goal of this article is to extend the geodetic mathematical model, which was developed for persistent scatterers, to a model which can exploit distributed scatterers. The main focus is on the integer least-squares (ILS) framework. This geodetic approach, originally designed for GPS applications [15, 17, 18], is a PSI algorithm developed for the estimation of parameters of interest (e.g., height and deformation) and for the estimation of phase cycle ambiguities [12, 21, 19]. The main advantages of this approach can be summarized as follows: 1) it considers the correlation between interferograms in parameter estimation, 2) it treats the unknown ambiguities as stochastic instead of deterministic, and 3) it allows the formal error propagation from observations (i.e., interferometric phases) to unknown parameters. The ILS mathematical model consists of a functional and a stochastic model. The functional model describes the relation between the observations and the unknown parameters, whereas the stochastic model formulates the stochastic properties of the observations. The main challenge is to include the decorrelation effect in these models. In order to adapt the mathematical model for DS, we apply following modifications: 1) changing the functional model from single master to multi-master configuration, 2) introducing the decorrelation effect in the stochastic model. The following section describes the ILS mathematical model in further detail.

2. ILS MATHEMATICAL MODEL FOR PERSISTENT SCATTERERS

In the ILS PSI processing chain, parameter estimation is based on a mathematical model whose general form is described by the Gauss-Markov model including functional

Proc. 'Fringe 2011 Workshop', Frascati, Italy,
19–23 September 2011 (ESA SP-697, January 2012)

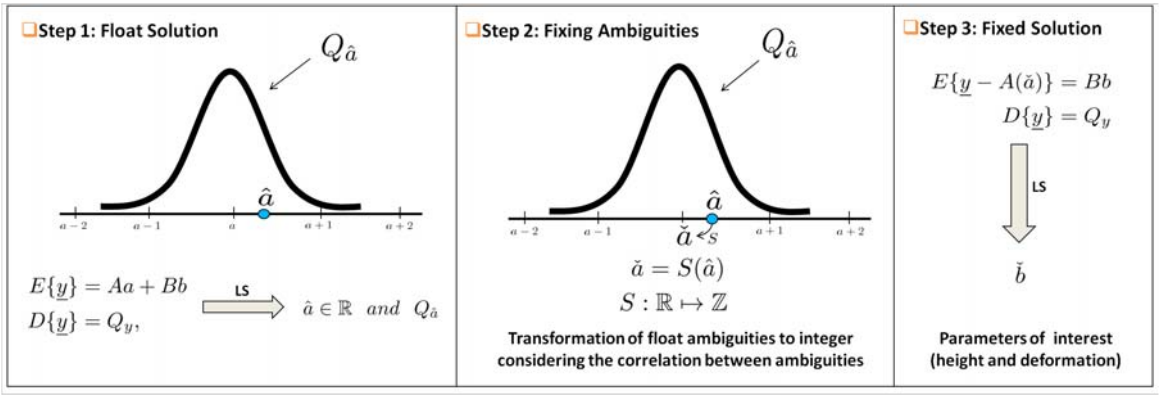

Figure 1. Schematic description of ILS algorithm in three steps.

and stochastic models:

$$\begin{aligned} E\{\underline{y}\} &= Aa + Bb \\ D\{\underline{y}\} &= Q_y, \end{aligned} \tag{1}$$

where the first part shows the functional model and the second part the stochastic model. $E\{.\}$ is the expectation operator, $D\{.\}$ the dispersion, y the vector of phase observations, a the vector of integer valued ambiguities, and b represents the unknown real valued parameters. A and B are design matrices which describe the functional relation between the observation and the integer- and real-valued unknown parameter vectors respectively. Q_y is the covariance matrix which describes the stochastic properties of the y vector. This mathematical model is based on the double-difference (DD) phase observations. That is, a spatial network among pixels has already been constructed and the estimation of the unknown parameters is performed per arc of the network.

The functional model for a single-master stack, where the master is indicated by a zero, and there are N slave acquisitions, is formulated as:

$$E\{\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix}\} = \begin{bmatrix} A_1 & B_1 \\ A2 & B_2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \tag{2}$$

$$\begin{aligned} E\{\underbrace{\begin{bmatrix} \psi^{01} \\ \vdots \\ \psi^{0N} \\ \Delta H^* \\ D^* \end{bmatrix}}_{y}\} = \underbrace{\begin{bmatrix} -2\pi & & \\ & \ddots & \\ & & -2\pi \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a^{01} \\ \vdots \\ a^{0N} \end{bmatrix}}_{a} \\ + \underbrace{\begin{bmatrix} \beta^{01} & \frac{-4\pi}{\lambda} B_T^{01} \\ \vdots & \vdots \\ \beta^{0N} & \frac{-4\pi}{B}{}^{0N}{}_T \\ 1 & \\ & 1 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} \Delta H \\ D \end{bmatrix}}_{b}, \end{aligned} \tag{3}$$

where ψ are the DD phase observations, ΔH the residual height, D the deformation parameter, λ the radar wavelength, and B_T^{0n} temporal baseline of the nth interferogram. β^{0n} is the height-to-phase conversion factor for the nth interferogram.

In the functional model, $(.)^*$ denotes the pseudo-observations needed to solve for the rank deficiency of the system [12]. Rank deficiency is caused by the fact that for each observed phase, an ambiguity needs to be estimated together with the parameters of interest. As a result, the number of unknowns exceeds the number of observations, and the system is under-determined. With the introduction of pseudo-observations, the mathematical model is regularized. Without *a priori* knowledge of the topography and deformation in the area, the pseudo-observations are set to zero.

The stochastic model contains the statistical characteristics of the observation vector y and is presented in the form of the covariance matrix Q_y. As the observation vector y comprises both DD phase observations and pseudo-observations, Q_y also has two parts:

$$D\{\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix}\} = \begin{bmatrix} Q_{y_1} & 0 \\ 0 & Q_{y_2} \end{bmatrix}, \tag{4}$$

where Q_{y_1} is the covariance matrix of the DD phase observations which mainly describes the stochastic characteristics of the atmospheric signal. y_2 and Q_{y_2} are the vector of pseudo-observations and its covariance matrix, respectively. Different atmospheric covariance functions can be used to fill the components of Q_{y_1} using the length of the arc. Values of Q_{y_2} can be assumed using *a priori* knowledge about the topography and the deformation mechanism, see [6]. The components of Q_{y_2} bound the solution space of the corresponding unknown parameters.

2.1. ILS algorithm

Using the mathematical model described earlier in this section, the unknown parameters can be estimated by least-squares estimation. The procedure is divided into three steps [15]:

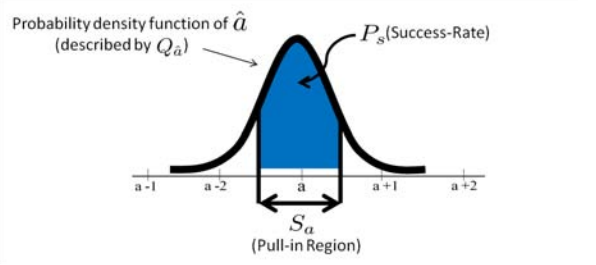

Figure 2. Schematic definition of success rate.

1. First, the float solution is estimated by disregarding the integerness of the ambiguities. Hence, standard least-squares adjustment is performed to estimate $\hat{a}$, $\hat{b}$ and the accompanying variance matrix

$$\begin{bmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}} \end{bmatrix}. \tag{5}$$

2. After obtaining the float solution, integer ambiguities $\breve{a}$ are estimated by mapping the float ambiguity vector $\hat{a}$ to the corresponding integer (fixed) ambiguity vector using a mapping operator $S : \mathbb{R} \mapsto \mathbb{Z}$, so:

$$\breve{a} = S(\hat{a}). \tag{6}$$

   In this mapping, the float ambiguity vector is mapped to the *nearest* integer ambiguity vector. However, the *nearest* must be measured in the metric of the covariance matrix of the float ambiguities taking into account their correlation and precision. There are different integer estimators to obtain fixed ambiguities. In [17] three admissible integer estimators were introduced: *integer rounding*, *integer bootstrapping*, and *integer least-squares*. The integer rounding estimator rounds the entries of the float ambiguity solution to their nearest integer values. The disadvantage of the integer rounding method is that it does not take the correlation between float ambiguities into account. In contrast, the integer bootstrapping estimator considers *some* of the ambiguity correlation [16, 18]. The integer least-squares estimator is the only estimator which accounts for all the correlation among float ambiguities [15].

   The integer least-squares estimator obtains the maximum likelihood solution for normally distributed data. The disadvantage of this technique is that its computational time is larger in comparison with the bootstrapping method. An extended bootstrap estimator in order to reduce the computational time was introduced in [11].

3. Once the fixed ambiguity $\breve{a}$ is obtained by one of the admissible estimators, the float solution of the parameters of interest $\hat{b}$ is updated using the fixed ambiguities. Fig. 1 shows the three steps of ILS in a schematic way.

2.2. Success Rate

During mapping of real-valued float ambiguities to fixed integer ambiguities (i.e., the second step of ILS), different real-valued ambiguity vectors are mapped to the same integer vector. The subset S_a, which contains all real-valued float ambiguities that will be mapped to the same integer vector a, is called the pull-in region of a. The shape of the pull-in region is dependent on the applied integer estimator. The success rate P_s is the probability that the float ambiguities are fixed to the correct integers. It is defined as:

$$P_s = \int_{S_a} pdf_{\hat{a}}(x)dx, \tag{7}$$

where $pdf_{\hat{a}}$ is the probability density function of float ambiguities which are described by $Q_{\hat{a}}$ in the case of normally distributed observations. The closed-form evaluation of success rate for the bootstrap method was given in [17, 11]. Fig. 2 presents the schematic definition of success rate.

Note that the narrower the PDF of float ambiguities, the higher the success rate. In other words, less noisy phase observations will result in a lower variance of float ambiguities, and consequently in a higher success rate. For reliable estimation of the parameters of interest, a high success rate (i.e., close to 1) is required.

3. STOCHASTIC MODEL FOR DS

In the stochastic model (4), Q_{y_1} describes the stochastic properties of DD phase observations. For PS, this covariance matrix contains mainly the effects of atmospheric signal and thermal noise. But for DS, this model should be extended to include the decorrelation effect in Q_{y_1}. We define Q_{y_1} as the sum of two components:

$$Q_{y_1} = Q_{sc} + Q_{coh}, \tag{8}$$

where Q_{sc} is the covariance matrix which describes the stochastic behavior of spatially correlated components (i.e., mainly atmospheric signal), and Q_{coh} describes the spatially uncorrelated components such as thermal noise and decorrelation. In practice, the magnitude of the spatial coherence γ is used as a measure of spatially uncorrelated components, assuming spatial ergodicity. The main objective here is to derive the variance and covariance values using the coherence values. For unbiased estimation of spatial coherence, improvements were presented in [3] and [22]. Knowing the absolute coherence of a particular pixel, the variance of single difference phase ϕ^{mn} (i.e., interferometric phase between image m and image n) can be derived as:

$$\sigma^2_{\phi^{mn}} = \int_{-\pi}^{\pi} \left[\phi^{mn} - E\left\{\phi^{mn}\right\}\right]^2 pdf_{\phi^{mn}}(\gamma^{mn})d\phi^{mn} \tag{9}$$

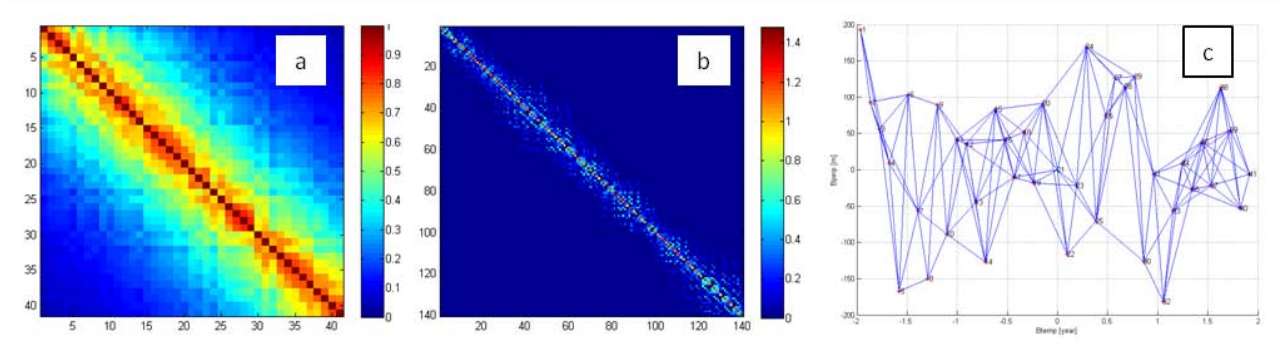

Figure 3. An example of a pixel coherence matrix (a) and a covariance matrix (b) for a particular baseline configuration with 140 interferograms made from 41 acquisitions (c).

where $pdf_{\phi^{mn}}(\gamma^{mn})$ is the probability density function of ϕ^{mn}, which is a function of absolute coherence γ^{mn}. A closed form expression for variance of single-look observations was derived in [1]. For multi-looked phases, the variance can be derived by numerical integration of (9) using the closed form expression of $pdf_{\phi^{mn}}(\gamma^{mn})$ presented in [10, 20]. By computing the variance of phase observations we actually derive the diagonal elements of covariance matrix Q_{coh_k} of single difference observations for pixel k.

We still need to compute the off-diagonal elements or covariances between interferometric phases. The main source of this covariance is the correlation between decorrelation noise of interferograms. The covariance between two interferometric phases ϕ^{om} and ϕ^{on} can be derived as:

$$Cov(\phi^{om}, \phi^{on}) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} [\phi^{om} - E\{\phi^{om}\}][\phi^{on} - E\{\phi^{on}\}] \dots \quad (10)$$
$$\dots pdf_{\phi^{om},\phi^{on}}(\gamma^{om}, \gamma^{on}, \gamma^{mn}) d\phi^{om} d\phi^{on},$$

where $pdf_{\phi^{om},\phi^{on}}(\gamma^{om}, \gamma^{on}, \gamma^{mn})$ is the joint probability distribution function of ϕ^{om} and ϕ^{on}, which is a function of three coherence values γ^{om}, γ^{on}, and γ^{mn}. This joint PDF was derived (only for single-look observation) in [14]. The integral (10) can be solved numerically in order to derive the covariance values. By computing the variance and covariance values using (9) and (10) we can construct the full covariance matrix for each pixel and for different baseline configurations. In other words, for each pixel, we can derive the full covariance matrix of an interferometric stack from its coherence matrix. Fig. 3 shows an example of coherence and covariance matrices for a particular baseline configuration.

The covariance matrix Q_{coh} now can be computed from the covariance matrices of two pixels (Q_{coh1} and Q_{coh2}) using the covariance propagation law:

$$Q_{coh} = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} Q_{coh1} & \\ & Q_{coh2} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \quad (11)$$

After computation of Q_{coh} and having Q_{sc}, the full covariance matrix of DD observations can be constructed using (8).

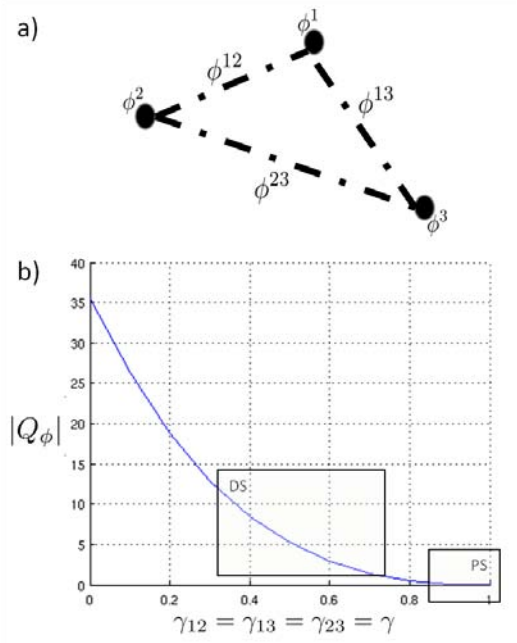

Figure 4. a) A simple set of three interferograms ϕ^{12}, ϕ^{13}, and ϕ^{23}. b) determinant of the covariance matrix $|Q_\phi|$ as a function of interferograms coherence γ.

4. FUNCTIONAL MODEL FOR DS

The functional model for PS (i.e., (3)) is written for single master baseline configuration. For PS, which by definition have high coherence or low decorrelation noise, changing the baseline configuration or using a more redundant configuration (e.g., set of small-baselines interferograms) does not improve the unwrapping or the success rate. In other words, for a PS in a stack of n acquisitions, we can only make a set of $n-1$ independent interferograms which contains all the stack information. Any other interferogram is simply a linear combination of

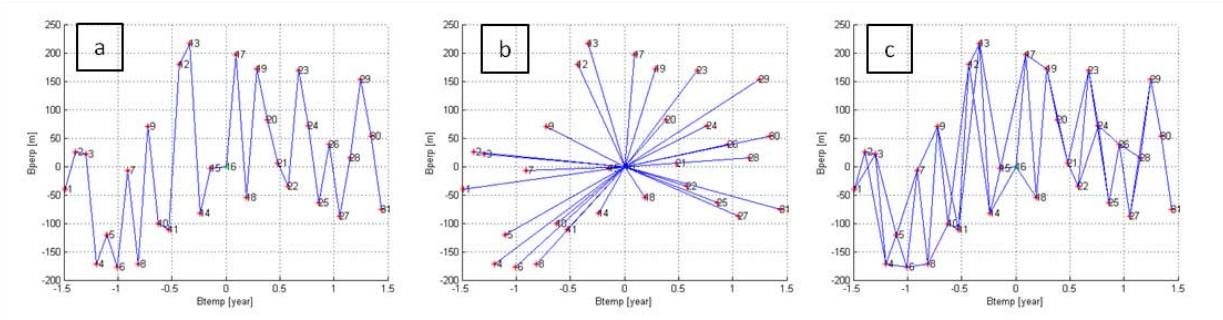

Figure 5. Baseline configuration used in the simulation: a) small temporal baselines, b) single master, and c)redundant small baselines.

other interferograms of that set. Therefore, adding other interferograms to the baseline configuration will result in a system of dependent equations which theoretically results in singularity in the covariance matrix of observations. However the question is whether this argument also holds for DS or not.

This question can be reformulated as whether there is any added value information in linearly dependent interferograms. In the case of totally dependent observations the covariance matrix of observations will be singular and hence has a zero determinant (i.e., $|Q_\phi| = 0$). We test this on a very simple case considering only three interferograms (Fig. 4a). We assume the same coherence γ for all the three interferograms, and we gradually increase γ from 0 to 1 and derive Q_ϕ and its determinant using the approach presented in Section 3. As seen in Fig. 4b, for high values of coherence (i.e., PS-like pixels), $|Q_\phi|$ becomes zero. This confirms our previous argument that linearly combined interferograms contain redundant information for PS. However, we can see that for medium coherence values (i.e., DS-like pixels), there is non-redundant or added value information in the observations. These results show that although adding a dependent interferogram does not directly bring any information about our signal of interest (i.e., height and deformation) into the model, it results in a better noise description in the stochastic model. So we can conclude that for DS, using baseline configurations other than single-master stack and adding more interferograms can improve the unwrapping performance, resulting in a higher success rate. Note that the high coherence combinations which cause singularity in the covariance matrix should be avoided.

5. SUCCESS RATE FOR DS

In order to test whether a more redundant network of interferograms can indeed increase the success rate for distributed scatterers, we compute the success rate (7) for a sample case. Note that for success rate computation, we do not need real observations and only require the covariance matrix of observations. We computed success rate for the sample case of 31 radar acquisitions with a revisit time of 35 days for three different baseline configurations: single master, small temporal baselines, and redundant small baselines (Fig. 5). Perpendicular baselines were simulated with normal distribution, with a standard deviation of $200m$. We assume an arc between a PS and DS with an arc-length of $200m$. To derive Q_{sc}, we only consider the contribution from atmosphere. We compute the components of Q_{sc} using the exponential covariance function for atmospheric signal with a standard deviation of $5mm$ and a range of $3km$. To derive the Q_{coh} for the DS, all the decorrelation sources (i.e., thermal noise, geometrical decorrelation, co-registration error) were modeled as in [7] and [1]. The DS is assumed to slowly (exponentially) decorrelate in time and lose its coherence completely after the *decorrelation range*. We computed the success rate for different decorrelation ranges. In order to see the effect of multi-looking, we also derive the success rate for the multi-looked case, with a multi-looking factor of 10. Note that in the multi-look case we use the joint PDF of single-looked phases as an approximation of joint PDF of multi-looked phases. The results are presented in Fig. 6.

We can see that the success rate for DS, unlike PS, is dependent on baseline configuration. In the presence of decorrelation noise, the choice of baseline configuration (using small baselines or a more redundant network of interferograms) can significantly improve the success rate. Multilooking on distributed scatterers can efficiently remove the decorrelation noise resulting in significant success rate improvement.

6. SUMMARY AND CONCLUSIONS

We have extended the geodetic theoretical framework of PS to a model capable of exploiting DS. We have derived the full covariance matrix for an interferogram stack from the (joint) probability distribution function of interferometric phases and the stack coherence matrix. We have demonstrated that there is non-redundant information in linearly-dependent interferograms. The correlation between decorrelation noise of connected interferograms is value-added information, which can be used for reducing the decorrelation effect and con-

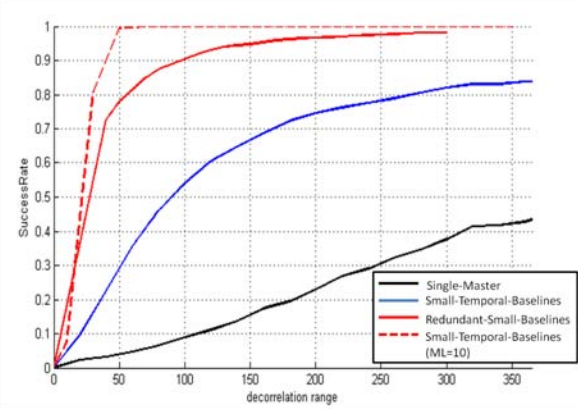

Figure 6. success rates for different baseline configurations and the effect of multi-looking (ML denotes the multi-looking factor).

sequently increasing the success rate leading to more reliable unwrapping for distributed scatterers. The success rate for DS, unlike PS, is dependent on baseline configuration. In the presence of decorrelation noise, using small baselines or a more redundant network of interferograms can significantly improve the success rate. It is important to note that the interferometric combinations which cause singularity in the covariance matrix should be avoided. Multilooking on distributed scatterers can result in significant improvement in success rate. Despite the promising results of this analytical study, future tests on simulated and real data have to show the reliability and the performance of the method on real applications.

Acknowledgment
The authors would like to thank Dr. G. Ketelaar for constructive and valuable discussions, and P. Mahapatra for her helpful review which improved the readability of this paper.

REFERENCES

[1] Richard Bamler and Philipp Hartl. Synthetic aperture radar interferometry. *Inverse Problems*, 14:R1–R54, 1998.

[2] Paolo Berardino, Gianfranco Fornaro, Riccardo Lanari, and Eugenio Sansosti. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11):2375–2383, 2002.

[3] A. Ferretti, A. Fumagalli, F. Novali, C. Prati, F. Rocca, and A. Rucci. A new algorithm for processing interferometric data-stacks: Squeesar. *IEEE Transactions on Geoscience and Remote Sensing*, 49(9):3460 –3470, 2011.

[4] Alessandro Ferretti, Claudio Prati, and Fabio Rocca. Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5):2202–2212, September 2000.

[5] Alessandro Ferretti, Claudio Prati, and Fabio Rocca. Permanent scatterers in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 39(1):8–20, January 2001.

[6] Ramon Hanssen. Stochastic modeling of time series radar interferometry. In *International Geoscience and Remote Sensing Symposium, Anchorage, Alaska, 20–24 September 2004*, pages cdrom, 4 pages, 2004.

[7] Ramon F Hanssen. *Radar Interferometry: Data Interpretation and Error Analysis*. Kluwer Academic Publishers, Dordrecht, 2001.

[8] Andrew Hooper. A multi-temporal InSAR method incorporating both persistent scatterer and small baseline approaches. *Geophysical Research Letters*, 35:L16302, 2008.

[9] Andrew Hooper, Howard Zebker, Paul Segall, and Bert Kampes. A new method for measuring deformation on volcanoes and other non-urban areas using InSAR persistent scatterers. *Geophysical Research Letters*, 31:L23611, doi:10.1029/2004GL021737, December 2004.

[10] Dieter Just and Richard Bamler. Phase statistics of interferograms with applications to synthetic aperture radar. *Applied Optics*, 33(20):4361–4368, 1994.

[11] Bert Kampes. *Displacement Parameter Estimation using Permanent Scatterer Interferometry*. PhD thesis, Delft University of Technology, Delft, the Netherlands, September 2005.

[12] Bert Kampes and Ramon F Hanssen. Ambiguity resolution for permanent scatterer interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 42(11):2446–2453, November 2004.

[13] Ricardo Lanari, Oscar Mora, Michele Manunta, Jordi J Mallorquí, Paolo Berardino, and Uegenio Sansosti. A small-baseline approach for investigating deformations on full-resolution differential SAR interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, 42(7):1377–1386, 2004.

[14] Mario Lucido, Federica Meglio, Vito Pascazio Pascazio, and Gilda Schirinzi. Closed-form evaluation of the second-order statistical distribution of the interferometric phases in dual-baseline sar systems. *IEEE Transactions on Geoscience and Remote Sensing*, 58(3):1698–1707, 2010.

[15] P J G Teunissen. Least-squares estimation of the integer GPS ambiguities. In *Invited Lecture, Section IV Theory and Methodology, IAG General Meeting, Beijing, China, august 1993*, 1993. Also in: Delft Geodetic Computing Centre, LGR Series, No. 6, 1994.

[16] Peter Teunissen. Success probability of integer GPS ambiguity rounding and bootstrapping. *Journal of Geodesy*, 72:606–612, 1998.

[17] Peter Teunissen. An optimality property of the integer least-squares estimator. *Journal of Geodesy*, 73:587–593, 1999.

[18] Peter Teunissen. GNSS ambiguity bootstrapping: Theory and application. In *International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation, Banff, Canada, 5–8 June 2001*, 2001.

[19] Peter Teunissen. On InSAR ambiguity resolution for deformation monitoring. *Artificial Satellites*, 41(1):17–22, 2006.

[20] R J A Tough, D Blacknell, and S Quegan. A statistical description of polarimetric and interferometric synthetic aperture radar. *Proceeedings of the Royal Society London A*, 449:567–589, 1995.

[21] Freek J van Leijen and Ramon F Hanssen. Persistent scatterer interferometry using adaptive deformation models. In *ESA ENVISAT Symposium, Montreux, Switzerland, 23–27 April 2007*, page pp, 2007.

[22] Yuanyuan Wang, Xiao Xiang Zhu, and Richard Bamler. Optimal estimation of distributed scatterer phase history parameters from meter-resolution sar data. In *International Geoscience and Remote Sensing Symposium, Vancouver, Canada, 24–29 July 2011*, pages 3468 –3471, 2011.