A New Method for Temporal Phase Unwrapping of Persistent Scatterers InSAR Time Series

Miguel Caro Cuenca, Andrew J. Hooper, and Ramon F. Hanssen, *Member, IEEE*

***Abstract*—The analysis of radar time series with persistent scatterer techniques usually relies on temporal unwrapping, because phase behavior can be often described by simple models. However, one of the major limitations of temporal algorithms is that they do not take advantage of spatially correlated information. Here, we focus on two types of information that can be spatially estimated, namely, observation precision and the probability density function of the model parameters. We introduce them in phase unwrapping using Bayesian theory. We test the proposed method using simulated data. We also apply them to a small area in the southern Netherlands and compare with conventional temporal unwrapping methods.**

***Index Terms*—Persistent scatterer InSAR (PSI), phase unwrapping.**

I. Introduction

SYNTHETIC aperture radar interferometry (InSAR) techniques use the phase differences of two radar images acquired over the same area to estimate heights and surface displacements [1]–[3]. However, conventional estimation methods such as least squares (see for example [4]) cannot be directly applied to InSAR measurements, because the observations—phase differences—are wrapped. In this context, a wrapped phase is the 2π-modulus of the absolute (or unwrapped) phase Φ. The absolute value Φ is estimated from the integer number of cycles a such that

$$\Phi = \phi + 2\pi a, \text{ with } \phi \in [-\pi, \pi) \text{ and } a \in \mathbb{Z}. \tag{1}$$

This process is known as phase unwrapping. Although the wrap operation $\phi = \mathbb{W}\{\Phi\}$ is straightforward, the inverse, i.e., estimating the number of cycles a, cannot be solved unless information is added to the problem, e.g., assuming additional conditions. In conventional InSAR, the interferograms are treated individually and phase unwrapping is performed by integrating wrapped phase gradients in the spatial dimension under the assumption that the absolute phase difference between neighboring pixels is generally less than π [5]–[7]. This assumption is here referred to as the continuous-phase criterion. The most basic evaluation of the continuous-phase criterion is performed by summing up the wrapped phase differences of neighboring pixels in a close loop. Nonzero results are called residues and indicate a violation of the continuous-phase criterion. Lines of gradients greater than π necessarily run between residues of opposite signs [8]. In order to perform phase unwrapping, residues of opposite signs are connected (unloaded) by minimizing a cost function, e.g., the L^{p}-norm of the differences between gradients of unwrapped and wrapped phases, [5]–[7]. The arcs connecting the residues constitute an obstacle through which phase integration cannot be performed.

Persistent scatterer InSAR (PSI) techniques are an extension of InSAR, which are employed to reduce the influence of atmospheric artifacts and decorrelation in the estimation process by detecting long-term coherent scatterers (persistent scatterers) in a stack of interferograms [9]–[11]. In PSI methodologies, phase unwrapping is usually carried out in the time domain. Similarly to the spatial unwrapping case, the continuous-phase criterion in this case consist of assuming that the time evolution of the phases can be described with a model.

The model parameters are obtained by searching the solution space for the values that minimize a cost function that depends on the differences between modeled and observed phases. Based on the assumption that most of the contributions to the phase, such as atmosphere and deformation, are spatially correlated, the temporal model can be simplified by taking the phase differences between pairs of nearby PS [12]. These are also referred to arcs or double differences. Once unwrapped, the parameters describing the model and the unwrapped phases are integrated with respect to a reference. To avoid the propagation of errors in the integration, a combination of temporal and spatial unwrapping strategies was proposed in [10] and [13] and referred to as spatio-temporal unwrapping. In these methods, a redundant network of PS is built. The arcs forming this network are unwrapped in time. The estimated integer number of cycles of each arc is then subject to statistical tests aiming at detecting closure errors. Then, outlier arcs and PS are removed before integrating spatially to a common reference.

Although the main tendency in PSI techniques is to perform phase unwrapping in time, other strategies where phase unwrapping is carried out in the spatial domain were also suggested. The stepwise 3-D algorithm of [14] unwraps the data first in time. This is not the final result of the unwrapping process but it is used to constrain the solution in the other two dimensions, see also [15].

In any case, it is clear that temporal phase unwrapping is a crucial part of these algorithms. However, this step is performed

Manuscript received December 8, 2010; revised February 7, 2011; accepted March 20, 2011. Date of publication June 7, 2011; date of current version October 28, 2011.

The authors are with the Department of Earth Observation and Space Systems, Delft University of Technology, 2629 Delft, The Netherlands (e-mail: m.carocuenca@tudelft.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2011.2143722

arc-wise and does not take fully advantage of spatially correlated information.

Here, we investigate how to extract spatially two different types of information with the focus on applying it to aid temporal phase unwrapping. This information is the phase variance and the probability density function (pdf) of the parameters describing the temporal model. The inverse of the phase variance is used as a weighting factor to reduce the influence of noisy observations. In addition to that, the estimated pdf of the parameters of interest is applied in combination with Bayesian theory to rule out solutions that are not likely based on the spatial information.

II. Phase Unwrapping Method

In this section, we investigate how to estimate spatially the phase variance of a PS and the pdf of the model parameters which are assumed to be heights and deformation rates. Then, we discuss how phase unwrapping can benefit from the results of these estimates.

The method is tested with simulated data in Section III-A. In Section III-B, we apply the new approach to real data and compare the results with previously developed algorithms, namely, the ensemble coherence maximization method [9], [17] and the bootstrap estimator [18]–[20]. We use the framework of DePSI (Delft PSI method) [21] for the validation, where unwrapped arcs are subject to statistical tests [13]. We use these tests to measure the performance of our algorithm, in particular the number of arcs and PS that are not rejected.

A. Phase Variance Estimation

The phase precision is estimated spatially under the assumption that noise variance of a PS can be described from its surrounding PS, at least in some extent. This assumption is supported by the fact that coherence, which measures the degree of correlation between two radar images, is spatially correlated. Therefore, we expect different noise levels depending on location. Furthermore, for the variance estimation process, we group pixels with regard to their temporal behavior. In this paper, we say that two PS resemble each other when they fulfill these two conditions, short distance and similar temporal stochastics. Only PS that resemble the PS in question are used for the variance estimation.

The inverse of the estimated variance is used to weight the observations during phase unwrapping. The incorporation of the precision in phase unwrapping was already suggested in [10]. Here, the difference is that we do not assume that all PS in an interferogram have the same variance and we estimate it individually i.e., one value per PS per interferogram. In [16], the observations are weighted with an estimate of the coherence. Under the assumption of ergodicity, coherence is estimated spatially using a window of variable size. No distinction between the pixels inside the window was made. In this paper, we categorize them based on their temporal behavior.

The proposed method is summarized in the flowchart of Fig. 1. The variance of PS k in any interferogram i, $(\hat{\sigma}_k^i)^2$, is calculated from the phase noise estimated for k and for the PS resembling k. This operation is performed per interferogram i. The magnitude of the signal reflected by a PS (amplitude) is not included in the calculations for simplification.

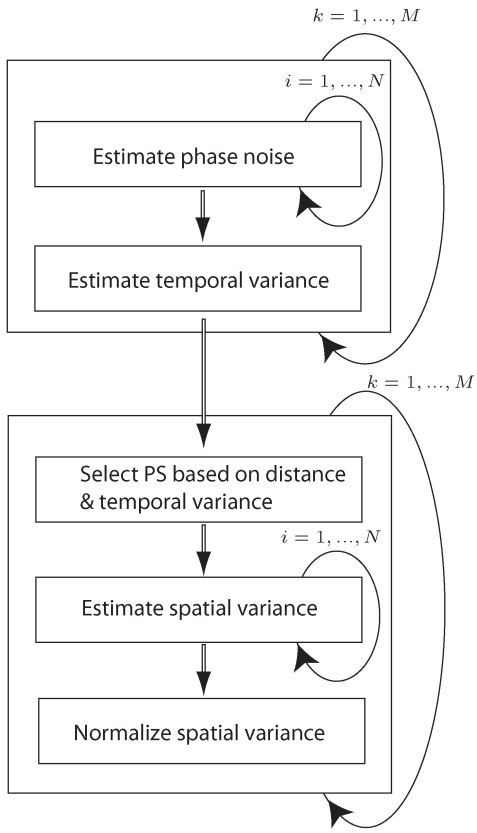

Fig. 1. Flowchart of the variance estimation process, where M is the number of PS and N the number of interferograms. First, the noise and the temporal variance is estimated for each PS. Second, the variance is estimated spatially using only PS resembling each other. The final result is normalized assuming that the mean in time of the spatial variance is equal to the temporal variance. Thin black arrows indicate loops over PS k or interferograms i.

The noise of a PS is estimated with a high-pass filter [22]. The spatially correlated terms such as atmosphere, surface deformation, and residual heights are obtained with bandpass filtering of the complex signal of surrounding PS.

The high-pass filter is then performed by subtracting the filtered phase $\phi^i_{\text{filtered},k}$ from the original phase ϕ^i_k. The result contains the non-spatially correlated or non-filtered surface deformation $(D^i_{\text{NSC},k})$ and residual heights $(H_{\text{NSC},k})$:

$$\phi^i_k - \phi^i_{\text{filtered},k} = \frac{-4\pi}{\lambda}\left(\frac{B^i_\perp}{R_k \sin\theta_k} H_{\text{NSC},k} + D_{\text{NSC}^i,k}\right) + \phi^i_{\text{noise},k} \quad (2)$$

where

- λ is the sensor wavelength.
- $B^i_\perp$ is the perpendicular baseline between master and acquisition i.
- R_k is the distance from the sensor to PS k.
- $\sin\theta_k$ is the sinus of the incidence angle θ_k.
- $\phi^i_{\text{noise},k}$ is the noise of interferogram i, which also includes unmodeled phase.

Substituting a linear model for non-spatially correlated surface displacements, (2) becomes

$$\phi_k^i - \phi_{\text{filtered},k}^i = \frac{-4\pi}{\lambda}\left(\frac{B_\perp^i}{R_k \sin\theta_k} H_{\text{NSC},k} + B_{\text{temp}}^i V_{\text{NSC},k}\right) + \phi_{\text{noise},k}^i \quad (3)$$

where B_{temp}^i is the temporal baseline between master and acquisition i and $V_{\text{NSC},k}$ is the rate of non-spatially correlated surface displacements of PS k.

Note that the original method of Hooper *et al.* [22] has been slightly modified to include non-spatially correlated displacements in the noise estimation [23].

Although the estimated noise, and therefore the subsequent estimation of the variance, depends on the filter settings, the absolute phase variance is of minor importance. The final goal of the algorithm is to weight the observations to aid phase unwrapping. Therefore, the algorithm focuses on estimating the variance factor that gives the relative weight between observations rather than on the absolute phase variance itself.

After removing the estimated contribution of non-spatially correlated residual heights $\hat{\phi}_{\text{NSC},H}^i$ and displacements $\hat{\phi}_{\text{NSC},D}^i$, the remaining phase is our estimate of the noise $\hat{\phi}_{\text{noise},k}^i$

$$\hat{\phi}_{\text{noise},k}^i = \phi_k^i - \phi_{\text{filtered},k}^i - \hat{\phi}_{\text{NSC},H}^i - \hat{\phi}_{\text{NSC},D}^i \quad (4)$$

where

$$\hat{\phi}_{\text{NSC},H}^i = \frac{-4\pi}{\lambda}\left(\frac{B_\perp^i}{R_k \sin\theta_k}\hat{H}_{\text{NSC},k}\right) \quad (5)$$

$$\hat{\phi}_{\text{NSC},D}^i = \frac{-4\pi}{\lambda}\left(B_{\text{temp}}^i \hat{V}_{\text{NSC},k}\right). \quad (6)$$

Temporal variability is measured with the variance estimated in time. For PS k, this operation is given by

$$\hat{\sigma}_{\text{time},k}^2 = \sum_{i=1}^{N} \frac{\left(\hat{\phi}_{\text{noise},k}^i - \hat{\mu}_{\text{noise},k}\right)^2}{N-1} \quad (7)$$

where N is the number of interferograms and $\hat{\mu}_{\text{noise},k}$ represents the sample mean of $\hat{\phi}_{\text{noise},k}^i$ estimated over time

$$\hat{\mu}_{\text{noise},k} = \sum_{i=1}^{N} \frac{\hat{\phi}_{\text{noise},k}^i}{N}. \quad (8)$$

We select the PS that lie at short distance with respect to PS k, and whose temporal variance is similar i.e., if the difference between the two is less than a threshold ξ:

$$\left|\hat{\sigma}_{\text{time},k}^2 - \hat{\sigma}_{\text{time},l}^2\right| < \xi \text{ with } l = 1, \ldots, m. \quad (9)$$

The maximum distance used for selecting PS and the threshold ξ are obtained, assuming that the temporal variance $\hat{\sigma}_{\text{time}}^2$ is stationary at short distances. We estimate a variogram spatially [24] where the observations are $\hat{\sigma}_{\text{time}}^2$ and the location is given by the corresponding PS coordinates. We employ as maximum distance twice the estimated range of the variogram (e.g., 1.5 km) and for the threshold ξ, its estimated sill (e.g., 0.05 rad^2).

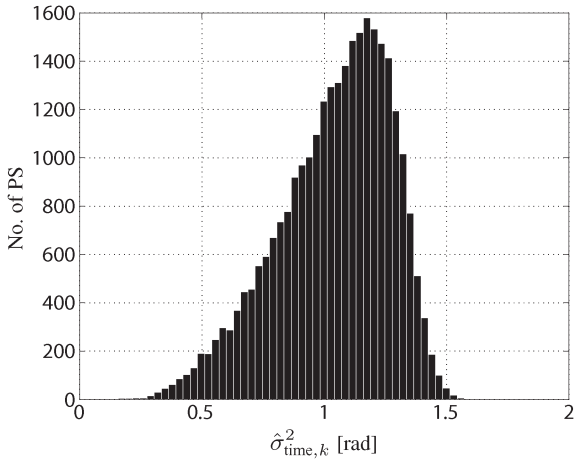

Fig. 2. Histogram of the temporal variance $\hat{\sigma}_{\text{time},k}^2$ estimated for a test area.

The selected PS are then said to resemble each other, and their estimated noise is used to compute the variance spatially. The variance of a generic PS k resembling m PS is given by

$$\left(\hat{\sigma}_k^i\right)^2 = \sum_{l=1}^{m} \frac{\left(\hat{\phi}_{\text{noise},l}^i - \hat{\mu}_{\text{noise},k}^i\right)^2}{m-1} \quad (10)$$

where $\hat{\mu}_{\text{noise},k}^i$ represents the sample mean calculated spatially with $l = 1, \ldots, m$ selected PS

$$\hat{\mu}_{\text{noise},k}^i = \sum_{l=1}^{m} \frac{\hat{\phi}_{\text{noise},l}^i}{m}. \quad (11)$$

This is repeated for all interferograms to build the vector $\hat{\sigma}_{\text{space},k}^2$ whose length is equal to the number of interferograms

$$\hat{\sigma}_{\text{space},k}^2 = \left[\left(\hat{\sigma}_{\text{space},k}^1\right)^2, \ldots, \left(\hat{\sigma}_{\text{space},k}^N\right)^2\right]. \quad (12)$$

Fig. 2 shows the histogram of $\hat{\sigma}_{\text{time},k}^2$ obtained from the test area described in Section III-B. From the histogram, we can deduce that PS with very high and very low values of $\hat{\sigma}_{\text{time},k}^2$ may not have nearby PS, because their number is very limited. Thus, for PS with low $\hat{\sigma}_{\text{time},k}^2$, the variance is not estimated spatially and $(\hat{\sigma}_k^i)^2$ is taken to be constant for all interferograms and equal to the value of $\hat{\sigma}_{\text{time},k}^2$. In principle, this assumption should not influence phase unwrapping of these PS very much because they are highly stable in time, i.e., $\hat{\sigma}_{\text{time},k}^2$ is low. On the other hand, for noisy PS (high $\hat{\sigma}_{\text{time},k}^2$) with no nearby PS, $(\hat{\sigma}_k^i)^2$ is not calculated at the first instance but computed as the mean of the variances obtained for the rest of the PS.

We correct $\hat{\sigma}_{\text{space},k}^2$ for the noise particular to k by assuming that the temporal variance of PS k, $\hat{\sigma}_{\text{time},k}^2$, represents the temporal mean of the variance estimated spatially

$$\left(\hat{\sigma}_k^i\right)^2 = \hat{\sigma}_{\text{time},k}^2 \frac{\left(\hat{\sigma}_k^i\right)^2}{\sum_{i=1}^{N}\left(\left(\hat{\sigma}_k^i\right)^2/N\right)} \quad (13)$$

$(\hat{\sigma}_k^i)^2$ is our final estimate of the variance of PS k in interferogram i, which can also be written in a vector form

$$\hat{\sigma}_k^2 = \left[\left(\hat{\sigma}_k^1\right)^2, \ldots, \left(\hat{\sigma}_k^N\right)^2\right] \tag{14}$$

with N the number of interferograms.

In summary, this method first estimates the noise of each PS. Then, it groups PS depending on distance and temporal behavior and uses their estimated noise to compute their variance.

Since our observations are the interferometric phases of an arc, we are interested in the variance of the differences of pairs of PS. For a generic arc lk spanning PS k and l, the estimated noise variance $\hat{\sigma}_{lk}^2$ is obtained from the propagation law of variances assuming independence

$$\hat{\sigma}_{lk}^2 = \hat{\sigma}_l^2 + \hat{\sigma}_k^2. \tag{15}$$

B. Phase Unwrapping Using Bayesian Theory

Bayes' rule uses the prior probability of some parameters of interest to estimate the conditional probability of these parameters given a set of observations. The application of Bayesian inference to extract information, namely heights, from interferometric phases was suggested in [25]. However, this method requires *a priori* knowledge of the pdf of the parameters of interest, e.g., heights. A similar approach is here reformulated but with the goal of providing a framework for phase unwrapping of PS time series. The method also estimates the pdf needed to apply Bayes's rule.

Although this is trivially extended to any other model, we start by assuming that the major contribution to the phase of arc lk are the height difference between PS k and l (H_{lk}), the displacement difference, which is here modeled as a linear function of time with velocity V_{lk} and a constant value M_{lk} to account for the correlation between interferograms, mainly produced because they are computed with respect to the same reference (master) image.

In order to unwrap the phases of arc lk, we search for the value of H_{lk}, V_{lk}, and M_{lk} that maximizes the conditional pdf

$$\wp\left(H_{lk}, V_{lk}, M_{lk} | \phi_{lk}^1, \ldots, \phi_{lk}^N\right) \tag{16}$$

where $\wp$ refers to the pdf and ϕ_{lk}^i (with $i = 1, \ldots, N$) are the interferometric phases of arc lk, also referred to as double differences, of N interferograms. For the sake of clarity, we drop the subscript referring to the arc from now on and replace it with interferogram number (ϕ_{lk}^i becomes ϕ_i). Then, the problem can be stated as follows:

$$[\hat{H}, \hat{V}, \hat{M}] = \max_{H,V,M} \left[\wp(H, V, M | \phi_1, \ldots, \phi_N)\right]. \tag{17}$$

Using Bayes' rule, $\wp(H, V, M|\phi_1, \ldots, \phi_N)$ becomes

$$\wp(X|\phi) = \frac{\wp(\phi|X)\wp(X)}{\wp(\phi)} \tag{18}$$

where ϕ is the vector of phase double differences $\phi = [\phi_1, \ldots, \phi_N]$ and X is the vector of unknowns $X = [H, V, M]$. The denominator in (18) can be ignored because it does not depend on H, V, or M.

The term $\wp(\phi|X)$ represents the pdf of the difference between ϕ and its expected value ϕ_0, which depends on the model parameters $X = [H, V, M]$, see (21). Since this difference $e = \phi - \phi_0$ contains mostly thermal noise and slave atmosphere, which is small because phase double differences are taken between nearby PS, we assume $\wp(\phi|X)$ to be Gaussian-distributed

$$\wp(\phi|X) = C_Q \cdot \exp\left(-\frac{1}{2} e^T Q_\phi^{-1} e\right) \tag{19}$$

with

$$C_Q = \frac{1}{\det(2\pi Q_\phi)}. \tag{20}$$

Q_ϕ is the variance-covariance (vc) matrix of the observations, i.e., phase double differences, which are assumed to be independent. Note that the correlation between interferograms due to the fact that all are with respect to the same master image is not included in the vc-matrix but estimated as the constant term M. Consequently, Q_ϕ is a diagonal matrix whose diagonal elements are the phase variances estimated with the method of Section II-A. The term ϕ_0 is also a vector $[\phi_1^0, \ldots, \phi_N^0]$. For interferogram i, the expected phase ϕ_i^0 is given by

$$\phi_i^0 = \frac{-4\pi}{\lambda}\left(\frac{B_\perp^i}{R \sin\theta} H + B_{\text{temp}}^i V\right) + M, \quad \text{with } i = 1, \ldots, N. \tag{21}$$

The same nomenclature as in (2) and (3) is used here.

The second term numerator of (18) can be obtained using *a priori* information, e.g., from leveling or SRTM measurements [25]. However, this is often not available for parameters other than heights. Here, we propose to estimate the pdf of the parameters of interest of an arc kl from the arcs around it.

1) Estimation of the pdf: The pdf of the parameters of interest $\wp\{X\}$ is estimated iteratively. Initially, for a given arc kl, we assume $\wp\{X\}$ to be a boxcar—or rect function—which is constant inside a bounded search space and zero outside. The search space is limited based on *a priori* knowledge. Then, we search for the solutions that maximize (18).

In the second iteration, we build the needed pdf from the estimates obtained for nearby arcs. The initial values obtained for the arc kl itself are not included in its pdf estimation.

At first instance, we describe the method for only one unknown e.g., heights H. Then, we extend it to include the other two, displacement rates V and a constant M.

We use indicator kriging [24] for estimating the pdf of the heights of any given arc. This estimate is denoted as $\hat{\wp}(H)$. The observations of indicator kriging are the estimated heights $\hat{H}$. These values, $\hat{H}$, are compared with a chosen list of thresholds $h_1 < h_2 < \cdots < h_W$. At a generic arc position p_0, taken as the middle point between the PS spanning the arc, we then evaluate

$$I(p_0, h_w) = \begin{cases} 1, & \text{if } \hat{H}(p_0) \le h_w \\ 0, & \text{otherwise} \end{cases} \tag{22}$$

for $w = 1, 2, \ldots, W$. This is repeated for all arcs. Then, we determine an indicator variogram from the $0|1$ observations. This is used to predict the cdf from which the pdf is calculated. First, $\hat{I}(p_0, h_w)$ is estimated at position p_0 using ordinary kriging:

$$\hat{I}(p_0, h_w) = \sum_{j=1}^{n} w_{0,j,w} I(p_j, h_w) \tag{23}$$

where $w_{0,j,w}$ is the weight of observation $I(p_j, h_w)$ evaluated at p_0 and obtained from the empirical variogram. The term n is the maximum number of observations used for the interpolation. Note that $j \neq 0$, i.e., the observation at p_0 is excluded from the pdf estimation.

The cdf estimated at p_0, $\hat{\varsigma}(p_0)$, is then given by

$$\hat{\varsigma}(p_0) = \left\{ \hat{I}(p_0, h_1), \hat{I}(p_0, h_2), \ldots, \hat{I}(p_0, h_K) \right\}. \tag{24}$$

The derivative of $\hat{\varsigma}$ yields the estimated pdf $\hat{\wp}$. The spacing between the thresholds h_w and h_{w+1} used in (22) are usually coarse. The finer step size is solved by 1-D interpolation at the required location. Smoothing of $\hat{\wp}$ may also be needed.

The same operations can be applied to the rest of the unknowns V and M. Then, $\hat{\wp}(H, V, M)$ is obtained assuming the unknowns are independent

$$\hat{\wp}(H, V, M) = \hat{\wp}(H)\hat{\wp}(V)\hat{\wp}(M). \tag{25}$$

III. Tests and Results

We tested the method for variance estimation and phase unwrapping in two ways. First using simulated data, where we compared known values with estimated solutions. Second, we applied the method to actual interferograms and measured the number of accepted arcs after statistical tests.

A. Test With Simulated Data

We created 20 synthetic interferograms whose phase contained atmosphere, height, and Gaussian-distributed noise. To simulate the effect of decorrelation, we simply assigned different noise levels to each interferogram.

The total number of pixels was 3136. The interferogram size was equivalent to 200 m in azimuth and 500 m in range direction. All arcs had the same pixel as reference, forming a starlike network. We considered all pixels as PS and did not perform any prior selection.

The atmosphere was simulated as isotropic 2-D fractal surfaces with a power law behavior which corresponds with the $[-2/3, -8/3, -5/3]$ power law, for short, medium, and large distance, respectively [3]. The height was simulated assuming a Gaussian pdf [25] with standard deviation of 5 m and zero mean. We begin by testing our method for variance estimation. Next, we focus on phase unwrapping using Bayesian theory.

1) Spatial Estimation of Phase Variance: To test our method for phase variance estimation, we first unwrapped the time series by weighting the phases with the inverse of the estimated variance. After that, since the actual simulated heights H were known, the number of wrongly unwrapped PS was calculated. We then repeated the phase unwrapping operations but in this case, we did not apply any weight. We compared the number of unwrapping errors when observations were weighted with $\hat{\sigma}_k^2$ with those when the observations were not weighted. For these two cases, we employed our Bayesian approach for phase unwrapping which required several iterations for calculating the pdf of the parameters of interest.

TABLE I
Number of Wrongly Unwrapped PS From Simulated Interferograms. Two Cases Are Compared: When the Wrapped Phases Are Weighted With the Inverse of the Estimated Variance $\hat{\sigma}_k^2$, and When the Phases Are Not Weighted. The Total Number of Pixels Was 3136

Iteration	weight =1/ $\hat{\sigma}_k^2$	No weight
0 (pdf=boxcar)	397	790
1 (updated pdf)	34	304
2 (updated pdf)	2	186
3 (updated pdf)	0	139

TABLE II
Number of Wrongly Unwrapped PS for Simulations With Different Noise Levels (Low, Medium, and High), See Text for Explanation. The Total Number of Pixels Was 3136

Iteration	low noise	medium noise	high noise
0 (pdf=boxcar)	4	397	1129
1 (updated pdf)	0	34	494
2 (updated pdf)	0	2	262
3 (updated pdf)	-	0	188

The results are summarized in Table I. The number of wrongly unwrapped PS after each iteration (updating the pdf) were always smaller when using the estimated variance than in the case when the observations were not weighted. Even more, four iterations proved to be enough to correct all wrongly unwrapped PS, while in the other case, more iterations were needed.

2) Bayesian Phase Unwrapping: To test our method for phase unwrapping, we created three stacks of 20 interferograms each. Different noise levels were assigned to each stack. The average standard deviation of the stacks were 0.36 rad (low-noise case), 0.47 rad (medium-noise case), and 1.10 rad (high-noise case). As before, unwrapping errors were found by comparing actual heights H with estimated ones $\hat{H}$. For all three interferograms stacks, we started with the same initial pdf. This was a boxcar that was equal to one within the bounds -40 m to $+40$ m and zero outside.

The results are summarized in Table II. They show that for the low-noise case, only four pixels were wrongly unwrapped and after the first update of the pdf, all errors were corrected. More iterations did not alter the results.

For the medium-noise case, the results of Table II are complemented with Fig. 3, where estimated heights are plotted against the corresponding real solution. In these graphs, correctly estimated heights should lie on a line with a slope equal to one, and most of them actually do. Solutions outside the central cluster represent unwrapping errors. Fig. 3(a) shows estimated versus real heights when $\hat{\wp}(H)$ was a boxcar. The initial number of 397 wrongly unwrapped PS dropped to 34 after a first update of the pdf. This can also be noticed by comparing Fig. 3(a) and

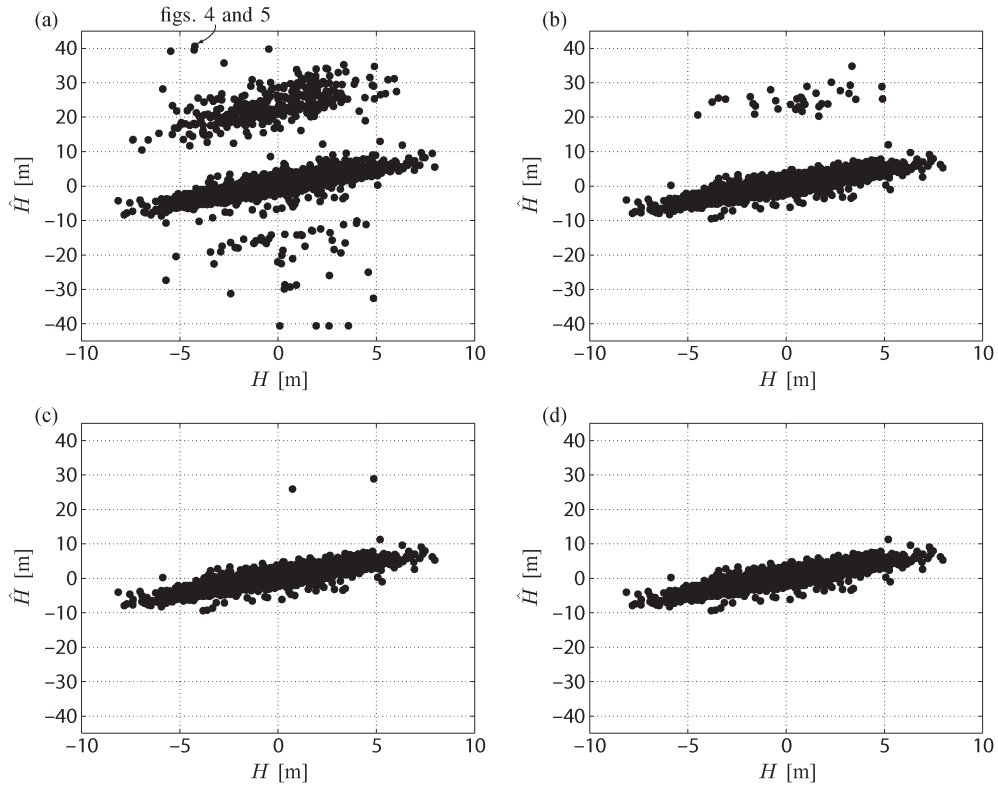

Fig. 3. Estimated $(\hat{H})$ versus actual value of the simulated height H for the medium-noise stack. The pixel marked with an thin arrow is further studied in Figs. 4 and 5. (a) Results after first estimation run: the initial pdf $\hat{\wp}(H)$ was a boxcar function. (b) Results after second estimation run: we employed an estimate of the pdf $\hat{\wp}(H)$. (c) Third run, the estimated pdf was updated $\hat{\wp}(H)$. (d) Fourth iteration, $\hat{\wp}(H)$ was updated once more.

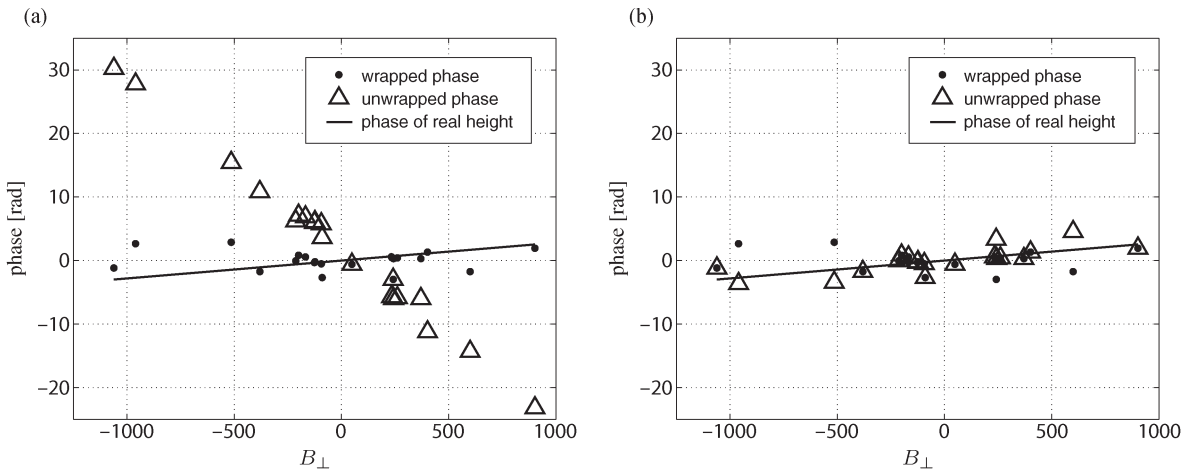

Fig. 4. (a) Wrapped and unwrapped phases of a PS whose height was initially wrongly estimated. (b) The unwrapped phase of the same pixel is corrected by updating $\hat{\wp}(H)$ from a boxcar to the value obtained using indicator kriging. See also Fig. 5. The phase corresponding the pixels real height is also shown, see legend.

(b). One more iteration reduced again the errors to only two, see Fig. 3(c). Three updates were required to correct them all as shown in Fig. 3(d). Finally, for high-noise interferograms also, major improvements were found after the first iteration. The number of wrongly unwrapped PS reduced from 1129 to 494, see Table II. Since the total number of pixels was 3136, this demonstrates that the algorithm proves to be efficient even if a third of the estimations are wrong. However, due to high noise, more than four iterations were needed to correct all the wrongly unwrapped PS.

For the medium-noise case, we analyzed the results further. We selected a pixel whose height was initially wrongly estimated and corrected after updating the pdf. The initial estimate of this pixel was 40.4 m while the true (simulated) solution was -4.2 m. The corresponding unwrapped phases are plotted against the perpendicular baseline $B_\perp$ in Fig. 4(a). They are shown as black triangles. The solid line represents the modeled phase due to true height (-4.2 m). After estimating $\hat{\wp}(H)$ and including it in the phase unwrapping process, a new height of -4.4 m was obtained, very close to the real value. The

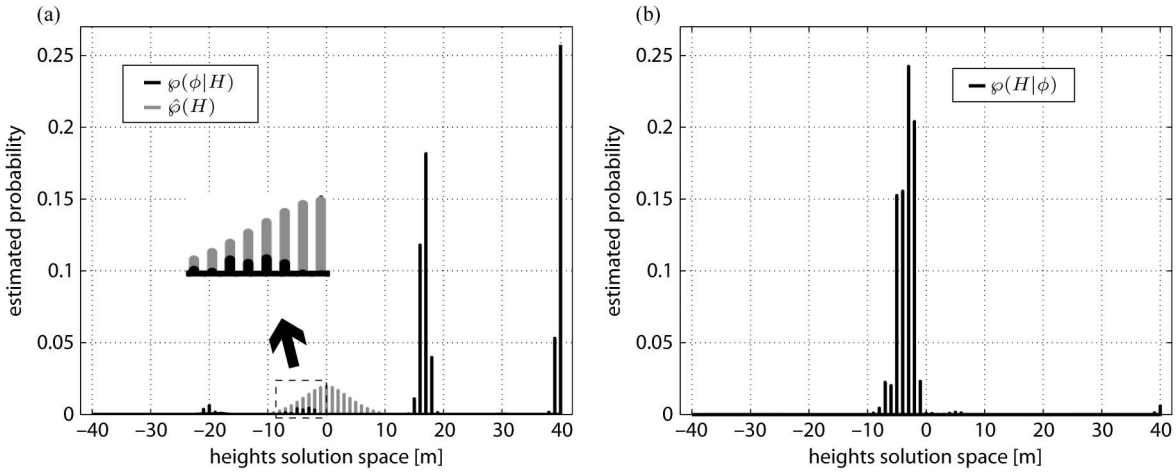

Fig. 5. Probabilities of the search space solutions obtained from simulated data. The step search, which was 1 m, created a discrete distribution (a) $\wp(\phi|H)$ is displayed in black. $\hat{\wp}(X)$ is a boxcar, not shown in the figure. In this case, $\wp(H|\phi)$ is equal to $\wp(\phi|H)$ inside the search space. The solution is given by $\wp(\phi|H)$ maximum value that is at 40.4 m. $\hat{\wp}(H)$, shown in gray, is estimated once all PS are unwrapped. The zoom shows also a local maximum in the neighborhood of the simulated (true) solution (-4.2 m). (b) $\hat{\wp}(H)$ is incorporated in the estimations with (18). $\wp(H|\phi)$ is displayed in black. Its maximum value gives a new solution of -4.4. m, nearer to the simulated height (-4.2 m).

corresponding unwrapped phases are shown in Fig. 4(b) also as black triangles. The improvement is clear. The differences between the unwrapped and phase due to height are caused by simulated noise and atmosphere.

For the same pixel, we also investigated in Fig. 5 how the maximum likelihood estimator, see (17), performed. In the first run, $\hat{\wp}(H)$ was a boxcar and the solution search was governed by $\wp(H|\phi)$, shown in black in Fig. 5(a). This was maximum at 40.4 m, therefore the selected solution. The conditional pdf $\wp(H|\phi)$ also had a local maximum in the neighborhood of the real solution -4.2 m, marked with a dashed rectangle and displayed as a zoom-in. The solution was searched in steps of 1 m, which produced a discrete $\wp(H|\phi)$. Once a maximum was found, the solution was refined by searching around the maximum in finer steps.

The estimated pdf $\hat{\wp}(H)$ is shown in gray in Fig. 5(a). Then, $\wp(\phi|H)$ was computed applying Bayes' rule, see (18). The resulting $\wp(\phi|H)$ is shown in Fig. 5(b). From this, we obtained the new corrected solution of -4.4 m.

From the simulations, we conclude that ambiguities estimated erroneously can be corrected using our algorithm.

B. Test With Real Data

We applied our methods for variance estimation and phase unwrapping to a time series of ERS1/2 images acquired over the Netherlands. We selected a small area in the southern Netherlands that was affected by motion due to water rebound in abandoned coal mines. The crop we analyzed was 15 km $\times$ 10 km wide and there were at least three different deformation modes, stable areas and two abandoned mines causing uplifting with different velocities.

In the DePSI [13], [21], a first-order network is built with the most coherent PS. The arcs forming the network are adjusted to a common reference and tested for outlier rejection [26]. The goal of this process is to remove spatially inconsistent arcs. Then, a second-order network is created by connecting the rest of the PS to the initial network at three different locations. These three arcs are unwrapped in time and integrated with respect to the initial common reference. Therefore, each second-order network PS has three independent solutions. PS with inconsistent—different—solutions are rejected.

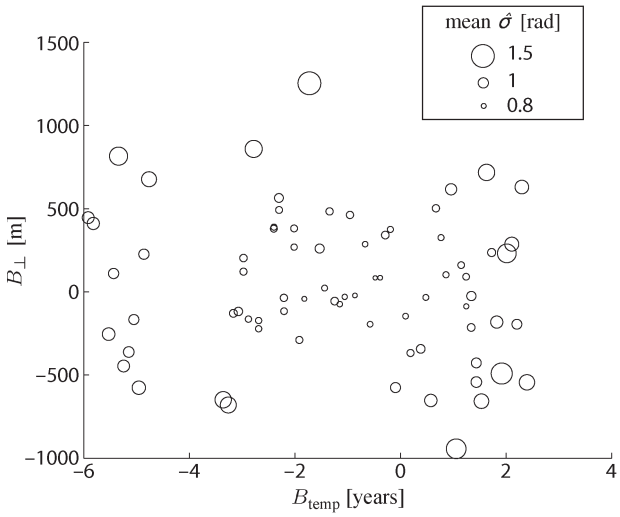

Fig. 6. Mean value of the estimated standard deviation for the arcs used in the initial network ($\hat{\sigma}$) plotted versus $B_\perp$ and B_{temp}. The size of the circle depends on the standard deviation.

1) Spatial Estimation of Phase Variance: We calculated the standard deviation of an arc as explained in Section II-A. The results are plotted against the temporal and perpendicular baseline in Fig. 6, where the area of the circles scales with the standard deviation. As expected, the variance increased with time and perpendicular baselines due to temporal and geometrical decorrelation, respectively.

From a total of 2051, the results showed that the number of accepted arcs was 1909 for our method. We compared this number with the results obtained using the variance component estimation (VCE) method proposed in [10], see Table III. This method calculates a posterior variance from the residuals (observed minus modeled phases) of the initial network. The

TABLE III
NUMBER OF FIRST-ORDER NETWORK ARCS AND PS ACCEPTED AFTER TESTS FOR THE PROPOSED METHOD (BAYES), BOOTSTRAP ESTIMATOR (BOOT.) AND COHERENCE MAXIMIZATION (PERIOD.). THE OBSERVATIONS WERE ALSO WEIGHTED WITH THE INVERSE OF THE VARIANCE, WHICH WAS ESTIMATED FROM THE METHOD HERE PROPOSED (WEIGHT $= 1/\hat{\sigma}_k^2$) AND EMPLOYING THE VCE METHOD OF [10], (WEIGHT $= 1/$VCE). FOR THE COHERENCE MAXIMIZATION METHOD, WE DID NOT WEIGHT THE OBSERVATIONS

	Bayes (weight=$1/\hat{\sigma}_k^2$)	Bayes (weight=1/VCE)	Boot. (weight=1/VCE)	Period. (No weight)
1st network (Total 2051 arcs)				
Accepted arcs	1909	1896	1876	1849
Selected PS (Total 28704)				
Accepted PS	5442	5383	4141	3660

number of accepted PS using VCE as weight was 1896. In both cases, we used the Bayesian approach to unwrap the arcs. We unwrapped the second-order network PS in the same manner. The results show that if we use as weight, the variance estimated spatially the number of total accepted PS is 5442 and when using the variance estimated from 5383. Although the differences are not high, the proposed method for variance estimation performed better in terms of accepted arcs and PS.

2) Bayesian Phase Unwrapping: We compared the results of our phase unwrapping method with those obtained by the bootstrap estimator and ensemble coherence maximization methods. This was performed in two ways. First, we compared the number of PS that were accepted by the statistical tests. Second, we inspected the estimated rates of all selected PS.

a) Comparison using statistical tests: We started by unwrapping the first-order network. The second-order network PS were unwrapped by connecting them to the first-order network at three different locations. The number of first-order network arcs that survived the test is different for each unwrapping method. Therefore, in order to have a fair comparison, we unwrapped the second-order network using the same first-order network for all methods. Table III summarizes the results of the statistical tests applied to each unwrapping method. The number of accepted arcs shows that our approach performed the best for both, first- and second-order networks. The method here proposed performs the best with 5442 compare with 4141 obtained with the bootstrap estimator and 3660 of the ensemble coherence maximization. Our new method also performed well when the observations were weighted with variance estimated with VCE method of [10].

b) Comparison using estimated rates: We also compared the deformation rates estimated with the proposed method with the values obtained using the bootstrap estimator and ensemble coherence maximization methods. The model also included height and master contribution but they are not here discussed. We employed the same initial search space for all cases. The initial search space for the velocities was restricted to the interval from -17.5 mm/yr to $+17.5$ mm/yr.

The results of the rates from all three methods are displayed in Fig. 7. After integrating the solutions with respect to a common reference PS, velocity values can be outside the initial search space. From visual inspection of Fig. 7, we see that the image with the least spatial noise corresponds to our new method shown in Fig. 7(a).

The estimates of the bootstrap estimator and ensemble coherence maximization method are plotted against the results

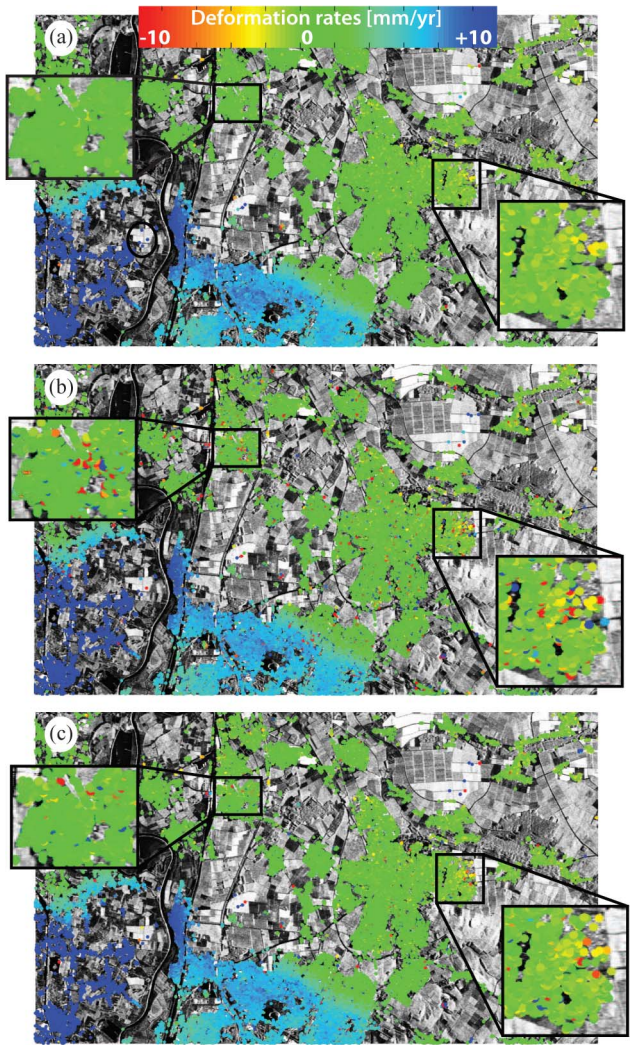

Fig. 7. Surface deformation rates estimated with (a) the method here proposed, (b) bootstrap estimator, and (c) ensemble coherence maximization. Black boxes show zooms of some of the areas where the differences between the methods are the greatest, after visual inspection.

obtained by our new approach in Fig. 8(b) and (c), respectively. If the estimates given by each method were the same, the values would lie on a diagonal line with slope one.

Most of the values form a line in Fig. 8(b) and (c) but some do not. Hence, different solutions were found by our method and the other two. The major differences are grouped in three parts of the plot. They correspond to the three modes of deformation

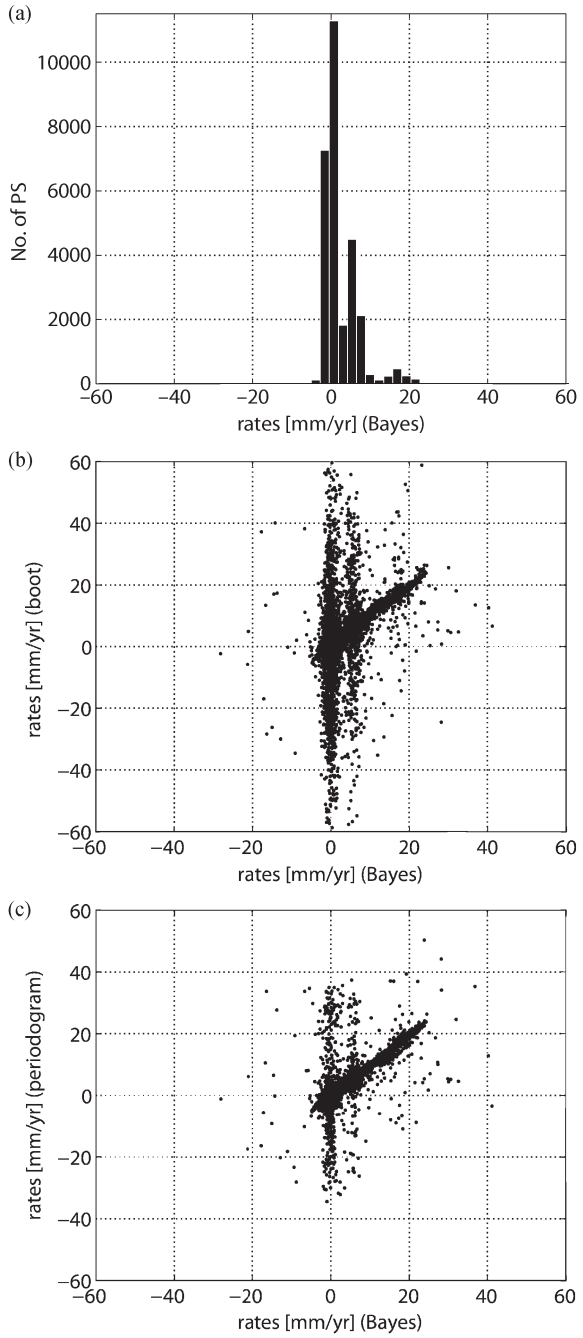

Fig. 8. (a) Histogram of the deformation rates estimated using the method here proposed (Bayes). (b) Rates estimated by bootstrap estimator (boot) versus our method (Bayes). (c) Rates estimated by ensemble coherence maximization (periodogram) versus our method (Bayes).

velocities, also visible in the histogram in Fig. 8(a). This shows the influence of $\hat{\varphi}(X)$ in the estimation process.

The introduction of $\hat{\varphi}(X)$ can affect the estimations in two ways. First, if the PS in question contains significant signal despite the noise, $\hat{\varphi}(X)$ will help to find the right solution by avoiding other maxima of $\wp(\phi|H)$, see (18). An example of this case is shown in Fig. 5, already discussed in Section III-A. Second, if the PS is completely noisy, the $\wp(\phi|H)$ plays almost no role in the estimations because it is expected to be uniform. Then, the solution is given by the maximum of $\hat{\varphi}(X)$. Although we believe that is very likely than the PS in question behaves as its nearby PS, the application of (18) with $\hat{\varphi}(X)$ carries the risk of assigning an estimate to the noisy PS that is mostly based on the surroundings. In any case, these pixels can be then detected based on spatially consistency tests or estimated noise.

We examined the probability $\wp(\phi|H)$ for some of the PS that had large differences between the estimation obtained by our new approach and the other two methods. We found in all cases that before including $\hat{\varphi}(X)$, there were local maxima near the final solution that was obtained by our method, similar to the situation shown in Fig. 5. Therefore, we conclude that the differences between the methods here examined are due to the first reason—$\hat{\varphi}(X)$ avoids wrong maxima—also proved with the simulations of Section III-A.

IV. Summary and Conclusion

We developed an algorithm that uses the spatial correlation between PS to improve phase unwrapping. This is done in two different ways, to calculate spatial variance and to constrain and weight the search space through Bayesian analysis. We tested the method using both simulated and real data. The results show a considerable improvement in the number of correctly unwrapped pixels in the case of synthetic data. With real data, the results were compared with those obtained by two other approaches. Our method shows a significant increase in the number of accepted arcs and PS after statistical tests, with at least 1200 more accepted PS than with the other two methods. From the results obtained from both simulated and real data, we conclude that the phase variance of PS and the pdf of their model coefficients can be calculated spatially and phase unwrapping can benefit from the two.

Despite the advantages, our method have some limitations. Since it performs spatial estimations, the approach requires sufficient data density. We also observed that the computational cost of our method was about three to four times higher than the other techniques here considered. This additional cost is caused by the extra spatial operations which are not performed by the other methods.

Acknowledgment

The authors would like to thank European Space Agency (ESA) for provision of data via the Cat-4048 project and Netherlands Research Center for Integrated Solid Earth Science (ISES) for supporting this research. The authors would also like to thank S. S. Esfahany for the valuable comments and discussions.

References

[1] D. Massonnet and K. L. Feigl, "Radar interferometry and its application to changes in the earth's surface," *Rev. Geophys.*, vol. 36, no. 4, pp. 441–500, Nov. 1998.

[2] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inverse Probl.*, vol. 14, no. 4, pp. R1–R54, Aug. 1998.

[3] R. F. Hanssen, *Radar Interferometry: Data Interpretation and Error Analysis*. Dordrecht, The Netherlands: Kluwer, 2001.

[4] P. J. G. Teunissen, *Adjustment Theory; An Introduction*, 1st ed. Delft, The Netherlands: Delft Univ. Press, 2000.

[5] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Sci.*, vol. 23, no. 4, pp. 713–720, Jul. 1988.

[6] D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. New York: Wiley, 1998.

[7] C. W. Chen, "Statistical-cost network-flow approaches to two-dimensional phase unwrapping for radar interferometry," Ph.D. dissertation, Stanford Univ., Stanford, CA, Jun. 2001.
[8] C. W. Chen and H. A. Zebker, "Two-dimensional phase unwrapping using statistical models for cost functions in nonlinear optimization," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 18, no. 2, pp. 338–351, Feb. 2001.
[9] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 8–20, Jan. 2001.
[10] B. M. Kampes, *Radar Interferometry: Persistent Scatterer Technique*. Dordrecht, The Netherlands: Springer-Verlag, 2006, ser. Remote Sensing and Digital Image Processing.
[11] A. Hooper, "Persistent scatterer radar interferometry for crustal deformation studies and modeling of volcanic deformation," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2006.
[12] C. Colesanti, A. Ferretti, F. Novali, C. Prati, and F. Rocca, "SAR monitoring of progressive and seasonal ground deformation using the Permanent Scatterers Technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 7, pp. 1685–1701, Jul. 2003.
[13] F. J. van Leijen, R. F. Hanssen, P. S. Marinkovic, and B. M. Kampes, "Spatio-temporal phase unwrapping using integer least-squares," in *Proc. 4th Int. Workshop ERS/Envisat SAR Interferometry, "FRINGE05"*, Frascati, Italy, Nov. 28–Dec. 2, 2005, p. 6.
[14] A. Hooper and H. Zebker, "Phase unwrapping in three dimensions with application to InSAR time series," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 24, no. 9, pp. 2737–2747, Sep. 2007.
[15] A. Hooper, "A Statistical-Cost approach to unwrapping the phase of InSAR time series," in *Proc. Adv. Sci. Appl. SAR Interferometry, "FRINGE09"*, Frascati, Italy, Nov. 30–Dec. 4, 2009, p. 6.
[16] F. De Zan and F. Rocca, "Coherent processing of long series of SAR images," in *Proc. Int. Geosci. Remote Sens. Symp.*, Seoul, Korea, Jul. 25–29, 2005, pp. 1987–1990.
[17] C. C. Counselman and S. A. Gourevitch, "Miniature interferometer terminals for earth surveying: Ambiguity and multipath with the Global Positioning System," *IEEE Trans. Geosci. Remote Sens.*, vol. GRS-19, no. 4, pp. 244–252, Oct. 1981.
[18] P. J. G. Teunissen and C. C. J. M. Tiberius, "Integer least-squares estimation of the GPS phase ambiguities," in *Proc. Int. Symp. Kinematic Syst. Geodesy, Geomatics Navig.*, Banff, AB, Canada, Aug. 30–Sep. 2, 1994, pp. 221–231.
[19] B. M. Kampes and R. F. Hanssen, "Ambiguity resolution for permanent scatterer interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 11, pp. 2446–2453, Nov. 2004.
[20] P. J. G. Teunissen, "On InSAR ambiguity resolution for deformation monitoring," *Artif. Satellites*, vol. 41, no. 1, pp. 17–22, 2006.
[21] V. B. H. Ketelaar, *Satellite Radar Interferometry, Subsidence Monitoring Techniques*, vol. 14. Dordrecht, The Netherlands: Springer-Verlag, 2009, ser. Remote Sensing and Digital Image Processing.
[22] A. Hooper, P. Segall, and H. Zebker, "Persistent scatterer interferometric synthetic aperture radar for crustal deformation analysis, with application to Volcan Alcedo, Galapagos," *J. Geophys. Res.*, vol. 112, p. B07 407, 2007.
[23] A. Schunert, M. Even, U. Sörgel, and K. Schulz, "Detection of localized surface deformation using a modified stamps algorithm," in *Proc. Adv. Sci. Appl. SAR Interferometry, "FRINGE09"*, Frascati, Italy, Nov. 30–Dec. 4, 2009, p. 6.
[24] H. Wackernagel, *Multivariate Geostatistics: An Introduction With Applications*. Berlin, Germany: Springer-Verlag, 1995.
[25] M. Eineder and N. Adam, "A maximum-likelihood estimator to simultaneously unwrap, geocode, and fuse SAR interferograms from different viewing geometries into one digital elevation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 24–36, Jan. 2005.
[26] P. J. G. Teunissen, *Testing Theory; An Introduction*, 1st ed. Delft, The Netherlands: Delft Univ. Press, 2000.

Miguel Caro Cuenca received the degree in general physics and applied physics from Granada University, Granada, Spain and Napier University, Edinburgh, U.K., respectively, in 2000, and the M.Sc. degree in microwave communication from Portsmouth University, Portsmouth, U.K., in 2001. He is currently working toward the Ph.D. degree at Delft University of Technology, Delft, The Netherlands.

After working in the radar industry under the Marie Curie research program, he joined Delft University of technology in 2006. His main interests concern advanced processing techniques for remote sensing data, in particular radar interferometry and their application to study of geophysical phenomena.

Andrew J. Hooper received the M.S. and Ph.D. degrees in geophysics from Stanford University, Stanford, CA, in 2002 and 2006, respectively, for his development of SAR persistent scatterer interferometry algorithms for geophysical applications.

Between 2002 and 2003, he worked at the German Space Center (DLR), Oberpfaffenhofen, Germany, on Doppler centroid determination for the TerraSAR-X satellite. From 2006 to 2008, he was a Research Scientist at the University of Iceland, Reykjavik, Iceland, developing more general SAR interferometry time series algorithms and applying them to model deformation at Icelandic volcanoes. Since 2008, he is Assistant Professor at Delft University of Technology, Delft, The Netherlands. His current research interests include imaging and modeling magma movement at volcanoes, slip on faults, and the solid Earth response to retreating glaciers.

Ramon F. Hanssen (M'04) received the M.Sc. degree in geodetic engineering and the Ph.D. (*cum laude*) degree from the Delft University of Technology, Delft, The Netherlands, in 1993 and 2001, respectively.

He was with the International Institute for Aerospace Survey and Earth Science (ITC), Stuttgart University, Stuttgart, Germany; the German Aerospace Center (DLR); Stanford University, Stanford, CA (Fulbright Fellow); and Scripps Institution of Oceanography in the field of microwave remote sensing, radar interferometry, signal processing, and geophysical application development. Since 2008, he has been an Antoni van Leeuwenhoek Professor in Earth observation with the Delft University of Technology, where, since 2009, he has been leading the research group on mathematical geodesy and positioning. He is the author of a textbook on radar interferometry.