TEMPORAL PHASE UNWRAPPING USING BAYESIAN INFERENCE

Miguel Caro Cuenca and Ramon F. Hanssen

m.carocuenca@tudelft.nl
Delft Institute of Earth Observation and Space Systems
Faculty of Aerospace Engineering
Delft University of Technology,
Delft, The Netherlands

ABSTRACT

When analyzing a stack of radar images using persistent scatterer techniques, one of the major limitations that temporal unwrapping methods have is that they frequently do not take advantage of the information provided by the rest of the detected scatterers. Here, we propose to use this information in two different ways. First, with an iterative method that utilizes the estimations from previous runs to constrain the solution using Bayesian inference. By doing so, solutions given by local maxima can be avoided. Second, we estimate the variance of the observations from the phase variability of the pixels around the PS in question. These variances are then used to weight the observations.

The method is first validated using simulated data. Then, we applied it to a small area in southern Netherlands and compare with conventional unwrapping methods (e.g. periodogram and bootstrapping). The method shows a decrease in the number of rejected arcs after statistical testing. It also reduces the spatial variability of the estimations and renders a smooth solution.
Spatially smooth estimations are , in principle, logic since the unknown parameters (deformation rates and heights) are usually spatially correlated and the uncorrelated part is normally small. However, it can happen that in the selection process included pixels that are not truly PS , i.e. we committed type II errors. These wrongly selected pixels whose phase contain mostly noise will have a solution artificially assigned that is similar to the surroundings and which makes it difficult to reject for further analysis. Other methods to detect these pixels should be used which estimates the noise of individual time series.

Key words: Persistent scatterer ; InSAR; temporal phase unwrapping.

1. INTRODUCTION

Phase unwrapping is the most crucial step in parameter estimation from time series of interferometric phases. Several methods have already been proposed such as periododogram (Ferretti et al. 2001) or integer least-squares estimator (Kampes & Hanssen 2004). However, they methods frequently focused the temporal characteristics of the arcs connecting two points. New algorithms have already been explored which do employ this information in the estimation of the spatial coherence, which is used to weight the observations, (De Zan & Rocca 2005).
Here, we investigate how to include this information provided by all detected scatterers for parameter estimation from interferometric time series. We propose to use it in two subsequent ways. First, in noise estimation from which we calculate the variance of single PS in an interferogram. Second, and most importantly, to optimize the solution search using Bayesian inference.

We applied this concept to both simulated and real data and compared the results with previously developed algorithms. To do so, we made use of the PS framework of DePSI (Delft PSI method) (Ketelaar 2008). On the whole, DePSI consists of establishing first an initial network with the most coherent PS and estimating the parameters of interest per arc, atmosphere among others. Then, the arcs are submitted to statistical test, such as closures errors for outlier removal (van Leijen et al. 2006). Once the atmosphere is removed, the rest of the PS are integrated to the initial network and the parameters estimated. These are also subject to outlier rejection.
These tests allows to measure the efficiency of an algorithm because it detects unwrapping errors. We use the numbers of detected outliers and unwrapping errors reported by this test to compare it with two of the most widely employed algorithms, periodogram (Ferretti et al. 2001) and bootstrapping, which is a simplification of integer least-squares estimator (Kampes & Hanssen 2004).

2. METHOD

We investigate how to improve phase unwrapping in two different manners. First, by estimating the variances of the observations that will be used to weight the observations. Then, we employ Bayesian inference to constrained the solutions.

2.1. Spatial variance

To estimate the noise of the observations (the phase of a PS in an interferogram) we start from two assumptions. First, persistent scatterers are not completely isolated, i.e. some of the surrounding pixels are also persistent. Second, interferometric phases are ergodic, i.e. spatial and ensemble averages are equivalent. Consequently, we are assuming that the noise of a given PS can be characterized using the rest of PS around. Notice also that our definition of PS is a scatterer whose signal to noise ratio is low, which means that information can be extracted from the phases despite of the noise.

At this stage, the initial selection of the PS was already done and and the rest of the pixels removed. The method for estimating the variance, which is based onHooper (2006) with the difference that we obtain the variance of an arc instead of a single PS, is as follows. We select the PS that lie at a distance shorter than the atmospheric decorrelation distance, i.e. $\sim$ 800 m, with respect to the pixel in question (j). After that, we calculate the complex average of the near PS and subtracted from their complex values. Finally, the estimated standard deviation of one pixel, $\hat{\sigma}_j$, is calculated from the remaining phases, r_j, which should contain mostly noise:

$$\hat{\sigma}_j^2 = \sum_{m=1}^{N} \frac{\left(r_j^m - \hat{r}_j\right)^2}{N}, \tag{1}$$

where r_j^m is the noise of the m^{th} PS close j and N is the total number of near PS.

However, since InSAR methodologies give relative measurements, we are interested in finding the variance of an arc, which represents double phase differences. Applying the propagation law of covariances we obtain the variance of the arc spanned two PS, which are denoted with the subscripts i and j:

$$\hat{\sigma}_{ji}^2 = A \cdot Q_{r_j,r_i} \cdot A', \tag{2}$$

where A is the design matrix of the double differences, $A = [1, -1]$, and Q_{r_j,r_i} is the variance-covariance matrix of the vectors r_j and r_i. Equation 2 can also be written in a simpler form:

$$\hat{\sigma}_{ji}^2 = \hat{\sigma}_j^2 + \hat{\sigma}_i^2 - 2\hat{C}_{ji}. \tag{3}$$

$\hat{C}_{ji}$ is the covariance between r_j and r_i, the off-diagonal element of the symmetric matrix Q_{r_j,r_i}. Notice that, r_j is a vector containing the estimated noise of the PS near j. Therefore, the covariance is calculated between two vectors, r_j and r_i.

If no scatterers are found at a short distance the variance is estimated as suggested by (Kampes 2006). It employs the residues of the estimation of some selected arcs to compute one variance per interferogram.

2.2. Phase unwrapping using Bayesian inference

We introduce Bayesian inference in the estimation of the parameters of interest from the wrapped phases. We assume that DEM error H, velocity V and master atmosphere, M and noise are the only contribution to the phase of an arc.

For a given arc ij, we search the value of H_{ij}, V_{ij} and M_{ij} that maximizes the conditional probability:

$$P(H_{ji}, V_{ji}, M_{ji} \mid \phi_{ji}^1, ..., \phi_{ji}^N), \tag{4}$$

where ϕ_{ji}^k (with $k = 1, ..., N$) are the interferometric phases of the arc ij in N interferograms. For the sake of simplicity, we drop the subscript referring to an arc and change it to indicate now interferogram number (ϕ_{ji}^k becomes ϕ_k).
In this algorithm, we apply the maximum-likelihood estimator, then the problem can be stated as follows:

$$[\hat{H}, \hat{V}, \hat{M}] = \text{argmax}\left[\text{pdf}(H, V, M \mid \phi_1, ..., \phi_N)\right], \tag{5}$$

where pdf stands for probability density function.

Using Bayesian theory, $\text{pdf}(H, V, M \mid \phi_1, ..., \phi_N)$ can be written as follows:

$$\text{pdf}(X \mid \phi_1, ..., \phi_N) = \frac{\text{pdf}(\phi_1, .., \phi_N \mid X)\,\text{pdf}(X)}{\text{pdf}(\phi)}, \tag{6}$$

where X is the vector of unknowns $X = \{H, V, M\}$. Since the denominator in eq. (6) is independent of H , V or M, it will be ignored.

We assumed that the phase differences between the PS forming an arc are Gaussian distributed. In principle, this is reasonable for short arcs i.e. maximum length of $\sim$ 2.5 km. Therefore, the first term on the right-hand side of eq. (6) is given by:

$$\text{pdf}(\phi_1, .., \phi_N \mid X) = C_Q \cdot \exp\left(-\frac{1}{2} e^T Q_\Phi^{-1} e\right), \tag{7}$$

where $e = \Phi - \Phi_0$ and

$$C_Q = \frac{1}{\det(2\pi Q_\Phi)}. \tag{8}$$

Φ is the vector of observations $[\phi_1, .., \phi_N]$ and the corresponding expected value $E\{\Phi\} = \Phi_0 = [\phi_1^0, .., \phi_N^0]$. In this case, ϕ_i^0 is given by:

$$\phi_0^i = \frac{-4\pi}{\lambda}\left(\frac{B_\perp^i}{R\sin\theta}H + B_{\text{temp}}^i V\right) + M \qquad (9)$$

where $B_\perp^i$ is the perpendicular baseline of interferogram i with respect to the master image, λ the wavelength, R the distance from the satellite to the scatter, θ the looking angle, and B_{temp}^i the temporal baseline with respect to master time.
In addition to that, Q_Φ is the variance covariance matrix of the observations, i.e. interferometric double differences, which are assumed to be independent. Consequently, Q_Φ is a diagonal matrix whose diagonal elements are obtained applying eq. (3) in each interferogram. In single master stack processing, this is equivalent of assuming that the noise is mostly produced by geometrical and temporal decorrelation, which is in principle valid for short arcs (low atmospheric contribution) and low thermal noise.

The second term numerator of eq. (6) can be obtained using a priori information, e.g. from leveling or SRTM measurements, (Eineder & Adam 2005). However, this can be extremely difficult for parameters other than heights. Therefore, we propose to derive this information from the technology that we have more at hand, i.e. radar interferometry.
This is in fact, the most relevant part of the algorithm: The incorporation and estimation of the pdf of the parameters of interest. We propose to compute it in an iterative manner. Initially, we assume a boxcar or rect function which is 1 inside the search space and 0 outside, for each of the pdf's of the unknowns. The limits of search space are based on a priory knowledge. Once the solutions are found for each arc we use this information to build the pdf of the unknowns. These are incorporated in eq. (6) assuming they are independent:

$$\text{pdf}(H, V, M) = \text{pdf}(H)\,\text{pdf}(V)\,\text{pdf}(M). \qquad (10)$$

The derivation of the pdf's on the right-hand side of eq. 10 is obtained from the histogram of the solutions with the bins centered at an initial coarse grid. The histogram should be normalized by the total amount of observations. We also applied a smoothing window to avoid peaks which could be produced because a limited number of measurements (arcs). The value of the pdf's in a fine grid is computed by interpolation. Currently, pdf(H, V, M) is calculated globally using the solutions given by the arcs of the initial network, which contains the most reliable PS. Then, during the densification step of the DePSI method, the individual pdf's are updated locally, within a grid cell whose size can change from 400^2 to 800^2 m^2. Other methods for calculating local pdf's are currently under investigation.

3. VALIDATION AND RESULTS

We validate the method in two ways. First using simulated data, where we compared known values with calculated solutions. Second, we tested the method with actual interferograms and measured the amount of accepted arcs after statistical tests.

3.1. Simulated data

We created 20 synthetic interferograms whose phase contained atmosphere, DEM error, and Gaussian distributed noise.
The atmosphere was reproduced from isotropic 2 dimensional 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, (Hanssen 2001). The DEM error was simulated assuming a Gaussian pdf (Eineder & Adam 2005) with standard deviation of 5 meters and zero mean.
Figure 1 displays the estimated solution for the DEM error against the simulated value. Figure 1A shows that there are four points that were unwrapped incorrectly. These were corrected after including the pdf of the initial solutions of DEM error using eq. 6. This is shown in fig. 1B. We also computed more iterations, but no improvements were found.

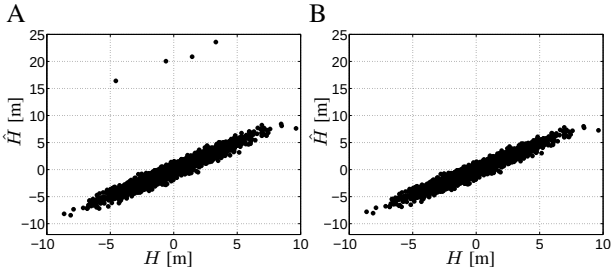

Figure 1. Estimated versus actual simulated DEM error. A) First run, pdf(X) is a boxcar function, B) Second run, with an updated pdf(X).

3.2. Real Data

We applied the method to a time series of ERS1/2 images acquired over the Netherlands. We selected a small area including the city of Sittard in the Dutch province of Limburg, in southern Netherlands. The region is affected by an interesting uplifting phenomena due to water recovery in abandoned coal mines. In the area we processed, which is 15x10 km^2 wide, there are at least three different deformation modes: stable areas, and two coal mines uplifting with different velocities, see fig. 4.

The test consist of comparing the number of accepted arcs after outlier rejection with other methods. In principle,

the test is very strict and if an arc is finally accepted is because it is very likely to be correctly unwrapped. Apart from that, no extra thresholding was included.

Table 1. Number of arcs belonging to the first order network (total number of arcs 2051) that were correctly unwrapped, for the spatial variance method and for VCE (Kampes 2006).

	Spatial variance	VCE
1st network (2051 arcs)		
Accepted arcs	1883	1850
Accept rate	91.8%	90.2%

3.2.1. Spatial variance

We calculated the standard deviation of an arc as explained in section 2.1. The results are plot against the temporal and perpendicular baseline in fig. 2. As expected, we found that the variance increase with time and perpendicular distance, both with respect to the master image. This is due to temporal and geometrical decorrelation, respectively.
We also compared the number of accepted arcs using our approach, i.e. spatial variance, with the number obtained using the Variance Component Estimation (VCE) proposed by Kampes (2006). This method calculates a posteriori variance from the residues of the estimations of the initial network. In order to have a fair comparison, we employed in both cases the Bayesian approach. The weight assigned to the observations was equal to the inverse of the variance, Q_{Φ}^{-1} in eq. (7). The results shows that the number of accepted arcs was 1883 for spatial variance. However, we obtained 1850 employing VCE, from a total of 2051 in both cases. This is summarized in table 1. Based on that numbers, we concluded that the variance seems to be correctly estimated. Therefore, the noise of a PS is well characterized by the rest of near PS, at least for our study area, which is highly urbanized.

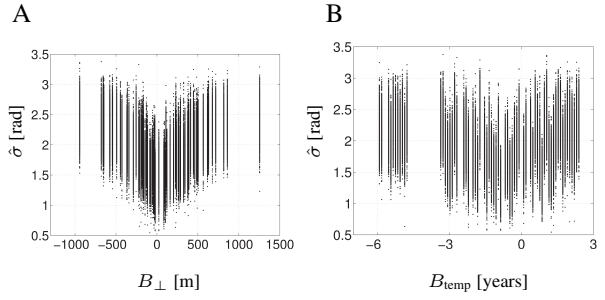

Figure 2. Estimated standard deviation for the arcs used in the initial network ($\hat{\sigma}$). (A) $\hat{\sigma}$ plotted against perpendicular distance between acquisitions $B_{\perp}$. (B) $\hat{\sigma}$ plotted against time difference between acquisitions B_{temp}.

3.2.2. Bayesian unwrapping

We compare the results of our method with the ones obtained by bootstrapping and periodogram in two different ways. First, using the number of PS that were accepted by the statistical tests. Second, by inspecting the estimated rates. In the last case, all pixels initially selected based on the coherence were output, i.e. no statistical test was carried out.

Comparison using statistical tests. In the Delft implementation of PSI (DePSI) (van Leijen et al. 2006) we initially build a first order network with the most coherent PS. The arcs forming the network are adjusted to a common reference and tested for outlier detection. The goal of this process is to remove wrongly unwrapped arcs. After this is finished, a second order network is created by connecting the rest of the PS to the initial network in three different points. We use those three arcs to decide if a given PS is correctly unwrapped.
From the number of accepted arcs, we found that Bayesian unwrapping performs the best for both, the first order network and all selected PS, with 91.8% and 17.6% accept rate, respectively, compared to 89.7% and 15.2% of bootstrapping and 85.8% and 14.5% of periodogram. The results are summarized in table 2.

Comparison using estimated rates. We estimated deformation rates, and also height differences and master atmosphere, for all points that were initially selected based on coherence without applying outlier rejection tests. The results of the rates are displayed in fig. 4. The area marked with a blue triangle was used as reference. The initial search space for the velocities was from -35 cm/yr to +35 cm/yr for the arcs, but after adjusting them to a reference the solutions can be larger. It can be seen that the image with less spatial noise corresponds to the Bayesian method (fig. 4A). We also plot the estimates of periodogram and bootstrapping methods against the results obtained by our approach in fig. 3B and C, respectively. In principle, if the estimations given by any method were the same the values should lie on a diagonal line. Effectively, we see that most of the values form a line in figs. 3B and C, but some do not. Hence, different solutions were found by the Bayesian method and the other two techniques. The major differences are grouped in three parts of the plot. They correspond to the three modes of deformation velocities, also visible in the histogram in fig. 3A. The reason is due to the pdf(X) we use in eq. 6.
The introduction of the term pdf(X) can affect the estimations in two ways. First, if the PS in question contains information despite the noise, pdf(X) will help to find the right solution by avoiding a local maxima of eq. 7. Second, if the point is completely noisy eq. 7 does not play a roll in the estimations and the solution is given my the maximum of pdf(X). This carries the danger of assigning an artificial estimate to the PS in question. Nevertheless, we believe that figs. 3B and C are due to the first

Table 2. Number of PS correctly unwrapped and corresponding success rate for the methods bootstrapping, Bayesian and periodogram. The second row included only the arcs selected for the first order network (total number of arcs 2051). The third row show the results for all PS that were initially selected (total number of PS 28704).

	Bayesian	Boot.	Period.
1st network (2051 arcs)			
Accepted arcs	1883	1840	1760
Accept rate	91.8 %	89.7%	85.8%
Selected PS (total 28704)			
Accepted PS	5047	4373	4175
Accept rate	17.6%	15.2%	14.5%

reason, — pdf(X) avoid local maxima —, because those pixels were initially selected based on their low noise, i.e. high coherence.

Apart from that, it must be said that although we employed initially the same search space for all cases, the method bootstrapping treats it differently. The limits are not fixed but flexible. Therefore, a solution can be searched outside the initial boundaries. This explains the noisy texture of figs. 4B and 3C.

4. SUMMARY AND CONCLUSION

We have developed a new unwrapping algorithm that uses the first estimated solutions of the PS with low noise to constrained the final estimations. This information was included in the process through Bayesian inference. We also computed the variance of each observation and use them as weight in the estimations. We found that using the spatial variance approach the number of rejected arcs was smaller when compared with VCE (Kampes 2006), which proofs that the noise was correctly estimated. Thus, we concluded that the variability of the phase can be well characterized using PS which are near by. This means that a persistent scatterer decorrelates similar to those ones lying closed by, at least for our study area, which is highly urbanized. In addition to that, the variance behaves as expected, i.e. it increases with temporal and perpendicular baselines due to temporal and geometrical decorrelation.

The Bayesian unwrapping method we developed shows an considerable increase on the arcs acceptance rate when compared with bootstrapping and periodogram. The reason is that we included extra information (the solutions of the most stable PS) which is not used by the other two methods. A decrease of unwrapping errors was also found with synthetic data.

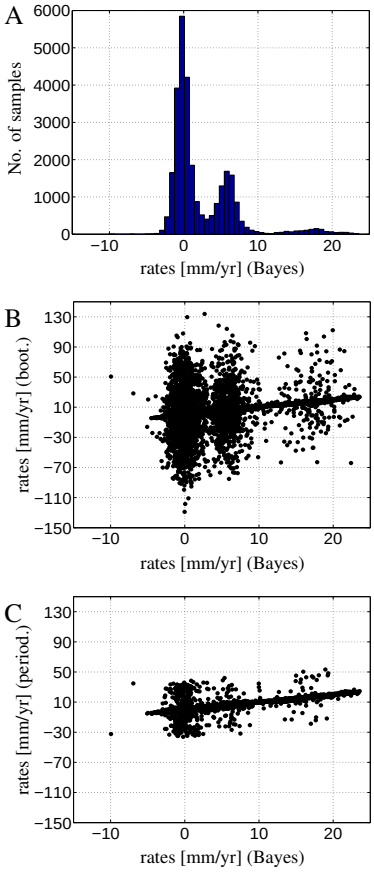

Figure 3. (A) histogram of the deformation estimated using Bayesian approach. (B) and (C) plot the deformation estimated by periodogram, and bootstrapping,respectively, against the Bayesian solutions.

Currently, the algorithm estimates a global initial pdf using the arcs of the initial network. Then, the pdf is updated locally by including the PS inside the grid cell that is being processed. This last step has however a minor influence. Therefore, other methods are currently being investigated, for example, using kriging for local pdf estimations (Goovaerts 1997).

Despite the advantages, there are two drawbacks that are worth mentioning. First, the processing time increases with respect to the other methods. The Bayesian approach took about 5 times more. The second disadvantage occurs for the pixels that contained not information, i.e. their phases represent only noise. For these PS, the term pdf($\phi_1, .., \phi_N \mid X$) in eq. 6 will not play an important role an they will be unwrapped following only pdf(X), and an artificial solution will be given. Despite the fact that it is very likely that the deformation or height of the noisy scatterers behaves similar to the rest of detected PS in a certain area, these objects should be identified and carefully processed in any further analysis. The detection of these types of pixels cannot be done any more based on spatial noise. Other methods that estimate the noise individually or temporally, must be employed, e.g. ensemble coherence. Nevertheless, if the initial selection is accurate enough the number of noisy pixels

should be minimal.

ACKNOWLEDGMENTS

We would like to thank to European Space Agency (ESA) for their support via the Cat-4048 project and the data supply for this research.

REFERENCES

De Zan, F. & Rocca, F. 2005, in International Geoscience and Remote Sensing Symposium, Seoul, Korea, 25–29 July 2005

Eineder, M. & Adam, N. 2005, IEEE Transactions on Geoscience and Remote Sensing, 43, 24

Ferretti, A., Prati, C., & Rocca, F. 2001, IEEE Transactions on Geoscience and Remote Sensing, 39, 8

Goovaerts, P. 1997, Geostatistics fo Natural Resources Evaluation, Applied Geostatistics Series (New York: Oxford University Press)

Hanssen, R. F. 2001, Radar Interferometry: Data Interpretation and Error Analysis (Dordrecht: Kluwer Academic Publishers)

Hooper, A. 2006, Persistent Scatterer Radar Interferometry for Crustal Deformation Studies and Modeling of Volcanic Deformation. PhD Thesis (Stanford University)

Kampes, B. M. 2006, Radar Interferometry: Persistent Scatterer Technique, Remote Sensing and Digital Image Processing, vol.12 (Dordrecht: Springer Verlag)

Kampes, B. M. & Hanssen, R. F. 2004, IEEE Transactions on Geoscience and Remote Sensing, 42, 2446

Ketelaar, V. B. H. 2008, Monitoring surface deformation induced by hydrocarbon production using satellite radar interferometry. Ph D. Thesis. (Delft, the Netherlands: Delft University of Technology)

van Leijen, F. J., Hanssen, R. F., Marinkovic, P. S., & Kampes, B. M. 2006, in Fourth International Workshop on ERS/Envisat SAR Interferometry, 'FRINGE05', Frascati, Italy, 28 Nov-2 Dec 2005, 6 pp.

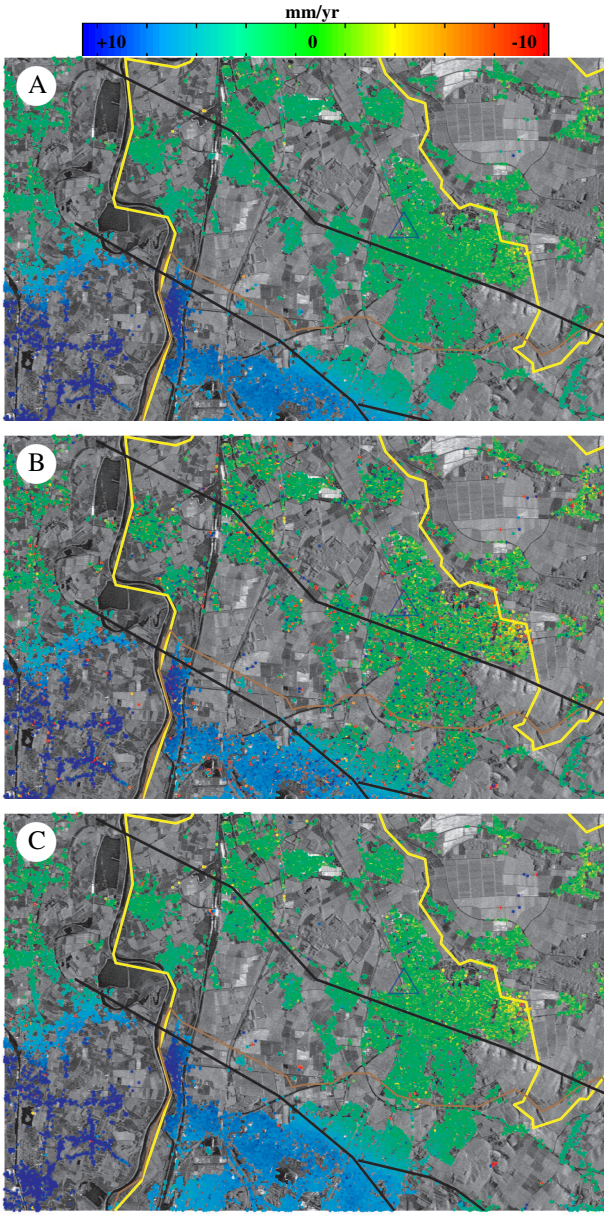

Figure 4. Estimated velocity. (A) Bayesian method, (B) bootstrapping, (C) periodogram. Country borders are displayed in yellow, in black the tectonic faults. The limits of the mine concessions are shown in brown.