

# ADVANCED INSAR PROCESSING IN THE FOOTSTEPS OF SQUEESAR

Markus Even<sup>(1)</sup>

<sup>(1)</sup> Fraunhofer IOSB, Gutleuthausstr. 1, 76275 Ettlingen, Germany, markus.even@iosb.fraunhofer.de

## ABSTRACT

Several years ago a promising approach for processing InSAR time series was introduced under the name SqueeSAR [1]. The successful application of this framework poses some delicate questions. This paper focuses on the problem that real data do rarely behave perfectly Gaussian. An augmentation of the stochastic model underlying the phase linking step is presented and the applicability under the assumption of complex elliptically symmetric distribution is discussed. Results from tests with two time series of TerraSAR-X HRS data are presented and preliminary conclusions drawn.

## 1. INTRODUCTION

At Fringe 2009 Fabrizio Novali introduced SqueeSAR, a refined approach to deformation analysis developed by Ferretti, Fumagalli, Novali himself, Prati, Rocca and Rucci [2]. The presented comparisons of results generated with SqueeSAR to such generated with PSInSAR exhibited an impressive increase of extracted information. This success is based on a unified spatio-temporal statistical analysis which can be seen to turn DS (distributed scatterers) into PS. Means of this conversion is a maximum likelihood estimator, derived and theoretically investigated by Monti Guarnieri and Tebaldini [3], that provides the complex signal common to a group of DS pixels obeying a joint circular complex normal distribution. In the framework of SqueeSAR [1], this estimator is made applicable by providing ideas for grouping DS candidate pixels in a neighbourhood of an investigated pixel, which is supposed to be statistically homogeneous and to allow for the estimation of the covariance matrix.

Despite the demonstrated capability of this approach to enhance coverage with high quality information considerably and although the use of related ideas in the SAR interferometry community increased in the last years, questions which arise naturally when one tries to put to work this approach are still not answered satisfyingly. These questions concern criteria for grouping pixels, estimators for the covariance matrix and the use of characteristic numbers, which allow recognizing low quality pixels and cutting short time consuming computations for them. Moreover good DS pixels have to be identified in order to be able to decide if they can be included in the start set of pixels on which an initial estimation is performed. This is of particular interest for rural scenes to bridge gaps in the PS net.

Because a deeper investigation is required the focus in this work is on the estimation of the covariance matrix for phase linking. In order to better understand how this should be done an extended stochastic model is stated accounting for different scale of backscattering among the grouped pixels, contamination by outliers and residual fringes. Furthermore the applicability of phase linking to some relevant classes of complex elliptically distributions other than the multivariate Gaussian is discussed. Preliminary results with several estimators for the covariance matrix and different parameter settings are shown.

For the purpose of exploring the effect of the different choices of estimator we implemented a testing environment. DS are processed in a detail of the investigated scene and then connected to a reference result of a PS analysis. Pixels are assessed according to the temporal coherence of their phase differences to nearby PS and versus ground truth. Test data are from a scene in Bavaria, where a LIDAR DEM is available and from the town of Lüneburg, Germany, where we use deformation series from levelling as ground truth. Both stacks are TerraSAR-X high resolution spotlight-mode (300MHz).

Before presenting our work I want to point out to the reader that the estimation of covariance respectively coherence matrices for InSAR time series analysis has been studied before [4], [5], [6], [7]. To our best knowledge the work of Francesco De Zan [4] (and therewith that of its PhD advisor Fabio Rocca) is the origin of the use of the covariance matrix in InSAR time series analysis. In order to extract the information of interest maximum likelihood estimation under Gaussian assumption for a low order parametric model of phase is applied. This is also the approach used in [5], [6], [7]. Our work is in part a study of the same estimators. What is different is that the usage for phase linking is clarified by giving an augmented stochastic model.

## 2. AUGMENTED STOCHASTIC MODEL

In order to clarify how phase linking can be used for data that are not well modelled by the Gaussian probability distribution, we suggest considering the following extended stochastic model for a DS. Given a neighbourhood  $\Omega$  of pixels we assume that the complex vectors of image values are realisations of random vectors  $\zeta_k$  that result by some modifications from independent identically circular complex normal

distributed random vectors  $z_k \sim CN(0, C, 0)$  ( $k \in \Omega$ ).  $C$  denotes the covariance matrix and we make the assumption that for its entries  $c_{nm} \in \mathbb{R}_{\geq 0}$  holds. The reason for this assumption will become clear later. (For the sake of simple notation we do not distinguish between the random vectors and their realisations in the following.) We allow for modification by scaling each pixel by a factor  $s_k \in \mathbb{R}_{>0}$ , by residual fringes  $\psi_{kn} \in \mathbb{R}$  and outliers  $o_{kn} \in \mathbb{C}$ . Thereby the images are enumerated by  $n = 1, \dots, N$ . The scaling factor accounts for varying backscatter from pixel to pixel. The residual fringes are thought to stem e.g. from a gradient in the deformation field, from uncompensated topography or small scale variations of atmospheric delay. We have to assume that they behave spatially smooth and that they can be filtered out by appropriate methods. The  $o_{kn}$  model outliers e.g. caused by presence of objects only in a few of the pixels and only some of the acquisitions and therefore are assumed to be zero for the larger part of  $k$  and  $n$ . As in the original model used by [3] all pixels in  $\Omega$  are assumed to have a common phase history  $\phi_n$  that accounts for deformation, atmospheric delay, large area DEM errors and other contributions that do not vary spatially. The  $\phi_n$  are the sought for unknowns during phase linking. For pixel  $k$  at time  $n$  we have

$$\zeta_{kn} = (s_k \cdot z_{kn} + o_{kn}) \cdot e^{i(\psi_{kn} + \phi_n)} \quad (1)$$

From this we can proceed in analogy to [3]. First we state the likelihood function

$$p(\zeta|\phi) = c \cdot \exp(-\tilde{\zeta}^H \cdot C(\phi, s, \psi)^{-1} \cdot \tilde{\zeta}) \quad (2)$$

for the vector  $\phi$  of phases  $\phi_n$ .  $c$  is a constant independent of  $\phi$ .  $\tilde{\zeta}$  equals  $\zeta - \mu(\phi, o, \psi)$ , where  $\zeta$  is the random vector composed of the vectors  $\zeta_k$ ,  $k=1, \dots, K$ ,  $\mu(\phi, o, \psi) = E[\zeta]$  is its expectation vector, which depends on  $\phi$ , on the vector  $o$  of all outliers  $o_{kn}$  and the vector  $\psi$  of all fringes  $\psi_{kn}$  and  $C(\phi, s, \psi)$  is the covariance matrix of  $\zeta$ , which depends on  $\phi$ , the vector  $s$  of all scaling factors  $s_k$  and  $\psi$ . After some transformations we obtain

$$\tilde{\zeta}^H \cdot C(\phi, s, \psi)^{-1} \cdot \tilde{\zeta} = \xi^H \cdot (|C|^{-1} \circ I_\Omega) \cdot \xi \quad (3)$$

$\xi$  is the vector with entries  $\xi_n = e^{i\phi_n}$ ,  $n = 1, \dots, N$  and  $I_\Omega$  is the matrix with entries

$$i_{nm} = \sum_{k \in \Omega} s_k^{-2} \cdot (e^{-i\psi_{kn}} \cdot \tilde{\zeta}_n) \cdot (e^{-i\psi_{kn}} \cdot \tilde{\zeta}_m)^H \quad (4)$$

Now we see that the assumption  $c_{nm} \in \mathbb{R}_{\geq 0}$  stated before is equivalent to assuming that the phase common to all pixels in  $\Omega$  is completely described by  $\phi$ . This augmented stochastic model provides us with the insight, how phase linking has to be applied when scale

of backscatter is varying, residual fringes are present and a limited number of outliers contaminate otherwise Gaussian distributed data: We estimate for every pixel in  $\Omega$  the scale of backscatter and normalize amplitudes accordingly. Furthermore we remove the residual fringes by some filtering procedure. Then a robust estimator for covariance matrices is applied. The result  $\hat{C}$  is an estimation of  $\# \Omega \cdot I_\Omega$  or  $C(\phi, 0, 0)$  respectively. That is the task of phase linking can be solved by minimizing

$$f(\phi) = \xi^H \cdot (|\hat{C}|^{-1} \circ \hat{C}) \cdot \xi \quad (5)$$

### 3. PHASE LINKING FOR CES DISTRIBUTIONS

Another possibility to extend the range of applicability of phase linking is to start with complex elliptically symmetric (CES) distributed random vectors  $z_k \sim CE(0, \Sigma, g)$  ( $k \in \Omega$ ).  $\Sigma$  is called scatter matrix and is equal to the covariance matrix up to a positive real constant.  $g$  is called density generator and has to obey some integrability condition [8]. CES distributions form a larger class that comprises e.g. complex normal, complex t-, complex K-, generalized Gaussian and inverse Gaussian distributions and are used to model radar clutter. More information on CES distributions and their properties can be found in [8]. For CES distributions the likelihood function for phase linking is given by

$$p(\zeta|\phi) = c \cdot g(\tilde{\zeta}^H \cdot \Sigma(\phi, s, \psi)^{-1} \cdot \tilde{\zeta}) \quad (6)$$

In case  $g$  is strictly monotonically decreasing we can proceed as in the case of complex circular normal distributions as the factor between scatter matrix and covariance matrix has no effect on the position of the minimum (cp. Eq. 3). Complex normal, complex t-, complex K-, complex generalized Gaussian and complex inverse Gaussian distributions are examples where  $g$  is strictly monotonically decreasing. Maximum likelihood estimators for the scatter matrices of CES distributions fulfil a certain implicit equation in case  $g$  is real continuously differentiable. For the distributions named above this equation can be solved uniquely by iteration (cp. [8], Theorems 6 and 7) and constitute so called M-estimators.

### 4. M-ESTIMATORS OF SCATTER

(As for Section 3 the main source for this section has been [8].) For each of the above CES distributions there exists a weight function  $\varphi$  such that the maximum likelihood estimator of scatter  $\hat{\Sigma}$  is the solution of the following equation

$$\hat{\Sigma} = \frac{1}{\#\Omega} \sum_{k \in \Omega} \varphi(\zeta_k^H \hat{\Sigma}^{-1} \zeta_k) \cdot \zeta_k \zeta_k^H \quad (7)$$

Examples are  $\varphi \equiv 1$  in the case of normal distributions or  $\varphi_{MLT}(t) = \frac{2N+\nu}{2t+\nu}$  in the case of the complex  $N$ -dimensional  $t$ -distribution with parameter  $\nu \in \mathbb{R}_{>0}$ . More generally Eq. 7 defines for appropriate  $\varphi$  an M-estimator of scatter (MES). An important property of an estimator is its robustness in the sense that its influence function is continuous and bounded. The boundedness implies that small deviations from the assumed probability distribution function do not cause arbitrary large estimation errors. If the influence function exists MES are robust iff the function  $\psi(t) = t \cdot \varphi(t)$  is continuous and bounded (cp. Theorem 8 in [8]). Furthermore the influence function exists if the function

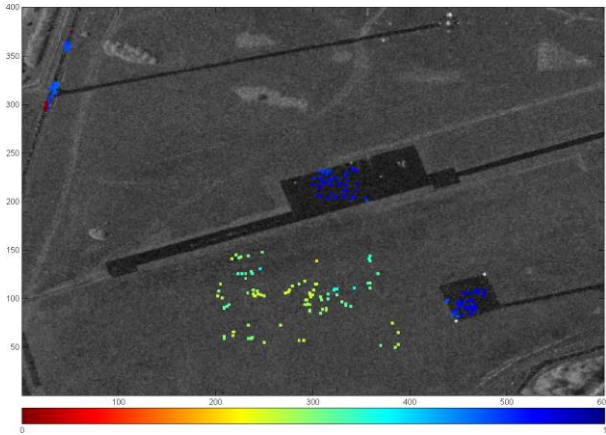


Figure 1. Pixels chosen for tests at Greding

given by  $t \cdot \psi'(t)$  is bounded.

From the named CES distributions only the  $t$ -distribution has a robust maximum likelihood estimator of scatter. As nonrobust estimators may severely fail if the distribution of the data deviates from the assumed distribution it is recommendable to look for robust alternatives. Robust estimators studied in this work are besides the  $MLT$  the Huber-MES (*Huber*) with parameter  $q$  given by

$$\varphi_{Huber}(t) = \begin{cases} \frac{1}{b(q)}, & t \leq c(q)^2 \\ \frac{c(q)^2}{t \cdot b(q)}, & t > c(q)^2 \end{cases} \quad (8)$$

and the  $S$ -estimator ( $SR$ ) with Rocke's weight function (cp. [9])

$$\varphi_{SR}(t) = \begin{cases} \frac{3}{4\gamma} \left( 1 - \left( \frac{t-1}{\gamma} \right)^2 \right), & 1 - \gamma \leq t \leq 1 + \gamma \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

For *Huber*( $q$ ) we determine  $c(q) = \frac{1}{2}(\chi_{2N}^2)^{-1}(q)$  and  $b(q) = \chi_{2(N+1)}^2(2c(q)^2) + 2c(q)^2 \frac{1-q}{N}$ , where  $\chi_m^2$  is the  $\chi^2$ -distribution with  $m$  degrees of freedom.

## 5. DATA

### 5.1. Greding (Bavaria)

This stack comprises 23 TerraSAR-X HRS (300 MHz) acquisitions. As ground truth a LIDAR-DTM of the Bavarian land surveying office (height accuracy better than 0.2 m, positional accuracy better than 0.5 m) was at our disposition. The DTM was transformed to the master raster. For tests we chose 250 pixels from a flat area where regions with different backscattering properties can be found (asphalt, plaster, concrete, meadow). Figure 1 shows the chosen pixels in front of a mean amplitude image. Their colour displays their phase triangulation coherence.

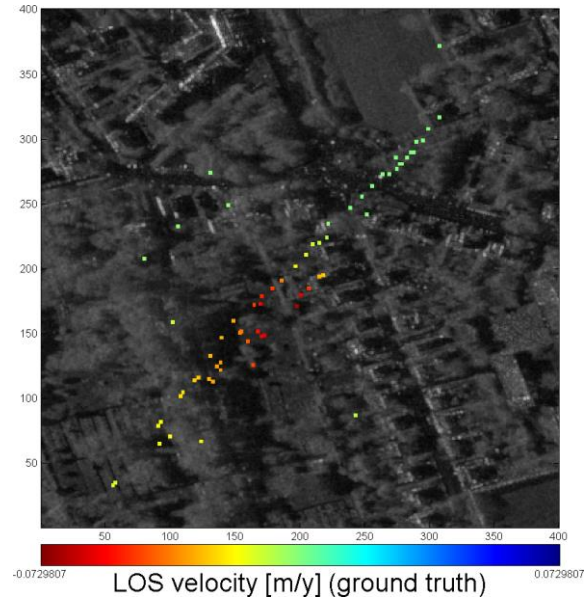


Figure 2. Pixels chosen for tests at Lüneburg

### 5.2. Lüneburg

As a second test case we used 26 TerraSAR-X HRS (300 MHz) acquisitions from the town of Lüneburg. The construction office of the town provided levelling data from which we extracted 199 measurement points that were useable for our purpose (cover of the period of interest, linear deformation). LOS velocities were computed and positions transformed to positions in the master raster. The deformation field was interpolated using natural neighbour interpolation. For the tests we restricted ourselves to an area of strong subsidence (Ochtmisser Kirchsteig), where ground truth was recorded every one or two months and also horizontal movements occurred. Because the interpolated deformation field apparently deviated considerably from the actual movement where measurement points of moderate velocity were interpolated with ones of zero velocity at a larger distance, we originally intended to use only the measurement points themselves for comparison with estimations. Unfortunately a mistake in the list of pixel coordinates elapsed our attention and

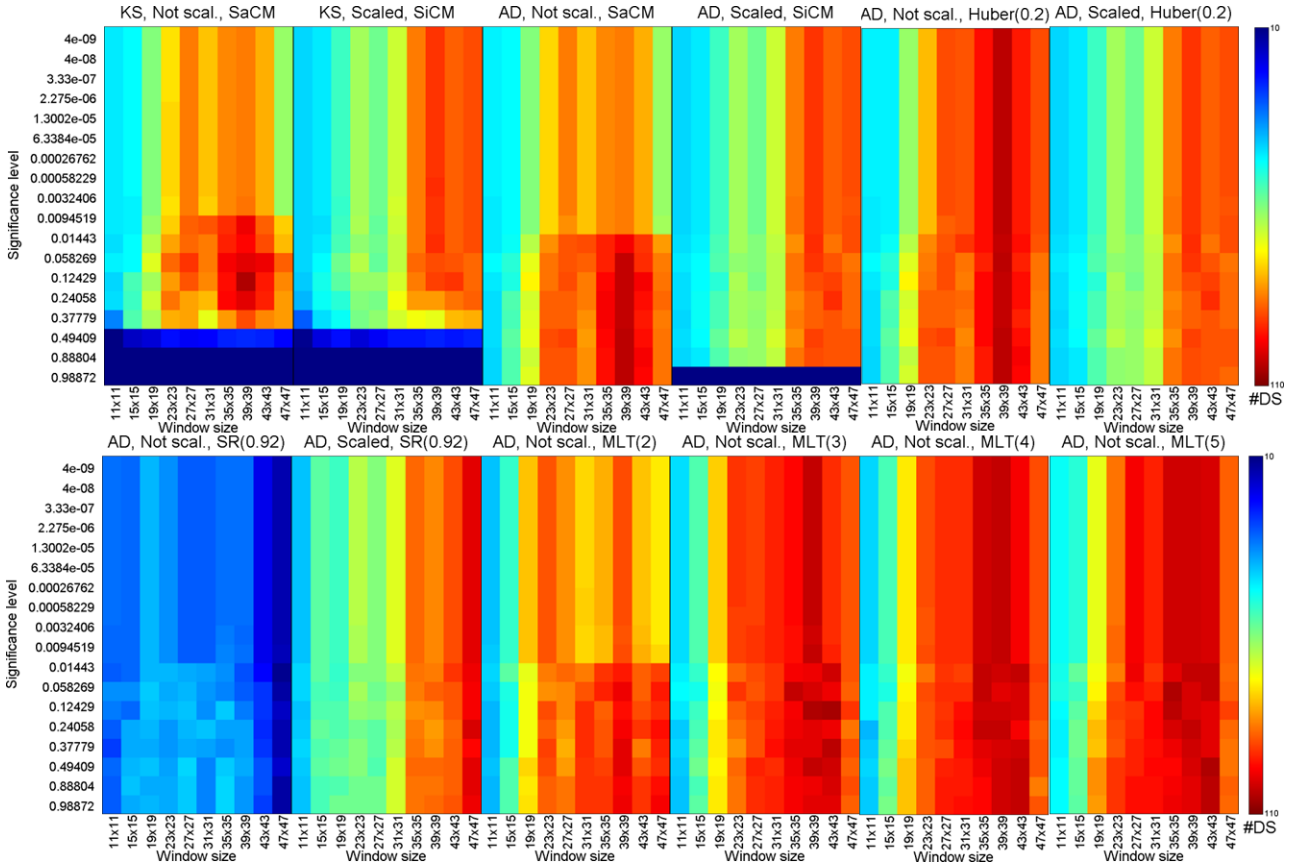


Figure 3. Testing results for Greeding

we performed tests for the pixels depicted in Figure 2. Pixel colours correspond to LOS velocity according to interpolated ground truth. The background image is the mean amplitude image over the stack.

## 6. TESTS

### 6.1. Testing environment for DS-processing

Put into one sentence DS-processing can be described as spatio-temporal filtering that turns DS into PS. DS-processing runs through several steps. First the neighbourhood  $\Omega$  of the pixel under consideration (called DS-candidate (DSC) in the following) has to be determined. Inside a search window centred on the considered pixel all other pixels are checked for similarity according to some criterion. In our implementation this is done by either the Kolmogorov-Smirnov (KS) or the Anderson-Darling (AD) two-sample test. A pixel is accepted as neighbour if the hypotheses that both pixels stem from the same distribution is not rejected given the chosen significance level. The second step is phase linking as described in Sections 2.-4. after which the phase of the DSC pixel is replaced by the estimated phase.

For testing we use a reference PS-analysis. The atmospheric phase screen estimated for the PS is extended to the position of the DSC and removed from its phase. Then the nearest PS in the vicinity of the DSC

is selected and the difference phase is evaluated. The estimated linear deformation trend or DEM error is compared to the ground truth described in Section 5. Because of the large number of test settings we restricted our investigation on the selections of pixels described in Section 5. As criterion of quality we counted the number of “good” pixels, where goodness is defined in Subsection 6.3. and 6.4. respectively.

### 6.2. Test settings

In our tests we investigated four estimators of scatter: the sample covariance matrix, which is the maximum likelihood estimator under assumption of the Gaussian distribution, *MLT*, *Huber* and *SR*. For each of them different parameter settings for the weight function, different significance levels, different sizes of search window and nonscaled as well as scaled amplitudes were studied. The tested significance levels correspond to those effective for the KS test. When writing this paper, results were not available for all possible combinations. Furthermore not all results that are available are shown.

### 6.3. Testing results for Greeding

Figure 3. shows the results for Greeding. Colour gives the number of good pixels, where high number is attained for red. The depicted blocks correspond each to

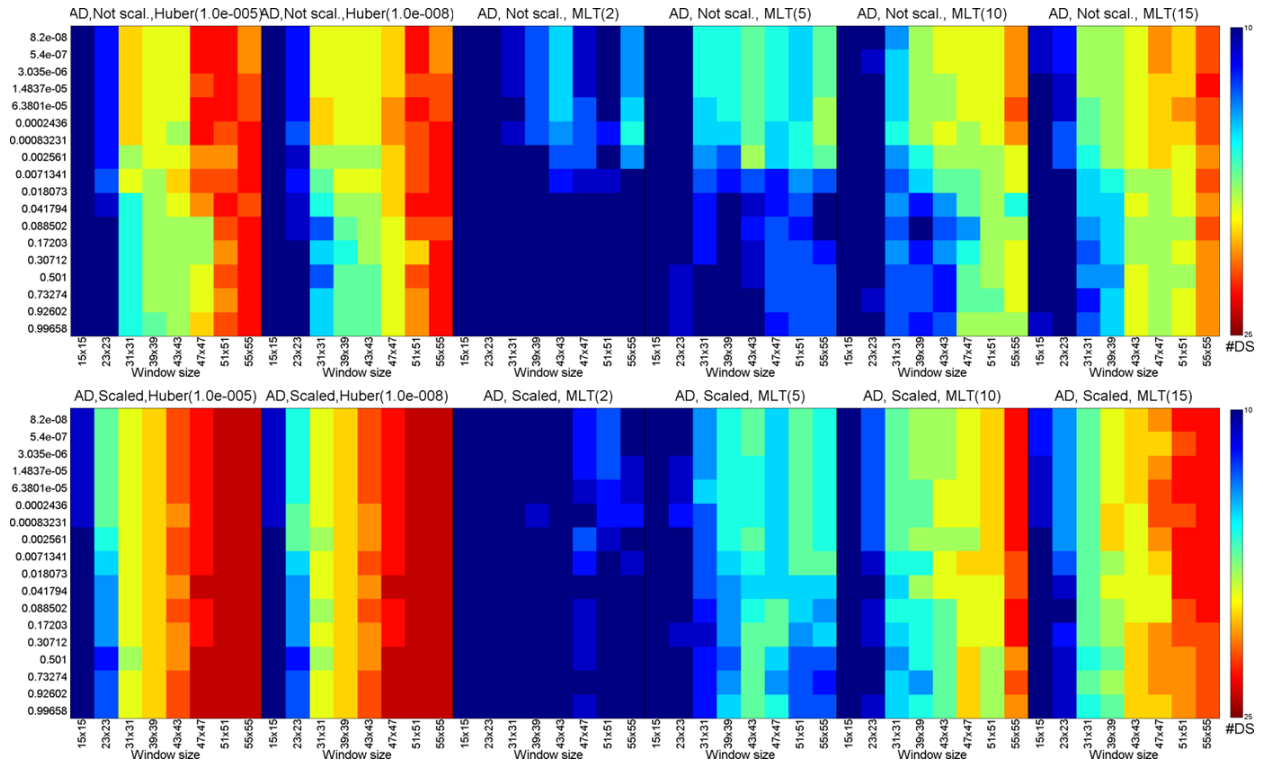


Figure 4. Testing results for Lüneburg with MLT and Huber

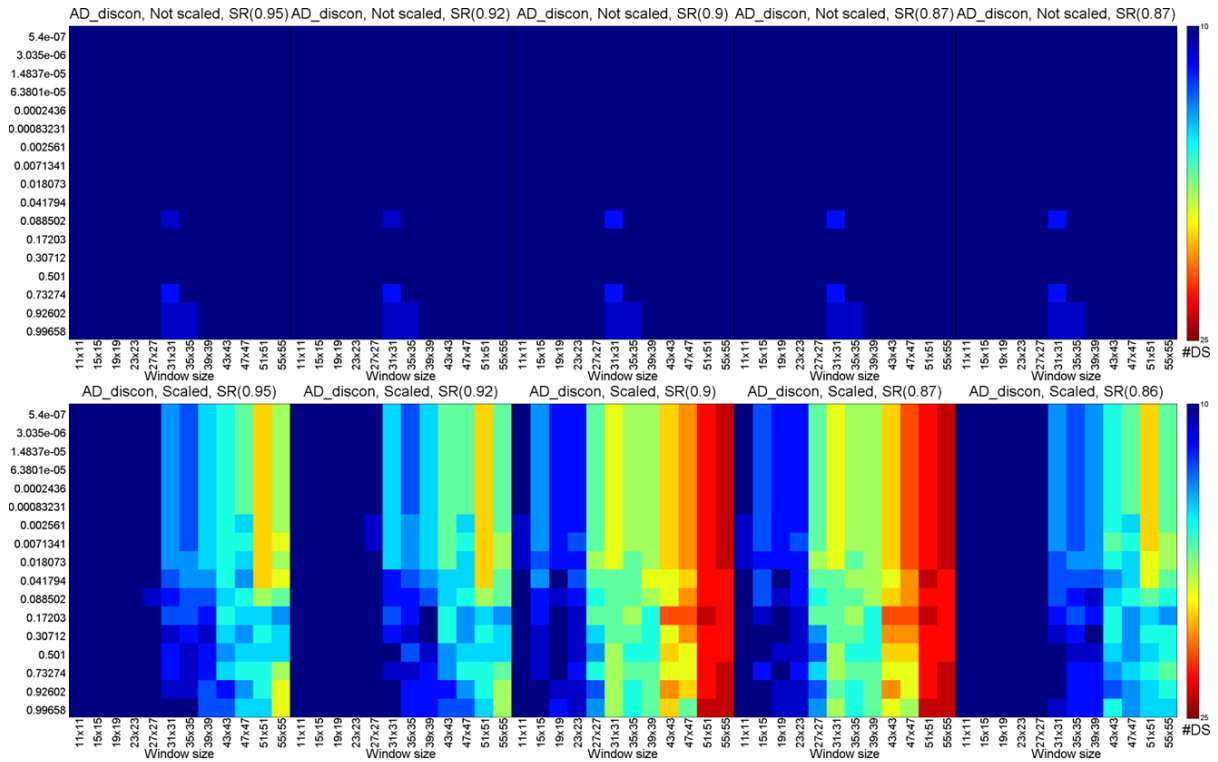


Figure 5. Testing results for Lüneburg with SR

a choice of estimator and a choice between scaling amplitudes or not scaling amplitudes. The sample covariance matrix is abbreviated either as *SaCM* or as

*SiCM*. The latter is used because for scaled amplitudes we obtain the so called sign covariance matrix. DSC pixels are considered “good” if the temporal coherence

$\gamma$  of the arc connecting it to the nearest PS is  $\geq 0.7$  and the deviation from ground truth is less than  $1m$ . Best results regarding “goodness” and insensitivity towards choice of significance level and choice of window size are achieved for the robust estimators  $MLT(4)$ ,  $MLT(5)$  and  $Huber(0.2)$  with nonscaled amplitudes. Results for  $MLT$  with scaled amplitudes were not available for Greiding yet.  $SR$  with scaled amplitudes gives similar results as  $Huber$  with scaled amplitudes. With nonscaled amplitudes  $SR$  fails completely. This is related to the fact that  $SR$  is, other than  $MLT$  and  $Huber$ , very sensitive to the matrix used as a starting point for iterations when solving Equation 7. According to [9]  $SR$  performs superior to all other estimators investigated by them. Unfortunately this aspirable performance is based on the usage of a starting matrix found by a procedure that forbids itself in our situation because of a high computing time.

#### 6.4. Testing results for Lüneburg

Results for Lüneburg are given in Figures 4. and 5. Being a “good” DSC pixel is now defined as having temporal coherence  $\gamma \geq 0.7$  for the arc connecting it to the nearest PS and having deviation from ground truth either less than  $0.002m/y$  or less than 10%. They are arranged in the same way as those for Greiding only that tests with nonscaled amplitudes are displayed in the upper row and tests with scaled amplitudes in the lower row. It is obvious that scaled amplitudes give significantly better results than nonscaled amplitudes. Best results are achieved with  $Huber(1.0e - 005)$  and  $Huber(1.0e - 008)$  with scaled amplitudes. We also tested the sample covariance matrix but again the robust estimators gave better results.

## 7. DISCUSSION AND SUMMARY

We presented an augmented stochastic model that clarifies how phase linking can be applied when the distribution of data deviates from the assumption of Gaussian distribution. Furthermore phase linking for complex elliptically symmetric distributions was discussed.

Testing results are yet incomplete and should be considered preliminary. Nevertheless they strongly indicate that scaling is more advantageous than not scaling, as the results for the stack of Lüneburg showed distinct improvements while for Greiding  $MLT$  and  $Huber(1.0e - 005)$  for scaled amplitudes, which gave the best results for Lüneburg, have not been evaluated yet. For both data stacks robust estimators  $MLT(v)$  and  $Huber(q)$  outperform the sample coherence matrix when suitably parameterized. As tests not presented here demonstrated that already a first order fringe correction as introduced in [10] improves results when a strong gradient of the deformation field is present, our proposed approach has already proven useful in all three

aspects that were integrated in our augmented model.

## 8. ACKNOWLEDGMENTS

We are grateful to the construction office of the town of Lüneburg for providing the levelling data used for our tests.

## 9. REFERENCES

1. Ferretti, A., Fumagalli, A., Novali, F., Prati, C., Rocca, F. & Rucci, A. (2011). *A New Algorithm for Processing Interferometric Data-Stacks: SqueeSAR*, IEEE TGARS, Vol. 49, No. 9, pp. 3460-3470.
2. Ferretti, A., Fumagalli, A., Novali, F., Prati, C., Rocca, F. & Rucci, A. (2009). *The second generation PSInSAR approach: SqueeSAR*, Int. Fringe Workshop, Frascati, Italy.
3. Monti Guarnieri, A. & Tebaldini, S. (2008). *On the Exploitation of Target Statistics for SAR Interferometry Applications*. IEEE TGRS, Vol 46, No.11, pp. 3436-3443.
4. De Zan, F. (2008). *Optimizing SAR interferometry for decorrelating scatterers*. PhD Thesis, Politecnico di Milano, Milan, Italy.
5. Schmitt, M., Schöneberger, J.L. & Stilla, U. (2014). *Adaptive Covariance Matrix Estimation for Multi-Baseline InSAR Data Stacks*. IEEE TGRS, Vol 52, No.11, pp. 6807-6817.
6. Schmitt, M. & Stilla, U. (2014). *Maximum Likelihood estimation for multi-aspect multi-baseline SAR interferometry of urban areas*. ISPRS Journal of Photogrammetry and Remote Sensing 87, pp. 68-77.
7. Wang, Y., Zhu, X. & Bamler, R. (2014). *Robust Covariance Matrix Estimation with Application to Volcano Monitoring using SAR image stacks*. EUSAR Proceedings (Berlin).
8. Ollila, E., Tyler, D.E., Koivunen, V. & Poor, H.V. (2012). *Complex Elliptically Symmetric Distributions: Survey, New Results and Applications*. IEEE TGRS, Vol 60, No.11, pp. 5597-5625.
9. Maronna, M.A., Martin, R.D. & Yohai, V.J., (2006). *Robust Statistics: Theory and Methods*. Wiley, New York.
10. Goel, K. & Adam, N. (2014). *A Distributed Scatterer Interferometry Approach for Precision Monitoring of Known Surface Deformation Phenomena*. IEEE TGRS, Vol 52, No.9, pp. 5454-5468.