Interferometric SAR Coherence Magnitude Estimation by Machine Learning

Nico Adam

***Abstract*—Current interferometric wide area ground motion services require the estimation of the coherence magnitude as accurately and computationally effectively as possible. However, a precise and at the same time computationally efficient method is missing. Therefore, the objective of this article is to improve the empirical Bayesian coherence magnitude estimation in terms of accuracy and computational cost. Precisely, this article proposes the interferometric coherence magnitude estimation by Machine Learning (ML).**

It results in a non-parametric and automated statistical inference. However, applying ML in this estimation context is not straightforward. The number and the domain of possible input processes is infinite and it is not possible to train all possible input signals. It is shown that the expected channel amplitudes and the expected interferometric phase cause redundancies in the input signals allowing to solve this issue. Similar to the empirical Bayesian methods, a single parameter for the maximum underlaying coherence is used to model the prior. However, no prior or any shape of prior probability is easy to implement within the ML framework.

The article reports on the bias, the standard deviation and the root mean square error (RMSE) of the developed estimators. It was found that ML estimators improve the coherence estimation RMSE from small samples ($2 \leq N < 30$) and for small underlaying coherence compared to the conventional and empirical Bayes estimators. For three interferometric samples ($N = 3$) and a zero coherence magnitude, the bias related to the sample estimator improves from 0.53 to 0.39 by $27.8\,\%$. Assuming the maximum underlaying coherence is 0.6, the bias is reduced by $33.0\,\%$ to 0.36 for the less strict and by $45.5\,\%$ to 0.29 for the strict prior.

The developed ML coherence magnitude estimators are suitable and recommended for operational InSAR systems. For the estimation, the ML model is extremely fast evaluated because no iteration, numeric integration or Bootstrapping is needed.

***Index Terms*—Coherence magnitude, degree of coherence, distributed scatterer in SqueeSAR or CESAR or phase linking, Gradient Boosted Trees, interferometric SAR (InSAR), Supervised Machine Learning**

I. Introduction

In recent years, SAR interferometry (InSAR) has developed rapidly and now allows continuous monitoring of subtle deformations of the Earth's surface with millimeter accuracy [1], [2], [3]. There is an increasing number of wide area operational services such as the European Ground Motion Service (EGMS) [4], [5], [6] and the Ground Motion Service Germany [7], [8], [9] that make the deformation maps freely available and thus widely visible. For their production, the coherence magnitude is an essential estimate, since this is the crucial weighting [10] in all estimation methods based on distributed scatterers. Due to the large amount of data and the significance, there is an actual need to estimate this parameter as accurately and computationally effectively as possible. Technically, the task is to estimate the population parameter coherence magnitude from a sample of size N. However, the challenges are the bias and variance of the estimate, which are large for small coherences and small sample sizes, and the high computational cost of using more precise methods.

This article proposes the interferometric coherence magnitude estimation by Machine Learning (ML). ML has made a lot of progress in recent years and has already found numerous applications in radar remote sensing [11], but not for the direct estimation of this parameter.

Practically, coherence magnitude estimation is a special kind of statistical inference. Conventional parametric methods are Maximum Likelihood Estimation (MLE) and Bayesian techniques. Basically, Bootstrapping is a non-parametric approach to statistical inference. And, indeed, ML can be considered to be another non-parametric and automated statistical inference.

Estimating the coherence magnitude has long been an active research topic. It stems from the fact that all state-of-the-art InSAR processing methods using distributed scatterers require precise coherence estimates [2], [3], [12], [13]. As pointed out by Zebker and Villasenor [14] as well as Just and Bamler [15] it is a proxy for the signal to noise ratio (SNR). Fundamental work on the underlying statistical models are contributed by Goodman [16] and Touzi *et al.* [17], [18] and is the basis for all conventional parametric methods. The sample estimator for the coherence magnitude is universal. Therefore, it is implemented by default in operational InSAR systems. It has been comprehensively studied by Touzi and Lopes [17] and its characteristics are well known.

Different techniques have been developed to improve the estimation accuracy. Touzi *et al.* [18] proposed the inversion of the functional relation between the first moment of the sample coherence magnitude estimate and the true coherence. For the applicability of this method, the authors state that the number of samples must be sufficiently large. Zebker and Chen [19] published a bias correction by fitting a polynomial to coherence estimates of simulated data as a function of the true correlation and the number of looks in the estimate. Another bias mitigation has been published by Abdelfattah and Nicolas [20] based on the logarithm of the sample coherence named second kind statistic. The first non-parametric approach has been published by Jiang *et al.* [21] with the Double Bootstrapping. It is computationally demanding and the double bias correction introduces extra estimation variability which can be observed by a high estimator RMSE. Recently, the empirical Bayesian method has been published [22]. For the first time, the inclusion of prior knowledge is demonstrated. The empirical Bayesian method improves the coherence magnitude estimation with respect to bias and standard deviation

measurably in a low estimation RMSE. Most improvements related to small sample sizes and low coherences which is advantageous in repeat pass InSAR.

Up to now, the coherence magnitude estimation based on Machine Learning has not been studied and published. Applying ML in this estimation context is not straightforward. The number and the domain of possible input processes is infinite and it is not possible to train all input signals. It has also not yet been shown that prior knowledge can be used to support and improve the ML prediction,

Therefore, specific objectives of the article are summarized as follows.

1) Provide the principle and methods of the ML estimation framework,
2) Demonstrate the coherence magnitude estimation by ML,
3) Demonstrate inclusion of prior information to support the estimation,
4) Characterize the estimation for small sample sizes by bias, standard deviation and root mean square error (RMSE),
5) Check if this technique is suitable in terms of performance for operational systems,
6) Compare the performance with the sample estimator but also with the empirical Bayesian methods,
7) Check whether the implementation is independent of the ML method used.

This article is organized as follows. Section II describes the methods i.e. the principle and components of the developed ML framework. Simulation results are provided for different types of prior and sample sizes in Section III. This section also includes a proof of concept using real Sentinel-1 data. In Section IV, the characteristics of these methods are discussed. Finally, Section V presents the conclusions.

II. Methods

The variable $x_{k,i} = a_{k,i}\exp(j\delta_{k,i})$ denotes the single look complex (SLC) SAR scene pixel with index $k = 1$ for the primary and $k = 2$ for the secondary scene. i is the pixel index within a statistically homogeneous area with N independent and identically distributed (i.i.d.) samples. If the scattering surface is rough with respect to the radar wavelength, the data are modeled by a stationary complex circular Gaussian (CCG) process as stated by Goodman [16] and Just and Bamler [15].

In practice, the sample coherence magnitude $\hat{\gamma}_s$ is the universal coherence estimator for CCG signals and corresponds according to Touzi *et al.* [18] to the MLE of the underlaying coherence magnitude γ

$$\hat{\gamma}_s e^{j\hat{\phi}_s} = \frac{\sum_{i=1}^{N} x_{1,i}\cdot x_{2,i}^*}{\sqrt{\sum_{i=1}^{N}|x_{1,i}|^2}\sqrt{\sum_{i=1}^{N}|x_{2,i}|^2}}. \tag{1}$$

The coherence magnitude has the domain $\{\gamma \mid 0 \le \gamma \le 1\}$.

This article intends to develop a new method for estimation of the coherence magnitude $\hat{\gamma}$ based on the random CCG processes $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$ with specific realizations $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$

$$\begin{aligned}\hat{\gamma} =& f(\boldsymbol{x_1}, \boldsymbol{x_2}) \\ =& f(x_{1,i=1}, x_{2,i=1}, \ldots, x_{1,i=N}, x_{2,i=N}).\end{aligned} \tag{2}$$

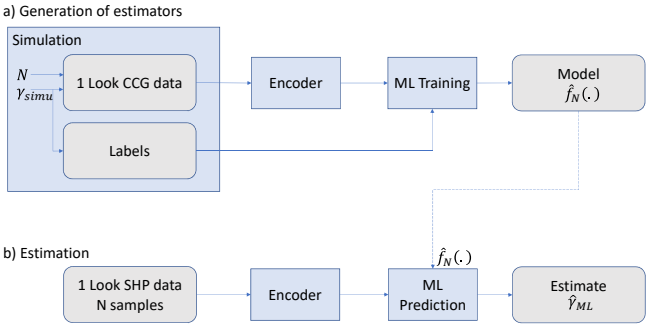


Fig. 1: Principle of the ML framework for coherence magnitude estimation.

The ML approach results in a non-parametric method. Because input samples are mapped to a continuous output value, this ML task corresponds to the regression problem in contrast to classification. Fig. 1 visualizes the basic principle and components of the development. Initially, non-parametric estimators $\hat{f}_N(x_{1,i=1}, x_{2,i=1}, \ldots, x_{1,i=N}, x_{2,i=N})$ are automatically generated from simulations and by supervised learning for any practically occurring number of samples N. In the operational system, this estimator $\hat{f}_N(.)$ is then used with N single look i.i.d. interferometric samples.

It implies to reverse the typically present oversampling and spectral weighting of the SAR data. This does not have to be done individually per statistically homogeneous area, but is better calculated only once for each SAR scene. Such preprocessing into single look complex SAR data is not a disadvantage of this technique in particular. It is also necessary for all other coherence magnitude estimation methods. In fact, all known estimators work with i.i.d. samples, where independence implies zero autocorrelation of samples within the primary and secondary channel. In case of autocorrelation, the spatial arrangement of the samples (for InSAR on the 2D grid) would have to be taken into account by the estimators. To illustrate typical effective number of looks, Sentinel-1 acquired with Interferometric Wide swath mode beam IW2 is chosen as an example. An area of 5 azimuth times 4 range samples corresponds to $N = 9$ and 6 azimuth times 7 range samples reduce to $N = 20$ independent samples.

Both system components are described in the following.

A. Generation of Estimators

For all required sample sizes N, ML provides a representative non-parametric model $\hat{f}_N(.)$. That means, there is no assumption about the function shape and the internal dependencies of the extracted features. As a result, a previously unknown number of internal parameters is required to represent the model and, accordingly, a lot of training data and computational effort are necessary for the learning. However, this does not pose a problem, since the corresponding data can be simulated in practically any quantity and the theoretically infinite number of possible variants of input data can be restricted in terms of quantity.

1) Simulation: As pointed out by Goodman [16] and Just and Bamler [15], we can limit ourselves to CCG signals for

medium resolution SAR. Starting point is the 2×2 covariance matrix $\mathbf{\Sigma}$ which describes the relation of the respective CCG processes $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$. It is defined by the simulation parameters a_1, a_2, which are the CCG processes' expected amplitudes, and the complex coherence $\rho = \gamma e^{j\phi_0}$. This term is substituted into (3) and the coherence magnitude γ_{simu} is substituted for γ and is also used as the ML label for the respective simulated data.

$$\mathbf{\Sigma} = \mathrm{cov}(\boldsymbol{X_1}, \boldsymbol{X_2}) = \begin{pmatrix} (a_1)^2 & a_1 a_2 \rho \\ a_1 a_2 \rho^* & (a_2)^2 \end{pmatrix} \tag{3}$$

The matrix above contains the expected intensities on the diagonal and the covariances on the off-diagonal. ϕ_0 is the true interferometric phase.

First of all, the square, positive definite and Hermitian covariance matrix $\mathbf{\Sigma}$ is decomposed

$$\mathbf{\Sigma} = \boldsymbol{A}\boldsymbol{A}^H. \tag{4}$$

The superscript H denotes the conjugate transpose of the complex matrix. Practically, this operation can be performed using singular value decomposition (SVD), Schur decomposition or Cholesky decomposition.

The SVD

$$\boldsymbol{U}\boldsymbol{W}\boldsymbol{V} = \mathrm{SVD}(\mathbf{\Sigma}) \tag{5}$$

results in $\boldsymbol{A} = \boldsymbol{U}\sqrt{\boldsymbol{W}}$, where $\sqrt{\boldsymbol{W}}$ denotes the element-by-element square root of the diagonal.

The Schur decomposition produces an orthonormal matrix $\boldsymbol{Q}$ and an upper triangular matrix $\boldsymbol{T}$

$$\boldsymbol{Q}\boldsymbol{T} = \mathrm{Schur}(\mathbf{\Sigma}). \tag{6}$$

This gives $\boldsymbol{A} = \boldsymbol{Q}\sqrt{\boldsymbol{T}}$, where $\sqrt{\boldsymbol{T}}$ is the element-wise square root of all matrix entries in the complex domain.

With the Cholesky decomposition

$$\boldsymbol{L}\boldsymbol{L}^H = \mathrm{Cholesky}_L(\mathbf{\Sigma}), \tag{7}$$

the product of a lower triangular $\boldsymbol{L}$ and its conjugate transpose matrix $\boldsymbol{L}^H$ arises. In this case $\boldsymbol{A} = \boldsymbol{L}$. In case, the library provides the upper triangular matrix $\boldsymbol{U}$ such that

$$\boldsymbol{U}^H\boldsymbol{U} = \mathrm{Cholesky}_U(\mathbf{\Sigma}) \tag{8}$$

then $\boldsymbol{A} = \boldsymbol{U}^H$.

Next, a complex matrix $\boldsymbol{Z} \in \mathbb{C}^{2\times N}$ of independent CCG random variables

$$\begin{aligned} z_{k,i} &= \mathfrak{Re}(z_{k,i}) + j\mathfrak{Im}(z_{k,i}) \\ \mathfrak{Re}(z_{k,i}) &\sim \mathcal{N}(0, 1/\sqrt{2}) \\ \mathfrak{Im}(z_{k,i}) &\sim \mathcal{N}(0, 1/\sqrt{2}) \end{aligned} \tag{9}$$

is created. $\mathcal{N}(0, 1/\sqrt{2})$ denotes the normal distribution with zero mean and standard deviation $1/\sqrt{2}$. The simulated interferometric data pair corresponds to the complex matrix $\boldsymbol{S} \in \mathbb{C}^{2\times N}$ calculated by

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{x_1}^T \\ \boldsymbol{x_2}^T \end{pmatrix} = \begin{pmatrix} — & x_{1,i} & — \\ — & x_{2,i} & — \end{pmatrix} = \boldsymbol{A}\boldsymbol{Z}. \tag{10}$$

This principle of transforming the covariance matrix $\mathbf{\Sigma}$ into an interferometric data pair $\boldsymbol{S}$ can be applied to the simulation of InSAR data stacks. The dimensions of the covariance and the CCG matrix $\boldsymbol{Z}$ need to be increased accordingly. For the described framework, it is not necessary and interferograms are simulated from independent 2×2 covariance matrices. All three decompositions were implemented and finally the SVD was used for the article.

2) Encoder: The encoder transforms the input data and has two preprocessing functions: a) reduce redundancies and b) convert the input into an advantageous data representation.

Ideally, the signal entering the ML training includes all appropriate features and recognizable patterns, and is a data representation without ambiguity or redundancy. This makes the ML training algorithm more precise and computationally efficient, and requires less computer memory.

a) Redundancy reduction: From the original simulation input parameters (3) $\{a_1 \mid 0 < a_1 < \infty\}$, $\{a_2 \mid 0 < a_2 < \infty\}$, $\{\gamma \mid 0 \leq \gamma \leq 1\}$ and $\{\phi_0 \mid -\pi < \phi_0 \leq \pi\}$ with their unrestricted domains and the unlimited possible combinations, we can see that there are an infinite number of possible input data sets. Generating and training all this is not realistic. As will be shown shortly, the expected channel amplitudes and the expected interferometric phase cause redundancies in the signal representation. Indeed the correlation coefficient $\rho_{\boldsymbol{X_1},\boldsymbol{X_2}}$ is independent of change of origin e.g. by real numbers b and d and scale of the data e.g. by real numbers a and c

$$\rho_{X_1,X_2} = \mathrm{corr}(\boldsymbol{X_1}, \boldsymbol{X_2}) = \mathrm{corr}(a\boldsymbol{X_1} + b, c\boldsymbol{X_2} + d). \tag{11}$$

This means that by scaling the amplitudes

$$x_{1,i} = \frac{x_{1,i}}{\max(|\boldsymbol{x_1}|)}\,,\, x_{2,i} = \frac{x_{2,i}}{\max(|\boldsymbol{x_2}|)} \text{ for } i = 1, \ldots, N \tag{12}$$

the data are restricted without loss of information to a domain $\{a_1 \mid 0 < a_1 \leq 1\}$, $\{a_2 \mid 0 < a_2 \leq 1\}$ known to the ML model.

Equation (1) shows that the coherence magnitude γ and the interferometric phase ϕ_0, which is optimally estimated by the sample estimator $\hat{\phi}_s$, are independent of each other. Hence, assuming a stationary phase signal i.e. residual topography, deformation and atmospheric phase screen are compensated, the expected interferometric phase ϕ_0 can be estimated from the statistically homogeneous pixels (SHPs)

$$\begin{aligned} \phi_0 &= \arg(\mathrm{E}\{\boldsymbol{X_1} \cdot \boldsymbol{X_2}^*\}) \\ &\approx \hat{\phi_s} = \arg(\sum_{i=1}^{N} x_{1,i} \cdot x_{2,i}^*). \end{aligned} \tag{13}$$

Since only the interferometric phase difference between each i.i.d. sample is used, the expected value can be compensated in the primary scene in advance

$$x_{1,i} = x_{1,i} \cdot \exp(-j\hat{\phi_s}). \tag{14}$$

This transformation eliminates the phase ambiguities $\delta_{k,i} + K2\pi$ and preserves the respective amplitude's Rayleigh PDF of the primary and secondary scene and the statistics of the interferometric phase differences (except for the mean).

As a result of the amplitude scaling and interferometric phase compensation, the number of possible input data has now been significantly reduced.

b) Data representation: The encoder converts the CCG input data because the data representation has an impact on the performance of the model. This is due to the fact that there is practically no direct regression from the input variables to the output value. Inside the ML model, attributes that are not visible from the outside are calculated. This internal automatic generation of features is manually supported by the encoding. Two examples for possible CCG data representations are $\{\mathfrak{Re}(\boldsymbol{x_1}), \mathfrak{Im}(\boldsymbol{x_1}), \mathfrak{Re}(\boldsymbol{x_2}), \mathfrak{Im}(\boldsymbol{x_2})\}$ and $\{|\boldsymbol{x_1}|, |\boldsymbol{x_2}|, \arg(\boldsymbol{x_1}\,\boldsymbol{x_2}^*)\}$. Tests have shown that the latter data representation, consisting of the sample amplitudes and expected interferometric phase compensated phase differences, is more advantageous than others. It is apparent that the data dimension is reduced and irrelevant information is removed. The data fully represent all required features and the ML methods can use them directly.

3) ML Training: The ML training learns the features with which the internal model is evaluated to return the coherence magnitude estimate. According to James *et al.* [23], the general form of ML regression is

$$\hat{\gamma} = f(.) + \epsilon. \tag{15}$$

ϵ is an inherent random error term and is named the irreducible error. In this application, it results from the random sampling and the limited sample size N but not from the noise in the data. Practically, every sample is differently representative and ϵ corresponds to unmeasured information which results in bias and variance of the coherence magnitude estimate. Touzi *et al.* [18] have proven that an unbiased estimator, which is a function of the sample coherence magnitude, cannot be found. It follows that the ML estimator will also have a bias and a variance. In other words, ϵ is independent of the input data $\boldsymbol{x_1}$, $\boldsymbol{x_2}$ and can only be mitigated by increasing the sample size N reducing unmeasured information.

ML provides procedures for estimating $f(.)$ based on training data and approximately represents it by $\hat{f}(.)$. Depending on the ML method, $\hat{f}(.)$ is represented differently, such as a Decision Tree, a Random Forest or a Neural Network. According to James *et al.* [23], the error from the approximation $\hat{f}(.)$ of a particular ML method is termed reducible error. It can be diminished by choosing an appropriate ML method and, if used, suitable Neural Network layers as well as optimizing the learning parameters such as the learning rate and the learning iteration count. In this article, Gradient Boosted Trees ML is implemented based on the XGBoost library with its C-API developed by Chen and Guestrin [24], [25].

All possible CCG input processes must be simulated for the ML training. To get as close to the real estimation scenario as possible, the amplitudes of the primary and secondary signals and the interferometric phases are modeled in such a way that the encoder works as it will later. In this article, the scenes' expected amplitudes are simulated with uniform likelihoods $a_1 \sim U(0, 2)$ and $a_2 \sim U(0, 2)$, and the expected interferometric phase with $\phi_0 \sim U(-\pi, \pi)$. For the training of an estimator, 10^8 independent interferograms are generated. In the course of the ML learning, the parameters of the model are tuned to perform best on the given training data. This suggests to add prior knowledge on the underlaying coherence magnitude by adjusting the training data set. In doing so, the fact that ML learns the model from the data is exploited. Training data are generated with a number of observations corresponding to the prior on the underlaying coherence. The assumption is that the ML parameter tuning then works better for these observed values than with the data, who has not or rarely seen the training. A single parameter γ_{max} is used to model the prior. In the following, this parameter is specified as a subscript at the respective method.

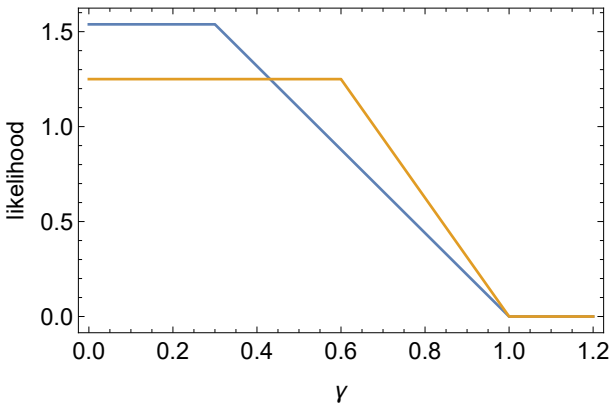


Fig. 2: Distribution of γ_{simu} for the less strict prior; blue graph: $\gamma_{max} = 0.3$; orange graph: $\gamma_{max} = 0.6$.

a) ML without prior (MLWP): Without prior information, training data are generated with the straight forward characteristic and γ_{simu} is sampled from the uniform distribution $\gamma_{simu} \sim U(0, 1)$. With 10^8 simulated interferograms, about 10^6 samples are generated in an interval $\gamma_{simu} \pm 0.005$.

b) ML less strict prior (MLLSP): Figure 2 shows the distribution of γ_{simu} for the less strict prior. The implementation is based on the inverse cumulative distribution function (CDF) sampling method. It provides one random variate $\gamma_{simu} \sim P_{prior}(\gamma_{max})$ from one random sample with distribution $u \sim U(0, 1)$. The corresponding CDF is

$$\mathrm{CDF} = \begin{cases} \frac{2\gamma}{\gamma_{max}+1} & 0 \le \gamma \le \gamma_{max} \\ \frac{(\gamma-2)\gamma+\gamma_{max}^2}{\gamma_{max}^2-1} & \gamma_{max} < \gamma \le 1. \end{cases} \tag{16}$$

This leads to the respective inverse CDF

$$\mathrm{CDF}^{-1} = \begin{cases} \frac{1}{2}(\gamma_{max}u + u) & u \le \frac{2\gamma_{max}}{\gamma_{max}+1} \\ 1 - \sqrt{\gamma_{max}^2 u - \gamma_{max}^2 - u + 1} & u > \frac{2\gamma_{max}}{\gamma_{max}+1}. \end{cases} \tag{17}$$

c) ML strict prior (MLSP): Figure 3 visualizes the distribution of the underlaying coherence magnitude γ for the strict prior. Consequently, the respective training data are generated with γ_{simu} sampled from the uniform distribution $\gamma_{simu} \sim U(0, \gamma_{max})$.

Practically, one ML model $\hat{f}_{N,p}(.)$ is generated for each prior type $\{p \mid MLWP, MLLSP_{\gamma_{max}}, MLSP_{\gamma_{max}}\}$ with every needed prior parameter $\{\gamma_{max} \mid 0.1, 0.2, ..., 0.9\}$. The utilized library XGBoost allows to persistently save each model into a JSON file [24], [25] for later operational estimation use.

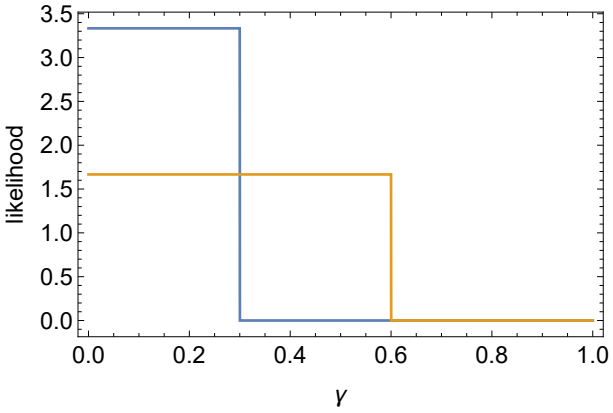


Fig. 3: Distribution of γ_{simu} for the strict prior; blue graph: $\gamma_{max} = 0.3$; orange graph: $\gamma_{max} = 0.6$.

B. Estimation of Coherence Magnitude

N interferometric samples are input to the operational coherence magnitude estimation. These data are transformed according to (12) and (14). I.e., $3 \times N$ real values, encoded by $\{\frac{|\boldsymbol{x_1}|}{\max(|\boldsymbol{x_1}|)}, \frac{|\boldsymbol{x_2}|}{\max(|\boldsymbol{x_2}|)}, \arg(\boldsymbol{x_1}\,\boldsymbol{x_2}^*\exp(-j\hat{\phi}_s)\}$, enter the ML prediction model $\hat{f}_{N,p}(.)$. Once again, all phases are the expected interferometric phase compensated phase differences. The model is extremely fast evaluated because no iteration, numeric integration or Bootstrapping is needed. The estimated coherence magnitude $\hat{\gamma}_p$ with $\{p \mid MLWP, MLLSP_{\gamma_{max}}, MLSP_{\gamma_{max}}\}$ is deterministic i.e. one and the same input data result in one and the same estimate.

III. Results

In this section the estimation characteristics obtained with Gradient Boosted Trees ML implemented using the XGBoost library [24], [25] are presented. Based on the fact that the estimation from a small sample size is the critical problem, priority is put on such test cases i.e. $N = 2$, $N = 3$ and $N = 9$.

The results show, the intuitively introduced Bayesian principle works. Any likelihood of prior can be implemented. In contrast to the empirical Bayesian approach [22], no insoluble integral has to be solved and replaced by computationally ineffective numerical integration.

In the following, the characteristics of the estimators are compared with each other and the universally applicable sample estimator (1) is taken as the reference. Generally, the bias $\gamma_{bias} = \mathrm{E}\{\hat{\gamma}_* - \gamma_{true}\}$, the standard deviation $\gamma_\sigma = \sqrt{\mathrm{E}\{(\hat{\gamma}_* - \mathrm{E}\{\gamma_*\})^2\}}$ and the RMSE $\gamma_{rmse} = \sqrt{\mathrm{E}\{(\hat{\gamma}_* - \gamma_{true})^2\}}$ for $\{\hat{\gamma}_* | \hat{\gamma}_s, \hat{\gamma}_{MLWP}, \hat{\gamma}_{MLLSP}, \hat{\gamma}_{MLSP}\}$ are relevant quality criteria of estimators. For the coherence magnitude estimation, these properties are functions of the underlaying true coherence magnitude γ_{true}. This is the reason, the parameters above are estimated 101 times for each plot with $\{\gamma_{true} \mid 0, 0.01, ..., 1\}$. 10^6 simulations are performed for each data point γ_{true} and each method is applied on the one and the same data set per analysis. For the test cases shown below, the prior parameter $\gamma_{max} = 0.6$ is chosen because it is a typical value in SAR interferometry. In the plots below, the MLSP curves end at an underlaying coherence of 0.6. It is apparent, a strict prior assumes zero probability outside of this range. However, it should be noted that the $\mathrm{MLSP}_{0.6}$ estimator provides also estimates outside of this strict range.

A. Test Case $N = 2$ Samples

The bias compared in Fig. 4a is reduced for small coherences by all ML methods. For a zero coherence, and compared to the sample estimator $\hat{\gamma}_s$, the MLWP reduces the bias from 0.6664 to 0.4374 i.e. by $34.4\,\%$, the $\mathrm{MLLSP}_{0.6}$ reduces the bias to 0.3860 i.e. by $42.1\,\%$, and the $\mathrm{MLSP}_{0.6}$ reduces the bias to 0.2967 i.e. by $55.5\,\%$. For the sample estimator, the bias becomes zero at an underlaying coherence of one. Not surprisingly, all newly developed ML estimators, are bias free at much smaller coherences. However, this is achieved at the expense of a larger bias for higher underlaying coherence magnitude values.

The standard deviation γ_σ is visualized in Fig. 4b. Again, the zero coherence is taken as an example. Compared to the sample estimator the MLWP reduces the standard deviation from 0.2358 to 0.1323 i.e. by $43.9\,\%$, the $\mathrm{MLLSP}_{0.6}$ reduces it to 0.0954 i.e. by $59.5\,\%$, and the $\mathrm{MLSP}_{0.6}$ reduces the standard deviation to 0.0545 i.e. by $76.9\,\%$.

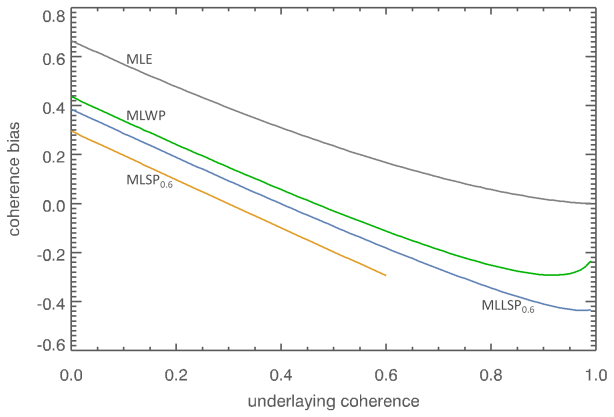
The RMSE best describes the estimator performance as it includes the bias and the variance of the estimators $\gamma_{rmse} = \sqrt{\gamma_{bias}^2 + \gamma_\sigma^2}$. The comparison of the RMSE in Fig. 4c confirms the observation from the empirical Bayesian coherence magnitude estimation [22] that the more information is used and the stricter the general prior, the more accurate the estimate will be. Compared to the conventional sample estimator, MLWP is more efficient for all underlaying coherence magnitudes up to 0.68, the $\mathrm{MLLSP}_{0.6}$ method up to 0.65 and the $\mathrm{MLSP}_{0.6}$ estimator up to 0.58.

B. Test Case N=3 Samples

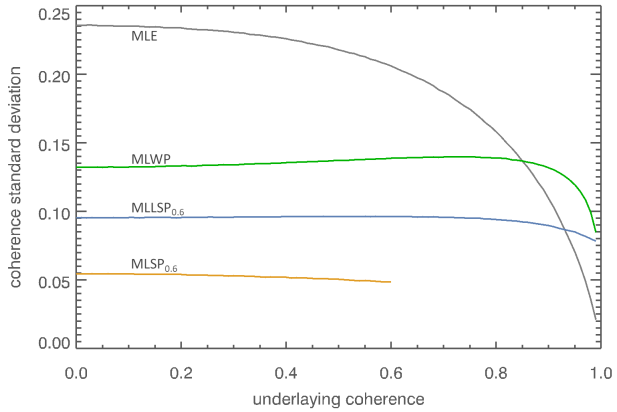
The properties and principles from the test case $N = 2$ are also confirmed in this configuration. The comparison of the bias (γ_{bias}) is visualized in Fig. 5a, of the standard deviation (γ_σ) in Fig. 5b and of the RMSE (γ_{rmse}) in Fig. 5c. For a zero coherence magnitude, the bias related to the sample estimator improves from 0.5333 to 0.3852 by $27.8\,\%$ for MLWP, to 0.3572 by $33.0\,\%$ for $\mathrm{MLLSP}_{0.6}$ and to 0.2905 by $45.5\,\%$ for $\mathrm{MLSP}_{0.6}$. The standard deviation is reduced from 0.2211 to 0.1528 i.e. by $30.9\,\%$ using MLWP, to 0.1285 i.e. by $41.9\,\%$ with $\mathrm{MLLSP}_{0.6}$ and to 0.08116 i.e. $63.3\,\%$ with $\mathrm{MLSP}_{0.6}$. Compared to the conventional sample estimator, MLWP is more efficient for all underlaying coherence magnitudes up to 0.62, the $\mathrm{MLLSP}_{0.6}$ method up to 0.61 and the $\mathrm{MLSP}_{0.6}$ estimator up to 0.55.

C. Test Case N=9 Samples

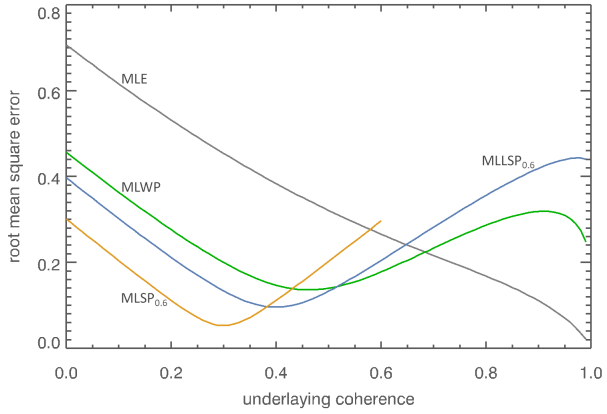
As the performance of the sample estimator improves with the number of samples, it can be expected that advantages are reduced for other methods. The visualizations of the bias in Fig. 6a, of the standard deviation in Fig. 6b and of the RMSE



(a)

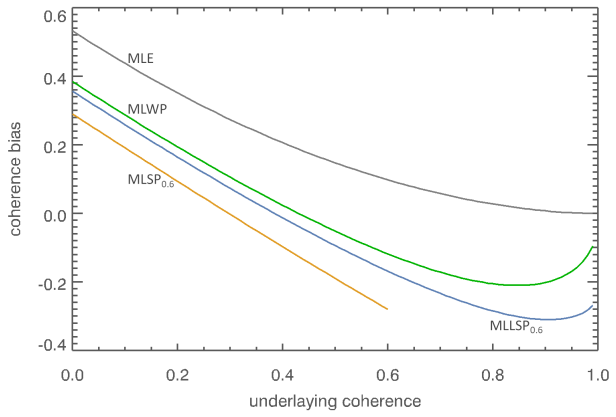


(b)

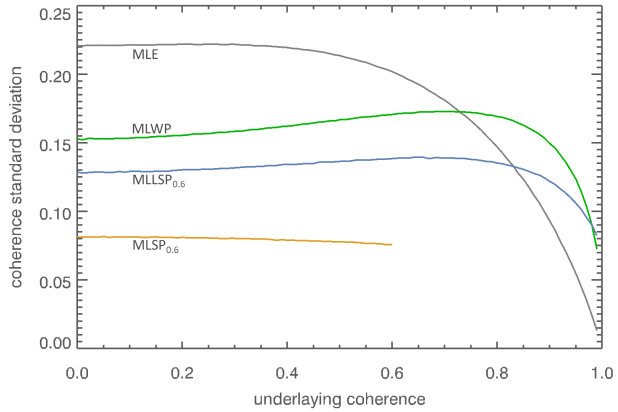


(c)

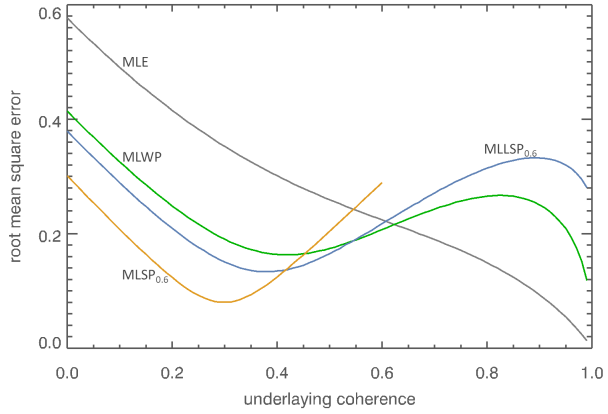
Fig. 4: Characteristic of estimators for $N = 2$ samples; gray: sample estimator (MLE), green: MLWP, blue: MLLSP$_{0.6}$ and orange: MLSP$_{0.6}$. (a) Estimation bias γ_{bias}. (b) Standard deviation γ_{σ}. (c) Root mean square error γ_{rmse}.



(a)



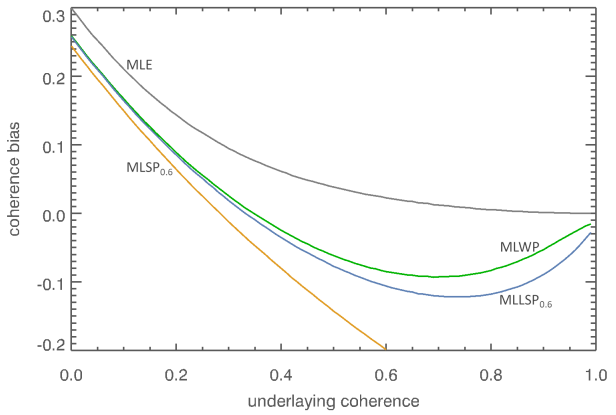
(b)



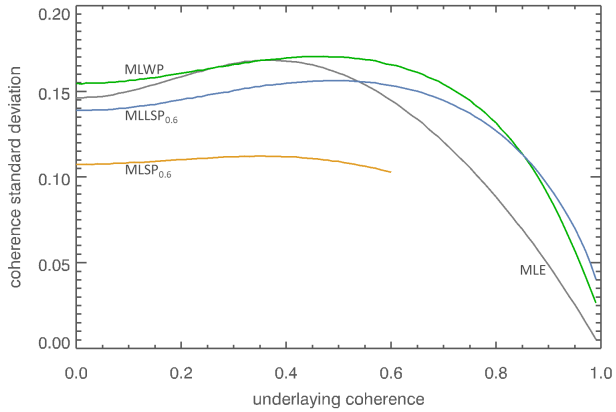
(c)

Fig. 5: Characteristic of estimators for $N = 3$ samples; gray: sample estimator (MLE), green: MLWP, blue: MLLSP$_{0.6}$ and orange: MLSP$_{0.6}$. (a) Estimation bias γ_{bias}. (b) Standard deviation γ_{σ}. (c) Root mean square error γ_{rmse}.

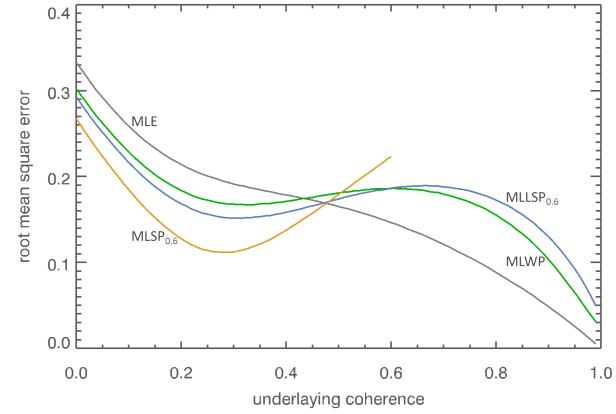
in Fig. 6c confirm this expectation. Accordingly, the reduction in bias is less pronounced. At zero coherence magnitude, the sample estimator has a bias of 0.30. The MLWP reduces the estimation bias by $13.4\,\%$ to 0.26. Also, the prior has less effect on the bias mitigation compared to test cases with fewer samples. The MLLSP$_{0.6}$ improves the bias by $14.0\,\%$ to 0.2576 and the MLSP$_{0.6}$ by $18.4\,\%$ to 0.2444. A similar characteristic is observed for the standard deviation.



(a)



(b)



(c)

Fig. 6: Characteristic of estimators for $N = 9$ samples; gray: sample estimator (MLE), green: MLWP, blue: MLLSP$_{0.6}$ and orange: MLSP$_{0.6}$. (a) Estimation bias γ_{bias}. (b) Standard deviation γ_σ. (c) Root mean square error γ_{rmse}.

For zero coherence, the standard deviation of MLWP even increases by $5.9\,\%$ from 0.1463 to 0.1548. Some prior helps to mitigate the random variation. The MLLSP$_{0.6}$ lessens the standard deviation by $4.9\,\%$ to 0.1390, and the MLSP$_{0.6}$ by $26.5\,\%$ to 0.1075. Nevertheless, the ML algorithms outperform the sample estimator for small coherence magnitude values. The MLWP is more efficient for all underlaying coherence magnitudes up to 0.43, the MLLSP$_{0.6}$ method up to 0.47 and the MLSP$_{0.6}$ estimator up to 0.48.

D. Sentinel-1 Application Demonstration

As a proof of concept, the estimator prototype is demonstrated using real Sentinel-1 data in Interferometric Wide swath mode. The primary scene has the orbit number 30741 and was acquired on January 10, 2020. After 12 days, the secondary scene was recorded. Their orbit number is 30916 and the observation geometry is characterized by an effective baseline of about $27\,\mathrm{m}$. Without going into details, the oversampling in the input data is reversed and the estimation window of 3×3 samples in range and azimuth (i.e. $N = 9$) overlaps from sample to sample.

Figure 7a visualizes the test case with 512×512 i.i.d. samples by the radar backscatter amplitude. The coherence magnitude from the sample estimator is visualized in Fig. 7b. Using identical estimation windows, the respective ML result is shown in Fig. 7d. In this example, the ML coherence magnitude is estimated locally adaptive with respect to the prior from MLWP, MLLSP$_{0.6}$, MLLSP$_{0.4}$ or MLSP$_{0.4}$. It can be seen that the estimation performance now depends not only on the window size but mainly on the prior and its strictness.

To give an intuitive idea of the effect of different priors and various parameters, Fig. 7c visualizes a composition of coherence estimates. In this figure from left to right, the result from the sample estimator, MLWP, MLLSP$_{0.6}$, MLLSP$_{0.4}$, MLSP$_{0.6}$, and MLSP$_{0.4}$ can be compared. Similar coherence magnitudes are observed for all but the last two columns. It follows that the less strict prior can robustly cope with an underlying coherence greater than the prior parameter γ_{max}.

IV. Discussion

In the section above, test cases using Gradient Boosted Trees are presented. The question arises whether other ML methods provide similar results. This is the reason, additional prototypes for the coherence magnitude estimation based on Neural Networks and Random Forests are assessed. By similar graphs, Fig. 8 demonstrates for the test case with $N = 3$ samples that the developed framework is robust with respect to a particular ML method. However, in the course of development it turned out that the encoder has a significant influence on the estimation performance. To illustrate this, Fig. 9 compares the RMSE for $N = 3$ samples with an unfavorable encoding $\{\Re\mathrm{e}(\boldsymbol{x_1}), \Im\mathrm{m}(\boldsymbol{x_1}), \Re\mathrm{e}(\boldsymbol{x_2}), \Im\mathrm{m}(\boldsymbol{x_2})\}$ between Gradient Boosted Trees ML, Neural Network ML and Random Forrest ML. The comparison with Fig. 8 shows that the ML methods have different robustness with respect to the encoding of the input data. For this application, Neural Networks are able to handle complex input data without encoding.

Unexpectedly, Fig. 5c shows a similarity of the ML methods in the RMSE characteristics with the empirical Bayesian methods [22, Fig. 8 (f)]. Some comparative plots will show which method performs better in terms of RMSE. The test

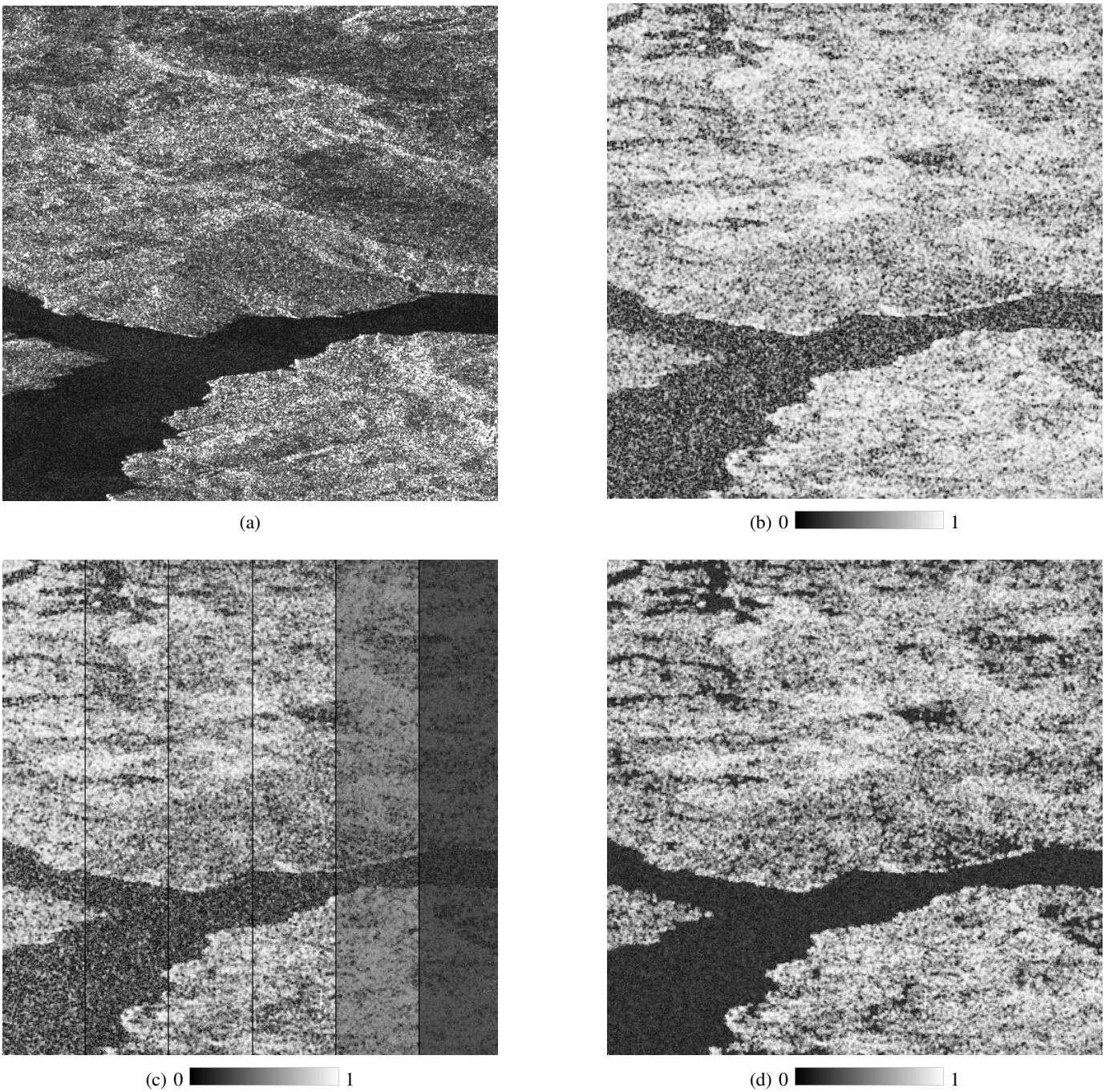


Fig. 7: Sentinel-1 test case with $N = 9$. (a) Radar backscatter amplitude. (b) Coherence magnitude from sample estimator (1). (c) Composite of coherence estimates from sample estimator, MLWP, $\text{MLLSP}_{0.6}$, $\text{MLLSP}_{0.4}$, $\text{MLSP}_{0.6}$, $\text{MLSP}_{0.4}$ from left to right. (d) ML coherence magnitude estimated locally adaptive from MLWP, $\text{MLLSP}_{0.6}$, $\text{MLLSP}_{0.4}$ or $\text{MLSP}_{0.4}$.

case with $N = 3$ serves as a demonstration. First, Fig. 10 compares the Bayesian and the ML estimation without prior. Second, Fig. 11 benchmarks the Bayesian and the ML estimation with less strict prior $\gamma_{max} = 0.6$, and third, Fig. 12 visualizes the Bayesian and the ML estimation with strict prior $\gamma_{max} = 0.6$. The examples demonstrate, ML improves the coherence magnitude estimation compared to the empirical Bayesian estimation for small coherences and small samples. In the field of SAR interferometry, the precise estimation of small coherence magnitudes is the challenge. This is one of the reasons, the newly developed ML methods are recommended for operational systems. Another is the computational efficiency compared to the empirical Bayesian estimation. On a laptop, the ML prototypes perform 10^5 estimates in less than 10 seconds. In contrast, the empirical Bayesian methods need more than half an hour for the same number of estimates. The computing performance is achieved at the expense of a high training effort. Each estimator took three days to train on a small laptop.

Above, the estimation characteristics are demonstrated for

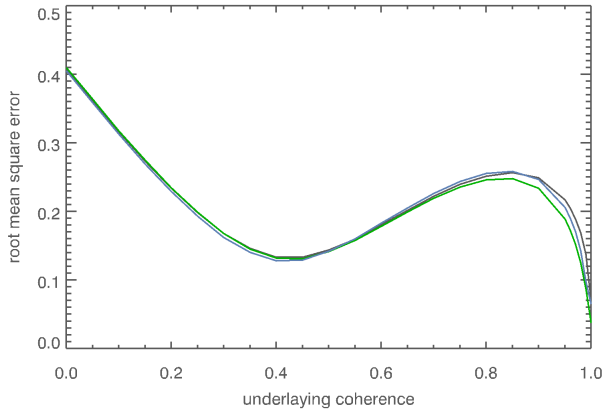


Fig. 8: Comparison of RMSE for $N = 3$ samples between Gradient Boosted Trees ML (gray), Neural Network ML (green), Random Forrest ML (blue).

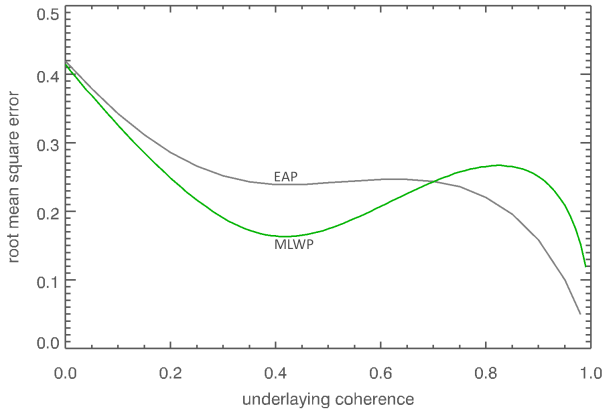


Fig. 10: Comparison of RMSE between estimation by expected *a posteriori* (EAP) in gray and MLWP in green i.e. both without prior for $N = 3$ samples.

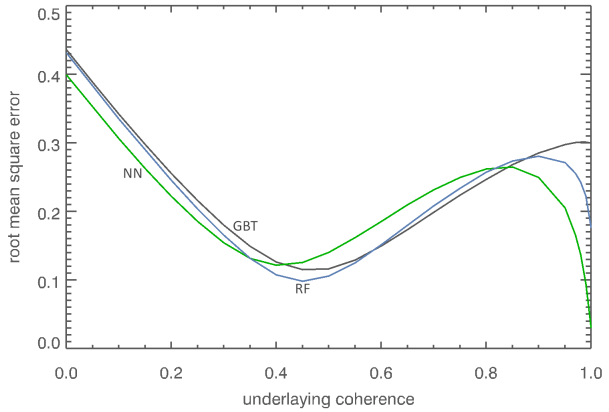


Fig. 9: Comparison of RMSE for $N = 3$ samples with unfavorable encoding $\{\mathfrak{Re}(\boldsymbol{x_1}), \mathfrak{Im}(\boldsymbol{x_1}), \mathfrak{Re}(\boldsymbol{x_2}), \mathfrak{Im}(\boldsymbol{x_2})\}$ between Gradient Boosted Trees (GBT) ML in gray, Neural Network (NN) ML in green and Random Forrest (RF) ML in blue.

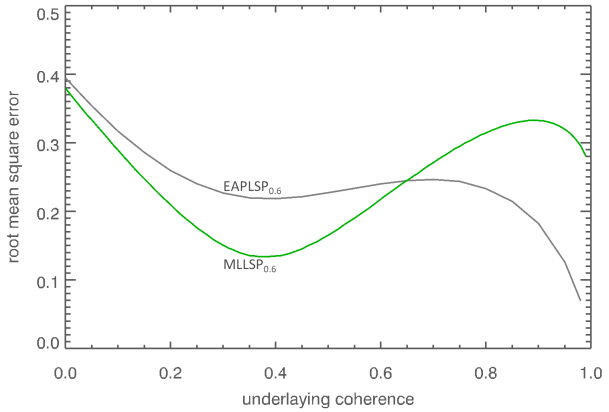


Fig. 11: Comparison of RMSE between estimation by expected *a posteriori* (EAPLSP$_{0.6}$) in gray and MLLSP$_{0.6}$ in green i.e. both with less strict prior for $\gamma_{max} = 0.6$ and $N = 3$ samples.

small sample sizes. For $N = 15$, the RMSE of the ML methods is still better compared to the sample estimator. The respective test case is visualized in Fig. 13. The MLWP is more efficient for all underlaying coherence magnitudes up to 0.40, the MLLSP$_{0.6}$ method up to 0.44 and the MLSP$_{0.6}$ estimator up to 0.46. However, the test case $N = 30$ demonstrates in Fig. 14, only the MLSP$_{0.6}$ performs better than the sample estimator. As indicated by Fig. 14 (c) and (f) in [22], the empirical Bayesian estimators are recommended for this test case. The plots show that a hybrid approach has to be developed in order to obtain the best estimate for an operational application. Such an approach should include the sample estimator, the ML methods, and the Bayesian methods [22]. As pointed out by a reviewer, the respective coherence magnitude estimate would be always better or equal to the sample estimator in terms of RMSE.

One reviewer argued that $N = 2$ is too low a sample count. It is a well-known fact; the concept of coherence is not relevant to individual samples requiring $N > 1$. The reason is that the coherence is a statistical quantity since an expected value has to be calculated as indicated by (3) in [22]. Practically, two samples are sufficient for the expected value to be meaningful. For this reason, the sample estimator (1) can be evaluated for $N = 2$ samples. Accordingly, if (1) can be evaluated meaningfully for two samples, then it can also be predicted by ML and is demonstrated in the section above. For the Bayesian estimation [22], it was surprising that $N = 2$ works because the used conditional probability density function (pdf) ([22, equation 12]) was reported by Touzi [17] to be valid for $N > 2$. The newly developed ML method is non-parametric and does not depend on this pdf. As a consequence, this restriction is not relevant. Interestingly, the practical demonstration of the empirical Bayesian estimator in [22] has shown that the conditional pdf also works for $N = 2$.

As already stated in [22] for the empirical Bayesian techniques, the demonstrated ML methods support typical InSAR scenarios. First, the MLWP improves the estimation without

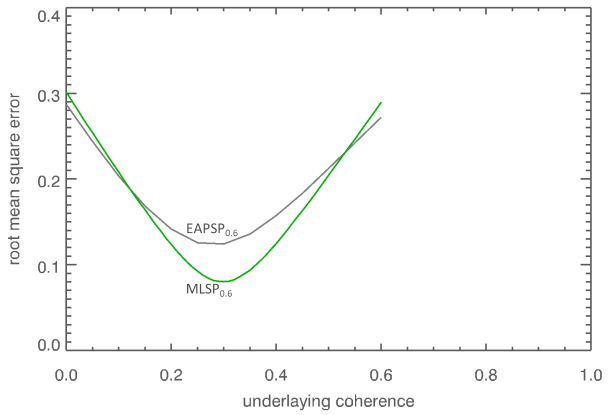


Fig. 12: Comparison of RMSE between estimation by expected *a posteriori* (EAPSP$_{0.6}$) in gray and MLSP$_{0.6}$ in green i.e. both with strict prior for $\gamma_{max} = 0.6$ and $N = 3$ samples.

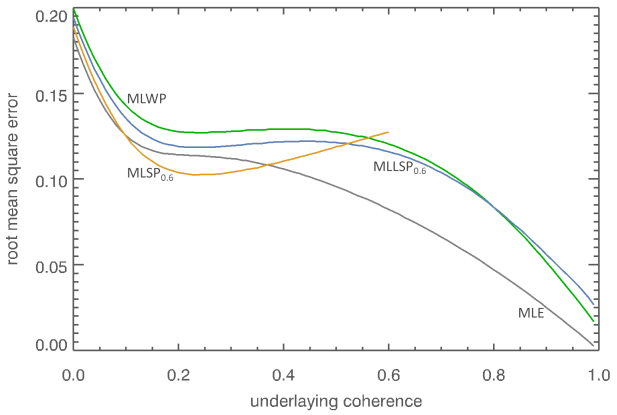


Fig. 14: Comparison of RMSE for $N = 30$ samples between sample estimator (gray), MLWP (green), MLLSP$_{0.6}$ (blue) and MLSP$_{0.6}$ (orange).

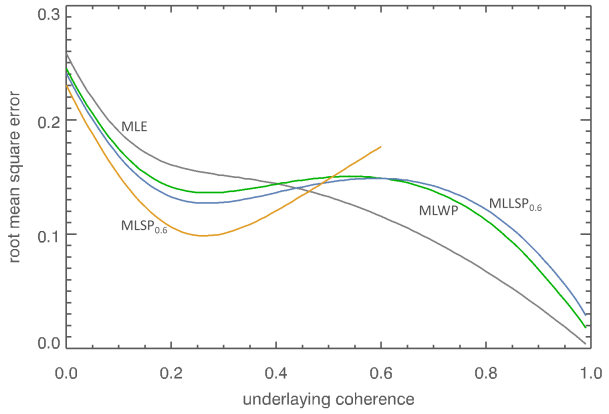


Fig. 13: Comparison of RMSE for $N = 15$ samples between sample estimator (gray), MLWP (green), MLLSP$_{0.6}$ (blue) and MLSP$_{0.6}$ (orange).

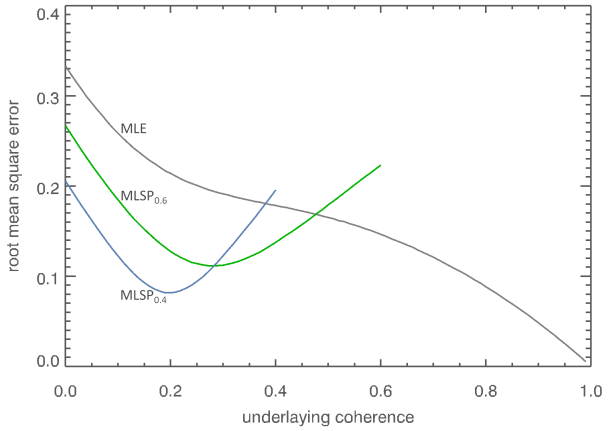


Fig. 15: Comparison of RMSE for $N = 9$ samples between sample estimator (gray), MLSP$_{0.6}$ (green) and MLSP$_{0.4}$ (blue).

prior knowledge and is generally applicable. Second, MLLSP and MLSP include an assumption on the maximum coherence magnitude γ_{max} of the underlaying true coherence. Such information is available in SAR interferometry based on stacks of interferograms. For example, $\hat{\gamma}_{max}$ can be estimated from an initial coherence matrix [3, Fig. 5] which can straight forward be converted into the best possible coherence as a function of acquisition time difference. Depending on the $\hat{\gamma}_{max}$ accuracy and the likelihood that the underlaying coherence is above $\hat{\gamma}_{max}$, the less strict or strict prior should be selected. The strict prior limits the estimates inside the assumed range and the less strict prior favors estimates in this range. In principle, any shape of prior can easily be implemented in the developed framework.

Practically, the greatest gain in precision is achieved with the strict prior. Not surprisingly, the more restrictive the prior, the better the estimation performance. Fig. 15 compares the estimators MLSP$_{0.6}$ and MLSP$_{0.4}$ using $N = 9$ as an example. It shows that the strict prior with a small γ_{max} should always be preferred. However, a reliable $\hat{\gamma}_{max}$ value for the application is crucial.

V. Conclusion

The developed ML coherence magnitude estimators are suitable and recommended for operational InSAR systems. First, they improve the estimation performance compared to the conventional sample estimator and to the empirical Bayesian estimators [22]. Especially, the estimation of small coherence magnitudes from a small sample is improved. Second, the framework supports any shape of Bayesian prior on the underlaying coherence magnitude. In this manuscript, the Bayesian prior is modeled with a single parameter (γ_{max}). Less strict and strict assumptions on the range of the underlaying coherence magnitude can be modeled and are demonstrated. Both types of prior correspond to typical InSAR scenarios. Third, the estimation is computationally extremely fast evaluated because no iteration, numeric integration or Bootstrapping is needed. Forth, the implementation is straight forward because

of the availability of many ML libraries. In this manuscript, the implementation utilizes XGBoost developed by Chen and Guestrin [24]. If a suitable encoder is used, the estimation results are independent of the ML method used.

A limitation of the newly developed ML methods is that the performance improves only for low coherences and for sample sizes $N < 30$.

The developed estimators are not limited to InSAR, but are generally applicable to coherence estimation problems from CCG processes.

ACKNOWLEDGMENT

I appreciate the work of the anonymous reviewers and I am grateful for the positive and very thorough comments and suggestions that helped improving the manuscript. The article contains modified Copernicus Sentinel data 2020, processed by ESA.

REFERENCES

[1] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 1, pp. 8–20, 2001.

[2] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti, "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 11, pp. 2375–2383, 2002.

[3] A. Ferretti, A. Fumagalli, F. Novali, C. Prati, F. Rocca, and A. Rucci, "A new algorithm for processing interferometric data-stacks: SqueeSAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 9, pp. 3460–3470, 2011.

[4] M. Frei, "White paper: European Ground Motion Service (EU-GMS) - A proposed Copernicus service element," 2017, pp. 1–24. [Online]. Available: https://land.copernicus.eu/user-corner/technical-library/egms-white-paper

[5] M. Crosetto, L. Solari, M. Mróz, J. Balasis-Levinsen, N. Casagli, M. Frei, A. Oyen, D. A. Moldestad, L. Bateson, L. Guerrieri, V. Comerci, and H. S. Andersen, "The evolution of wide-area DInSAR: From regional and national services to the European Ground Motion Service," *Remote Sensing*, vol. 12, no. 12, 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/12/2043

[6] M. Costantini, F. Minati, F. Trillo, A. Ferretti, E. Passera, A. Rucci, J. Dehls, Y. Larsen, P. Marinkovic, M. Eineder, R. Brcic, R. Siegmund, P. Kotzerke, A. Kenyeres, V. Costantini, S. Proietti, L. Solari, and H. S. Andersen, "EGMS: Europe-wide ground motion monitoring based on full resolution InSAR processing of all Sentinel-1 acquisitions," in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, 2022, pp. 5093–5096.

[7] A. C. Kalia, "User driven products in the context of the Ground Motion Service Germany," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 1688–1691.

[8] A. C. Kalia, M. Frei, and T. Lege, "Status of the operational Ground Motion Service Germany," in *EUSAR 2018; 12th European Conference on Synthetic Aperture Radar*, vol. 12, 2018, pp. 1–3. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8438197

[9] T. Lege, A. C. Kalia, and N. Adam, "Application of Persistent Scatterer Interferometry (PSI) indicates a significant sediment sink for the Ems estuary," in *Coastal Structures 2019*, ser. Hydraulic Engineering Repository (HENRY), N. Goseberg and T. Schlurmann, Eds., vol. 2019. Bundesanstalt für Wasserbau, Dezember 2019, pp. 793–802. [Online]. Available: https://elib.dlr.de/132689/

[10] N. Cao, H. Lee, and H. C. Jung, "Mathematical framework for phase-triangulation algorithms in distributed-scatterer interferometry," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1838–1842, 2015.

[11] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep Learning in remote sensing: A comprehensive review and list of resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 8–36, 2017.

[12] G. Fornaro, S. Verde, D. Reale, and A. Pauciullo, "CAESAR: An approach based on covariance matrix decomposition to improve multibaseline–multitemporal interferometric SAR processing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2050–2065, 2015.

[13] H. Ansari, F. D. Zan, and R. Bamler, "Sequential estimator: Toward efficient InSAR time series analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 10, pp. 5637–5652, September 2017.

[14] H. Zebker and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 5, pp. 950–959, 1992.

[15] D. Just and R. Bamler, "Phase statistics of interferograms with applications to synthetic aperture radar," *Appl. Opt.*, vol. 33, no. 20, pp. 4361–4368, Jul 1994. [Online]. Available: http://opg.optica.org/ao/abstract.cfm?URI=ao-33-20-4361

[16] J. W. Goodman, "Some fundamental properties of speckle∗," *J. Opt. Soc. Am.*, vol. 66, no. 11, pp. 1145–1150, Nov 1976. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=josa-66-11-1145

[17] R. Touzi and A. Lopes, "Statistics of the stokes parameters and of the complex coherence parameters in one-look and multilook speckle fields," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 2, pp. 519–531, 1996.

[18] R. Touzi, A. Lopes, J. Bruniquel, and P. Vachon, "Coherence estimation for SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 1, pp. 135–149, 1999.

[19] H. Zebker and K. Chen, "Accurate estimation of correlation in InSAR observations," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 2, pp. 124–127, 2005.

[20] R. Abdelfattah and J.-M. Nicolas, "Interferometric SAR coherence magnitude estimation using second kind statistics," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 7, pp. 1942–1953, 2006.

[21] M. Jiang, X. Ding, and Z. Li, "Hybrid approach for unbiased coherence estimation for multitemporal InSAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2459–2473, 2014.

[22] N. Adam, "Empirical Bayesian estimation of the interferometric SAR coherence magnitude," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 6306–6323, 2022. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9835014

[23] G. James, D. Witten, T. Hastie, and R. Tibshirani, *Linear Model Selection and Regularization*. New York, NY: Springer US, 2021, pp. 15–57. [Online]. Available: https://doi.org/10.1007/978-1-0716-1418-1_2

[24] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 785–794. [Online]. Available: http://doi.acm.org/10.1145/2939672.2939785

[25] T. Chen, M. Benesty, Y. Tang, N. Zhu, J. Yuan, H. Cho, R. Mitchell, H. Liu, T. He, V. Khotilovich, B. Xu, S. Lebedev, S. Lundberg, and E. Smirnov, "XGBoost eXtreme Gradient Boosting library." [Online]. Available: https://github.com/dmlc/xgboost

Nico Adam received the Diploma degree in electrical engineering and telecommunication science from the University of Rostock, Germany in 1995. Since 1995, he has been with the German Aerospace Center (DLR), Oberpfaffenhofen, Germany, where he manages SAR interferometry projects at DLR's Remote Sensing Technology Institute. Relevant projects are the Shuttle Radar Topography Mission (DLR, NASA, ASI), TerraSAR-X (DLR), Terrafirma (ESA) and the Sentinel-1 Ground Motion Monitoring Service Germany. In 2017, he was honored as "DLR Senior Scientist" for his application-oriented work within projects. His research interests include signal processing in the persistent scatterer interferometry framework, the use of numerical weather models to mitigate wave propagation effects in InSAR and the development of algorithms and software for advanced remote sensing radar systems.