Phase-Based Similarly Decorrelated Pixel Selection and Phase-Linking in InSAR Using Circular Statistics

Shuyi Yao, *Student Member, IEEE*, and Timo Balz, *Senior Member, IEEE*

***Abstract*— Circular statistics is the mathematical theory for dealing with variables distributed on a circle. The interferometric phase of distributed targets can be modeled with circular statistics due to its wrapped and pseudorandom nature. In this study, we introduce a novel adaptive neighborhood selection (ANS) method and a novel phase-linking (PL) method for distributed scatterer (DS) interferometry, both based on circular statistics principles. The proposed ANS method enables the direct selection of pixels with similar SAR interferometry (InSAR) decorrelation behaviors, called similarly decorrelated pixels (SDP), from the interferometric phases. The proposed PL method: 1) shows significant resistance to the potential departure from the fully developed speckle assumption in the SAR observations (also known as non-Gaussianity) compared to methods that rely on this assumption; 2) does not introduce substantial implementation complexity, computational cost, or numerical solution challenges compared to methods that elaborately model the non-Gaussianity, e.g., through a product model; and 3) can achieve higher consistent wrapped phase estimation precision compared to other methods solely based on interferometric phases. In addition to validating the proposed methods through simulation experiments, we also found that a combination of the two proposed methods can produce interferograms with minimal noise in a real data experiment.**

***Index Terms*— Adaptive neighborhood selection (ANS), circular statistics, distributed scatterer (DS), interferogram filtering, non-Gaussianity, phase-linking (PL), SAR interferometry (InSAR), synthetic aperture radar (SAR), trigonometric moment, von Mises distribution.**

I. Introduction

SAR interferometry (InSAR) has become a powerful tool for estimating motion on the Earth's surface with unprecedented accuracy and resolution. The core of InSAR analysis and applications lies in estimating parameters from the interferometric phases. Thus, the final results will certainly benefit from accurate phase estimates. Distributed targets are widespread in natural scenes, and susceptible to noise and interference. The original interferometric phase observed from a distributed scatterer (DS) is often noisy. To adequately reduce the phase noise and reliably estimate the phase, an appropriate stochastic phase model is required and a set of samples that can fit that stochastic model needs to be selected. Once the samples and the stochastic model are determined, a set of consistent wrapped phases can be estimated by employing the phase closure constraint [1]. For example, in SqueeSAR [2], the Gaussian speckle assumption, also known as the fully developed speckle assumption [3], is adopted to derive its stochastic model for its phase estimation process, i.e., the phase triangulation algorithm (PTA), and a two-sample statistical test working on the amplitude time series is used to select the statistically homogenous pixel (SHP). Selecting a set of pixels with some kind of homogeneity (referred to as "brother pixels") and then estimating the consistent wrapped phases is a typical two-step framework in DS-InSAR [4]. In recent years, different methods have been proposed for DS phase estimation within this framework [5], [6], [7], [8].

We refer to the process of selecting brother pixels as adaptive neighborhood selection (ANS) and the subsequent consistent wrapped phase estimation as phase-linking (PL). In ANS, the amplitude or intensity time series is widely used as the a metric to determine whether two pixels are statistically similar because: 1) they could be used as a proxy for phases and 2) they are linear variables and do not suffer from wrapping as phases do. However, the amplitude or intensity inherently represents the energy reflected from the target, whereas the noise distribution of the DS phases is mainly determined by the interference of elementary scatterer motion (temporal decorrelation). Thus, the amplitude or intensity does not necessarily represent the phase noise distribution [8]. Currently, we lack methods that directly select the pixels with similar phase noise distributions [9] and we lack comparisons between the results of these two pixel selection schemes.

The Gaussian speckle model is widely adopted in various PL methods, which takes the following assumptions [10], [11]:

1) The amplitude and the phase of every elementary scatterer are statistically collective independent.
2) The phase of every elementary scatterer is uniformly distributed in $[-\pi, \pi)$.
3) No single scatterer dominates the others in a resolution cell.
4) The number of elementary scatterer approaches infinite.

Therefore, the mathematical derivation results in a multivariate complex circular Gaussian (CCG) distribution for one SAR pixel [12]. We therefore refer to this SAR stochastic model as the CCG model. PL methods based on the CCG model rely on its corresponding assumptions, so their accuracy will

Manuscript received 10 April 2024; revised 4 June 2024; accepted 29 June 2024. Date of publication 2 July 2024; date of current version 10 July 2024. This work was supported by the National Natural Science Foundation of China under Grant 42250610212. *(Corresponding author: Timo Balz.)*

The authors are with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: yaoshuyi@whu.edu.cn; balz@whu.edu.cn).

Digital Object Identifier 10.1109/TGRS.2024.3422171

decrease if these assumptions differ from reality for the reason of, say, a small resolution cell with few elementary scatterers. There are few methods to account for this deviation using an additional independent real product factor [3], [13]. These factors can introduce additional computational cost due to non-analytical probability density function (pdf) and challenge the numerical optimization process. The EMCF-SBAS method [7] provides a simpler way of PL without assuming any particular distribution.

It is possible to group the pixels and estimate the consistent wrapped phases without using any amplitude data. In ANS, this means selecting pixels directly from their phase noise distributions and avoiding the disadvantages caused by the potential discrepancy between amplitude statistics and phase statistics. In the PL, dropping the amplitude can lead to more generalized models, as the correlation between interferometric amplitudes and phases no longer needs to be assumed, though at the cost of losing some potentially useful information. The optimal DS-InSAR processing under this condition is still an open issue. Since the phases are wrapped and can therefore be seen as distributed on a circle, the circular statistics, also known as the directional statistics [14], of mathematical sciences can be a promising tool. For example, the Kuiper's test, which concerns whether two groups of circular data have similar distribution, has been adopted in the ANS [9] and the PL of EMCF-SBAS [7] approximates to weight the phase of each interferogram with the mean resultant length.

In this article, we present a new ANS method and a new weighting scheme for the PL. Both parts are based on the theory of circular statistics. In our ANS method, instead of identifying pixels with similar amplitude statistics, we focus on the pixels that have a similar interferometric phase component corresponding to the target decorrelation. We refer to these pixels as similarly decorrelated pixels (SDP). The disadvantage caused by the potential discrepancy between the statistical properties of amplitude and phase is overcome in our ANS method. A clustering approach is used for our ANS, yielding a more efficient selection compared to applying two-sample tests which are designed for circular data. In our PL method, we weight the interferometric phases with the inverse of the variance of their circular sample means according to the von Mises distribution, and a debiasing is applied. This approach can lead to the resistance to non-Gaussianity in intensities and thus to the robustness, without introducing significant implementation complexity and computational burden. Compared to other PL methods that focus on the phases, e.g., EMCF-SBAS, the proposed PL method has a more solid mathematical basis and can therefore achieve a higher theoretical accuracy.

This article is organized as follows. Section II introduces the interferometric signal model. Section III describes the proposed ANS method. The proposed PL method is introduced in Section IV. Section V is the experiment with real data. The conclusions are summarized in Section VI.

II. Interferometric Signal Model

Let y_m and y_n be complex signals received from two well-aligned SAR pixels at different times over a distributed target. Assuming that the number of elementary scatterers is very large and the speckle is fully developed, then $[y_m, y_n]^{\mathrm{T}}$ follows a bivariate CCG distribution. By assuming the spatial samples in a window are i.i.d. distributed, the corresponding joint maximum likelihood estimator (MLE) of the coherence magnitude $\gamma_{m,n}$ and the true interferometric phase $\psi^0_{m,n}$ is

$$\widehat{\gamma}_{m,n}e^{\jmath\widehat{\psi}^0_{m,n}} = \frac{\sum_{p=1}^{N_p}|y_m(p)||y_n(p)|e^{\jmath(\psi_m(p)-\psi_n(p))}}{\sqrt{\sum_{p=1}^{N_p}|y_m(p)|^2\sum_{p=1}^{N_p}|y_n(p)|^2}} \tag{1}$$

where $\jmath = (-1)^{1/2}$ and N_p denotes the number of pixels. From the numerator in (1), we can see that the estimate of $\psi_{m,n}$ can be regarded as the sample mean of $e^{\jmath(\psi_m(p)-\psi_n(p))}$ weighted by the interferometric amplitude $A_{\mathrm{int}}(p) = |y_m(p)||y_n(p)|$, indicating that the pixels with larger amplitude tend to have smaller residuals.[1] Furthermore, $A_{\mathrm{int}}(p)$ is the optimal weight rather than any other increasing function of $A_{\mathrm{int}}(p)$, and the phase uncertainty can be accurately described by the underlying $\gamma_{m,n}$, for example, through the Cramér-Rao bound [15] $\mathrm{var}(\psi)_{\mathrm{cr}} = (1-\gamma^2_{m,n})/2N_p\gamma^2_{m,n}$. However, this quantitative coupling is directly associated with the fully developed speckle assumption. It is shown that in modern SAR systems, departures from the fully developed speckle assumption can occur frequently [16], causing the CCG model to become inaccurate.

The departures from the CCG model, also known as the non-Gaussianity, can occur in two ways. One is the so-called texture with clear boundaries between different regions of the target. In this case, adaptive filters focusing on the intensity or amplitude can be appropriate. In the other case, there is an intrinsic spatial fluctuation of the target signal from pixel to pixel. This kind of non-Gaussianity can be attributed to the insufficiency of elementary scatterers in a pixel due to the improvement of spatial resolution. For interferometry aimed at surface motion estimation, the non-Gaussianity could be addressed by more elaborate adaptive filters, especially focusing on the outlier suppression [17], [18]. However, it can be expected that the adaptive filters aimed at selecting CCG samples can easily drop pixels in such areas. Another way to account for the non-Gaussianity is to use the product models [3], [13], [19], [20], which assume the received SAR signal

$$y = z \times x \tag{2}$$

where z is a real factor that can be constrained via a distribution [3] or on its moment(s) [19] or without constraint [13], and x is the component that fits the CCG model. It can be deduced that such models: 1) assume that the non-Gaussianity factor z does not affect the decorrelation of the interferometric phase because of the independence between z and x, which is not verified in terms of the scattering mechanism and 2) still invoke the central limit theorem because of the term x.

[1] The phase residual is defined as one minus the cosine of the difference of the phase derived inversely from the estimated parameter between the original observed phase. In (1), the estimated parameter is the multilooked phase $\widehat{\psi}^0_{m,n}$, so the phase derived inversely from it is also $\widehat{\psi}^0_{m,n}(p) = \widehat{\psi}^0_{m,n}$, and the original observed phase is $\arg[\exp(\jmath(\psi_m(p) - \psi_n(p)))] = \psi_{m,n}(p)$. Therefore, the phase residual is $1 - \cos(\widehat{\psi}^0_{m,n} - \psi_{m,n}(p))$.

The potential failure of the first assumption can distort the actual coupling of interferometric amplitude and phase, and the second can make an inaccurate assumption. In addition, numerical methods (e.g., integration) are often required for estimates based on such models, which increases the computational burden and challenges the estimation.

Instead of exploiting the coupling between interferometric amplitude and phase distorted by the non-Gaussianity, we adopt a simple model that focuses only on the interferometric phase. We assume that the interferometric phases are wrapped normal distributed around their true values due to the random change of the relative positions between the elementary scatterers and the antenna [21], and can therefore be approximated with the von Mises distribution [22], regardless of whether the joint amplitude-phase distribution can fit the CCG model or not. This interferometric phase model is hereafter referred to as the VM model. The pdf of the von Mises distribution and wrapped normal distribution are

$$f_{\mathrm{VM}}(\psi;\psi^0,\kappa)=\frac{1}{2\pi I_0(\kappa)}\exp\left(\kappa\cos\left(\psi-\psi^0\right)\right) \tag{3}$$

$$f_{\mathrm{WN}}(\psi;\psi^0,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\sum_{k=-\infty}^{\infty}\exp\left\{\frac{-\left(\psi-\psi^0+2\pi k\right)^2}{2\sigma^2}\right\} \tag{4}$$

where I_0 is the modified Bessel function of the first kind and order zero, and κ is referred to as the concentration parameter. κ acts as the inverse of the variance, so a random angle (noisy interferometric phase ψ) will be more dispersed for a smaller κ. It should be noted that we define all phase or angle values in the interval $[-\pi,\pi)$. Without an infinite sum in the density function, the von Mises distribution is easier to use than the wrapped normal distribution in many cases [23].

We show the compatibility of the VM model with the CCG phase model by plotting their pdfs together (see Fig. 1). The pdf of the interferometric phase under the CCG model (single-look) is

$$\begin{aligned}&f_{\mathrm{CCG}}(\psi;\psi^0,\gamma)\\&=\frac{\left(1-\gamma^2\right)}{2\pi}\frac{1}{1-\gamma^2\cos^2\left(\psi-\psi^0\right)}\\&\times\left(\frac{\gamma\cos\left(\psi-\psi^0\right)\arccos\left(-\gamma\cos\left(\psi-\psi^0\right)\right)}{\sqrt{\left(1-\gamma^2\cos^2\left(\psi-\psi^0\right)\right)}}+1\right).\end{aligned} \tag{5}$$

It is shown that both f_{VM} and f_{CCG} are unimodal in the field of definition $[-\pi,\pi)$ (and so is f_{WN}). We compared the pdfs of two phase distributions with the same degree of uncertainty, measured by an index that does not depend on any particular distribution, the mean resultant length R. The mean resultant length is defined as

$$R=\left|E\left[e^{\jmath\psi}\right]\right|=E\left[\cos\left(\psi-\psi^0\right)\right]. \tag{6}$$

To obtain the von Mises pdf with respect to a certain coherence value, the following relationship is adopted [21]:

$$R=\frac{\pi}{4}\gamma F\left(\frac{1}{2},\frac{1}{2};2;\gamma^2\right) \tag{7}$$

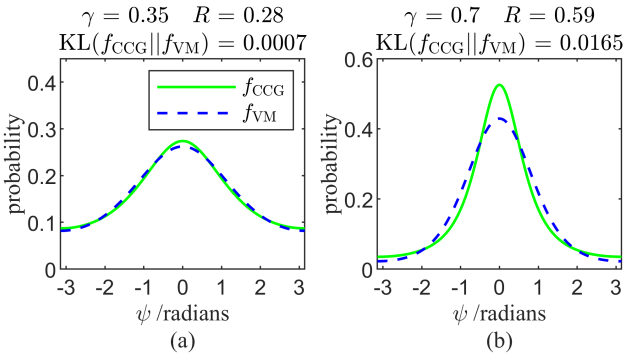

Fig. 1. Interferometric phase pdfs with (a) $R=0.28$ and (b) $R=0.59$. The Kullback–Leibler divergence between two curves $\mathrm{KL}(f_{\mathrm{CCG}}||f_{\mathrm{VM}})$ is calculated numerically.

where $F()$ is the hypergeometric function. It is shown that both R and γ take values from 0 to 1 and R is monotone increasing with γ [21]. Fig. 1 shows that both these two pdfs are quite similar at medium-low coherence, and the difference between them increases with coherence. This increasing difference can be recognized by the increasing Kullback–Leibler divergence $\mathrm{KL}(f_{\mathrm{CCG}}||f_{\mathrm{VM}})$. However, as the coherence increases, the difficulty of estimating the phase and its uncertainty decreases. Furthermore, these two pdfs both converge to the uniform distribution when $R=\gamma=0$. This compatibility indicates that the VM model can also work on the CCG distributed SAR observations.

III. SDP Clustering

We assume that N SAR images acquired over the same area have been co-registered and M interferograms have been generated. We use $\psi_i(p)\in[-\pi,\pi)$ to denote the phase of a pixel with coordinates $p\equiv(x,y)$ in the ith interferogram. Thus, $\psi_i(p)$ can be seen as a random angle, i.e., a circular random variable. To identify SDPs with interferometric phase time series, the spatial low-pass (SLP) signal component must first be removed, since its high-pass behavior in time can overwhelm the phase component corresponding to the decorrelation. Same as in [9] and [24], we estimate the SLP component using the standard circular mean estimator in circular statistics

$$\psi_{i,\mathrm{SLP}}=\arg\left\{\sum_{p\in\Omega_b}\exp(-\jmath\psi_i(p))\right\} \tag{8}$$

where Ω_b denotes a boxcar spatial neighborhood. It is shown that this estimator is robust to potential outliers [25] compared to the mean of linear samples. The SLP component can also be eliminated by spatial differentiation [8]. In Ω_b, the phase components corresponding to the target decorrelation of each pixel are the residuals after subtraction $\Psi_{\mathrm{res}}(p)=W\{\Psi(p)-\Psi_{\mathrm{SLP}}\}$, where $\Psi=[\psi_1,\psi_2,\ldots,\psi_M]$ denotes the vector containing the temporal samples and $W\{\}$ represents wrapping the input values into $[-\pi,\pi)$.

We determine which pixels are similarly decorrelated by analyzing the statistical properties of $\psi_{\mathrm{res}}(p)$. According to the signal model described in Section II, pixels that are grouped into SDPs mean that they have similar κ in each interferogram

(we assume that the phase is locally stationary so that pixels within a boxcar window always have similar ψ^0). As with the treatment of linear data such as amplitudes and intensities in SHP identification, two-sample statistical tests could be an option [2], and nonparametric tests designed for circular samples, such as Kuiper's test [9], [26] and Watson's-U^2 [27] test, could be used. A test relying on a particular distribution is not suitable for $\Psi_{\text{res}}(p)$, since according to the adopted signal model, $\Psi_{\text{res}}(p)$ contains correlated samples from a set of von Mises distributions with mixture parameters.

Instead of using two-sample statistical tests, we present a different approach to determine the adaptive neighborhood from the phases. Before describing our approach, we need to introduce the characteristic function and the trigonometric moments of a circular variable. Given $\psi \in [-\pi, \pi)$ is a random angle, its characteristic function is

$$\phi_p = E\left[e^{\jmath p\psi}\right] = \int_{-\pi}^{\pi} (e)^{\jmath p\psi} dF(\psi), \quad p = 0, \pm 1, \pm 2, \ldots \tag{9}$$

where $F(\psi)$ is the probability distribution function. We shall have

$$\phi_p = \alpha_p + \jmath\beta_p \tag{10}$$

$$\alpha_p = E\left[\cos p\psi\right], \quad \left|\alpha_p\right| < 1 \tag{11}$$

$$\beta_p = E\left[\sin p\psi\right], \quad \left|\beta_p\right| < 1 \tag{12}$$

where α_p and β_p are defined as the trigonometric moments. It is shown that the probability distribution of a random angle is determined by its characteristic function, and the sequence $\{(\alpha_p, \beta_p): p = 0, \pm 1, \ldots\}$ is equivalent to the characteristic function of ψ [14]. Thus, the distribution of a random angle is determined by its trigonometric moments.

We assume that the pixels in a boxcar window have N_c decorrelation behaviors, and can form N_c clusters in a N_m dimensional space, where N_m is the number of orders of the trigonometric moments. Therefore, we identify the SDP by clustering according to the sample moments of Ψ_{res}. We choose the first four-order moments for the clustering ($N_m = 4$), as the location, dispersion, skewness, and kurtosis, which control the shape of a probability distribution, are controlled up to the fourth moment. Since the SPL component is preremoved, we assume $\widehat{\beta_p} = E[\sin p\psi_{\text{res}}] = 0$, and thus $\widehat{\phi_p} = \widehat{\alpha_p} = (1/M)\sum_{i=1}^{M} \cos p\psi_{i,\text{res}}$. In our approach, we use the k-means clustering method [28] and we estimate N_c as the number of near-zero eigenvalues of the normalized random-walk Laplacian matrix [29], [30]. The cluster containing the central pixel forms the adaptive neighborhood.

We designed a simulation experiment to evaluate the performance of different phase-based ANS methods. Kuiper's test, Watson's-U^2 test, and our SDP clustering approach were tested. A total of 25 SAR images were simulated and 300 interferograms were generated. We assumed that there are two different decorrelation behaviors in a window [see Fig. 2(a)], one with an exponential decay and the other with almost complete decorrelation in each interferogram. The decorrelation model is

$$\gamma_{m,n} = p_0 \exp\left(-\delta t_{m,n}/\tau\right) + p_\infty \tag{13}$$

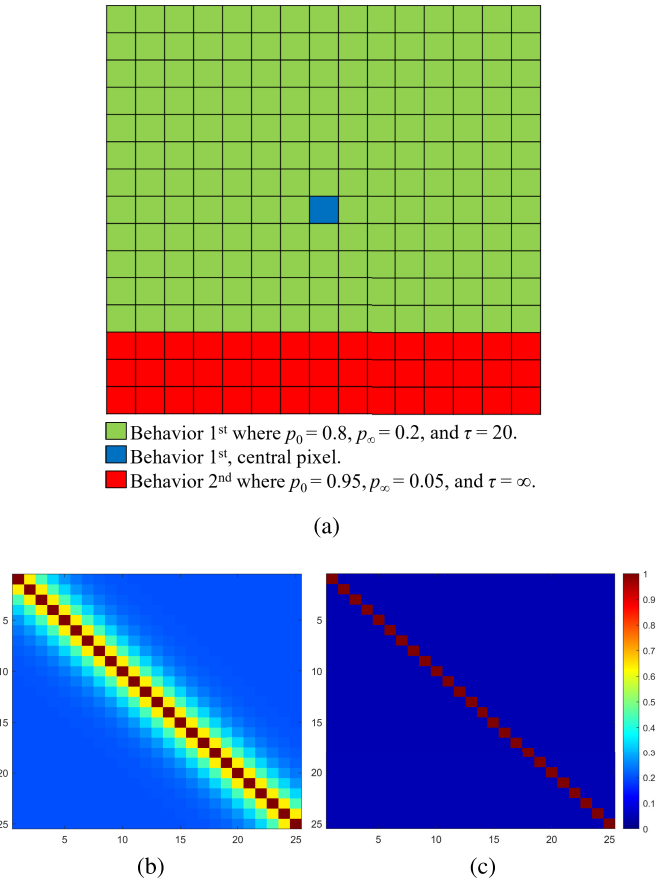

Fig. 2. (a) Simulated multilooking window with two decorrelation behaviors. (b) Coherence magnitude matrix of the first decorrelation behavior. (c) Coherence magnitude matrix of the second decorrelation behavior.

TABLE I
COMPARATION OF THREE ANS METHODS

Methods		Precision	Recall	Harmonic mean
Kuiper's test	$\alpha = 0.05$	0.4346	0.6152	0.5094
	$\alpha = 0.01$	0.4329	0.5276	0.4756
Watson's-U^2 test	$\alpha = 0.05$	0.4398	0.6654	0.5296
	$\alpha = 0.01$	0.4372	0.5710	0.4953
SDP clustering		0.5955	0.9592	0.7348

The Harmonic mean is defined as $2/(1/\text{Precision} + 1/\text{Recall})$.

where p_0 and p_∞ are the proportions of short-term coherent scatterers and the long-term persistently coherent scatterers, respectively, $p_0 + p_\infty = 1$, $\delta t_{m,n}$ is the temporal baseline, and τ is the decorrelation rate [31]. We set $p_0 = 0.8$, $p_\infty = 0.2$, and $\tau = 20$ in the first decorrelation behavior, which gives that the coherence with a temporal baseline of 12 days is 0.64 and with a temporal baseline greater than 204 days is 0.2. For the second decorrelation behavior, we assumed that $p_0 = 0.95$, $p_\infty = 0.05$, and $\tau = \infty$, resulting in that the coherence magnitude in each interferogram is 0.05 [see Fig. 2(b) and (c)]. We assumed that the temporal baseline between each pair of adjacent acquisitions was 12 days. We repeated the test 2000 times and summarized their precision, recall, and harmonic mean in Table I. As the alternative hypothesis of a two-sample statistical test is typically that the samples are from different distributions, we took points identified as nonbrother pixels as positive

samples in the calculation of precision and recall. Thus, the precision indicates the ability to include more brother pixels correctly, and the recall indicates the ability to ensure that the identified brother pixels actually have similar distributions. The harmonic mean, also known as the $F1$ score, balances between precision and recall and provides a single metric to assess the ANS performance. According to the recall, our approach shows an overwhelming advantage in excluding the heterogeneous pixels against Kuiper's test and Watson's-U^2 test. Furthermore, the precision of our method is also superior to that of Kuiper's test and Watson's-U^2 test.

IV. Phase-Linking

PL, also known as phase optimization, is the process of estimating $N-1$ optimal consistent wrapped interferometric phases $\boldsymbol{\theta} = [0, \theta_1, \theta_2, \ldots, \theta_{N-1}]$ by employing the phase closure constraint [2], [5], [32]. The first element of $\boldsymbol{\theta}$ is the zero-reference. The first PL algorithm is an MLE associated with the multivariate CCG distribution [33]. By replacing the complex covariance matrix with the coherence matrix, PTA has almost the same objective function as in [33], and they can be considered as using the same model. As shown in [34], PL can be seen as the minimization of the weighted residuals between the parameterized phases of each interferogram and their initial (multilooked) values. In this format, the objective function can be expressed as

$$\arg\min_{\boldsymbol{\theta}} \left\{ -\sum_{m=1}^{N}\sum_{n>m}^{N} (w)_{m,n} \cos\left(\theta_m - \theta_n - \widehat{\psi}^0_{m,n}\right) \right\}. \tag{14}$$

The objective function can also be further parameterized with additional terms, for example, accounting for the uncertainty in $w_{m,n}$ [19], [35].

After SDP clustering, an adaptive neighborhood Ω_a is determined. We assume that for each interferogram, Ω_a contains i.i.d. von Mises samples, and the MLE of the multilooked phase and the sample mean resultant length $\overline{R}$ in the ith interferogram is

$$\overline{R}_i e^{\jmath\overline{\psi}_i} = \overline{R}_i e^{\jmath\widehat{\psi}^0_i} = \sum_{p\in\Omega_a} \exp(\jmath\psi_i(p)). \tag{15}$$

Equation (15) provides another way to estimate ψ^0 that does not rely on the CCG model, which is different from (1). It is shown that the variance of $\overline{\psi}$ is [14]

$$\begin{aligned}\operatorname{var}\left(\overline{\psi}\right) &= \frac{1}{\kappa A(\kappa) L} + \frac{3\kappa\left(1-A(\kappa)^2\right) - 5A(\kappa)}{\kappa^2 A(\kappa)^3 L^2} + O\left(L^{-2}\right) \\ &\approx \frac{1}{\kappa A(\kappa) L}\end{aligned} \tag{16}$$

where L is the number of independent samples and

$$A(\kappa) = \frac{I_1(\kappa)}{I_0(\kappa)} = R. \tag{17}$$

Thus, in our PL method, we determine that

$$w = 1/\operatorname{var}\left(\overline{\psi}\right) = \kappa A(\kappa) L. \tag{18}$$

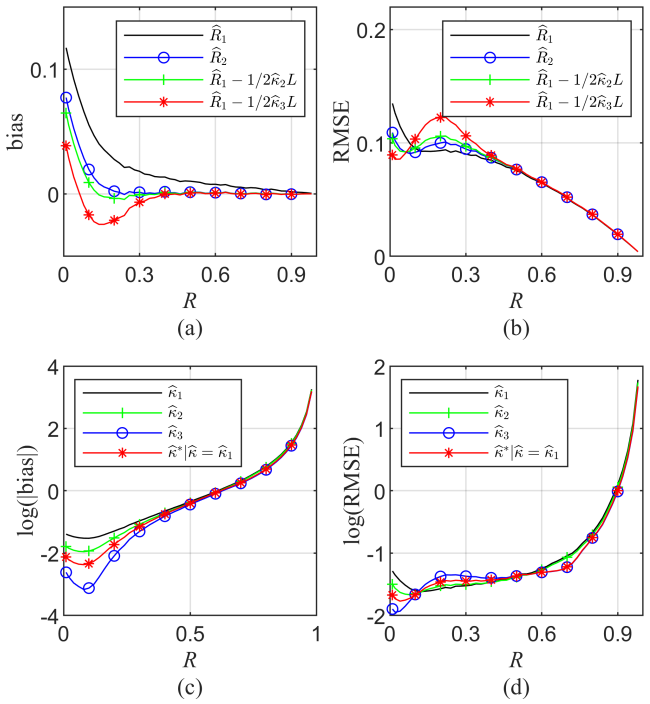

Fig. 3. (a) Bias of $\widehat{R}$, (b) RMSE of $\widehat{R}$, (c) absolute bias of $\widehat{\kappa}$, and (d) RMSE of $\widehat{\kappa}$. (c) and (d) are on a logarithmic scale for a better visualization. $L = 49$.

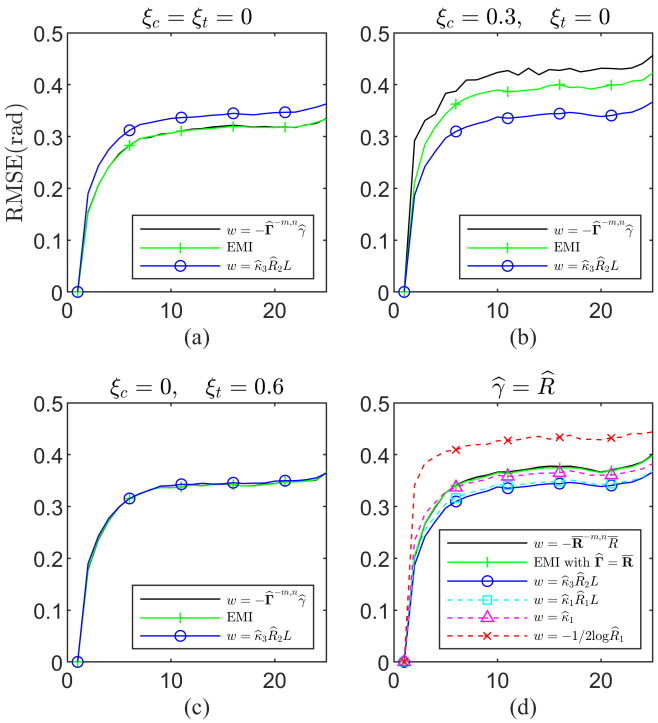

Fig. 4. RMSE of the estimated consistent phase in a simulated SLC stack. $N = 25$, $M = (N-1)N/2 = 300$, $L = 49$. (a) Gaussian speckle case and (b) and (c) are two non-Gaussian cases generated with the KM model [3]. In (d), we compare several phase-based approaches.

The term $\kappa A(\kappa) L$ can also be interpreted as the Fisher information

$$\mathbf{J} = -E\begin{bmatrix} \dfrac{\partial^2 \sum \ln f_{\mathrm{VM}}}{\partial {\psi^0}^2} & \dfrac{\partial^2 \sum \ln f_{\mathrm{VM}}}{\partial \psi^0 \partial \kappa} \\ \dfrac{\partial^2 \sum \ln f_{\mathrm{VM}}}{\partial \kappa \partial \psi^0} & \dfrac{\partial^2 \sum \ln f_{\mathrm{VM}}}{\partial \kappa^2} \end{bmatrix}$$

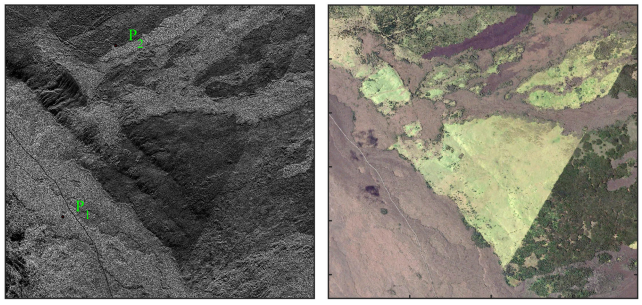

Fig. 5. (a) Time-averaged amplitude of 19 TerraSAR-X data (©DLR, 2022) and (b) optical image (from ©Google Earth) of the study area. Both are in the SAR coordinates.

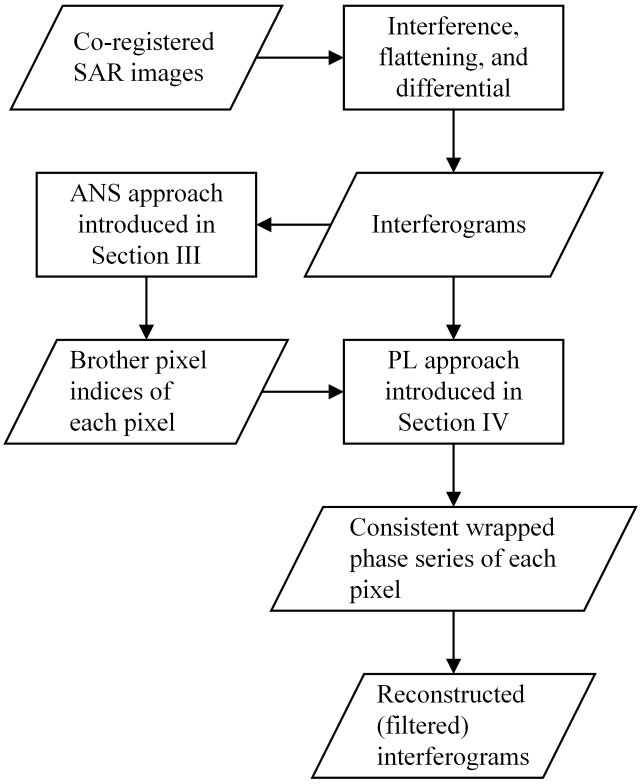

Fig. 6. Flowchart of data processing.

$$= L\begin{bmatrix} \kappa A(\kappa) & 0 \\ 0 & 1 - A(\kappa)^2 - A(\kappa)/\kappa \end{bmatrix}. \tag{19}$$

Therefore, the weighting factor $\kappa A(\kappa)L$ is similar to $(2L\gamma^2)/(1-\gamma^2)$ in the CCG model [5]. It should be noted that our PL method (as well as weighting the phases with $(2L\gamma^2)/(1-\gamma^2)$) ignores the correlations between different interferograms, which causes an inherent departure from the possible optimal estimate.

The estimation of R and k is required for getting w. It is shown that $\overline{R}$ is a biased estimate of R and the bias can be approximated by [36]

$$E\left[\overline{R} - R\right] \approx \frac{1}{2\kappa L}. \tag{20}$$

To estimate κ, the inverse of $A(\kappa)$ is required. Best and Fisher [37] gives an approximation of $A^{-1}(R)$ with good accuracy

$$\widehat{\kappa} = A^{-1}(R) \approx \begin{cases} 2R + R^3 + \left(5R^5\right)/6, & R < 0.53 \\ -0.4 + 1.39R + 0.43/(1-R), & 0.53 \le R \le 0.85 \\ \left(R^3 - 4R^2 + 3R\right)^{-1}, & R > 0.85. \end{cases} \tag{21}$$

However, $\widehat{\kappa}$ is also biased due to the approximation of $A^{-1}(R)$, Best and Fisher suggested an estimator of κ with bias reduction [37]

$$\widehat{\kappa}^* = \begin{cases} \max\left(\widehat{\kappa} - 1/2\widehat{\kappa}L, 0\right), & \widehat{\kappa} < 2 \\ (L-1)^3\widehat{\kappa}/\left(L^3 + L\right), & \widehat{\kappa} \ge 2. \end{cases} \tag{22}$$

It should be noted that a debias processing may cause an increase of the variance. To ensure a small overall root mean square error (RMSE) of $\widehat{R}$ and $\widehat{\kappa}$, we present a way to obtain $\widehat{w}$

1) Getting $\widehat{R}_1$ with (15) $\widehat{R}_1 = \overline{R}$.
2) Getting $\widehat{\kappa}_1$ with $\widehat{\kappa}_1 = A^{-1}(\widehat{R}_1)$, where $A^{-1}(x)$ is approximated by (21).
3) The debias of $\widehat{R}$ with (20): $\widehat{R}_2 = \widehat{R}_1 - 1/2\widehat{\kappa}_1 L$.
4) Getting $\widehat{\kappa}_2$ with $\widehat{\kappa}_2 = A^{-1}(\widehat{R}_2)$, where $A^{-1}(x)$ is approximated by (21).
5) The debias of $\widehat{\kappa}_2$ with (22): $\widehat{\kappa}_3 = (\widehat{\kappa}^* | \widehat{\kappa} = \widehat{\kappa}_2)$.
6) Calculating the weight factor $\widehat{w} = \widehat{\kappa}_3 \widehat{R}_2 L$.

Once the weights of each interferogram have been calculated, (14) can be optimized via a quasi-Newton algorithm as in PTA.

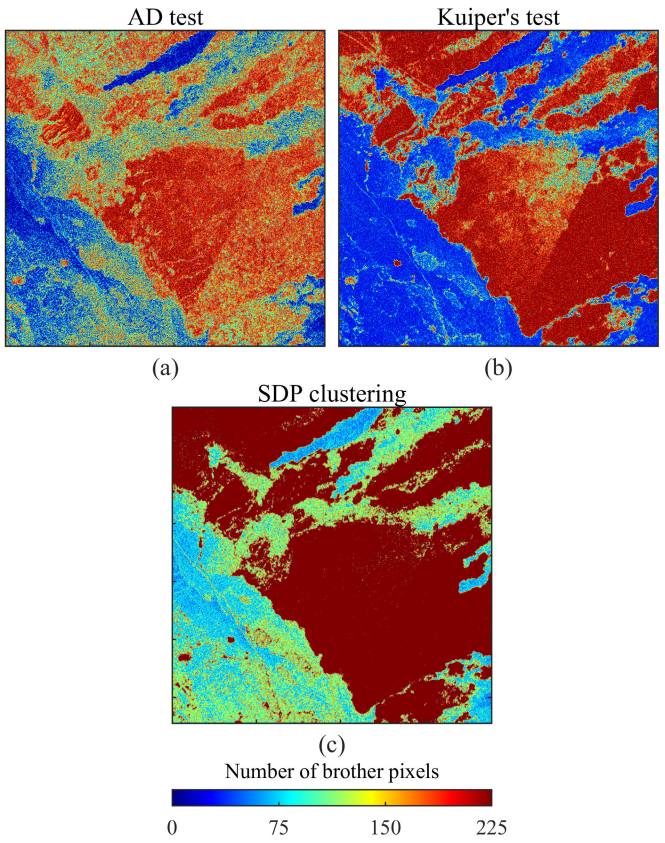

Fig. 7. Number of brother pixels identified by (a) AD test, (b) Kuiper's test, and (c) SDP clustering.

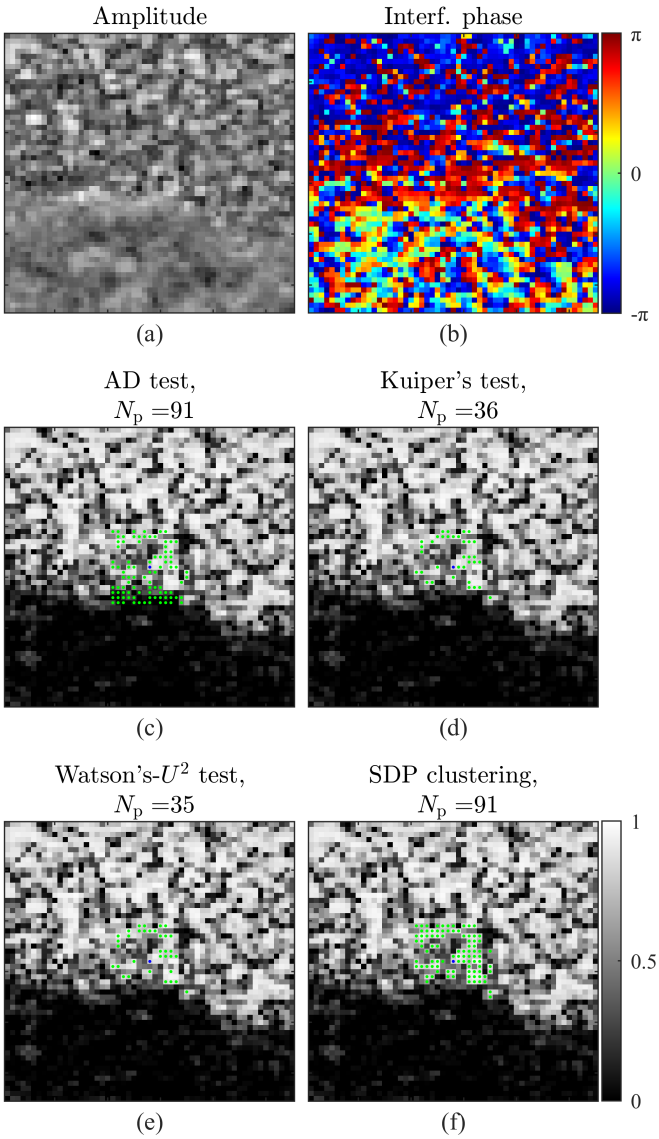

Fig. 8. Identified brother pixels of pixel P_1. (a) Time-averaged amplitude on a logarithmic scale. (b) One of the interferogram in the stack. (c)–(f) Brother pixels identified by four methods, with blue denoting the centra pixel and green denoting the identified brother pixel. The significant level used in (c)–(e) is 0.05. The base map of (c)–(f) is $\widehat{\alpha_1}$.

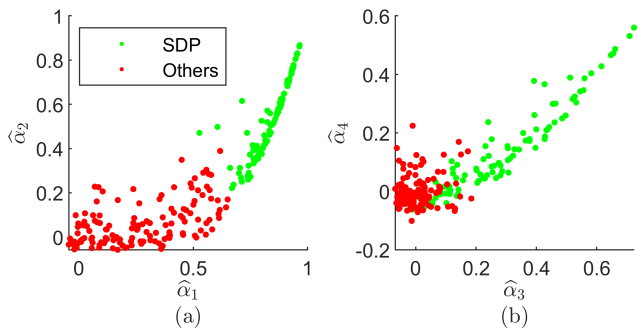

Fig. 9. Clustered scatterplot of the sample trigonometric moments of pixel P_1. (a) $\widehat{\alpha_1}$ and $\widehat{\alpha_2}$ and (b) $\widehat{\alpha_3}$ and $\widehat{\alpha_4}$.

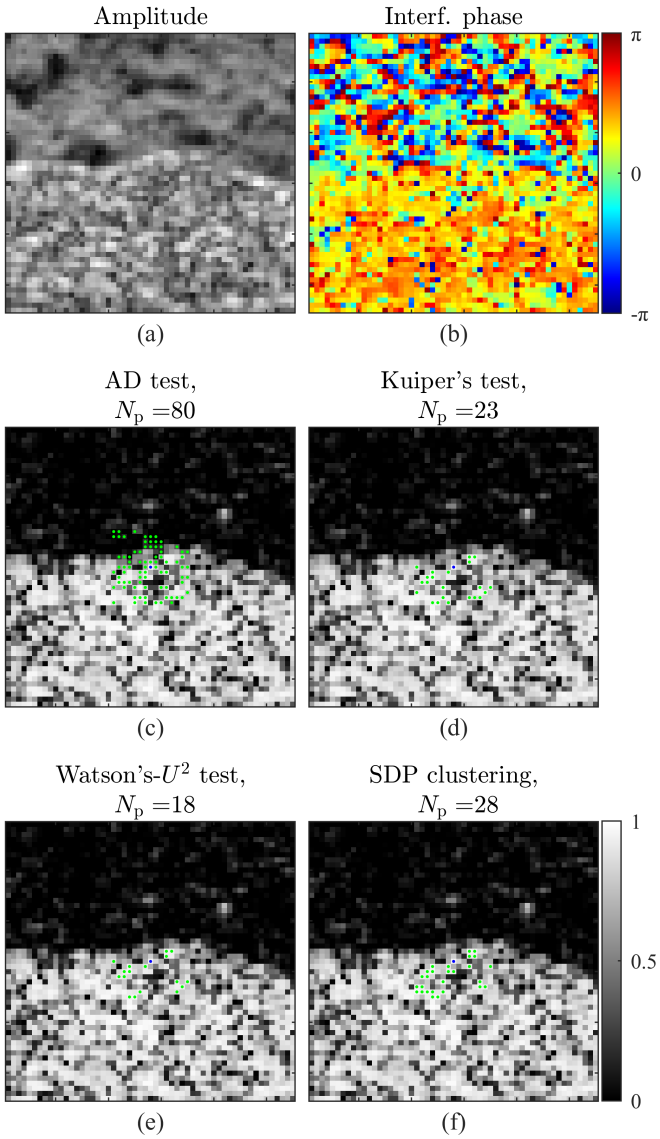

Fig. 10. Identified brother pixels of pixel P_2. (a) Time-averaged amplitude on a logarithmic scale, (b) one of the interferogram in the stack, and (c)–(f) brother pixels identified by four methods, with blue denoting the centra pixel and green denoting the identified brother pixel. The significant level used in (c)–(e) is 0.05. The base map of (c)–(f) is $\widehat{\alpha_1}$.

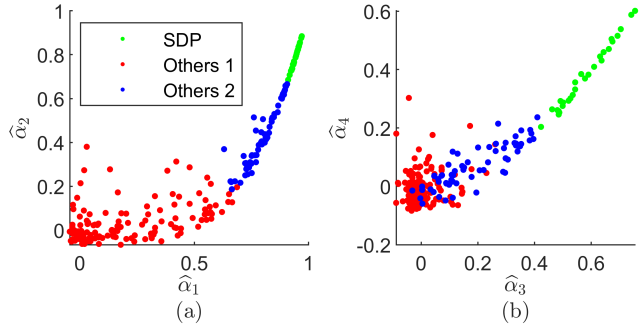

Fig. 11. Clustered scatterplot of the sample trigonometric moments of pixel P_2. (a) $\widehat{\alpha_1}$ and $\widehat{\alpha_2}$ and (b) $\widehat{\alpha_3}$ and $\widehat{\alpha_4}$.

We simulated 10 000 replications for varying R to evaluate $\widehat{R}_2$ and $\widehat{\kappa}_3$ (see Fig. 3). The sample interval of R is 0.02 and the range is $[0.01, 0.98]$. Fig. 3(a) and (c) shows that using $\widehat{R}_2$ and $\widehat{\kappa}_3$ can significantly reduce the estimation bias. When R is very small, both bias and RMSE are reduced by using $\widehat{R}_2$ and $\widehat{\kappa}_3$. However, the RMSE of $\widehat{R}_2$ and $\widehat{\kappa}_3$ is increased in the interval $R \in [0.1, 0.4]$.

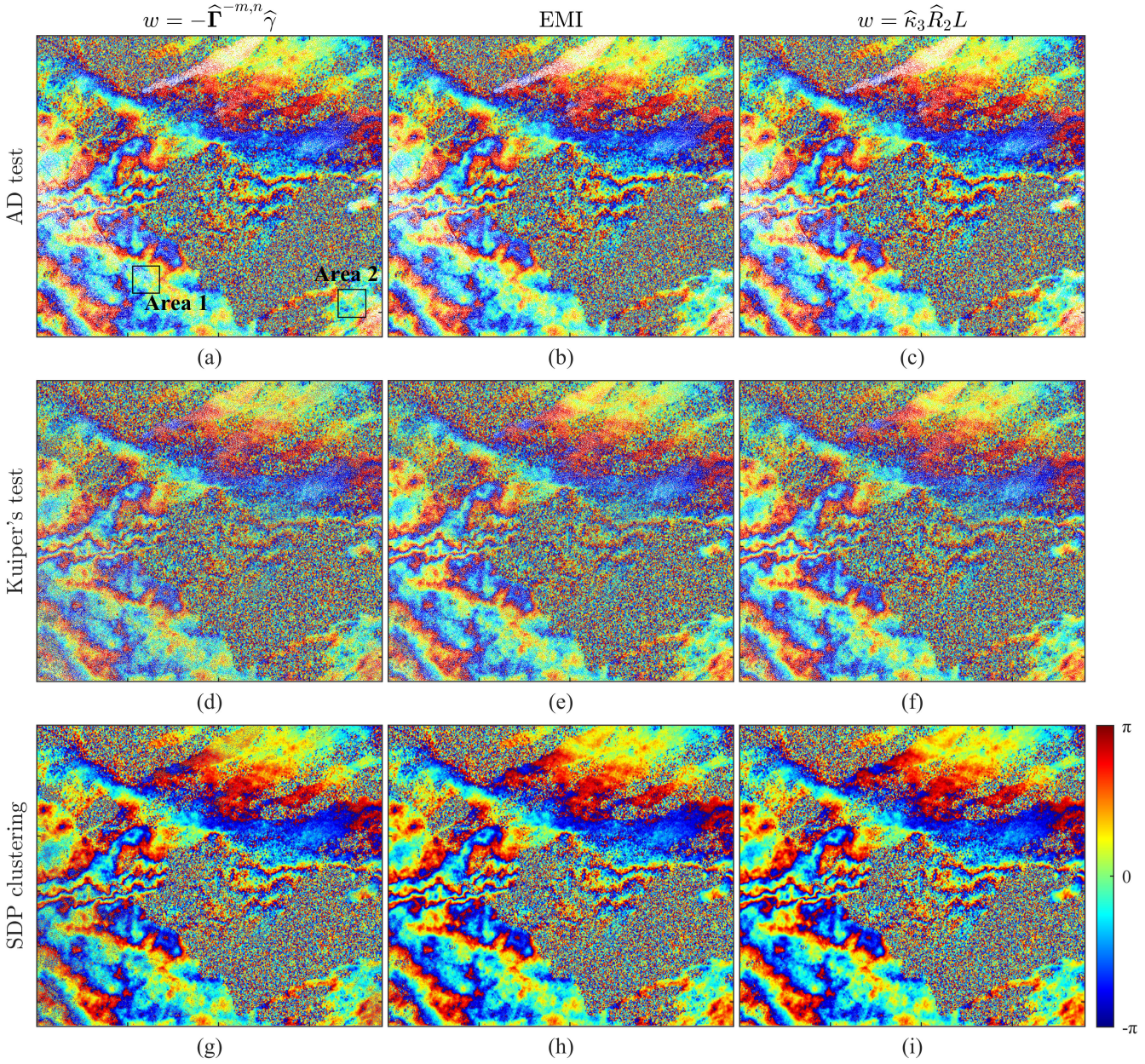

Fig. 12. Longest-temporal baseline interferograms reconstructed with consistent wrapped phases estimated by different ANS and PL methods. Different rows indicate different ANS methods and different columns indicate different PL methods. Rows 1 to 3 (top to bottom): AD test, Kuiper's test and SDP clustering, respectively. Columns 1 to 3 (left to right): PTA, EMI and the proposed PL method, respectively.

We demonstrate that the consistent phase estimation can benefit from our PL approach when the non-Gaussianity presents through simulated SLC stack [see Fig. 4(a)–(c)]. The signal $\mathbf{y}$ was simulated with a product model [see (2)]. The non-Gaussian component is formed with $z = (z_c \times z_t)^{1/2}$, where 1) z_c and z_t are independent; 2) z_c is time-invariant and z_t is realized independently in each scene; and 3) both z_c and z_t follow the Gamma distribution with shape and scale parameter ξ^{-1}, yielding $E[z_c] = E[z_t] = 1$ and $\text{var}[z_c] = \xi_c$, $\text{var}[z_t] = \xi_t$. The above settings result in the KM model presented in [3]. For the component x, which fits the CCG model, we assumed the same decorrelation process as in Fig. 2(b).

In Fig. 4(a)–(c), we compare the PTA method ($w = -\widehat{\Gamma}^{-m,n}\widehat{\gamma}$, where $\widehat{\Gamma}$ represents the sample coherence magnitude matrix which is real symmetric, and the superscript $^{-m,n}$ denotes the m row n column element in the inverse matrix $\widehat{\Gamma}^{-1}$), the eigendecomposition-based maximum-likelihood estimator of interferometric phase (EMI) [19] and our PL method ($w = \widehat{\kappa}_3\widehat{R}_2 L$). It is not surprising to see that in the Gaussian speckle case, the PTA and EMI methods have lower phase RMSE, since they are both designed for the CCG model and exploit the coupling between the interferometric amplitudes and phases according to (1). In this case, our method increased the phase RMSE by an average of 10%. However, when this coupling is distorted by the non-Gaussianity caused by the time-invariant component z_c, where $\text{var}[z_c] = 0.4$ corresponds to a small heterogeneity [3], our PL method reduced the phase RMSE by an average of 30% and 17% compared to PTA and EMI, respectively. We observed that the time-variant component z_t

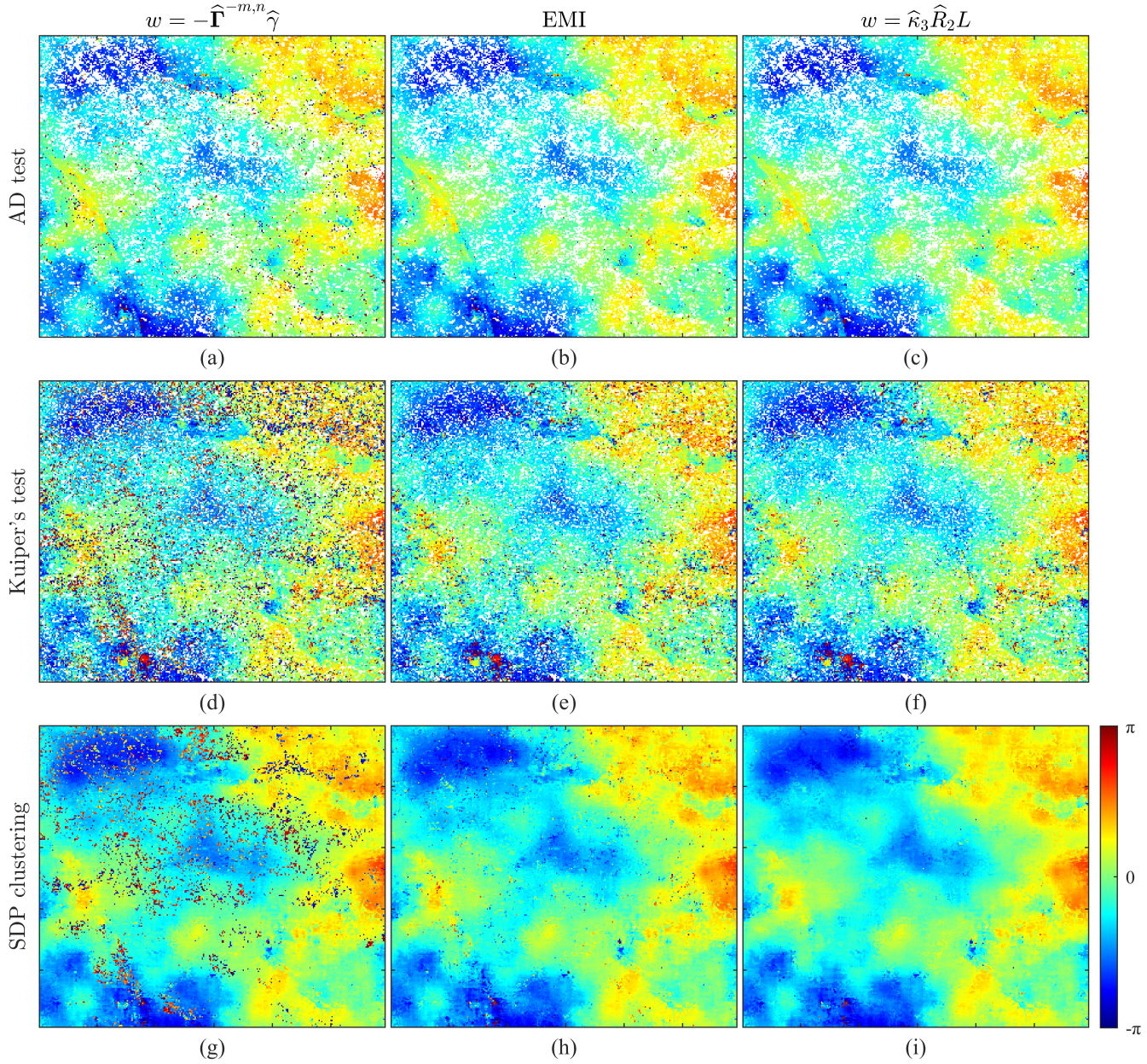

Fig. 13. Reconstructed longest-temporal baseline interferograms of area 1 identified in Fig. 12(a). Different rows indicate different ANS methods and different columns indicate different PL methods. Rows 1 to 3 (top to bottom): AD test, Kuiper's test and SDP clustering, respectively. Columns 1 to 3 (left to right): PTA, EMI and the proposed PL method, respectively.

has less effect on the PL, as our PL method can have the comparable accuracy with PTA and EMI when ξ_t expands to 0.6, which corresponds to a medium level of heterogeneity. We expect that the superiority of our PL method over CCG model-based methods to increase as the level of non-Gaussianity increases. In Fig. 4(d), PTA and EMI are modified into a phase-based version by replacing $\widehat{\gamma}$ with $\overline{R}$ [21], and three other phase-based estimators are tested, showing that our PL method is superior to other estimators when only the phase information is used.

V. Experiment on Real Data

A. Study Area and Data Processing

The selected study area is a region within a radius of 3.5 km situated in the Hawaii Volcanoes National Park as shown in Fig. 5. The various features make the area well suited to test the proposed method, especially in terms of the ANS. We collected $N = 19$ TerraSAR-X SAR images covering this area over the period from 8 December 2021 to 8 July 2022. The size of each co-registered image is 2720 pixels in range by 3800 pixels in azimuth. The time interval between each adjacent pair is either 11 or 22 days. We generated the interferograms with all possible combinations, yielding $M = N(N-1)/2 = 171$. The topographic and orbital fringes are subsequently removed from each interferogram. To fully evaluate the impact of the ANS and PL methods, we compared the consistent wrapped phases estimated from all possible combinations of ANS and PL. The ANS methods involved are: Anderson-Darling (AD) test, Kuiper's test, and the proposed SDP clustering method and the PL methods

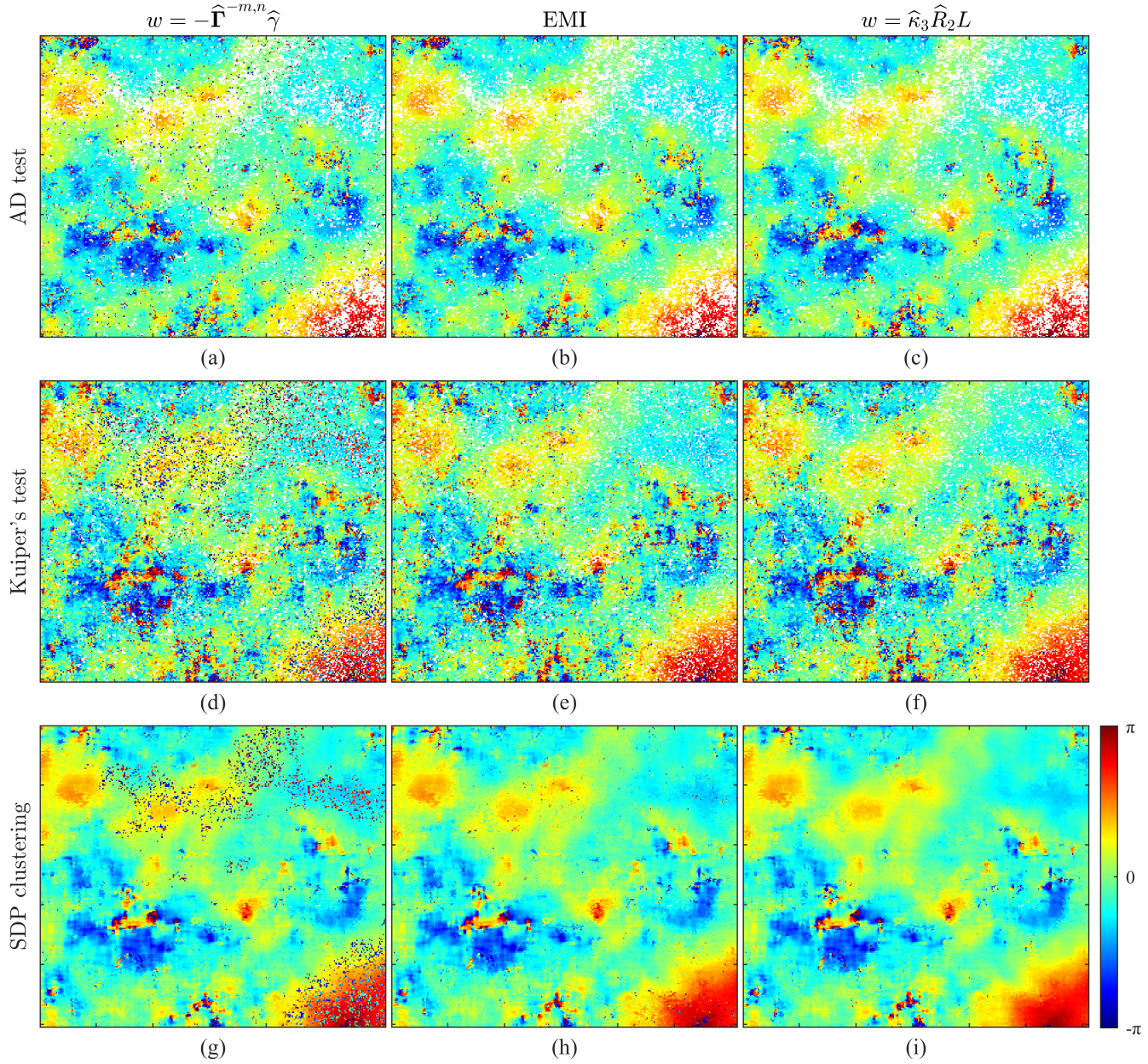

Fig. 14. Reconstructed longest-temporal baseline interferograms of area 2 identified in Fig. 12(a). Different rows indicate different ANS methods and different columns indicate different PL methods. Rows 1 to 3 (top to bottom): AD test, Kuiper's test and SDP clustering, respectively. Columns 1 to 3 (left to right): PTA, EMI and the proposed PL method, respectively.

involved are: PTA, EMI, and using the proposed PL weights $\widehat{w} = \widehat{\kappa}_3\widehat{R}_2 L$. The window size is 15 × 15 for all ANS and PL processing. Pixels with <20 brother pixels were discarded, and displayed as blank in the reconstructed interferograms. The deformation are not retrieved, since we focus on the impact on the consistent wrapped phase estimation. Fig. 6 is the data processing flowchart.

B. Results and Discussion

Fig. 7 shows the number of brother pixels identified by different selection methods. It can be seen that the number of brother pixels identified by all methods has a visible correlation with the SAR and optical texture. However, the texture in the lower center right of the SAR and optical images is missing from Fig. 7(c), which may be due to the poor phase quality in these areas, preventing the trigonometric moments from forming distinct clusters.

Fig. 8 shows an example of identified brother pixels of pixel P_1 [see Fig. 5(a)] by different approaches. We also tested the Watson's-U^2 test on this pixel. The first-order sample moment $\widehat{\alpha}_1$ acts as a phase stability indicator like temporal coherence, and the two distinct gray levels in the base map of Fig. 8(c)–(f) indicate two different scattering mechanisms. We can see that the AD test [38] working on the amplitude time series confuses these two mechanisms [see Fig. 8(c)], which could be explained either by the insufficient power of the test or by the deviation of the amplitude statistics from the phase statistics. The Kuiper's test and Watson's-U^2 test working on ψ_{res} show very similar results, and did not confuse the mechanisms, but resulted in a significant reduction of the

sample size as many pixels inside the boundary were rejected [see Fig. 8(d) and (e)]. Our clustering can both distinguish different decorrelation mechanisms and preserve sufficient samples as shown in Fig. 8(f). It should be noted that although our SDP clustering approach outperforms other approaches in both the simulated experiment and a real-data case, the failure of approaches based on circular two-sample statistical tests is different. In this real data case, these approaches show a lack of ability to include a sufficient number of brother pixels, rather than an inability to exclude wrong pixels as shown in the simulated experiment. Fig. 9 shows that pixels can form different clusters according to their moments, which allows us to separate them using clustering methods such as K-means. In this example, the pixels around P_1 have been split into two clusters, one with higher phase stability (as indicated by the higher $\widehat{\alpha}_1$) and the other with lower. Fig. 10 shows another example of pixel P_2, but the surrounding pixels have been split into three clusters with high, medium, and low phase stability, respectively, as shown in Fig. 11. In Fig. 10, we can see that in this example, the different scattering mechanisms are again confused by the ANS method based on the AD test and the pixel selected by different phase-based methods are similar.

Fig. 12 shows the reconstructed wrapped interferograms of the longest temporal baseline (242 days). To compare the results of different methods more clearly and visibly, two representative areas boxed in Fig. 12(a) are selected for focused views as shown in Figs. 13 and 14. In the subgraph (i) of Figs. 12–14, we can see that the combination of the proposed ANS and PL approaches obtained the phases with minimized noise. These results demonstrate the validity of the two proposed methods, and indicate that the selected samples can fit the stochastic model assumed in the PL approach. The noticeable noise in the subgraphs (g) and (h) of Figs. 12–14 indicates that applying the proposed SDP clustering method alone cannot guarantee an improved result. The reason is that the proposed SDP clustering approach aims to select pixels with similar phase noise components, which cannot guarantee that the samples have similar CCG parameters. However, the proposed PL method can accommodate different types of samples, as we can see that the phases retrieved by our PL method always have minimized noise, regardless of how the samples are selected. This is due to the compatibility of the VM model with the CCG phase model, as shown in Fig. 1. Since the proposed PL method requires an iterative solution, the computational cost of our PL method is comparable to that of PTA, except that there is no matrix inversion in our PL method, and it does not suffer from the ill-condition. The numerical solution of our PL method can also benefit from an initialization closer to its optima [19], for example, through EMI.

VI. Conclusion

In this article, we present a novel ANS method and a novel PL method that focus on the exploitation of interferometric phase statistics. These methods are based on a simple but generalized and practical interferometric signal model that does not rely on the CCG assumption: the interferometric phases are the only observation and the interferometric amplitudes are discarded. The circular statistics of mathematical science is the foundation of the two proposed methods. The proposed ANS method does not rely on any particular phase distribution, and the proposed PL method assumes that the interferometric phases are von Mises distributed and uncorrelated between different interferograms. Simulated and real data results preliminarily show the interest and superiority of the proposed methods and the model used. Future work will focus on developing the quality number and exploiting the correlation between interferograms in terms of circular statistics.

Acknowledgment

The authors would like to thank the editors and the anonymous reviewers for their helpful comments. The TerraSAR-X data were provided by the German Aerospace Center (DLR) via the super sites initiative.

References

[1] J. Eppler and B. T. Rabus, "Adapting InSAR phase linking for seasonally snow-covered terrain," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4305313.

[2] A. Ferretti, A. Fumagalli, F. Novali, C. Prati, F. Rocca, and A. Rucci, "A new algorithm for processing interferometric data-stacks: SqueeSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3460–3470, Sep. 2011.

[3] S. Zwieback and F. J. Meyer, "Repeat-pass interferometric speckle," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6736–6750, Aug. 2021.

[4] M. Even and K. Schulz, "InSAR deformation analysis with distributed scatterers: A review complemented by new advances," *Remote Sens.*, vol. 10, no. 5, p. 744, May 2018.

[5] S. Samiei-Esfahany, J. E. Martins, F. van Leijen, and R. F. Hanssen, "Phase estimation for distributed scatterers in InSAR stacks using integer least squares estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5671–5687, Oct. 2016.

[6] C. Wang et al., "A new likelihood function for consistent phase series estimation in distributed scatterer interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5227314.

[7] A. Pepe, Y. Yang, M. Manzo, and R. Lanari, "Improved EMCF-SBAS processing chain based on advanced techniques for the noise-filtering and selection of small baseline multi-look DInSAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4394–4417, Aug. 2015.

[8] A. B. Narayan, A. Tiwari, R. Dwivedi, and O. Dikshit, "A novel measure for categorization and optimal phase history retrieval of distributed scatterers for InSAR applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5843–5849, Oct. 2018.

[9] A. Pepe, P. Mastro, and C. E. Jones, "Adaptive multilooking of multitemporal differential SAR interferometric data stack using directional statistics," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6706–6721, Aug. 2021.

[10] D.-X. Yue, F. Xu, A. C. Frery, and Y.-Q. Jin, "Synthetic aperture radar image statistical modeling: Part one-single-pixel statistical models," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 1, pp. 82–114, Mar. 2021.

[11] R. F. Hanssen, *Radar Interferometry: Data Interpretation and Error Analysis*, vol. 2. Cham, Switzerland: Springer, 2001.

[12] J.-S. Lee, K. W. Hoppel, S. A. Mango, and A. R. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 5, pp. 1017–1028, Mar. 1994.

[13] P. V. H. Vu, A. Breloy, F. Brigui, Y. Yan, and G. Ginolhac, "Robust phase linking in InSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5211111.

[14] K. V. Mardia, P. E. Jupp, and K. Mardia, *Directional Statistics*, vol. 2. Hoboken, NJ, USA: Wiley, 2000.

[15] E. Rodriguez and J. M. Martin, "Theory and design of interferometric synthetic aperture radars," *IEE Proc. F Radar Signal Process.*, vol. 139, no. 2, pp. 147–159, Apr. 1992.

[16] S. H. Yueh, J. A. Kong, J. K. Jao, R. T. Shin, and L. M. Novak, "K-distribution and polarimetric terrain radar clutter," *J. Electromagn. Waves Appl.*, vol. 3, no. 8, pp. 747–768, 1989.

[17] M. Jiang, X. Ding, and Z. Li, "Hybrid approach for unbiased coherence estimation for multitemporal InSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2459–2473, May 2014.

[18] Y. Wang and X. X. Zhu, "Robust estimators for multipass SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 968–980, Feb. 2016.

[19] H. Ansari, F. De Zan, and R. Bamler, "Efficient phase estimation for interferogram stacks," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 4109–4125, Jul. 2018.

[20] K. D. Ward, "Compound representation of high resolution sea clutter," *Electron. Lett.*, vol. 17, no. 16, pp. 561–563, Aug. 1981.

[21] C. Lopez-Martinez and X. Fabregas, "Modeling and reduction of SAR interferometric phase noise in the wavelet domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 12, pp. 2553–2566, Dec. 2002.

[22] T. Hillen, K. J. Painter, A. C. Swan, and A. D. Murtha, "Moments of von Mises and Fisher distributions and applications," *Math. Biosci. Eng.*, vol. 14, no. 3, pp. 673–694, 2017.

[23] U. J. Lund, *Regression and Goodness of Fit for Directional Data*. Santa Barbara, CA, USA: Univ. California, 1998.

[24] A. Hooper, H. Zebker, P. Segall, and B. Kampes, "A new method for measuring deformation on volcanoes and other natural terrains using InSAR persistent scatterers," *Geophys. Res. Lett.*, vol. 31, no. 23, pp. L23611–L23615, Dec. 2004.

[25] T. Wehrly and E. Shien, "Influence curves of estimators for directional-data," *Biometrika*, vol. 68, no. 1, pp. 334–335, 1981.

[26] N. H. Kuiper, "Tests concerning random points on a circle," *Indagationes Mathematicae Proc.*, vol. 63, no. 1, pp. 38–47, 1960.

[27] G. S. Watson, "Goodness-Of-Fit tests on a circle," *Biometrika*, vol. 48, nos. 1–2, p. 109, Jun. 1961.

[28] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 129–137, Mar. 1982.

[29] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[30] U. von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, Dec. 2007.

[31] F. Rocca, "Modeling interferogram stacks," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3289–3299, Oct. 2007.

[32] S. Zwieback et al., "A statistical test of phase closure to detect influences on DInSAR deformation estimates besides displacements and decorrelation noise: Two case studies in high-latitude regions," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5588–5601, Sep. 2016.

[33] A. M. Guarnieri and S. Tebaldini, "On the exploitation of target statistics for SAR interferometry applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3436–3443, Nov. 2008.

[34] N. Cao, H. Lee, and H. C. Jung, "Mathematical framework for phase-triangulation algorithms in distributed-scatterer interferometry," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 9, pp. 1838–1842, Sep. 2015.

[35] S. Li et al., "An adaptive weighted phase optimization algorithm based on the sigmoid model for distributed scatterers," *Remote Sens.*, vol. 13, no. 16, p. 3253, Aug. 2021.

[36] M. A. Stephens, "Tests for the von Mises distribution," *Biometrika*, vol. 56, no. 1, p. 149, Mar. 1969.

[37] D. J. Best and N. I. Fisher, "The BIAS of the maximum likelihood estimators of the von mises-Fisher concentration parameters," *Commun. Statist. Simul. Comput.*, vol. 10, no. 5, pp. 493–502, Jan. 1981.

[38] F. W. Scholz and M. A. Stephens, "K-sample Anderson-darling tests," *J. Amer. Stat. Assoc.*, vol. 82, no. 399, p. 918, Sep. 1987.