

Letter

Identification of Statistically Homogeneous Pixels Based on One-Sample Test

Keng-Fan Lin * and Daniele Perissin

Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA; perissin@purdue.edu

* Correspondence: francislin@purdue.edu; Tel.: +1-949-629-8180

Academic Editors: Xiaofeng Li and Prasad S. Thenkabail

Received: 21 October 2016 ; Accepted: 29 December 2016; Published: 4 January 2017

Abstract: Statistically homogeneous pixels (SHP) play a crucial role in synthetic aperture radar (SAR) analysis. In past studies, various two-sample tests were applied on multitemporal SAR data stacks under the assumption of having stationary backscattering properties over time. In this letter, we propose the *Robust T-test* (TR) to improve the effectiveness of test operation. The TR test reduces the impact of temporal variabilities and outliers, thus helping to identify SHP with assurances of similar temporal behaviors. This method includes three steps: (1) signal suppression; (2) outlier removal; and (3) one-sample test. In the experiments, we apply the TR test on both simulated and real data. Different stack sizes, types of distributions, and hypothesis tests are compared. Results of both experiments signify that the TR test outperforms conventional approaches and provides reliable SHP for SAR image analysis.

Keywords: statistically homogeneous pixels; outlier removal; one-sample test

1. Introduction

Statistically homogeneous pixels (SHP) represent pixels that share similar statistical characteristics over time. In the Synthetic Aperture Radar (SAR) community, using a stack of images to identify SHP in space has been extensively used for various SAR applications, such as adaptive filtering [1–6] (multitemporal configurations) [7,8] (single-temporal configurations), complex coherence estimation [9–12], and image segmentation [13]. As processing images based on SHP precludes the participation of irrelevant observations, successful SHP identification plays a key role in the SAR image analysis.

A generic approach for SHP identification is through hypothesis tests. Pixels showing no significant differences from each other will be incorporated into the same SHP family. In general, the hypothesis tests can be divided into two categories: *nonparametric* and *parametric* tests. Nonparametric tests perform acceptably over a wide range of distributions as they do not assume the distribution of the data. However, these tests only utilize part of the information provided by the sample [14], which makes them less effective than the parametric ones. To attain the same *power* (power is the probability of correctly rejecting the null hypothesis.), nonparametric tests usually require larger sample sizes than the parametric tests. In contrast, parametric tests are usually defined based on certain assumptions. They are expected to perform better when the assumptions hold [2].

Statistically speaking, the effectiveness of the hypothesis tests relies on how the assumptions are in line with the data. Nonparametric tests can provide relatively robust results when no appropriate assumptions can be made, while parametric tests can perform effectively when the data characteristics are in conformity with the tests. Indeed, SAR amplitude was commonly assumed to be *Rayleigh* distributed (several theoretical models have also been suggested, such as Nakagami, Gamma, K , and G^0 distributions [15–18].); however, empirically non-Rayleigh behaviors were often observed in past

studies [19]. To avoid incorrect presumptions of SAR data characteristics, nonparametric two-sample tests (e.g., Kolmogorov–Smirnov (KS) and Anderson–Darling (AD) tests) were more extensively used than the parametric ones for SHP identification [2,9,10]. As these tests rely on the stationarity of the backscattering properties, they may fail in the presence of temporal variations or outliers.

In theory, these hypothesis tests handle only stochastic processes. Any inclusions of meaningful signals (e.g., temporal changes) will be interpreted as the stochastic processes and potentially hamper the effectiveness of the hypothesis testing. Standard tests usually assume that the backscattering properties remain stationary over time. Although reducing the sample size may help the data characteristics meet this assumption, committing type I and II errors becomes an undesired trade-off [14]. As there is no clear winner between these choices, we should develop an approach that improves the robustness of the test operation while keeping the large sample size.

In this letter, we propose the *Robust T-test* (TR) to reduce the impact of temporal variabilities and outliers. The TR test helps identify SHP with assurances of similar temporal behaviors, which improves the reliability of SHP in SAR analysis.

2. Methodology

In the following section, we introduce the procedure of the TR test based on the multitemporal configuration. As illustrated in Figure 1, this method is comprised of three steps: (1) signal suppression; (2) outlier removal; and (3) one-sample test.

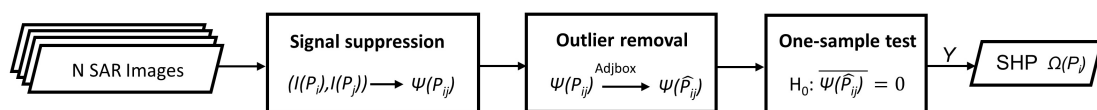


Figure 1. Research flowchart. The main procedures include three steps: (1) signal suppression; (2) outlier removal; and (3) one-sample test.

2.1. Signal Suppression

Conventionally, two-sample tests were directly applied to pairs of time series to determine whether or not each pair was retrieved from the same distribution. In such a case, the test operation deals with signals and stochastic processes concurrently. The handling of non-random signals potentially impedes the effectiveness of the test procedure. To enhance the performance of the hypothesis testing, one should remove signals from the data before any test operations are applied. In this study, we propose a strategy to eliminate signals.

Assume N SAR images were acquired in the same geometry and have been well coregistered. The amplitude time series of a pixel P can be represented as

$$I(P) = [I_1(P), I_2(P), \dots, I_N(P)], \quad (1)$$

where $I(P)$ is a sample comprised of temporally ordered N amplitude values from 1st to N th images.

To detail the theoretical basis of signal suppression, we first analyze the components of a temporal observation $I_t(P)$. We will ignore the additive systematic errors in the following discussions and consider the observed amplitude as the product of speckles and scene signals [20]

$$I_t(P) = I_t^n(P) I_t^S(P), \quad (2)$$

where the superscripts n and S indicate the speckle noise and the scene signal, respectively.

Assume two samples, denoted as $I(P_i)$ and $I(P_j)$, need to be tested. Instead of applying a two-sample test on $I(P_i)$ and $I(P_j)$ directly, we take the logarithm on each of their temporal observations to transform the multiplicative noise into an additive one [21]

$$\begin{aligned}\psi_t(P) &= \ln(I_t(P)) \\ &= \ln(I_t^n(P)) + \ln(I_t^s(P)),\end{aligned}\quad (3)$$

and then subtract $\psi_t(P_j)$ from $\psi_t(P_i)$ for each of the temporal observations, acquiring a new sample

$$\psi(P_{ij}) = \psi(P_i) - \psi(P_j), \quad (4)$$

where $\psi(P_{ij})$ is a N-element time series (vector) that represents the difference between $\psi(P_j)$ and $\psi(P_i)$.

If the original two samples $I(P_i)$ and $I(P_j)$ are statistically homogeneous, most signals in the observations will be canceled after Equation (4) is applied (so as the original distribution in the data). In other words, only random errors should remain in $\psi(P_{ij})$. Since random errors are usually assumed Gaussian distributed with the mean equal to zero, the result leads to

$$\mathbf{E}(\psi(P_{ij})) = 0. \quad (5)$$

2.2. Outlier Removal

Considering that abnormal values will influence the test operation, an outlier removal process is applied to improve the effectiveness of the hypothesis testing. In past studies, a standard approach for outlier removal was through *boxplot* [22], which was effective when the data were normally distributed. Since the amplitudes are usually considered as skewed distributions [19], $\psi(P_{ij})$ might tend to be skewed (especially for small sample size). We therefore apply *adjusted boxplot* [23] to effectively remove abnormal values in different situations.

The adjusted boxplot utilizes *medcouple* (MC) [24] as a robust measure to the skewness of a distribution. Consider $\tilde{\psi}(P_{ij})$ as a sorted sample of $\psi(P_{ij})$, where $\tilde{\psi}_1(P_{ij}) \leq \tilde{\psi}_2(P_{ij}) \leq \dots \leq \tilde{\psi}_N(P_{ij})$. We first define a kernel function h for all $\tilde{\psi}_a(P_{ij}) \neq \tilde{\psi}_b(P_{ij})$

$$h(\tilde{\psi}_a(P_{ij}), \tilde{\psi}_b(P_{ij})) = \frac{(\tilde{\psi}_b(P_{ij}) - Q_2) - (Q_2 - \tilde{\psi}_a(P_{ij}))}{\tilde{\psi}_b(P_{ij}) - \tilde{\psi}_a(P_{ij})}, \quad (6)$$

where Q_2 is the median of $\tilde{\psi}(P_{ij})$. For the special case that $\tilde{\psi}_a(P_{ij}) = \tilde{\psi}_b(P_{ij}) = Q_2$, the kernel h is defined in a different way (see details in [23]). In real SAR applications, the special case may not play a role due to the presence of measuring errors.

The MC of $\psi(P_{ij})$ is defined as the median of all $h(\tilde{\psi}_a(P_{ij}), \tilde{\psi}_b(P_{ij}))$ values for which $\tilde{\psi}_a(P_{ij}) \leq Q_2 \leq \tilde{\psi}_b(P_{ij})$. Consequently, observations not within the interval $[L, U]$ will be considered as outliers

$$[L, U] = \begin{cases} [Q_1 - 1.5e^{-4MC}(Q_3 - Q_1), \\ Q_3 + 1.5e^{3MC}(Q_3 - Q_1)], & MC \geq 0, \\ [Q_1 - 1.5e^{-3MC}(Q_3 - Q_1), \\ Q_3 + 1.5e^{4MC}(Q_3 - Q_1)], & MC < 0, \end{cases} \quad (7)$$

where Q_1 and Q_3 are the first and third quartiles of $\tilde{\psi}(P_{ij})$.

The adjusted boxplot does not require any assumptions on the distributional type. Therefore, it is adaptive to various distributions. The effectiveness of adjusted boxplot on amplitude data can be referred to [9]. In this study, we apply this procedure on $\psi(P_{ij})$ to remove abnormal values prior to the hypothesis testing. We denote the sample after the outlier removal process as $\psi(\hat{P}_{ij})$.

2.3. One-Sample Test

After the signal suppression and outlier removal processes have been performed, a *one-sample t-test* is applied on $\psi(\hat{P}_{ij})$. One-sample *t-test* compares the sample mean to a known population mean without knowing the variability/dispersion of the population. It has been considered as a *robust procedure* among the parametric tests. Slight departures from normality may only result in minor deviations from the ideal. Therefore, the *p-value* quoted may be slightly in error if there is a moderate violation of the normality assumption [14]. Since random errors are considered normally distributed, the *t-test* becomes a suitable measure to test if $\psi(\hat{P}_{ij})$ mainly contains random errors

$$\begin{cases} H_0 : \bar{\mu} = 0, \\ H_1 : \bar{\mu} \neq 0, \end{cases} \quad (8)$$

where $\bar{\mu}$ is the population mean of $\psi(\hat{P}_{ij})$, and the corresponding *t*-statistic can be computed through

$$t = \frac{\mu}{s/\sqrt{n}}, \quad (9)$$

where μ , s and n are the sample mean, standard deviation, and sample size of $\psi(\hat{P}_{ij})$.

As a consequence, two pixels P_i and P_j will be considered as SHP only if $\psi(\hat{P}_{ij})$ contains no meaningful signals. Handling of only the stochastic processes potentially improves the effectiveness of the hypothesis testing. Under the proposed procedure, SHP of a generic pixel P (denoted as $\Omega(P)$) can be identified regardless of meaningful signals or outliers.

3. Experiments and Discussion

As previously mentioned, we apply the TR test on both simulated and real data. Assuming the existence of abnormal values in the data, we compare the TR test with other standard tests based on different hypothesis tests and stack sizes. Decoupling of outlier rejection from TR-test is not considered in this study.

3.1. Monte Carlo Simulation

Monte Carlo simulation is used for validating the proposed procedure. We simulate two samples based on inverse transform sampling under the desired distribution. A unit mean gamma distribution [16] is applied to the samples to serve as the speckle effect. We consider the two samples as the temporal observations of two pixels. For each of the following four scenarios, different stack sizes, distribution types, and hypothesis tests are analyzed:

(i) **Temporal change: No; Outlier: No**

The two samples possess neither temporal changes nor outliers. Three types of distributions are analyzed. The distributional parameters are shown in Table 1.

(ii) **Temporal change: No; Outlier: Yes**

The two samples possess outliers without temporal changes. The distributional parameters are the same as Table 1; however, both samples include 5% outliers. The magnitudes of outliers are set to be $\mu + 5\sigma$, where μ and σ represent the mean and standard deviation of the sample.

(iii) **Temporal change: Yes; Outlier: No**

A temporal change is designed in the first sample. The time of change is always at $N/2$ for different sample sizes, meaning that observations before and after $N/2$ follow different distributional parameters. The corresponding parameters are shown in Table 2.

(iv) **Temporal change: Yes; Outlier: Yes**

A temporal change is designed in the first sample. The distributional parameters are the same as Table 2; however, both samples include 5% outliers. The magnitudes of outliers are set to be $\mu + 5\sigma$.

Table 1. Distributional parameters for case (i) and (ii). λ , k , θ , Ω , m , u , and σ represent the scale, shape, scale, spread, shape parameters, log mean, and log standard deviation, respectively.

	Rayleigh	Gamma	Nakagami	Lognormal	Inverse Gaussian	Exponential
Sample 1	$\lambda = 0.20$	$k = 1.00$ $\theta = 0.20$	$\Omega = 1.00$ $\mu = 0.20$	$\sigma = 1.00$ $\mu = 0.20$	$m = 1.00$ $\mu = 0.20$	$\mu = 1.00$
Sample 2	$\lambda = 0.24$	$k = 1.00$ $\theta = 0.26$	$\Omega = 1.00$ $\mu = 0.25$	$\sigma = 1.00$ $\mu = 0.50$	$m = 1.00$ $\mu = 0.23$	$\mu = 1.50$

Table 2. Distributional parameters for case (iii) and (iv). λ , k , θ , Ω , m , u , and σ represent the scale, shape, scale, spread, shape parameters, log mean, and log standard deviation, respectively. Subscripts b and a indicate the parameters before and after the temporal change, respectively.

	Rayleigh	Gamma	Nakagami	Lognormal	Inverse Gaussian	Exponential
Sample 1	$\lambda_b = 0.20$ $\lambda_a = 0.24$	$k = 1.00$ $\theta_b = 0.20$ $\theta_a = 0.26$	$\Omega = 1.00$ $\mu_b = 0.20$ $\mu_a = 0.25$	$\sigma = 1.00$ $\mu_b = 0.20$ $\mu_a = 0.50$	$m = 1.00$ $\mu_b = 0.20$ $\mu_a = 0.23$	$\mu_b = 1.00$ $\mu_a = 1.50$
Sample 2	$\lambda = 0.24$	$k = 1.00$ $\theta = 0.26$	$\Omega = 1.00$ $\mu = 0.25$	$\sigma = 1.00$ $\mu = 0.50$	$m = 1.00$ $\mu = 0.23$	$\mu = 1.50$

In this experiment, we compare the proposed method with three types of nonparametric tests, including the KS, AD, and *Cramèr–von Mises* (CM) tests. Also, two parametric tests, including Bhattacharyya distance (under Gaussian assumption) and generalized likelihood test (GLRT) (under Rayleigh assumption) are compared. The significance level is set to be 1% for all tests, and the number of simulations is 10,000 for all cases. As the distributional parameters are known, the power can be computed accordingly, as shown in Figures 2–5.

In case (i), where no temporal changes or outliers are present, the TR test provides higher power than all other tests over different distributions. Although GLRT performs well under the Rayleigh distribution among the parametric/nonparametric tests, it is less effective than the proposed TR test. The results signify the importance of the signal suppression (even if the signals are simply related to the primitive backscattering energy rather than the changes/outliers), as it helps the test deal only with the stochastic processes.

In case (ii), where outliers exist in both samples, the TR test still performs well thanks to the adjusted boxplot. Conversely, other tests show less effectively than they had in case (i) due to the absence of outlier removal procedure. The results demonstrate the importance of removing abnormal values prior to the test operation.

In case (iii), where we have temporal changes in the first sample, the TR test shows discriminative power compared to the other tests. The signal suppression becomes pivotal especially in scenes with temporal changes. These changes contain meaningful signals that can seriously bias the hypothesis testing. As the TR test applies the signal suppression before the hypothesis testing, it removes the signals from both backscattering energy and temporal changes, thus augmenting its test power.

In case (iv), where samples include both temporal changes and outliers, the TR test provides reliable results. As aforementioned, the TR test suppresses signals and outliers in the measurements and deals with the signal randomness. These procedures make the SHP identification not only adaptive to the temporal changes, but also resistant to the abnormal values in the observations.

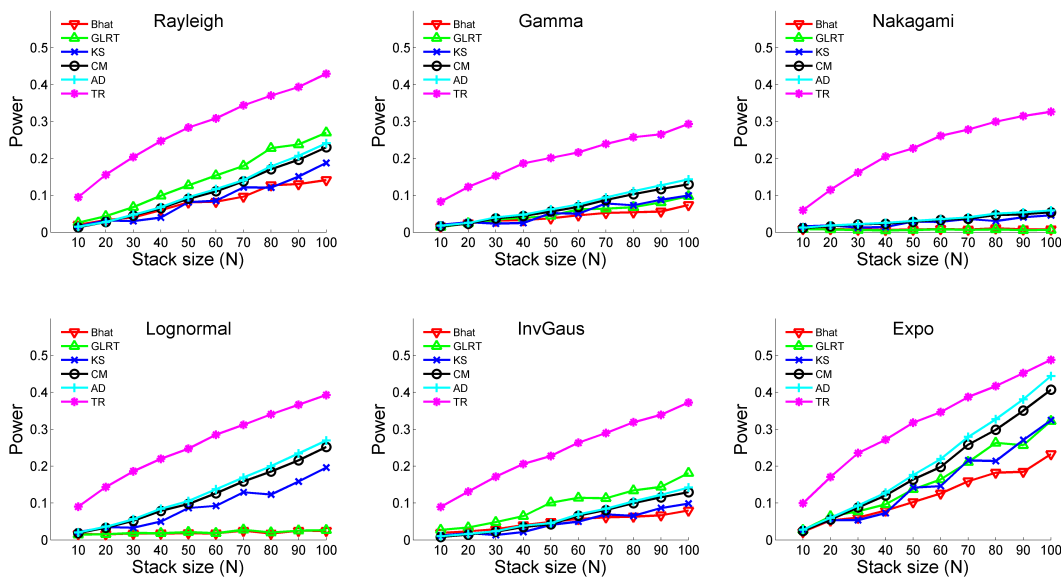


Figure 2. The power of hypothesis tests in case (i).

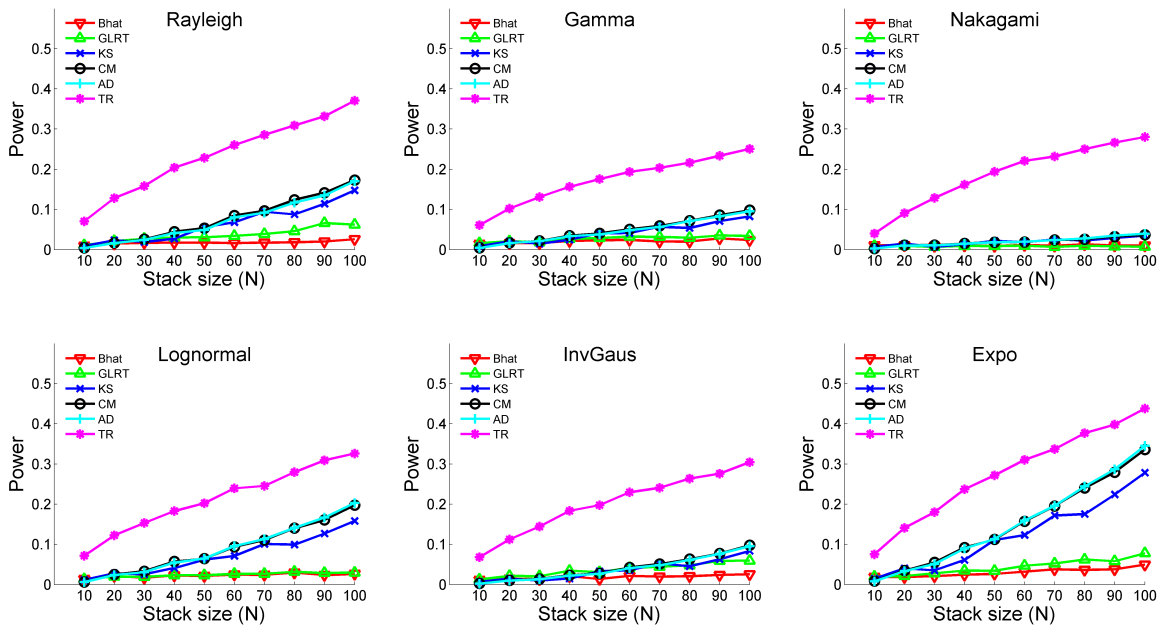


Figure 3. The power of hypothesis tests in case (ii).

Overall, the TR test reveals discriminative power in different distributions when compared with other tests. In the parametric tests, GLRT works well under Rayleigh scenario in case (i), which is as expected. Bhattacharyya criterion fails in most cases because the normality does not exist in all of the simulated scenarios. In the nonparametric tests, the AD and CM tests provide similar power in most cases. Both conventional parametric and nonparametric tests fail to provide reliable results compared to the TR test because they are not capable of taking meaningful signals into account.

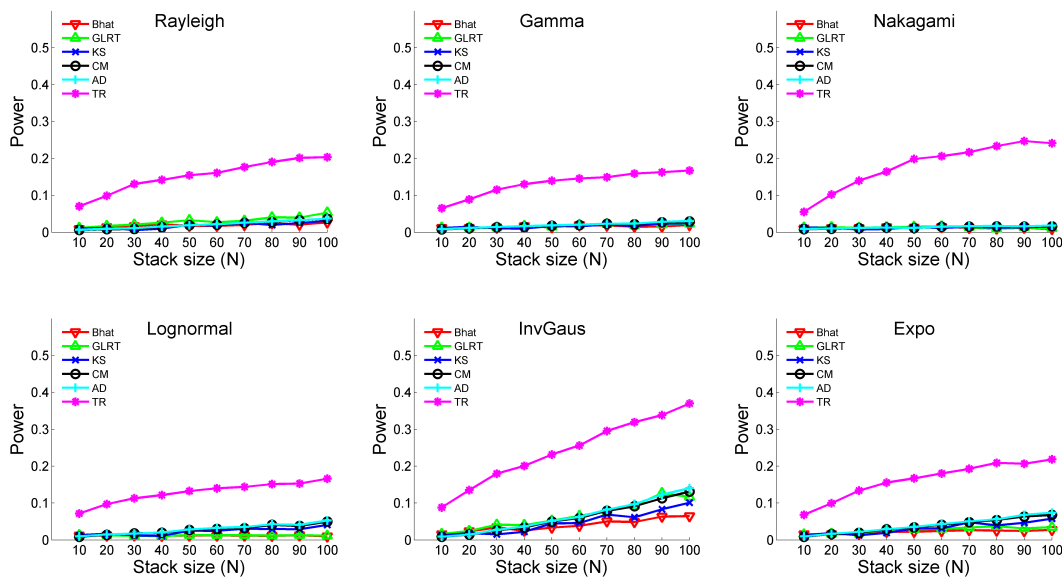


Figure 4. The power of hypothesis tests in case (iii).

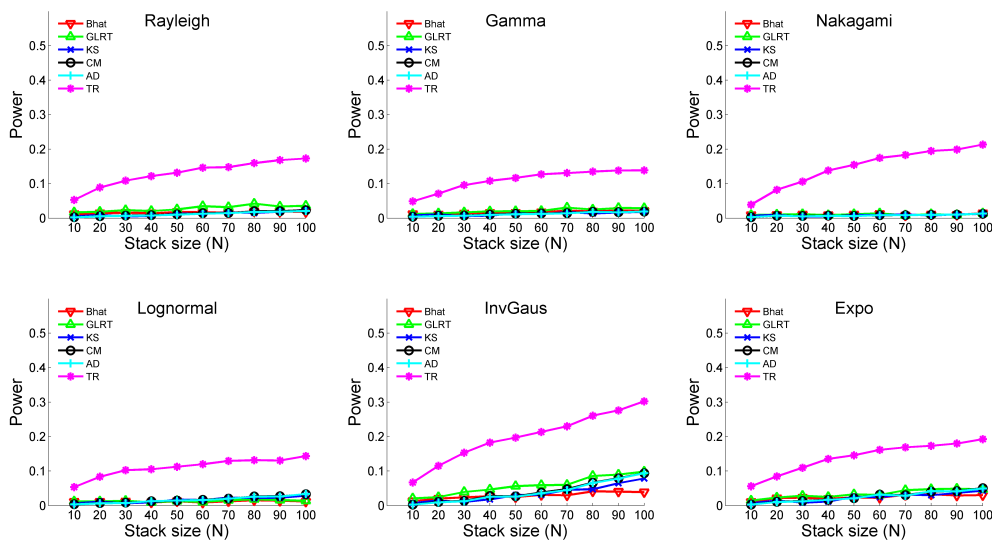


Figure 5. The power of hypothesis tests in case (iv).

3.2. Experiments with SAR Data Stack

To detail the effectiveness of the proposed TR test in real applications, we select a harbor area in Hong Kong as our test site. The SAR Data stack was acquired at $24^{\circ}19'49.24''N$, $114^{\circ}8'0.16''E$ with approximately $2.5 \text{ km} \times 2.5 \text{ km}$ spatial coverage (image size: 1500×1250 pixels), as shown in Figure 6. This data stack consists of seventy-five TerraSAR-X images acquired in the ascending orbit with the VV (Vertical-transmit-Vertical-receive) polarization. We select Area 1 and Area 2 for detailed comparisons as these areas involved several human activities (e.g., transport of containers) in the entire time span. Since these temporal changes may alter the backscattering signals received by the antenna, building up a single model for the parametric tests is not a simple task. Therefore, we compare the TR test with the AD test to show fundamental differences between the proposed and the conventional approaches. As previously mentioned, SHP can be utilized in various SAR applications. In this experiment, we use amplitude denoising as an example to verify the proposed method.

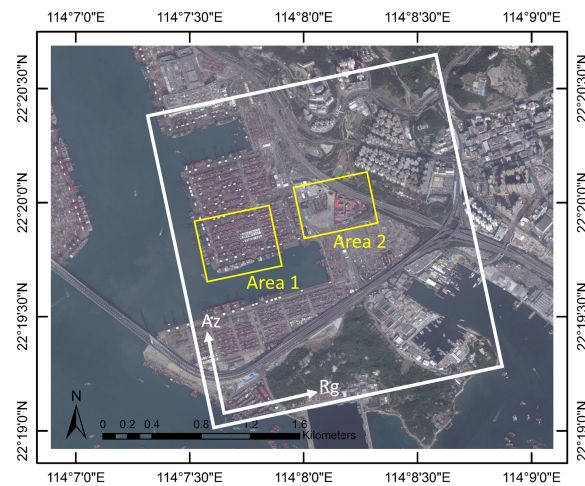


Figure 6. The area of interest (white box) in Hong Kong retrieved from World Imagery (Esri, Redlands, CA, USA). Two areas were selected for detailed comparisons.

Figure 7 shows the number of SHP identified at each pixel using different tests and stack sizes. Figure 8 shows the histograms corresponding to Figure 7. Results in Figure 7a–c indicate that the cumulative distribution function (CDF) becomes more meaningful as the stack size increases. The boundary delineation of different ground features becomes more significant as well. However, the performances of the two tests are significantly different. With only 10 images (Figure 7a,d), the AD test has limited capabilities to distinguish different ground features, whereas the TR test has good delineations of different backscattering properties.

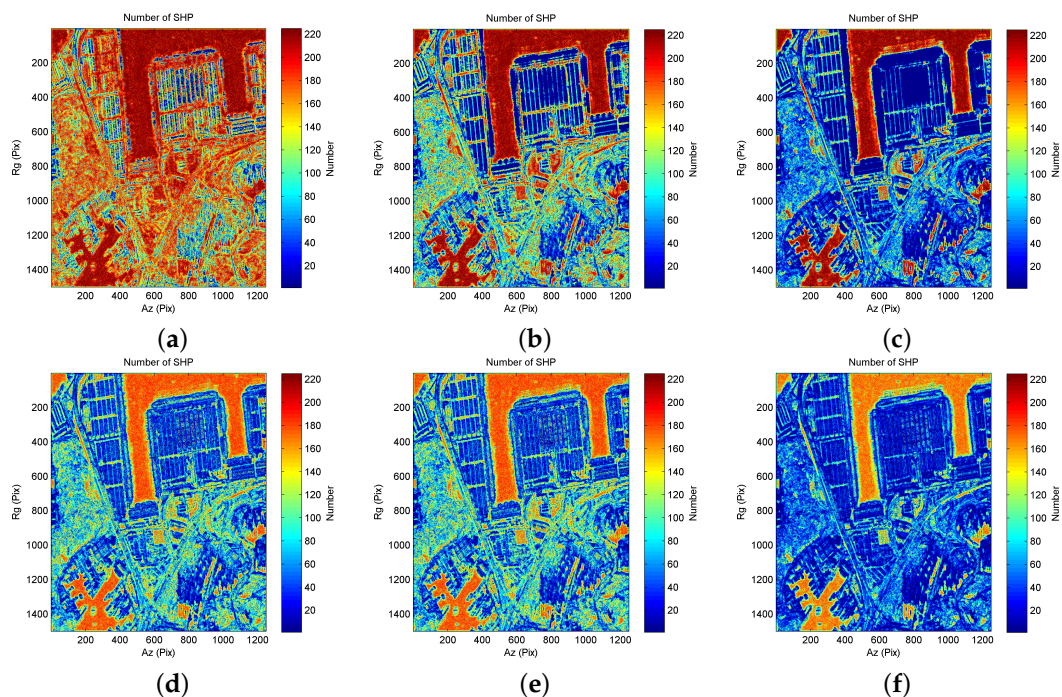


Figure 7. Number of statistically homogeneous pixels (SHP) identified at each pixel based on different tests and stack sizes. The significance level is 5%. (a–c) represent the SHP number of the Anderson-Darling (AD) test using 10, 30, 75 images, respectively; (d–f) correspond to the SHP number of Robust-T (TR) test using 10, 30, 75 images, respectively.

Figure 8a signifies that the AD test has a lower rejection rate and a higher number of SHP on the right tail of the histogram. A total of 37,497 pixels (nearly 2%) identifies the SHP with a full window size (window size is 15×15), i.e., 225 pixels. Conversely, Figure 8d shows that the TR test has a higher rejection rate and a lower number of SHP on the right tail of the histogram. None of the pixels identify the SHP with a full window size. Although the power cannot be directly obtained from this experiment, we may surmise that the higher rejection rate in the TR test correlates to better delineations between different ground features. On the other hand, it is clear that the stack size has significant influences on the shape of the histograms in the AD test (as demonstrated in Figure 8a–c). The results can be explained by the fact that the stack size affects the buildup of CDF and can potentially improve the power in some areas, which leads to alterations of SHP number. Note, however, that the shape of the histograms is little influenced by the stack size in the TR test (as shown in Figure 8d–f). The subtle impact on the histograms can possibly relate to the robustness of the proposed approach since the TR test does not rely on the information built up from the CDF.

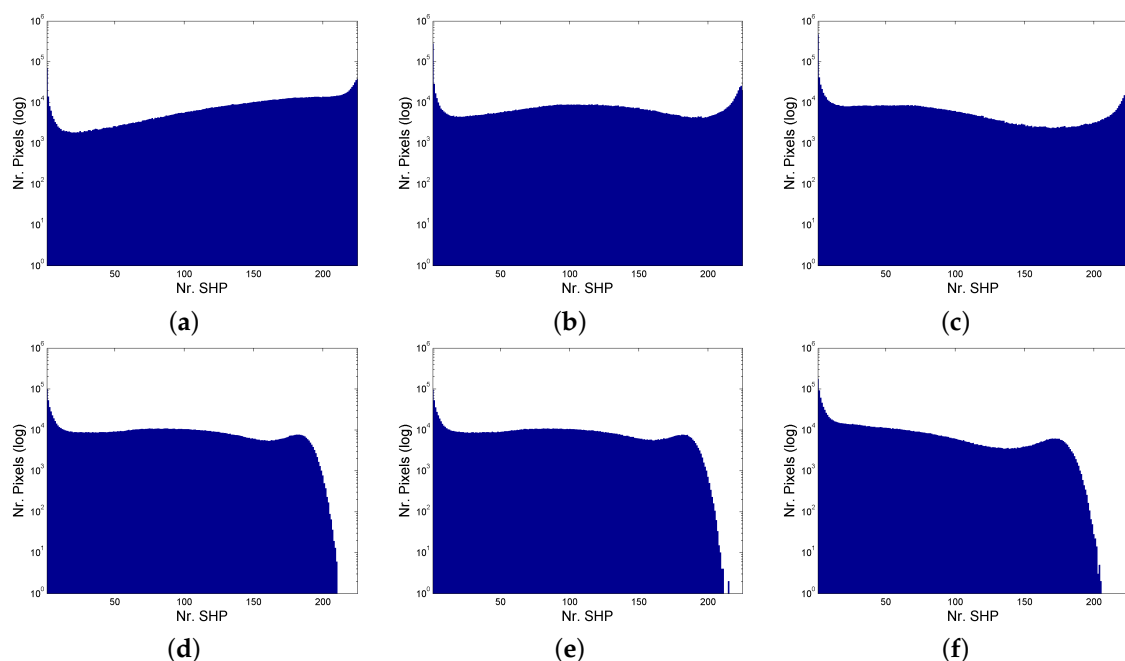


Figure 8. Histograms of the SHP number. Horizontal axis is the number of SHP based on a 15×15 window. Vertical axis is the pixel counts at the corresponding SHP number; (a–f) correspond to Figure 7a–f, respectively.

Figure 9 shows the denoised amplitude (amplitude values of a single SAR acquisition) and reflectivity (temporally averaged amplitude values of the entire SAR data stack) maps of Area 1. The denoised amplitude and reflectivity maps are acquired by averaging all amplitudes in SHP groups (i.e., $\Omega(P)$) for given amplitude map and reflectivity map, respectively. Figure 9d shows that the AD test (10 images) identifies a limited number of SHP in the container areas, which leads to underfiltered maps as shown in Figure 9h,l. The delineations of the container boundaries are blurred due to the low rejection rate. Indeed, building up a meaningful CDF requires enough observations. With only a limited number of temporal observations, the AD test can easily fail to correctly reject the null hypothesis. As a result, the unchanged areas are overfiltered and the changed areas are underfiltered. The denoised results are not considered ideal in this case. Moreover, Figure 9e indicates that the AD test (75 images) identifies much less SHP in the changed areas than the previous case (i.e., AD test with 10 images). The boundary delineations become clearer at the cost of the high rejection rate. Few pixels are denoised in the changed areas as illustrated in Figure 9i. Contrarily, the proposed TR test has much better delineations between different ground features than the AD test, especially for small data

stacks (see Figure 9f). Figure 9j,n also demonstrate that the amplitude and reflectivity maps are better denoised compared with Figure 9h,l. These results show that the statistics of the TR test can provide meaningful SHP groups even with small stack sizes. Increasing the stack sizes does not hamper the SHP identification either. The contrast can thus be enhanced while leaving limited pixels unfiltered, as shown in Figure 9j,k.

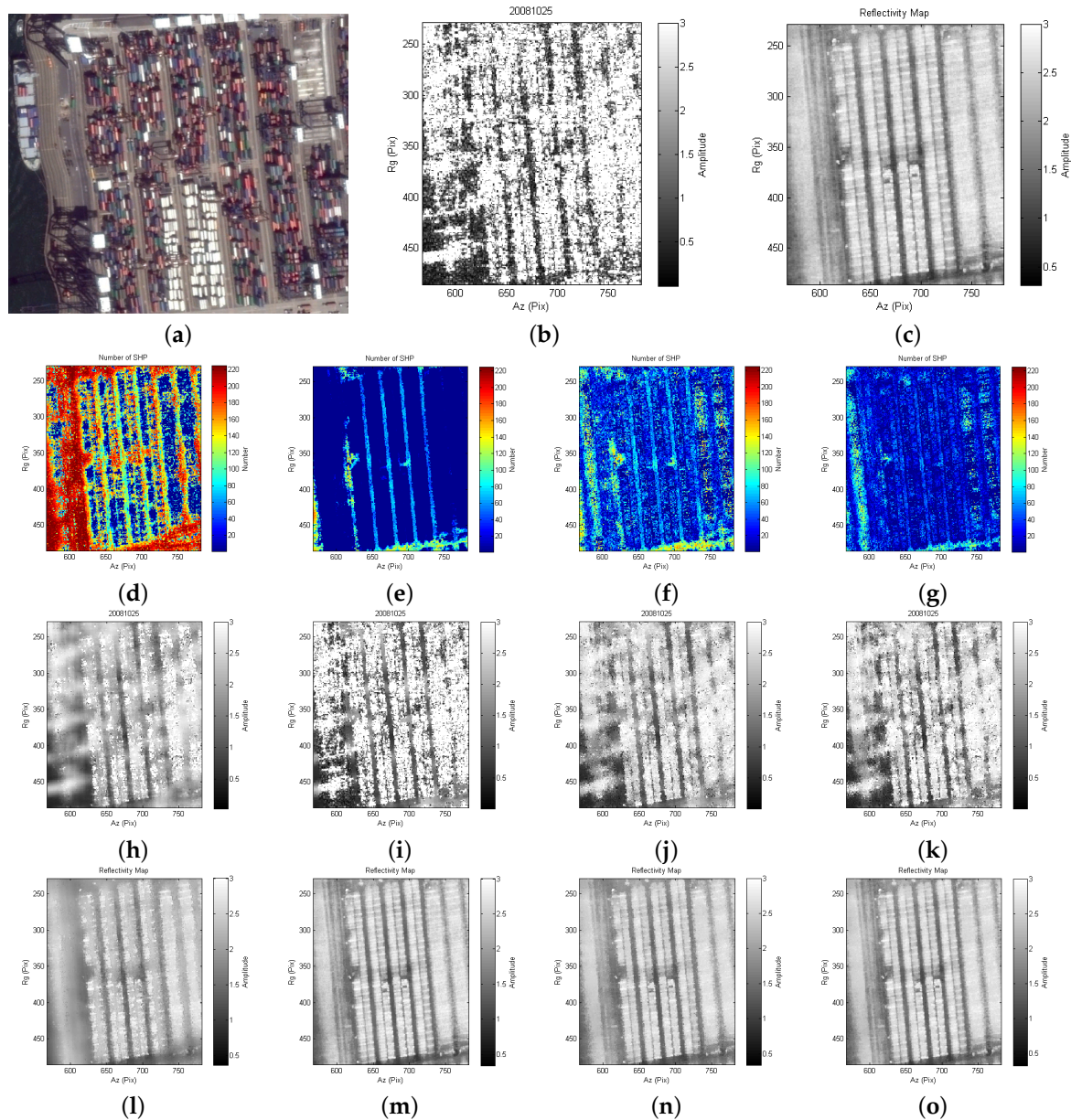


Figure 9. Amplitude denoising of Area 1 based on different approaches and stack sizes. (a) an optical image retrieved from Google Earth; (b) an unfiltered amplitude map acquired on the 25 October 2008; (c) a reflectivity map generated by the original data stack (75 images); (d–g) SHP number correspond to AD^{10imgs}, AD^{75imgs}, TR^{10imgs}, and TR^{75imgs}, respectively; (h–k) denoised amplitude maps correspond to AD^{10imgs}, AD^{75imgs}, TR^{10imgs}, and TR^{75imgs}, respectively; and (l–o) denoised reflectivity maps correspond to AD^{10imgs}, AD^{75imgs}, TR^{10imgs}, and TR^{75imgs}, respectively.

Figure 10 shows the denoised amplitude and reflectivity maps of Area 2. From Figure 10d, one can observe that the AD test (10 images) identifies many uncorrelated pixels as SHP, and the boundaries are not ideal due to the small correct rejection rate limited by the stack size. A significant amount of

overfiltered pixels can be found in Figure 10h,l. Furthermore, a high rejection rate occurs in the changed areas as the stack size increases (see Figure 10e). Similar to Area 1, the changed areas are mostly underfiltered in the AD test (see Figure 10i), whereas distinct ground features are well delineated in the proposed TR test (see Figure 10f). Compared with the results in Figure 10h,l, the results in Figure 10j,n are better denoised. One can also observe better denoised results in Figure 10k than in Figure 10j as the stack sizes increase.

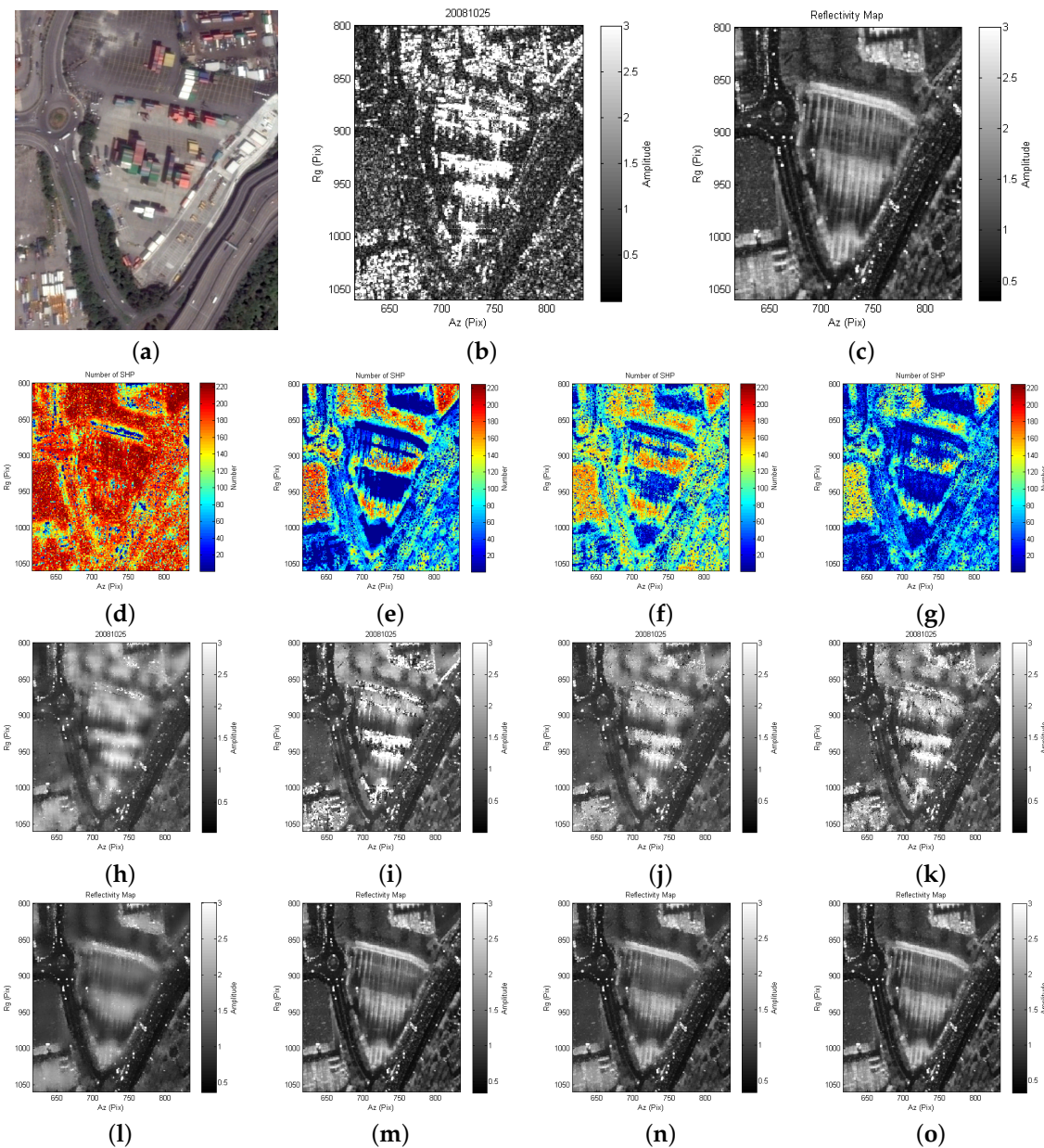


Figure 10. Amplitude denoising of Area 2 based on different approaches and stack sizes. (a) an optical image retrieved from Google Earth; (b) an unfiltered amplitude map acquired on 25 October 2008; (c) a reflectivity map generated by the original data stack (75 images); (d–g) SHP number corresponds to AD^{10img} , AD^{75img} , TR^{10img} , and TR^{75img} , respectively; (h–k) denoised amplitude maps correspond to AD^{10img} , AD^{75img} , TR^{10img} , and TR^{75img} , respectively; and (l–o) denoised reflectivity maps correspond to AD^{10img} , AD^{75img} , TR^{10img} , and TR^{75img} , respectively.

4. Conclusions

In this letter, we present a new testing method for SHP identification. The method considers the impacts of signals and outliers in SHP identification. By suppressing the meaningful signals and removing the outliers, the approach makes the test operation more effective than the conventional ones. Consequently, the test operation outperforms the conventional methods. According to the experimental results, we draw the following conclusions:

- Removing signals and outliers before conducting any hypothesis tests is advantageous. By reducing the impacts of these measurements, the hypothesis testing can deal only with the stochastic processes. This not only makes the parametric tests applicable, but also augments the power of the test operation.
- Considering temporal variabilities is pragmatically necessary, especially when dealing with data stacks crossing through a large temporal spacing. The proposed approach helps to identify SHP family even with temporal variations.
- Since having large sample sizes can lower the probability of conducting type I and II errors, the proposed approach becomes useful in mitigating the impact of temporal variabilities while keeping the large temporal spacing.
- The difference of two time series (i.e., $\psi(P_{ij})$) keeps the temporal sequence that would be lost when building sample CDF in tests like KS or AD. This also helps to improve the effectiveness of hypothesis tests.

For future works, the proposed procedure can be directly applied to different applications, such as adaptive filtering, unbiased coherence estimation, image segmentation, etc.

Acknowledgments: We thank Chong Gu (Department of Statistics, Purdue University) as his advice has been a great help in this study. His willingness to give his time so generously has been very much appreciated. In addition, the authors would like to acknowledge anonymous reviewers for their helpful comments in completing the revised manuscript.

Author Contributions: K.-F.L. and D.P. conceived and designed the proposed methodology. K.-F.L. implemented the methodology and analyzed the experiments under D.P.'s supervision. The manuscript was written by K.-F.L.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ferretti, A.; Fumagalli, A.; Novali, F.; Prati, C.; Rocca, F.; Rucci, A. A new algorithm for processing interferometric data-stacks: SqueeSAR. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3460–3470.
2. Parizzi, A.; Brcic, R. Adaptive InSAR stack multilooking exploiting amplitude statistics: A comparison between different techniques and practical results. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 441–445.
3. Goel, K. Advanced Stacking Techniques and Applications in High Resolution SAR Interferometry. Ph.D. Thesis, Technical University of Munich, Munich, Germany, 2014.
4. Goel, K.; Adam, N. A distributed scatterer interferometry approach for precision monitoring of known surface deformation phenomena. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 5454–5468.
5. Schmitt, M.; Stilla, U. Adaptive multilooking of airborne single-pass multi-baseline InSAR stacks. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 305–312.
6. Sica, F.; Reale, D.; Poggi, G.; Verdoliva, L.; Fornaro, G. Nonlocal adaptive multilooking in SAR multipass differential interferometry. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 1727–1742.
7. Deledalle, C.A.; Denis, L.; Tupin, F. NL-InSAR: Nonlocal interferogram estimation. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1441–1452.
8. Deledalle, C.A.; Denis, L.; Tupin, F.; Reigber, A.; Jäger, M. NL-SAR: A unified nonlocal framework for resolution-preserving (Pol)(In) SAR Denoising. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2021–2038.
9. Jiang, M.; Ding, X.; Li, Z. Hybrid approach for unbiased coherence estimation for multitemporal InSAR. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2459–2473.

10. Jiang, M.; Ding, X.; Li, Z.; Tian, X.; Wang, C.; Zhu, W. InSAR coherence estimation for small data sets and its impact on temporal decorrelation extraction. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 6584–6596.
11. Schmitt, M.; Schönberger, J.L.; Stilla, U. Adaptive covariance matrix estimation for multi-baseline InSAR data stacks. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 6807–6817.
12. Jiang, M.; Ding, X.; Hanssen, R.F.; Malhotra, R.; Chang, L. Fast statistically homogeneous pixel selection for covariance matrix estimation for multitemporal InSAR. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1213–1224.
13. Carvalho, E.A.; Ushizima, D.M.; Medeiros, F.N.; Martins, C.I.O.; Marques, R.C.; Oliveira, I. SAR imagery segmentation by statistical region growing and hierarchical merging. *Digit. Signal Process.* **2009**, *20*, 1365–1378.
14. Walpole, R.E.; Myers, R.H. *Probability and Statistics for Engineers and Scientists*, 5th ed.; Macmillan: New York, NY, USA, 1993; pp. 671–672.
15. Tison, C.; Nicolas, J.M.; Tupin, F.; Maître, H. A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 2046–2057.
16. Goodman, J.W. Statistical properties of laser speckle patterns. In *Laser Speckle and Related Phenomena*; Springer: Heidelberg, Germany, 1975; pp. 9–75.
17. Jao, J. Amplitude distribution of composite terrain radar clutter and the κ -distribution. *IEEE Trans. Antennas Propag.* **1984**, *32*, 1049–1062.
18. Frery, A.C.; Muller, H.J.; Yanasse, C.C.F.; Sant’Anna, S.J.S. A model for extremely heterogeneous clutter. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 648–659.
19. Moser, G.; Zerubia, J.; Serpico, S.B. SAR amplitude probability density function estimation based on a generalized Gaussian model. *IEEE Trans. Image Process.* **2006**, *15*, 1429–1442.
20. Touzi, R. A review of speckle filtering in the context of estimation theory. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2392–2404.
21. Mascarenhas, N.D.A. An overview of speckle noise filtering in SAR images. In Proceedings of the First Latino-American Seminar on Radar Remote Sensing—Image Processing Techniques, Buenos Aires, Argentina, 2–4 December 1996; Volume 407, p. 71.
22. Wilcox, R.R. *Introduction to Robust Estimation and Hypothesis Testing*, 3rd ed.; Elsevier Science: Burlington, MA, USA, 2011.
23. Hubert, M.; Vandervieren, E. An adjusted boxplot for skewed distributions. *Computat. Stat. Data Anal.* **2008**, *52*, 5186–5201.
24. Brys, G.; Hubert, M.; Struyf, A. A robust measure of skewness. *J. Comput. Graph. Stat.* **2004**, *13*, 996–1018.



© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).