Article

Underlying Topography Inversion Using TomoSAR Based on Non-Local Means for an L-Band Airborne Dataset

Xing Peng [1], **Youjun Wang** [1], **Shilin Long** [1], **Xiong Pan** [1], **Qinghua Xie** [1], **Yanan Du** [2], **Haiqiang Fu** [3], **Jianjun Zhu** [3] and **Xinwu Li** [4,*]

1   School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China; pengxing@cug.edu.cn (X.P.); youjunwang@cug.edu.cn (Y.W.); lsl_cug@cug.edu.cn (S.L.); pxjlh@cug.edu.cn (X.P.); xieqh@cug.edu.cn (Q.X.)
2   School of Geographical Sciences, Guangzhou University, Guangzhou 510006, China; yndu@gzhu.edu.cn
3   School of Geoscience and Info-Physics, Central South University, Changsha 410083, China; haiqiangfu@csu.edu.cn (H.F.); zjj@csu.edu.cn (J.Z.)
4   Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
*   Correspondence: lixw@aircas.ac.cn

Abstract: The underlying topography is an important part of the three-dimensional structure of forests, and is used for a variety of applications, such as hydrology and water resource management, civil engineering projects, and forest resource surveying. Due to the three-dimensional imaging ability and strong penetration, the tomographic synthetic aperture radar (TomoSAR) with a long wavelength has been shown to be a useful tool to estimate the underlying topography. At present, most of the current methods use the local means method to estimate the sample covariance matrix, in which the vertical backscattering power is estimated. However, these methods cannot easily obtain high-precision underlying topography, and often lose some detailed information. In this paper, to solve this problem, a non-local means method is introduced to estimate the optimal covariance matrix by combining weighted neighborhood pixels. To validate the feasibility and effectiveness of this proposed method, a BioSAR 2008 campaign L-band dataset acquired from the northern forests of Sweden was used to inverse the underlying topography. The results show that the accuracy of the underlying topography retrieved by the proposed method is improved by more than 30% when compared with the traditional method.

Keywords: underlying topography; tomographic synthetic aperture radar (TomoSAR); covariance matrix (CM); local means (LM); nonlocal means (NLM)

1. Introduction

The underlying topography, as an important parameter of forest resource surveying, not only affects the spatial distribution of forest resources, but is also closely related to the stability of forest ecosystems [1,2]. The traditional aerial survey or optical remote sensing approaches can only obtain the height information of the forest canopy, and they cannot obtain the real underlying topography of forests [3–5]. Tomographic synthetic aperture radar (TomoSAR) [6–9], and especially the long-wavelength SAR systems, can penetrate the forest canopy to the ground and record the backscattering information from the forest vertical structure [10], which provides us with the possibility of underlying topographic mapping [11,12]. TomoSAR is an extension of the traditional two-dimensional (2D) imaging to three-dimensional (3D) imaging by collecting several images at different heights [13]. In other words, TomoSAR can be used to obtain the scattering echo along the elevation by the computed tomography for each resolution cell [14]. Therefore, it has been used for underlying topography inversion by effectively separating different scatterers along the vertical direction in the same resolution cell [15–18].

Over forest areas, there are a lot of distributed scatterers, and their vertical backscattering power is contained in the amplitude and phase information of the covariance matrix

(CM). Accordingly, the tomographic 3D imaging of forests is processed using the CM [19,20]. For example, the frequently used tomographic methods such as beamforming [6], adaptive beamforming (Capon) [8], and the multiple signal classification (MUSIC) [9] algorithm all obtain tomograms on the basis of CMs [21,22]. Therefore, the estimation accuracy of the CM directly determines the performance of tomographic focusing. In fact, it is impossible to acquire a true CM, due to the lack of observations. To address this difficulty, the true CM is generally replaced by a sample covariance matrix (SCM). At present, the SCM is usually estimated by the local means (LM) method [23–25], which statistically averages all the pixels in a sliding window. This method is simple and easy to implement. However, it can only obtain a correct estimation when all the pixel statistics within the window are consistent. If there are some differences, the estimation of the CM will be inaccurate, which easily leads to the mixing and superposition of different scattering mechanisms. Furthermore, the resolution of the tomographic spectrum will be coarse, resulting in a loss of detailed information and a reduction in the underlying topography estimation accuracy. For example, a slope can become flat ground after being processed by the LM method.

To overcome the above problems, the non-local means (NLM) method was applied to perform three-dimensional focusing, which has been already studied in [26,27]. NLM takes advantage of the neighborhood pixels in a sliding window, that is also the search window, around the center pixel. It also calculates the similarity between the center pixel and the neighborhood pixels to obtain the weight. Based on the weight, the optimal CM is calculated. Compared to the LM method, the NLM method can preserve the true scattering characteristics in each resolution cell after averaging. Based on the estimated CM, the reconstructed tomograms can retain the details well, especially over undulating terrain. However, reference [26] only considered radiometric similarity between two pixels but ignored the spatial similarity, which probably resulted in homogeneous area misinterpretation or excessive smoothness. Moreover, it is necessary to manually select a homogeneous area prior to filtering. The methodology in [27] can solve the above problem by adaptively considering both the radiometric similarity and the spatial similarity between two pixels. This method is based on the complex Gaussian distribution to calculate the similarity distance. However, over forest areas, the scattering process is complex. In particular, the pure volume scattering hypothesis no longer applies under the observation of SAR sensors with long wavelength such as L band and P band. Besides the volume scattering, there are ground scattering and double-bounce scattering. This means that it is not suitable to use complex Gaussian distribution to express statistical characteristics. In addition, it is still not yet fully clear how NLM performs in underlying topography estimation. In view of this, this paper proposes an improved NLM in TomoSAR to estimate the underlying topography. This method adopted the affine invariant distance to replace the distance used in [27], since it is invariant to any affine transformation of the matrices. This distance does not assume a particular statistical distribution. Thus, the proposed NLM method is suitable for the application of forests and can improve the accuracy of underlying topography estimation.

The rest of this work is formed as follows. Section 2 first presents the tomographic SAR model, and briefly describes the principles of the NLM method. Three classical tomographic estimators based on the NLM method are then described. Section 3 gives a description of the research area and the experimental dataset, and then presents and analyzes the experimental results. A further discussion about the tomograms in other channels and forest height estimation is shown in Section 4. Finally, the conclusions of this paper are drawn in Section 5.

2. Non-Local Means (NLM) Tomographic Synthetic Aperture Radar (TomoSAR) Method

2.1. SAR Tomography Model

TomoSAR is a 3D extension of the traditional 2D SAR imaging [28]. The technique makes multiple observations of the same target at different heights in the normal direction (i.e., cross-range c) of the traditional azimuth x-range r plane. It can achieve 3D imaging of

target objects by obtaining another synthetic aperture in the height dimension (as shown in Figure 1).

Figure 1. Geometry and spectral estimation of the tomographic synthetic aperture radar (TomoSAR) configuration.

Through multiple SAR observations at different times with slightly different incidence angles over the same place, N SAR images can be acquired. Some preprocessing steps which contain co-registration, deramping, and phase calibration with respect to a common master image must be done [29]. For an arbitrary pixel (x, r), the focused complex value $u_n(x, r)$ for the nth acquisition is expressed as [6,30]:

$$u_n(x, r) = \int \beta(x, r, z) \exp(jk_z z) dz, \quad n = 1, \cdots, N \tag{1}$$

where $\beta(x, r, z)$ is the complex reflectivity along the vertical direction z, j is the complex unit, and $k_z = 4\pi b_{\perp n} / \lambda r \sin(\theta)$ is known as the vertical number. $b_{\perp n}$ is the perpendicular baseline between the master image and this nth image, λ represents the wavelength, r is the slant distance between the master image and the target, and θ is the incidence angle of the radar.

As the coherent noise in forest areas is strong, a multi-look approach is generally needed to suppress it. If we assume that the number of looks is L, then the tomographic SAR model can be rewritten as:

$$g(l) = A(z)\beta(l) + e(l), l = 1, \cdots, L \tag{2}$$

where $g(l)$ represents the vector composed of the N measurements; $\beta(l)$ denotes the unknown vector with D complex reflectivity elements; $e(l)$ is a $N \times 1$ noise vector; and $A(z)$ indicates a $N \times D$ steering matrix with $A(z) = [a(z_1), a(z_2), \cdots, a(z_D)]$, where $z_d(d = 1, \cdots, D)$ represents the discrete height position.

For the distributed scatterers, SAR tomography is interested in the reflectivity power along the vertical direction, which can be given by:

$$P(x, r, z) = E\left(|\beta(x, r, z)|^2\right) \tag{3}$$

where $E(\cdot)$ denotes the statistical expectation operator.

Equation (3) can be actually regarded as a problem of spectral estimation. To date, a lot of spectral estimation approaches have been proposed to implement tomographic focusing, which can be categorized into three groups: (1) non-parametric spectral estimation [6,13,14]; (2) parametric spectral estimation [9,15]; and (3) compressive sensing [16,19,31]. Among these methods, the most widely used classical spectral estimation methods are Beamforming, Capon, and MUSIC.

2.1.1. Beamforming

Beamforming was the earliest algorithm used to solve the overlay problem [32,33]. The basic idea is that this "finite impulse response filter" allows signals of a specific spatial frequency to pass without distortion, and attenuates the signals of other frequencies. The power of the reflectivity estimated by beamforming is written as [6,34]:

$$\hat{P}_{\mathrm{BF}} = \frac{a(z_d)^H R a(z_d)}{N^2} \tag{4}$$

where the SCM R is defined by:

$$R = \frac{1}{L} \sum_{l=1}^{L} g(l) g(l)^H \tag{5}$$

where $(\cdot)^H$ denotes the Hermitian operator.

2.1.2. Adaptive Beamforming (Capon)

Some of the shortcomings of beamforming have been solved by the Capon approach, including the low height resolution, irregular sampling, and serious sidelobes. Its basic principle is similar to that of beamforming, but it does not assume that the received signal is the spatial white noise distribution whose variance is the unit matrix I, which is replaced by R. The reflectivity power estimated by Capon can be expressed as [8,35]:

$$\hat{P}_{\mathrm{CP}} = \frac{1}{a(z_d)^H R^{-1} a(z_d)} \tag{6}$$

2.1.3. Multiple Signal Classification (MUSIC)

The MUSIC algorithm decomposes R into signal and noise by eigen decomposition. The reflectivity power can be given by [9,33]:

$$\hat{P}_{\mathrm{MUSIC}} = \frac{1}{a(z_d)^H U U^H a(z_d)} \tag{7}$$

where U represents the eigenvector of the signal.

2.2. The NLM Algorithm

According to the analysis in the previous section, beamforming, Capon, and MUSIC all estimate the backscattering power spectrum from the CM. In order to obtain a CM that is closest to the real values, we utilize the NLM method to optimize the estimation. NLM first uses a search window (e.g., 11×11) with the target pixel as the center, and then uses a matching window (e.g., 3×3) to calculate the SCM of the center pixel and its neighborhood pixels. After this, the similarity weight, which is composed of the spatial and radiometric similarity between each neighborhood pixel and the center pixel, is calculated by traversing the entire search window. Finally, the weighted SCM of all the neighborhood pixels is used to describe the CM of the center pixel. The NLM method not only makes maximum use of the neighborhood information to ensure the authenticity of the estimation for the center pixel, but it also eliminates the interference information, such as noise or heterogeneous

information in the neighborhood, by means of the weighting. This improves the estimation accuracy of the center pixel.

The value of the center pixel in the estimation window W can be estimated using the NLM method, which is given by [36]:

$$\hat{F}(x_0) = \frac{\sum_{x_i \in W} w_{x_0, x_i} C_{x_i}}{\sum_{x_i \in W} w_{x_0, x_i}} \tag{8}$$

where x_0 indicates the target pixel (the central pixel in the estimation window), index x_i denotes the neighborhood pixel, and C_{x_i} represents the observation value of the neighborhood pixel. The weights are expressed as:

$$w_{x_0, x_i} = f_s(x_0, x_i) f_r(x_0, x_i) \tag{9}$$

where the spatial similarity $f_s(x_0, x_i)$ is defined by [27]:

$$f_s(x_0, x_i) = \delta_{\gamma_s}(||x_0 - x_i||) \tag{10}$$

and radiometric similarity $f_r(x_0, x_i)$ is defined by:

$$f_r(x_0, x_i) = \delta_{\gamma_r} \left[\frac{1}{P^2} \sum_{p \in \boldsymbol{P}} d^2 \left(C_{x_i + p}, C_{x_0 + p} \right) \right]^{1/2} \tag{11}$$

where $\boldsymbol{P}$ is a 2D vector about positions of pixels with reference to the matching window center, and P^2 is the total pixel number in the matching window [27]. Both the spatial and radiometric similarity kernels can be estimated from [27]:

$$\delta_\gamma(x) = \exp\left[-\left(\gamma^{-1} x \right)^2 \right] \tag{12}$$

where γ is the user-defined scaling factor; γ_s is the spatial extent of the filter, and can refer to the window size used in other filters such as the Lee filter [37]; and γ_r controls the amount of filtering based on the radiometric similarity between these two pixels. They are usually designated as $\gamma_s = 3$ and $\gamma_s = 0.9$ [36]. The function $d(C_{x_0}, C_{x_i})$ is the distance between two Hermitian matrices. This distance is defined as the affine invariant (AI) distance, because it keeps the same under any affine transformation of the matrices, and can be expressed as [27,38]:

$$d_{\text{AI}} \left(C_{x_i}, C_{x_j} \right) = \left|\left| \log \left(C_{x_j}^{-1/2} C_{x_i} C_{x_j}^{-1/2} \right) \right|\right|_F \tag{13}$$

where log is a logarithm function and $||\cdot||_F$ is the Frobenius norm.

Using this method for the CM is similar to "filtering", i.e., assigning different weights to filter out the pixels with low similarity around the target pixel and selecting the pixels with high similarity to jointly estimate the optimal CM value of the target pixel.

2.3. The Traditional Spectral Estimation Methods Based on NLM

The three classical spectral estimation methods (beamforming, Capon, and MUSIC) are all calculated by using the SCM obtained by the LM method. In the proposed approach, NLM is applied to the CM estimation of the tomographic SAR model to improve its estimation accuracy and the inversion accuracy of the underlying topography. The steps are as follows:

(1) Solve the SCM $\boldsymbol{R}$ of all the pixels in the area of interest.
(2) Specify the size of the search window W and matching window P.

(3) Calculate the spatial similarity $f_s(x_0, x_i)$ and radiometric similarity $f_r(x_0, x_i)$ of the matching window between the central pixel x_0 and neighboring pixel x_i. (A pixel (m, n) in the research region is located at x_0 within its search window W).

(4) Calculate the weight of the neighborhood pixel x_i based on the spatial and radiometric similarity.

(5) Calculate the optimal weighted CM of the center pixel by the using of the SCM of all the neighborhood pixels (except for the center pixel) in the search window and their corresponding weights.

(6) Substitute the estimated CM into the spectrum estimation formula to estimate the pixel's spectrum.

(7) Traverse the whole study area, and repeat steps (3) to (6) to obtain the spectra over the whole area.

In addition, the details of the classical spectral estimators based on NLM are provided in Table 1.

Table 1. Details of the classical spectral estimators based on NLM.

Initialization	$R = \frac{1}{L} \sum\limits_{l=1}^{L} g(l)g(l)^H$
Traverse	
	repeat
	$f_s(x_0, x_i) = \delta_{\gamma_s}(\lVert x_0 - x_i \rVert)$
	$f_r(x_0, x_i) = \delta_{\gamma_r}\left[\frac{1}{P^2} \sum\limits_{p \in P} d^2\left(R_{x_i+p}, R_{x_0+p}\right) \right]^{1/2}$
	$w_{x_0, x_i} = f_s(x_0, x_i) f_r(x_0, x_i)$
	$\hat{R}(x_0) = \frac{\sum_{x_i \in W} w_{x_0, x_i} R_{x_i}}{\sum_{x_i \in W} w_{x_0, x_i}}$
	$\hat{P}_{\text{BF}} = \frac{a(z_d)^H \hat{R} a(z_d)}{N^2} \quad \hat{P}_{\text{CP}} = \frac{1}{a(z_d)^H \hat{R}^{-1} a(z_d)} \quad \hat{P}_{\text{MUSIC}} = \frac{1}{a(z_d)^H u u^H a(z_d)}$
Until (finish)	

The principle of optimal CM estimation based on NLM is shown in Figure 2. The central pixel x_0 is the weighted average of all the neighborhood pixels x_i in the search window W. Neighbors with a high similarity are given larger weights w_{x_0, x_i^1} and w_{x_0, x_i^2}, while the neighborhood pixels with large differences are given smaller weights w_{x_0, x_i^3}.
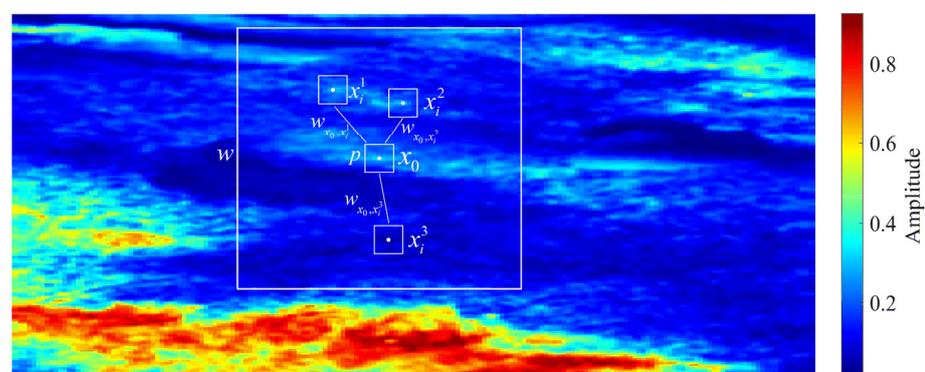
Figure 2. The principle of optimal CM estimation based on NLM (p represents the size of the matching window, and w denotes the size of the search window).

3. Experiments and Results

For the purpose of investigating the feasibility and validity of the proposed method, TomoSAR was applied to a real airborne SAR dataset to obtain the underlying topography, and a comparison was made with LiDAR measurements.

3.1. Study Area and Dataset

This paper selected a boreal forest in Krycklan, located in the north of Sweden, as the study area, which is made up of coniferous trees such as Scots pine and Norway spruce, and a few birch trees. The average annual temperature is about 1 °C and the average annual precipitation is around 600 mm. Moreover, the terrain topography is hilly, varying from 190 m to 290 m. The average tree height is about 18 m and the maximum tree height is around 30 m [39,40].

In order to address some important specific requirements of the European Space Agency (ESA) BIOMASS Earth resources exploration program, a stack of six L-band SAR images in fully polarimetric mode was obtained over the study area by the German Aerospace Centre (DLR) during the ESA BioSAR 2008 campaign on 15 October 2008. The BioSAR 2008 project was carried out as a cooperation between the ESA, the DLR, the Swedish Defense Research Agency (FOI), the Swedish University of Agricultural Sciences (SLU), the Biosphere Remote Sensing Research and Education Centre (CESBIO), and the Polytechnic of Milan in Italy. They undertook the preprocessing steps for this dataset, including co-registration, flat-earth phase removal and phase error calibration [40]. Six uniform perpendicular baselines were distributed along the vertical direction. The interval between two adjacent baselines was about 6 m. The incidence angle varied from 25° at the near range to 55° at the far range [40]. The vertical resolution varied in the range of 6 m to 25 m. Parameters of the E-SAR airborne system and the information of perpendicular baselines are presented in Tables 2 and 3, respectively [40].

Table 2. The parameters of the E-SAR airborne system [40].

Items	Parameters
Wavelength	0.23 m (L-band)
Polarimetric channel	HH + HV + VV
Incidence angle	25–55°
Center slant range	3900 m
Range resolution	2.12 m
Azimuth resolution	1.20 m

Table 3. The baseline information for the InSAR pairs [40].

Identifier	Acquisition Date	Baseline (m)
08biosar0201 × 1		0
08biosar0203 × 1		−6
08biosar0205 × 1		−12
08biosar0207 × 1	15 October 2008	−18
08biosar0209 × 1		−24
08biosar0211 × 1		−30

Moreover, in order to validate the results of TomoSAR, a laser radar S/N425 TopEye system onboard a helicopter platform was used to generate a point cloud dataset over this study area on 5 August and 6 August 2008. From these point cloud data, the digital terrain model (DTM) was obtained.

3.2. Experimental Results and Analysis

3.2.1. Comparison of the Tomograms

A profile along range (the red dotted line, as seen in Figure 3) was selected as an example for tomogram analyzing. This profile was fixed at the 300th azimuth resolution cell.

According to Tebaldini et al. [11], the backscattering reflectivity of the HH polarization channel mainly comes from the ground, and the backscattering power of the HV polarization channel is mainly concentrated on the canopy volume scattering. Figure 4 shows the spectrum estimation results in the HH polarization channel obtained by the

classical spectral estimation methods (beamforming, Capon, and MUSIC) based on the LM and NLM methods, respectively. It was found that the tomograms estimated by the three methods based on the LM method failed to separate the ground and canopy backscattering contributions in many places (e.g., the three red circles in Figure 4a,c,e) for which the ground contributions were weak, mistreating the scattering phase center of the canopy as the ground scattering phase center. However, the tomograms based on the NLM method could basically distinguish the scattering phase centers of the ground and the canopy, such as those marked by the three yellow circles in Figure 4b,d,f. In addition, the tomograms estimated by the NLM method were clearer and more continuous than those estimated by the LM method, which was mainly due to the more real and accurate CM estimated by the NLM method. As a result of the NLM method considering more non-local neighborhood information, the target pixel reflects its own optimal information by looking for pixels with higher spatial and radiometric similarity. This avoids the erroneous information which can result from a direct and sloppy adoption of the LM method, such as noise or pixels that differ greatly from their characteristics.
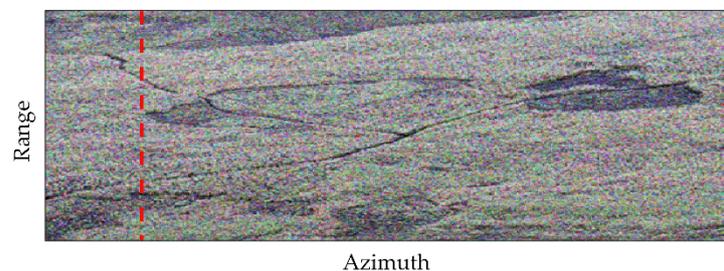
Figure 3. The selected range profile (red dotted line) on the SAR Pauli map.

Figure 5 shows the underlying topography obtained from the tomogram estimation results. Obviously, the spectral estimation method based on the LM method failed to estimate the phase center of the ground in many parts, and mistook the height of the canopy as the height of the ground, especially in the methods of beamforming and Capon. Two important parameters in the NLM method were selected as $p = 3$ and $w = 15$, and the optimal value of w was selected through a large number of experiments (see Figure 6). The selection of the above two parameters depend on the study area. In general, several profiles or plots are selected to optimize the parameters before the application for whole areas. According to the experience in this paper, the range of p is 3–5, and the range of w is 10–25.

In order to quantitatively investigate the effectiveness of the proposed method, the estimation accuracies of different methods were calculated by the root mean square error (RMSE) values. According to Table 4, after the application of the NLM method, the accuracy of the three classical spectral estimation methods of beamforming, Capon, and MUSIC was significantly improved by 34.87%, 38.28%, and 31.61%, respectively.

Table 4. RMSE comparison of the different methods used for the profile analysis.

Item	Beamforming (m)	Capon (m)	MUSIC (m)
LM	3.24	2.87	1.55
NLM	2.11	1.77	1.06
Improvement	34.87%	38.28%	31.61%
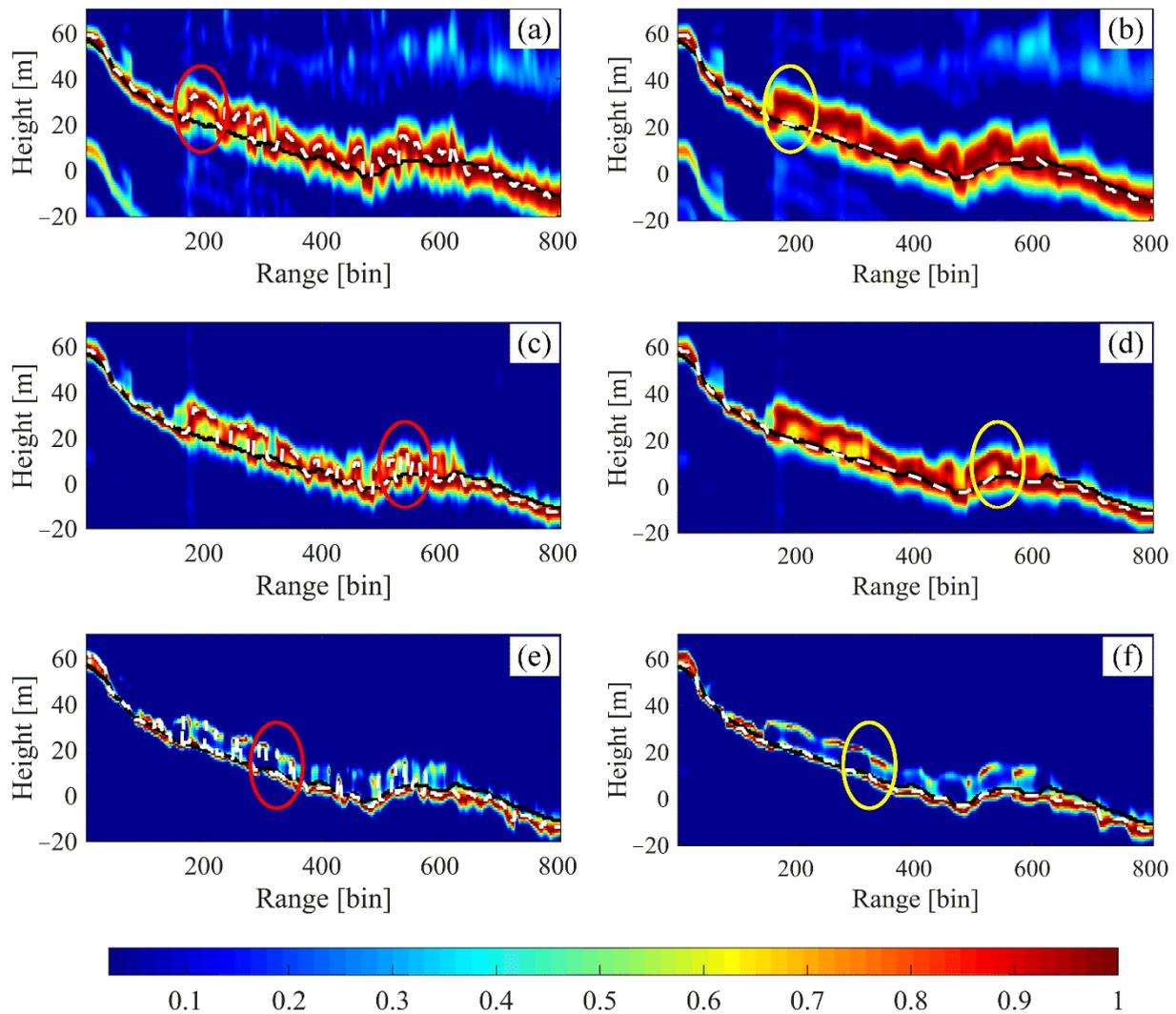
Figure 4. Tomograms of the HH polarization channel at the selected profile estimated by the different TomoSAR methods: (**a**), (**c**), (**e**) are, respectively, the results of local means (LM)-based beamforming, adaptive beamforming (Capon), and multiple signal classification (MUSIC); and (**b**), (**d**), (**f**) are the results of NLM-based beamforming, Capon, and MUSIC, respectively. The solid black line is the LiDAR DTM, and the dashed white line is the estimated values.

Furthermore, we analyzed the computational impact on the used NLM method quantitatively, as shown in Table 5. For the selected range profile, the running time of the NLM methods increased by approximately 14 times that of the LM methods for the same computer (a desktop PC with an i7-8700 6-core processor).

Table 5. Running time of different methods.

Item	Beamforming (s)	Capon (s)	MUSIC (s)
LM	26	25	24
NLM	369	371	374
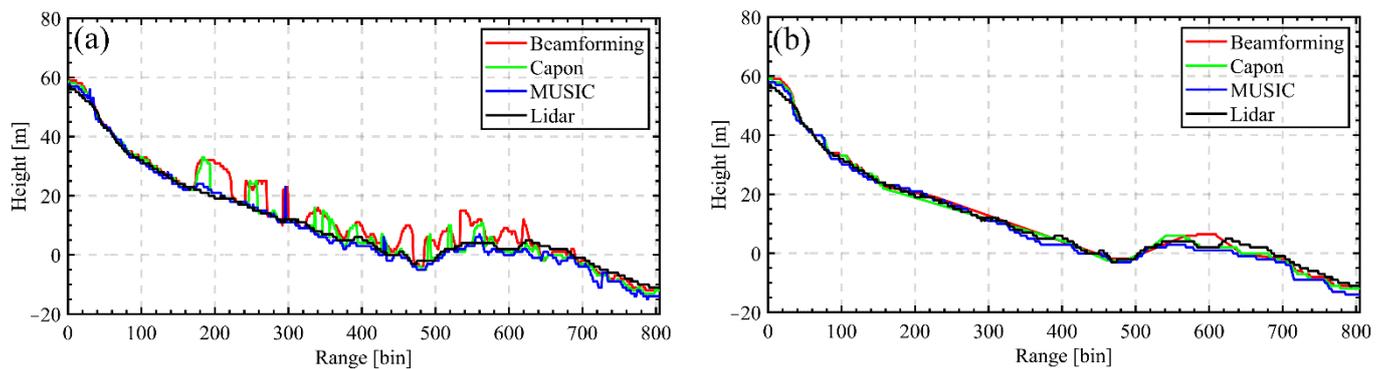
Figure 5. The underlying topography at the selected profile: (**a**) the results of the LM-based TomoSAR method; and (**b**) the results of the NLM-based TomoSAR method.
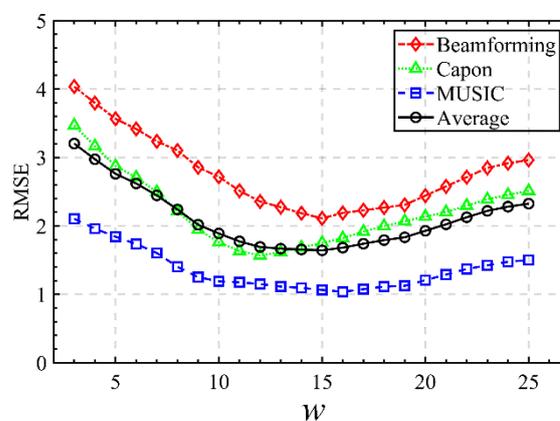
Figure 6. The effect of the search window size w on the RMSE.

3.2.2. Inversion of the Underlying Topography

For verifying the universality of the proposed method, we applied the method to the whole study area for the inversion of the underlying topography, and compared the results with those of the LM method.

From Figure 7, it can be observed that the underlying topography estimated based on the LM and NLM methods is close to the LiDAR DTM measurements. However, there are many areas where the underlying terrain has been overestimated (e.g., the six black circles in Figure 7) in the results based on the LM method. However, the results based on the NLM method (e.g., the six white circles in Figure 7) are less likely to show this. The main reason for this is that the LM method fails to distinguish the scattering phase center between the ground and the canopy in some areas, which leads to the canopy height being mistaken as the ground height. This is the same as the tomogram analysis. In addition, the NLM method also shows a better inversion performance in places with drastic terrain height changes, and it preserves more detailed information (e.g., the three white rectangles in Figure 7). Therefore, the NLM method is applicable for use in areas with large terrain fluctuation, and can achieve a high inversion accuracy. Table 6 lists the RMSEs of the two methods for the inversion of the underlying topography, which proves quantitatively that the NLM method improves the accuracy of the inversion of the underlying terrain by more than 30% when compared with the LM method.
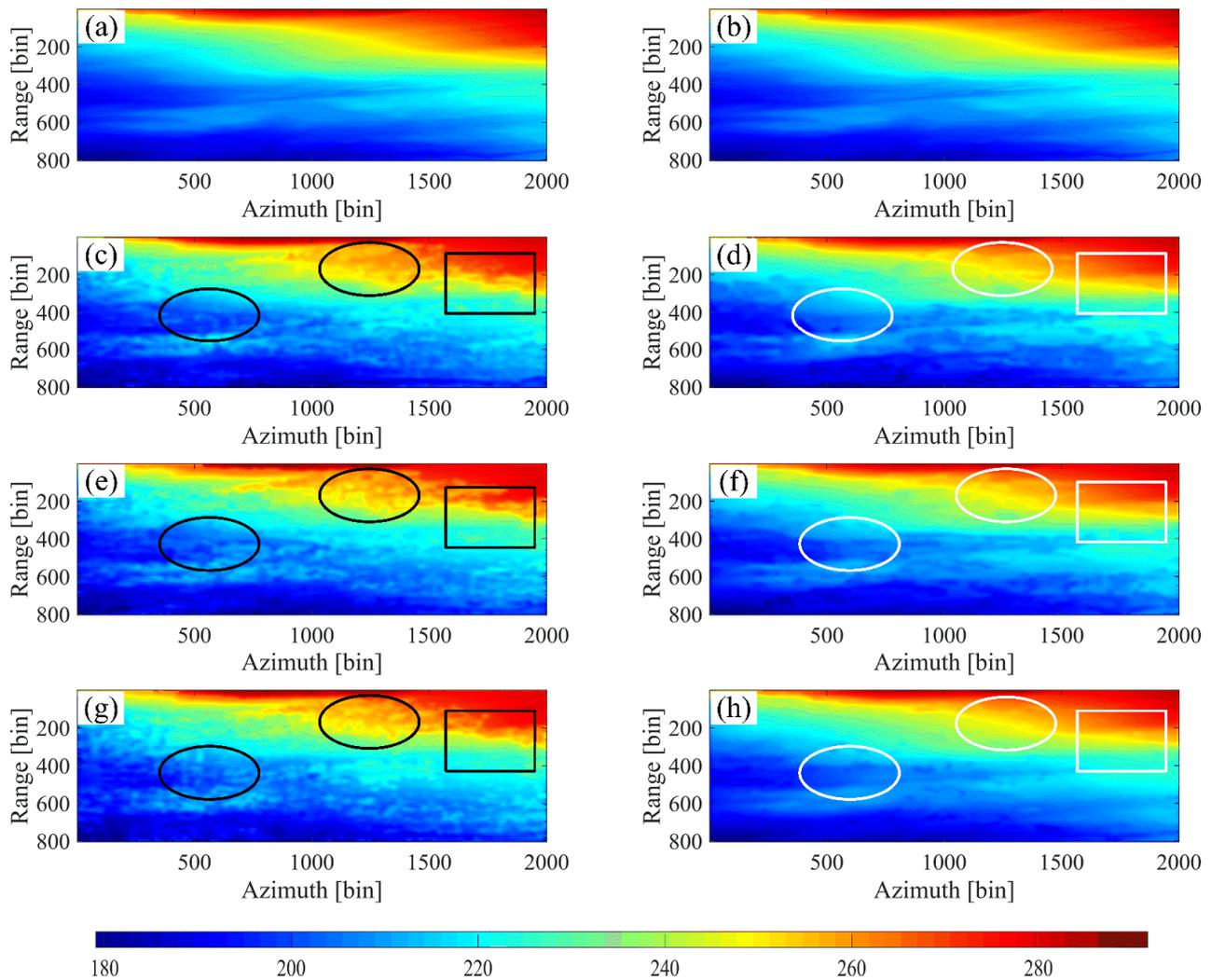
Figure 7. Underlying topography from LiDAR and TomoSAR: (**a**,**b**) are the LiDAR DTM; (**c**), (**e**), (**g**) are, respectively, the beamforming, Capon, and MUSIC results based on the LM method; and (**d**), (**f**), (**h**) are, respectively, the beamforming, Capon, and MUSIC results based on the NLM method.

Table 6. Comparison of the RMSEs of the underlying topography inversion results.

	Beamforming (m)	Capon (m)	MUSIC (m)
LM	2.85	2.56	1.61
NLM	1.83	1.67	1.12
Improvement	35.78%	34.76%	30.43%

Figure 8 shows the relationships between the LiDAR measurements and the inversion results obtained by different methods. Moreover, we calculated the correlation coefficients between the LiDAR measurements and the estimation results (see in Figure 8). The results show that the estimation results of the NLM method (as shown in Figure 8d–f) were more relevant to LiDAR measurements than those of the LM method (as shown in Figure 8a–c). It suggests that the inversion results of the NLM method are more reliable.
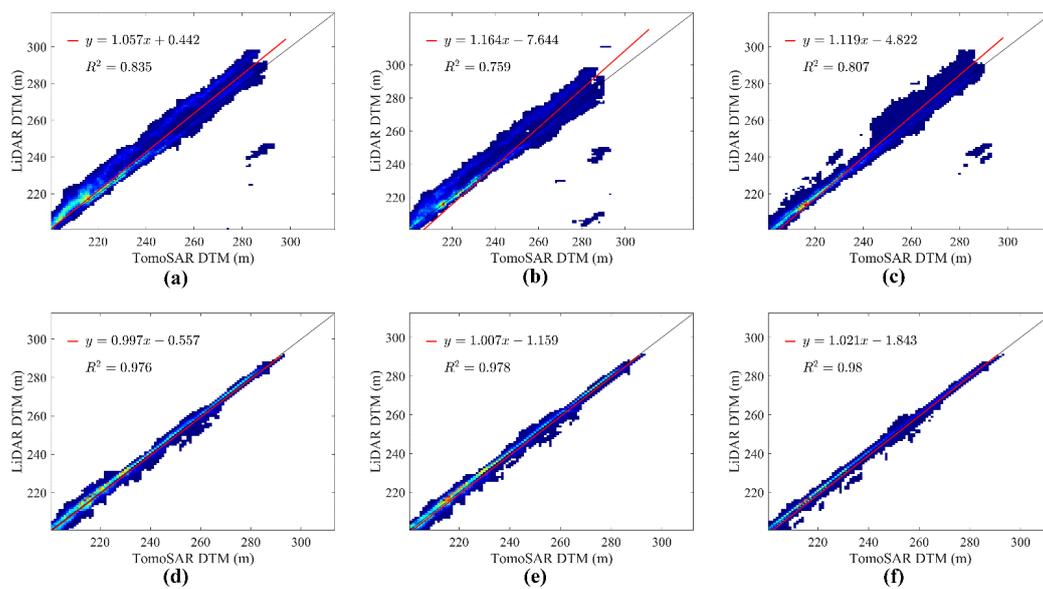
Figure 8. The 2D joint distribution between the LiDAR and TomoSAR DTM of different methods: (**a**) LiDAR DTM and DM by LM beamforming; (**b**) LiDAR DTM and DTM by LM Capon; (**c**) LiDAR DTM and DTM by LM MUSIC; (**d**) LiDAR DTM and DTM by NLM beamforming; (**e**) LiDAR DTM and DTM by NLM Capon; (**f**) LiDAR DTM and DTM by NLM MUSIC.

4. Discussion

4.1. Optimal Covariance Matrix Estimation

From Figure 9, it can be seen that the SCM calculated using the LM method is significantly different from the optimal CM obtained by the NLM method. Figure 9a represents the amplitude information of the SCM estimated by the traditional LM method. Its main diagonal elements have almost identical values, which is the result of using the local mean values. However, the amplitude information of the CM estimated by the NLM method (see Figure 9b) clearly distinguishes the proportions of the different information. Furthermore, there are also some differences between the two methods with regard to phase information (see Figure 9c,d). Thus, the scattering phase centers of the ground and the canopy can be detected separately. This is also the main reason why the accuracy of the NLM method is improved compared with that of the traditional LM method. Therefore, the accurate estimation of CM information plays an important role in underlying topography inversion. The estimation accuracy of the NLM method is better than that of the LM method, but its computation time of neighborhood search is higher, which may limit the application of this method for the for the large scale or global scale monitoring. Moreover, there may exist the potentially distortive effect of NLM approach on radiometric information, which may be caused by the topography. In the future, our work will focus on improving the computational efficiency of this method and removing the topography-induced radiometric distortion as much as possible.

4.2. Tomograms in the HV and VV Channels

Tomograms of the aforementioned selected profile in the HV channel and VV channel were also estimated. The estimated spectra obtained by the traditional methods can basically only detect the backscattering phase center of the canopy in the high tree region (approximately the 200th to 600th pixels), as shown in Figure 10. Furthermore, this is mainly embodied in the beamforming and Capon methods. In contrast, although the method presented in this paper shows a certain improvement effect in the high tree area, there are still some areas where the backscattering phase center of the ground is not detected, especially for beamforming. This is mainly because the backscattering reflectivity of the HV polarization channel is mainly concentrated in the canopy. The inversion performances of the proposed method and the traditional method in the VV polarization channel are shown

in Figure 11. It can be found that the backscattering power level of the VV polarization channel is between that of the HH polarization channel and that of the HV polarization channel, which is the same as the conclusion made by Tebaldini et al. [11]. In other words, the backscattering power from the canopy in the VV polarization channel is stronger than that from the HH polarization channel, and the backscattering power from the ground is stronger than that from the HV polarization channel.
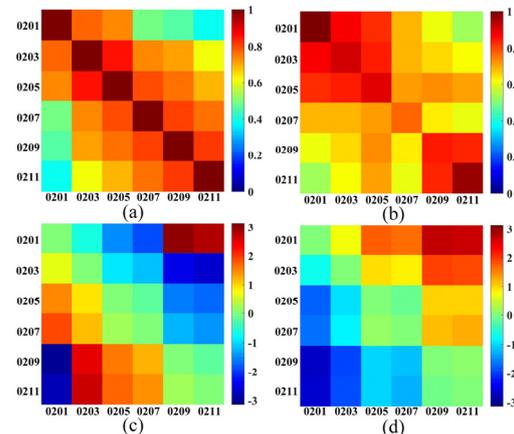
Figure 9. Comparison of the amplitude and phase of the CM for the different methods: (**a**), (**c**) are, respectively, the amplitude and phase of the CM of the LM method; and (**b**), (**d**) are, respectively, the amplitude and phase of the CM of the NLM method.
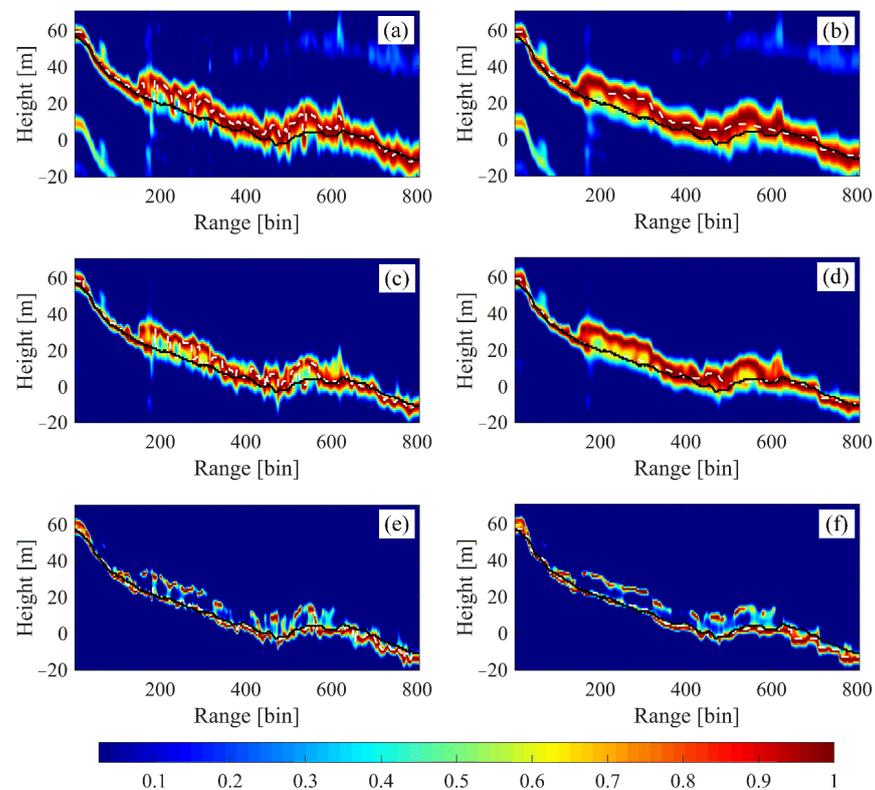
Figure 10. Tomograms of the HV polarization channel at the selected profile estimated by the different TomoSAR methods: (**a**), (**c**), and (**e**) are, respectively, the results of LM-based beamforming, Capon, and MUSIC; and (**b**), (**d**), (**f**) are, respectively, the results of NLM-based beamforming, Capon, and MUSIC. The solid black line is the LiDAR DTM, and the dashed white line is the estimated values.
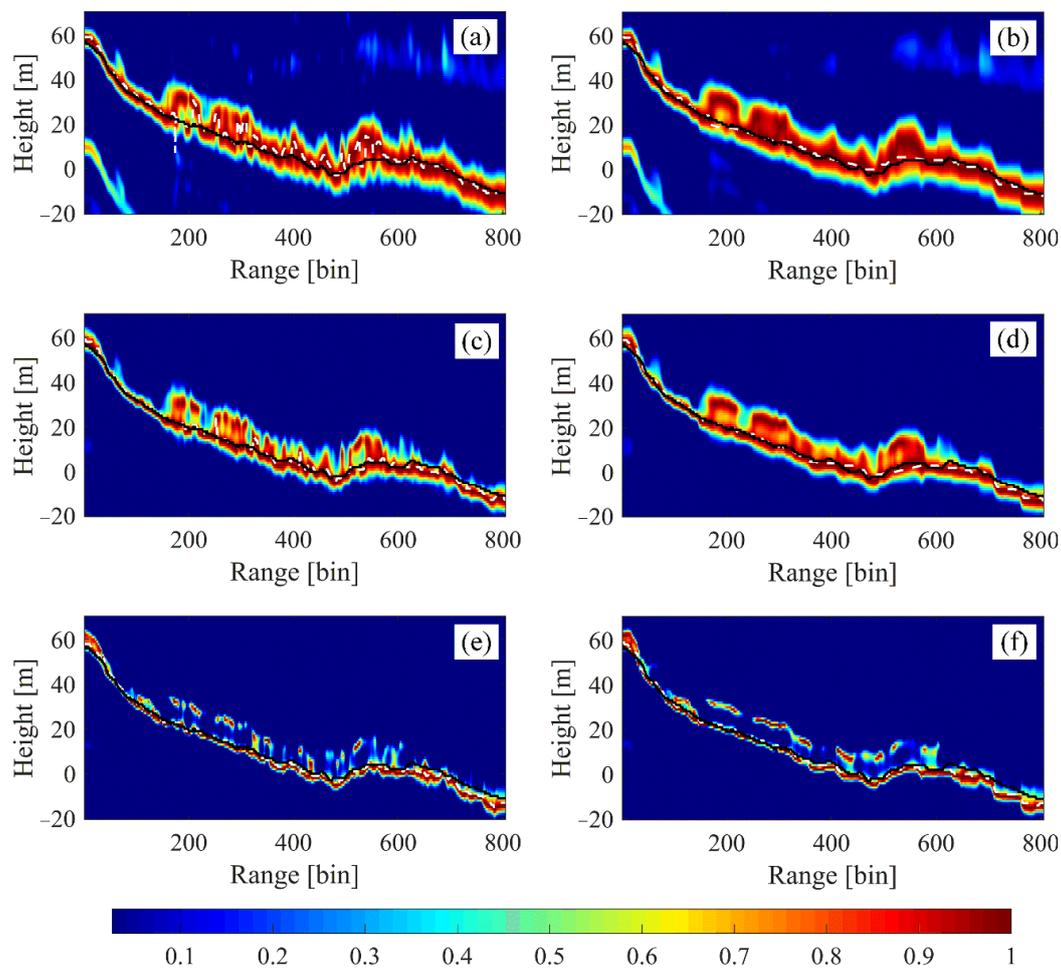
Figure 11. Profile line spectrum estimation results for the VV polarization channel: the solid black line is the LiDAR DTM, and the dashed white line is the estimated values; (**a**), (**c**), and (**e**) are, respectively, the estimation results of LM-based beamforming, Capon, and MUSIC; and (**b**), (**d**), (**f**) are the estimation results of NLM-based beamforming, Capon, and MUSIC, respectively.

4.3. Forest Height Estimation

Besides the underlying topography, forest height is another significant parameter of forest resource surveying; furthermore, it is very important to estimate the forest biomass. In addition, the backscattering power of the canopy is mainly concentrated in the HV polarization channel, as we mentioned above. In order to clearly show the forest height estimation at the selected profile, the ground scattering phase was removed, as shown in Figure 12. In some areas (e.g., the three dotted circles), the traditional methods may cause canopy underestimation by mistaking the backscattering phase center of the ground for that of the canopy. However, the canopy height can be accurately estimated at the three corresponding solid circles in the results of the proposed method. In addition, the tomograms obtained by beamforming and Capon are similar, but the resolution of the latter is higher. The tomogram obtained by MUSIC is not as complete as those of the previous two methods. In the future, we will try to solve this problem by using the data of multiple polarization channels at the same time, and we will investigate in detail the performance of the proposed method in forest height inversion.
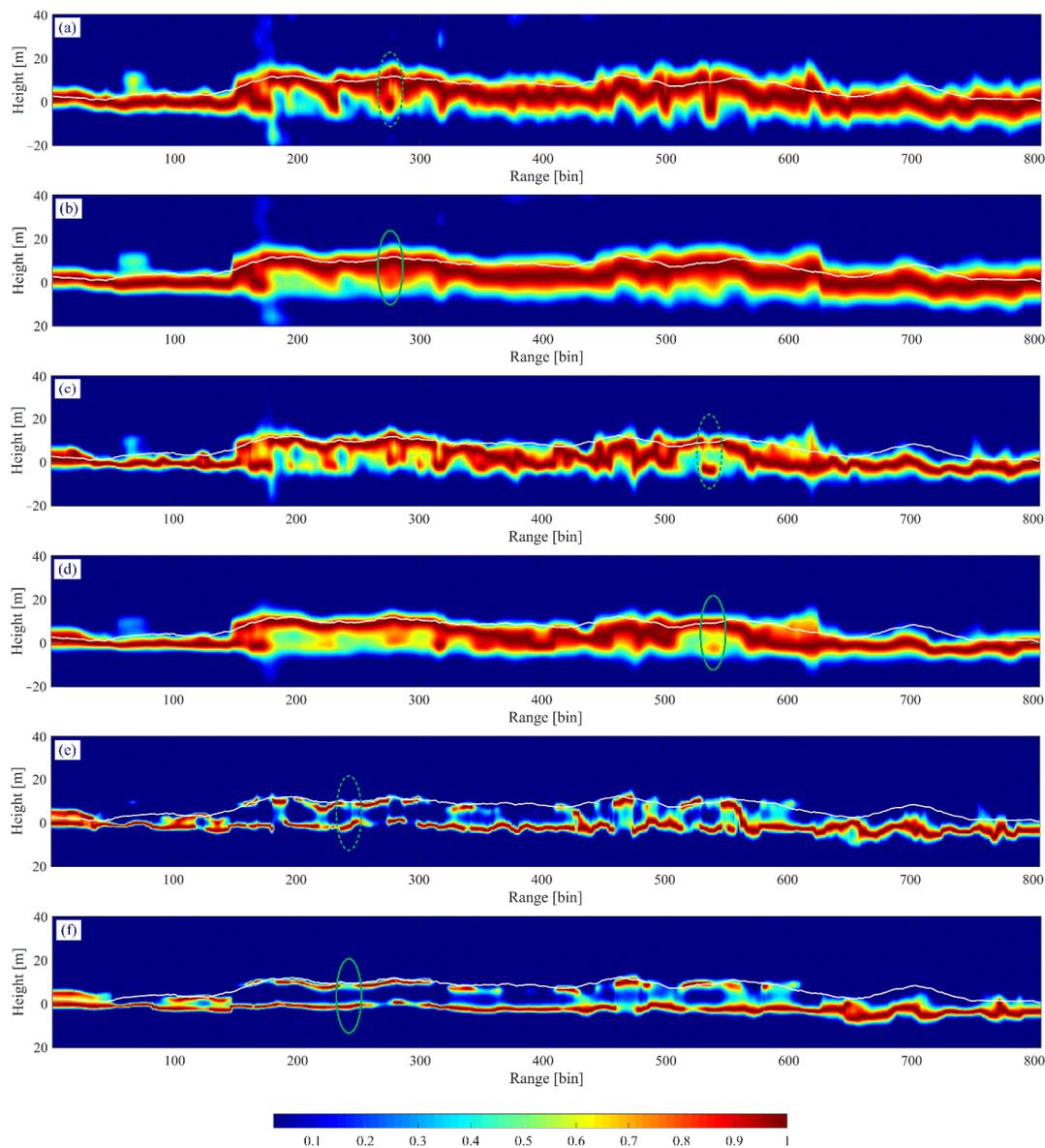
Figure 12. Profile line spectrum estimation results for the HV polarization channel (terrain removed): the solid white line is the LiDAR canopy height model; (**a**), (**c**), and (**e**) are, respectively, the estimation results of LM-based beamforming, Capon, and MUSIC; and (**b**), (**d**), and (**f**) are the estimation results of NLM-based beamforming, Capon, and MUSIC.

5. Conclusions

In this paper, we utilized the idea of NLM filtering into the estimation of the CM in tomographic estimation, to solve the problems caused by the inaccurate estimation of the CM, such as the mixing and superposition of different scattering mechanisms, the inaccurate tomographic spectrum, the loss of detail information, and the low estimation accuracy of the terrain information. The three classical spectral estimation methods (i.e., beamforming, Capon, and MUSIC) based on the NLM method are mainly used to perform tomographic focusing. Specifically, these three methods first calculate the SCM of all the pixels in the target region. Secondly, for each target pixel, in its search window, the similarity, including the spatial and radiometric similarity between each neighborhood pixel and the target pixel, is calculated through a matching window. The similarity weight between each neighborhood pixel and the center pixel is then calculated using the spatial and radiometric similarity. Finally, the optimal CM of the target pixel is estimated from the SCM of each neighborhood pixel and its corresponding weights. For the purpose

of verifying the feasibility and effectiveness of this method, we used BioSAR 2008 data obtained from the boreal forest of northern Sweden for the experimental verification. Tomogram analysis and underlying topography inversion were carried out. Moreover, in this paper, we have made a comparison between the consequences of the traditional LM method and the results of the NLM method. The consequences show that the NLM method performs significantly better than the LM method. The NLM method can not only better separate the scattering phase centers between the ground and the canopy, but it also shows a better inversion performance in the places where the topographic relief changes greatly. Moreover, the inversion accuracy of the NLM method is improved by more than 30% when compared with that of the LM method.

In conclusion, the NLM method considers giving different weights to the non-local neighborhood pixels with different similarity to the target pixel, to suppress the interference information and achieve more accurate estimation of the target pixel value. However, this makes the calculation slower and the estimation efficiency is reduced, which is something we need to investigate and improve upon in the future.

Author Contributions: Conceptualization, X.P. (Xing Peng); methodology, X.P. (Xing Peng) and Y.W.; software, Y.W.; validation, S.L., X.P. (Xiong Pan), Q.X., and Y.D.; formal analysis, H.F. and J.Z. and X.L.; writing—original draft preparation, Y.W.; writing—review and editing, X.P. (Xing Peng); funding acquisition, X.P. (Xing Peng) Q.X.,Y.D. and X.L. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Dinh, H.T.M.; Toan, T.L.; Rocca, F.; Tebaldini, S.; d'Alessandro, M.M.; Villard, L. Relating P-Band Synthetic Aperture Radar Tomography to Tropical Forest Biomass. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 967–979. [CrossRef]
2. Aghababaee, H.; Ferraioli, G.; Ferro-Famil, L.; Huang, Y.; d'Alessandro, M.M.; Pascazio, V.; Schirinzi, G.; Tebaldini, S. Forest SAR Tomography: Principles and Applications. *IEEE Geosci. Remote Sens. Mag.* **2020**, *8*, 30–45. [CrossRef]
3. Dinh, H.T.M.; Le Toan, T.; Rocca, F.; Tebaldini, S.; Villard, L.; Réjou-Méchain, M.; Phillips, O.L.; Feldpausch, T.R.; Dubois-Fernandez, P.; Scipal, K. SAR tomography for the retrieval of forest biomass and height: Cross-validation at two tropical forest sites in French Guiana. *Remote Sens. Environ.* **2016**, *175*, 138–147. [CrossRef]
4. Ngo, Y.-N.; Ho Tong Minh, D.; Moussawi, I.; Villard, L.; Ferro-Famil, L.; d'Alessandro, M.M.; Tebaldini, S.; Albinetv, C.; Scipal, K.; Le Toan, T. Afrisar-Tropisar: Forest Biomass Retrieval by P-Band Sar Tomography. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 8675–8678.
5. Peng, X.; Li, X.; Wang, C.; Zhu, J.; Liang, L.; Fu, H.; Du, Y.; Yang, Z.; Xie, Q. SPICE-Based SAR Tomography Over Forest Areas Using a Small Number of P-band Airborne F-SAR Images Characterized by Non-Uniformly Distributed Baselines. *Remote Sens.* **2019**, *11*. [CrossRef]
6. Reigber, A.; Moreira, A. First demonstration of airborne SAR tomography using multibaseline L-band data. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 2142–2152. [CrossRef]
7. Gini, F.; Lombardini, F.; Montanari, M. Layover solution in multibaseline SAR interferometry. *IEEE Trans. Aerosp. Electron. Syst.* **2002**, *38*, 1344–1356. [CrossRef]
8. Lombardini, F.; Reigber, A. Adaptive spectral estimation for multibaseline SAR tomography with airborne L-band data. In Proceedings of the 2003 IEEE International Geoscience and Remote Sensing Symposium (IEEE Cat. No. 03CH37477), Toulouse, France, 21–25 July 2003; pp. 2014–2016.

9.   Nannini, M.; Scheiber, R.; Moreira, A. Estimation of the minimum number of tracks for SAR tomography. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 531–543. [CrossRef]

10. Yu, H.; Zhang, Z. The Performance of Relative Height Metrics for Estimation of Forest Above-Ground Biomass Using L-and X-Bands TomoSAR Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 1857–1871. [CrossRef]

11. Tebaldini, S.; Rocca, F. Multibaseline polarimetric SAR tomography of a boreal forest at P-and L-bands. *IEEE Trans. Geosci. Remote Sens.* **2011**, *50*, 232–246. [CrossRef]

12. d'Alessandro, M.M.; Tebaldini, S. Digital terrain model retrieval in tropical forests through P-band SAR tomography. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6774–6781. [CrossRef]

13. El Moussawi, I.; Ho Tong Minh, D.; Baghdadi, N.; Abdallah, C.; Jomaah, J.; Strauss, O.; Lavalle, M. L-Band UAVSAR tomographic imaging in dense forests: Gabon forests. *Remote Sens.* **2019**, *11*, 475. [CrossRef]

14. Yang, X.; Tebaldini, S.; d'Alessandro, M.M.; Liao, M. Tropical Forest Height Retrieval Based on P-Band Multibaseline SAR Data. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 451–455. [CrossRef]

15. Zhang, Q.; Huang, Y.; Schwaebisch, M.; Mercer, B.; Wei, M. Forest height estimation using single-pass dual-baseline L-band PolInSAR data. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 7055–7058.

16. Liang, L.; Li, X.; Gao, X.; Guo, H. Multibaseline polarimetric synthetic aperture radar tomography of forested areas using wavelet-based distribution compressive sensing. *J. Appl. Remote Sens.* **2015**, *9*, 095048. [CrossRef]

17. Pardini, M.; Tello, M.; Cazcarra-Bes, V.; Papathanassiou, K.P.; Hajnsek, I. L-and P-band 3-D SAR reflectivity profiles versus lidar waveforms: The AfriSAR case. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 3386–3401. [CrossRef]

18. Labriere, N.; Tao, S.; Chave, J.; Scipal, K.; Le Toan, T.; Abernethy, K.; Alonso, A.; Barbier, N.; Bissiengou, P.; Casal, T. In situ reference datasets from the TropiSAR and AfriSAR campaigns in support of upcoming spaceborne biomass missions. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 3617–3627. [CrossRef]

19. Cazcarra-Bes, V.; Pardini, M.; Tello, M.; Papathanassiou, K.P. Comparison of tomographic SAR reflectivity reconstruction algorithms for forest applications at L-band. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 147–164. [CrossRef]

20. Huang, Y.; Ferro-Famil, L.; Reigber, A. Under-foliage object imaging using SAR tomography and polarimetric spectral estimators. *IEEE Trans. Geosci. Remote Sens.* **2011**, *50*, 2213–2225. [CrossRef]

21. Ramachandran, N.; Saatchi, S.; Tebaldini, S.; d'Alessandro, M.M.; Dikshit, O. Evaluation of P-Band SAR Tomography for Mapping Tropical Forest Vertical Backscatter and Tree Height. *Remote Sens.* **2021**, *13*, 1485. [CrossRef]

22. Huang, Y.; Zhang, Q.; Ferro-Famil, L. Forest Height Estimation Using a Single-Pass Airborne L-Band Polarimetric and Interferometric SAR System and Tomographic Techniques. *Remote Sens.* **2021**, *13*, 487. [CrossRef]

23. Ferro-Famil, L.; Huang, Y.; Pottier, E. Principles and applications of polarimetric SAR tomography for the characterization of complex environments. In Proceedings of the VIII Hotine-Marussi Symposium on Mathematical Geodesy, Rome, Italy, 17–21 June 2015; pp. 243–255.

24. Guliaev, R.; Pardini, M.; Papathanassiou, K. A Comparison of Function Bases for Polarization Coherence Tomography in Forest Scenarios. In Proceedings of the European Conference on Synthetic Aperture Radar, EUSAR 2021, Virtual, 28 March–1 April 2021.

25. Cazcarra-Bes, V.; Pardini, M.; Papathanassiou, K. Definition of Tomographic SAR Configurations for Forest Structure Applications at L-Band. *IEEE Geosci. Remote Sens. Lett.* **2020**. [CrossRef]

26. Aghababaee, H.; Ferraioli, G.; Schirinzi, G.; Sahebi, M.R. The role of nonlocal estimation in SAR tomographic imaging of volumetric media. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 729–733. [CrossRef]

27. D'Hondt, O.; López-Martínez, C.; Guillaso, S.; Hellwich, O. Nonlocal filtering applied to 3-D reconstruction of tomographic SAR data. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 272–285. [CrossRef]

28. Asopa, U.; Kumar, S. UAVSAR Tomography for Vertical Profile Generation of Tropical Forest of Mondah National Park, Gabon. *Earth Space Sci.* **2020**, *7*. [CrossRef]

29. Li, Z.; Zhang, P.; Qiao, H.; Zhao, C.; Zhou, J.; Huang, L. Advances in Information Extraction of Surface Parameters Using Tomographic SAR. *J. Radars* **2021**, *10*, 116–130. [CrossRef]

30. Fornaro, G.; Serafino, F.; Soldovieri, F. Three-dimensional focusing with multipass SAR data. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 507–517. [CrossRef]

31. Aguilera, E.; Nannini, M.; Reigber, A. Wavelet-based compressed sensing for SAR tomography of forested areas. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 5283–5295. [CrossRef]

32. Homer, J.; Longstaff, I.; Callaghan, G. High resolution 3-D SAR via multi-baseline interferometry. In Proceedings of the IGARSS'96 1996 International Geoscience and Remote Sensing Symposium, Lincoln, NE, USA, 31 May 1996; pp. 796–798.

33. Stoica, P.; Moses, R.L. *Spectral Analysis of Signals*; Prentice Hall: Upper Saddle River, NJ, USA, 2005.

34. She, Z.; Gray, D.; Bogner, R.; Homer, J. Three-dimensional SAR imaging via multiple pass processing. In Proceedings of the IEEE 1999 International Geoscience and Remote Sensing Symposium, IGARSS'99 (Cat. No. 99CH36293), Hamburg, Germany, 28 June–2 July 1999; pp. 2389–2391.

35. Kumar, S.; Joshi, S.K.; Govil, H. Spaceborne PolSAR tomography for forest height retrieval. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 5175–5185. [CrossRef]

36. D'Hondt, O.; Guillaso, S.; Hellwich, O. Iterative bilateral filtering of polarimetric SAR data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1628–1639. [CrossRef]

37. Lee, J.-S.; Grunes, M.R.; De Grandi, G. Polarimetric SAR speckle filtering and its implication for classification. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 2363–2373. [CrossRef]

38. Barbaresco, F. Interactions between symmetric cone and information geometries: Bruhat-tits and siegel spaces models for high resolution autoregressive doppler imagery. In Proceedings of the LIX Fall Colloquium on Emerging Trends in Visual Computing, Palaiseau, France, 18–20 November 2008; pp. 124–163.

39. European Space Agency. *Technical Assistance for the Development of Airborne SAR and Geophysical Measurements during the BioSAR 2008 Experiment*; Final Report; European Space Agency: Paris, France, 2009.

40. Peng, X.; Li, X.; Du, Y.; Xie, Q. Forest Height Estimation from a Robust TomoSAR Method in the Case of Small Tomographic Aperture with Airborne Dataset at L-Band. *Remote Sens.* **2021**, *13*, 2147. [CrossRef]