A Deep Learning Solution for Height Estimation on a Forested Area Based on Pol-TomoSAR Data

Wenyu Yang, *Graduate Student Member, IEEE*, Sergio Vitale, *Member, IEEE*, Hossein Aghababaei, *Senior Member, IEEE*, Giampaolo Ferraioli, *Senior Member, IEEE*, Vito Pascazio, *Senior Member, IEEE*, and Gilda Schirinzi, *Senior Member, IEEE*

***Abstract*— Forest height and underlying terrain reconstruction is one of the main aims in dealing with forested areas. Theoretically, synthetic aperture radar tomography (TomoSAR) offers the possibility to solve the layover problem, making it possible to estimate the elevation of scatters located in the same resolution cell. This article describes a deep learning approach, named tomographic SAR neural network (TSNN), which aims at reconstructing forest and ground height using multipolarimetric multibaseline (MPMB) SAR data and light detection and ranging (LiDAR)-based data. The reconstruction of the forest and ground height is formulated as a classification problem, in which TSNN, a feedforward network, is trained using covariance matrix elements as input vectors and quantized LiDAR-based data as the reference. In our work, TSNN is trained and tested with P-band MPMB data acquired by ONERA over Paracou region of French Guiana in the frame of the European Space Agency's campaign TROPISAR and LiDAR-based data provided by the French Agricultural Research Center. The novelty of the proposed TSNN is related to its ability to estimate the height with a high agreement with LiDAR-based measurement and actual height with no requirement for phase calibration. Experimental results of different covariance window sizes are included to demonstrate that TSNN conducts height measurement with high spatial resolution and vertical accuracy outperforming the other two TomoSAR methods. Moreover, the conducted experiments on the effects of phase errors in different ranges show that TSNN has a good tolerance for small errors and is still able to precisely reconstruct forest heights.**

***Index Terms*— Deep learning (DL), forest height, polarimetry, synthetic aperture radar (SAR), tomography, underlying topography.**

Manuscript received 18 August 2022; revised 13 January 2023; accepted 28 April 2023. Date of publication 8 May 2023; date of current version 26 May 2023. This work was supported by the Italian Space Agency (ASI) through "MEFISTO—Metodi di Elaborazione di dati SAR multi-Frequenza per Il monitoraggio del disseSTO idrogeologico" within the Framework of BANDO ASI DC-UOT-2019-017 "Studio di nuovi metodi e tecniche basati sull'utilizzo di dati sar multimissione/multifrequenza" under Grant ASI N. 2021-10-UO CUP n. F65F21000630005. *(Corresponding author: Giampaolo Ferraioli.)*

Wenyu Yang is with the Dipartimento di Ingegneria, Università degli Studi di Napoli "Parthenope," 80143 Naples, Italy (e-mail: wenyu.yang001@studenti.uniparthenope.it).

Sergio Vitale, Vito Pascazio, and Gilda Schirinzi are with the Dipartimento di Ingegneria, Università degli Studi di Napoli "Parthenope," 80143 Naples, Italy, and also with the National Inter-University Consortium for Telecommunications (CNIT), 80126 Naples, Italy (e-mail: sergio.vitale@uniparthenope.it; vito.pascazio@uniparthenope.it; gilda.schirinzi@uniparthenope.it).

Hossein Aghababaei is with the Department of Earth Observation Science, Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede 7514AE, The Netherlands (e-mail: h.aghababaei@utwente.nl).

Giampaolo Ferraioli is with the Dipartimento di Scienze e Tecnologie, Università degli Studi di Napoli "Parthenope," 80143 Naples, Italy, and also with the National Inter-University Consortium for Telecommunications (CNIT), 80126 Naples, Italy (e-mail: giampaolo.ferraioli@uniparthenope.it).

Digital Object Identifier 10.1109/TGRS.2023.3274395

I. Introduction

As a fundamental part of forest characterization, the estimation of the forest height seeks to offer a significant indicator of the productivity of forests and of the biomass level [1]. In this regard, synthetic aperture radar (SAR) and light detection and ranging (LiDAR) are competitive systems providing information on the vertical distribution and internal structure of the vegetation [2]. In particular, SAR tomography (TomoSAR) is a technique with the capability of providing high-resolution 3-D reflectivity profiles along the azimuth, range, and elevation coordinates by synthesizing an additional aperture in the elevation direction [3]. The 3-D image focusing is performed through the coherent combination of multibaseline (MB) data, allowing the retrieval of the power backscattered from the distributed targets along the vertical direction. The TomoSAR reconstruction of the forest height and the underlying topography relies heavily on the discrimination and exact positions of the phase centers of the scattering from the canopy and ground. For this purpose, polarimetric TomoSAR, which is sensitive to the shape, direction, and dielectric properties of the scatterers, has been widely used for the height reconstruction of forested areas [4], [5].

There are several TomoSAR inversion algorithms, ranging from the classical Fourier-based methods to super-resolution techniques. Among them, the conventional Fourier-based algorithms may be affected by grating lobes because of uneven and few baseline sampling [6], [7]. This problem has been addressed by baseline interpolation [8] or by using super-resolution techniques, such as the Capon adaptive filtering [2], [9], Multiple Signal Classification (MUSIC) [10], singular value decomposition analysis [7], and compressive sensing [11], [12]. The complex vertical reflectivity profiles of forested areas and the lack of a larger number of uniformly distributed MB data limit the vertical resolution of these techniques. Under these circumstances, the location of the ground and canopy from the reconstructed vertical profiles becomes a difficult task. To overcome this problem, decomposition methods have been proposed as a promising solution in the localization of ground and canopy heights [4], [13], [14], [15]. These techniques are usually based on the identification and separation of different scattering mechanisms that occur in forested areas. Thus, according to the expected scattering mechanisms (i.e., ground-trunk double bounce and volumetric scattering through the canopy), the contributions of ground and canopy can be determined in the estimated sample covariance matrix of the polarimetric MB datasets.

Another system providing measurements sensitive to 3-D forest structure parameters at high spatial resolutions is LiDAR. Measurements acquired by LiDAR are based on the round trip time taken by each laser pulse to travel the distance between the sensor and the target [16], [17]. Research has demonstrated that LiDAR can characterize the structural complexity and associated functional properties of natural landscapes relevant to ecological investigations by providing vertical and volumetric profiles of forest vegetation. LiDAR metrics have proven useful for determining the forest canopy height and structure such as the hundredth percentile of the cumulative waveform energy relative to the ground (RH100) [18], [19]. LiDAR data typically have superior vertical accuracies and are not impacted by TomoSAR-related errors, such as speckle noise and temporal decorrelation. However, LiDAR acquisitions are constrained at a local scale because of the small swath and the noticeable workload involved in the acquisition and processing of LiDAR data per unit area, which is higher than the one required by SAR imagery.

LiDAR data are commonly used to assess the ability of SAR tomography to monitor and estimate forest structure parameters. Anyway, it has to be considered that when low-frequency SAR systems (such as P- and L-band systems) are used, the penetration of radiation into the canopy causes the height of the effective scattering phase centers to be within the canopy itself so that LiDAR and TomoSAR canopy surface reconstructions can exhibit remarkable differences.

More recently, a number of studies have sought to generate forest height estimates by exploiting the synergy of LiDAR and MB PolSAR data, with the objective of producing wide-coverage forest height maps with high spatial resolution and vertical accuracy [20]. In [21], support vector machine (SVM) has been employed to extrapolate LiDAR-based canopy height utilizing TomoSAR inverted parameters. In addition, several researchers have attempted to improve the performance of polarimetric synthetic aperture radar interferometry (PolInSAR)-based forest height estimation by extracting prior information (ground phase, mean extinction, and so on) from the LiDAR metrics [20], [22], [23]. However, these methods still have a strong dependency on the simplified models.

Due to the ability of neural networks to build a hierarchy of abstract representations of the data and, so, to act as a nonlinear function able of representing very complicated mathematical models [24], deep learning (DL)-based methods have become a fundamental methodology for different remote sensing image processing tasks [25] as well as for SAR image processing [26]. Following this trend, some methods start to exploit the potential of DL for the TomoSAR 3-D reconstruction. A first preliminary DL approach has been proposed for TomoSAR applications in urban areas [27]. In this method, the estimation of the scatterers' position is formulated as a classification problem, with the classes indicating all the discretized possible positions, within the elevation extension of the scene, of the single scatterers that are possibly present in each range–azimuth resolution cell. In the context of polarimetric SAR tomography, a method named PolGAN has been proposed in [20]. Instead of using the inversion of the data model, PolGAN reformulates the forest height estimation as an image pan-sharpening task to generate forest height estimation with high spatial resolution and vertical accuracy, based on PolInSAR and LiDAR inputs together with feedback provided by the discriminators of PolGAN. The promising results of these solutions inspired us to explore the DL potential on the height reconstruction task. In this article, the multisource (LiDAR and TomoSAR) forest height and ground height reconstruction is reformulated as a classification task aiming at the estimation of forest height and terrain topography adopting a DL approach, without using any inversion model. In particular, we develop a simple multilayer perceptron [28] architecture that is targeted for the classification of multipolarimetric multibaseline (MPMB) SAR images at a pixel level. The proposed architecture is trained for extracting features and recognizing patterns within the vector containing the values of a single MPMB SAR pixel and exploiting them for retrieving the corresponding LiDAR-based height value. The LiDAR data are quantized in order to obtain a set of height reference classes the single MPMB SAR pixels belong to, and the inputs of the network are the elements of the covariance matrix characterizing the MPMB SAR images. In this article, starting from the same MPMB SAR data, the proposed architecture has been trained for both forest height and ground height reconstruction using the LiDAR-based canopy height and ground height values as references, respectively. Moreover, the ability of the network in working in the absence of data calibration, an analysis of the robustness with respect to the covariance window sizes and with respect to phase errors within the MPMB SAR images, together with a comparison with state-of-the-art methods have been carried out on the Paracou region. This article is organized as follows. In Section II, the methodology is presented in terms of the proposed workflow, network architecture, and training. The description of the testing study area and the analysis of the experimental results are presented in Section III.

II. Methodology

Forest and ground height estimation is formulated as a classification task: a solution based on a fully connected neural network is proposed, named tomographic SAR neural network (TSNN). In Section II-A, the details about the input data and the proposed workflow are provided. The TSNN architecture is described in Section II-B. Finally, the training of TSNN is discussed in II-D.

A. Data and Workflow

Based on the TomoSAR principle, in order to estimate the 3-D reflectivity function along the coordinates of azimuth x, range r, and elevation s, a stack of $3N$ single-look fully polarimetric complex SAR images are acquired. Some preprocessing operations (such as coregistration, phase flattening, and calibration) have to be applied before the tomographic processing.

After such preprocessing operations, for a fixed range–azimuth pixel, the MPMB $3N \times 1$ data vector $\boldsymbol{y}$ is defined as

$$\begin{aligned} \boldsymbol{y} &= \begin{bmatrix} \boldsymbol{y}_1, & \boldsymbol{y}_2, & \ldots & \boldsymbol{y}_N \end{bmatrix}^T \\ \boldsymbol{y}_i &= \begin{bmatrix} S_i^{\mathrm{HH}}, & \sqrt{2} S_i^{\mathrm{HV}}, & S_i^{\mathrm{VV}} \end{bmatrix} \end{aligned} \tag{1}$$

where S_i^{PQ} denotes the backscattering using transmitted and received pairs of polarizations $P, Q = \{H, V\}$, and T is the transpose operator.

In forest applications, the scattering response by the involved distributed media has a random behavior, due to the presence of multiple interacting scatterers in random positions. Then, height estimation is usually based on the analysis of the data covariance matrix $\mathbf{R}$ of size $3N \times 3N$, which in the assumption of a zero mean data vector is given by

$$\mathbf{R} = E\{\boldsymbol{y}\boldsymbol{y}^\dagger\} = \begin{bmatrix} \mathbf{C}_{11} & \Omega_{12} & \cdots & \Omega_{1N} \\ \Omega_{21} & \mathbf{C}_{22} & \cdots & \Omega_{2N} \\ \cdots & \cdots & \ddots & \vdots \\ \Omega_{N1} & \Omega_{N2} & \cdots & \mathbf{C}_{NN} \end{bmatrix} \tag{2}$$

where the superscript $\dagger$ denotes the Hermitian transpose and $E\{\cdot\}$ is the statistical expectation operator.

In (2), the covariance matrix $\mathbf{R}$ of the data vector $\boldsymbol{y}$, defined by (1), is decomposed in $N \times N$ submatrices, each of size 3×3. The diagonal submatrices $\mathbf{C}_{ii}$ are related to polarimetry, whereas the off-diagonal submatrices Ω_{ij} are 3×3 complex cross correlation matrices that include both polarimetric and interferometric information [13].

It is noted that the computation of the matrix $\mathbf{R}$ involves a statistical expectation operation to be performed on different realizations of the data vector $\boldsymbol{y}$. Since different realizations of the random data vector are not available, the statistical expectation is performed by means of a spatial averaging operation on a set of neighboring range–azimuth pixels lying in an assigned window surrounding the considered pixel.

For each range–azimuth pixel of the MPMB image stack, we take the $3N$ diagonal elements of the matrix $\mathbf{R}$ and the elements in the first row representing both polarimetric and interferometric information as the input to TSNN. In particular, the complex elements of the first row are separated into real parts and imaginary parts. Discarding the imaginary unit, we get the real input feature vector $\mathbf{I}$ of a range–azimuth pixel with the size $M \times 1$ ($M = 3N + (3N - 1) \times 2$).

The LiDAR-based height values are taken as ground truth. First, in order to have a spatial resolution comparable with SAR data one, a spatial averaging operation is conducted on the LiDAR-based canopy height model (CHM) and Digital Terrain Model (DTM). The size of the adopted averaging window is the same as the one used in the SAR covariance matrix computation. Then, the filtered LiDAR-based is quantized with a 1-m step, generating ζ labels $S = [s_1, s_2, \cdots, s_\zeta]$ for the classification network. Indeed, the quantized CHM and DTM values represent the reference height classes of forest and ground to be predicted for each range–azimuth pixel.

Fig. 1 represents the data flowchart, where TSNN-CHM is TSNN trained with quantized CHM as labels and TSNN-DTM is TSNN trained with quantized DTM as labels.

B. TSNN Architecture

In order to extract the necessary mutual information from MPMB SAR data, a fully connected architecture has been designed for the proposed TSNN, whose description is provided in the following.

Supposing that the lth fully connect layer is composed with N^l neurons, each neuron is equipped with learnable weights w_{jk}^l and the bias b_j^l, so the output of the jth neuron of the lth layer is

$$z_j^l = \sum_{k=1}^{N^{l-1}} w_{jk}^l z_k^{l-1} + b_j^l, \quad j = 1, 2, \ldots, N^l \tag{3}$$

where N^{l-1} is the number of neurons of the $(l-1)$th layer. Thus, the output vector of the lth layer is

$$\mathbf{z}^l = \mathbf{w}^l \mathbf{z}^{l-1} + \mathbf{b}^l \tag{4}$$

where the tensor $\mathbf{w}^l$ is with the dimension $(N^l \times N^{l-1})$ and $\mathbf{b}^l$ is with the dimension $(N^l \times 1)$.

Let $\mathbf{W}^l = (\mathbf{w}^l, \mathbf{b}^l)$ be the set of parameters of the lth layer. The relative function performed can be described as

$$f_l\left(\mathbf{z}^{(l-1)}, \mathbf{W}^l\right) = \sigma\left(\mathbf{w}^l \mathbf{z}^{l-1} + \mathbf{b}^l\right) \tag{5}$$

where σ is the activation function and, therefore, the overall fully connected network function can be described as [27]

$$f(\mathbf{I}, \mathbf{W}) = f_L\left(f_{L-1}\left(f_1\left(\mathbf{x}, \mathbf{W}^{(1)}\right), \ldots, \mathbf{W}^{(L-1)}\right), \mathbf{W}^l\right) \tag{6}$$

with L representing the number of layers. The output of the network is represented by a vector of scores for each possible class

$$\hat{H} = f(\mathbf{I}, \mathbf{W}) = \left[h_1, h_2, \ldots, h_\zeta\right]. \tag{7}$$

Here, the proposed TSNN consists of nine fully connected layers with 400 neurons for the first eight layers, while for the last one, the number of neurons matches the number of considered classes. The structure is shown in Fig. 2 where FC is a fully connected layer. All the layers, but the last, are followed by the rectified linear unit (ReLU) activation function that ensures the stable training procedure and fast convergence [29]. The core element of the proposed network is the fully connected layer followed by ReLU, which builds a hierarchical set of features. The resulting feature maps are used to discriminate classes with the last fully connected output layer.

C. Loss Function

Multiclass cross-entropy function [30] is used as a loss function when adjusting network weights $\mathbf{W}$ during training. The smaller the loss the better the model. A perfect model is equipped with a cross-entropy loss equal to 0. The cross

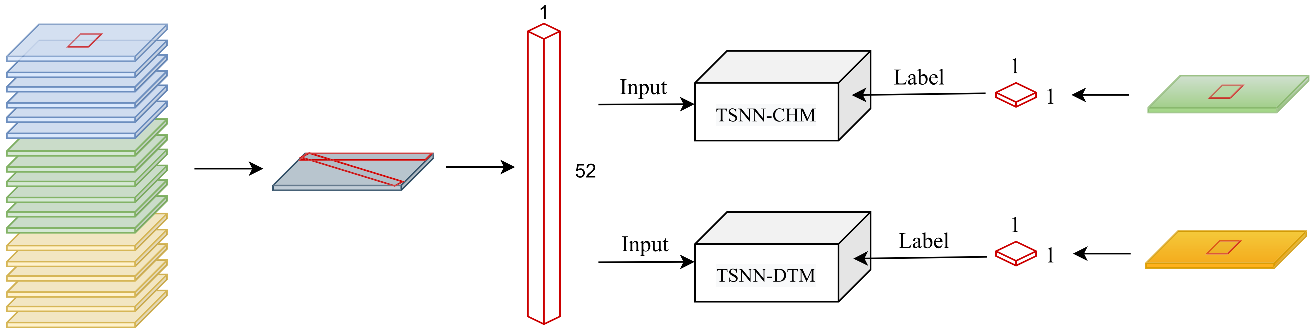


Fig. 1. Workflow of the proposed TSNN. For each range–azimuth pixel of the MPMB image stack, the $3N$ diagonal elements of the matrix R and the elements in the first are stacked and considered as input of the network. The complex elements are separated as real and imaginary parts and stacked as well. The reference labels consist of the corresponding quantized LiDAR height values: CHM in case of training targeted for prediction of forest height and DTM in case of training targeted for prediction of ground height.

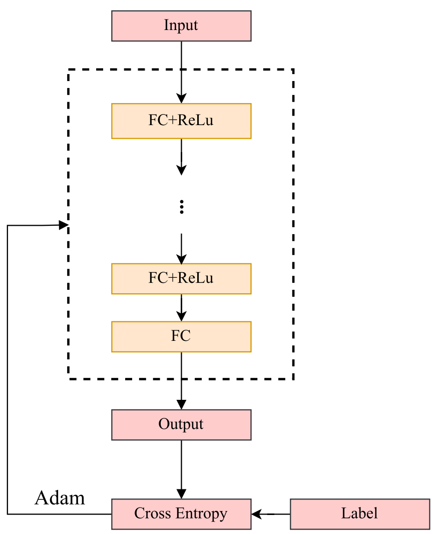


Fig. 2. Architecture of TSNN. The architecture is composed of nine fully connected layers each followed by the ReLU activation function but the last. Each layer is composed of 400 neurons, except the last whose neurons need to match the number of considered classes.

entropy between two statistical distributions $\mathbf{p}$ and $\mathbf{q}$ is defined as

$$\mathrm{CE}(\mathbf{p}, \mathbf{q}) = -\sum_{x} p(x) \log q(x). \tag{8}$$

In this case, $\mathbf{p}$ is the reference distribution containing a unitary value at the position corresponding to the ith class and $\mathbf{q}$ is the estimated a posteriori probability distribution for each class obtained by applying the softmax function to the output $\hat{H}$. Thus, the estimated distribution can be written as $\mathbf{q} = [e^{h_j}/\sum_j e^{h_j} \mid j = 1, \ldots, \zeta]$ and, $\mathbf{p} = [0, \ldots, 1, \ldots, 0]$ with a unitary values at the ith position. Thus, the cross-entropy loss is computed as

$$L_{\mathrm{CE}} = -\log\left(\frac{e^{h_i}}{\sum_j e^{h_j}}\right). \tag{9}$$

D. Network Training

TSNNs are implemented with the Pytorch framework and optimized by minimizing (9) through the Adam optimizer [31] with the parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. We start training on GeForce GTX 1080Ti GPU with 12 GB of memory. The batch size is set as 32 and the learning rate is set as 10^{-4} and is fixed for all 200 epochs. The Xavier initialization [32] has been considered for the weights of each fully connected layer.

Given the characteristics of the available dataset, the class imbalance issue has been addressed. Indeed, generally, class imbalance crucially affects the classification performance: when the frequency of certain classes is much more than the other classes, the classification network may get biased toward the prediction showing much better performance on the majority classes and worse performance on the minority classes even though the loss function presents pretty low value. To handle such a problem, we apply weighting factors to undersample the majority classes and oversample the minority ones. This allows us to provide equal importance to all classes during the training phase [33], [34].

In TomoSAR applications, generally, a phase calibration problem has to be addressed to recover the coherence among the acquisitions by calibrating phase errors caused by atmosphere propagation delays or residual platform motion. In forested areas, due to the difficulty in locating reference targets with stable phases or to the lack of specific assumptions about the phase calibration function, phase calibration entails intricate estimation procedures and frequently yields inaccurate results [35]. In order to test the proposed TSNNs' tolerance to phase miscalibration, the training has been performed by using both calibrated and noncalibrated input data to get TSNNs that are compared on the testing datasets. The testing results are shown in Section III.

III. Experiments

A. Study Area and Dataset

We address the problem of the forest canopy and ground height reconstruction over the forest stand of Paracou region of French Guiana shown in Fig. 3. The stack of data is composed of six fully polarimetric P-band SAR images acquired by the ONERA SETHI airborne system on August 24, 2009, in the frame of the ESA's campaign TROPISAR. The flight lines were in a vertical direction and the Fourier vertical resolution in about 20 m. The flight baselines and acquisition parameters

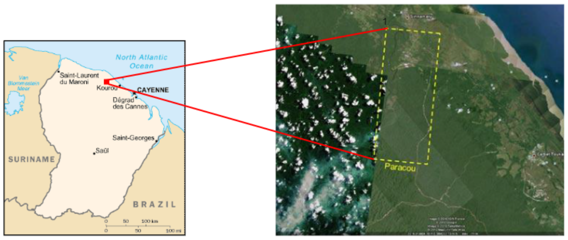


Fig. 3. (Left) Geographic location of Paracou site. (Right) Image coverage.

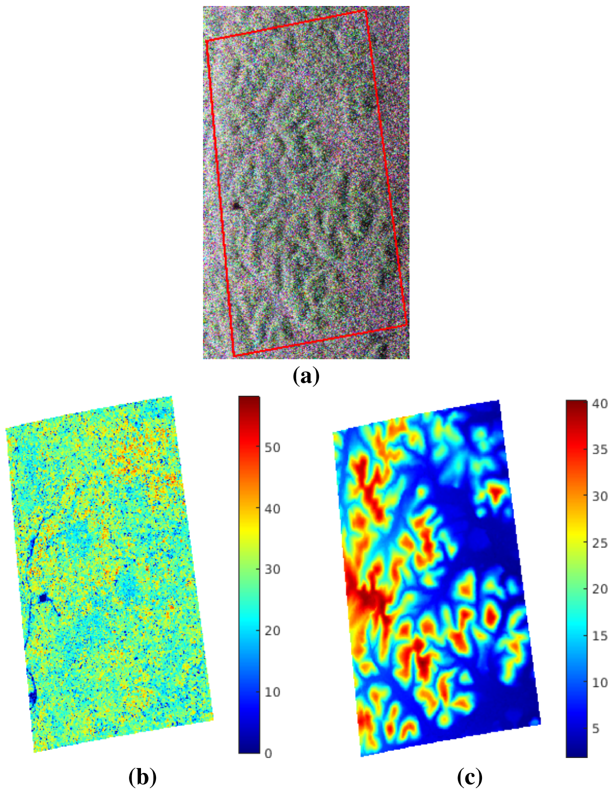


Fig. 4. SAR data and LiDAR-based data. (a) Pauli RGB image of the master acquisition. (b) LiDAR-based CHM. (c) LiDAR-based DTM of ROI of Paracou region.

TABLE I
FLIGHT PLANES AND ACQUISITION PARAMETERS

Baseline Length [m]		System Parameters	
1	0	Wavelength	0.7542 m
2	-14.4879	Flight height	3962 m
3	-30.1163	Incidence angle	35.061°
4	-43.7343	Range resolution	1 m
5	-60.0632	Azimuth resolution	1.245 m
6	-74.9683	Polarization	Full-Pol

of the SAR data are shown in Table I and the Pauli RGB image of the master acquisition of the region of interest (ROI) is shown in Fig. 4(a) (red: HH-VV, green: 2 HV, and blue: HH + VV). Within the site, wide vegetation with a variety of tree species can be recognized. The vegetation height ranges between 0 and 60 m, and the terrain topography is fairly flat with an elevation between 0 and 40 m. The territory

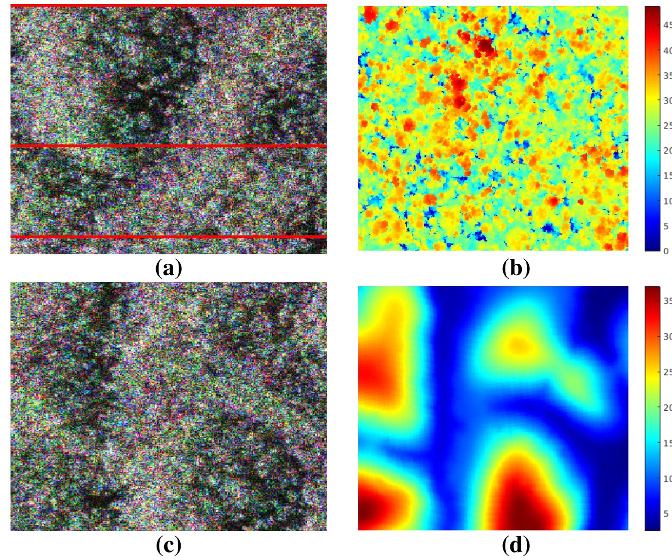


Fig. 5. (a) Pauli RGB image and (b) LiDAR-based CHM of testing patch1. (c) Pauli RGB image and (d) LiDAR-based DTM of testing patch2.

is representative of forest height and underlying topography reconstruction.

For the considered site, the LiDAR-based data are provided by the French Agricultural Research Center for International Development and the Guyafor project. The LiDAR-based DTM was generated by triangular interpolation (TIN) of the ground data. The LiDAR-based CHM was generated by subtracting ground elevation from the raw point cloud z-values and extracting maximum height with a 1-m resolution grid. The LiDAR-based data under WGS84 UTM zone 22 were projected into SAR coordinates based on an ASCII file provided by TROPISAR, providing the forest height and ground height of a geographic position associated with an SAR image position [36]. More details about LiDAR-based data processing are given in [37]. In our study, LiDAR-based data (Fig. 4(b) and (c) for forest and ground, respectively) are used as ground truth to train and verify the accuracy of the reconstructed results.

The number of baselines of TROPISAR is 6, that is, $N = 6$. The covariance matrix $\mathbf{R}$ in (2) size is 18×18, so the input feature vector size $M \times 1$ of each range–azimuth pixel is 52×1. We select testing patch1 [shown in Fig. 5(a)] with the size 300×300 pixels from the ROI of Paracou region as the testing dataset for TSNN-CHM. The rest data are randomly divided into the training dataset (80%) and validating dataset (20%) to train and validate TSNN-CHM. Likewise, testing patch2 [shown in Fig. 5(c)] with the same size is selected as the testing dataset for TSNN-DTM, and the test dataset is randomly divided into the training dataset and validating dataset for TSNN-DTM. Besides, in order to test TSNNs' performance of classification on several classes, the selected testing patches for TSNNs extend a wide range of height covering more height classes, as shown in Fig. 5(b) and (d).

B. Analysis of the Results

In this section, the results of the proposed method are analyzed and discussed. To demonstrate the accuracy and

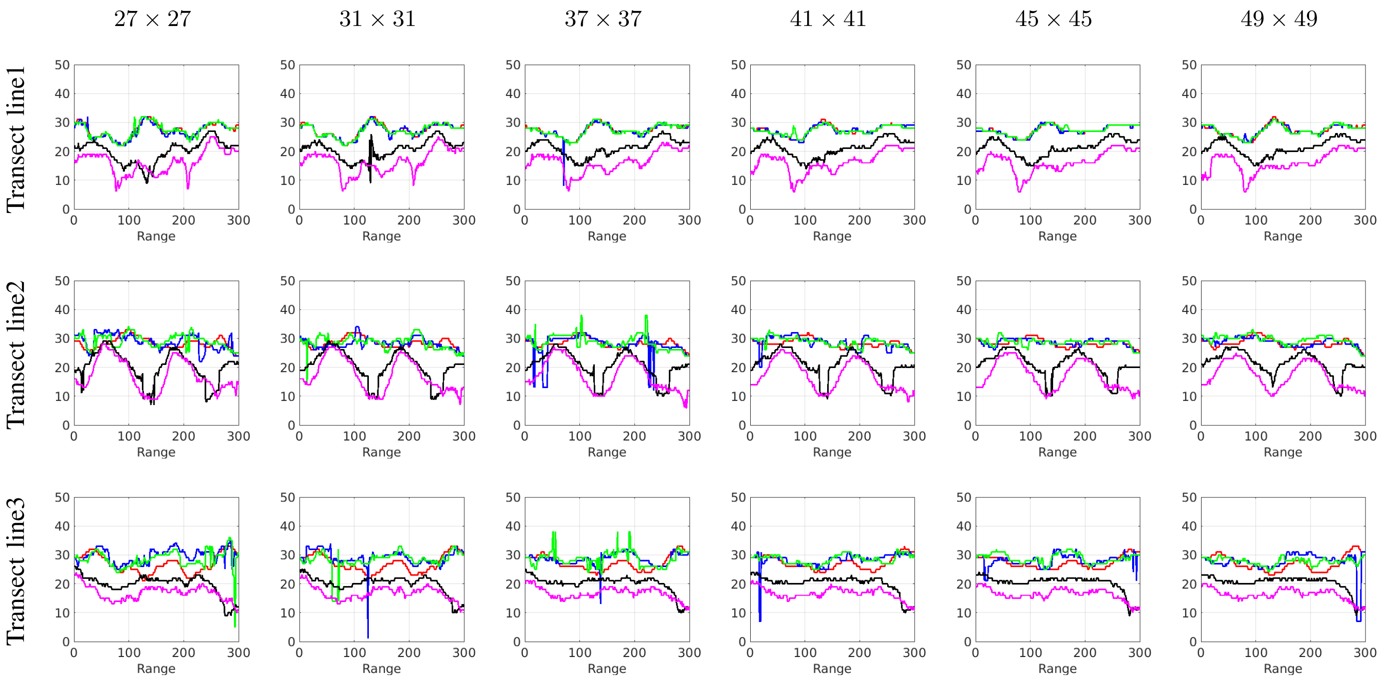


Fig. 6. Profile plots comparison between LiDAR-based CHM (red lines) and forest height estimation over three different transect lines represented in Fig. 5 using TSNN-CHM-C (blue lines), TSNN-CHM-NC (green lines), SKP (black lines), and GLRT (magenta lines).

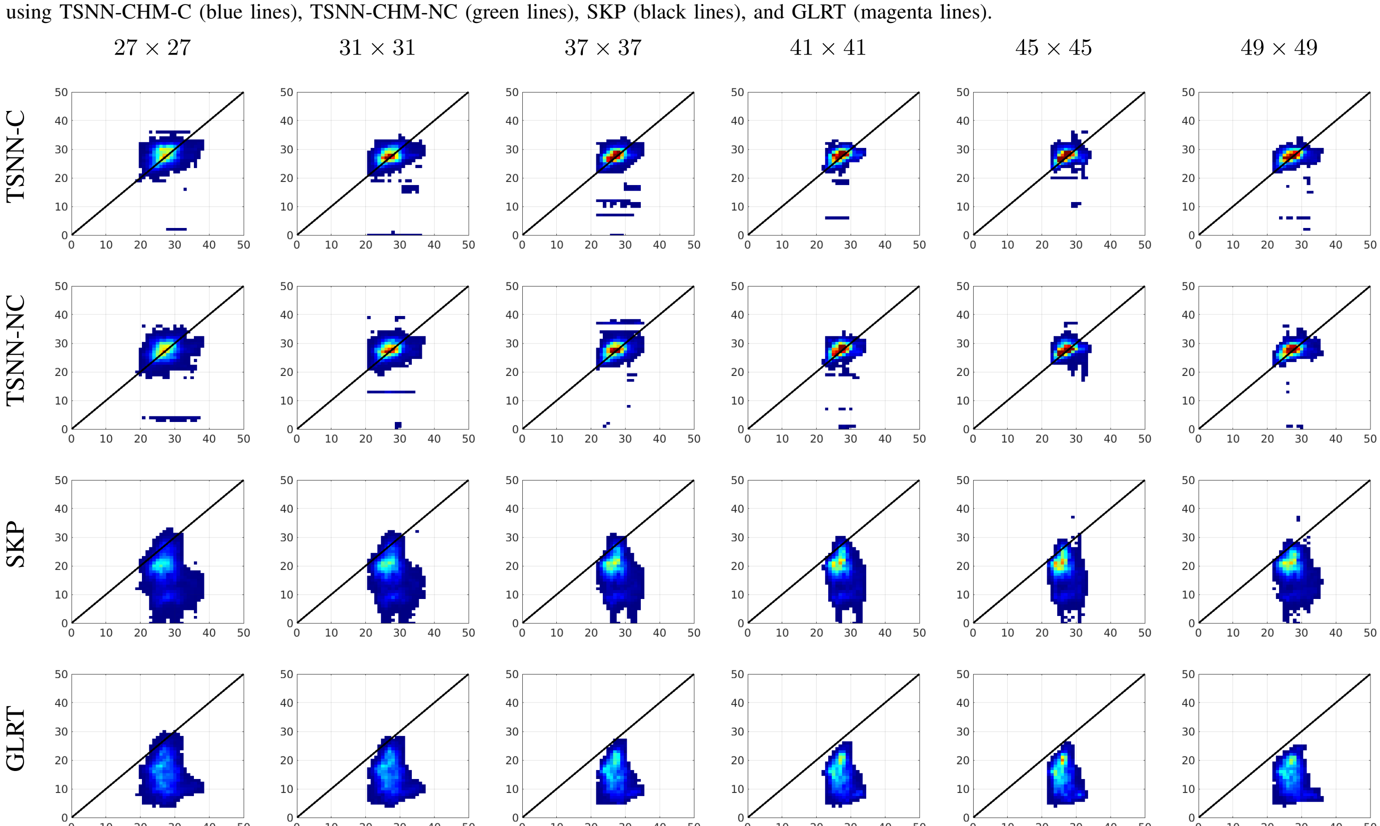


Fig. 7. Joint distribution between the LiDAR-based CHM and forest height predicted by TSNN-CHM-C (first row), TSNN-CHM-NC (second row), SKP (third row) [38], and GLRT method (last row) [15] on different covariance window sizes.

efficiency of the proposed TSNNs, we test them on testing patches and compare them with LiDAR-based ground truth. Moreover, the comparison with the calibrated trained network (see Section II-D) and two representative tomographic methods, namely, the sum of Kronecker product (SKP) method [38] and generalized likelihood ratio test (GLRT) method [15], is carried out.

Typically, TomoSAR methods perform the identification and separation of scattering mechanisms, by relying on the estimation of the target correlation properties measured on the MPMB SAR image stack. To this end, spatial multilook has to be carried out to estimate the covariance matrix, removing speckle noise while reducing spatial resolution. To assess the performance of the TSNNs over different levels of speckle

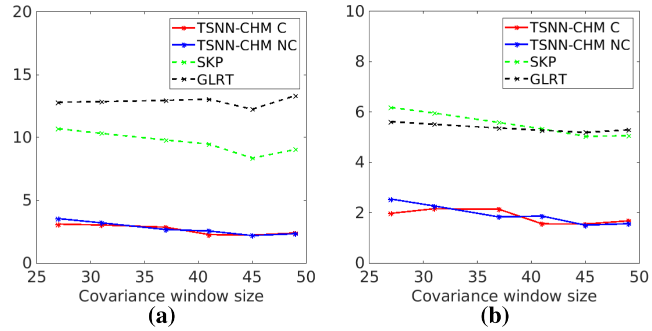


Fig. 8. Quantitative results comparison. (a) RMSE and (b) Std of the ground heights estimated by TSNN-CHMs, SKP, and GLRT methods with respect to LiDAR-based CHM, based on different covariance window sizes.

noise and the sensitivity of the network to the window size, the experiments are carried out using sample covariance matrices estimated by different window sizes. In particular, the window sizes are set to 27×27, 31×31, 37×37, 41×41, 45×45, and 49×49 pixels, to estimate covariance matrices and get the input vectors, as explained in Section II-A. Meanwhile, the mean filters with the corresponding window sizes are conducted for LiDAR-based data to match the estimation resolution. Then, LiDAR-based CHM and DTM are quantized with step 1 m as labels.

The TSNN-CHMs, calibrated (TSNN-CHM-C) and noncalibrated (TSNN-CHM-NC), as well as SKP and GLRT methods, are implemented on testing patch1 shown in Fig. 5 (first row). In Fig. 6, the predicted forest heights by the employed methods with respect to the window sizes of covariance matrices over the three selected test transect lines (transect line1, transect line2, and transect line3) shown in Fig. 5 are presented. Along the azimuth lines, the vegetation layer is very dense, and multiple interactions from the ground and canopy are expected to perceive. In our implementation, the forest TomoSAR heights are generated by subtracting the estimated ground height from the estimated canopy height by the employed methods.

From the results, it can be seen that TomoSAR forest heights are underestimated by SKP and GLRT methods with respect to the LiDAR-based CHM, which are actually due to the difference between the imaging geometry of the LiDAR and SAR system. The retrieved backscattering of contribution from the canopy surface by TomoSAR methods is situated in the middle of volumetric scattering. Although performing some vertical shifts can address such an underestimation, proper determination of the amount of shift needs additional measurements [12]. Thus, measuring the forest height from only SAR data with the best agreement to LiDAR-based CHM is a challenging task. However, with the proposed TSNN-CHMs (both calibrated and noncalibrated), a better agreement with LiDAR-based CHM is achieved. It is interesting to note that with the increase of the window size, the predicted forest height profiles by TSNN-CHMs are closer to the LiDAR-based CHM profiles: this is especially visible with transect line3.

In order to further demonstrate the performance of the proposed TSNN-CHM, the joint distributions between reconstructed forest height and LiDAR-based CHM of testing patch1 are presented in Fig. 7. Forest heights estimated by TomoSAR methods have the previously mentioned underestimation issue, while TSNN-CHMs forest heights are mostly situated on the black line, representing ideal estimation. The bigger the covariance window size is, the TSNN-CHMs forest heights focus densely on the black line meaning a better agreement with LiDAR-based CHM. The estimation of the data covariance matrix deals with speckle noise, thus improving signal-to-noise ratio (SNR) and leading to higher accuracy in height estimation. It is understandable that SNR improvements are paid at the cost of degrading the spatial resolution of reconstructed height.

It is important to highlight that in Figs. 6 and 7, both trained networks (i.e., TSNN-CHM-C and TSNN-CHM-NC) have about the same performance on forest height prediction. This is also supported in Fig. 8(a) and (b). The root-mean-square error (RMSE) and the standard deviation (Std) of the forest heights estimated by TSNN-CHM-NC have similar performance. However, the results achieved by both TSNN-CHM-C and TSNN-CHM-NC outperform SKP and GLRT. The analysis shows that the proposed TSNN-CHM gives a new way to reconstruct forest height without relying on the vertical shift compensation and phase error calibration.

Furthermore, Fig. 9 presents the reconstructed tomographic heights of testing patch2 shown in Fig. 5. From the results, SKP, GLRT, and TSNN-DTMs calibrated (TSNN-DTM-C) and noncalibrated (TSNN-DTM-NC) all have shown a good agreement with LiDAR-based DTM. This agreement can be verified from the joint distribution between the LiDAR-based DTM and retrieval of the ground heights in Fig. 10, while a better agreement between TSNN-DTMs and LiDAR-based DTM can be observed: in Fig. 10, the ground heights predicted by TSNN-DTMs rely more on the black line (representing ideal estimation). In terms of the same covariance window size, both TSNN-DTMs have better agreement with LiDAR-based DTM compared with the estimated ground height by SKP and GLRT methods. When the window size becomes bigger, TSNN-DTMs ground heights gradually shrink to the black line presenting a better estimation of ground height. In addition, TSNN-DTM-NC shows better prediction performance and fewer mistakes than TSNN-DTM-C. Fig. 11(a) and (b) shows the quantitative results comparison of the ground heights estimated by TSNN-DTM-C, TSNN-DTM-NC, SKP, and GLRT method with respect to LiDAR-based DTM. TSNN-DTM-NC better estimates terrain topography compared with others.

The presented results show the capability of the proposed TSNNs in estimating canopy and ground height within the considered forested area, using MPMB SAR data. The achieved results are comparable and even better than state-of-the-art TomoSAR methods, with particular reference to the canopy height underestimation issue. The latter, while largely affecting classical TomoSAR methods, does not influence the proposed estimation method. Moreover, the proposed method is able to provide the correct height estimation without the need for phase calibration.

In the following, the performance of TSNNs trained with noncalibrated SAR data when dealing with different phase errors will be discussed.

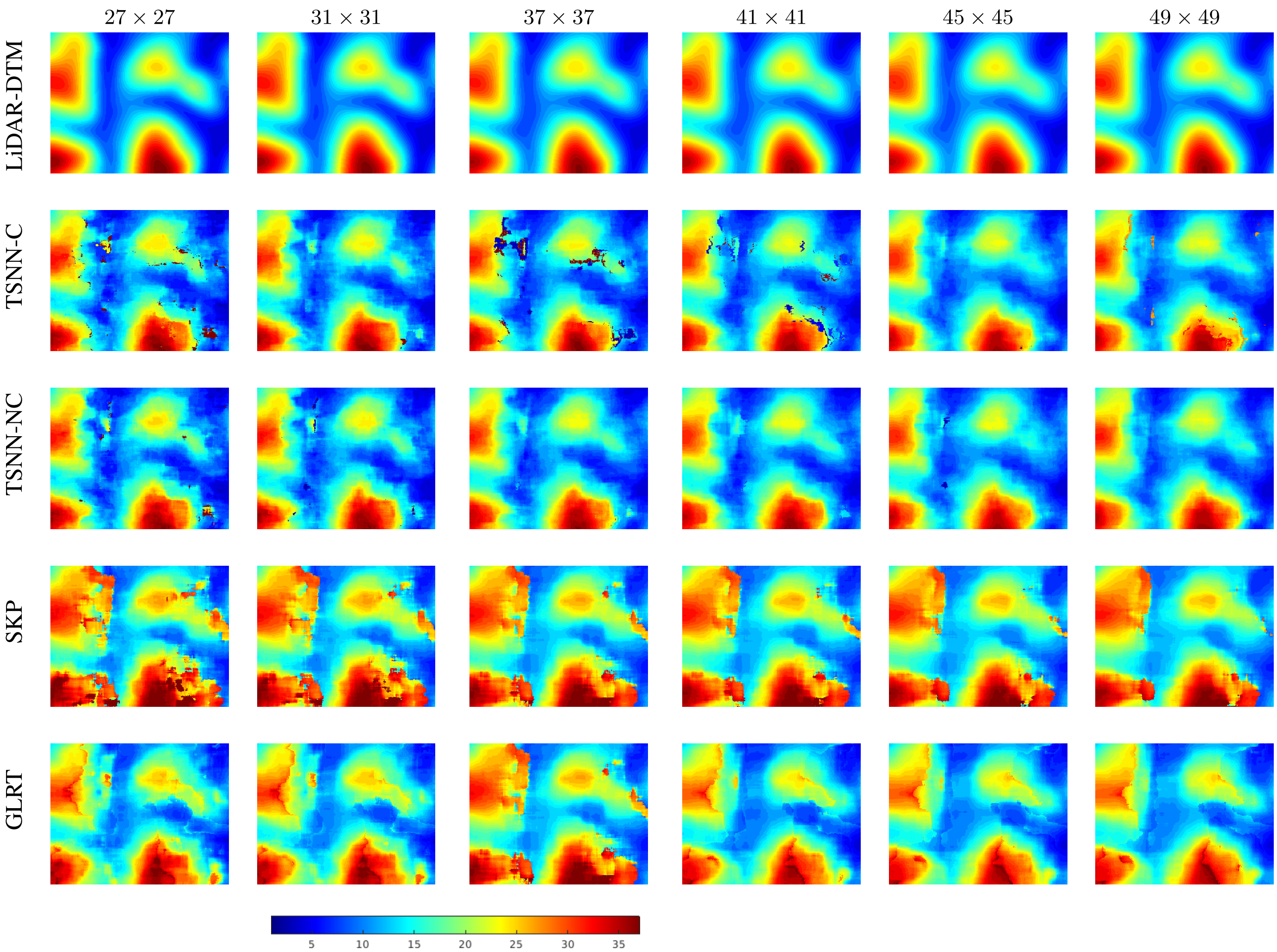


Fig. 9. Comparison within reference LiDAR-based DTM (first row) and ground height estimated by: from second to last row, TSNN-DTM-C, TSNN-DTM-NC, SKP [38], and GLRT method [15] on different covariance window sizes.

C. Robustness With Respect to Phase Errors

In Section III-B, the capability of the method in providing effective results even in the case of noncalibrated data has been shown. The TSNN-DTM-NC and TSNN-CHM-NC have shown similar or even better results compared to the calibrated TSNNs. The network is somehow able to automatically compensate for the phase offset corrupting the data or to automatically learn how the handle such features. It is worth noting that the testing areas (patch1 and patch2) included pixels different from the ones selected for the training. By the way, the testing pixels were selected nearby the pixels where the training was conducted. This helped the algorithm in taking care of the phase calibration issue while producing the solution: it is expected that the phase offsets of the testing areas were not much different from the ones of the training data learned by the network, allowing an effective solution on patch1 and patch2. To analyze TSNN-DTM-NC and TSNN-CHM-NC tolerance for phase error, the proposed method has been evaluated under controlled conditions by applying random phase errors on testing SAR data. The idea is to evaluate the robustness of the method when the input data are characterized by phase offset different from the one corrupting the training data. This is the case, for example, of a testing area far away from the training one. With such aim, available TomoSAR data (patch1 and patch2) have been corrupted using an N-dimensional phase error vector as

$$\boldsymbol{\varphi} = [\varphi_1 \quad \varphi_2 \quad \cdots \quad \varphi_N]. \tag{10}$$

The tests have been conducted using TSNN-DTM-NC and TSNN-CHM-NC with the covariance window size set to 49 × 49, under the hypothesis that the phase offset is different for the different baseline, but, fixing a baseline, they were the same for each polarization and across the testing patch.

The data were corrupted with various levels of phase error. In particular, random phase errors were simulated and multiplied to the data. The random phase error for each baseline is generated in three different scenarios: small error case when the additive phase is uniformly distributed within the interval $[-\pi/16,\ \pi/16]$, medium error case when the additive phase is uniformly distributed within the interval $[-\pi/8,\ \pi/8]$, and large error case when the additive phase is uniformly distributed within the interval $[-\pi/4,\ \pi/4]$. The plot of the generated phase offsets is shown in Fig. 12.

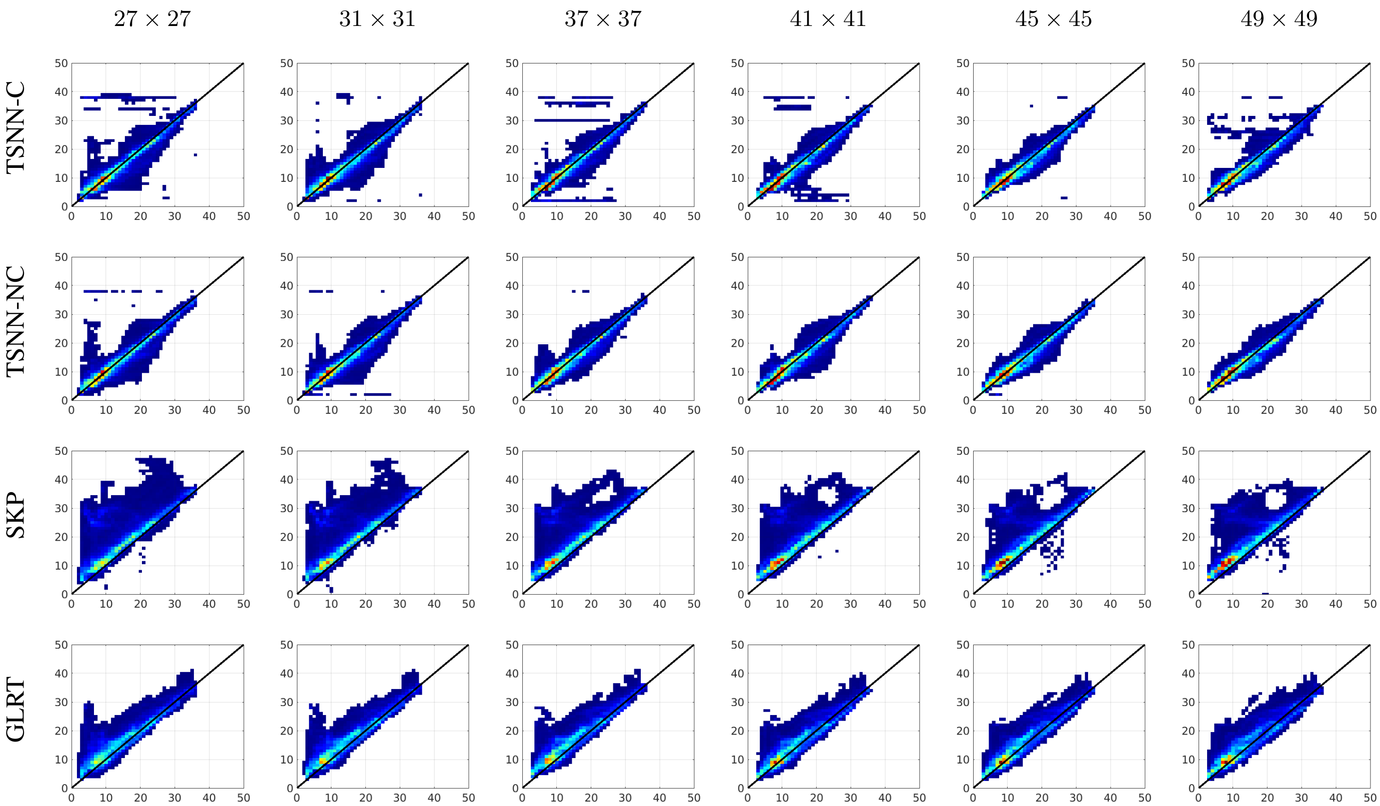


Fig. 10. Joint distribution between the LiDAR-based DTM and ground height predicted by TSNN-DTM-C (first row), TSNN-DTM-NC (second row), SKP (third row) [38], and GLRT method (last row) [15] on different covariance window sizes.

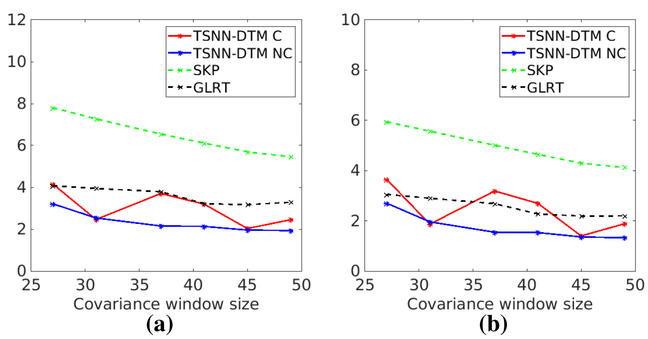


Fig. 11. Quantitative results comparison. (a) RMSE and (b) Std of the forest heights estimated by TSNN-DTMs, SKP, and GLRT methods with respect to LiDAR-based DTM based on different covariance window sizes.

The predicted forest heights by TSNN-CHM-NC are shown in Fig. 13, where the first column gives the comparison reference without phase error applying on testing patch1. Based on the phase errors ranging in $[-\pi/16,\ \pi/16]$, the profiles of forest height by TSNN-CHM match well with LiDAR-based CHM along the three transect lines. Regarding the medium error case, i.e., phase errors with range $[-\pi/8,\ \pi/8]$, the agreement between forest height by TSNN-CHM and LiDAR-based CHM is still interesting except for transect line1 where the results go a little worse. Furthermore, for the situation where phase errors are in the range $[-\pi/4,\ \pi/4]$, even though fewer agreements are shown along three transect lines, the maximum error of forest heights by TSNN-CHM with respect to LiDAR-based CHM is limited around 5 m indicating good robustness of TSNN-CHM. Moreover, it can be noted from the joint distributions between LiDAR-based CHM and forest height by TSNN-CHM-NC of three phase error vectors (shown in Fig. 14) that well concentrate on the black lines (representing the ideal estimation) showing similar performance with the situation without applying phase error: the robustness of the method is evident.

The reconstructed ground heights of patch2 corrupted by the phase errors of Fig. 12 are shown in Fig. 15, where the first column gives the comparison reference without phase error applied on testing patch2. In the case of small error, i.e., phase errors ranging in $[-\pi/16,\ \pi/16]$, TSNN-DTM-NC has a similar good prediction performance with the case of no phase error. Under the situation of medium errors, i.e., phase errors ranging in $[-\pi/8,\ \pi/8]$, the reconstructed ground heights are still generally reliable with some isolated errors. When the testing SAR image data are applied to the large error case, i.e., phase errors ranging in $[-\pi/4,\ \pi/4]$, TSNN-DTM-NC exhibits more areas of incorrect estimation. However, by inspecting the joint distribution between the reconstructed ground heights and LiDAR-based DTM reported in Fig. 16, the overall good performance of TSNN-DTM-NC is evident, with errors appearing in more extreme conditions.

To get a flavor of the accuracy assessment of the proposed method, ten realizations of the experiment are conducted for each case (small, medium, and large errors). Table II shows the mean values of RMSE and its Std, between the retrieved forest heights by TSNN-CHM-NC and LiDAR-based

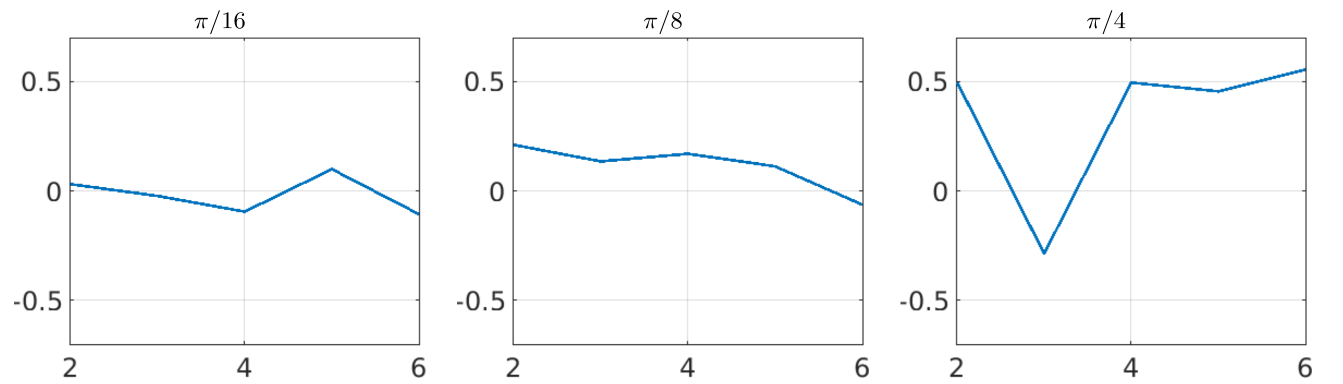


Fig. 12. Phase error vectors ranging in $[-\pi/16,\ \pi/16]$, $[-\pi/8,\ \pi/8]$ and $[-\pi/4,\ \pi/4]$.

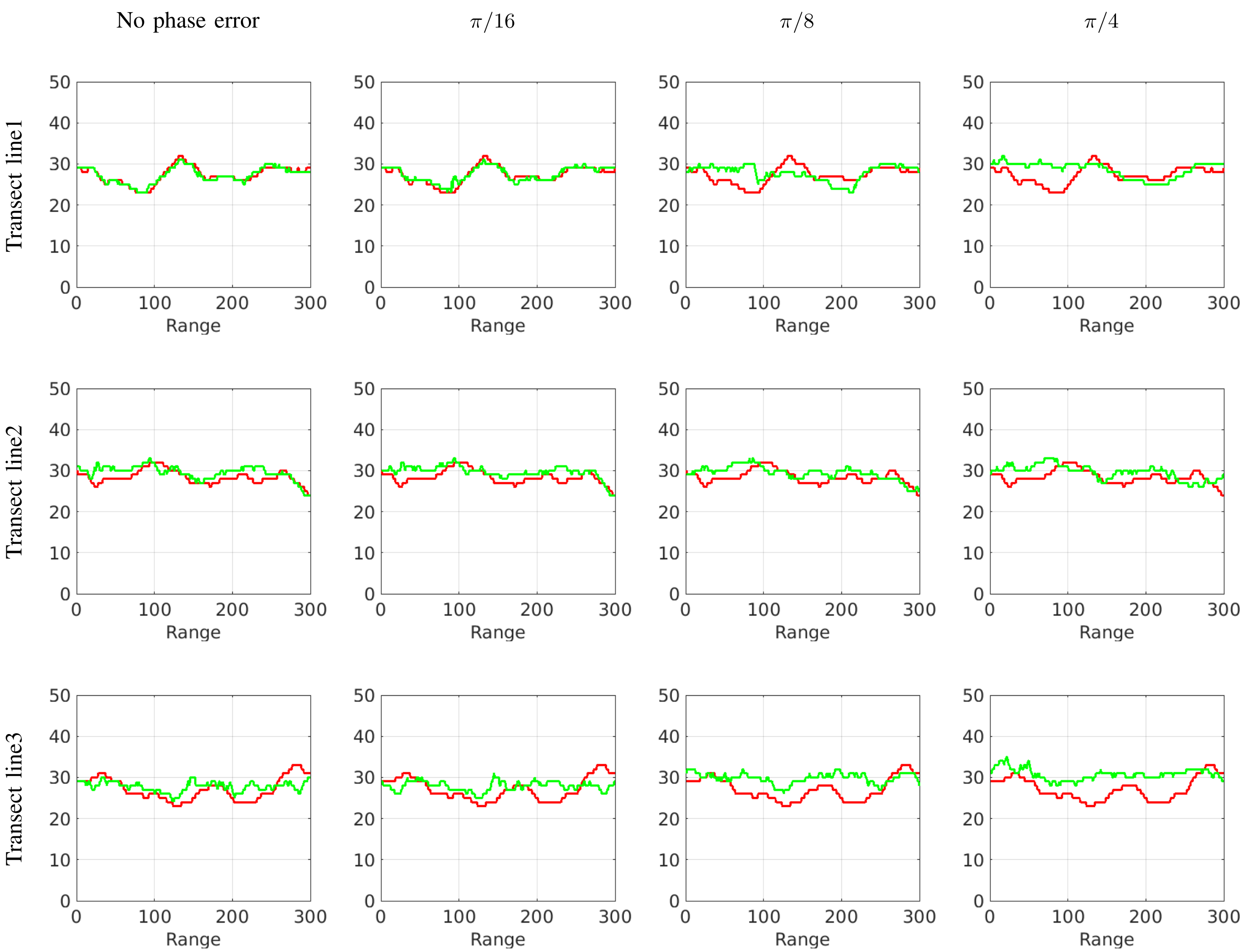


Fig. 13. Profile plots comparison between LiDAR-based CHM (red lines) and forest heights by TSNN-CHM-NC (green lines) along three transect lines.

CHM reference. Table II indicates that small phase errors in TomoSAR data do not have much influence on TSNN-CHM-NC's ability to extract forest backscattering features and reconstruct the forest height. The mean values of RMSE and Std between the reconstructed ground heights by TSNN-DTM-NC and LiDAR-based DTM reference of ten experiments are listed in Table III. Again, the results are effective. Compared with TSNN-CHM-NC, TSNN-DTM-NC has a little lower tolerance for phase errors and performs some mis-measurements on ground heights. Indeed, TSNN-DTM-NC and TSNN-CHM-NC trained with noncalibrated SAR data have good robustness to medium/small phase errors and, even in the case of large errors, the results are generally satisfactory. This aspect of the proposed method is of fundamental importance considering that the feasibility and applicability of all conventional tomographic methods largely depend on the phase calibration step.

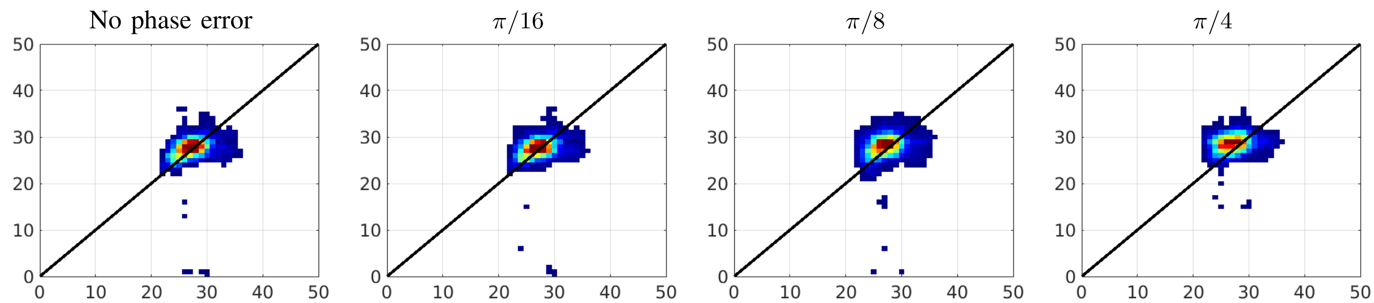


Fig. 14. Joint distribution between LiDAR-based CHM and forest height by TSNN-CHM-NC over testing patch1.

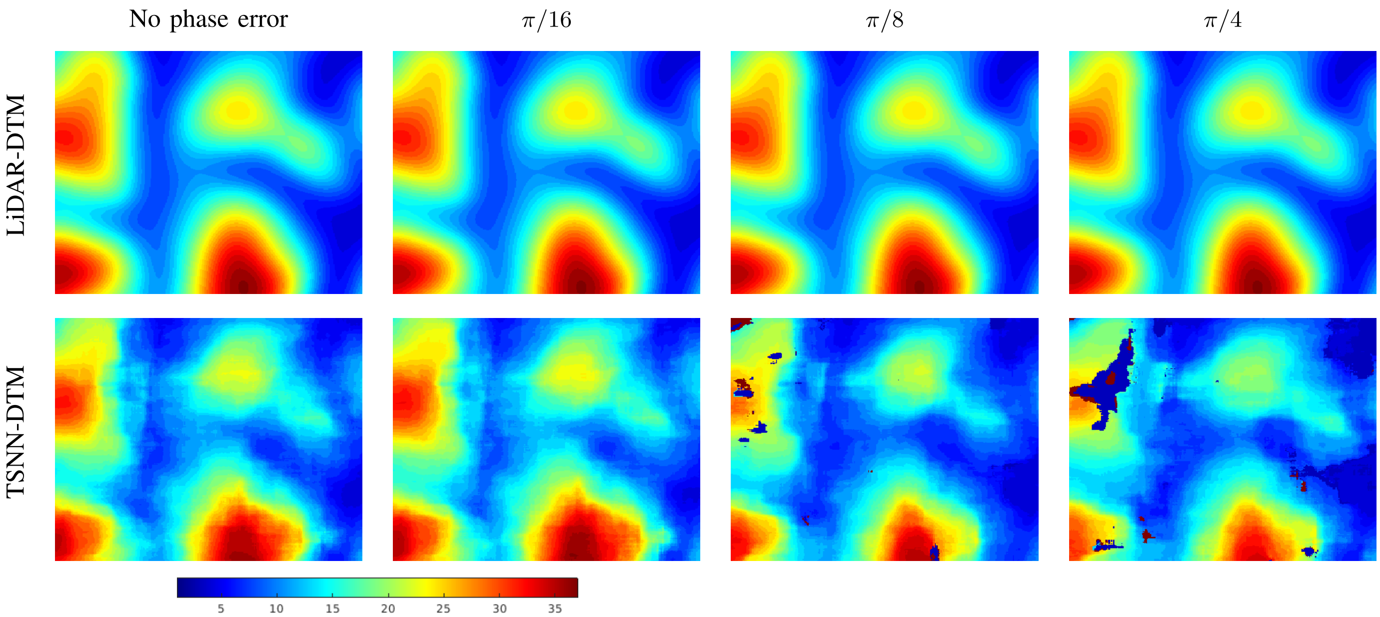


Fig. 15. Comparison within LiDAR-based DTM and ground heights by TSNN-DTM-NC.

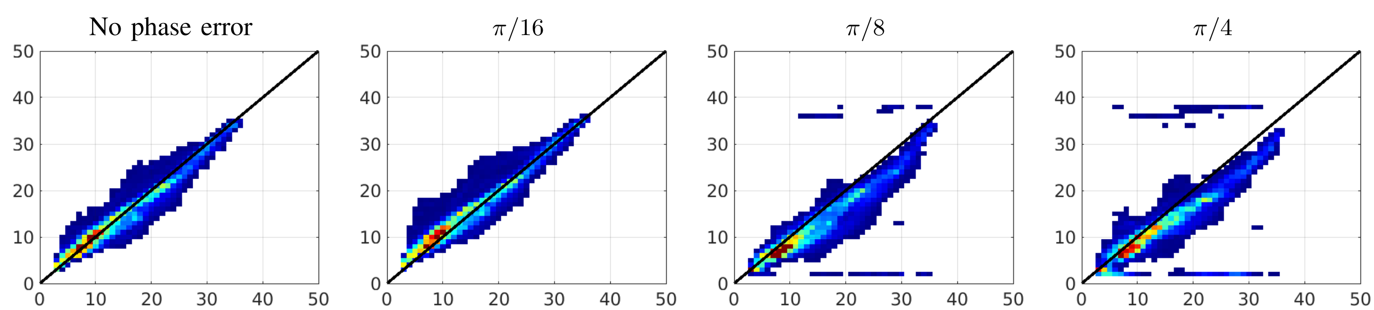


Fig. 16. Joint distribution between LiDAR-based DTM and ground height by TSNN-DTM-NC over testing patch2.

TABLE II

QUANTITATIVE RESULTS COMPARISON OF TSNN-CHM-NC WITHIN DIFFERENT PHASE ERRORS

	No phase error	$\pi/16$	$\pi/8$	$\pi/4$
RMSE	2.3328	2.4254	2.496	3.2645
Std	1.5614	1.5963	1.6106	2.3670

TABLE III

QUANTITATIVE RESULTS COMPARISON OF TSNN-DTM-NC WITHIN DIFFERENT PHASE ERRORS

	No phase error	$\pi/16$	$\pi/8$	$\pi/4$
RMSE	1.9328	2.2621	3.0425	6.3224
Std	1.3222	1.5509	2.2168	4.6844

D. Test on PolInSAR Configuration

PolInSAR has shown significant success in the estimation of canopy height and underlying ground elevation [39], [40]. In this section, we exploit the versatility, robustness, and potential of the proposed TomoSAR method (MPMB) within PolInSAR (MPMB) framework. In particular, the TSNN architecture has been retrained limiting the number of SAR images to two. Therefore, the size of the covariance matrix is 6 ×

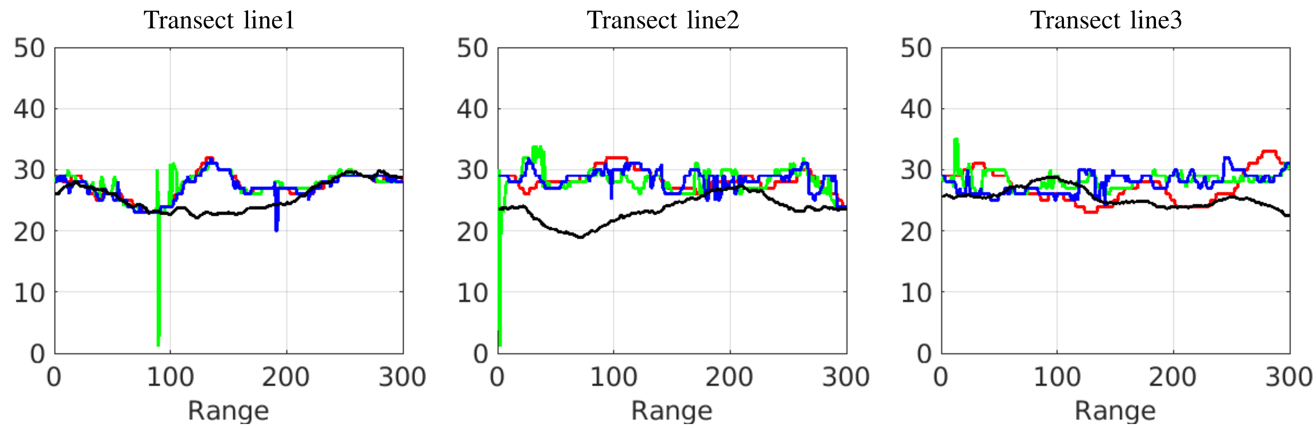


Fig. 17. Profile plots comparison between LiDAR-based CHM (red lines) and forest height estimation over three different transect lines represented in Fig. 5 using TSNN-CHM-C (blue lines), TSNN-CHM-NC (green lines), and RVoG (black lines).

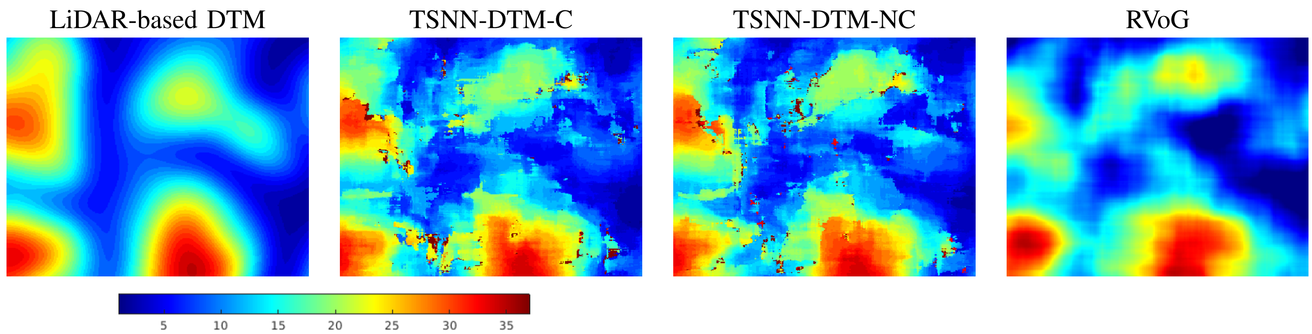


Fig. 18. Comparison within LiDAR-based DTM, ground height by TSNN-DTM-C, ground height by TSNN-DTM-NC, and ground height by RVoG.

6 and the input vector of each range–azimuth pixel is 16×1. The construction of the LiDAR reference is the same described in Section II-A. The training has been conducted both with calibrated and not calibrated data for both forest and ground height retrieval. In order to assess the performance, a comparison with the random volume over ground (RVoG) [41] PolInSAR method has been carried out on testing patch1 and patch2 (Fig. 5). In Fig. 17, the forest heights estimated by two methods over the three testing transect lines selected in Fig. 5 are illustrated, along with the comparison with the LiDAR-based CHM. It can be noted that the proposed method achieves significant agreement with LiDAR-based CHM where the height error is limited within meters. However, the forest height by RVoG is seriously underestimated with a height error of around 20 m. Furthermore, in Fig. 18, the ground heights reconstructed by two methods are shown. The results still show a good agreement of the proposed solution with the LiDAR-based DTM, while RVoG is characterized by a severe overestimation with low spatial resolution.

IV. Conclusion

In this article, fully connected neural networks (TSNNs) were proposed for forest height and underfoliage ground elevation mapping using MB polarimetric SAR images. The problem has been formulated as a classification task: TSNNs were trained by taking elements of the sample covariance matrix of MB polarimetric SAR data as input vectors and the quantized LiDAR-based CHM and DTM as reference labels. The proposed method has been trained and tested using MPMB SAR data acquired by the ONERA SETHI sensor over Paracou region of French Guiana in the frame of the European Space Agency's campaign TROPISAR and LiDAR-based data provided by the French Agricultural Research Center. Experiments based on difference covariance sizes demonstrate the proposed method's efficiency in comparison with existing classical tomographic techniques. Two main aspects need to be underlined. First, the proposed technique is able to estimate canopy and ground height with an accuracy comparable to and even better than state-of-the-art TomoSAR methods, especially with reference to the canopy height underestimation issue, typical of classical TomoSAR methods. In addition, the proposed method is able to provide the correct height estimation without the need for phase calibration, avoiding the dependence of results on such problematic processing steps. Finally, the conducted experiments showed the robustness of the proposed method for the presence of phase error in MPMB SAR data, even different from the ones characterizing the training data.

We deem that TSNN has been trained on a specific dataset: TropiSAR datasets over Paracou region. The method provides an accurate solution for reconstruction in a situation with the same parameters as TropiSAR data and similar characteristics as in the Paracou region. Changes in system parameters, such as baseline length, flight height, and incident angles, could greatly affect the performance of the TSNN. However, the results show the ability of the proposed solution in extracting mutual information from MPMB data necessary for predicting the height LiDAR values. Thus, including additional forested areas in the training dataset will increase the generalization ability of the solution.

REFERENCES

[1] S. Englhart, V. Keuck, and F. Siegert, "Aboveground biomass retrieval in tropical forests—The potential of combined X- and L-band SAR data use," *Remote Sens. Environ.*, vol. 115, no. 5, pp. 1260–1271, May 2011.

[2] H. Aghababaei et al., "Forest SAR tomography: Principles and applications," *IEEE Geosci. Remote Sens. Mag.*, vol. 8, no. 2, pp. 30–45, Jun. 2020.

[3] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.

[4] S. Tebaldini and F. Rocca, "Multibaseline polarimetric SAR tomography of a boreal forest at P- and L-bands," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 232–246, Jan. 2012.

[5] Y. Huang, J. Levy-Vehel, L. Ferro-Famil, and A. Reigber, "Three-dimensional imaging of objects concealed below a forest canopy using SAR tomography at L-band and wavelet-based sparse estimation," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 9, pp. 1454–1458, Sep. 2017.

[6] S. Tebaldini, "Single and multipolarimetric SAR tomography of forested areas: A parametric approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2375–2387, May 2010.

[7] G. Fornaro, F. Serafino, and F. Soldovieri, "Three-dimensional focusing with multipass SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 507–517, Mar. 2003.

[8] F. Lombardini and M. Pardini, "First experiments of sector interpolated SAR tomography," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2010, pp. 21–24.

[9] P. Lopez-Dekker and J. J. Mallorqui, "Capon- and APES-based SAR processing: Performance and practical considerations," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2388–2402, May 2010.

[10] F. Lombardini, F. Cai, and D. Pasculli, "Three dimensional focusing with multipass SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 960–968, Apr. 2013.

[11] A. Budillon, A. Evangelista, and G. Schirinzi, "Three-dimensional SAR focusing from multipass signals using compressive sampling," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 488–499, Jan. 2011.

[12] X. Li, L. Liang, H. Guo, and Y. Huang, "Compressive sensing for multibaseline polarimetric SAR tomography of forested areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 153–166, Jan. 2016.

[13] H. Aghababaee and M. R. Sahebi, "Model-based target scattering decomposition of polarimetric SAR tomography," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 972–983, Feb. 2018.

[14] M. Pardini and K. Papathanassiou, "On the estimation of ground and volume polarimetric covariances in forest scenarios with SAR tomography," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1860–1864, Oct. 2017.

[15] H. Aghababaee, A. Budillon, G. Ferraioli, V. Pascazio, and G. Schirinzi, "On the separation of ground and canopy scatterings using single polarimetric multi-baseline SAR tomography," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 975–978.

[16] J. O. Sexton, T. Bax, P. Siqueira, J. J. Swenson, and S. Hensley, "A comparison of LiDAR, radar, and field measurements of canopy height in pine and hardwood forests of southeastern North America," *Forest Ecol. Manage.*, vol. 257, no. 3, pp. 1136–1147, Feb. 2009.

[17] M. Pardini, M. Tello, V. Cazcarra-Bes, K. P. Papathanassiou, and I. Hajnsek, "L- and P-band 3-D SAR reflectivity profiles versus LiDAR waveforms: The AfriSAR case," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3386–3401, Oct. 2018.

[18] J. Anderson et al., "The use of waveform LiDAR to measure northern temperate mixed conifer and deciduous forest structure in new hampshire," *Remote Sens. Environ.*, vol. 105, no. 3, pp. 248–261, Dec. 2006.

[19] C. A. Silva et al., "Comparison of small- and large-footprint LiDAR characterization of tropical forest aboveground structure and biomass: A case study from central Gabon," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3512–3526, Oct. 2018.

[20] Q. Zhang, L. Ge, S. Hensley, G. I. Metternicht, C. Liu, and R. Zhang, "PolGAN: A deep-learning-based unsupervised forest height estimation based on the synergy of PolInSAR and LiDAR data," *ISPRS J. Photogramm. Remote Sens.*, vol. 186, pp. 2716–2924, Jan. 2022.

[21] M. Pourshamsi, M. Garcia, M. Lavalle, and H. Balzter, "A machine-learning approach to PolInSAR and LiDAR data fusion for improved tropical forest canopy height estimation using NASA AfriSAR campaign data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3453–3463, Oct. 2018.

[22] J. Yang, X. Fu, Y. Hu, Y. Huang, X. Ding, and J. Paisley, "PanNet: A deep network architecture for pan-sharpening," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 5449–5457.

[23] W. Qi et al., "Improved forest height estimation by fusion of simulated GEDI LiDAR data and TanDEM-X InSAR data," *Remote Sens. Environ.*, vol. 221, pp. 621–634, Feb. 2019.

[24] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[25] X. X. Zhu et al., "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.

[26] X. X. Zhu et al., "Deep learning meets SAR: Concepts, models, pitfalls, and perspectives," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 4, pp. 143–172, Dec. 2021.

[27] A. Budillon, A. C. Johnsy, G. Schirinzi, and S. Vitale, "SAR tomography based on deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 3625–3628.

[28] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, 1958.

[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1106–1114.

[30] K. P. Murphy, *Machine Learning: A Probabilistic perspective*. Cambridge, MA, USA: MIT Press, 2012.

[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent.*, 2015, pp. 7–9.

[32] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *J. Mach. Learn. Res.*, vol. 9, pp. 249–256, May 2010.

[33] B. Liu and G. Tsoumakas, "Dealing with class imbalance in classifier chains via random undersampling," *Knowl.-Based Syst.*, vol. 192, Mar. 2020, Art. no. 105292.

[34] I. Cordón, S. García, A. Fernández, and F. Herrera, "Imbalance: Oversampling algorithms for imbalanced classification in R," *Knowl.-Based Syst.*, vol. 161, pp. 329–341, Dec. 2018.

[35] H. Aghababaee, G. Fornaro, and G. Schirinzi, "Phase calibration based on phase derivative constrained optimization in multibaseline SAR tomography," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6779–6791, Nov. 2018.

[36] Z. Liao, B. He, X. Bai, and X. Quan, "Improving forest height retrieval by reducing the ambiguity of volume-only coherence using multi-baseline PolInSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 8853–8866, Nov. 2019.

[37] G. Vincent et al., "Accuracy of small footprint airborne LiDAR in its predictions of tropical moist forest stand structure," *Remote Sens. Environ.*, vol. 125, pp. 23–33, Oct. 2012.

[38] S. Tebaldini, "Algebraic synthesis of forest scenarios from multibaseline PolInSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 12, pp. 4132–4142, Dec. 2009.

[39] R. N. Treuhaft and P. R. Siqueira, "Vertical structure of vegetated land surfaces from interferometric and polarimetric radar," *Radio Sci.*, vol. 35, no. 1, pp. 141–177, Jan. 2000.

[40] S. R. Cloude and K. P. Papathanassiou, "Three-stage inversion process for polarimetric SAR interferometry," *IEE Proc.-Radar, Sonar Navigat.*, vol. 150, no. 3, pp. 125–134, Jun. 2003.

[41] K. P. Papathanassiou and S. R. Cloude, "Single-baseline polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2352–2363, Nov. 2001.

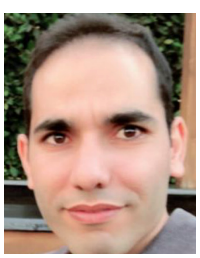
Wenyu Yang (Graduate Student Member, IEEE) was born in Cangzhou, China. She received the B.S. and M.S. degrees in information and communication science from the Beijing Institute of Technology, Beijing, China, in 2017 and 2020, respectively. She is currently pursuing the Ph.D. degree in information and communication technology and engineering with the Department of Engineering, University of Naples Parthenope, Naples, Italy.

Her research topics focus on applying deep learning methods to polarimetry synthetic aperture radar (SAR) and SAR tomography images for canopy and terrain topography reconstruction.

Sergio Vitale (Member, IEEE) received the master's degree (summa cum laude) in telecommunication engineering from the University of Naples Federico II, Naples, Italy, in May 2017, and the Ph.D. degree in information and communication technology and engineering from the Department of Engineering, University of Naples Parthenope, Naples, in 2021.

He is currently a Post-Doctoral Researcher at the Department of Engineering, University of Naples Parthenope. His research activities focus on the definition and analysis of deep learning solutions for image processing applied to remote sensing. He mainly focuses on the study and definition of convolutional neural networks (CNNs) for speckle denoising of synthetic aperture radar (SAR) image, SAR tomography, and pansharpening of image from optical sensors.

Dr. Vitale won the 2017 Best Italian Remote Sensing Thesis Prize for the IEEE Geoscience Remote Sensing South Italy Chapter in 2018. Since 2021, he has been a member of the Topic Advisor Panel of *Sensors* (MDPI) and a Guest Editor for the Special Issue of *Sensors* (MDPI) and the Special Issue of *Remote Sensing* (MDPI).



Hossein Aghababaei (Senior Member, IEEE) is currently an Assistant Professor at the Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands. He enthusiastically shares his knowledge of concepts, theories, and methods of radar remote sensing and machine learning in several courses at the University of Twente and is actively involved in capacity-building projects. ESA AlignSAR is the latest project he is involved in. His research interests are polarimetric and multichannel synthetic aperture radar (SAR) image processing, radar interferometry, and tomography.

Mr. Aghababaei is a member of several scientific committees, including the Nederlands Aardwetenschappelijk Congress, and an Associate Editor of IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Giampaolo Ferraioli (Senior Member, IEEE) was born in Lagonegro, Italy, in 1982. He received the B.S. and M.S. degrees and the Ph.D. degree in telecommunication engineering from the Università degli Studi di Napoli Parthenope, Naples, Italy, in 2003, 2005, and 2008, respectively.

He has been a Visiting Scientist at the Département TSI, Télécom ParisTech, Paris, France. He is currently an Associate Professor with the Università degli Studi di Napoli Parthenope. His main research interests deal with statistical signal and image processing, radar systems, synthetic aperture radar, image restoration, and magnetic resonance imaging.

Dr. Ferraioli is a member of the Technical Liaison Committee of IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING. He serves on the Editorial Board of *Remote Sensing* (MDPI). In 2009, he won the IEEE 2009 Best European Ph.D. Thesis in Remote Sensing Prize, sponsored by the IEEE Geoscience and Remote Sensing Society. He serves as an Associate Editor for IEEE GEOSCIENCE AND REMOTE SENSING LETTERS and IEEE JOURNAL ON MINIATURIZATION FOR AIR AND SPACE SYSTEMS.

Vito Pascazio (Senior Member, IEEE) graduated (summa cum laude) in electronic engineering from the Università di Bari, Bari, Italy, in 1986. He received the Ph.D. degree in electronic engineering and computer science from the Department of Electronic Engineering, Università di Napoli Federico II, Naples, Italy, in 1990.

In 1990, he was at the Research Institute on Electromagnetics and Electronic Devices (IRECE), Italian National Council of Research (CNR), Naples. He is currently a Full Professor and the Chair of the Department of Engineering, Università di Napoli Parthenope, Naples. He was a Visiting Scientist at the Laboratoire des Signaux et Systèmes, École Supérieure d'électricité (Supelec), Gif-sur-Yvette, France, from 1994 to 1995, and the Universite de Nice Sophia-Antipolis, Nice, France, from 1998 to 1999. He was also the Director of the National Laboratory of Multi-Media Communications, Italian Consortium of Telecommunications (CNIT), Naples. He has published more than 180 technical articles. His research interests include remote sensing, image processing, and linear and nonlinear estimation, with particular emphasis on image computing and processing, and reconstruction of microwave and radar images.

Dr. Pascazio was awarded the Philip Morris Prize for Scientific and Technological Research in 1990. He was the General Co-Chairman of IGARSS 2015 conference.

Gilda Schirinzi (Senior Member, IEEE) received the M.Sc. degree in electronic engineering from the University of Naples Federico II, Naples, Italy, in 1983.

She joined the Electronic Engineering Department, University of Naples Federico II, as a Research Fellow. From 1985 to 1986, she was with the European Space Agency, ESTEC, Noordwijk, The Netherlands. In 1988, she joined the Istituto di Ricerca per l'Elettromagnetismo e i Componenti Elettronici (IRECE), Italian National Council of Researches, Naples. In 1992, she was appointed as the Head of the Electromagnetics Division, IRECE, where she became a Senior Researcher in 1997. In 1998, she joined the University of Cassino, Cassino, Italy, as an Associate Professor of telecommunications, where she became a Full Professor in 2005. Since 2008, she has been with the Telecommunication Group, University of Naples "Parthenope," Naples. She has taught signal theory, electrical communications, microwave remote sensing systems, and image processing. Her research interests include synthetic aperture radar (SAR) signal processing and coding, SAR interferometry and tomography, microwave imaging techniques, and image and signal processing for remote sensing applications