Global Coverage of Sentinel-1 and Spaceborne LiDAR: A Data-Driven Foundation for Forest Height Estimation

Hossein Aghababaei[1], Giampaolo Ferraioli[2], Jaan Praks[3], Erkki Tomppo [4], Gilda Schirinzi,[5]

[1] University of Twente, Hallenweg 8, 7522 NH Enschede, Netherlands - h.aghababaei@utwente.nl
[2] Universita degli Studi di Napoli "Parthenope," 80143 Naples, Italy - giampaolo.ferraioli@uniparthenope.it
[3] Aalto University, Maarintie 8, 02150, Espoo, Finland - jaan.praks@aalto.fi
[4] University of Helsinki, Latokartanonkaari 7, P.O. Box 27, Helsinki FI-00014, Finland - erkki.tomppo@helsinki.fi
[5] Universita degli Studi di Napoli "Parthenope," 80143 Naples, Italy - gilda.schirinzi@uniparthenope.it

Keywords: Sentinel-1, Lidar, Polarimetric-Interferometric data, Data-driven models, Forest Height.

Abstract

While polarimetric interferometric SAR techniques provide a strong theoretical framework for forest height retrieval, their application using C-band Sentinel-1 data is challenging due to repeat-pass acquisition geometry and strong temporal decorrelation. In this study, we assemble a globally distributed dataset combining Sentinel-1 interferometric observations with spaceborne LiDAR forest height measurements from the GEDI and ICESat-2 missions. More than 1800 Sentinel-1 interferometric image pairs were processed and spatially matched with LiDAR observations across tropical, temperate, and boreal forest regions. Sentinel-1 Single Look Complex data were used to derive interferometric coherence and polarimetric–interferometric observables, enabling statistical analysis of their relationship with forest structural properties. The results reveal physical relationships between Sentinel-1 coherence and canopy height across multiple forest biomes, indicating that Sentinel-1 interferometric measurements, under near-zero spatial baseline conditions, retain measurable sensitivity to vegetation structure despite temporal decorrelation effects. These findings provide a conceptual basis for exploiting similar repeat-pass interferometric observations from new low-frequency SAR missions such as NISAR and upcoming ROSE-L for forest height mapping. In addition, the assembled dataset provides a global benchmark for developing and evaluating data-driven approaches for forest height estimation using Sentinel-1 observations.

1. Introduction

Forests are critical components of the Earth's biosphere, playing a central role in regulating the global carbon cycle, moderating climate, and sustaining biodiversity. Among the structural attributes of forests, *canopy or forest height* is particularly informative, as it encapsulates vertical forest structure and serves as a proxy for aboveground biomass, productivity, and disturbance regimes. Accurate and spatially consistent forest height information underpins global efforts in carbon accounting, ecosystem monitoring, and sustainable land management (Le Toan et al., 2011).

Remote sensing techniques based upon Synthetic Aperture Radar (SAR), particularly Polarimetric Interferometric SAR (PolInSAR), have demonstrated exceptional capability for estimating forest height and vertical structure across large areas (Cloude and Papathanassiou, 1998). By exploiting both the polarization diversity and interferometric information between SAR acquisitions, PolInSAR enables the separation of scattering contributions from the canopy and the ground surface. Physical models such as the Random-Volume-over-Ground (RVoG) framework (Cloude and Papathanassiou, 2003) provide a theoretical link between interferometric coherence and forest structural parameters, modeling the observed complex coherence as a combination of a temporally coherent ground return and a volume-scattered canopy component. Under ideal conditions, particularly for single-pass acquisitions where temporal decorrelation is negligible, RVoG-based inversion allows reliable retrieval of forest height. Successful applications have been demonstrated using single-pass missions such as DLR's TanDEM-X (Olesk et al., 2016) and airborne SAR systems with very short repeat intervals (Aghababaei et al., 2020), where the assumption of coherence preservation holds. In contrast, the Copernicus Sentinel-1 mission operates in C-band with a repeat-pass geometry, providing open-access data and full global coverage, but at the cost of considerable temporal decorrelation. Canopy motion, vegetation growth, and surface moisture dynamics between acquisitions strongly degrade coherence, violating the assumptions of RVoG and limiting the applicability of physically based PolInSAR inversion techniques.

The presence of temporal decorrelation fundamentally challenges the assumptions underlying physically based PolInSAR models, whose validity depends on the preservation of interferometric coherence between acquisitions. To account for these effects, the interferometric coherence can be described using an extended RVoG formulation (Cloude and Papathanassiou, 2003):

$$\tilde{\gamma} = e^{j\phi_g} \left(\gamma_v \gamma_t + \frac{\mu}{1+\mu} \left(1 - \gamma_v \gamma_t\right) \right) \tag{1}$$

In this expression, the measured complex coherence $\tilde{\gamma}$ is modeled as a weighted combination of the volume coherence γ_v, which depends on forest vertical structure, and the ground contribution, represented by the ground to volume reflectivity ratio μ. The term γ_t accounts for temporal decorrelation, describing the loss of coherence due to temporal changes between acquisitions, while $e^{j\phi_g}$ introduces the ground-phase term. This formulation highlights how the temporal decorrelation factor γ_t acts multiplicatively on the volume component, effectively reducing the measure coherence in repeat-pass systems such as Sentinel-1.

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-467-2026 | © Author(s) 2026. CC BY 4.0 License.

Despite this unavoidable effect in multi-baseline SAR data acquired through repeat-pass geometry, recent studies (Lavalle et al., 2023) have shown that Sentinel-1 coherence can retain measurable sensitivity to forest biophysical properties. In particular, (Lavalle et al., 2023) proposed a model-based framework that exploits temporal time series of interferometric coherence and backscatter to retrieve forest structural parameters using an extended Random-Motion-over-Ground (RMoG) model. Unlike the formulation in Eq. (1), where temporal decorrelation is introduced as a multiplicative factor applied to the volume coherence, the RMoG model explicitly represents temporal decorrelation as a function of acquisition time separation and scatterer motion statistics. In this framework, interferometric coherence is primarily governed by temporal decorrelation effects, while geometric decorrelation can be considered negligible under near-zero baseline conditions typical of Sentinel-1 acquisitions. The RMoG model therefore expresses the interferometric coherence as a combination of ground and vegetation contributions, parameterized by the ground-to-volume scattering ratio μ, vegetation height h, extinction coefficient κ, and motion-induced decorrelation terms associated with canopy and ground dynamics. Overall, the inversion of the RMoG model is an ill-posed problem, as the number of unknown parameters exceeds the number of available observations. A key aspect of the approach in (Lavalle et al., 2023) is the use of coherence time series acquired at multiple temporal baselines, which helps stabilize the estimation by exploiting the temporal behavior of coherence while assuming that certain model parameters remain constant across acquisitions. This, in particular, enables the separation of slowly varying parameters (e.g., the extinction coefficient κ and motion-related terms) from spatially varying parameters such as forest height. To further constrain the inversion, the method incorporates external information in the form of sparse LiDAR measurements. The overall parameter retrieval is performed through a multi-stage inversion strategy. The study demonstrates that, despite strong temporal decorrelation at C-band, the temporal evolution of coherence contains sufficient information to support model-based retrieval of forest height and related biophysical variables.

Unlike the approach in (Lavalle et al., 2023), which relies on single-polarization observations and external constraints such as sparse LiDAR measurements to regularize the inversion, the present study investigates the intrinsic information content of Sentinel-1 dual-polarimetric interferometric data for forest structure estimation. In particular, we aim to exploit the full polarimetric–interferometric observables available from Sentinel-1 (VV and VH channels) to assess their capability to constrain forest structural parameters within the RMoG framework. Under repeat-pass C-band acquisition conditions, temporal decorrelation significantly degrades interferometric coherence, limiting the direct applicability of physically based inversion models in forested environments. As a result, the retrieval of parameters such as forest height, extinction coefficient, and scattering contributions becomes inherently ill-conditioned. In this study, we address this limitation by exploiting the full polarimetric capacity of Sentinel-1 data, i.e., by jointly using both VV and VH channels to better constrain the inversion problem.

The global availability, high revisit frequency, and open-access nature of Sentinel-1 data provide a unique opportunity to systematically investigate whether meaningful structural information can still be extracted under these constraints, without relying on external sources of information. The objective of this study is therefore not to claim full physical retrievability of forest parameters from Sentinel-1 data, but rather to quantify the extent to which the RMoG model, when driven by dual-polarimetric interferometric observables, can provide stable and physically interpretable estimates of forest height. This analysis serves as a foundation for understanding the limitations and potential of C-band repeat-pass interferometry, and for guiding the development of hybrid approaches that combine physical modeling with data-driven refinement. To this end, the present study assembles and analyzes a globally distributed Sentinel-1 dataset, aiming to support the development of data-driven models while systematically investigating the extent to which C-band repeat-pass observations can reliably be used within the RMoG framework.

The remainder of this paper is organized as follows. Section 2 describes the dataset construction and preprocessing steps, together with the adopted methodology. Section 3 presents the experimental results and analysis, highlighting the performance of the proposed approach across different forest conditions. Finally, Section 4 concludes the paper and discusses potential directions for future work.

2. Methodology

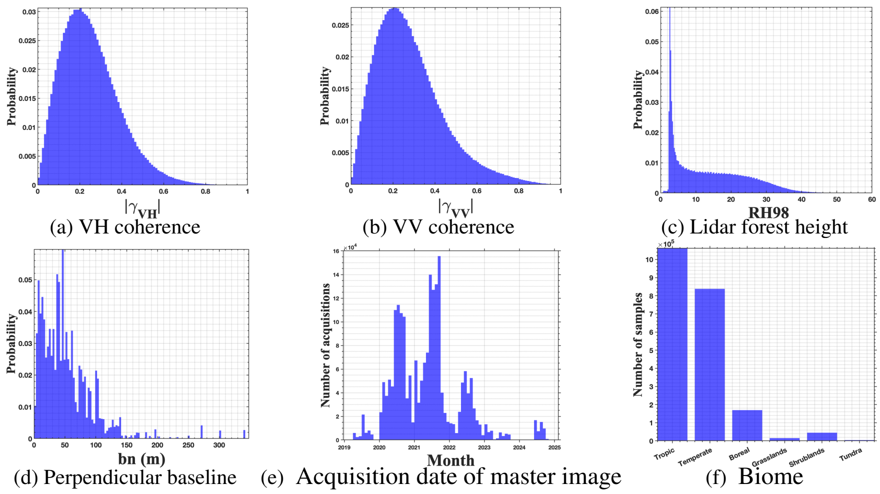
This section is organized into three subsections. The first subsection describes the globally collected Sentinel-1 and LiDAR dataset, including its specifications, preprocessing steps, and statistical characteristics. Subsection 2.2 introduces the methodological framework for implementing the RMoG model under dual-polarimetric Sentinel-1 repeat-pass acquisition conditions, with emphasis on the derivation of polarimetric interferometric observables and the associated inversion strategy. Finally, Subsection 2.3 presents the data-driven approach, which is designed to complement the physical model by learning residual relationships between the Sentinel-1 observables and reference forest height measurements.

2.1 DATA

This study is based on a globally distributed dataset constructed by combining Sentinel-1 interferometric SAR observations with spaceborne LiDAR forest height measurements. The integration of these two complementary data sources enables the analysis of the relationship between radar interferometric observables and forest structural parameters at a global scale. Sentinel1 provides consistent, repeat-pass C-band observations with global coverage, while LiDAR measurements from GEDI and ICESat-2 offer accurate canopy height references. By coregistering these datasets in space and time, a large-scale benchmark dataset is established for both physical modeling and data-driven analysis of forest height.

2.1.1 SAR data Sentinel-1 Single Look Complex (SLC) data are used to derive interferometric coherence and polarimetric interferometric observables. For each interferometric pair, standard preprocessing steps including co-registration and phase flattening are applied to preserve the complex-valued information required for interferometric analysis. The use of SLC data enables access to both amplitude and phase information, which is essential for interferometric analysis. For each pixel, the dual-polarization interferometric measurements are represented using the 4×4 polarimetric–interferometric covariance matrix $\mathbf{C}_4$,

$$\mathbf{C}_4 = \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \end{bmatrix} \begin{bmatrix} \mathbf{k}_1^T & \mathbf{k}_2^T \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{\Omega}_{12}^T \\ \mathbf{\Omega}_{12} & \mathbf{C}_{22} \end{bmatrix} \tag{2}$$



(a) VH coherence (b) VV coherence (c) Lidar forest height

(d) Perpendicular baseline (e) Acquisition date of master image (f) Biome

Figure 1. Statistical summary of the SAR–LiDAR benchmark dataset.

where $\mathbf{k}_i$ is a 2×1 complex vector representing the SAR measurements in VV and VH polarizations for the i-th acquisition. This formulation jointly captures polarimetric and interferometric information, providing a compact and physically consistent representation of the scattering mechanisms. The covariance matrix serves as the basis for deriving coherence measures and polarimetric–interferometric features used in both the RMoG-based modeling and the subsequent data-driven analysis.

2.1.2 LiDAR data Reference forest height measurements are derived from the GEDI Level-2A and ICESat-2 ATL08 spaceborne LiDAR products. These datasets provide globally distributed canopy height estimates derived from waveform LiDAR observations. The RH98 metric is used as the reference canopy height, as it provides a robust estimate of the upper canopy structure. LiDAR observations are temporally matched with Sentinel-1 acquisitions to minimize temporal inconsistencies. In addition, quality filtering based on mission-provided flags is applied to retain only high-confidence measurements, ensuring the reliability of the reference data used for model development and evaluation.

2.1.3 Data Analysis The co-registered SAR scenes are partitioned into patches of 245×256 pixels. Each patch contains four complex-valued SAR channels corresponding to the dual-polarization interferometric observations, along with auxiliary layers describing acquisition geometry, including incidence angle, slant range, and vertical wavenumber k_z. Each patch is associated with LiDAR-derived canopy height (RH98), which serves as the reference variable. To characterize the statistical properties of the dataset, the distributions of interferometric coherence in VV and VH polarizations, vertical wavenumber k_z, and Sentinel-1 acquisition dates are analyzed. These variables describe the range of interferometric decorrelation conditions and acquisition geometries present in the dataset. Figure 2 summarizes these statistical characteristics. The distributions are computed using the central pixel of each patch to reduce spatial correlation effects. The results demonstrate that the dataset spans a wide range of forest structural conditions across tropical, temperate, and boreal regions, as well as diverse acquisition geometries, making it suitable for global-scale analysis of the relationship between Sentinel-1 interferometric observables and forest height.

Detailed descriptions of the dataset construction, preprocessing workflow, and patch extraction strategy are available through [1] and the associated public data repository [2].

2.2 Zero-Baseline RMoG Model

Under the repeat-pass acquisition geometry of Sentinel-1, the perpendicular baseline is typically small (see Fig. 2), and geometric decorrelation can therefore be considered negligible. In this regime, interferometric coherence is primarily governed by temporal decorrelation effects associated with vegetation motion and variations in canopy and ground scattering. This behavior can be described using the Random Motion over Ground model (Lavalle et al., 2012), which provides a physically interpretable framework linking coherence to forest structural parameters.

Assuming negligible ground motion and a constant ground-to-volume scattering ratio between master and slave acquisitions, the interferometric coherence magnitude for polarization channel w can be approximated as:

$$|\gamma_w| = \frac{\mu_w + \gamma_v}{\mu_w + 1}, \tag{3}$$

where μ_w denotes the ground-to-volume scattering ratio and γ_v represents the vegetation volume coherence, which is directly related to canopy structure through parameters such as forest height h and extinction coefficient κ.

In the case of dual-polarimetric Sentinel-1 observations (e.g. VV and VH), the number of independent observables remains limited relative to the number of unknown model parameters, rendering the inversion problem inherently ill-posed. However, the availability of two polarization channels enables the construction of a polarimetric coherence locus, from which optimized interferometric coherences can be derived following (Cloude and Papathanassiou, 1998). In particular, two extreme coherence values, denoted as γ_{max} and γ_{min}, are obtained by projecting the polarimetric covariance matrix onto optimal scattering mechanisms. The optimized coherence γ_{max} corresponds to a configuration in which the ground contribution is minimized, and can therefore be approximated as $|\gamma_{max}| \approx \gamma_v$. This

[1] https://www.sentinel2height.nl/Downloads/benchmark-database
[2] `https://ftp.itc.nl/pub/sentinel/`

provides a direct observable primarily sensitive to vegetation volume decorrelation and, consequently, to forest structural properties. In contrast, γ_{min} is associated with stronger ground contributions and provides complementary information about the scattering mixture. Together, these observables partially compensate for the limited information content of single-polarization interferometric measurements. Based on this formulation, a numerical inversion strategy is adopted to estimate the key RMoG parameters controlling vegetation volume decorrelation. In particular, the extinction coefficient κ and forest height h are treated as primary unknowns. Analysis of the global SAR–LiDAR dataset indicates that the motion-induced decorrelation term exhibits a consistent relationship with the optimized coherence measurements, allowing it to be approximated or implicitly accounted for. This reduces the dimensionality of the inversion problem and enables a stable estimation of κ and h using the available dual-polarimetric observables.

This framework directly addresses the central research question of this study, namely, to what extent forest structural parameters can be inferred from Sentinel-1 repeat-pass interferometric data without relying on external constraints. By explicitly exploiting the dual-polarimetric information content, the approach seeks to maximize the physically meaningful information retrievable from C-band coherence measurements under strong temporal decorrelation conditions. Despite these improvements, the resulting forest height estimates may remain limited in accuracy due to the intrinsic information loss caused by temporal decorrelation and the simplified assumptions of the model. While the RMoG-based inversion is able to reproduce the dominant spatial patterns of forest height, fine-scale structural variations are not fully captured. To address these limitations, a data-driven refinement strategy is introduced in the following section.

2.3 Data-Driven Method

To address the limitations of the physical model described previously, a data-driven refinement strategy is adopted to improve the accuracy of forest height estimation. Rather than replacing the physically based RMoG framework with a purely data-driven predictor, a hybrid residual learning approach is employed. In this setting, the deep learning model is trained to estimate the systematic discrepancy between the forest height derived from the physical model and the reference height obtained from LiDAR observations. This formulation leverages the complementary strengths of both approaches: the physical model provides a structurally consistent but coarse estimate of forest height, while the data-driven component learns to correct residual errors arising from model simplifications, temporal decorrelation effects, and unmodeled scattering mechanisms. By focusing on residual estimation instead of direct prediction, the method preserves the physical interpretability of the solution while improving its accuracy.

The model is trained using a globally distributed dataset consisting of more than two million co-registered Sentinel-1 and spaceborne LiDAR samples. Each training sample is represented by a feature vector that includes the initial forest height estimate obtained from the RMoG inversion, interferometric coherence measurements, selected polarimetric–interferometric observables derived from the covariance matrix, and acquisition geometry parameters such as incidence angle and vertical wavenumber. This feature design ensures that both physical model outputs and raw SAR observables are jointly exploited during training.

The correction model is implemented using a fully connected residual neural network tailored for tabular geospatial data. Each input sample is represented by a 17-dimensional feature vector, which is first projected into a higher-dimensional latent space to enhance representational capacity. The network then processes the features through a sequence of residual blocks, enabling the modeling of complex nonlinear relationships between SAR observables and forest height discrepancies while maintaining stable training behavior. To account for variability in forest structure, acquisition conditions, and biome-dependent scattering characteristics, a mixture-of-experts architecture is incorporated. In this framework, multiple expert subnetworks independently estimate residual corrections, while a gating mechanism dynamically combines their outputs based on the input feature space. This allows the model to adapt to different forest regimes and acquisition scenarios without requiring explicit stratification of the data.

Following global training, a lightweight fine-tuning step is applied to adapt the model to regional conditions, further improving local prediction accuracy. The final forest height estimate is obtained by adding the predicted residual correction to the initial physical-model estimate. This hybrid formulation ensures that the resulting predictions remain physically consistent while benefiting from the flexibility and expressiveness of data-driven learning.

3. Experimental Section

To evaluate the proposed methodology, both the model-based RMoG inversion and the data-driven refinement approach are implemented over a study area located in the Kotka forest region in Finland. This site is characterized by heterogeneous boreal forest structure and provides a suitable testbed for assessing the performance of Sentinel-1-based forest height estimation under realistic conditions. For this experiment, four interferometric pairs acquired during the summer period of 2021 are used:

- Pair 1: 24 June 2021 – 06 July 2021 (baseline = 60.18 m)
- Pair 2: 06 July 2021 – 18 July 2021 (baseline = 118.5 m)
- Pair 3: 12 July 2021 – 24 July 2021 (baseline = 7.75 m)
- Pair 4: 18 July 2021 – 30 July 2021 (baseline = 85.9 m)

A spatial subset of approximately 1000×2000 pixels is extracted from the full scene to enable detailed analysis while maintaining computational efficiency. These interferometric pairs provide multiple realizations of coherence under varying baseline conditions, enabling the analysis of temporal decorrelation effects and their impact on forest structure retrieval. For each interferometric pair, dual-polarimetric coherence measurements are used to construct the polarimetric coherence locus, from which the optimized coherences γ_{max} and γ_{min} are derived. As described in the previous section, γ_{max} corresponds to a projection that minimizes the ground contribution and is therefore primarily sensitive to vegetation volume decorrelation, while γ_{min} is associated with stronger ground scattering contributions. Figure 2 presents the spatial distribution of the optimized interferometric coherences for the four Sentinel-1 pairs. The first row shows γ_{max}, while the second row shows γ_{min} for the corresponding acquisitions. The coherence values are displayed using a consistent color scale limited to the range [0,

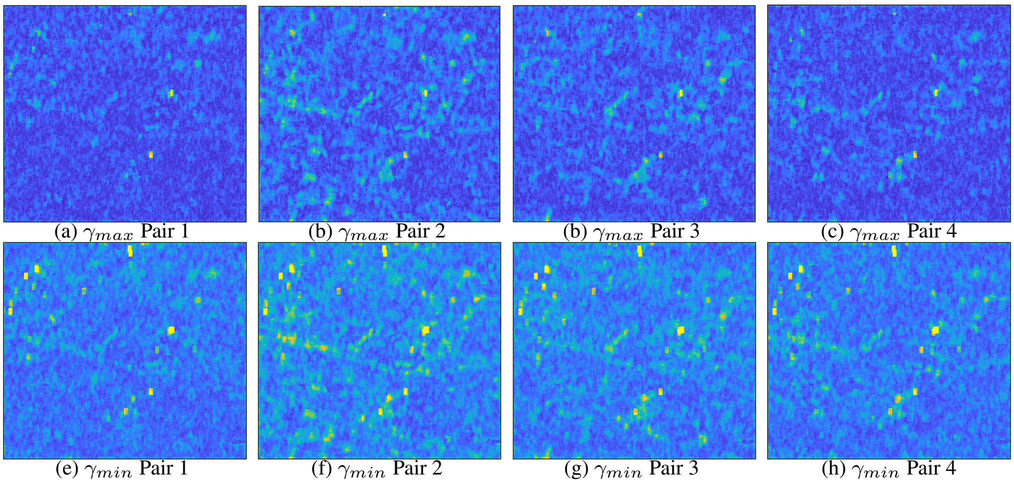


Figure 2. Spatial distribution of optimized interferometric coherences for the Kotka study area. The first row (a–d) shows γ_{max} and the second row (e–h) shows γ_{min} for four Sentinel-1 interferometric pairs. The color scale is limited to the range [0, 0.7].

0.7]. The γ_{max} maps exhibit generally lower coherence values across the scene, reflecting the dominance of vegetation-induced temporal decorrelation when the ground contribution is suppressed. In contrast, the γ_{min} maps show higher coherence values, as the ground scattering component—being more temporally stable—contributes more strongly to the interferometric signal. This contrast between γ_{max} and γ_{min} highlights the ability of dual-polarimetric observations to separate ground and volume scattering contributions, thereby providing complementary information for the estimation of forest structural parameters. In particular, γ_{max} serves as a key observable for constraining vegetation-related decorrelation effects, while γ_{min} provides additional context on the ground-to-volume scattering ratio.

Both the model-based RMoG inversion and the data-driven refinement approach are implemented and evaluated over the Kotka study area. For the physical model, all four available interferometric pairs are jointly utilized to stabilize the inversion process. The use of multiple coherence observations allows for a more robust estimation of vegetation-related parameters by reducing the impact of noise and temporal decorrelation variability across individual acquisitions. In contrast, the data-driven model is trained on the global dataset and applied using a single interferometric pair (Pair 1: 24 June–06 July 2021), in order to assess the capability of the learning-based approach to compensate for the limited information content of single-baseline C-band observations.

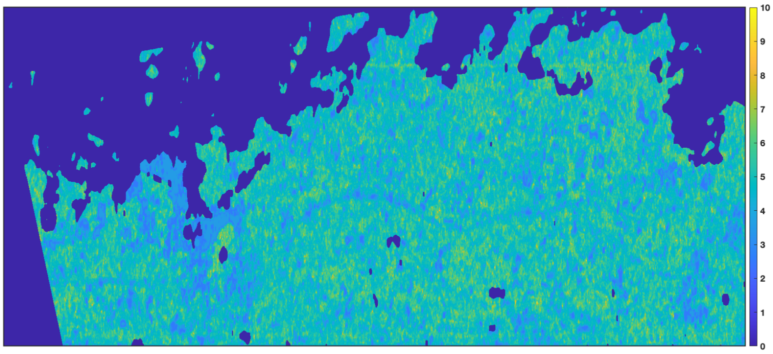
Figure 3 presents the forest height estimates obtained using both the RMoG-based physical model and the proposed hybrid approach with deep learning refinement, alongside the LiDAR reference. The results obtained from the RMoG inversion demonstrate that, despite the strong temporal decorrelation present in Sentinel-1 repeat-pass data, the model is able to capture the dominant spatial patterns of forest structure. In particular, regions of relatively low and high forest height are consistently identified, and the large-scale spatial variability shows a clear correspondence with the LiDAR reference. However, the RMoG-based estimates exhibit notable limitations in terms of accuracy and spatial detail. The reconstructed height map lacks fine-scale structural variability. Moreover, the model tends to underestimate forest height in densely tall vegetated regions. These limitations are reflected in the quantitative evaluation, where the RMoG model achieves a mean absolute error (MAE) of approximately $3.8, \mathrm{m}$, a root mean square error (RMSE) of $4.6, \mathrm{m}$, and a correlation coefficient of $R = 0.45$. Importantly, these results do not imply that the RMoG model is ineffective. On the contrary, it provides a physically consistent baseline that can captures the general spatial distribution of forest height. The application of the data-driven residual correction significantly improves the quality of the forest height estimates. The refined height map in Fig. 3 shows enhanced spatial consistency and closer agreement with the LiDAR reference, particularly in areas of tall forest structure. Quantitatively, the combined physical model and deep learning approach achieves a mean absolute error (MAE) of approximately $2.31, \mathrm{m}$, a root mean square error (RMSE) of $2.99, \mathrm{m}$, and a correlation coefficient of $R = 0.74$, demonstrating a substantial improvement over the physical model results.

It is worth noting that the employed neural network architecture is relatively simple and computationally efficient. Nevertheless, due to training on a large-scale, globally distributed dataset, the model is able to generalize well and learn meaningful correction patterns. This highlights the effectiveness of combining a physically grounded model with a lightweight data-driven refinement strategy for forest height estimation using Sentinel-1 data.

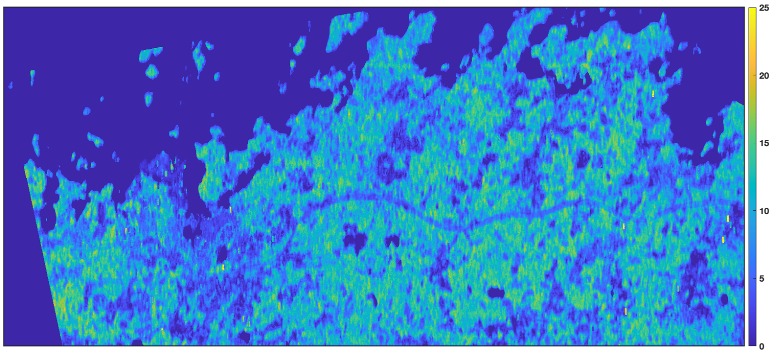
4. Conclusion

This study presented a framework for forest height estimation from Sentinel-1 repeat-pass interferometric observations, combining a simplified physical RMoG model with a data-driven residual correction strategy. A globally distributed SAR–LiDAR dataset was constructed by integrating Sentinel-1 measurements with GEDI and ICESat-2 canopy height references, enabling a systematic analysis of the relationship between interferometric coherence and forest structure across diverse conditions.

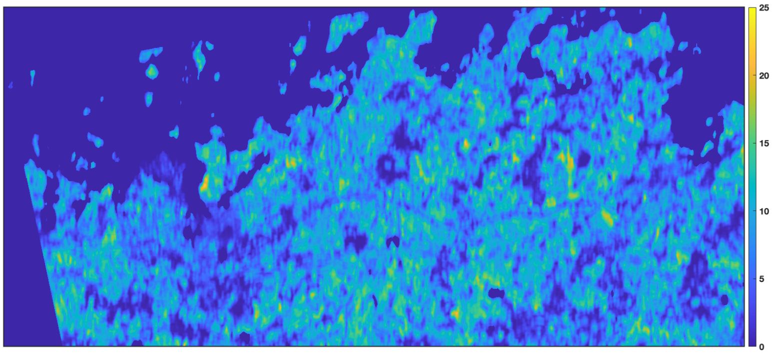
The results confirm that, despite strong temporal decorrelation, Sentinel-1 coherence retains meaningful sensitivity to forest height under near-zero spatial baseline conditions. The RMoG-based approach provides a physically consistent baseline capable of capturing the dominant spatial patterns of forest structure, although it remains limited in resolving fine-scale variability and in accurately estimating height in tall forest regions. The proposed residual learning strategy effectively addresses these lim-



(a) RMoG-based forest height estimate



(b) Hybrid (RMoG + deep learning) estimate



(c) LiDAR reference canopy height

Figure 3. Forest height estimation results for the Kotka study area. (a) Forest height derived from the RMoG-based physical model, (b) refined forest height after applying the deep learning residual correction, and (c) reference canopy height derived from LiDAR measurements. The RMoG model captures the dominant spatial patterns but lacks fine-scale detail and tends to underestimate height in dense forest regions, while the hybrid approach improves spatial consistency and agreement with the reference.

itations, leading to improved accuracy and enhanced spatial consistency of the retrieved forest height maps.

Overall, this work demonstrates that combining physically grounded models with lightweight data-driven refinement offers a practical and scalable solution for forest height estimation using operational C-band SAR data without relying on external data sources. The presented dataset and methodology provide a foundation for future developments, particularly in the context of upcoming low-frequency SAR missions such as ROSE-L, where improved coherence stability may further enhance the retrieval of forest structural parameters.

References

Aghababaei, H. et al., 2020. Forest SAR tomography: Principles and applications. *IEEE Geoscience and Remote Sensing Magazine*, 8(2), 30–45.

Cloude, S. R., Papathanassiou, K. P., 1998. Polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5), 1551–1565.

Cloude, S. R., Papathanassiou, K. P., 2003. Three-stage inversion process for polarimetric SAR interferometry. *IEE Proceedings - Radar, Sonar and Navigation*, 150(3), 125–134.

Lavalle, M., Simard, M., Hensley, S., 2012. A Temporal Decorrelation Model for Polarimetric Radar Interferometers. *IEEE Transactions on Geoscience and Remote Sensing*, 50(7), 2880-2888.

Lavalle, M., Telli, C., Pierdicca, N., Khati, U., Cartus, O., Kellndorfer, J., 2023. Model-Based Retrieval of Forest Parameters From Sentinel-1 Coherence and Backscatter Time Series. *IEEE Geoscience and Remote Sensing Letters*, 20, 1-5.

Le Toan, T., Quegan, S., Davidson, M. W. J., Balzter, H., Paillou, P., Papathanassiou, K., Ulander, L. et al., 2011. The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sensing of Environment*, 115(11), 2850–2860.

Olesk, A., Praks, J., Antropov, O., Zalite, K., Arumäe, T., Voormansik, K., 2016. Interferometric SAR coherence models for characterization of hemiboreal forests using TanDEM-X data. *Remote Sensing*, 8(9), 700.