URSI AP-RASC, Sydney, Australia, 17 – 22 August 2025

Forest Height Estimation Training Dataset for SAR Interferometry: The Sentinel2Height Project

Jaan Praks[1], Erkki Tomppo[2], Hossein Aghababaei[3], Giampaolo Ferraioli[4], Gilda Schirinzi[4]
[4]Department of Electronics and Nanoengineering, Aalto University, Finland
[2]Department of Forest Sciences University of Helsinki, Finland
[3]Department of Earth Observation Science, ITC, University of Twente, The Netherlands
4Universita degli Studi di Napoli "Parthenope", Naples, Italy

Extended Abstract

Accurate forest height estimation is essential for monitoring forest health, carbon storage, timber resources, biodiversity, and ecosystem dynamics. It provides critical insights for sustainable forest management and climate change mitigation. While Interferometric Synthetic Aperture Radar (InSAR), particularly single-pass interferometry with a suitable baseline, has demonstrated significant potential in this domain [1, 2], small-baseline repeat-pass interferometry still faces considerable challenges due to temporal decorrelation. One approach to mitigating this issue is the use of machine learning (ML) algorithms, which are increasingly favored for their scalability and ability to process large datasets efficiently.

Machine learning, leveraging the global coverage of Sentinel-1 Synthetic Aperture Radar (SAR), offers a promising solution for automating forest height estimation. However, the effectiveness of AI-driven approaches depends on access to high-quality benchmark datasets with comprehensive annotated reference data. To address this need, we introduce a new benchmark dataset that integrates pre-processed Sentinel-1 polarimetric interferometric data with forest height annotations derived from airborne and spaceborne LiDAR. This dataset spans diverse forest types worldwide, incorporates varying temporal baselines, and includes geocoded polarimetric-interferometric covariance matrices, providing a robust foundation for developing advanced machine learning models. By integrating SAR and LiDAR data, researchers can gain deeper insights into canopy structure and biomass distribution, further enhancing the accuracy of forest height estimations.

In this presentation, we introduce Sentinel2Height: A Synergic AI-SAR Tool for Elevated Insights into Forestry, a collaborative project aimed at advancing AI-SAR methodologies for forest monitoring and management. Led by the University of Twente (The Netherlands) in partnership with Università degli Studi di Napoli “Parthenope” (Italy), the University of Helsinki (Finland), and Aalto University (Finland), the project is funded by the European Space Agency (ESA). A key objective of Sentinel2Height is assembling a vast and diverse training dataset to support the development and validation of machine learning models for forest height extraction from Sentinel-1 interferometric data. By providing a comprehensive and well-annotated dataset, the project aims to enhance the accuracy and reliability of AI-driven forest height estimation on a global scale.

Through the development of machine learning-based solutions for forest height estimation, Sentinel2Height seeks to support more effective conservation efforts, facilitate better-informed policy decisions, and contribute to improved resource management strategies worldwide.

References

[1] J. Praks, O. Antropov, and M. T. Hallikainen. “LIDAR-Aided SAR Interferometry Studies in Boreal Forest: Scattering Phase Center and Extinction Coefficient at X- and L-Band”. In: *IEEE Transactions on Geoscience and Remote Sensing* 50.10 (Oct. 2012), pp. 3831–3843. ISSN: 0196-2892. DOI: 10.1109/TGRS.2012.2185803.

[2] Aire Olesk et al. “Interferometric SAR Coherence Models for Characterization of Hemiboreal Forests Using TanDEM-X Data”. In: *Remote Sensing* 8.9 (2016), p. 700. ISSN: 2072-4292. DOI: 10.3390/rs8090700. URL: http://www.mdpi.com/2072-4292/8/9/700.