Towards the understanding of the C-Band temporal signature of boreal forest through physiology parameters retrieval from sentinel-1 time series and machine learning

Thomas Di Martino, Régis Guinvarc'h, Laetitia Thirion-Lefevre, Elise Colin

TOWARDS THE UNDERSTANDING OF THE C-BAND TEMPORAL SIGNATURE OF BOREAL FOREST THROUGH PHYSIOLOGY PARAMETERS RETRIEVAL FROM SENTINEL-1 TIME SERIES AND MACHINE LEARNING

Thomas Di Martino[1,2], *Régis Guinvarc'h*[1], *Laetitia Thirion-Lefevre*[1], *Elise Colin*[2]

[1] SONDRA, ONERA, CentraleSupélec, Université Paris-Saclay, 91190 Gif-sur-Yvette, France
[2] ONERA, Traitement de l'Information et Systèmes, Université Paris-Saclay, 91123 Palaiseau, France

ABSTRACT

The C-Band radiometric signature of boreal forests is highly seasonal, with apparent correlations to temperature changes. Within these seasonal components, we assume that information related to tree height can be extracted. We apply a one-dimensional Convolutional Neural Network to assess this assumption, intending to retrieve tree height measured by Airborne Laser Scanning from C-Band Sentinel-1 time series. A study site in the Parc National des Grands Jardins, in Québec, Canada, was selected for this analysis. Prediction-wise, we reach an R2 score of 0.45 and an RMSE of 1.84m, following a 4-fold cross-validation, which exhibits a non-negligible influence of the tree height parameter on boreal forest radiometric response in C-Band Synthetic Aperture Radar, despite the presumed fast saturation of this wavelength, when observing forested environments. In addition to performance metrics, we use a gradient-based explainability tool to diagnose the most contributing periods of the input time series to predict tree height to better correlate the seasonal conditions of this parameter's influence on the forests' radiometry.

Index Terms— Deep Learning, Inversion, SAR, Time Series, Explainable AI

1. INTRODUCTION

The extraction of forest parameters relies on long wavelength configurations such as P-band or L-band SAR [1], as they offer better penetration of the canopy layer. However, access to P-band and L-band SAR data is limited. The current SAR satellite offering the best spatial and temporal coverage is ESA's Sentinel-1, which on-boards a C-band sensor. The understanding of the interaction of the electromagnetic wave of C-Band Synthetic Aperture Radar (SAR) imagery with complex environments such as boreal forests is still partial because of its more limited penetration capacities. The few studies analyzing the interaction of C-band wave with a forested environment alleviated the limitations mentioned above by relying on temporal approaches [2, 3]. An example of such a study can be found in [4], where the seasonal variations of the radiometric response of the forest are correlated with the soil state, with higher radiometry during summer and lower radiometry during winter under freezing conditions. The amplitude of summer-to-winter radiometric changes appears to be a direct function of the physiology of the tree [5]. Thus, with the problem of inversion in mind, while a single C-Band SAR acquisition may be insufficient, a shift towards multi-temporal SAR is expected to lead to better results [6]. In addition, with the launch of the new generation of Sentinel-1 C-Band satellites, unparalleled worldwide open-source coverage is available, with a short revisit time ranging from 12 to 6 days for some regions of the globe. This unprecedented amount of temporal information has the potential to ease the understanding of the relationship between the physiology of a forested area and its C-Band temporal profile.

This paper uses a deep learning approach to assess the retrievability of a commonly used forest parameter, tree height, from Sentinel-1 time series. We first present the study area and illustrate the seasonalities of the C-Band radiometry of forests by introducing Sentinel-1 and airborne laser scanning data. We then detail the model used, particularly a 1D Convolutional Neural Network. In the recent literature, 1D convolutions were shown to be excellent feature extractors when processing Sentinel-1 time series, even in unsupervised conditions [7]. The presentation of the application performance of the network to the retrieval of tree height from Sentinel-1 time series is accompanied by a study on its explainability, using the Grad-CAM algorithm, to study date contribution to the prediction.

2. MATERIALS AND METHODS

2.1. Study area: Parc National des Grands Jardins, Québec, Canada

Located in the Québec province of Canada, the Parc National des Grands Jardins is a unique ecosystem with a boreal forest similar to the taiga while remaining accessible for experimentation. Thus, considering the need to monitor boreal forests at scale, this park is prone to experimentation and measurements. In this paper, the ground truth forest measurements used consist of a rasterized 30m resolution tree height data product, originating from airborne LiDAR acquisitions, and illustrated in Fig. 1.

The tree height metric estimates the height of the top of the canopy, from the ground, in meters. The point measure is then averaged over resolution cells of 30m in length. The Parc appears highly diverse in tree height: the western side has lower tree heights, represented by light color. This may relate to more sparse vegetation and younger vegetation. Oppositely, the eastern side of the Parc contains many areas with up to 10m of height and more. In addition, in the center of the image, a vast area appears unwooded, deforested by a forest fire from the late 90s. In addition to LiDAR data, Sentinel-1 multitemporal imagery was also retrieved.

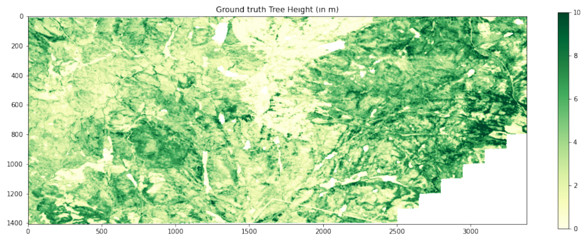

Fig. 1: Tree Height product illustration

2.2. Sentinel-1 multitemporal data

To exploit the seasonality of the radiometric signature of boreal forest, we use two years' worth of ground range detected Sentinel-1 data, which include the years 2018 and 2019. 56 acquisitions, with an average revisit time of 12 days, constitute the multitemporal stack of Sentinel-1 data. Each image consists of a decibel-scaled intensity value of both Vertical-Vertical and Vertical-Horizontal polarizations. A snippet of the average temporal image of the area layered over the LiDAR acquisition is displayed in Fig. 2. While some unwooded areas' tree height appears retrievable from the Sentinel-1 temporal average image, particularly the burnt area, it becomes less clear for others. Thus, a more thorough analysis of the Sentinel-1 time series temporal patterns is required. In Fig. 3, we display the average signature of both VV and VH polarizations as a function of time, alongside daily 25th and 75th radiometry percentiles and daily minimum temperatures. Variations between seasons appear, with high radiometry in the summer, and lower radiometry in the winter, for both VV and VH polarizations. In addition, these variations of radiometry appear to correlate with temperature variations. To assess the relationship between σ_0 and the temperature, we compute Pearson's correlation coefficient evaluation of temperature variations with VV, VH, and VV-VH in log scale. The correlation is subtle with VV time series, with a Pearson's R coefficient of 0.22 and rather more present with VH (0.56), considered volume backscattering, and often prevails in importance when monitoring vegetation. The most evident correlation lies in the relationship of the VV-VH time series and temperatures, with a negative correlation factor of -0.84: this high negative Pearson coefficient indicates that when the temperature drops, the gap between VV and VH widens. This seems to indicate a more substantial overall decrease in VH signal during temperature drops, in comparison to the decrease in VV signal.

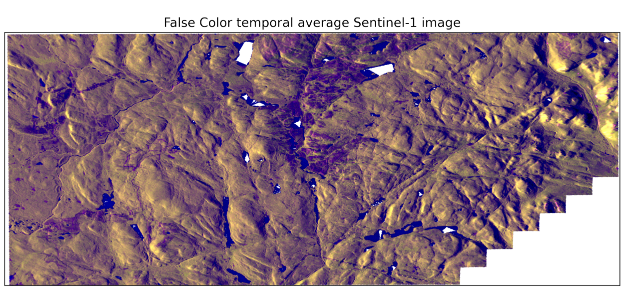

Fig. 2: False color Sentinel-1 image (R: VV, G: VH, B: VV-VH) with the respective value intervals $\{[-15,-5],[-20,-10],[0,15]\}$

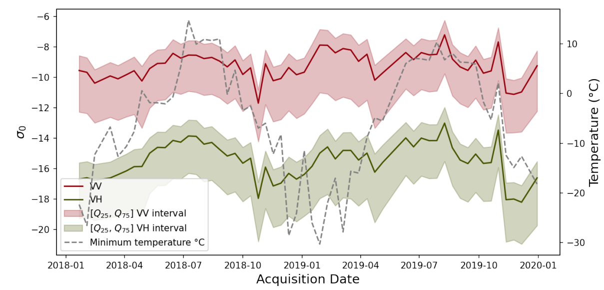

Fig. 3: Average temporal profile over the Parc des Grands Jardins

While we expect these temporal patterns to be connected to the physiological information of forests so that their inversion becomes possible to some degree, its demonstration requires a powerful modeling tool. In this paper, we rely on a 1D convolutional neural network that we now introduce.

2.3. The model, the Convolutional Neural Network

2.3.1. *Architecture*

A convolutional neural network (CNN) is a specific kind of artificial neural network using learned convolution layers to extract patterns in data. A familiar context for the application of CNNs is image processing, which relies on 2D convolutions to extract spatial features from images. In our case, the input data are time series, which are 1-dimensional sequences. They thus rather induce the use of 1D convolutions. The complete detail of the used architecture can be found in Tab. 1.

In particular, temporal feature extraction relies on three

Table 1: Architecture of the 1D Convolutional Neural Network used in this study

Layer	(Count, size, stride, padding)	Output size
Input time series	-	58×2
1D Convolution	(64, 7, 1, 1)	54×64
Max Pooling 1D	(-, 2, 2, -)	27×64
1D Convolution	(128, 5, 1, 0)	23×128
Max Pooling 1D	(-, 2, 2, -)	11×128
1D Convolution	(256, 3, 1, 0)	11×256
Max Pooling 1D	(-, 2, 2, -)	5×25
Flatten	-	1280
Fully Connected	-	128
Fully Connected	-	64
Fully Connected	-	32
Fully Connected	-	1

stacks of convolutions and pooling, all 1D-parameterized. The activation function in all layers is the Exponential Linear Unit (ELU) [8]. Four separate fully-connected layers then process the features before making a tree height prediction.

2.3.2. *Explainability, and Grad-CAM*

To extend the outcomes of applying a CNN to this regression task, we rely on a gradient-based explainability tool called Grad-CAM [9]. Initially developed for image processing and classification, we adapt its functioning and equations to consider the sequential nature of our input data and the regression task at hand. With this tool, we can extract date contribution to the regression prediction and better understand important dates where height information radiometrically manifests itself.

2.3.3. *Training setup*

The proposed network's training involves using various set of hyper-parameters, which include a learning rate set to $1e-3$, a batch size to 1024, and an epoch count to 20. In addition, the network is trained using the Adam optimizer. The choice of parameters was empirically optimized. In addition, a 4-fold cross-validation scheme was used for validation, with a handmade slice of the dataset in 4 separate spatially coherent areas. Using spatially coherent splitting prevents any data leakage between sets coming from the correlation of neighboring time series, which could be caused by Sentinel-1 preprocessing steps such as multi-looking or despeckling.

3. RESULTS

Following the previously mentioned 4-fold cross-validation scheme, we obtain an R2 score between the ground truth and the predictions of 0.45 and an RMSE of 1.84m.

A first visualization of the proposed CNN network's prediction results is illustrated in Fig. 4, and overall, the predictions appear accurate. In particular, unwooded areas are well retrieved: the forest fire in the center of the image is retrieved, and areas of lower/higher height appear distinguish-

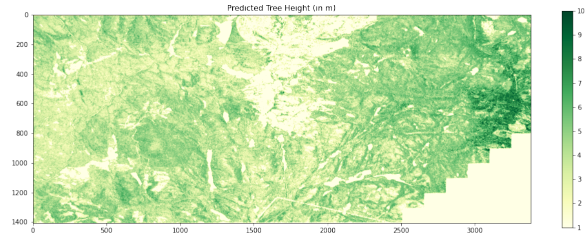

Fig. 4: Tree Height prediction map

able. Parts of the northeastern road network, mid-western waterways, and the high-voltage powerline's pathway are distinguishable.

However, the first apparent difficulties are in the high-density areas, particularly where we can find >10m trees: the prediction of the CNN underestimates the true height. This observation is made clearer by the density-colored prediction scatter plot in Fig. 5. Various explanations can justify the

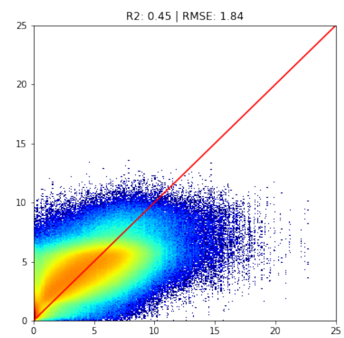

Fig. 5: Tree Height density-colored prediction scatter plot

difficulties encountered by the network: high environments are also high in biomass density. Aside from underestimating high trees, this scatter plot also sheds light on the overestimation of the smallest. This underestimation showcases a similar difficulty in dealing with extremums which are, in this case, highly sparse vegetated environments. Thus, a combination of Sentinel-1 time series and a 1D CNN is not operational for inversion purposes. However, it still successfully extracts groups of vegetation types with similar heights.

4. ANALYSIS AND EXPLAINABILITY

Now that we have answered the question "*does it work ?*", we orient our analysis towards "*How does it work?*". We plot the Grad-CAM-based date importance using a color-coded average Sentinel-1 profile over the Parc. To assess particular radiometric events which play a significant role in the differentiation between tree heights, we illustrate 20 S1 color-coded time series, averaged iteratively over the groups of data points, gathered in buckets of 5% of heights' percentiles. Their respective Grad-CAM coloring is thus calculated separately.

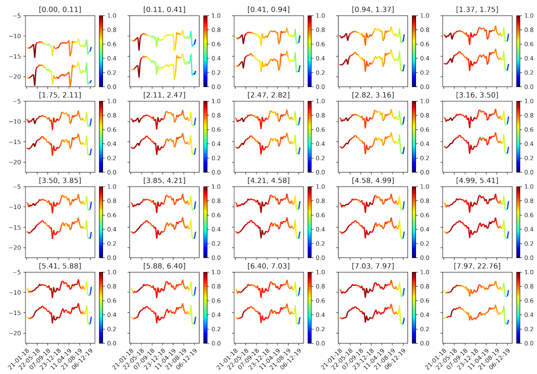

Fig. 6: Grad-CAM colored average radiometric time series, split into groups of tree height

These time series are displayed in Fig. 6. The main observation regarding a significant difference in temporal radiometry between the various tree height groups lies in three distinct signal drops. Their intensity appears inversely proportional to tree height cover, and these drops almost entirely disappear in high tree scenarios. These three drops are the most important dates for tree height regression, particularly for smaller trees. In addition to helping to understand the difference between various tree height radiometries, Fig. 6 also sheds light on the confusion between groups of higher trees: their average signature appears highly similar in C-Band, which disturbs their regression.

5. CONCLUSION

This paper assessed the retrievability of tree height from C-Band SAR time series. Acknowledging the task's difficulty by the theorized saturation of C-band wavelength to biomass levels, we relied on a state-of-the-art prediction algorithm using a 1D Convolutional Neural Network. We achieve performances of 0.45 as an R2 score which, while not strong enough for pure inversion, exhibits the potential of multitemporal C-Band SAR to be used for inversion of tree height alongside other imagery, such as optical imagery or different wavelengths of SAR data. In particular, thanks to its high temporal resolution, Sentinel-1 has the unique capacity among open data earth observation sensors to monitor forest seasonality finely.

6. REFERENCES

[1] M.C. Dobson, F.T. Ulaby, T. LeToan, A. Beaudoin, E.S. Kasischke, and N. Christensen, "Dependence of radar backscatter on coniferous forest biomass," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 2, pp. 412–415, 1992.

[2] Juval Cohen, Kimmo Rautiainen, Jaakko Ikonen, Juha Lemmetyinen, Tuomo Smol²ander, Juho Vehvilêinen, and Jouni Pulliainen, "A modeling-based approach for soil frost detection in the northern boreal forest region with C-band SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 2, pp. 1069–1083, 2019.

[3] Maurizio Santoro, Christian Beer, Oliver Cartus, Christiane Schmullius, Anatoly Shvidenko, Ian McCallum, Urs Wegmüller, and Andreas Wiesmann, "Retrieval of growing stock volume in boreal forest using hyper-temporal series of envisat asar scansar backscatter measurements," *Remote Sensing of Environment*, vol. 115, no. 2, pp. 490–507, Feb 2011.

[4] E. Rignot, Jo Way, Kyle Mcdonald, Leslie Viereck, Cynthia Williams, Phyllis Adams, Cheryl Payne, William Wood, and Jiancheng Shi, "Monitoring of environmental conditions in Taiga forests using ERS-1 SAR," *Remote Sensing of Environment*, vol. 49, pp. 145–154, 08 1994.

[5] J.T. Pulliainen, L. Kurvonen, and M.T. Hallikainen, "Multitemporal behavior of l- and c-band sar observations of boreal forests," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 2, pp. 927–937, Mar 1999.

[6] Moritz Bruggisser, Wouter Dorigo, Alena Dostálová, Markus Hollaus, Claudio Navacchi, Stefan Schlaffer, and Norbert Pfeifer, "Potential of Sentinel-1 C-band time series to derive structural parameters of temperate deciduous forests," *Remote Sensing*, vol. 13, no. 4, 2021.

[7] Thomas Di Martino, Régis Guinvarch, Laetitia Thirion-Lefevre, and Elise Colin Koeniguer, "Beets or cotton? blind extraction of fine agricultural classes using a convolutional autoencoder applied to temporal sar signatures," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

[8] Djork-Arné Clevert, Thomas Unterthiner, and Sepp Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," *arXiv preprint arXiv:1511.07289*, 2015.

[9] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, oct 2019.