Vegetation Canopy Height Retrieval Using L1 and L5 Airborne GNSS-R

J. F. Munoz-Martin, *Member, IEEE*, D. Pascual, *Member, IEEE*, R. Onrubia, *Member, IEEE*, H. Park, *Senior Member, IEEE*, A. Camps, *Fellow, IEEE*, C. Rüdiger, *Senior Member, IEEE*, J. P. Walker, *Fellow, IEEE*, and A. Monerris, *Senior Member, IEEE*

***Abstract*— Vegetation canopy height (CH) is one of the important remote-sensing parameters related to forests' structure, and it can be related to the biomass and the carbon stock. Global navigation satellite system-reflectometry (GNSS-R) has proved capable to retrieve vegetation information at a moderate resolution from space (20–65 km) using L1 C/A signals. In this study, data retrieved by the airborne microwave interferometric reflectometer (MIR) GNSS-R instrument at L1 and L5 are compared to the Global Forest CH product, with a spatial resolution of 30 m. This work analyzes the waveforms (WFs) measured at both bands, and the correlation of the waveform width and the reflectivity values to the CH product. A neural network algorithm is used for the retrieval, showing that the combination of the reflectivity and the waveform width allows to estimate the CH information at a very high resolution, with a root-mean-square error (RMSE) of 4.25 and 4.07 m at L1 and L5, respectively, which is an error about 14% of the actual CH.**

***Index Terms*— Airborne, artificial neural network (ANN), canopy height (CH), global navigation satellite system-reflectometry (GNSS-R), L1 and L5, vegetation.**

I. Introduction

GLOBAL navigation satellite system-reflectometry (GNSS-R) is a relatively new and valuable technique to remotely sense several essential climate variables (ECVs) [1], [2]. In the past years, L1 GNSS-R data over land has been analyzed to retrieve soil moisture [3]–[5], above-ground biomass (AGB) [6], [7], or canopy height (CH) [7], showing promising results in all three cases.

Manuscript received August 9, 2021; revised October 15, 2021 and November 5, 2021; accepted November 24, 2021. Date of publication November 29, 2021; date of current version January 6, 2022. This work was supported in part by the Spanish Ministerio de Ciencia, Innovación y Universidades (MCIU) and European Union European Regional Development Fund (EU ERDF) Project "Sensing With Pioneering Opportunistic Techniques" under Grant RTI2018-099008-B-C21/AEI/10.13039/501100011033, in part by the "CommSensLab–Universitat Politecnica de Catalunya (UPC)" Excellence Research Unit Maria de Maeztu [Ministerio de Asuntos Económicos y Transformación Digital (MINECO)] under Grant MDM-2016-600, and in part by the grant for recruitment of early-stage research staff of the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR)—Generalitat de Catalunya [Fondo Europeo de Desarrollo Regional (FEDER)], Spain, under Grant FI-DGR 2018. *(Corresponding author: J. F. Munoz-Martin.)*

J. F. Munoz-Martin, D. Pascual, R. Onrubia, H. Park, and A. Camps are with the CommSensLab, Department of Signal Theory and Communications, Universitat Politecnica de Catalunya–BarcelonaTech, 08034 Barcelona, Spain, and also with IEEC/CTE-UPC, 08034 Barcelona, Spain (e-mail: joan.francesc@tsc.upc.edu; adriano.jose.camps@upc.edu).

C. Rüdiger and A. Monerris are with the Bureau of Meteorology, Melbourne, VIC 3001, Australia.

J. P. Walker is with the Department of Civil Engineering, Monash University, Melbourne, VIC 3800, Australia.

Digital Object Identifier 10.1109/LGRS.2021.3131263

GNSS-R using L1 C/A signals has been widely studied from both low-altitude (i.e., airplane or balloon) or high-altitude (i.e., satellite) platforms to retrieve vegetation-related parameters. As it is shown in [8], the presence of dense vegetation is linked to a decrease in the GNSS-R signal-to-noise ratio (SNR), showing that the reflectivity can be linked to vegetation information. The same correlation is studied in [9], where the GNSS-R reflectivity and SNR have a certain degree of correlation with the leaf area index and the CH. Other studies show that polarimetric GNSS-R from a high-altitude balloon could be used to infer AGB [10]. In [7], it is shown that GNSS-R data collected by the U.K. TechDemoSat-1 (TDS-1) mission and the cyclone GNSS (CyGNSS) mission have a certain degree of correlation to the soil moisture active passive (SMAP) vegetation optical depth (VOD). This last work proposes an artificial neural network (ANN) algorithm to retrieve both AGB and CH using CyGNSS reflectivity measurements. Aggregating different data points into a regular grid of 5×5 km^2, a root-mean-square error (RMSE) of 6.5 m is achieved. However, the actual single-pass resolution of GNSS-R for vegetation monitoring is larger than 5×5 km^2. In [6], it is shown that the trailing edge of the GNSS-R L1 C/A waveform over land collected by CyGNSS has a strong correlation with both SMAP VOD at a spatial resolution of ∼20 km. As the GNSS-R reflection is dominated by incoherent scattering, mostly when long integration times are used, the spatial resolution is given by the large glistening zone [11]–[13].

The widening of the trailing edge shown in [6] is a direct consequence of the signal wavefront of a mixture of the volume scattering produced in the vegetation canopy, and the bare soil reflection, which is then incoherently averaged in the receiver [1], [14]. The received signal for a bi-static radar configuration is a combination of the volume scattering produced at the vegetation canopy, and the bare soil, as sketched in Fig. 1.

For narrow bandwidth signals, such as the GPS L1 C/A signal, the volume scattering produces a widening of the L1 C/A auto-correlation function (ACF) (ACF width ∼300 m). However, this phenomenon has been never studied for GPS L5 signals, where the signal bandwidth is ten times larger (the spatial resolution is ∼30 m). In such case, other studies [15], [16] with wide-band radars show that the retrieved waveform could exhibit multiple peaks due to reflections in different layers (e.g., soil reflection, soil-trunk reflection, and soil-volume scattering, etc.).

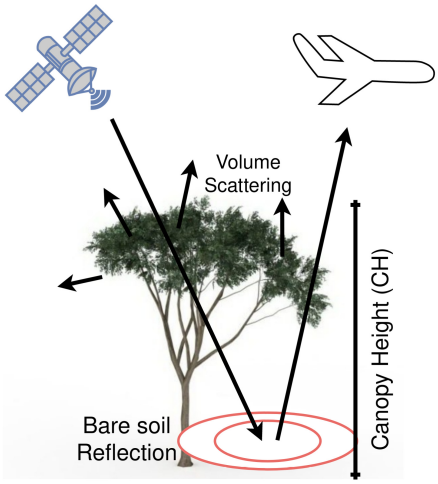

Fig. 1. Sketch of the GNSS-R scattering geometry of over rain-forest vegetation. The GNSS transmitted signal suffers volume scattering in the vegetation canopy.

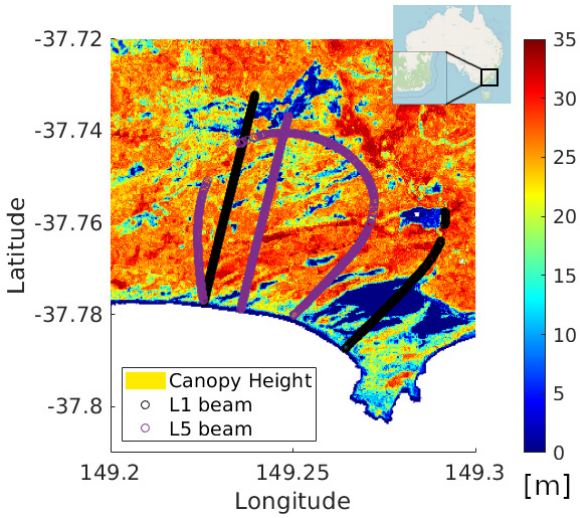

Fig. 2. L1 and L5 MIR specular reflection point location overlaid by the CH map retrieved from GFCH [21].

This work presents several algorithms to retrieve CH over land, and it compares the performance between the GPS L1 C/A and the GPS L5 signals, for the first time. The letter is organized as follows: Section II presents the region where the study has been conducted, and the dataset used. Section III presents the methodology, detailing the correlation between the CH and different GNSS-R observable at L1 C/A and L5. Then, Section IV presents an ANN algorithm proposed to estimate CH from airborne GNSS-R observables, providing a higher resolution than actual GNSS-R products. Finally, Section V discusses the conclusions of the letter.

II. Data Description and Study Region

The microwave interferometric reflectometer (MIR) [17] is a dual-band and dual-constellation (L1/E1 and L5/E5a) airborne GNSS-R receiver. The MIR instrument flew over the Bass Strait, Australia, in June 2018. Although the objective of the flight was mainly to acquire GNSS-R data over the ocean, it also entered into the mainland. In particular, the airplane flew over the "Croajingolong National Park," which is entirely covered by rainforest. Fig. 2 presents two selected GNSS-R tracks at L1 and L5 over the rainforest area. As shown in Chapter 11 (pp. 282–283) from [18], the topography of the area varies from 0 m near the coastline up to 120 m, but with low-moderate gradients (∼6-m height per second of flight time).

In order to maximize the GNSS-R spatial resolution, the GNSS-R data is incoherently averaged with a relatively short integration time of 100 ms, preventing waveform blurring [19]. As the plane velocity is v_{plane} ∼75 m/s, and the size of the first Fresnel zone at an altitude of h ∼1500 m is ∼30 m [20], the selected incoherent integration time prevents that multiple Fresnel zones are integrated within a single observation, maximizing the spatial resolution, notably at L5, where the ACF width is 30 m.

The vegetation information in Fig. 2 is from the global forest canopy height (GFCH) [21] product. This product provides a very fine spatial resolution of 30 m, and it is generated from the combination of LIDAR and multi-spectral data. Note that the CH information is linked to the VOD product, as it is shown in Fig. 3 from [22]. Thus, this high-resolution product has been selected to assess the capabilities of GNSS-R to retrieve a high-resolution vegetation product. The GFCH product is 2-D linearly interpolated into the specular point positions, providing a total amount of 981 co-located CH measurements for the L1 case, and 1655 co-located CH measurements for the L5 case.

III. Methodology

The interaction with different layers of the canopy produces two phenomena. First, the amplitude of the reflected signal is attenuated. As shown in the studies of [23], the attenuation of the reflected signal depends on the local incidence angle and the VOD. Second, due to scattering the leading and trailing edges of the GNSS-R waveforms (WFs) widen, which can be correlated with the CH, and as shown in [6], it can also be indirectly linked to the AGB.

A. Waveform Analysis

Fig. 3 presents two examples of GPS L1 C/A WFs retrieved by MIR over the Australian rainforest. As it can be seen, the width of the two WFs varies. Waveform 1 in Fig. 3, where the CH in the surrounding area is ∼13.2 m, presents a WF width of ∼280 m, while for Waveform 2 from Fig. 3, the CH is ∼30.8 m, and its WF width is ∼397 m. In this case, the WF width is measured by the distance between the $1/e$ amplitude (0.3678), measured after WF normalization, as sketched in Fig. 3. To have a better precision estimating the actual width, the WFs are re-sampled to 128 MS/s (2.3-m delay-bin size) from the initial 32 MS/s (9.4-m delay-bin size) value using a Fourier optimum interpolation method [24].

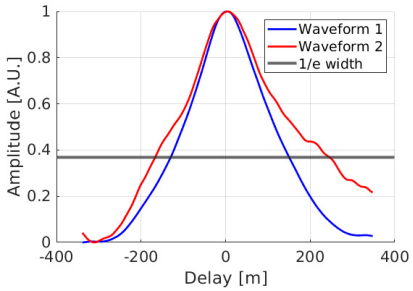

Fig. 3. GPS L1 C/A waveform examples. Waveform 1 (in blue) with a ground-truth CH of ∼13.2 m, producing a WF width of 280 m, and Waveform 2 (in red) with a ground-truth CH of ∼30.8 m, producing a WF width of ∼397 m.

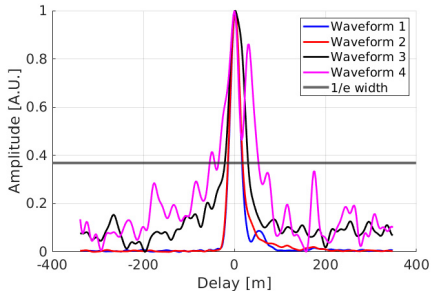

Fig. 4. Four GPS L5 WFs example for different CH: 17.63, 23.86, 30.61, and 30.81 m for WFs 1–4, respectively.

For the GPS L5 case, as the ACF is ten times narrower than in L1 C/A, the retrieved waveform presents either multiple peaks or a peak widening, similar to the L1 C/A case. Multiple peaks appear in ∼20% of the retrieved WFs over vegetated areas. Fig. 4 shows four different WFs detected by MIR for different CHs. Waveform 1 corresponds to a CH of 17.63 m, Waveform 2 to a CH of 23.86 m, and WFs 3 and 4 to a CH of 30.61 and 30.81 m, respectively. In this case, the waveform widths at $1/e$ are 36.62, 41.19, 54.93, and 96.13 m for WFs 1–4, respectively. Note that the larger the CH, the larger the WF width, but when two peaks appear, as in Waveform 4, the measured WF width is largely increased. In the L5 case, due to the presence of multiple peaks, the $1/e$ criterion cannot be used to measure the WF width, as some WFs are contaminated by nearby peaks. In this case, and only for L5, the WF width is computed at half its maximum (0.5) instead of being measured at $1/e$.

B. WF Width and Reflectivity Correlations With CH

The WF width and the reflectivity at L1 and L5 are retrieved for the points presented in Fig. 2. The CH information, which is gridded at 30 m, is 2-D linearly interpolated on top of the specular point location. As seen in Fig. 5, the CH and the reflectivity do correlate to the CH. The variable that correlates less is the WF width at L5. The Pearson correlation coefficient (R) with respect to the CH information is 0.26 for the WF width at L1, and −0.26 for the reflectivity at L1. At L5, the WF width present a similar correlation ∼0.20, and the reflectivity at L5 has a correlation coefficient of −0.46. In addition, the correlation coefficient between the reflection incidence angle

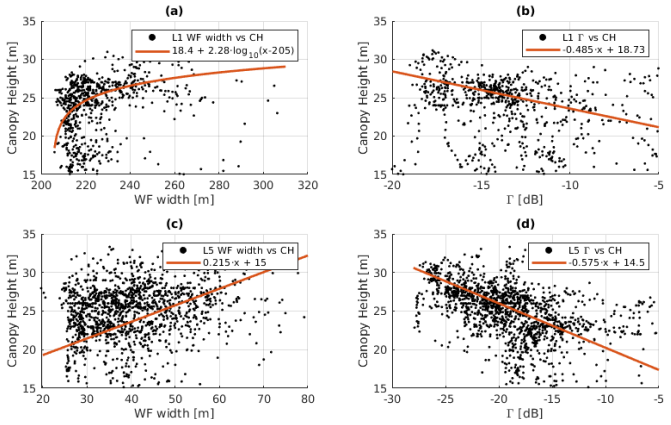

Fig. 5. Scatter plots and line-fit for (a) L1 WF width and CH, (b) L1 reflectivity (Γ) and CH, (c) L5 WF width and CH, and (d) L5 reflectivity (Γ) and CH.

and the CH is lower than 0.05 at both bands, showing that there is not a direct correlation between these magnitudes.

Even though the correlation coefficient is very low, the combination of WF width, reflectivity, and local incidence angle can provide additional information. Notwithstanding that the incidence angle is uncorrelated with the CH, the vegetation attenuation of the reflected signal depends on the local incidence angle. Multi-regression analysis is sometimes used to perform a first investigation of a multivariate problem [25]. In this case, a canonical correlation algorithm has been used to minimize the error between a three-column matrix and one-column output (WF width, reflectivity, and incidence angle, versus the ground-truth CH data) by multiplying the terms by a canonical vector. The algorithm uses singular value decomposition (SVD) and reduces the three-column matrix into a "U" vector, and the one-column matrix into a "V" vector, which are a linearly dependent of the WF width, reflectivity, and incidence angle, and of the CH one-column matrix. The results of the canonical correlation are shown in Fig. 6. As it can be seen, the correlation coefficient between the U and V vectors is 0.48 at L1 and 0.51 at L5. This increase in the correlation coefficient with respect to the pair-wise case is showing that just a linear combination of these three variables allows to retrieve CH information with a moderate correlation, slightly higher at L5. The canonical vectors that linearly combine the three-column inputs (i.e., the WF width, reflectivity, and incidence angle inputs) into the "U" vectors are $a = [-0.035, 0.134, 0.125]$, and $b = [0.046, -0.187, 0.070]$, respectively, for L1 and L5. The scaling coefficient for the output vector (the one mapping the CH to a "V" vector) is −0.132 for L1 and 0.162 for L5.

C. ANN for Multivariate Analysis

As shown in the canonical correlation analysis, the WF width and the reflectivity are correlated with the CH magnitude. However, the correlation could be further improved using other data-driven techniques. ANNs are one type of data-driven algorithms that have been broadly used to retrieve different ECV using GNSS-R data.

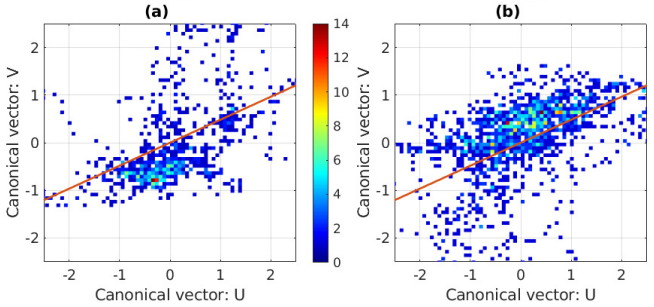

Fig. 6. Scatter plot of the canonical correlation vectors (U and V) at (a) L1 and (b) L5.

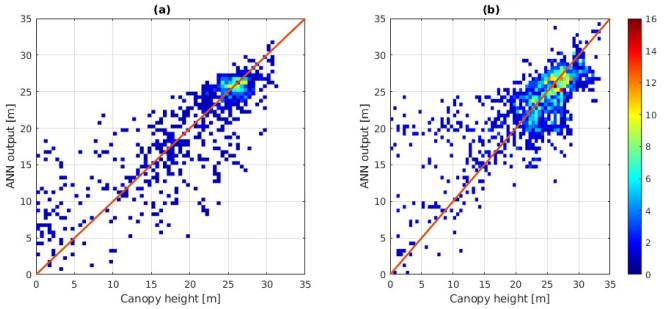

Fig. 7. Scatter plots between the ANN output and the CH ground truth for (a) L1 and (b) L5.

To address the retrieval of CH using GNSS-R information, an ANN is proposed for each band (L1 and L5). The network topology selected is very simple, using a two-layer network with seven neurons per layer (total of 14 neurons). All neurons use the sigmoid transfer function. The dataset is split into 50% for training, 15% for validation, and finally 35% for testing. In order to prevent over-fitting and to provide a generalized output, the Bayesian regularization method [26] is used. Furthermore, the networks are pruned after training to prevent over-fitting [27].

IV. Results

Fig. 7 presents the results of the ANN, and the data retrieved is compared to the ground-truth CH from GFCH. For the L1 case, the Pearson correlation coefficient is $R = 0.82$, and for the L5 case, $R = 0.75$, and the RMSE is 4.25 m at L1 and 4.07 m at L5. As it can be seen in Fig. 8, most of the discrepancies are found when transitioning from the rainforest into non-vegetated areas (i.e., sand dunes). In these transition cases, the errors are larger, notably at L5, where in some low-vegetation areas, there is almost no error (e.g., sample count 213 and 576), but some areas (e.g., sample counts 664 and 726) present a very large error with respect to the ground-truth data. In the first two low-vegetated areas, even though the canopy has decreased, it is around 2.5 and 6 m, respectively. However, in the other two cases, the reflection is occurring in the bare soil, where the ground-truth data is showing a vegetation height below the meter. As it can be seen in Fig. 9, the dune-vegetation transition is showing a large error. Two phenomena can drive this error: first, the reflectivity of dry bare soil (e.g., sand) is very low, and the ACF can also

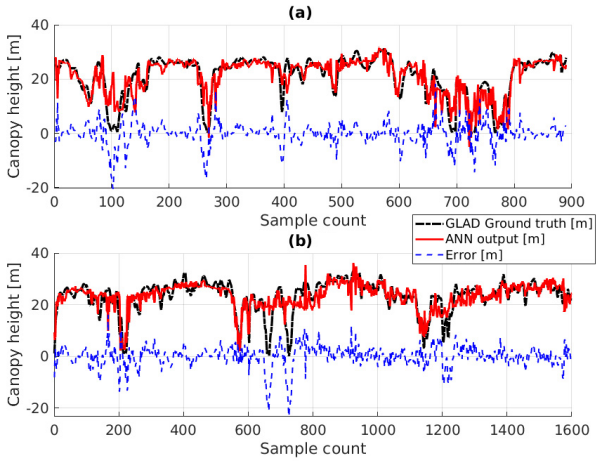

Fig. 8. CH evolution with flight time, comparison between the GFCH ground truth and the ANN estimation. (a) L1. (b) L5.

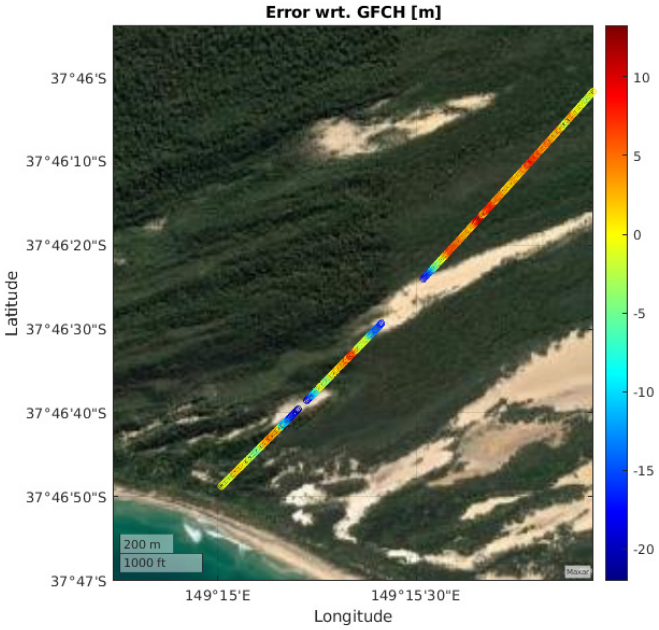

Fig. 9. Map detail of a dune-vegetation transition, where a large error is present in the retrieval algorithm.

contain different peaks due to reflections in nearby canopy (30–100 m from the specular point, similar to [20]) or other nearby elements in the reflection scenario.

V. Conclusion

This study has analyzed the correlation between the CH to airborne L1 and L5 GNSS-R data. It is presented that some L5 WFs exhibit multiple peaks. It has been shown that both the WF width data and the reflectivity data are correlated with the CH data and that ANN algorithms can "merge" these datasets to retrieve CH. A first evaluation has been performed using the canonical correlation method and then two ANNs have been implemented (one per each band). The result shows similar performances at L1 and L5, with an RMSE of 4.25 m at L1 and

4.07 m at L5. Furthermore, dune-vegetation transitions have been discussed, where multiple reflections in nearby dunes or vegetated areas may distort the recovery algorithm.

GNSS-R is able to estimate CH with a relatively good agreement with the GFCH product. It is worth to mention that other approaches to retrieve CH, such as the Polarimetric SAR that will be used in ESA's BIOMASS mission [28], are targeting accuracies on the order of 30% of the CH for CH values CH larger than 10 m (i.e., 9 m for a CH of 30 m, 3 m for a CH of 10 m), similar to the performances shown by MIR.

ACKNOWLEDGMENT

The authors would like to thank the reviewers and editors for their grateful comments to improve this manuscript.

REFERENCES

[1] V. U. Zavorotny, S. Gleason, E. Cardellach, and A. Camps, "Tutorial on remote sensing using GNSS bistatic radar of opportunity," *IEEE Geosci. Remote Sens. Mag.*, vol. 2, no. 4, pp. 8–45, Dec. 2014.

[2] GCOS. *What are Essential Climate Variables?*. Accessed: Jan. 18, 2021. [Online]. Available: https://gcos.wmo.int/en/essential-climate-variables/abouth

[3] V. Senyurek, F. Lei, D. Boyd, M. Kurum, A. C. Gurbuz, and R. Moorhead, "Machine learning-based CYGNSS soil moisture estimates over ISMN sites in CONUS," *Remote Sens.*, vol. 12, no. 7, p. 1168, Apr. 2020, doi: 10.3390/rs12071168.

[4] C. Chew and E. Small, "Description of the UCAR/CU soil moisture product," *Remote Sens.*, vol. 12, no. 10, p. 1558, May 2020, doi: 10.3390/rs12101558.

[5] Q. Yan, W. Huang, S. Jin, and Y. Jia, "Pan-tropical soil moisture mapping based on a three-layer model from CYGNSS GNSS-R data," *Remote Sens. Environ.*, vol. 247, Sep. 2020, Art. no. 111944, doi: 10.1016/j.rse.2020.111944.

[6] H. Carreno-Luengo, G. Luzi, and M. Crosetto, "Above-ground biomass retrieval over tropical forests: A novel GNSS-R approach with CyGNSS," *Remote Sens.*, vol. 12, no. 9, p. 1368, Apr. 2020, doi: 10.3390/rs12091368.

[7] E. Santi *et al.*, "Remote sensing of forest biomass using GNSS reflectometry," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 2351–2368, 2020.

[8] A. Camps *et al.*, "Sensitivity of GNSS-R spaceborne observations to soil moisture and vegetation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 10, pp. 4730–4742, Oct. 2016.

[9] M. Zribi *et al.*, "Potential applications of GNSS-R observations over agricultural areas: Results from the GLORI airborne campaign," *Remote Sens.*, vol. 10, no. 8, p. 1245, Aug. 2018, doi: 10.3390/rs10081245.

[10] H. Carreno-Luengo, A. Amèzaga, D. Vidal, R. Olivé, J. Munoz, and A. Camps, "First polarimetric GNSS-R measurements from a stratospheric flight over boreal forests," *Remote Sens.*, vol. 7, no. 10, pp. 13120–13138, 2015.

[11] Y. Wang and Y. J. Morton, "Coherent GNSS reflection signal processing for high-precision and high-resolution spaceborne applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 831–842, Jan. 2021.

[12] J. F. Munoz-Martin *et al.*, "Untangling the incoherent and coherent scattering components in GNSS-R and novel applications," *Remote Sens.*, vol. 12, no. 7, p. 1208, Apr. 2020, doi: 10.3390/rs12071208.

[13] S. Gleason, A. O'Brien, A. Russel, M. M. Al-Khaldi, and J. T. Johnson, "Geolocation, calibration and surface resolution of CYGNSS GNSS-R land observations," *Remote Sens.*, vol. 12, no. 8, p. 1317, Apr. 2020, doi: 10.3390/rs12081317.

[14] P. Ferrazzoli, L. Guerriero, N. Pierdicca, and R. Rahmoune, "Forest biomass monitoring with GNSS-R: Theoretical simulations," *Adv. Space Res.*, vol. 47, no. 10, pp. 1823–1832, May 2011, doi: 10.1016/j.asr.2010.04.025.

[15] P. Liang, L. E. Pierce, and M. Moghaddam, "Radiative transfer model for microwave bistatic scattering from forest canopies," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 11, pp. 2470–2483, Nov. 2005, doi: 10.1109/TGRS.2005.853926.

[16] M. Kurum, R. H. Lang, P. E. O'Neill, A. T. Joseph, T. J. Jackson, and M. H. Cosh, "L-band radar estimation of forest attenuation for active/passive soil moisture inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 9, pp. 3026–3040, Sep. 2009, doi: 10.1109/TGRS.2009.2026641.

[17] R. Onrubia, D. Pascual, J. Querol, H. Park, and A. Camps, "The global navigation satellite systems reflectometry (GNSS-R) microwave interferometric reflectometer: Hardware, calibration, and validation experiments," *Sensors*, vol. 19, no. 5, p. 1019, Feb. 2019, doi: 10.3390/s19051019.

[18] D. Pascual, "Design and performance analysis of advanced GNSS-R instruments back-end," Ph.D. dissertation, Dept. Signal Theory Commun., Universitat Politècnica de Catalunya, Barcelona, Spain, 2020. [Online]. Available: https://www.tdx.cat/handle/10803/670632

[19] H. Park *et al.*, "Retracking considerations in spaceborne GNSS-R altimetry," *GPS Solutions*, vol. 16, no. 4, pp. 507–518, 2012, doi: 10.1007/s10291-011-0251-7.

[20] J. F. Munoz-Martin *et al.*, "Experimental evidence of swell signatures in airborne L5/E5a GNSS-reflectometry," *Remote Sens.*, vol. 12, no. 11, p. 1759, May 2020, doi: 10.3390/rs12111759.

[21] P. Potapov *et al.*, "Mapping global forest canopy height through integration of GEDI and landsat data," *Remote Sens. Environ.*, vol. 253, Feb. 2021, Art. no. 112165, doi: 10.1016/j.rse.2020.112165.

[22] A. G. Konings, M. Piles, N. Das, and D. Entekhabi, "L-band vegetation optical depth and effective scattering albedo estimation from SMAP," *Remote Sens. Environ.*, vol. 198, pp. 460–470, Sep. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425717302961

[23] H. Park, A. Camps, J. Castellvi, and J. Muro, "Generic performance simulator of spaceborne GNSS-reflectometer for land applications," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3179–3191, 2020.

[24] R. Bracewell, *The Fourier Transform and Its Applications*, 2nd ed. Tokyo, Japan: McGraw-Hill, 1978.

[25] W. J. Krzanowski, *Principles of Multivariate Analysis: A User's Perspective*. New York, NY, USA: Oxford Univ. Press, 1988.

[26] F. Burden and D. Winkler, "Bayesian regularization of neural networks," in *Methods in Molecular Biology*. Totowa, NJ, USA: Humana Press, 2008, pp. 23–42, doi: 10.1007/978-1-60327-101-1_3.

[27] X. Ying, "An overview of overfitting and its solutions," in *Proc. J. Phys., Conf.*, vol. 1168, Feb. 2019, Art. no. 022022, doi: 10.1088/1742-6596/1168/2/022022.

[28] S. Quegan *et al.*, "The European space agency BIOMASS mission: Measuring forest above-ground biomass from space," *Remote Sens. Environ.*, vol. 227, pp. 44–60, Jun. 2019, doi: 10.1016/j.rse.2019.03.032.