

Enhanced Mapping of Wildfire-Burned Areas using GNSS-Reflectometry Raw Intermediate Frequency Data

Archana Kannan^{*1}, Amer Melebari¹, Grigorios Tsagkatakis², Angel Farguell³, Kurtis Nelson⁴, Adam Kochanski³, Vinay Ravindra⁵, Sreeja Nag⁵, Chris Ruf⁶, and Mahta Moghaddam¹

¹University of Southern California, Los Angeles, CA, USA

²Foundation for Research and Technology - Hellas (FORTH), Institute of Computer Science, Heraklion, Greece

³San Jose State University, CA, USA and Utah State University, UT, USA

⁴EROS Center at the United States Geological Survey, Sioux Falls, SD, USA

⁵Bay Area Environmental Research Institute at NASA Ames Research Centre, Moffett Field, CA, USA

⁶University of Michigan, MI, USA

* kannana@usc.edu

Abstract—Mapping wildfire-burned areas with high spatiotemporal accuracy is vital for various operational, environmental, and meteorological applications. This paper provides a new methodology for mapping burned areas using the delay-Doppler maps (DDMs) from the NASA global navigation satellite system-reflectometry (GNSS-R) mission Cyclone Global Navigation Satellite System (CYGNSS), using a machine learning (ML) model. In addition to the nominal mode, we study the capabilities of including high-resolution DDMs, which have a lower incoherent integration time than the nominal mode and are derived from the raw intermediate frequency (raw IF) mode of data acquisition. We first analyze second-order textural properties such as dissimilarity, entropy, and correlation of the gray-level co-occurrence matrix (GLCM) calculated from nominal and raw IF DDMs. Then, we use the texture features of DDMs along with soil and terrain properties to make binary burned area classifications with various nominal-raw IF training splits. Textural analysis on raw IF DDMs revealed a strong discrimination between burned and not-burned pixels. By including the raw IF DDMs while training, we were able to achieve a 15% improvement in F1 score and Matthew's correlation, compared to training only with nominal mode DDMs. These results represent the first mapping of wildfire-burned areas using spaceborne GNSS-R raw IF data, and in particular, its combination with the use of higher-order statistics.

Index Terms—GNSS-Reflectometry, CYGNSS, Raw IF mode, Wildfires.

I. INTRODUCTION

In recent years, approximately one billion acres have been burned due to wildfires worldwide each year [1]. Wildfires tend to have significant long-term adverse effects on ecosystems and societies alike. There is a need for precisely monitoring these wildfire-burned areas to study the impacts of climate change and the resilience required for both wildland and urban environments.

Mapping wildfire-burned areas is a complicated task involving numerous environmental, meteorological, and geophysical parameters. Due to its scalability and non-contact observation scheme, remote sensing is one of the best approaches for mapping wildfire-burned areas. Many physics-based, statistical, and machine learning (ML)-based models have been explored for burned area mapping [2]. Many of these methods often provide coarse spatiotemporal resolution burned area

products. But most of the applications, such as monitoring loss of forest cover, fire and smoke emission simulations, and fire-spread modeling, require fine spatiotemporal resolution observations [3]–[5].

Within the microwave remote sensing domain, global navigation satellite system-reflectometry (GNSS-R) provides unique advantages for mapping burned areas. A GNSS-R receiver measures the surface-reflected global navigation satellite system (GNSS) microwave signals in the form of a delay-Doppler map (DDM). The Cyclone Global Navigation Satellite System (CYGNSS) mission, the NASA GNSS-R constellation of satellites, consists of eight satellites in low Earth orbit (LEO) with an inclination of 35° [6]. Each satellite can simultaneously measure four DDMs, which gives CYGNSS an opportunity for sub-daily high temporal resolution measurements. Wildfire burned area mapping has been implemented using peak reflectivity of DDM, specular point (SP) incidence angle, DDM signal-to-noise ratio (SNR), along with either vegetation properties [7] or soil properties [8] as inputs to ML-based models.

The nominal mode of CYGNSS, which is most commonly used for ocean and land surface applications [9], consists of 17×11 delay-Doppler bins and at 0.5 s incoherent integration time. CYGNSS also captures raw intermediate frequency (raw IF) [10] signal tracks over specific parts of Earth as required, which provides shorter incoherent integration times and a finer DDM resolution, consisting of 69×111 delay-Doppler bins. There are very limited studies carried out with raw IF mode with applications restricted to altimetry [11] and inland water detection [12]. In this work, we not only use raw IF data for mapping wildfire-burned areas for the first time, but we also investigate the use of second-order statistics, including textural information such as dissimilarity, entropy, and correlation of the gray-level co-occurrence matrix (GLCM) formed from nominal and raw IF DDMs. We then study the potential of using raw IF in combination with the nominal mode DDMs for mapping wildfire-burned areas using the extreme-gradient boosting (XGBoost) ML model [13].

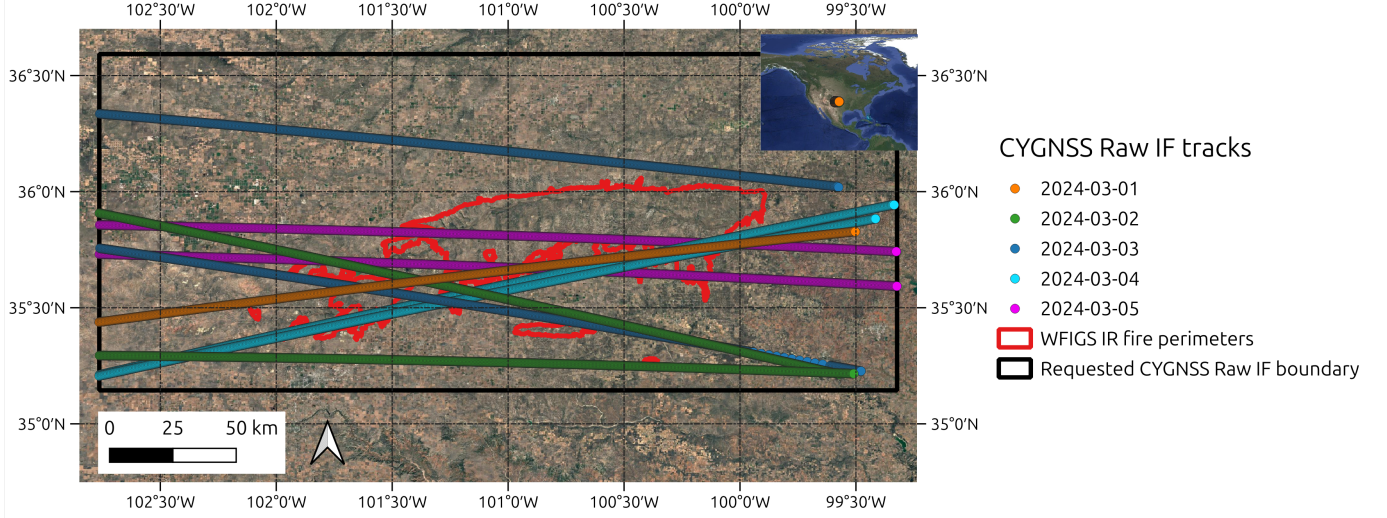


Fig. 1. Panhandle, Texas, USA study area with the CYGNSS raw IF tracks collected and 2020-2024 wildfire infrared perimeters from WFIGS.

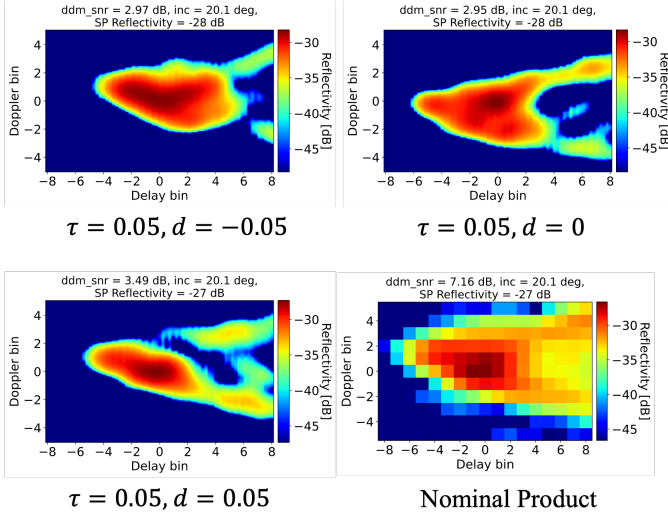


Fig. 2. Example of raw IF mode DDMs collected on March 1, 2024, at 0.05 seconds incoherent integration time at different time shifts from nominal mode. τ : Incoherent integration time in seconds. d : Timeshift from the nominal mode in seconds.

II. STUDY AREA AND MATERIALS

For studying the effects of raw IF of CYGNSS for mapping wildfire-burned areas, Panhandle, an area in northern Texas, was chosen. In February and March of 2024, this area had the largest wildfire in Texas history, burning over 960,000 acres.

As part of the distributed spacecraft with heuristic intelligence (D-SHIELD) to monitor wildfire spread for responsive control framework [5], raw IF tracks over the Texas Panhandle area were requested from the CYGNSS mission team. Fig.1 shows the nine raw IF tracks captured during March 1st to 5th of 2024. The calibrated raw IF data were processed at 0.05, 0.1, 0.25, and 0.5 s incoherent integration times. For this experiment, only the 0.05 s raw IF data were used. Fig.2 shows an example of nominal mode (17×11 bins) and 0.05 s

incoherent integration time raw IF (69×111 bins) DDMs at different time shifts from nominal mode.

The CYGNSS nominal mode data were also processed for the 2024 Panhandle fires, as well as 12 other smaller wildfires that happened within the chosen boundary from 2020 to 2023. Fig.1 shows the infrared perimeters of these fires.

From both nominal and raw IF DDMs, the peak reflectivity (Γ), SP incidence angle (θ_i), DDM SNR, 3×5 delay-Dopplers bins surrounding the SP, were chosen as inputs to the XGBoost model. Additionally, results of DDM texture analysis were taken as features to ML model. Section III-A gives more details on the texture analysis.

In addition to CYGNSS inputs, ancillary data layers such as soil moisture (SM), vegetation water content (VWC), surface temperature (LST), soil properties such as clay fraction (CF) and bulk density (BD), and elevation (E) were taken from Soil Moisture Active Passive (SMAP) [14] and Shuttle Radar Topography Mission (SRTM) [15]. For validation, a set-theoretic union of three different burned area products from MODIS [16], LANDSAT [17], and ESA CCI [18] burned area datasets was used. All the training and validation datasets were processed according to the framework provided in [8].

III. METHODS

A. Texture analysis on DDMs

In order to better investigate previously unexplored information within DDMs, the second-order statistics were calculated from the GLCM [19]. A GLCM contains the distribution of co-occurring pixel values at a given offset, specified with distance and angles. The distances and angles suitable for DDMs are estimated based on optimizing the point biserial correlation coefficient between textural feature and binary target class (burned or not burned). From the GLCM, dissimilarity, entropy, and correlation are extracted using (1), (2), and (3). In

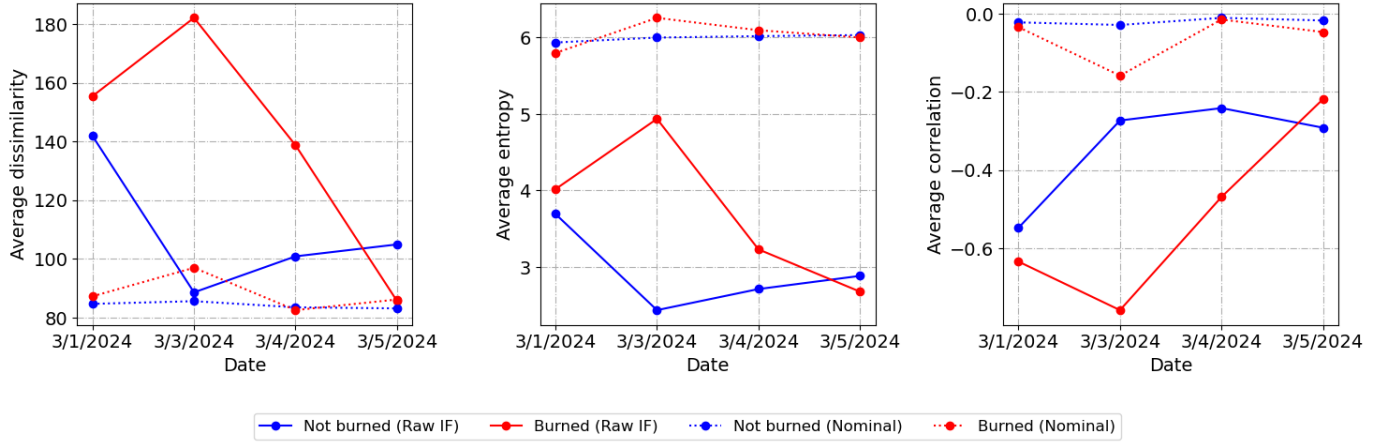


Fig. 3. Results of texture analysis on nominal and Raw IF modes of DDMs. Plots represent the average dissimilarity, entropy, and correlation extracted from GLCM, color differentiated based on burned vs not burned. Percentage of raw IF tracks burned on 3/1/2024: 12.31%, 3/3/2024: 4.3%, 3/4/2024: 6.33%, 3/5/2024: 14.6%.

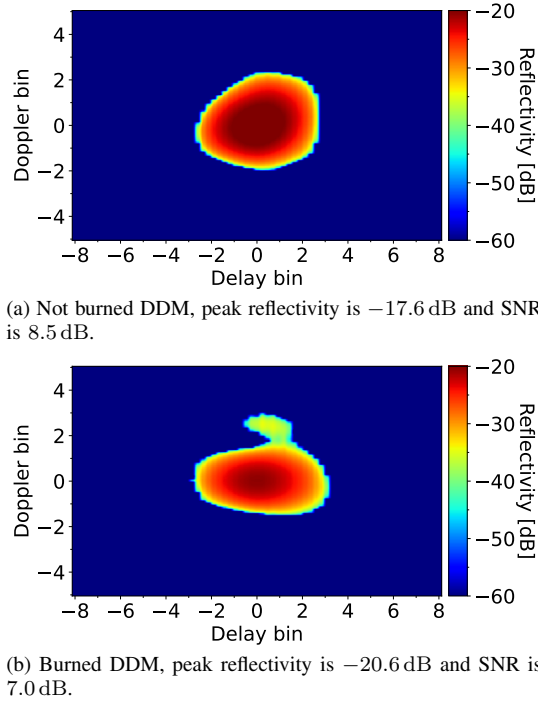


Fig. 4. An example of not burned vs burned raw IF DDM with high correlation in GLCM.

all of the equations, $p(i, j)$ indicates the frequency of i and j pixels occurring together, and N is the total number of pixels in GLCM.

$$\text{Dissimilarity} = \sum_{i=1}^N \sum_{j=1}^N |i - j| \cdot p(i, j) \quad (1)$$

$$\text{Entropy} = - \sum_{i=1}^N \sum_{j=1}^N p(i, j) \cdot \log_e(p(i, j)) \quad (2)$$

$$\text{Correlation} = \frac{1}{\sigma_i \sigma_j} \sum_{i=1}^N \sum_{j=1}^N (i - \mu_i)(j - \mu_j) \cdot p(i, j) \quad (3)$$

Where:

$$\mu_i = \sum_{i=1}^N \sum_{j=1}^N i \cdot p(i, j), \quad \mu_j = \sum_{i=1}^N \sum_{j=1}^N j \cdot p(i, j) \quad (4)$$

and σ_i, σ_j are the standard deviations of i and j respectively.

The top three of each textural analysis with high point biserial correlation coefficient were taken as training inputs to the ML model.

B. Classification model

The ML model, XGBoost, used in this paper, is trained using the preprocessed features and binary labels from validation datasets defined as

$$\mathcal{D} = \{(\mathbf{x}_i, y_i) \mid i = 1, 2, \dots, n\} \quad (5)$$

where $\mathbf{x}_i = [\theta_i, \text{SNR}, \Gamma_{\max}, \Gamma_{1, \dots, 15}, \text{dissimilarity}_{1,2,3}, \text{entropy}_{1,2,3}, \text{correlation}_{1,2,3}, \text{SM}, \text{VWC}, \text{LST}, \text{CF}, \text{BD}, E, \text{month}, \text{latitude}, \text{longitude}]$; $y_i \in \{0, 1\}$.

The XGBoost model is represented by the sum of decision trees, and the classification is based on the probabilities calculated using the sigmoid function. The model is trained by minimizing binary logistic loss. The model classifies the output based on a threshold of 0.5.

Due to the highly imbalanced nature of fire data, we partitioned the dataset into training (80%) and testing (20%) subsets, ensuring stratification based on burned labels. To test the efficiency of including raw IF, we added different percentages of raw IF to the nominal node during training. To avoid data leakage, we removed data covering fire for which raw IF was collected. Additionally, the nominal-raw IF percentage split was done based on stratifying the limited number of burned classes. The details of the different split scenarios are tabulated in Table I.

Different metrics were computed to assess the performance of the classification task. This includes accuracy, F1 score, precision, recall, confusion matrix, and Matthew's correlation coefficient (MCC). The definitions of these performance metrics are given in Fig.5 and Table I.

IV. RESULTS AND DISCUSSION

The average dissimilarity, entropy, and correlation from the GLCM of nominal and raw IF DDMs for the 4 days in March during which raw IF tracks were collected are shown in Fig.3. From the figure, we can see that for both modes, dissimilarity, entropy, and absolute correlation are higher for burned DDMs than not-burned DDMs for most cases. Also, the study site, both burned and non-burned areas, comprises grassland and cropland land cover. So, the contribution to the texture variations is mainly due to wildfire. In the dissimilarity, the variation in texture is much more prominent in raw IF mode compared to the nominal mode DDMs (28.78% vs. 4.75%). Similarly, we see lower entropy and a higher difference, 26.79%, between the burned and not-burned DDMs in raw IF mode. This compares to a higher entropy and lower 0.69% variation in nominal mode. Furthermore, we see a higher negative correlation in burned cases of raw IF DDMs. An example of burned vs. not burned raw IF DDM with high correlation is shown in Fig.4. Though a similar high-low trend is seen in the nominal mode, the correlations are closer to zero, making it difficult to distinguish just based on the correlation. From the second-order analysis, it is evident that using the texture of raw IF DDMs can significantly improve the burned area mapping.

The results of XGBoost binary classification with different nominal-raw IF training splits are shown in Fig.5 and Table I. We can see a 6% improvement in both F1 score and MCC when we introduce just 5% of raw IF during training. As we increase the percentage of raw IF in training, there is a slight drop in the F1 score after reaching 10 % of raw IF data in the training dataset. This could be attributed to the reduction of nominal mode data in the training. Overall, with the optimal split of nominal and raw IF during training, we were able to achieve 97.12% accuracy with 71.48% F1 score and 70.40% MCC. By including raw IF in training, we also saw a reduction in false negatives, indicated by improved precision.

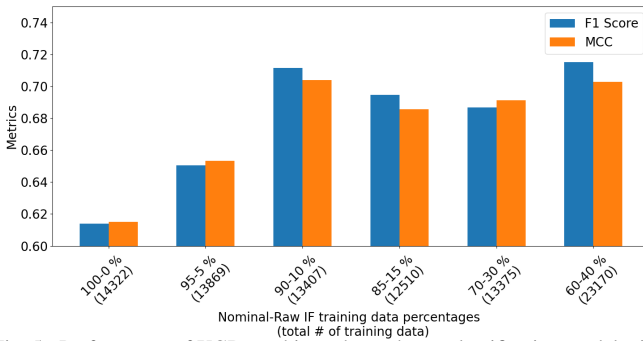


Fig. 5. Performance of XGBoost binary burned area classification model with different nominal-raw IF training splits.

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}; \text{MCC} = \frac{\text{TP} \cdot \text{TN} - \text{FP} \cdot \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}$$

V. CONCLUSION

This paper provides the first observations of wildfire-burned areas using spaceborne GNSS-R raw IF data. The second-order textural analysis, such as dissimilarity, entropy, and

TABLE I

PERFORMANCE METRICS OF XGBOOST BINARY CLASSIFICATION MODEL. THE FIRST COLUMN REPRESENTS THE NOMINAL-RAW IF SPLIT IN THE TRAINING DATASET.

Acc.: Accuracy, Pre.: Precision, Rec.: Recall, CM: Confusion matrix, TP: True positives, FP: False positives, TN: True negatives, FN: False negatives.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}; \text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}; \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}};$$

$$\text{CM} = \begin{bmatrix} \text{TP} & \text{FP} \\ \text{FN} & \text{TN} \end{bmatrix}$$

Split	# training	Acc.	Pre.	Rec.	CM	
100-0%	14322	96.45	80.15	49.75	3353 102	25 101
95-5%	13869	96.78	84.25	52.97	3359 95	20 107
90-10%	13407	97.12	82.46	62.56	3351 76	27 127
85-15%	12510	96.95	80	61.38	3348 78	31 124
70-30%	13375	97.09	88.87	56.15	3363 89	15 114
60-40%	23170	97.03	78.69	65.51	3342 70	36 133

correlation on GLCM from DDMs, shows that information from raw IF mode can aid in distinguishing burned areas with high temporal resolution and low latency. Supplementing the nominal mode data with raw IF in training the XGBoost ML model improved the classification by 15% and reduced burn misclassification in a highly imbalanced dataset. Such enhanced-accuracy spatiotemporal resolution can be a critical tool for providing timely burned-area maps to operational agencies in support of their disaster response efforts, as well as for post-fire damage assessment and resilient recovery.

REFERENCES

- [1] M. W. Jones, D. I. Kelley, C. A. Burton, F. Di Giuseppe, M. L. F. Barbosa, E. Brambleby, A. J. Hartley, A. Lombardi, G. Mataveli, J. R. McNorton *et al.*, "State of wildfires 2023–2024," *Earth System Science Data*, vol. 16, no. 8, pp. 3601–3685, 2024, doi:10.5194/essd-16-3601-2024.
- [2] E. Chuvieco, F. Mouillot, G. R. Van der Werf, J. San Miguel, M. Tanase, N. Koutsias, M. García, M. Yebra, M. Padilla, I. Gitas *et al.*, "Historical background and current developments for mapping burned area from satellite earth observation," *Remote Sensing of Environment*, vol. 225, pp. 45–64, 2019, doi:10.1016/j.rse.2019.02.013.
- [3] A. Khairoun, F. Mouillot, W. Chen, P. Ciais, and E. Chuvieco, "Coarse-resolution burned area datasets severely underestimate fire-related forest loss," *Science of the Total Environment*, vol. 920, p. 170599, 2024, doi:10.1016/j.scitotenv.2024.170599.
- [4] J. Mandel, B. Tumi, S. A. Taleghan, A. Farguelli, J. Haley, A. Kannan, A. Kochanski, C. Da, S. Nag, K. Hilburn *et al.*, "Enhanced fire detection and simulation using multi-satellite data assimilation and CYGNSS reflectometry in the coupled atmosphere-fire model wrf-sfire," *AGU24*, 2024.
- [5] S. Nag, V. Ravindra, R. Levinson, M. Moghaddam, K. Nelson, J. Mandel, A. Kochanski, A. F. Caus, A. Melebari, A. Kannan *et al.*, "Distributed spacecraft with heuristic intelligence to monitor wildfire spread for responsive control," in *IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2024, pp. 699–703, doi:10.1109/IGARSS53475.2024.10640833.
- [6] C. Ruf, S. Asharaf, R. Balasubramaniam, S. Gleason, T. Lang, D. McKague, D. Twigg, and D. Waliser, "In-Orbit Performance of the Constellation of CYGNSS Hurricane Satellites," *Bulletin of the American Meteorological Society*, vol. 100, no. 10, pp. 2009–2023, 10 2019, doi:10.1175/bams-d-18-0337.1.

- [7] E. Santi, M. Clarizia, D. Comite, L. Dente, L. Guerriero, and N. Pierdicca, "Detecting fire disturbances in forests by using GNSS reflectometry and machine learning: A case study in angola," *Remote Sensing of Environment*, vol. 270, p. 112878, 2022, doi:10.1016/j.rse.2021.112878.
- [8] A. Kannan, A. Melebari, G. Tsagkatakis, K. Nelson, V. Ravindra, S. Nag, and M. Moghaddam, "Mapping wildfire burned area using GNSS-reflectometry in densely vegetated regions with complex topography: A machine learning approach," in *IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2024, pp. 2072–2076, doi:10.1109/IGARSS53475.2024.10641487.
- [9] H. Carreno-Luengo, J. A. Crespo, R. Akbar, A. Bringer, A. Warnock, M. Morris, and C. Ruf, "The CYGNSS mission: On-going science team investigations," *Remote Sensing*, vol. 13, no. 9, p. 1814, 2021, doi:10.3390/rs13091814.
- [10] H. Carreno-Luengo, C. S. Ruf, S. Gleason, and A. Russel, "A new multi-resolution CYGNSS data product for fully and partially coherent scattering," *IEEE Trans. Geosci. Remote Sens.*, 2023, doi:10.1109/tgrs.2023.3318639.
- [11] W. Li, E. Cardellach, F. Fabra, S. Ribó, and A. Rius, "Assessment of spaceborne GNSS-R ocean altimetry performance using CYGNSS mission raw data," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 238–250, 2019, doi:10.1109/tgrs.2019.2936108.
- [12] Y. Zhang, Z. Yan, S. Yang, W. Meng, Y. Han, and Z. Hong, "Feasibility study on qinghai lake boundary detection using CYGNSS raw IF data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 2024, doi:10.1109/jstars.2024.3384235.
- [13] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794, doi:10.1145/2939672.2939785.
- [14] D. Entekhabi, E. G. Njoku, P. E. O'Neill, K. H. Kellogg, W. T. Crow, W. N. Edelstein, J. K. Entin, S. D. Goodman, T. J. Jackson, J. Johnson, J. Kimball, J. R. Piepmeier, R. D. Koster, N. Martin, K. C. McDonald, M. Moghaddam, S. Moran, R. Reichle, J. C. Shi, M. W. Spencer, S. W. Thurman, L. Tsang, and J. Van Zyl, "The Soil Moisture Active Passive (SMAP) Mission," *Proc. IEEE*, vol. 98, no. 5, pp. 704–716, 2010, doi:10.1109/JPROC.2010.2043918.
- [15] Earth Resources Observation and Science (EROS) Center, "Shuttle Radar Topography Mission 1 Arc-Second Global," 2018, doi:10.5066/F7PR7TFT.
- [16] L. Giglio, L. Boschetti, D. P. Roy, M. L. Humber, and C. O. Justice, "The collection 6 MODIS burned area mapping algorithm and product," *Remote sensing of environment*, vol. 217, pp. 72–85, 2018, doi:10.1016/j.rse.2018.08.005.
- [17] T. J. Hawbaker, M. K. Vanderhoof, G. L. Schmidt, Y.-J. Beal, J. J. Picotte, J. D. Takacs, J. T. Falgout, and J. L. Dwyer, "The LANDSAT burned area algorithm and products for the conterminous united states," *Remote Sensing of Environment*, vol. 244, p. 111801, 2020, doi:10.1016/j.rse.2020.111801.
- [18] E. Chuvieco, M. Pettinari, J. Lizundia-Loiola, T. Storm, M. Padilla Parrellada, E. F. C. C. Initiative *et al.*, "MODIS fire_cci burned area pixel product, version 5.1," *Centre for Environmental Data Analysis*, vol. 1, pp. 943–956, 2018.
- [19] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, no. 6, pp. 610–621, 1973, doi:10.1109/TSMC.1973.4309314.