IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-6032-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IGARSS53475.2024.10641487

MAPPING WILDFIRE BURNED AREA USING GNSS-REFLECTOMETRY IN DENSELY VEGETATED REGIONS WITH COMPLEX TOPOGRAPHY: A MACHINE LEARNING APPROACH

Archana Kannan [*,1], *Amer Melebari*[1], *Grigorios Tsagkatakis*[2], *Kurtis Nelson*[3], *Vinay Ravindra*[4]
Sreeja Nag[4], *and Mahta Moghaddam*[1]

[1]Ming Hsieh Department of Electrical and Computer Engineering,
University of Southern California, Los Angeles, CA, USA;
[2]Foundation for Research and Technology - Hellas (FORTH),
Institute of Computer Science, Heraklion, Greece;
[3]EROS Center at United States Geological Survey, Sioux Falls, SD, United States;
[4]Bay Area Environmental Research Institute at NASA Ames Research Centre,
Moffett Field, United States
[*]kannana@usc.edu

ABSTRACT

Accurate assessment of areas burned in wildfires is vital for various monitoring, management, and spread modeling applications. Wildfires, especially in forested regions, pose immense challenges for precise mapping due to the inherent dynamics of fuel types and terrain complexities. While remote sensing, particularly satellite imagery, offers an approach to studying burned areas, reliance on such satellite sources introduces challenges in characterizing burned areas amidst dense vegetation and environmental variations. This paper presents a mapping of forested burned areas utilizing global navigation satellite system–reflectometry (GNSS-R) from Cyclone Global Navigation Satellite System (CYGNSS) with ancillary observations from Soil Moisture Active Passive (SMAP) mission and Shuttle Radar Topography Mission (SRTM) using machine learning approaches. We validate the results with existing burned area products and provide maps of representative California fires within CYGNSS coverage. Assimilation of GNSS-R data into the model provides near real-time and high temporal resolution, enabling rapid response and mitigation efforts to fire events.

***Index Terms*—** Forest fires, GNSS-Reflectometry, Machine Learning, CYGNSS

1. INTRODUCTION

Wildfires are a complex and detrimental phenomenon with widespread impacts on ecosystems and human communities. Over recent decades, the scale and frequency of wildfires have exhibited a notable increase, posing unprecedented challenges [1]. Forest fires are particularly more pronounced in the western US, especially in the forests of California [2]. Various studies have been conducted to understand the dynamics of wildfires, given California's susceptibility to recurring fire events [2, 3, 4].

Mapping the wildfire-burned areas with low latency is necessary for fire progression monitoring and emission modeling. Accurate mapping of burned areas is a complex task involving a multitude of environmental and meteorological parameters. Leveraging a large number of observations from remotely sensed sources [5], numerous studies [6] have explored effective methodologies for mapping burned areas. The integration of machine learning techniques, driven by advancements in data-centric approaches [7], has emerged as a powerful tool for precise and automated mapping of burned areas from satellite data. Burned areas have been mapped using the Sentinel-1 and Sentinel-2 satellites [8, 9, 10, 11]. Landsat [12, 13] and MODIS [14] were also used to map burned areas using machine learning approaches. However, these approaches often have poor temporal resolution and high latency for wildfire progression applications.

The global navigation satellite system–reflectometry (GNSS-R) provides a unique measurement of the Earth's surface in the form of delay-Doppler maps (DDMs). Unlike optical measurements, GNSS-R can measure through clouds and other atmospheric obstructions to the visible light. As spaceborne GNSS-R systems are compact and economically viable, numerous GNSS-R receivers are in orbit, providing high temporal resolution compared to conventional synthetic aperture radar (SAR) systems. GNSS-R DDMs has been used in several land and ocean applications [15], including surface soil moisture estimation and wildfire detection agricultural regions [16].

In this paper, we study the effectiveness of utilizing

GNSS-R from Cyclone Global Navigation Satellite System (CYGNSS) for training random forests and XGBoost machine learning (ML) algorithms to map burned areas from California wildfires, with a specific focus on regions with dense vegetation and intricate topographical variations.

2. METHODOLOGY FOR BURNED AREA MAPPING

The machine learning models employed in this study are trained to utilize data layers sourced from CYGNSS [17] with ancillary data from Soil Moisture Active Passive (SMAP) [18] and Shuttle Radar Topography Mission (SRTM) [19] as covariates, with the objective of binary classification to determine burned or not-burned pixels. Given the different sources used for the training data, our analysis was bound to data temporally overlapping with CYGNSS observations. To ensure consistency in the measurements of the specular point (SP), only data within a $30\,\mathrm{m}$ radius was considered from ancillary data sources.

We utilized two different decision tree-based machine learning algorithms for mapping burned areas in California: Random Forests [20] and XGBoost [21]. Random forest is an ensemble technique combining multiple decision tree results with random feature selection to improve accuracy. XGBoost, a boosting algorithm, sequentially constructs a robust model by combining weak learners. It uses a gradient-boosting framework, minimizing errors by optimizing the loss function at each iteration. Both models provide inherent interpretability through feature importance metrics.

Due to the imbalanced nature of fire data spanning the years 2019 to 2022, we partitioned the dataset into training ($70\,\%$), validation ($10\,\%$), and test ($20\,\%$) subsets, ensuring stratification based on burned labels. The training phase involved hyperparameter tuning, feature selection, and cross-validation to optimize model performance and generalizability. The burned area binary classification results were validated using existing burned area products.

3. WILDFIRES AND DATA SOURCES

3.1. Study area

In this study, we investigate the incidences of wildfires in California from 2019 to 2022, focusing on regions with dense vegetation and high fuel content. Our analysis is limited to areas within the coverage range of CYGNSS (within 40-degree latitudes). The baseline fire perimeter data was obtained from the Cal Fire dataset [22]. The geographic distribution of these fires during the specified study period is illustrated in Fig. 1.

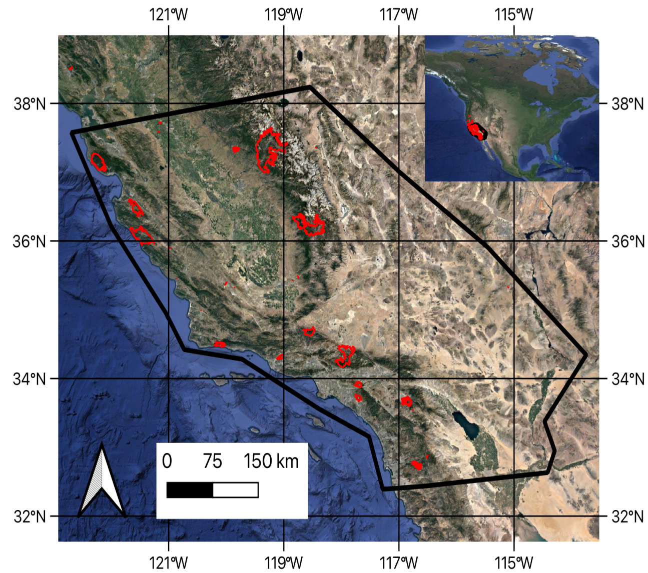


Fig. 1. Forest fires over California region within CYGNSS coverage for the years 2019 to 2022.

3.2. Materials

CYGNSS observations are the main inputs to the ML estimator. The reflectance changes due to a fire event. Hence, the DDM peak reflectivity was taken as the primary input. Also, the SP signal-to-noise ratio (SNR) and SP incidence angles were incorporated into our model to include interdependencies of the CYGNSS variables. Due to the correlation between fire occurrences and dryness factors, we integrate critical environmental parameters, specifically soil moisture (SM) and vegetation water content (VWC), into our analysis. Also, fires contribute to elevated surface temperature (TS) hence, the TS parameter was also integrated into our analysis. All these auxiliary layers were sourced from the SMAP mission. Considering the complexity of the terrain under examination, elevation data obtained from the SRTM serves as an additional layer to integrate spatial variations.

3.3. Validation data sources

A comprehensive approach was employed to validate the predicted burned area product, incorporating products derived from multiple satellite sources. We utilized three different burned area products, namely, the MODIS/Terra+Aqua Burned Area Monthly L3 Global $500\,\mathrm{m}$ SIN Grid [23], MODIS Fire_cci Burned Area Pixel (version 5.1) [24], and the Landsat Collection 2 Level-3 Burned Area [25] products, available at $500\,\mathrm{m}$, $250\,\mathrm{m}$, and $30\,\mathrm{m}$ spatial resolution, respectively. The MODIS and ESA CCI products are available as monthly composites, and the LANDSAT product is available bimonthly. Except ESA CCI product, which is available only until 2020, the datasets are current. Given the disparities between these sources, a set-theoretic union of these products was used for burned area assessment.

4. THE BURNED AREA PRODUCT

To assess the binary classification performance, accuracy and F1 score were considered. Accuracy provides an overall measure of correct predictions. Since the data was imbalanced, the F1 score was a better metric for measuring the effectiveness of classification. These metrics [26] are computed using (1) and (2), where TP and TN are the true positives and true negatives, FP and FN are the false positives and false negatives.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$F1score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{2}$$

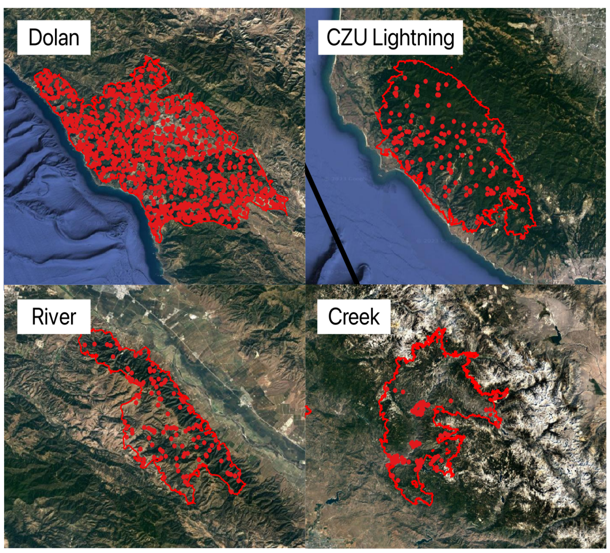


Fig. 2. Burned area mapped using CYGNSS and XGBoost for largest fires in 2020. Red dots indicate burning predicted by ML model and the red boundary is the burn perimeter from the Cal Fire dataset.

Fig. 2 illustrates examples of burned area mapping based on CYGNSS SP. The performance metrics of the training and testing datasets and a detailed assessment of the top four significant fires, Dolan, CZU Lightning, River, and Creek complex fires, are presented in Table 1 and Table 2. The confusion matrices for these four fires are visually depicted in Fig. 3. The Random Forest and XGBoost algorithms demonstrate good performance in large fire regions, with XGBoost demonstrating slightly better results for smaller fires.

Table 1. Performance metrics of all fires (2019 – 2022) in California with coverage from CYGNSS. *Acc.: Accuracy*.

	Random forest		**XGBoost**	
	Acc.	**F1 score**	**Acc**	**F1 score**
Train	100	1	100	1
Test	97.1590	0.9804	**98.1060**	**0.9869**

Table 2. Performance metrics of largest fires in California in 2020 with coverage from CYGNSS. *Acc.: Accuracy*.

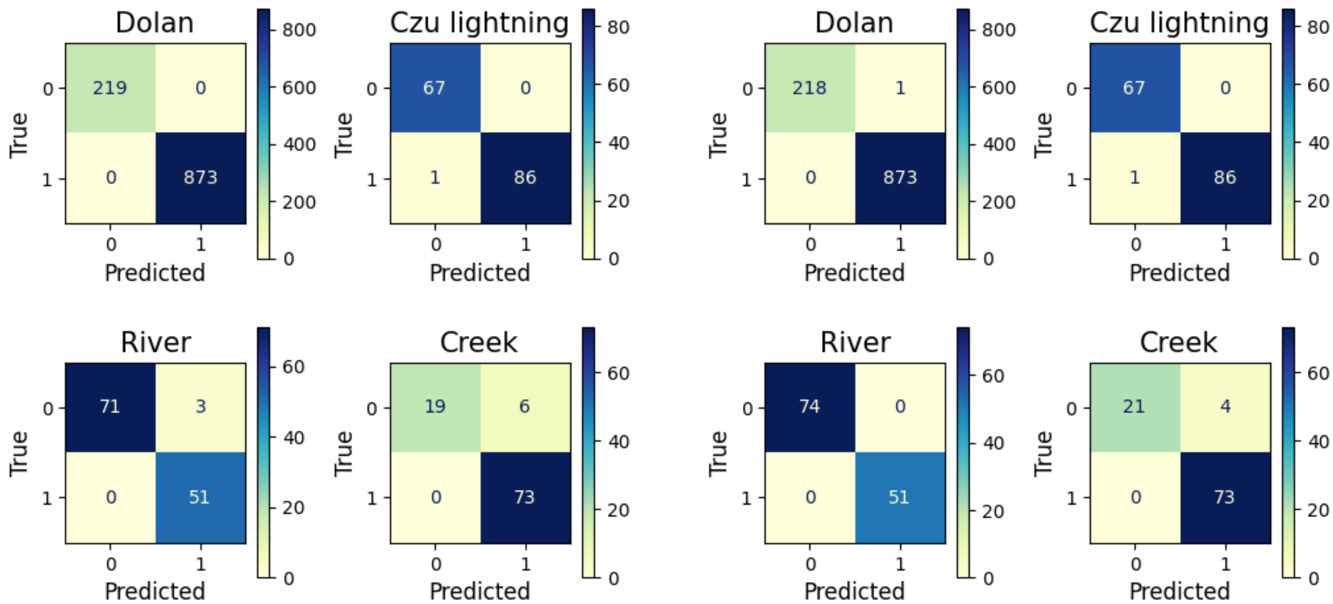
Fires	**Random forest**		**XGBoost**	
	Acc.	**F1 score**	**Acc**	**F1 score**
Dolan	100	1	99.9084	0.9994
CZU Lightning	99.3506	0.9942	99.3506	0.9942
River	97.6	0.9714	100	1
Creek	93.8775	0.9605	95.9183	0.9733

Moreover, this study includes a confidence level analysis for predictions generated from the XGBoost model using logistic regression. Fig. 4 illustrates an example, showcasing the calculated confidence level for the Blue Ridge, Lake, and Alisal fires. This supplementary analysis adds a layer of insight into the spatiotemporal uncertainty of the model's classifications.

5. DISCUSSION

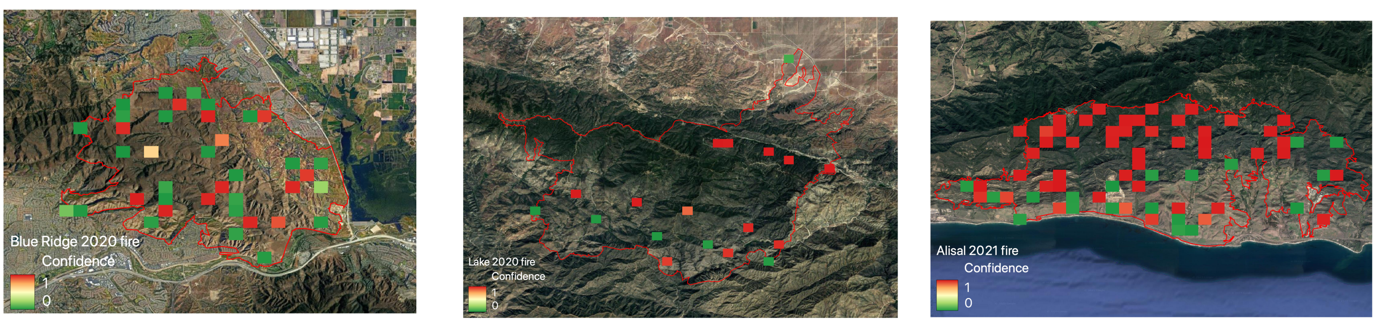
Comparing the performance of the two models (Fig. 3), XGBoost was marginally better than Random Forests. This enhanced performance can be attributed to the intrinsic strengths of XGBoost, specifically its ability to handle complex relationships within the data through a sequential and adaptive boosting process. Notably, in the context of small fires, XGBoost exhibits significantly better performance compared to random forests. This advantage can be attributed to the XGBoost's capacity to mitigate overfitting and enhance predictive accuracy, especially in imbalanced datasets. Additionally, the improved performance for small fires could be influenced by the relatively limited number of training samples, wherein regularization mechanisms of XGBoost contribute to more effective learning from sparse data, further substantiating its suitability for scenarios with reduced training instances.

Analyzing the confusion matrices of both models, it can be seen that, during the fire period, the unburned pixels were misclassified as burned (FP) pixels. During fire events, various environmental factors may lead to variations resembling burned conditions, causing misclassifications. The sensitivity of the models to these variations during fire periods also contributes to an increased likelihood of false positives.



(a) Random forests (b) XGBoost

Fig. 3. Confusion matrices of the largest four fires in California in 2020 with coverage from CYGNSS using two ML models.



(a) Blue Ridge fire (b) Lake fire (c) Alisal fire

Fig. 4. Confidence levels of three different burned areas. The red boundary is the fire perimeter from Cal Fire. Colored squares represent the confidence level of the pixel being burned, with red indicating high confidence and green indicating low confidence.

The uncertainty analysis exhibited different confidence levels for various fires, as demonstrated by a few examples showcased in Fig. 4. Fire regions undergoing simultaneous vegetation and topographical changes showed decreased confidence in burned area classifications. This trend was seen in the Blue Ridge and Lake fires that occurred in 2020. Also, the high frequency of fire training pixels in densely vegetated areas biased the ML model toward categorizing fires in such environments. The uncertainty in the results can also be attributed to complex topography, which is present in most areas in this study, posing a challenge, as GNSS-R observations are sensitive to surface roughness at multiple scales [27]. Another pattern was seen in fires closer to coastal areas. The pixels near the coastlines exhibited lower confidence in being classified as burned. This trend could be attributed to the proximity of CYGNSS SP to water bodies, introducing large reflectivity variations.

6. CONCLUSION

This paper provides the first reported results for mapping burned areas in densely vegetated regions with complex topography using observations from CYGNSS and SMAP. Our initial results for the chosen study area show an accuracy and F1 score of 98% and 0.98, respectively, using the XGBoost algorithm trained with an imbalanced dataset. Incorporating GNSS-R into the binary classification model provides high temporal resolution and low latency burned area products. Future work will explore the usage of entire DDM in addition to peak reflectivity to include burn severity predictions.

7. REFERENCES

[1] Virginia Iglesias, Jennifer K Balch, and William R Travis, "Us fires became larger, more frequent, and more widespread in the 2000s," *Science advances*, vol. 8, no. 11, pp. eabc0020, 2022.

[2] Anthony LeRoy Westerling, "Increasing western us forest wildfire activity: sensitivity to changes in the timing of spring," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 371, no. 1696, pp. 20150178, 2016.

[3] Patrick T Brown, Holt Hanley, Ankur Mahesh, Colorado Reed, Scott J Strenfel, Steven J Davis, Adam K Kochanski, and Craig B Clements, "Climate warming increases extreme daily wildfire growth risk in california," *Nature*, vol. 621, no. 7980, pp. 760–766, 2023.

[4] Francis K Dwomoh, Roger F Auch, Jesslyn F Brown, and Heather J Tollerud, "Trends in tree cover change over three decades related to interannual climate variability and wildfire in california," *Environmental Research Letters*, vol. 18, no. 2, pp. 024007, 2023.

[5] Emilio Chuvieco and Russell G Congalton, "Application of remote sensing and geographic information systems to forest fire hazard mapping," *Remote sensing of Environment*, vol. 29, no. 2, pp. 147–159, 1989.

[6] Emilio Chuvieco, Florent Mouillot, Guido R Van der Werf, Jesús San Miguel, Mihai Tanase, Nikos Koutsias, Mariano García, Marta Yebra, Marc Padilla, Ioannis 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.

[7] M Arif, KK Alghamdi, SA Sahel, SO Alosaimi, ME Alsahaft, MA Alharthi, and M Arif, "Role of machine learning algorithms in forest fire management: A literature review," *J. Robot. Autom*, vol. 5, pp. 212–226, 2021.

[8] Miguel A Belenguer-Plomer, Mihai A Tanase, Angel Fernandez-Carrillo, and Emilio Chuvieco, "Burned area detection and mapping using sentinel-1 backscatter coefficient and thermal anomalies," *Remote Sensing of Environment*, vol. 233, pp. 111345, 2019.

[9] Daniela Stroppiana, Gloria Bordogna, Matteo Sali, Mirco Boschetti, Giovanna Sona, and Pietro Alessandro Brivio, "A fully automatic, interpretable and adaptive machine learning approach to map burned area from remote sensing," *ISPRS International Journal of Geo-Information*, vol. 10, no. 8, pp. 546, 2021.

[10] Sanjiwana Arjasakusuma, Sandiaga Swahyu Kusuma, Yenni Vetrita, Indah Prasasti, and Rahmat Arief, "Monthly burned-area mapping using multi-sensor integration of sentinel-1 and sentinel-2 and machine learning: Case study of 2019's fire events in south sumatra province, indonesia," *Remote Sensing Applications: Society and Environment*, vol. 27, pp. 100790, 2022.

[11] Changhui Lee, Seonyoung Park, Taeheon Kim, Sicong Liu, Mohd Nadzri Md Reba, Jaehong Oh, and Youkyung Han, "Machine learning-based forest burned area detection with various input variables: A case study of south korea," *Applied Sciences*, vol. 12, no. 19, pp. 10077, 2022.

[12] Resul Çömert, Dilek Küçük Matci, and Uğur Avdan, "Object based burned area mapping with random forest algorithm," *International Journal of Engineering and Geosciences*, vol. 4, no. 2, pp. 78–87, 2019.

[13] Somnath Bar, Bikash Ranjan Parida, and Arvind Chandra Pandey, "Landsat-8 and sentinel-2 based forest fire burn area mapping using machine learning algorithms on gee cloud platform over uttarakhand, western himalaya," *Remote Sensing Applications: Society and Environment*, vol. 18, pp. 100324, 2020.

[14] Varun Mithal, Guruprasad Nayak, Ankush Khandelwal, Vipin Kumar, Ramakrishna Nemani, and Nikunj C Oza, "Mapping burned areas in tropical forests using a novel machine learning framework," *Remote Sensing*, vol. 10, no. 1, pp. 69, 2018.

[15] Nereida Rodriguez-Alvarez, Joan Francesc Munoz-Martin, and Mary Morris, "Latest Advances in the Global Navigation Satellite System—Reflectometry (GNSS-R) Field," *Remote Sensing 2023, Vol. 15, Page 2157*, vol. 15, no. 8, pp. 2157, 4 2023.

[16] E Santi, MP 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, pp. 112878, 2022.

[17] Christopher S. Ruf, Scott Gleason, Zorana Jelenak, Stephen Katzberg, Aaron Ridley, Randall Rose, John Scherrer, and Valery Zavorotny, "The CYGNSS Nanosatellite Constellation Hurricane Mission," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2012, pp. 214–216.

[18] Dara Entekhabi, Eni G. Njoku, Peggy E. O'Neill, Kent H. Kellogg, Wade T. Crow, Wendy N. Edelstein, Jared K. Entin, Shawn D. Goodman, Thomas J. Jackson, Joel Johnson, John Kimball, Jeffrey R. Piepmeier, Randal D. Koster, Neil Martin, Kyle C. McDonald, Mahta Moghaddam, Susan Moran, Rolf Reichle, J. C. Shi, Michael W. Spencer, Samuel W. Thurman, Leung Tsang, and Jakob Van Zyl, "The Soil Moisture Active Passive (SMAP) Mission," *Proceedings of the IEEE*, vol. 98, no. 5, pp. 704–716, 2010.

[19] Earth Resources Observation and Science (EROS) Center, "Shuttle Radar Topography Mission 1 Arc-Second Global," 2018.

[20] Leo Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[21] Tianqi Chen and Carlos 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.

[22] California Department of Forestry and Fire Protection, "Cal fire wildfire perimeters and prescribed burns," 2023.

[23] Luis Giglio, C Justice, L Boschetti, and D Roy, "Modis/terra+ aqua burned area monthly l3 global 500m sin grid v061," *NASA EOSDIS Land Processes DAAC: Washington, DC, USA*, 2021.

[24] E Chuvieco, ML Pettinari, J Lizundia-Loiola, T Storm, M Padilla Parellada, ESA Fire Climate Change Initiative, et al., "Modis fire_cci burned area pixel product, version 5.1," *Centre for Environmental Data Analysis*, vol. 1, pp. 943–956, 2018.

[25] Todd J Hawbaker, Melanie K Vanderhoof, Gail L Schmidt, Yen-Ju Beal, Joshua J Picotte, Joshua D Takacs, Jeff T Falgout, and John L Dwyer, "The landsat burned area algorithm and products for the conterminous united states," *Remote Sensing of Environment*, vol. 244, pp. 111801, 2020.

[26] Mohammad Hossin and Md Nasir Sulaiman, "A review on evaluation metrics for data classification evaluations," *International journal of data mining & knowledge management process*, vol. 5, no. 2, pp. 1, 2015.

[27] Amer Melebari, James D. Campbell, Erik Hodges, and Mahta Moghaddam, "Improved Geometric Optics with Topography (IGOT) Model for GNSS-R Delay-Doppler Maps Using Three-Scale Surface Roughness," *Remote Sensing*, vol. 15, no. 7, pp. 1880, 3 2023.