

Verification and Validation of SmokeViz: An Ensemble-Based Deep Learning Model for Identifying Wildfire Smoke Plumes in Satellite Imagery

001
002
003

Anonymous WPO NOFO submission

004

Paper ID 17349

1. Problem Statement

In part due to public policy, the average levels of fine particulate matter ($PM_{2.5}$) in the US have generally been declining over the past few decades [1]. Despite those improvements, the contribution of wildfire smoke to $PM_{2.5}$ concentrations in the US has been calculated to have more than doubled between 2010 to 2020, accounting for up to half of the overall $PM_{2.5}$ exposure in Western regions [2]. Increases in $PM_{2.5}$ due to wildfire smoke are concerning since ambient $PM_{2.5}$ exposure is a leading environmental risk factor for adverse health effects and premature mortality [3]. These risks underscore the necessity for efficient and effective monitoring methods to mitigate the adverse health impacts associated with wildfire smoke.

Satellite imagery, equipped with state-of-the-art sensors, such as the Advanced Baseline Imager (ABI) on the Geostationary Operational Environmental Satellites (GOES) [4], have revolutionized environmental monitoring. Compared to orbiting satellites such as the Suomi or Sentinel satellites, geostationary satellites maintain constant observation over a fixed area. GOES offers the advantage being able to reliably and consistently capture the dynamic behavior of wildfire smoke plumes. In turn, GOES capabilities can provide critical insights into the concentration and movement of smoke particulates, facilitating real-time assessments of air quality.

Current fire detection algorithms rely primarily on calculating the fire radiative power (FRP) of active fires using near-infrared (NIR) satellite bands [5]. When fires have thick smoke plumes, the smoke can obfuscate a satellite's view of an active fire, making FRP impossible to calculate. If combined with FRP calculations, a real-time wildfire smoke plume analysis product can deliver additional insight into an active wildfire's activity level.

Generally, integrating satellite imagery into wildfire smoke monitoring provides real-time data that can improve the timeliness of public health planning and response. By mapping the spread and density of smoke, health authorities can issue prompt warnings, implement evacuation protocols, and deploy resources effectively to mitigate health risks. Furthermore, long-term data gathered from satellite observations can aid in understanding the broader impacts of wildfire smoke on public health, influencing policy decisions and preventive measures.

To address the need for a timely smoke analysis product, we present SmokeViz, an ensemble based deep learning method to identify and classify smoke plumes in satellite imagery. The model is a result

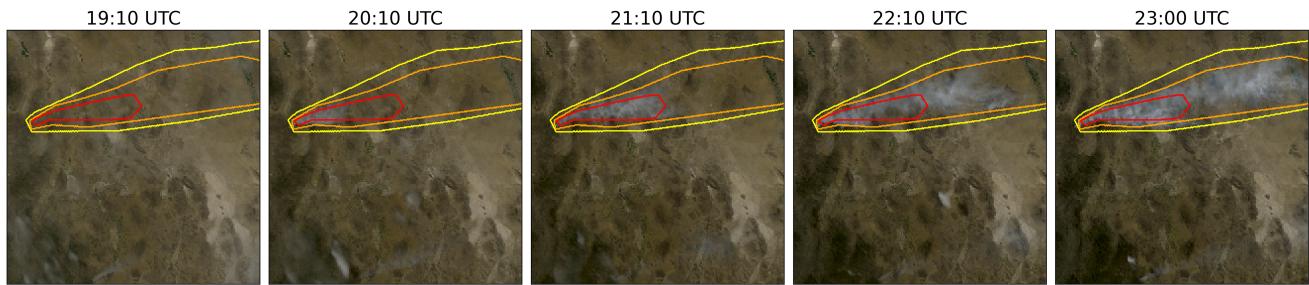


Figure 1. True color GOES-East imagery from May 5th, 2022, Southeast New Mexico (31.38°N , 107.87°W) during the start of the Foster Fire. The red, orange and yellow lines represent the heavy, medium and low density HMS smoke annotations that span 19:10–23:00 UTC.

of a successful project at the Global Systems Laboratory at the National Oceanic and Atmospheric Administration (NOAA) to build a large scale dataset of smoke plumes in GOES imagery. Keeping in mind the correlation between both the quality and quantity of data with deep learning model performance, we introduced the largest known smoke plume dataset, containing over 180,000 samples. The resulting dataset was used to train an ensemble of deep learning models that decorrelate errors by varying the inductive biases and give a more robust prediction than any singular model.

We propose bringing SmokeViz to NOAA's Fire Weather Testbed to evaluate it's capabilities to (1) replace or enhance NOAA's current human annotated smoke product and (2) improve latency in NOAA's smoke dispersion models by using it as a smoke analysis product. The following subsections elaborate on the current needs for SmokeViz in these two application areas.

044 1.1. Replacing the Hazard Mapping Systems Smoke Product

045 The National Oceanic and Atmospheric Administration (NOAA) manages environmental sattelite
046 programs such as the Hazard Mapping System (HMS) program that uses an aggregation of satellite data to
047 generate active fire and smoke data [6, 7]. The HMS smoke product [6, 7] is generated by human visual
048 inspection of GOES satellite imagery for smoke identification. Trained satellite analysts use movement
049 characteristics to help identify smoke by scanning through a time series of satellite images. When visual
050 inspection indicates smoke, the analyst will draw a polygon that corresponds to the geolocation and den-
051 sity of smoke. HMS smoke analysis data gives the coordinates of the smoke perimeter as a polygon and
052 classifies the smoke by density within a given time window. The time windows can range from instant-
053 aneous (same start/end time) to lengths of 22 hours, with an average time window around 3 hours. While
054 the true bounds of the smoke can change within the larger time spans (figure 1), the analyst is making
055 an approximation that should reflect the smoke coverage over the duration of the time window. The
056 density information is qualitatively determined by each analyst based on the apparent smoke opacity in
057 the satellite imagery and categorized as either light, medium or heavy as seen in figure 2a [8]. The HMS
058 smoke product is released on a rolling but undefined schedule ranging from one to multiple times a day
059 as observation conditions permit and is heavily limited by the availability of analysts and their time.

060 Due the primarily its coarse time resolution and latency between the event and the corresponding an-
061 notation, the HMS smoke product is not an ideal candidate for a smoke analysis product for real-time
062 smoke dispersion forecasting. Current end users of the HMS smoke product include the Environmen-
063 tal Protection Agency's AirNow program, NOAA's hurricane center and Mexico's National Weather
064 Service. SmokeViz offers the current capabilities with additional advancements to the HMS product,

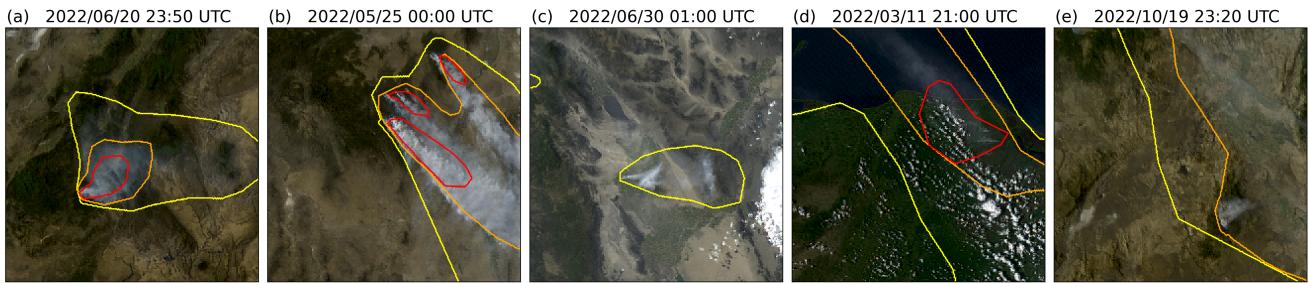


Figure 2. HMS smoke annotations overlaid on GOES imagery, the yellow, orange and red contours indicate the extent of light, medium and heavy density smoke. (a) and (b) show a canonical examples of a smoke plumes. (c)-(e) show observable variations in the density labels.

some of the immediate items of improvement include higher time resolution and real-time analysis on streaming GOES data.

Additionally, through the fire weather testbed, we aim to verify that SmokeViz is able to provide more consistency in smoke density labels, better detail of smoke borders and detect smoke plumes missed by analysts.

1.2. Introducing a Smoke Analysis Product for Data Assimilation

The High-Resolution Rapid Refresh (HRRR) model and prototype Rapid Refresh Forecast System (RRFS) produce short-range weather forecasts for CONUS and Alaska, including predictions of wild-fire smoke. For each hour, the model is initialized using meteorological data assimilation, combining weather observations with a 3-D model background. While these models use satellite-detected FRP to determine location and smoke emissions associated with wildfires, there is currently no assimilation of smoke observations. We propose to use a deep learning derived smoke identification product to create a smoke analysis to initialize HRRR or RRFS.

Currently, HRRR and RRFS have a significant latency for representing the plumes from recently started wildfires figure 3. HRRR relies on polar-orbiting satellites, which only pass over a point on Earth's surface twice per day. RRFS uses detections from geostationary satellites, which can detect fires with much lower latency, but depends on the Regional ABI and VIIRS fire Emissions (RAVE) algorithm which takes some time to run. In addition, formal data assimilation of either satellite aerosol optical depth (AOD) or surface-based smoke observations takes considerable computational time and resources. For these reasons, a deep learning based smoke analysis could significantly improve short-range model forecasts in rapidly evolving situations featuring new fire starts.

2. Methods and Activities

While we will work directly with the fire weather testbed to develop a thorough and fair proccess for SmokeViz evaluation, we provide a summary of key aspects we will address.

$$IoU_{\text{overall}} = \frac{\sum_{\substack{i=\text{light} \\ i=\text{heavy}}} |y_i \cap y_i^*|}{\sum_{\substack{i=\text{heavy} \\ i=\text{light}}} |y_i| \cup |y_i^*|} \quad (1)$$

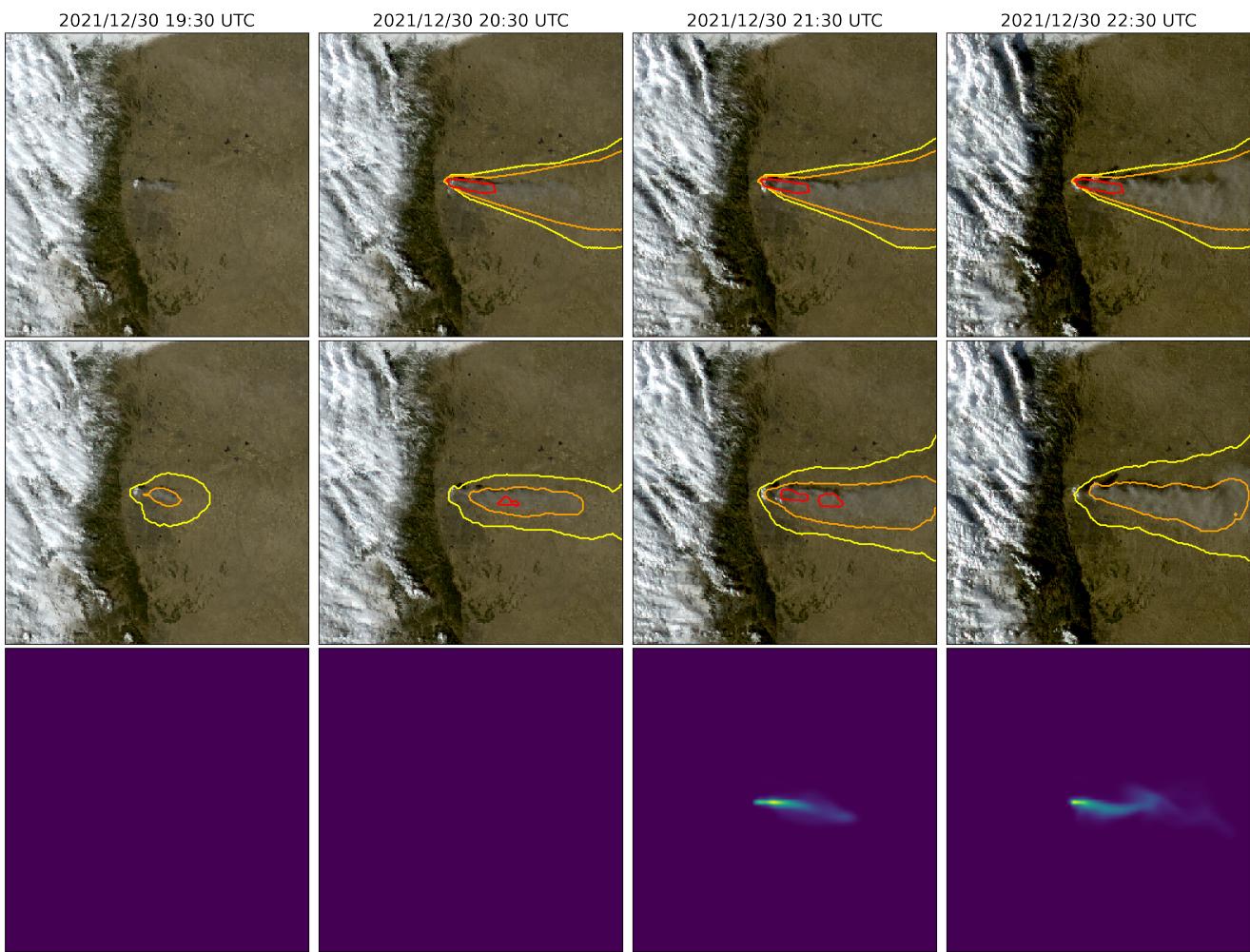


Figure 3. Top row shows the HMS smoke annotations, middle is SmokeViz and bottom row is HRRR smoke for the start of the Marshall fire in Boulder, CO on December 29th, 2021.

- 090 1. **Identify all the relevant end users.** We have the HMS analysts that would directly use the product
091 and HRRR/RRFS smoke dispersion model developers, but would want to determine if including
092 further down the line users would be relevant to the evaluation. The HMS smoke product and HRRR
093 forecast is publically published online so that it is difficult to track all users, but we will develop a
094 strategy for identifying and communicating with users.
095 2. **Establish the needs for all end users.** While we have the needs established for those end users that
096 are collaborators on this proposal, we need to perform an assessment on current further down the line
097 end users of the HMS smoke product.
098 3. **Survey how to improve training material and documentation.** We have a tutorial available on how
099 to use the SmokeViz model on specified GOES imagery, but would want to do further assessments
100 on how we can make it more user-friendly to not only the scientific community but also the general
101 public. This would be an iterative process that would start before the official testbed evaluation
102 event. Success of the training materials will be included as part of the testbed evaluation.
103 4. **Determine evaluation metrics.** In the SmokeViz paper that is currently in the review process, we use

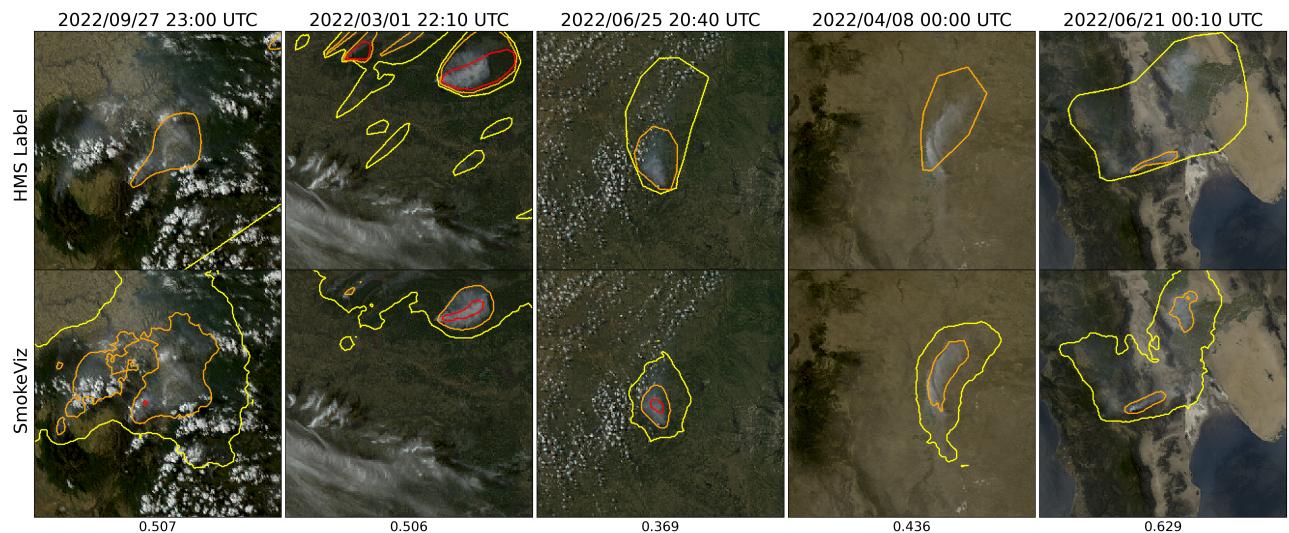


Figure 4. Examples of HMS annotations (top row) vs SmokeViz model output (bottom row) on SmokeViz dataset samples. The overall IoU score is reported at the bottom of each column.

intersection over union (IoU) to quantify the overlap between SmokeViz and HMS analyst annotations (equation 1 and figure 4). The results reported in that paper gives an IoU score of 0.62 between the two datasets. IoU is a commonly used evaluation metric for object semantic segmentation but can have nuances, especially when your ground truth (HMS annotations) may have inconsistencies. Having a lower IoU value between the SmokeViz output and HMS annotations does not necessarily mean that SmokeViz performed poorly.

We anticipate that the addition of SmokeViz as a smoke analysis product into the HRRR smoke dispersion forecasting will be more straight forward to evaluate. While HRRR smoke has not been formally validated, there has been a case study performed on how well HRRR smoke forecasting aligned with air quality ground measurements for the 2018 Camp Fire [9]. Through a similar analysis, we can quantify if there is an improvement with how well the HRRR smoke forecast performs when compared to PM_{2.5} measurements once the SmokeViz analysis product is introduced.

5. **Develop evaluation scenarios.** The SmokeViz model was trained on smoke plumes from 2018-2021 and 2024. 2023 data served to validate the model and 2022 was used for final performance testing. Since neither of those years were used for training the model they could be included for testbed evaluation but it would be important to add in events from 2025 that the model has never touched before. We would need to include scenarios from varying solar zenith angles during the day and over different seasons of the year. Since the SmokeViz model domain covers all of North America, there should be scenarios that span multiple geolocations within those boundaries. While currently untested, we anticipate SmokeViz to have similar performance in the Southern hemisphere and would include samples to test that hypothesis.
6. **Perform evaluation.** We will hold the evaluation in Boulder, Colorado at NOAA Fire Weather Testbed. Participants will include HMS representatives and proposal collaborators from College Park, Maryland and the HRRR Smoke developers and collaborators local to the Boulder area along with any additional relevant virtual participants identified. We anticipate that the evaluation will take place over two days where the focus of the first day will be on training how to use SmokeViz and

130 starting the HMS smoke product evaluation. The second day be focused on further HMS smoke
131 product comparisons and the HRRR Smoke dispersion modeling evaluation.
132

7. Implement SmokeViz evaluation feedback.

133 3. Products/Outputs

134 The project will produce an operationally-ready smoke analysis product that can be used for near-
135 realtime DA for air quality models. Additionally, a comprehensive report on the operational utility
136 of the SmokeViz product will be delivered in order to support operational deployment.

137 Potential end-user adopters of the SmokeViz product would include the scientists that create and run
138 the air quality and smoke plume dispersion models along with down-the-line air quality forecasters and
139 decision makers.

140 4. Impacts/Benefits/Outcomes

141 The expected result will be an efficient deep learning based smoke analysis product that describes the
142 current extent and categorical density of smoke. Used as a smoke DA product, SmokeViz is expected to
143 significantly advance short-range model predictive capabilities for rapidly evolving wildfires.

144 5. Schedule

Schedule		
Date	Start RL	milestone
Q1	5	
Aland Islands	AX	ALA

145 The current RL for this project is RL 5 where we have performed evaluation and testing for SmokeViz
146 on streaming GOES imagery. At the end of this project timeline, we will work with a transition partner
147 to reach readiness level 8.
148

149 6. Outreach and Education

150 We've developed a series of Jupyter Lab notebooks that walk users through the SmokeViz dataset hosted
151 by NOAA on AWS that includes HMS analyst annotations and the corresponding GOES imagery. The
152 notebooks demonstrate how the user can interact with the curated dataset to train their own deep learning
153 models and how to apply the tested SmokeViz model to incoming GOES imagery in order to detect active
154 smoke plumes.

155 The SmokeViz dataset and model serve as an excellent benchmark for learning how to work with
156 large environmental datasets and deep learning models. We've put a lot of thought and effort into the
157 reproducibility of the research along with readability of the code and tools. Not only can users use the
158 codebase for smoke plume related projects but also as a toolkit for other remote sensing work involving
159 satellite imagery and geospatial shape files.

160 This proposal is focused on getting SmokeViz operational in the United States for detecting and
161 monitoring wildfires in North America. We will reach out to facilitate virtual participation from Mexico's
162 National Weather Service for the SmokeViz evaluation. Verification of SmokeViz for North America
163 will give insights on developing additional deep learning wildfire smoke algorithms for other parts of
164 the world that may not have the required resources to have analysts constantly monitor smoke plumes.

7. Diversity and Inclusion

The original SmokeViz dataset and model project began during the 2022 NOAA/CIRES summer internship program. This program was a CIRES diversity and inclusion initiative to give underrepresented graduate students opportunities to work with NOAA scientists. As a member of multiple underrepresented groups in the scientific community, PI Rey Koki was able to develop collaborations for SmokeViz directly through that internship opportunity. Since the end of that internship, Rey has since been hired as a CIRES research associate and continues work with SmokeViz and related projects at NOAA. Rey Koki identifies as queer, is half Pacific Islander and is the only member in their direct or extended family to hold a 4-year college degree, yet alone a graduate degree.

Having a direct understanding of how fostering equity in science benefits the scientific community as a whole, Rey Koki is committed to continued efforts towards equity-centered approaches in their work. Related to this project, Rey Koki and collaborator Jebb Stewart developed an internship project with SmokeViz, where they mentored a summer 2024 Hollings undergraduate intern, Annabel Wade, who successfully implemented ensembling methods for SmokeViz. Rey Koki and Jebb Stewart continue to make efforts to facilitate research opportunities for students to gain experience working with SmokeViz.

8. Data and/or Software Management Plan

References

- [1] J. E. Aldy, M. Auffhammer, M. Cropper, A. Fraas, and R. Morgenstern, “Looking back at 50 years of the clean air act,” *Journal of Economic Literature*, vol. 60, no. 1, pp. 179–232, 2022. 1
- [2] M. Burke, A. Driscoll, S. Heft-Neal, J. Xue, J. Burney, and M. Wara, “The changing risk and burden of wildfire in the united states,” *Proceedings of the National Academy of Sciences*, vol. 118, no. 2, p. e2011048118, 2021. 1
- [3] E. Gakidou, A. Afshin, A. A. Abajobir, K. H. Abate, C. Abbafati, K. M. Abbas, F. Abd-Allah, A. M. Abdulle, S. F. Abera, V. Aboyans, *et al.*, “Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: a systematic analysis for the global burden of disease study 2016,” *The Lancet*, vol. 390, no. 10100, pp. 1345–1422, 2017. 1
- [4] S. J. Goodman, T. J. Schmit, J. Daniels, and R. J. Redmon, *The GOES-R series: a new generation of geostationary environmental satellites*. Elsevier, 2019. 1
- [5] W. Schroeder, P. Oliva, L. Giglio, and I. A. Csiszar, “The new viirs 375 m active fire detection data product: Algorithm description and initial assessment,” *Remote Sensing of Environment*, vol. 143, pp. 85–96, 2014. 1
- [6] D. McNamara, G. Stephens, M. Ruminski, and T. Kasheta, “The hazard mapping system (hms) - noaa’s multi-sensor fire and smoke detection program using environmental satellites,” *Conference on Satellite Meteorology and Oceanography*, 01 2004. 2
- [7] W. Schroeder, M. Ruminski, I. Csiszar, L. Giglio, E. Prins, C. Schmidt, and J. Morisette, “Validation analyses of an operational fire monitoring product: The hazard mapping system,” *International Journal of Remote Sensing*, vol. 29, no. 20, pp. 6059–6066, 2008. 2
- [8] NOAA, “Hazard mapping system fire and smoke product.” 2
- [9] F. K. Chow, K. A. Yu, A. Young, E. James, G. A. Grell, I. Csiszar, M. Tsidulko, S. Freitas, G. Pereira, L. Giglio, *et al.*, “High-resolution smoke forecasting for the 2018 camp fire in california,” *Bulletin of the American Meteorological Society*, vol. 103, no. 6, pp. E1531–E1552, 2022. 5