

# **Fire Weather Testbed Evaluation of SmokeViz: A Deep Learning Wildfire Smoke Plume Analysis Product for Data Assimilation**

FY2025 Weather Program Office Research Programs Announcement - Testbeds

**NOAA-OAR-WPO-2025-28601**

**Fire Weather Testbed**

**Lead PI:** Rey Koki, Associate Scientist in Machine Learning  
Cooperative Institute for Research in Environmental  
Sciences (CIRES), NOAA GSL affiliate  
phone: 303-478-9911, email: [rey.koki@noaa.gov](mailto:rey.koki@noaa.gov)

**Collaborator:** Jebb Q. Stewart, Weather Visualization and AI  
Research Development (WIZARD) Branch Chief, NOAA GSL  
phone: 307-215-3917, email: [jebb.q.stewart@noaa.gov](mailto:jebb.q.stewart@noaa.gov)

**Collaborator:** Eric James, Senior Associate Scientist, NOAA GSL  
phone: 720-340-1167, email: [eric.james@noaa.gov](mailto:eric.james@noaa.gov)

**Collaborator:** Wilfrid Schroeder, Physical Scientist, NOAA NESDIS  
phone: 202-341-7763, email: [wilfrid.schroeder@noaa.gov](mailto:wilfrid.schroeder@noaa.gov)

## Abstract

### Fire Weather Testbed Evaluation of SmokeViz: A Deep Learning Wildfire Smoke Plume Analysis Product for Data Assimilation

**Project Goal:** In order to address the need to advance predictive capabilities for wildfire smoke dispersion and air quality, we present an innovative and computationally efficient smoke analysis product, SmokeViz.

**Problem/Opportunity Statement:** Currently, wildfire smoke plume models such as HRRR and RRFS have a significant latency for representing the plumes from recently started wildfires. This project addresses the operational need to address model latency for wildfire smoke by providing a timely and reliable smoke observation product for data assimilation into air quality prediction models.

**Methodology and Timeline:** Over the proposed two year timeline, we will validate SmokeViz's capabilities to identify smoke over North America. Then we will integrate and evaluate SmokeViz as input to HRRR-Smoke as a data assimilation product within an operational context.

**Primary Project Products/Outputs:** The project will produce an operationally-ready smoke analysis product that can be used for near-realtime data assimilation for air quality models.

**Expected Results, Outcomes, and Benefits:** The expected result will be a deep learning based smoke analysis product that describes the current extent and categorical density of smoke. Used as a smoke data assimilation product, SmokeViz is expected to significantly advance short-range model predictive capabilities for rapidly evolving wildfires.

**Intended Beneficiaries and Recipients:** Expected end-users SmokeViz would include HMS analysts and scientists that create and run the air quality and smoke plume dispersion models along with down-the-line air quality forecasters and decision makers.

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

## 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 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.

### 1.1. Replacing the Hazard Mapping Systems Smoke Product

The National Oceanic and Atmospheric Administration (NOAA) manages environmental sattelite programs such as the Hazard Mapping System (HMS) program that uses an aggregation of satellite data to generate active fire and smoke data [6, 7]. The HMS smoke product [6, 7] is generated by human visual

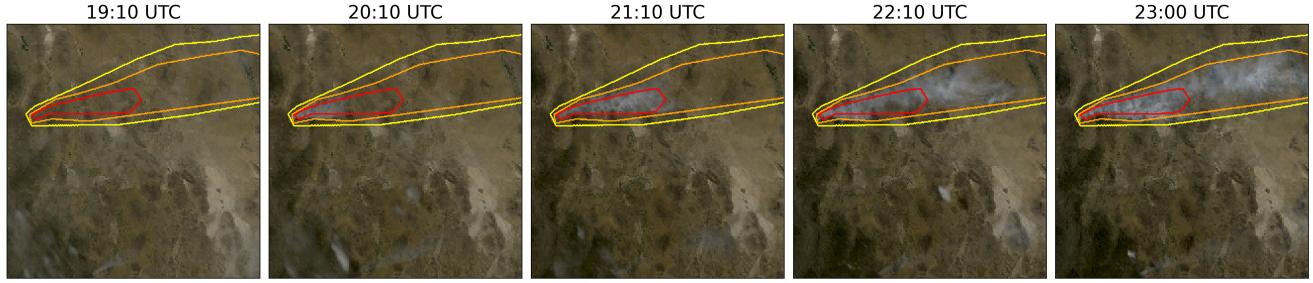


Figure 1. True color GOES-East imagery from May 5th, 2022, Southeast New Mexico ( $31.38^{\circ}\text{N}$ ,  $107.87^{\circ}\text{W}$ ) during the start of the Foster Fire. The red, orange and yellow lines represent the heavy, medium and low density HMS smoke annotations that span 19:10–23:00 UTC.

inspection of GOES satellite imagery for smoke identification. Trained satellite analysts use movement characteristics to help identify smoke by scanning through a time series of satellite images. When visual inspection indicates smoke, the analyst will draw a polygon that corresponds to the geolocation and density of smoke. HMS smoke analysis data gives the coordinates of the smoke perimeter as a polygon and classifies the smoke by density within a given time window. The time windows can range from instantaneous (same start/end time) to lengths of 22 hours, with an average time window around 3 hours. While the true bounds of the smoke can change within the larger time spans (figure 1), the analyst is making an approximation that should reflect the smoke coverage over the duration of the time window. The density information is qualitatively determined by each analyst based on the apparent smoke opacity in the satellite imagery and categorized as either light, medium or heavy as seen in figure 2a [8]. The HMS smoke product is released on a rolling but undefined schedule ranging from one to multiple times a day as observation conditions permit and is heavily limited by the availability of analysts and their time.

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

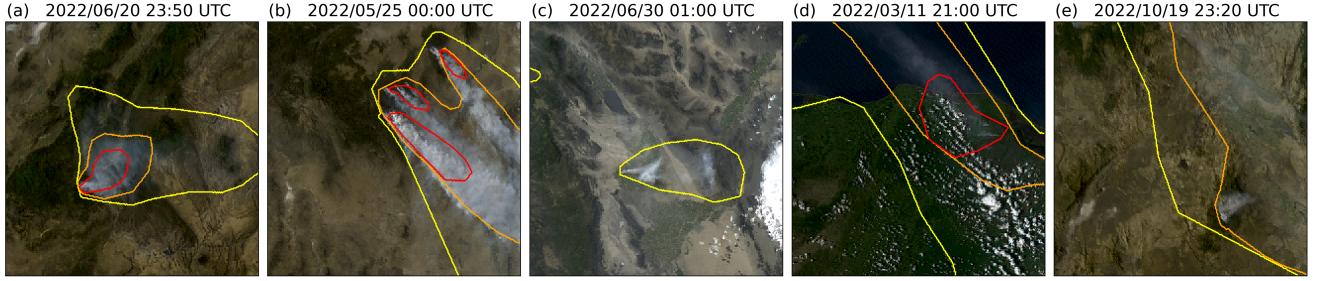


Figure 2. HMS smoke annotations overlaid on GOES imagery, the yellow, orange and red contours indicate the extent of light, medium and heavy density smoke. (a) and (b) show a canonical examples of a smoke plumes. (c)-(e) show observable variations in the density labels.

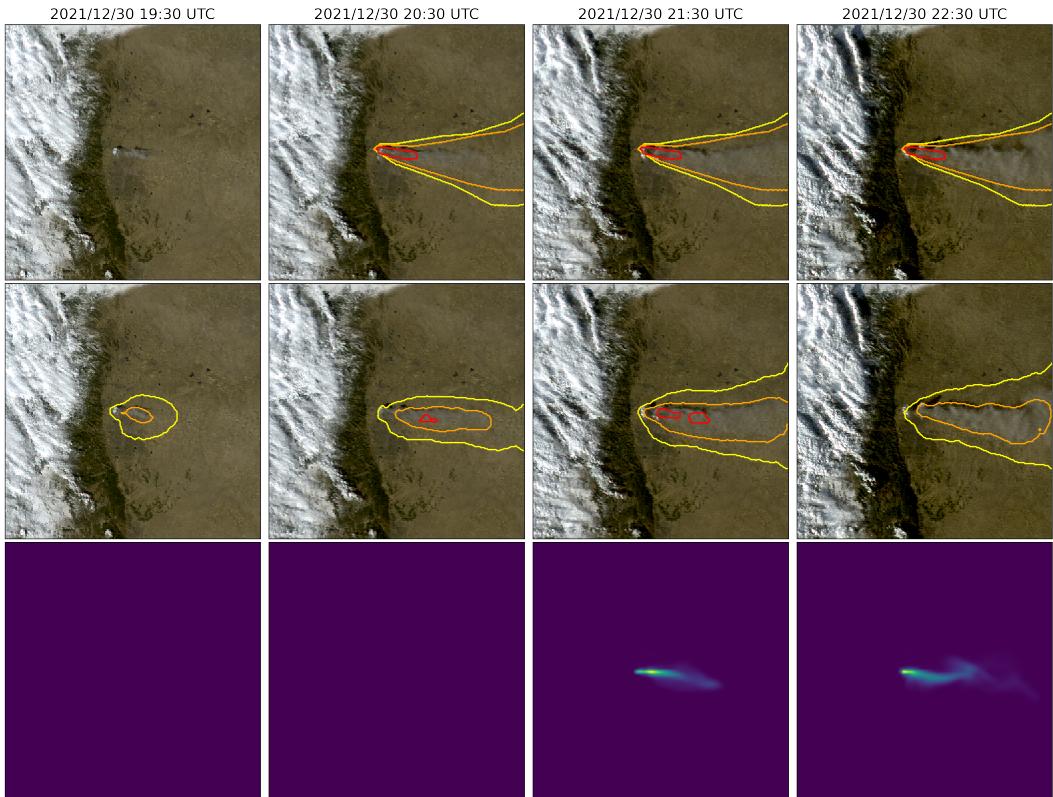


Figure 3. Top row shows the HMS smoke annotations, middle is SmokeViz and bottom row is HRRR smoke for the start of the Marshall fire in Boulder, CO.

started wildfires figure 4. 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.

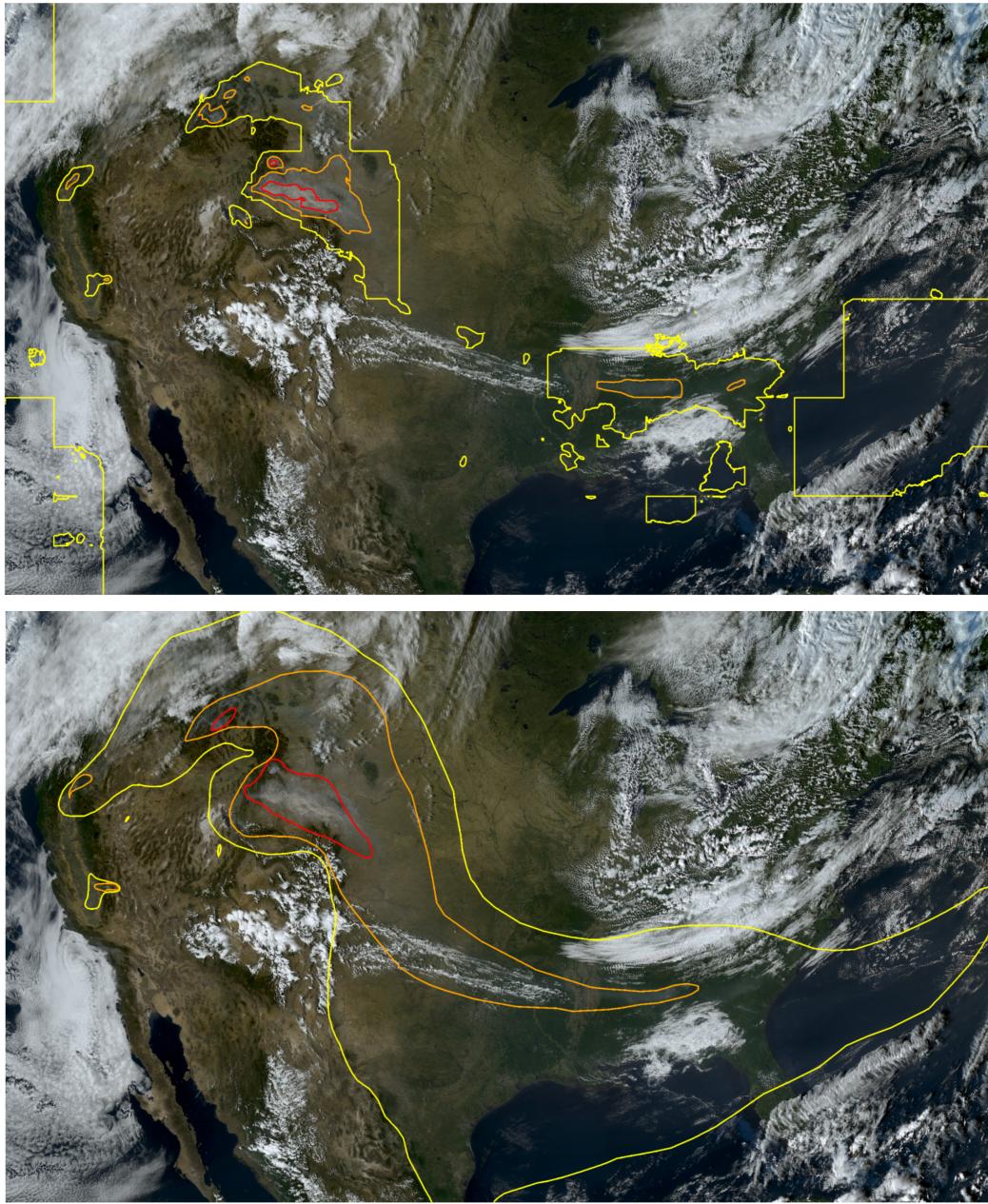


Figure 4. Top row shows the SmokeViz output on CONUS, bottom show HMS analyst annotations for October 15th, 2024 21:00.

## 2. Methods and Activities

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

- 1. Identify all the relevant end users.** We have identified the HMS analysts that would directly use the product and HRRR/RRFS smoke dispersion model developers as end-users but would want to determine if including further down the line users would be relevant to the evaluation. Further potential

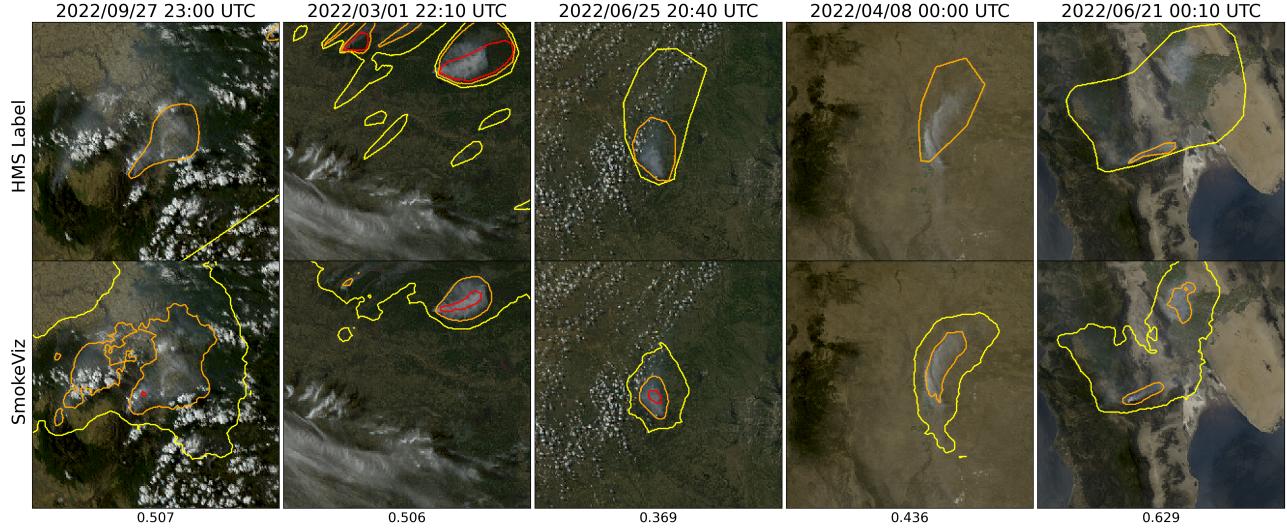


Figure 5. Examples of HMS annotations (top row) vs SmokeViz model output (bottom row) on SmokeViz dataset samples. The overall IoU score is reported at the bottom of each column.

end-user adopters of the SmokeViz product would include scientists that would want a smoke plume data assimilation product for air quality models along with down-the-line air quality forecasters and decision makers. The HMS smoke product and HRRR forecast output are publically available online so that it is difficult to track all users, but we will develop a strategy for identifying and communicating with users.

2. **Establish the needs for current/potential end users.** While SmokeViz offers some advantages to the HMS smoke analyst product, we need to ensure that we are not missing any use cases if we propose to replace/modify the HMS smoke product. For instance, if a down-the-line user needs an approximation of smoke existing within a polygon over a large time window, we can facilitate that need, but would need to first identify it. For further advancements of using SmokeViz for data assimilation uses, we need to establish how we can best fit the requirements for temporal and spatial resolution.
3. **Survey how to improve training material and documentation.** We have a tutorial available on how to use the SmokeViz model on specified GOES imagery, but would want to do further assessments on how we can make it more user-friendly to not only the scientific community but also the general public. This would be an iterative process that would start before the official testbed evaluation event. Success of the training materials will be included as part of the testbed evaluation.
4. **Integrate SmokeViz as a smoke analysis product for HRRR Smoke.** SmokeViz was developed with the goal of creating a real-time smoke analysis product on streaming satellite imagery with comparable results to the HMS smoke product. Due to design structure emulating the HMS annotations, no additional steps are needed for SmokeViz to be in a comparable format to the HMS smoke product. On the other hand, integrating SmokeViz into the HRRR Smoke forecasting model will require some minimal development. The framework for data assimilation into HRRR Smoke is already established but some development will be required to ensure effective integration of SmokeViz as a smoke analysis product.
5. **Determine evaluation metrics.** In the SmokeViz paper that is currently in the review process, we use

intersection over union (IoU) to quantify the overlap between SmokeViz and HMS analyst annotations (equation 1 and figure 5). 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.

$$IoU = \frac{|A \cap B|}{|A| \cup |B|} \quad (1)$$

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<sub>2.5</sub> measurements once the SmokeViz analysis product is introduced.

6. **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.
7. **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 starting the HMS smoke product evaluation. The second day be focused on further HMS smoke product comparisons and the HRRR Smoke dispersion modeling evaluation.
8. **Obtain formal testbed evaluation.** We will coordinate with the Fire Weather Testbed to obtain a formal testbed evaluation after the event has occurred.
9. **Implement SmokeViz evaluation feedback.** After obtaining feedback from the testbed evaluation, we will strategize on how to address any concerns and implement the feedback into the SmokeViz product.
10. **Demonstrate SmokeViz outside of testbed scenarios.** Once the evaluation feedback has been addressed, we will demonstrate the updated SmokeViz in operations at the HMS analyst operations and for the HRRR Smoke forecasting model and obtain any further feedback.
11. **Deliver final SmokeViz product.** After the last stage of feedback, we will demonstrate the final updated version of SmokeViz in operations and deliver the finalized documentation for the transition plan to operation.

### **3. Products/Outputs**

The primary product this project will produce is an operationally-ready smoke analysis product that can be used for near-realtime data assimilation for smoke dispersion and air quality models. Along with the smoke analysis product itself, a comprehensive report on the operational utility of the SmokeViz product will be delivered in order to support operational deployment.

### **4. Impacts/Benefits/Outcomes**

The expected result will be an efficient deep learning based smoke analysis product that describes the current extent and categorical density of smoke. Used as a smoke data assimilation product, SmokeViz is expected to significantly advance short-range model predictive capabilities for rapidly evolving wildfires. Additionally, we anticipate SmokeViz to reduce HMS analyst smoke plume related workload so that they can instead focus on other areas of interest for satellite environmental observations.

The scope of this proposal is focused on getting SmokeViz operational specifically for the United States to detect and monitor wildfires in North America. But verification of SmokeViz for North America will give insights on further deep learning wildfire smoke algorithms for other parts of the world that may not have the required resources for analysts to continuously monitor smoke plumes.

### **5. Schedule**

The current RL for this project is RL 5 where we have performed evaluation and testing for SmokeViz on streaming GOES imagery. At the end of this project timeline, we will work with a transition partner to reach readiness level 8. The timeline for the project spans from August 2025 through July 2027.

Schedule		
RL	date	deliverables
5	now	SmokeViz product has been evaluated on streaming GOES imagery. Prototype for HMS analyst usage has been developed for Continental US.
5	Q4 2025	Identify all relevant end users and establish current and potential future needs. Survey how to improve training material and documentation.
5	Q1-Q2 2026	Iterative process for improving training material and documentation. HRRR Smoke analysis product integration. Create evaluation plan.
6	Q3 2026	Determine evaluation scenarios and perform evaluation within testbed.
6	Q4 2026	After SmokeViz evaluation is delivered by testbed, integrate feedback.
7	Q1 2027	Demonstrate SmokeViz in operations at HMS and for HRRR Smoke forecasting. Identify any final changes to be made.
7	Q2 2027	Implement changes identified during operations demonstration.
8	Q3 2027	Demonstrate the final version of SmokeViz at HMS and for HRRR Smoke forecasting and finalize documentation for transition.

### **6. Outreach and Education**

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

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

## 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

The dataset for SmokeViz is already maintained and available through NOAA AWS at <https://noaa-gsl-experimental-pds.s3.amazonaws.com/index.html#SmokeViz/>. Along with the model itself, the user-friendly Jupyter notebook series for working with the SmokeViz dataset and building deep learning models is available on github [https://github.com/reykoki/Smoke\\_Semantic\\_Segmentation\\_Learning\\_Notebook](https://github.com/reykoki/Smoke_Semantic_Segmentation_Learning_Notebook). Running the SmokeViz model takes low computational power and can easily be run on the entire CONUS on standard laptop CPUs in a matter of seconds. Therefore, we do not anticipate the need to store all SmokeViz outputs for public access since it is easily generated with the publically available model.

## 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. [3](#)
- [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. [3](#)
- [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. [3](#)
- [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. [3](#)

- [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. [3](#)
- [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. [3](#)
- [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. [3](#)
- [8] NOAA, “Hazard mapping system fire and smoke product.” [4](#)
- [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. [8](#)