131 Hartwell Avenue
Lexington, Massachusetts
02421-3126
USA
Tel: +1 781 761-2288
Fax: +1 781 761-2299
www.aer.com

FINAL REPORT

Applying Machine Learning Techniques to Track Smoke Plumes: Phase II

TCEQ Contract No. 582-19-90498
Work Order No. 582-22-31966-013
Revision 2.0

Prepared by:
Archana Dayalu, Rick Pernak, and Matthew Alvarado
Atmospheric and Environmental Research, Inc. (AER)
131 Hartwell Ave.
Lexington, MA 02421-3126

Correspondence to: malvarad@aer.com

Prepared for:
Erik Gribbin
Texas Commission on Environmental Quality
Air Quality Division
Building E, Room 342S
Austin, TX 78711-3087

June 30, 2022

Document Change Record

Revision	Revision Date	Remarks
0.1	14 Jun 2022	Internal Version for Review
1.0	15 Jun 2022	Draft Submitted to the TCEQ
2.0	30 Jun 2022	Final Version Submitted to the TCEQ

TABLE OF CONTENTS

List of Figures

List of Tables

List of Acronyms

AER – Atmospheric and Environmental Research
AOD – Aerosol Optical Depth
BC2 – Black and Brown Carbon campaign
CO – Carbon monoxide
CNN – Convolutional Neural Network
DNN – Deep Neural Network
EASD – Early Adopter Synthetic Data
GOES – Geostationary Operational Environmental Satellite
HCHO – Formaldehyde
HGB – Houston-Galveston-Brazoria region
HMS – Hazard Mapping System Fire and Smoke Product
ML – Machine Learning
NASA – National Aeronautics and Space Administration
NOAA – National Oceanic and Atmospheric Administration
NO_2 – Nitrogen dioxide
NOAA – National Oceanic and Atmospheric Administration
PCA – Principal Component Analysis
PC – Principal Component
QAPP – Quality Assurance Project Plan
TCEQ – Texas Commission on Environmental Quality
TEMPO – Tropospheric Emissions: Monitoring of Pollution instrument
UV – Ultraviolet
UVAI – Ultraviolet Aerosol Index

Executive Summary

Detecting the transport of smoke plumes from the initial fire location to populated areas usually requires a human analyst to track the extent and transport of the smoke. However, recent advances in artificial intelligence and computer vision would allow this analysis to be performed automatically. In ML techniques, DNNs use multiple hidden layers and hyperparameters to classify complex data. DNNs consist of input datasets ("truth"/ "label" and "predictor"/ "feature"), output dataset (predicted quantity), and a suite of hyperparameters that tune the model including modulating behavior at each hidden layer. DNN parameters and hyperparameters can be readily tuned to optimize the predictive model (i.e, optimize the agreement of DNN features with the DNN labels).

In this study, AER expanded the GOES radiance-based CNN approach developed in previous work (Phase I) to incorporate data from TROPOMI and the upcoming TEMPO mission in the neural network training. We analyzed the changes our new data and approach made to the smoke plume predictions from Phase I. We compared the DNN-based smoke plume predictions with the NOAA HMS truth data over a subset of key dates that included heavy smoke events studied during the TRACER-AQ campaign. Finally, we evaluated the improved smoke plume tracking models with the results of the surface BC2 campaign.

Our Phase II work incorporated several key differences from Phase I methods with the goal to greatly improve performance, speed and, ultimately, the utility and real-world applicability of smoke plume tracking ML techniques. Our streamlined Phase II processing has enabled more sophisticated model evaluation through (i) elimination of the manual spatial domain selection window; (ii) its ability to efficiently ingest GOES data so multiple dates can be readily evaluated; and (iii) its option to include multiple feature dimensions (temporal slices and/or multiple variables including radiance bands from GOES, UVAI from TROPOMI, and trace gases from TEMPO). We successfully demonstrated full DNN output for each of TEMPO+GOES (proof-of-concept only), TROPOMI+GOES and GOES_Only (for science and societal application).

Our Phase II work resulted in the following broad conclusions:

- NOAA HMS data is a valid choice for truth data based on comparison with smoke-relevant ground-based BC2 observations
- A DNN using TROPOMI and/or GOES data produces results consistent with NOAA HMS when evaluated against BC2 observations
- A PCA analysis on UVAI 340-380nm and GOES bands 1-8 and 15 indicates that PC1 through PC3 explain ~86% of the data variance. Optimal GOES bands used for smoke prediction are 1,2,3, and 15 and feature as important in the first two of ten PCs. TROPOMI UVAI is features as important in PC3 and, to a lesser extent, PC2.
- Importance of using optimal GOES bands in a DNN >> importance of including adjacent GOES timestamps in that DNN.
- DNN performance >> naïve Bayes performance from Phase I
- Many outstanding issues/questions from Phase I resolved

1. Introduction

1.1 Project Objectives

The purpose of this project was to extend the work accomplished by AER in Phase I Work Order Number 582-21-22400-007 (Phase I). In Phase I, AER used a naïve Bayes model along with the NOAA HMS analyst-derived smoke mask and radiance observations from GOES to identify smoke plumes. In this subsequent project (Phase II), AER developed an improved machine learning smoke mask algorithm that optimized the use of GOES data while also including the capability to use data from The TROPOspheric Monitoring Instrument (TROPOMI) and Tropospheric Emissions: Monitoring of Pollution (TEMPO) satellite instruments in the neural network training. In addition, Phase II included a smoke plume tracking model evaluation component using surface observations from the BC2 surface data. In summary, AER's Phase II objectives were to deliver (i) an improved smoke plume detection algorithm relative to Phase I, (ii) the code and data used to develop the Phase II model, and (iii) a final report and user's guide summarizing its use in tracking the transport of biomass burning smoke plumes.

Task 3 of the Work Order – "Using Computer Vision Techniques to Identify and Track Smoke Plumes – was concerned with all the Phase II technical development and results. In this task, we streamlined all Phase I Jupyter Notebook python-based code into standard python-based code. We also developed Phase II base python code where NOAA HMS truth data was incorporated into a DNN framework along with GOES radiance data, TROPOMI UVAI data, and TEMPO NO2 and HCHO EASD. All the base python code from Phase I and Phase II were then ingested into end-user Jupyter Notebook interfaces where an end-user can run Phase I and Phase II versions of smoke plume tracking using GOES-TEMPO, GOES-TROPOMI, and GOES-only data inputs. Further objectives of Task 3 were to (i) optimize the training datasets provided to the DNN to and (ii) evaluate the Phase II model against BC2 observational surface data of smoke-relevant species from the HGB region.

The schedule of deliverables for this project is given in **Error! Reference source not found.**.

Table 1. Projected Schedule for TCEQ Work Order No. 582-22-31966-013

Milestones	Planned Date
Task 1 – Work Plan	
1.1: TCEQ-approved Work Plan	February 3, 2022
1.2: TCEQ-approved QAPP	February 3, 2022
Task 2 – Monthly Progress Reports	
Task 3 – Using Computer Vision Techniques to Identify and Track Smoke Plumes	
3.1: Documentation, Scripts, Data Files	May 31, 2022
Task 4 – Draft and Final Reports	
4.1: Draft Report	June 15, 2022
4.2: Final Report	June 30, 2022

1.2 Background

Various satellite observations provide valuable information on the locations of fires and transport of smoke. However, there are multiple products that use different techniques to identify smoke plumes which can result in disagreement on the extent of the area covered by biomass burning smoke. In many cases, detecting the transport of smoke plumes from the initial fire location to populated areas usually requires a human analyst to track the extent and transport of the smoke. Fortunately, recent advances in artificial intelligence and computer vision would allow this analysis to be performed automatically. For example, Ba et al. (2019) used satellite observations from the Moderate Resolution Imaging Spectroradiometer (MODIS) to train a convolutional neural network (CNN) model called SmokeNet to classify scenes as smoke, dust, etc. Similarly, Larsen et al. (2020) used a deep fully convolutional neural network to predict fire smoke in satellite imagery of Australia in near-real time. The algorithm had a high classification accuracy and precision and could be applied to any geographic region.

One difficulty is in establishing a reference, or "truth", dataset as smoke detection and tracking smoke plume transport is not directly measurable. The NOAA HMS smoke plume data set, used as the ML "truth" dataset in this study, is a powerful operational product that is used by many air quality and health agencies to understand and quantify the potential impact of biomass burning smoke plumes on communities within the U.S. In order to produce the HMS dataset, human analysts examine a series of CONUS images and products, including GOES imagery (typically in visible wavelengths) for a given day and individually outline regions over which they perceive smoke plumes. This approach is powerful, as there are substantial variations in the size, shape, color, and evolution of smoke plumes, and human perception is capable of incorporating large amounts of data and quickly processing them in order to classify smoke plumes from, for example, clouds. However, there is ultimately individual subjectivity within this process, and it is extremely difficult to hard-code and develop machine tools to readily duplicate the series of subjective choices than an individual analyst might make (Brown-Steiner et al., 2021).

There are multiple smoke-relevant data sets available for training a smoke plume DNN. In this study, we primarily used raw GOES radiance data across all bands (as opposed to the imagery itself that is used in the HMS smoke plume estimation). When available, we also incorporated TROPOMI UVAI data into our ML framework in addition to the GOES data. We also demonstrated how TEMPO NO2 and HCHO data can be incorporated into a TEMPO-GOES ML model version; TEMPO is scheduled for launch in 2023 but there are currently EASD available for a small number of dates for planning purposes.

1.3 Report Outline

This Final Report highlights major activities and key findings, provides pertinent analysis, describes encountered problems and associated corrective actions, and details relevant statistics including data, parameter, or model completeness, accuracy and precision. It satisfies Deliverables 4.2 of the Work Plan for Work Order No. 582-22-31966-013:

Deliverable 4.2:	Final Report
Deliverable 4.2 Due Date:	June 30, 2022

The major results of Task 3 are summarized in Section 2. Section 3 discusses the quality assurance findings for this project following the procedures from the project Quality Assurance Project Plan (QAPP). Section 4 summarizes our conclusions and Section 5 makes recommendations for future work based on the results of this project.

2 Using Computer Vision Techniques to Identify and Track Smoke Plumes

2.1 Introduction

In Machine Learning (ML), Deep Neural Networks (DNN) use multiple hidden layers and hyperparameters to classify complex data. DNNs consist of input datasets (“truth”/“label” and “predictor”/“feature”), output dataset (predicted quantity), and a suite of hyperparameters that tune the model including modulating behavior at each hidden layer. DNN parameters and hyperparameters can be readily tuned to optimize the predictive model (i.e, optimize the agreement of DNN features with the DNN labels).

In this study, AER expanded the GOES radiance-based CNN approach developed in previous work to incorporate data from TROPOMI and the upcoming TEMPO mission in the neural network training. We analyzed the changes our new data and approach made to the smoke plume predictions from Phase I. We compared the DNN-based smoke plume predictions with the NOAA HMS truth data over a subset of key dates that included heavy smoke events studied during the TRACER-AQ campaign. Finally, we evaluated the improved smoke plume tracking models with the results of the surface BC2 campaign.

Our Phase II work incorporated several key differences from Phase I methods with the goal to greatly improve performance, speed and, ultimately, the utility and real-world applicability of smoke plume tracking ML techniques. To this end, in Phase II we eliminated significant sources of computing time overhead by leveraging our Amazon Web Services (AWS) resources to access more efficiently the large and numerous GOES data files. We also eliminated the Phase I manual training/testing region selection and replaced it with an automated random training/testing splitting of data across the study domain. In addition, Phase II demonstrated the ability to incorporate additional smoke-relevant data in addition to GOES radiances (i.e., TROPOMI UVAI, TEMPO NO2, TEMPO HCHO). Finally, the Phase II enhancements have enabled direct use of observational surface data to evaluate quality of both the truth dataset and the DNN model smoke predictions. This section demonstrates how the enhanced computing efficiency, model automation, and expansion enables the Phase II model to readily be deployed for important science and societal biomass burning applications.

2.2 Data Description

We briefly describe the datasets below. For more detail, please refer to Brown-Steiner et al. (2021). For detailed instructions on data access and download, please refer to the User Guide accompanying this Work Order (Deliverable 3.1).

2.2.1 NOAA HMS

To make the HMS Fire and Smoke product, National Environmental Satellite, Data, and Information Service (NESDIS) satellite analysts manually generate a daily operational list of fire locations and outline areas of smoke. After identifying fire locations, HMS analysts use imagery from multiple NOAA and NASA satellites to identify the geographic extent of smoke plumes. Smoke detection is done primarily with visible-band geostationary GOES imagery, which has high temporal coverage (typically every 10 min), occasionally assisted by GOES infrared imagery and polar orbiting satellite imagery (Brown-Steiner et al., 2021).

2.2.2 GOES ABI L1bRadC Bands 1-16.

The GOES L1b RadC (CONUS) radiances product provides top-of-atmosphere (TOA) outgoing radiances for 16 wavelength bands. The table below details the main features of the L1b product associated with each band and, based on our Phase II work, whether the band is optimal for smoke detection. Given the limited time span and geographic extent of our project, we designate bands as either “optimal”, “likely useful” or “uncertain”.

2.2.3 TROPOMI UVAI

TROPOMI Ultraviolet Aerosol Index (UVAI) is measured at two wavelength windows: 354-388nm (OMI Heritage) and 340-380nm (TOMS Heritage). Positive UVAI values indicate absorbing aerosols like smoke and dust; negative values indicate non-absorbing aerosols. Details on data access and processing for this study are provided in the accompanying User Guide.

2.2.4 TEMPO EASD

In preparation for the launch of the Tropospheric Emissions: Monitoring of Pollution (TEMPO) instrument scheduled for launch in January 2023, TEMPO synthetic data is available in NetCDF file format from NASA’s Short-term Prediction Research and Transition Center (SPoRT). In this study, inclusion of TEMPO EASD served as a placeholder for future analysis when real TEMPO data become available. 6-minute temporal resolution synthetic data are available for nitrogen dioxide (NO2), formaldehyde (HCHO) and aerosols for three dates: 2020-07-16 and 2020-07-17 for NO2 and HCHO; 2020-08-26 for aerosols. In this project, we focused on NO2 and HCHO using synthetic data from 2020-07-16. There was no smoke detected by HMS for the time window of the aerosol coverage, so we did not include aerosol EASD in this study.

2.2.5 BC2 Surface Measurements

BC2 surface measurements of smoke-relevant species were used in HMS “truth” dataset and DNN model evaluation. BC2 measurement locations are shown in Figure 1. Further details including data access are provided in the accompanying User Guide.

2.2.6 TRACER-AQ Measurements

The TRACER-AQ campaign did not have sufficient smoke-relevant data available to use in our observational data-based model evaluation. We therefore relied only on BC2 data for model evaluation. However, the TRACER-AQ data influenced selection of our study dates – the TRACER-AQ Gulfstream V aircraft platform was deployed on several dates with heavy regional smoke, and we incorporated a subset of these dates in our study temporal domain. See Table 3.

Table 2. GOES L1b Radiance Bands evaluated in this study.

Band ID	Optimal for Smoke?	Wavelength (mm)	Resolution (km)	Grid Spacing
1	Yes	0.47/Vis Blue	1	3000x5000
2	Yes	0.64/Vis Red	0.5	6000x10000
3	Yes	0.86/NearIR Veg	1	3000x5000
4	Likely	1.37/NearIR Cirrus	2	1500x2500
5	Likely	1.6/NearIR Snow, Ice	1	3000x5000
6	Likely	2.2/NearIR Cloud Ice	2	1500x2500
7	Likely	3.9/IR Shortwave	2	1500x2500
8	Likely	6.2/IR Upper Level Water Vapor	2	1500x2500
9	Uncertain	6.9/IR Mid Level Water Vapor	2	1500x2500
10	Uncertain	7.3/IR Lower Level Water Vapor	2	1500x2500
11	Uncertain	8.4/Cloud Top Phase	2	1500x2500
12	Uncertain	9.6/IR-Ozone	2	1500x2500
13	Uncertain	10.3/IR-Clean	2	1500x2500
14	Uncertain	11.2/IR-Standard	2	1500x2500
15	Yes	12.3/IR-Dirty	2	1500x2500
16	Uncertain	13.3/IR-CO2	2	1500x2500

M

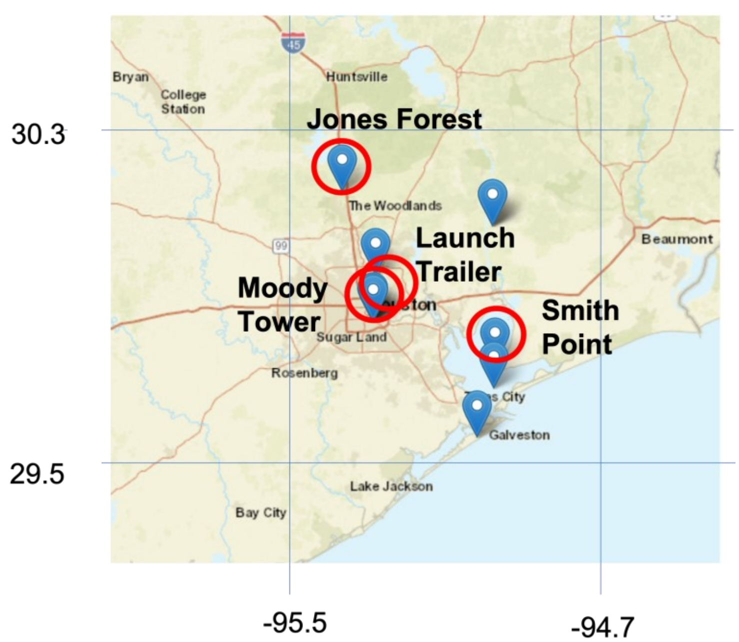


Figure 1. Map of BC2 surface sites (Houston Area). Highlighted sites (red circles) are those that had with available measurements for our selected study dates.

2.3 Methods

2.3.1 Code, Model, and Computing Structures

All data was analyzed and processed using a combination of Python and R. A schematic of the data sets and scripts used, along with their location in the flow of the ML framework, is shown in the figure below. Drawing upon code developed in Phase I, there are also three separate DNN models: versions for TEMPO+GOES data,

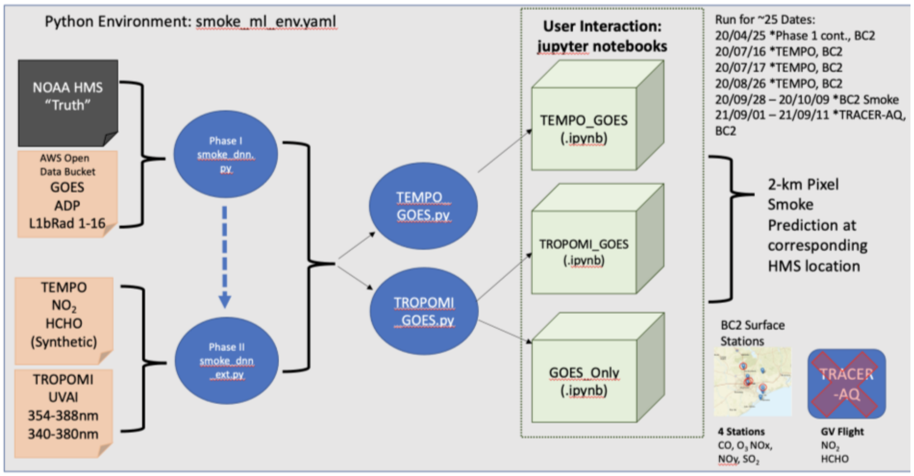


Figure 2. Summary of ML framework components. BC2 (direct) and TRACER-AQ (indirect, via study date selection) data are final steps in the process where the ML smoke prediction in the Houston area is compared to smoke-relevant measurements.

TROPOMI+GOES data, and GOES-only data. The latter is for instances where neither TROPOMI nor TEMPO data was available. See Table 3.

Similar to Phase I, NOAA HMS "truth" polygons are converted to GOES-grid-based point data and are assigned a binary smoke label (0=no smoke detected by HMS, 1 = smoke detected). The NOAA HMS qualitative smoke density estimates are ignored for our purposes.

The Phase II GOES+TEMPO and GOES+TROPOMI extension processes included data trimming, data quality filtering and the splitting of synthetic TEMPO data into training and testing components. In addition, the GOES AWS OpenData bucket was directly mounted which significantly increased the efficiency of our ML process by eliminating the need to copy and store the numerous large radiance files.

We incorporate pre-processed TROPOMI UVAI data as csv files for each date with all data quality filtered including removal of missing values. See accompanying User Guide for details on TROPOMI UVAI data downloading and processing.

We have further implemented an option to incorporate multiple time slices as well as multiple predictor variables: e.g., multiple radiance bands + TEMPO nitrogen dioxide + TEMPO formaldehyde (+ a placeholder for the TEMPO aerosol product, when that becomes available). In the Results section, we evaluate the relative importance of adjacent time slices over optimal predictor variables.

Finally, the Phase II DNN was designed such that most of the code complexity (including most functions) is handled by the base python code. The user interface consists of three Jupyter Notebooks (the TEMPO+GOES placeholder for upcoming TEMPO mission data, TROPOMI+GOES, and GOES-only). All post-processing of DNN output is done in R. See User Guide for more details.

2.3.2 Study Spatial and Temporal Domain Selection

We subset the spatial domain to a broadly TCEQ-relevant region (approx. bounds: 15N, 38N, -110W, -82W). Our spatial domain is shown in Figure 3 along with NOAA HMS daily smoke plumes and TROPOMI UVAI data for an example date 2020-10-01. Smoke presence was evaluated for 21 dates (Table 3).

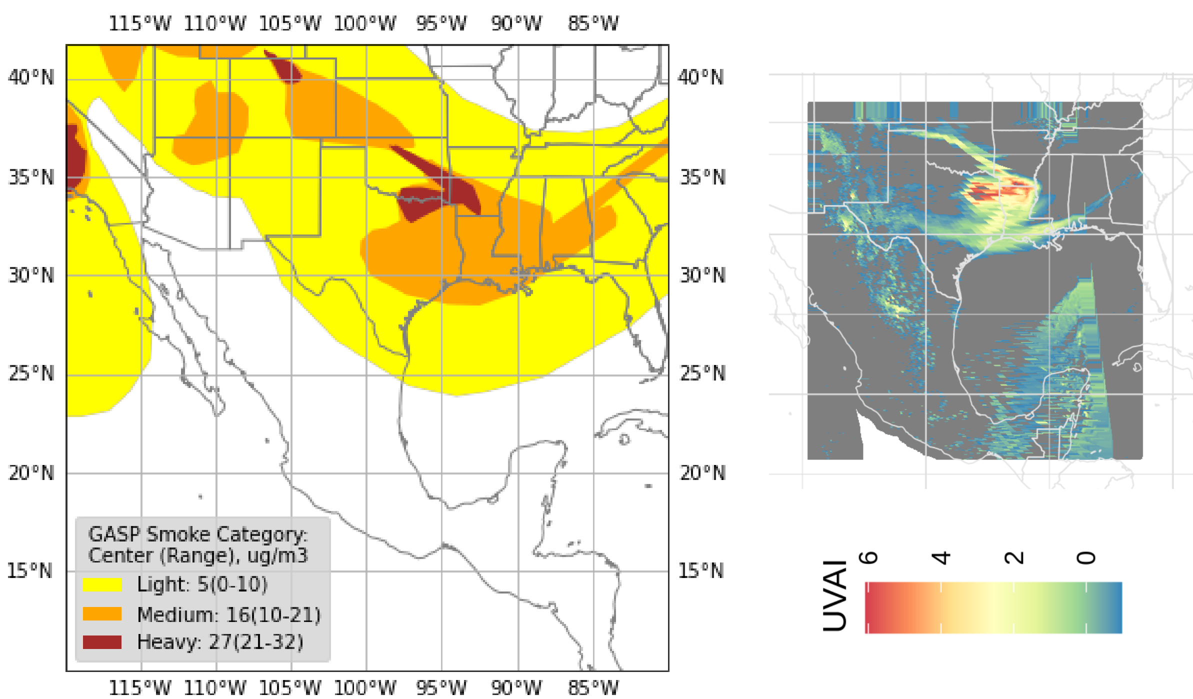


Figure 3. Study spatial domain overlaid with (left) HMS smoke plume polygons and (right) processed TROPOMI UVAI data. Date shown: 2020-10-01.

2.3.3 Model Optimization: Hyperparameter Tuning and PCA

We use the Sequential model from the keras python library. The Sequential model is used when creating a linear stack of layers with their corresponding weights. Our dataset comprising of images in their vectorized formats best works with the Sequential model where each layer has exactly one input tensor and one output tensor (Brown-Steiner et al., 2021). We follow common machine learning practice by setting our training, testing, and validation data fractions to 70%, 20%, and 10%, respectively, of the total data set. Our hyperparameter space consists of 11 settings that are described in Table 4 and described in more detail in Brown-Steiner et al. (2021).

As discussed later in the Results section, we also implemented a basic Principal Component Analysis to eliminate redundancy (and therefore improve computing efficiency/reduce overfitting) in GOES radiance predictor variables.

Table 3. List of dates for smoke plume ML model. TEMPO data is EASD.

Date	Explanation	HMS "Truth" shows smoke in HGB?	Available Data Sets
20200425	Continuity with Phase I	No	TROPOMI, GOES, HMS, BC2
20200716	TEMPO data (HCHO, NO2)	No	TROPOMI, TEMPO, GOES, HMS, BC2
20200717	TEMPO data (HCHO, NO2)	No	TROPOMI, TEMPO, GOES, HMS, BC2
20200826	TEMPO data (Aerosols)	No	TROPOMI, TEMPO, GOES, HMS, BC2
20200928	BC2 + HMS Smoke Events in Houston	No	TROPOMI, GOES, HMS, BC2
20200929	BC2 + HMS Smoke Events in Houston	Yes (Smith Point Only)	TROPOMI, GOES, HMS, BC2
20201001	BC2 + HMS Smoke Events in Houston	Yes	TROPOMI, GOES, HMS, BC2
20201002	BC2 + HMS Smoke Events in Houston	Yes	TROPOMI, GOES, HMS, BC2
20201003	BC2 + HMS Smoke Events in Houston	Yes	TROPOMI, GOES, HMS, BC2
20201004	BC2 + HMS Smoke Events in Houston	Yes	TROPOMI, GOES, HMS, BC2
20201005	BC2 + HMS Smoke Events in Houston	Yes	TROPOMI, GOES, HMS, BC2
20201006	BC2 + HMS Smoke Events in Houston	Yes	TROPOMI, GOES, HMS, BC2
20201007	BC2 + HMS Smoke Events in Houston	Yes	TROPOMI, GOES, HMS, BC2
20201008	BC2 + HMS Smoke Events in Houston	No	TROPOMI, GOES, HMS, BC2
20201009	BC2 + HMS Smoke Events in Houston	No	TROPOMI, GOES, HMS, BC2
20210901	TRACER-AQ GV Flt 1	Yes	GOES, HMS, BC2
20210903	TRACER-AQ GV Flt 2	Yes	GOES, HMS, BC2
20210908	TRACER-AQ GV Flt 3	Yes	GOES, HMS, BC2
20210909	TRACER-AQ GV Flt 4	Yes	GOES, HMS, BC2
20210910	TRACER-AQ GV Flt 5	Yes	GOES, HMS, BC2
20210911	TRACER-AQ GV Flt 6	Yes	GOES, HMS, BC2

Table 4. List of key hyperparameters in the DNN Phase II model. Where more than one element in a hyperparameter's setting exists, the number of elements reflects the number of hidden layers in the DNN.

Hyperparameter	Description	Value
nNodesH	Number of nodes per hidden layer.	[16, 16]
lossRateH	Penalty for a bad prediction.	[3e-3, 3e-3]
drop	Dropout fraction randomly zeroes nodes to simulate a larger network.	[0, 0]
activateH	Activation function used for hidden layers. Rectified Linear Unit (ReLu) is the most widely used activation function in almost all convolutional neural networks and deep learning.	['relu', 'relu']
nNodesFC	Number of fully connected nodes.	1
activateFC	Activation function for fully connected neural network. 'Sigmoid' maps output values from 0-1 and is appropriate for Boolean-type applications.	'sigmoid'
optimizer	The RMSprop optimizer uses an adaptive learning rate instead of treating the learning rate as a hyperparameter. This means that the learning rate changes over time.	'rmsprop'
loss	We trained a binary classifier to do a pixel-wise classification as Smoke/No Smoke. We use the binary cross-entropy as our loss function. The binary classification problem posed is – "what is the probability of there being smoke in the pixel?"	'binary_crossentropy'
nEpochs	Number of model iterations.	25
batchSize	Base-12; used to balance speed of model while minimizing the tendency to over-generalize.	96

2.3.4 Model Output and Evaluation

Given the nature of the TEMPO EASD, the TEMPO+GOES Jupyter Notebook was run for demonstration purposes only. However, we used the output from TROPOMI+GOES and GOES_Only model runs (via the Jupyter notebook interfaces) for the science applications discussed in the Results section. We ran the TROPOMI+GOES and GOES_Only models using multiple conformations of the hyperparameter and predictor space. We then evaluated the most important GOES bands using PCA. The final optimized model was selected based on the hyperparameter settings that consistently resulted in highest agreement between the NOAA HMS truth and TROPOMI+GOES predictors. A PCA was performed on this final optimized model to extract the most important predictors.

The DNN models (either TROPOMI+GOES or GOES_Only) were run for the study dates in Table 3, with the GOES_Only version run for dates without TROPOMI data. The Jupyter Notebook outputs DNN results as csv files for later analysis in R. The csv files consist of truth and predictor values for each gridcell along with predicted smoke (as a probability of smoke – i.e., mapped to values ranging from 0-1 by the sigmoid activation function). Quality of predicted and "truth" smoke designations were then evaluated based on binning smoke-relevant BC2 surface measurements on smoke- and non-smoke days. The accompanying User Guide provides detailed instructions. We discuss key results in the following section.

2.4 Key Results

2.4.1 Phase II Streamlining

By directly accessing GOES data from the OpenData AWS portal we have significantly increased the efficiency of our DNN and its capacity to be used for science and societal applications. In addition, our streamlined Phase II processing has enabled more sophisticated model evaluation through (i) elimination of the manual spatial domain selection window; (ii) its ability to efficiently ingest GOES data so multiple dates can be readily evaluated; and (iii) its option to include multiple feature dimensions (temporal slices and/or multiple variables including radiance bands from GOES, UVAI from TROPOMI, and trace gases from TEMPO). We successfully demonstrated full DNN output for each of TEMPO+GOES (proof-of-concept only), TROPOMI+GOES and GOES_Only (for science and societal application).

2.4.2 Model Optimization: Hyperparameter tuning and PCA results

Our final selected model was Model 2 from Table 5 and with a Train/Test score of 64%, Model 2 represents an improvement of 11% over the Phase 1 (Naïve Bayes) version of our smoke prediction. We found that eliminating a drop rate, doubling the number of nodes for each of the two hidden layers, while decreasing the batch size to 96 was the most efficient way to obtain the highest performance. Notably, performance was fairly insensitive to loss rate (penalty for a bad prediction).

Prior to applying the final Model 2 for our science applications, we wanted to remove feature redundancies. Given that GOES data processing incurs significant computational time and space, we wanted to minimize the number of GOES bands used for maximum information; the results of this analysis are summarized in Table 6 and the steps are described below.

Our first step was to run Model 2 individually for each GOES band (1-16) along with both TROPOMI UVAI bands (340-380nm, 354-388nm); all other settings were held constant. The mean Phase 1 train/test score across all 16 bands was 51%, or about the same as a coin toss (Table 6). We then established a selection criterion where a

Table 5. Hyperparameter Tuning Step 1: Identification of key hyperparameters. Blank cells indicate values identical to Model 1-Base. Shaded green models were selected for use in final model selection in Step 2. Data Used: TROPOMI, GOES, and NOAA HMS from 2022-10-01.

	MODEL									
	1-Base	**1a**	**1b**	**1c**	**1d**	**1e**	**1f**	**1g**	**1h**	**1i**
	[8, 8]							[16, 16]		
	[3e-3,3e-3]					[3e-2 3e-2]	[3e-4 3e-4]			
drop	[0.5,0.6]				[0,0]					
activate	[‘relu’,‘relu’]	[‘elu’, ‘elu’]								
Fully connected										
nNodesFC	1									
activateFC	'sigmoid’									
Fitting and Training										
optimizer	'rmsprop’		'adam’							
loss	'binary cross- entropy’									
metrics	[‘accuracy’]									
nEpochs	25			10					50	
batchSize	256									96
GOES Band	6									
Train Score	0.6	0.6	0.59	0.59	0.62	0.6	0.6	0.61	0.61	0.6
Test Score	0.59	0.59	0.58	0.58	0.61	0.59	0.59	0.6	0.6	0.6
Phase 1 Test	0.53									
Phase 1 Train	0.52									

GOES band was deemed smoke-relevant if the associated DNN Model 2 performance was >10% over the Phase 1 mean. Based on this criterion, GOES bands 1-8 and 15 each resulted in a DNN/Phase 2 model performance >61%.

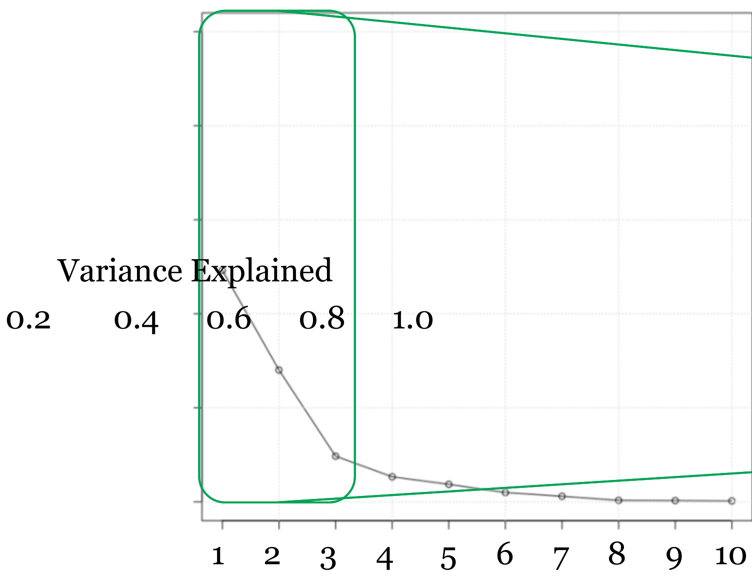
Table 6. Summary of Model 2 performance on 2020-10-01 using available TROPOMI UVAI data GOES bands. GOES bands in bold font rows are likely most smoke-relevant based on the DNN performance. Italicized results are various combinations of the selected smoke-relevant GOES bands. Highlighted green row indicates optimized model.

Phase 1 Ref: Train Score	Phase 1 Ref: Test Score	Phase 2 DNN Train Score	Phase 2 DNN Test Score	GOES Band Used	GOES File(s) Used
0.5	**0.5**	**0.62**	**0.62**	**1**	**1,2,3,4**
0.5	**0.5**	**0.64**	**0.64**	**2**	**1,2,3,4**
0.5	**0.5**	**0.64**	**0.64**	**3**	**1,2,3,4**
0.5	**0.5**	**0.63**	**0.63**	**4**	**1,2,3,4**
0.5	**0.5**	**0.64**	**0.64**	**5**	**1,2,3,4**
0.53	**0.52**	**0.64**	**0.63**	**6**	**1,2,3,4**
0.53	**0.53**	**0.65**	**0.64**	**7**	**1,2,3,4**
0.52	**0.52**	**0.64**	**0.64**	**8**	**1,2,3,4**
0.51	0.51	0.55	0.55	9	1,2,3,4
0.51	0.5	0.56	0.56	10	1,2,3,4
0.5	0.5	0.57	0.56	11	1,2,3,4
0.5	0.5	0.54	0.54	12	1,2,3,4
0.5	0.5	0.61	0.61	13	1,2,3,4
0.5	0.5	0.59	0.59	14	1,2,3,4
0.5	**0.5**	**0.66**	**0.66**	**15**	**1,2,3,4**
0.5	0.5	0.57	0.56	16	1,2,3,4
0.5	***0.5***	***0.72***	***0.72***	***1-8,15***	***3***
0.5	***0.5***	***0.72***	***0.71***	***1,2,3,15***	***3***
0.5	***0.5***	***0.68***	***0.68***	***1,2,3***	***3***

We re-ran Model 2 combining information from all nine of the selected bands which resulted in a model performance of 72% on 2022-10-01 (versus a Phase 1 model performance of 50%). However, processing nine GOES bands is still time and computationally expensive, so we narrowed down the necessary bands further by performing a PCA on the nine-band data set. The goal here was to remove redundancy – i.e., eliminating the need to process features (GOES bands in particular) that had little to no impact on ultimate model smoke predictive capacity. The relative contribution of the 10 principal components (PC) to the variance of the dataset is shown in Figure 4. Of note, PC1, PC2, and PC3 explained 49%, 28%, and 9.7% of the data variance respectively. Within PC1 and PC2, GOES bands 1,2,3, and 15 were among the most relevant. Also worth noting, TROPOMI UVAI was not identified as a prominent feature until PC3 and, to a lesser extent, PC2. Therefore, in instances where TROPOMI UVAI data were not yet available (i.e., the 2021 study dates in

September and October) relying on GOES data only for smoke prediction is a reasonable approach.

Noting that GOES band 1 explains most of the variance in smoke predictions, we used this to explore the relative importance of incorporating adjacent time slices vs. optimal GOES bands b1, b2, b3, and b15. In other words, would the model perform as well or better if band 1 and adjacent time slices were used rather than bands 1,2,3, 15 and a single time slice? The purpose, again, was to minimize compute effort and maximize performance as processing the four bands for each time slice takes nearly 30 minutes. Table 6 and Figure 5 clearly illustrate that using optimal PC1 and PC2 bands 1,2,3 and 15 is more valuable than using a single band 1 that explains the most variance. For applications, a user can include the four optimal GOES bands and a single time slice, thereby saving considerable computing overhead.



	PC1 Weights	PC2 Weights	PC3 Weights
UVAI	0.0276	0.150	-0.972
GOES b1	0.420	-0.0185	-0.0594
GOES b2	0.421	0.0896	-0.0154
GOES b3	0.381	0.192	-0.0290
GOES b4	0.342	-0.223	-0.00213
GOES b5	0.206	0.509	0.0821
GOES b6	0.287	0.428	0.130
GOES b7	-0.165	0.521	0.122
GOES b8	-0.352	0.127	-0.112
GOES b15	-0.327	0.391	-0.0106

Figure 4. Results from PCA of TROPOMI+GOES features. Table provides output from PC1-PC3; for clarity we have not shown PCs 4-10. UVAI is 340-380nm.

2.4.3 Model Application and Evaluation

Other than noting successful end-to-end completion of TEMPO+GOES, the limited TEMPO EASD data prevented us from conducting further evaluation. In this section we focus entirely on the TROPOMI+GOES and GOES Only DNN formulations. We trained and tested the TROPOMI+GOES and GOES Only models against the HMS “truth” for the study dates shown in Table 3.

For our evaluation with BC2 surface station observational data, we used the DNN csv file output for each study date. A screenshot of the DNN model output is shown in Figure 6 for clarity. For each date, we find the nearest point in the DNN testing output csv file to each of the four BC2 sites shown in Figure 1. Then, we provide the

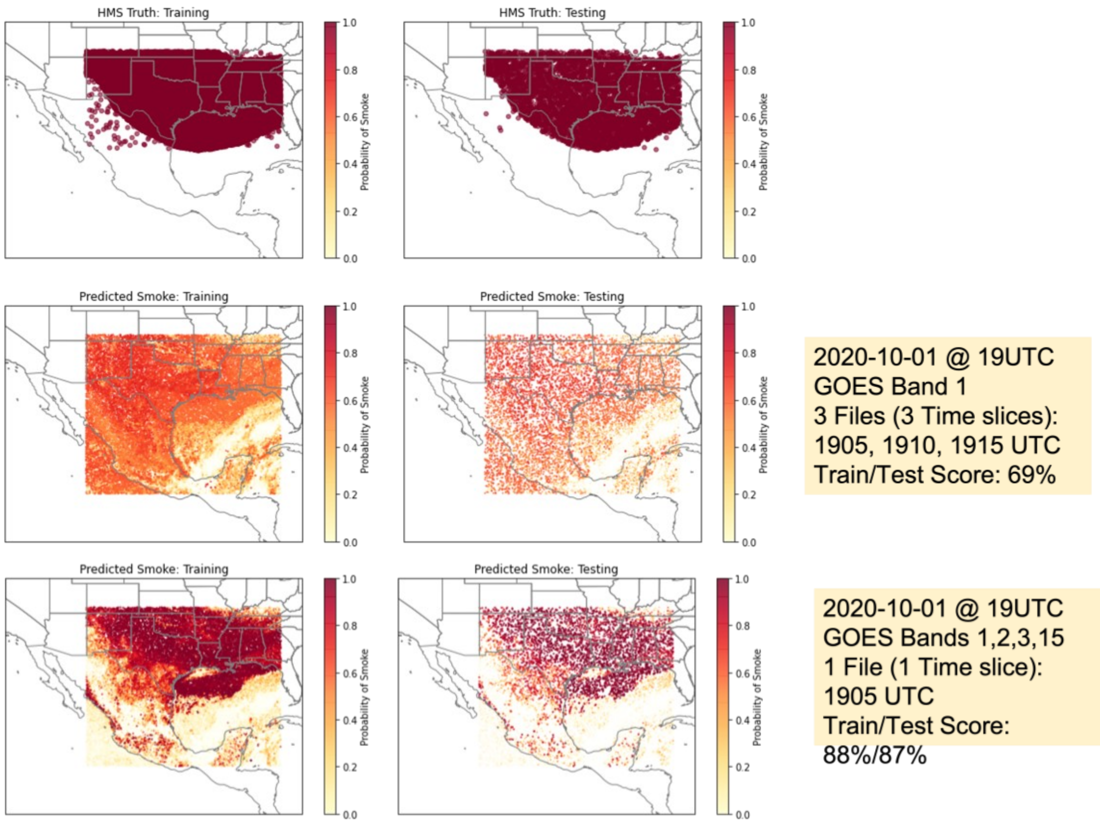


Figure 5. Illustration of the value of optimal bands over time slices (adjacent files).

smoke category for the pixel nearest to each site and date for (i) HMS "truth" and (ii) DNN prediction. If the HMS truth value is "1", the smoke category for that site and date is "smoke"; otherwise, it is non-smoke. If the DNN predicted smoke probability is >0.6, the smoke category for that site and date is "smoke"; otherwise, it is non-smoke. We then aggregate all the data for each site and divide into "smoke" and "non-smoke" days for both HMS and DNN prediction. Table 7 summarizes these results for the four BC2 sites. We find that both the HMS and DNN smoke categorizations are consistently associated with significantly higher values of smoke-relevant variables measured at the BC2 sites, suggesting that HMS and DNN have some smoke predictive power. However, the DNN tends to predict more "non-smoke" events such that BC2 measurements on DNN-predicted smoke days are skewed lower than on HMS-predicted smoke days. But we also note that these differences between HMS and DNN categorizations are typically not significant.

```
In [1480]: train_df
```

Out[1480]:

	Lat	Lon	uvai_340_380	GOES_b1_f2	HMS_Smoke	Pred_Smoke
0	34.971226	-107.984398	-2.341297	73.952438	True	0.698723
1	26.818151	-106.304253	-1.990572	67.455589	False	0.839928
2	33.397305	-105.560417	-2.345750	78.825073	True	0.685930
3	36.470234	-100.128937	-1.716757	86.946136	True	0.891949
4	32.347477	-89.517426	-2.109215	65.831375	True	0.799095
...	...	...	...	...	...	...
34995	32.121281	-97.980377	-1.599325	78.012970	True	0.897024
34996	29.409077	-107.711258	-1.982500	82.885605	False	0.861059
34997	21.261831	-104.919060	-2.290073	90.194565	False	0.698295
34998	25.989136	-102.037025	-1.732639	86.134033	False	0.890308
34999	29.725601	-103.667755	-1.019700	115.369858	True	0.781470

35000 rows × 6 columns

Figure 6. Screenshot of DNN output. In this example, the DNN was tested for a single GOES band (band 1) at file 2 (GOES_b1_f2) and TROPOMI UVAI at 340-380nm. Additional GOES bands are appended column-wise with the same naming convention (GOES_bX_fY). Files are output as csv to enable further analysis.

Table 7. BC2 Surface Station data from four sites evaluated according to predictions of smoke by HMS ("truth") and the DNN (predicted). For each study date and BC2 site, and smoke estimate (HMS, DNN) daily averages of data were calculated for "smoke" or "nonsmoke" days. Within each smoke category, differences in BC2 data across HMS and DNN-based predictions are also shown. 95% CIs derived from t-test. Significant differences are in ***bold italics****.*

All units: ppb	CO	O3	NO	NO2	NOy	SO2
Jones Forest						
HMS Mean: Smoke Days	142	40.5				
HMS Mean: Nonsmoke Days	111	28.7				
HMS Smoke-Nonsmoke (95% CI)	***30.7 (17.7, 43.7)***	***11.8 (11.1,12.6)***				
DNN Mean: Smoke Days	135	40.0				
DNN Mean: Nonsmoke Days	116	29.1				
DNN Smoke-Nonsmoke (95% CI)	***18.2 (5.17, 31.2)***	***11.0 (10.2, 11.8)***				
DNN-HMS Smoke Days (95% CI)	-0.498 (-1.34,0.34)	-0.498 (-1.34, 0.340)				
DNN-HMS Nonsmoke Days (95% CI)	0.377 (-0.336, 1.09)	***0.377 (0.336, 1.09)***				
Launch Trailer						
HMS Mean: Smoke Days	248	33.8	4.54	13.3	18.4	
HMS Mean: Nonsmoke Days	151	20.7	3.15	10.8	11.6	
HMS Smoke-Nonsmoke (95% CI)	***97.1 (86.5,108)***	***13.1 (12.04, 14.2)***	***1.38 (0.955, 1.82)***	***2.51 (1.81, 3.21)***	***6.78 (6.06, 7.51)***	
DNN Mean: Smoke Days	246	34.2	4.11	12.2	17.8	
DNN Mean: Nonsmoke Days	162	21.6	3.60	11.8	12.7	
DNN Smoke-Nonsmoke (95% CI)	***83.8 (73.2, 94.4)***	***12.6 (11.5, 13.8)***	***0.514 (0.0708,0.958)***	0.352 (-0.373, 1.08)	***5.06 (4.31, 5.80)***	
DNN-HMS Smoke Days (95% CI)	-2.56 (-13.0,7.91)	0.412 (-0.938, 1.76)	-0.427 (-0.932, 0.0784)	***-1.10 (-1.92, -0.278)***	-0.635 (-1.48, 0.207)	
DNN-HMS Nonsmoke Days (95% CI)	***10.7 (0.0772, 21.3)***	***0.868 (0.130,1.61)***	***0.449 (0.0880, 0.810)***	***1.06 (0.473, 1.65)***	***1.09 (0.480, 1.71)***	
Moody Tower						
HMS Mean: Smoke Days	183	38.7	3.45	12.2	17.4	0.576
HMS Mean: Nonsmoke Days	133	22.0	1.89	6.78	10.0	0.210
HMS Smoke-Nonsmoke (95% CI)	***49.9 (43.3, 56.5)***	***16.7 (15.7, 17.7)***	***1.56 (1.19, 1.93)***	***5.46 (4.99, 5.93)***	***7.37 (6.69, 8.05)***	***0.367 (0.309, 0.425)***
DNN Mean: Smoke Days	178	39.2	3.21	12.1	16.7	0.470
DNN Mean: Nonsmoke Days	142	23.2	2.19	7.36	11.2	0.318
DNN Smoke-Nonsmoke (95% CI)	***36.1 (29.1,43.0)***	***16.0 (15.0, 17.1)***	***1.02 (0.631, 1.42)***	***4.74 (4.24, 5.24)***	***5.49 (4.78, 6,19)***	***0.152 (0.0935, 0.211)***
DNN-HMS Smoke Days (95% CI)	-5.54 (-13.9, 2.85)	0.414 (-0.839, 1.67)	-0.233 (-0.718, 0.252)	-0.147 (-0.757, 0.463)	-0.722 (-1.54, 0.0921)	***-0.106 (-0.178, -0.0341)***
DNN-HMS Nonsmoke Days (95% CI)	***8.31 (3.63,13.0)***	***1.14 (0.420, 1.87)***	***0.303 (0.0647, 0.541)***	***0.577 (0.259, 0.894)***	***1.16 (0.611, 1.70)***	***0.108 (0.0683, 0.149)***

Smith Point							
HMS Mean: Smoke Days		43.3	0.410	3.90		
HMS Mean: Nonsmoke Days		30.5	0.239	2.27		
HMS Smoke-Nonsmoke (95% CI)		***12.8 (11.9,13.8)***	***0.171 (0.0763, 0.265)***	***1.62 (1.23, 2.01)***		
DNN Mean: Smoke Days		41.5	0.358	3.54		
DNN Mean: Nonsmoke Days		31.1	0.273	2.47		
DNN Smoke-Nonsmoke (95% CI)		***10.3 (9.37, 11.2)***	0.0852 (-0.00692, 0.177)	***1.07 (0.714, 1.43)***		
DNN-HMS Smoke Days (95% CI)		***-1.89 (-2.81, -0.969)***	-0.0517 (-0.155, 0.0519)	-0.356 (-0.856, 0.144)		
DNN-HMS Nonsmoke Days (95% CI)		0.651 (-0.279, 1.58)	0.0339 (-0.0480, 0.116)	***0.194 (0.0228, 0.365)***		

2.5 Challenges and Problems Encountered

- TEMPO EASD data sparsity (NO2, HCHO for two days only that are not particularly reflective of smoky conditions over Texas). TEMPO EASD Aerosol data was only for a few hours on one day that (i) did not overlap with either NO2 or HCHO EASD dates (ii) had no smoke. Therefore, we could only test TEMPO+GOES with NO2 and HCHO as features.
- The TRACER-AQ campaign did not have sufficient smoke-relevant data available to use in our observational data-based model evaluation. We therefore relied only on BC2 data.
- ***Issue:*** TROPOMI UVAI data has significant "missingness" for the original selected 4 dates (2020-04-25, 2020-07-16, 2020-07-17, 2020-08-26). The low hourly sample size for the four dates selected makes it difficult to incorporate in a meaningful way in our current ML set up. ***Solution:*** We extended our ML smoke plume tracking dates of interest to 19 days beyond the original 4 proposed and subset over regions relevant to Texas to manage compute time. There is now sufficient data coverage for smoke-relevant dates (see example date in Figure 1)
- ***Issue:*** Multiple product grid resolutions makes a GOES+TEMPO+TROPOMI version of our model computationally expensive; we will explore ways to circumvent this. ***Solution:*** We have two versions of the model, GOES+TEMPO and GOES+TROPOMI. Given that TEMPO synthetic data only exists for three dates, this is unlikely to substantially change any results.
- ***Issue:*** Sequential processing of features led to inconsistency in feature and label location indices. The result was that feature and label locations were getting scrambled and did not refer to the same spatial point; in other words, a given row of features and a label were scrambled geographically therefore producing meaningless results. ***Solution:*** We refined our index selection method and such that TROPOMI (TEMPO) features are randomly split into training and testing; the associated locations are then mapped to the nearest HMS-to-pixel location (to get the label location) and then passed to the GOES feature selection (to get the nearest GOES feature).

3 Quality Assurance

The processing and analysis scripts used in this project will be inspected by a team member not involved in their creation for accuracy. All automated calculations and at least 10% of manual calculations will be inspected for correctness. This meets the requirement of Level III QAPPs that 10% of the data must be inspected.

As the quality of the information, including secondary data, was not be evaluated by EPA, the below disclaimer applies to all project deliverables:

> *Disclaimer: The information contained in this report or deliverable has not been evaluated by EPA for this specific application, i.e. the use of computer vision techniques to identify and track smoke plumes.*

4 Conclusions

Here we summarize the conclusions of our project, with reference to the corresponding report section.

- NOAA HMS data is a valid choice for truth data based on comparison with smoke-relevant ground-based BC2 observations
- The TROPOMI+GOES and GOES Only DNN produces results consistent with NOAA HMS when evaluated against BC2 observations
- A PCA analysis on UVAI 340-380nm and GOES bands 1-8 and 15 indicates that PC1 through PC3 explain ~86% of the data variance. Optimal GOES bands used for smoke prediction are 1,2,3, and 15 and feature as important in the first two of ten PCs. TROPOMI UVAI is features as important in PC3 and, to a lesser extent, PC2.
- Importance of optimal GOES bands >> importance of adjacent GOES timestamps
- DNN performance >> naïve Bayes performance from Phase I
- Many outstanding issues/questions from Phase I resolved

5 Recommendations for Further Study

Based on the results of this work, we make the following recommendations for further study:

- Eventual multi-product model (TEMPO+TROPOMI+GOES) when real TEMPO data becomes available, including TEMPO aerosol data.
- Multi-regional multi-season approach: relevant GOES bands and other features (TROPOMI UVAI, TEMPO data) might vary in importance based on type of fire/smoke (e.g. agricultural vs. forest, including distance from fire).
- Aggregate training data sets across multiple dates rather than training and testing on the same date. This will include more randomization in the training and testing (and also a larger sample size). The testing can then be done for individual dates of interest.

6 References

Ba, R., Chen, C., Yuan, J., Song, W. and Lo, S., 2019. SmokeNet: Satellite smoke scene detection using convolutional neural network with spatial and channel-wise attention. Remote Sensing, 11(14), p.1702.

Brown-Steiner, B., Dayalu, A., Ashok, A., Alvarado, M. (2021), *Wildfire Smoke Tracking Algorithm*, Brief Technical Memo to Texas Commission on Environmental Quality (TCEQ) for Work Order No. 21-22400-007 under TCEQ Contract No. 582-19-90498, June 30.

Larsen, A., Hanigan, I., Reich, B.J., Qin, Y., Cope, M., Morgan, G. and Rappold, A.G., 2020. A deep learning approach to identify smoke plumes in satellite imagery in near-real time for health risk communication. Journal of Exposure Science & Environmental Epidemiology, pp.1-7.