Pixel-Based Smoke Detection Using Machine Learning for the Next Generation Geostationary Satellite Imagery

Aaron Kaulfus[1], Muthukumaran Ramasubramanian[1], Iksha Gurung[1], Manil Maskey[2], Rahul Ramachandran[2], Udaysankar Nair[1]

(1) University of Alabama in Huntsville
(2) NASA/Marshall Space Flight Center

2020 American Meteorological Society Annual Meeting. January 12-16. Boston, MA

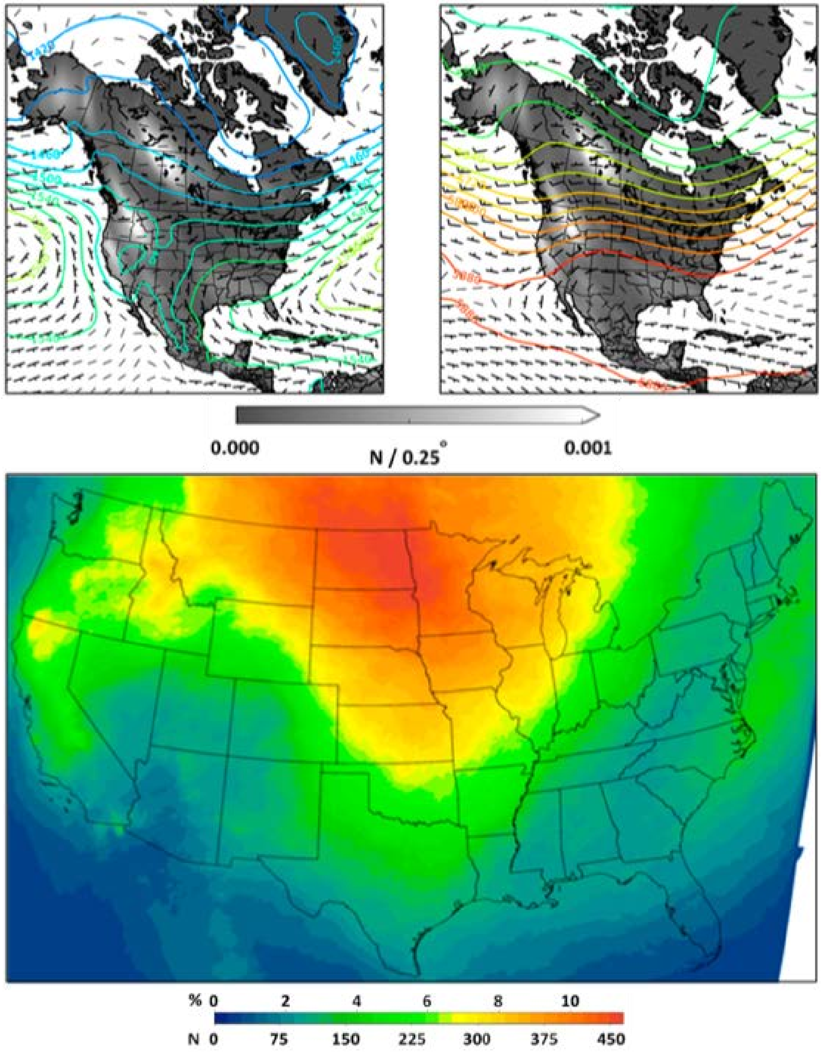
Introduction

Biomass burning smoke has numerous detrimental environmental and ecological impacts including:

- Respiratory and cardiovascular illnesses
- Radiation budget
- Nutrient availability

Impacts realized both near source and potentially thousands of kilometers downwind depending on:

- Fire duration
- Amount and type of biomass burned
- Meteorological and fuel conditions
- Vertical distribution in the atmosphere

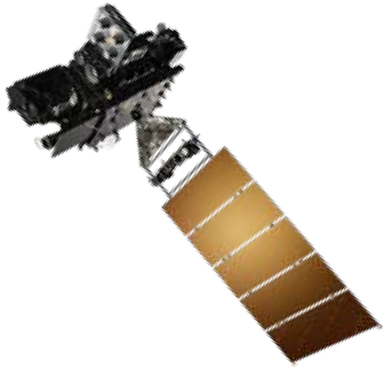


Spatial distribution of MODIS fire occurrence and mean atmospheric motion (top) and HMS smoke frequency for summer 2006-2015 (bottom). From Kaulfus et al. 2017 Figure 2.

Objective

Deploy a machine learning model for smoke detection in satellite remote sensing observations

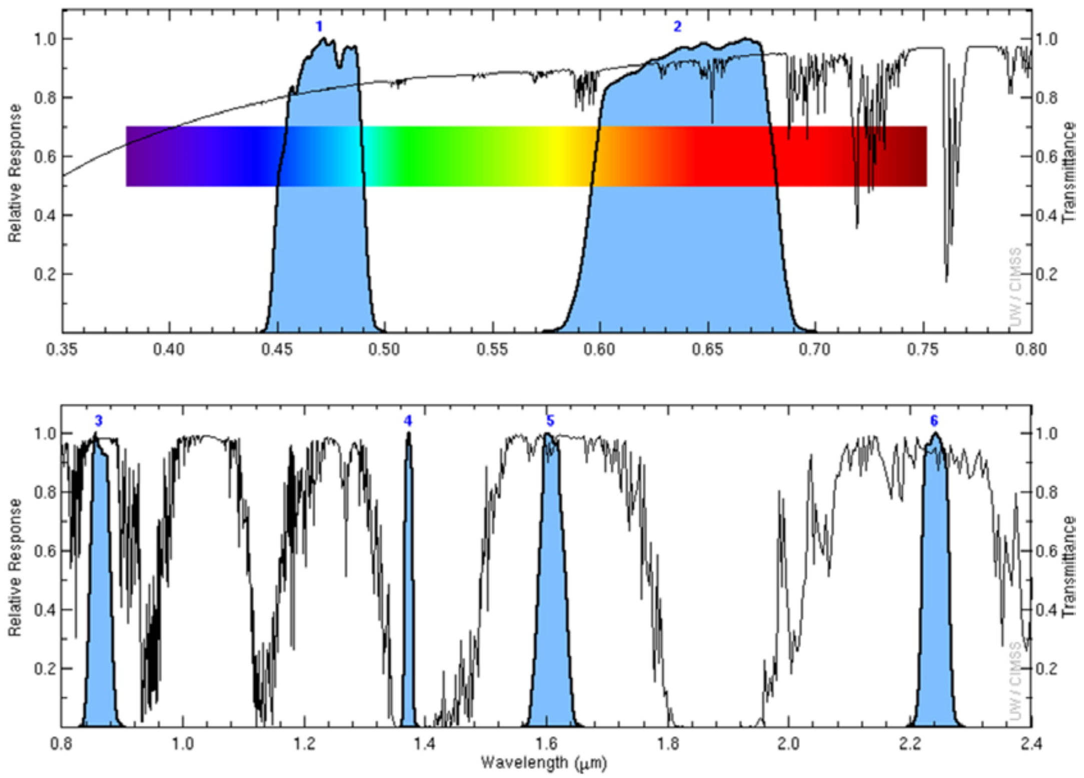
- Leverage observations from the new generation of geostationary satellite to overcome spatial and temporal limitations in current smoke detection techniques
- Develop classification alternative to existing multispectral or subjective manual methods
- Leverage cloud computing resources
  - Scalable to large data volumes
  - Computationally efficient
- Develop near-real time capabilities



Data

Geostationary Operational Environmental Satellite (GOES) 16 ABI visible and near-IR

- Bands 1-6 (0.47, 0.64, 0.86, 1.37, 1.6 and 2.2 µm)
  - Smoke aerosols reflect shortwave radiation
  - Additional bands for capturing signiture spectrum of atmospheric aerosols and surface features
- L1B radiance data from AWS
  - Spatially resample to 1 kilometer and convert to reflectance



ABI spectral response plot of the visible and near-IR bands. From Schmit, T.J., P. Griffith, M.M. Gunshor, J.M. Daniels, S.J. Goodman, and W.J. Lebair, 2017: A Closer Look at the ABI on the GOES-R Series. Bull. Amer. Meteor. Soc., **98**,681–698, https://doi.org/10.1175/BAMS-D-15-00230.1

Smoke Extent Truth Dataset

Informed by National Oceanic and Atmospheric Administration (NOAA) Hazard Mapping System (HMS) smoke product

- Manual quality control by subject matter expert to ensure all smoke in a single GOES 16 image is labeled

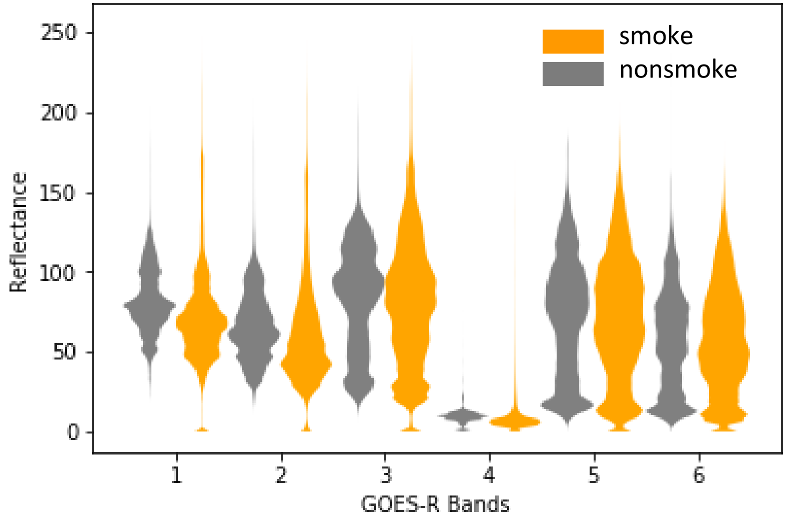
Model trained on ~1 million smoke pixels

- Over low and high background reflectance (land and ocean)
- Full range of sun angles
- Range of optical thicknesses

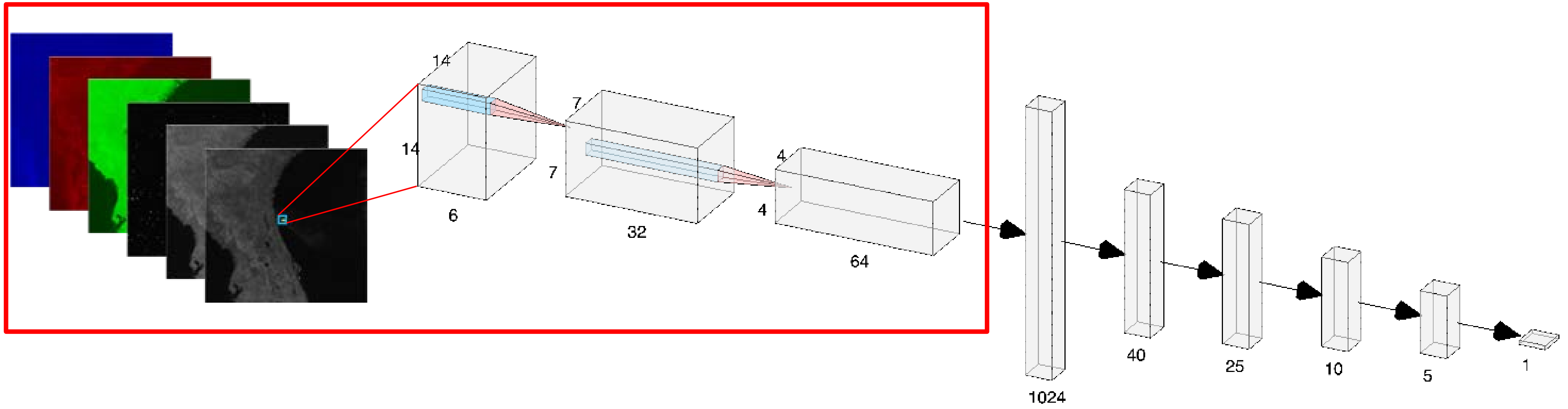
Training and testing dataset includes null features including

- Other aerosols
- Snow and ice
- Clouds

GOES 16 Band pseudo-RGB with nearest in time HMS shapefiles (magenta and purple) with subject matter quality controlled shapefile (blue).

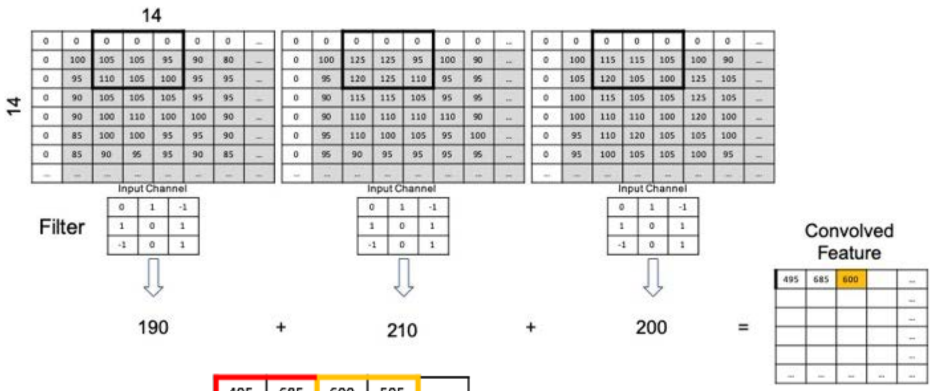


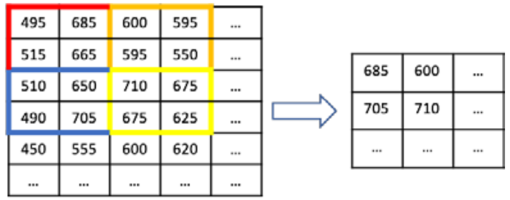
Model Architecture



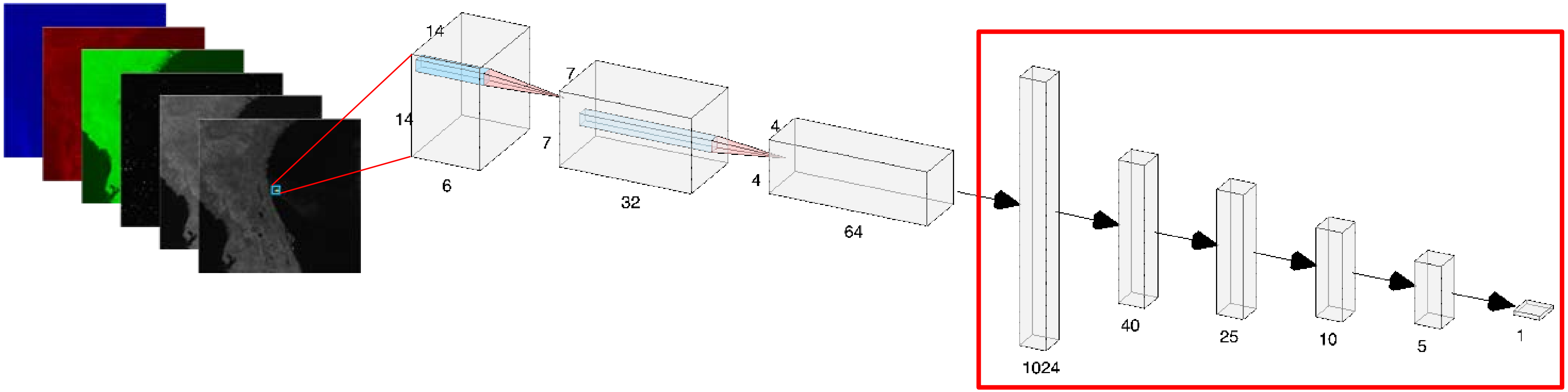
Apply a pixel-based Convolutional Neural Network (CNN)

- Input (N*2)*(N*2) neighborhood of reflectance values surrounding a center pixel (sample)
- 3 convolutional layers
- Each convolutional layer followed by max-pooling layer
- Convolutional outputs are flattened into vectors



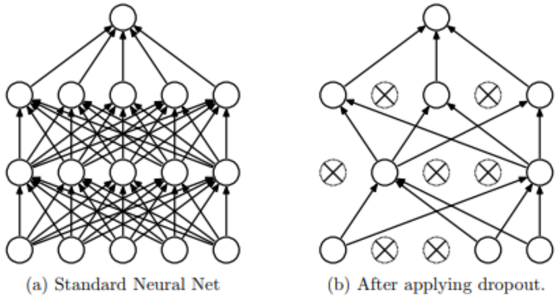


Model Architecture

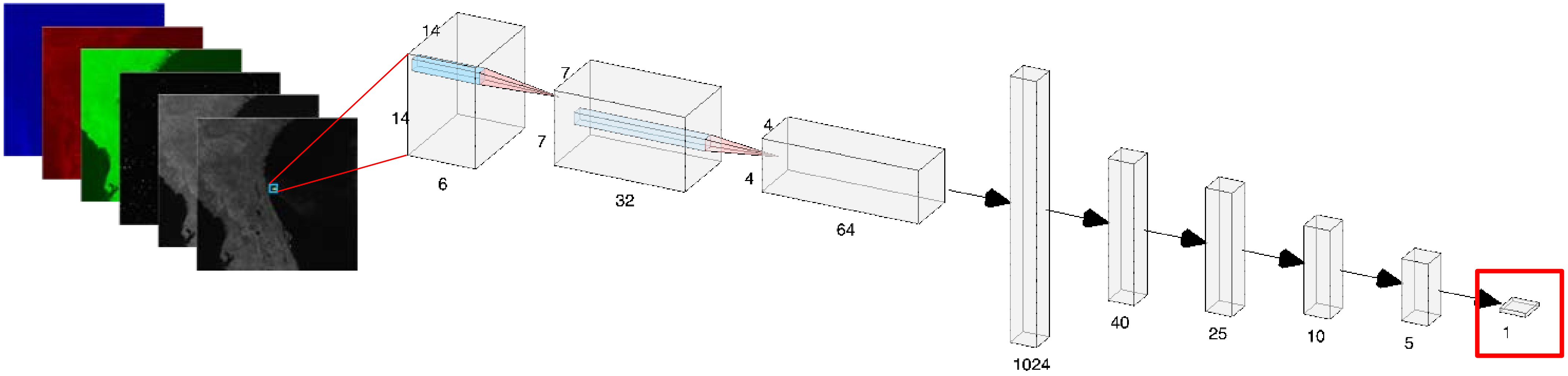


Apply a pixel-based Convolutional Neural Network (CNN)

- 4 fully connected layers with activation function calculation g(Wx + b)
  - x is the flattened input vector
  - W is the weight matrix
  - b is the bias vector
- Dropout for each fully connected layer



Model Architecture



The model outputs the probability, ranging from 0 - 1 as determined by a sigmoid function, that a pixel is smoke contaminated

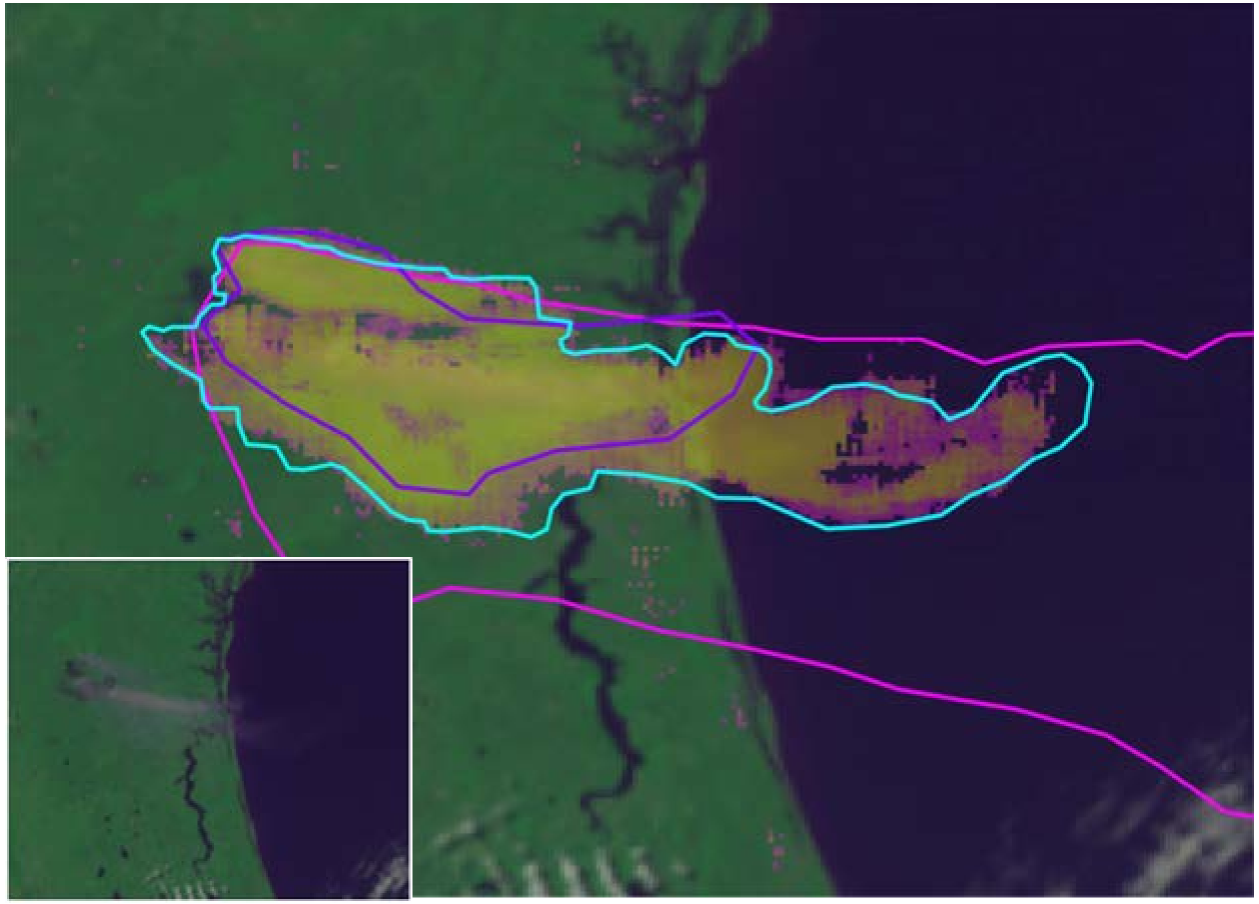
$$p(x) = \frac{1}{1+e^{-x}}$$

p > 0.5 threshold applied to define smoke

Dataset Conceptual Validation

Predictions more closely resemble quality controlled shapefiles

Model probabilities resemble visually observed optical thicknesses



GOES 16 pseudo-RGB with contoured model predictions (shading), HMS shapefiles (magenta and purple), and subject matter expert quality controlled shapefile (cyan)

Results

N=7	Precision	Recall	F1-Score	Accuracy
All	0.852	0.590	0.697	0.918
Land	0.883	0.559	0.684	0.916
Water	0.741	0.770	0.770	0.925

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + TN + TP + FN}$$

$$F1\ Score = 2 * \frac{Precision\ * Recall}{Precision + Recall}$$

A neighborhood size of 7 provided best tradeoff between quality and quantity of smoke predictions

- Best model has low false positive rate which drives high precision
- Prefer conservative identification over incorrect classification
- High accuracy artifact of large number of true negatives

Overall, better predictive capability of smoke over water

- Degraded precision driven by relative increase in false positives

Results

Discriminate smoke from variety of cloud types

- Cumulus, cirrus, coastal stratus

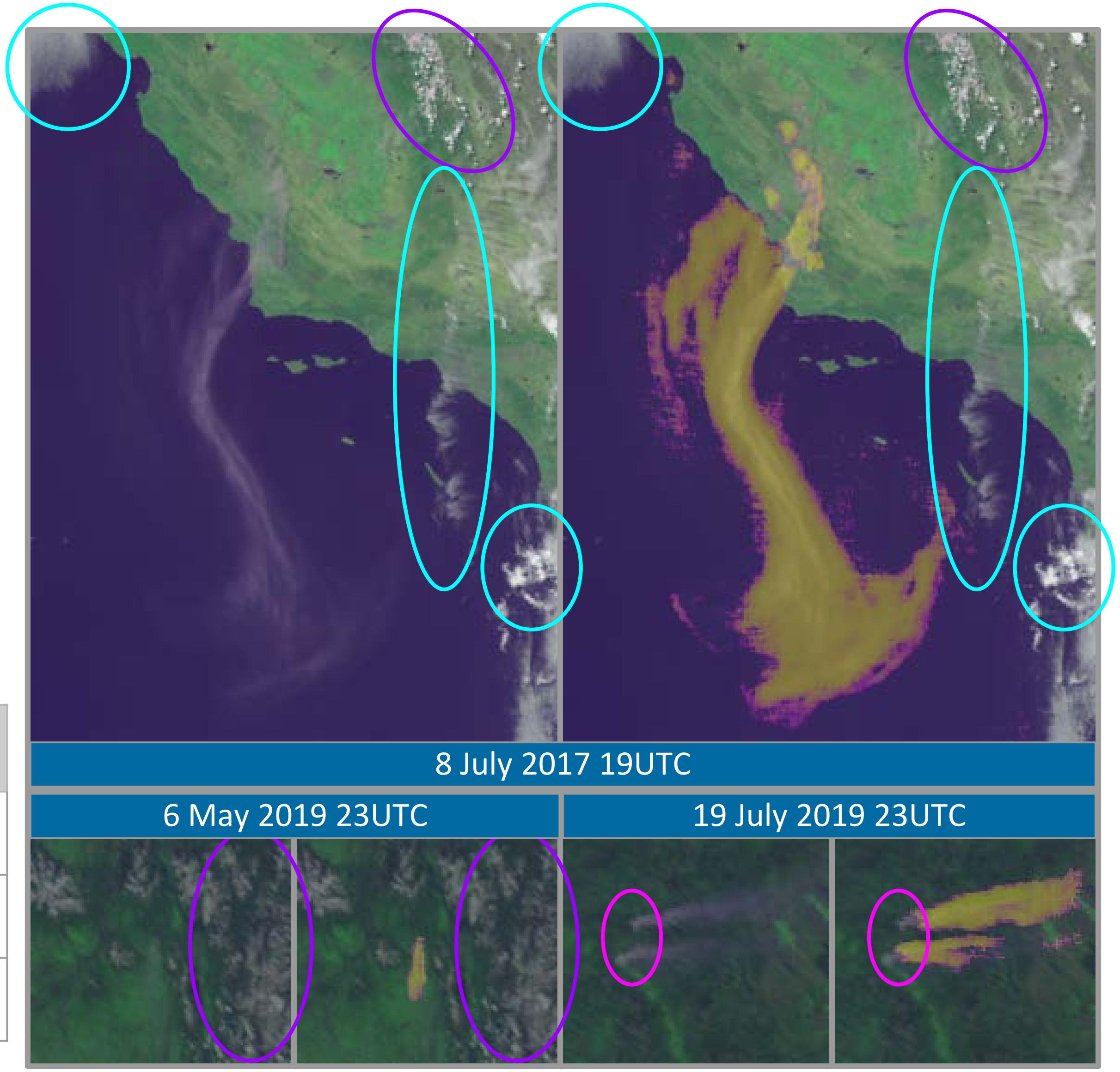
Discriminates land surface snow/ice from smoke

- Snow capped mountains

Struggle with thick smoke

- Near source; pyrocumulus

	Precision	Recall	F1-Score	Accuracy
8 Jul	0.932	0.735	0.822	0.938
6 May	1.0	0.173	0.295	0.880
19 Jul	0.995	0.475	0.643	0.870



Detection Challenges

Optically thin smoke over high reflectance surfaces

Smoke not detected at very low sun angles

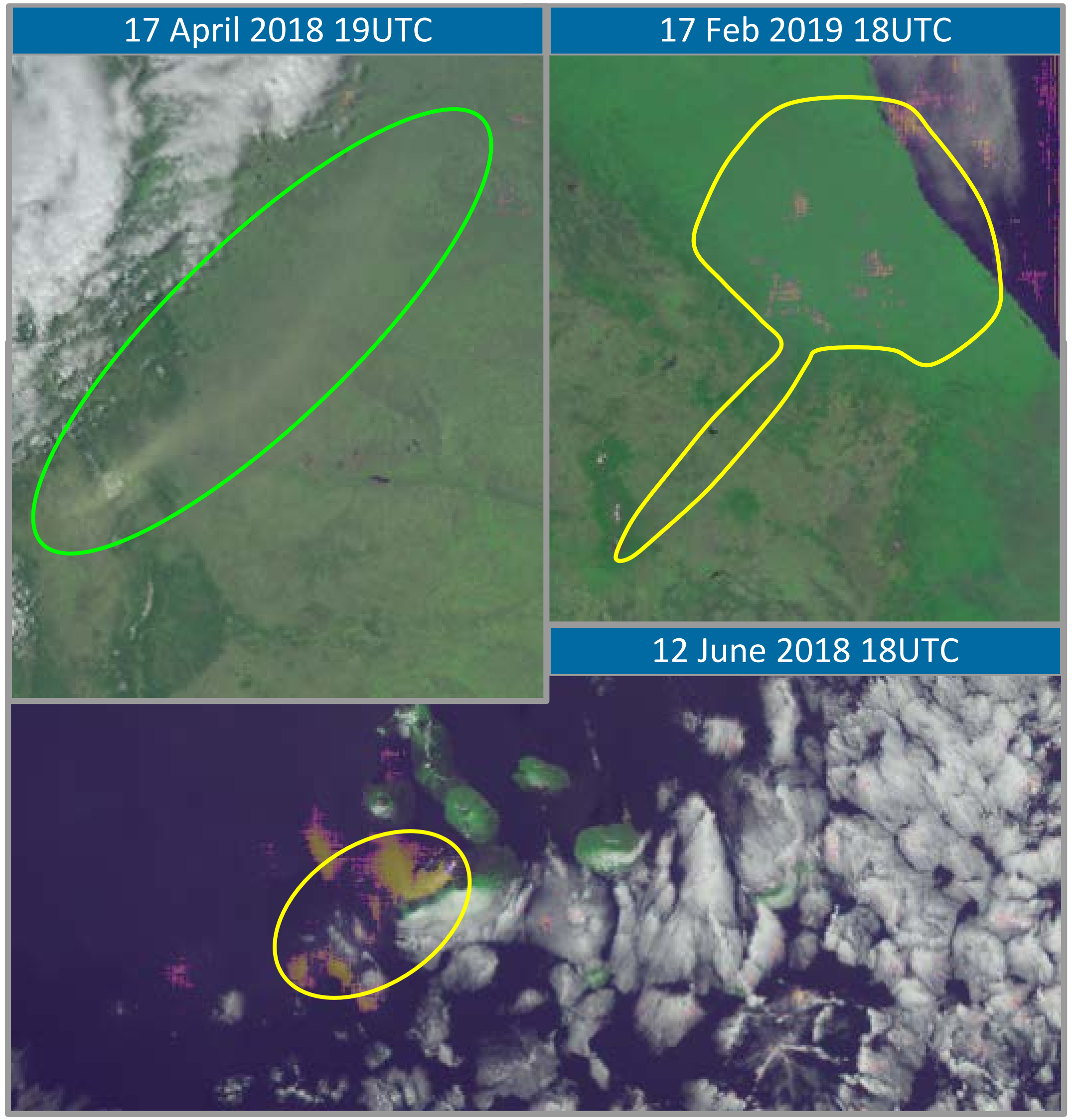
- Compounded by low optical thickness over relatively high reflective surface
- Probability of being smoke is low for few pixels that are identified

	Precision	Recall	F1-Score	Accuracy
14 Jun	0.990	0.412	0.582	0.825
20 May	0.997	0.040	0.077	0.812

Atmospheric Aerosols

A quality detection model must be able to distinguish smoke from other atmospheric aerosols

- Dust
  - Commonly found in regions influenced by smoke
  - Typical particle size spectrum different than smoke therefore model learns differences
- Volcanic Ash
  - Similar characteristics to smoke; mixed performance in testing

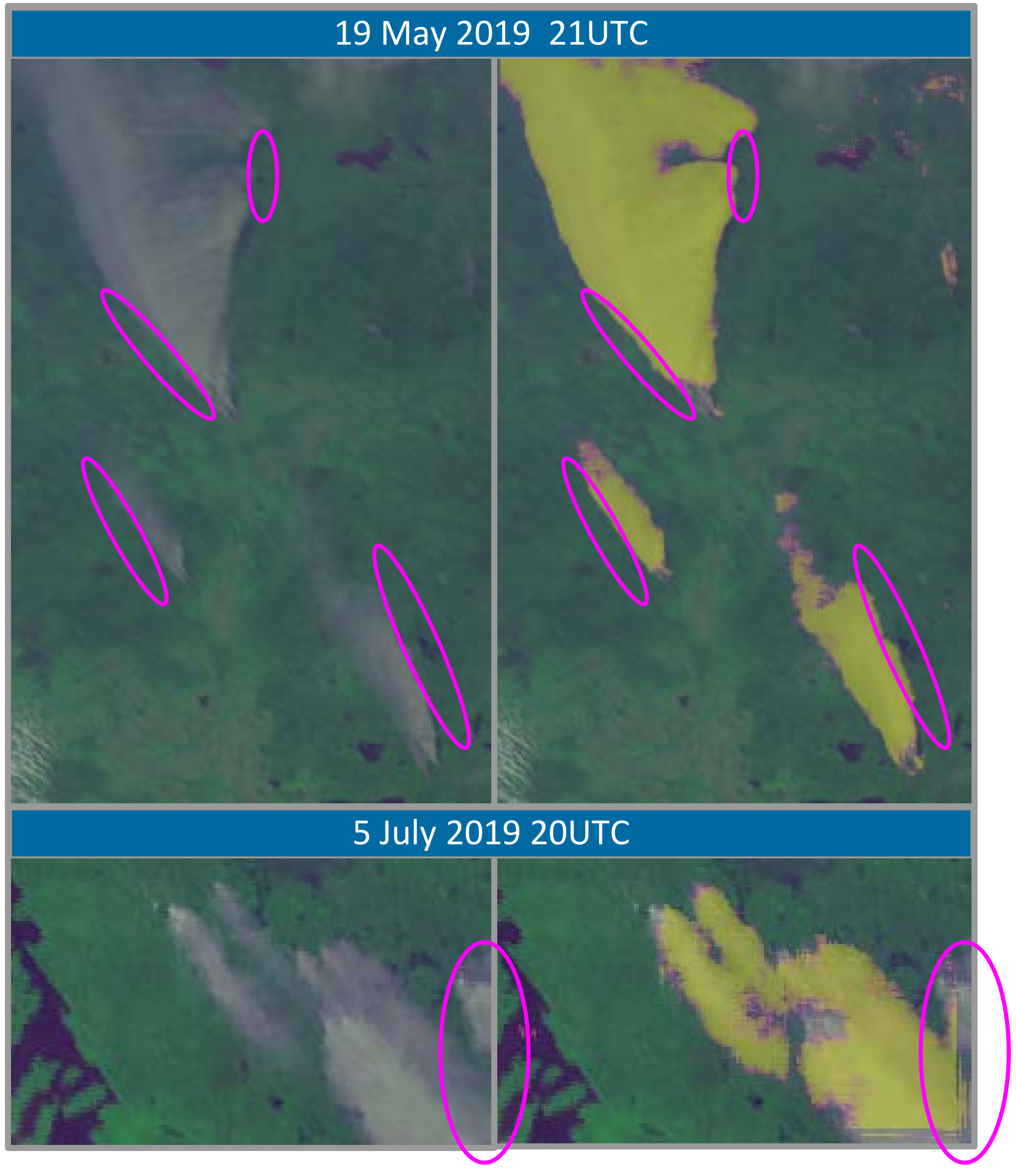


Model Design Errors

Over- and underprediction of smoke due to model design

- Artifact of a large neighborhood size
- Overprediction at plume edges results in non-zero floor in number of false positives (decreases precision)
- Underprediction at plume edges results in non-zero floor to false negatives (decreases recall)

	Precision	Recall	F1-Score	Accuracy
19 May	0.986	0.609	0.753	0.869
5 Jul	0.995	0.557	0.714	0.811



Ongoing Efforts

Expand training data to account for identified weaknesses

- Low sun angles
- Thin smoke over arid regions
- Extremely thick smoke
- Thin clouds

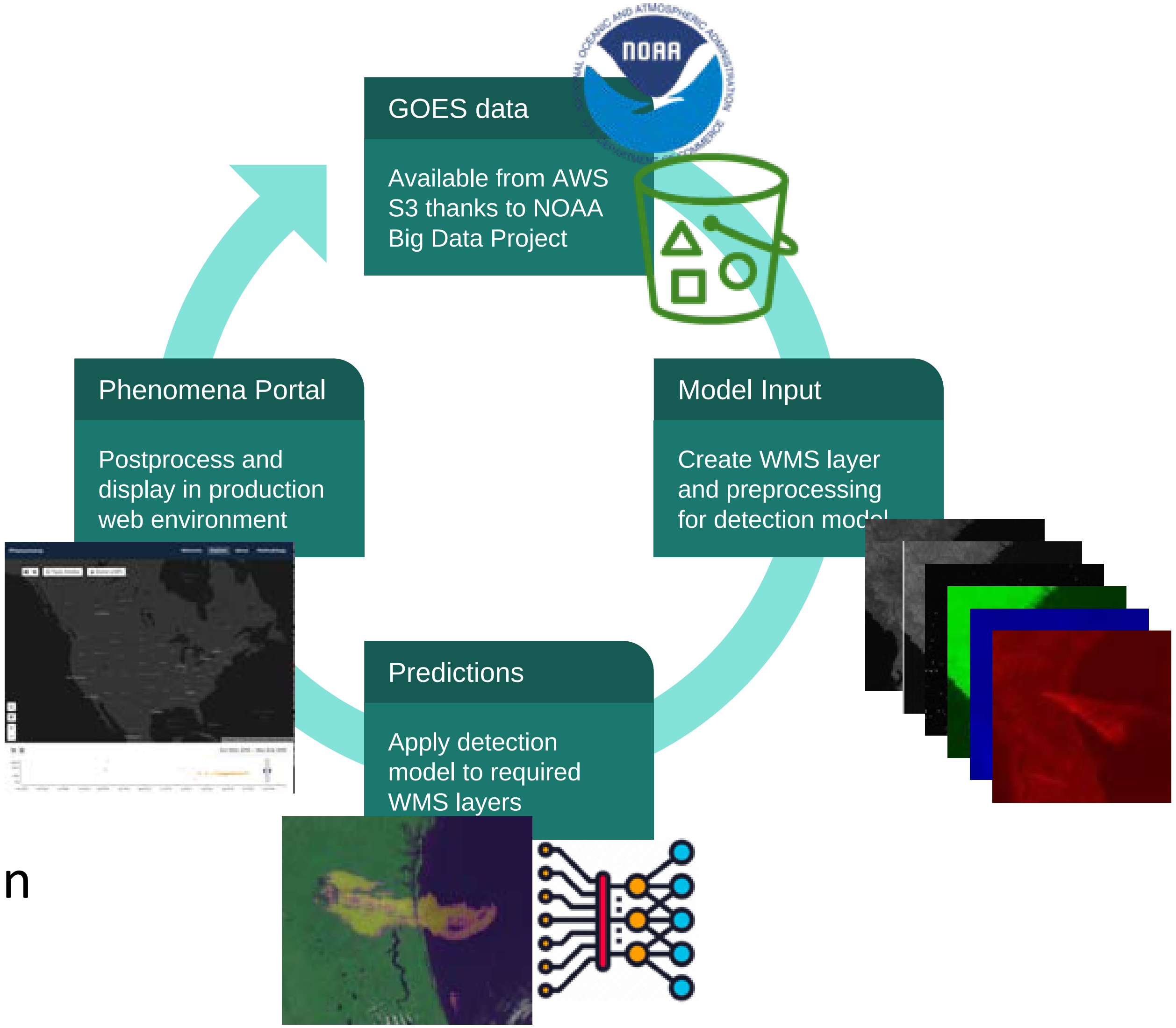
Refinement of the machine learning model

- Confirmation of N=7 as best performing model
  - Explore trade-off between neighborhood size and prediction capabilities
- Systematic approach for selecting initial model weights
- Stepwise band selection considering all 16 GOES ABI bands
- Robust model validation
  - Band exclusion to identify contribution to feature learning
  - Visualization and quantification of model learned features

Operational Capabilities

Currently testing implementation of an end-to-end analysis and visualization pipeline to a NRT production environment

- Model predictions available within full disk GOES 16 10 minute operational interval
  - ~2 min after data availability on AWS
- Plumes visualized with geojson representation of plume extents made available for download in the Phenomena Portal (http://phenomena.surge.sh)
- Fully deployed in the cloud using Amazon S3 and Cloud Computing Services



Questions?

ak0033@uah.edu

Interagency Implementation and Advanced Concepts Team

Backup

IMPACT
Interagency Implementation and Advanced Concepts Team

Development Testing

N	Precision	Recall	F1-Score	Accuracy
1	0.654	0.328	0.437	0.897
3	0.650	0.384	0.483	0.900
5	0.724	0.449	0.554	0.912
7	0.835	0.419	0.558	0.919
9	0.639	0.498	0.560	0.905

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + TN + TP + FN}$$

$$F1\ Score = 2 * \frac{Precision\ * Recall}{Precision + Recall}$$

- The F1 Scores, balance between Precision and Recall, for N=5,7,9 is comparable
  - Trade-off between quality and quantity of smoke predictions
- Best model has low false positive detection rate which drives high precision
  - Prefer conservative identification over incorrect classification
- Accuracy artifact of large number of True Negatives