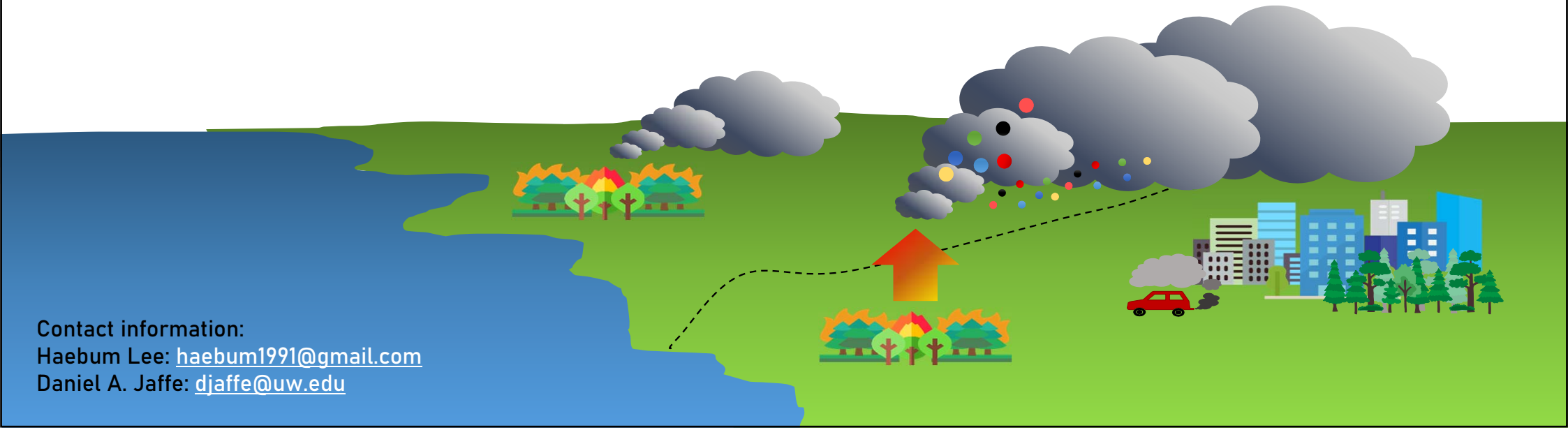
Jaffe research group
School of Science, Technology, Engineering and Mathematics (STEM)
University of Washington (UW)

WRAP/WESTAR Exceptional Events workgroup virtual meeting: GAM tool for smoke O_3 analysis, UW

Haebum Lee and Daniel A. Jaffe

January 15, 2025



Contact information:
Haebum Lee: haebum1991@gmail.com
Daniel A. Jaffe: djaffe@uw.edu

Our goals: Quantifying the impact of wildfire emissions on surface O_3

Generalized additive model (GAM)

Quantify smoke O_3

$O_{3,Obs} \uparrow = O_{3,Local} + O_{3,Smoke} \uparrow ?$

Smoke plume

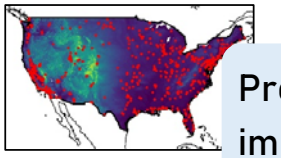
Due to smoke?

$O_{3,Obs} \uparrow$ smoke day? (HMS + PM2.5)

Wildfires

Single site

Scale up



CONUS

In previous work we have… (Gong et al., 2017; McClure et al., 2018; Lee and Jaffe, 2023)

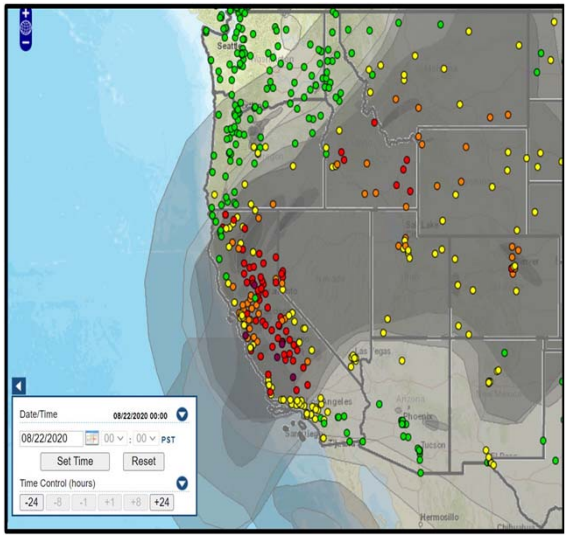
- Used surface $PM_{2.5}$ and HMS (hazard mapping system) smoke product to identify "smoke days" at individual monitoring sites for every day in different urban areas (SLC, Houston, Boise, Denver, etc.)
- Use Generalized Additive Models (GAMs) to predict O_3 for non-smoke days based on meteorology and other predictors
- From the GAM residual, estimate the enhancement in O_3 due to smoke (smoke O_3) for smoke days

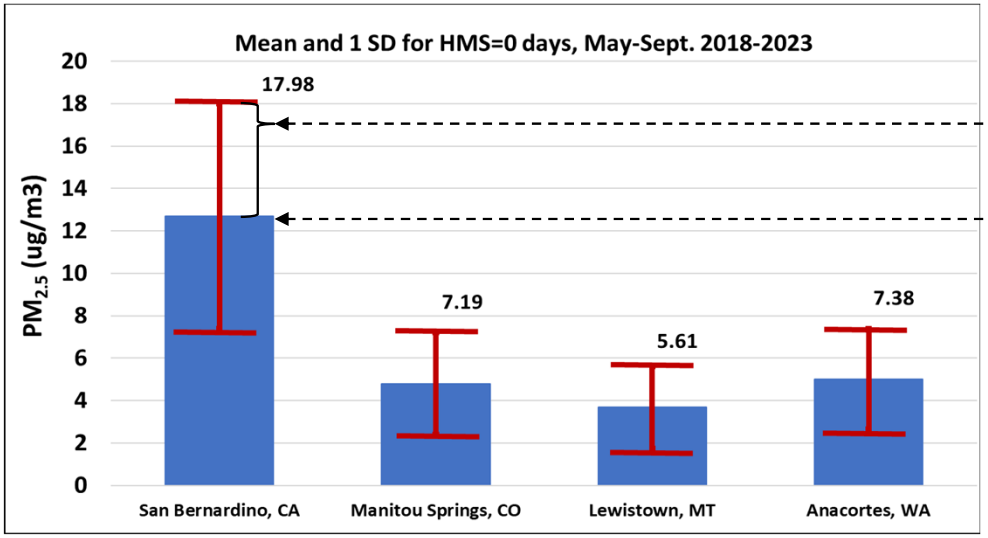
Previous studies have primarily focused on quantifying the impact of wildfires on O_3 at single-site or localized regional scales. In this work, we apply this method to every AQS site in the US with sufficient data.

- GAM version 1 (Lee and Jaffe, 2024): 606 sites (May–Sept, 2018–2023)
- GAM version 2 (this study): 802 sites (Apr–Oct, 2019–2023)

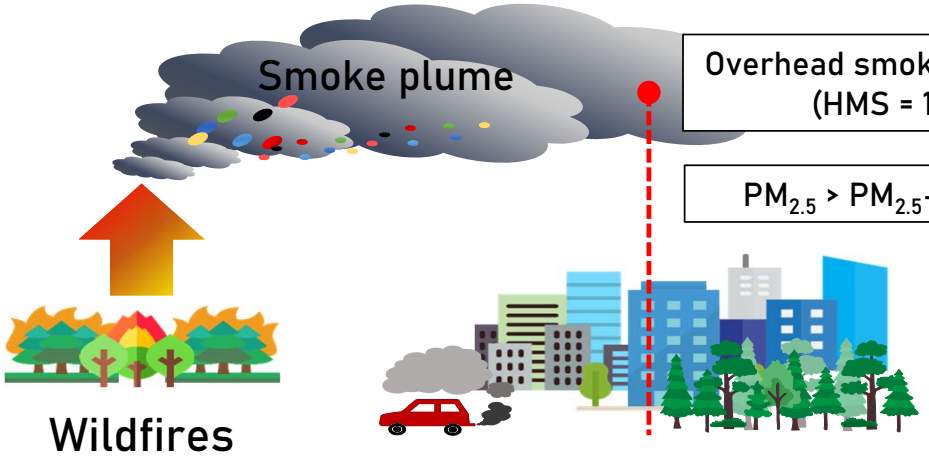
Use HMS smoke and surface $PM_{2.5}$ to identify smoke days

HMS Smoke plume + $PM_{2.5}$-criteria





$PM_{2.5,1\ SD}$ + $PM_{2.5,mean}$ = $PM_{2.5}$-criteria (17.98 µg m^{-3})



"Smoke day"
= Day with detected
*HMS smoke (HMS = 1) and $PM_{2.5}$ > *$PM_{2.5}$-criteria*

*$PM_{2.5}$-criteria

= Mean + 1 SD of HMS = 0 days (Lee and Jaffe, 2024)

= Median + 1 MAD (median absolute deviation) of HMS = 0 days (this study)

How to quantify the O_3 due to smoke?
→ Use a statistical or machine learning model to predict MDA8 O_3 w/wo smoke.

Generalized additive model (GAM)

$$g(y_i) = \beta_0 + e_i + \sum_{j=1}^{n} s(x_{ji})$$

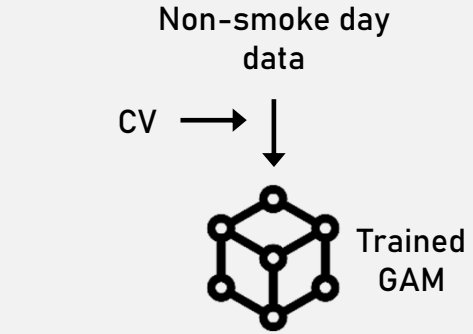
Main advantages

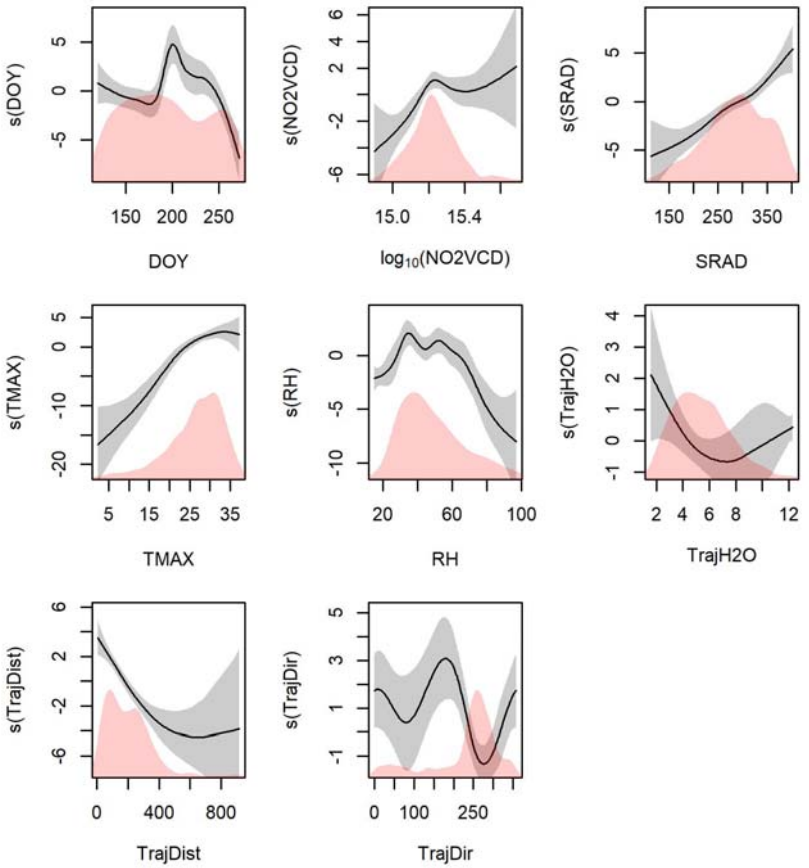
- Intuitive interpretations
- Incorporate linear, non-linear and categorical relationships

Typical predictors for O_3

- Daily maximum temp (TMAX)
- Relative humidity (RH)
- Solar radiation
- Wind direction/speed
- Trajectory distance/direction
- …

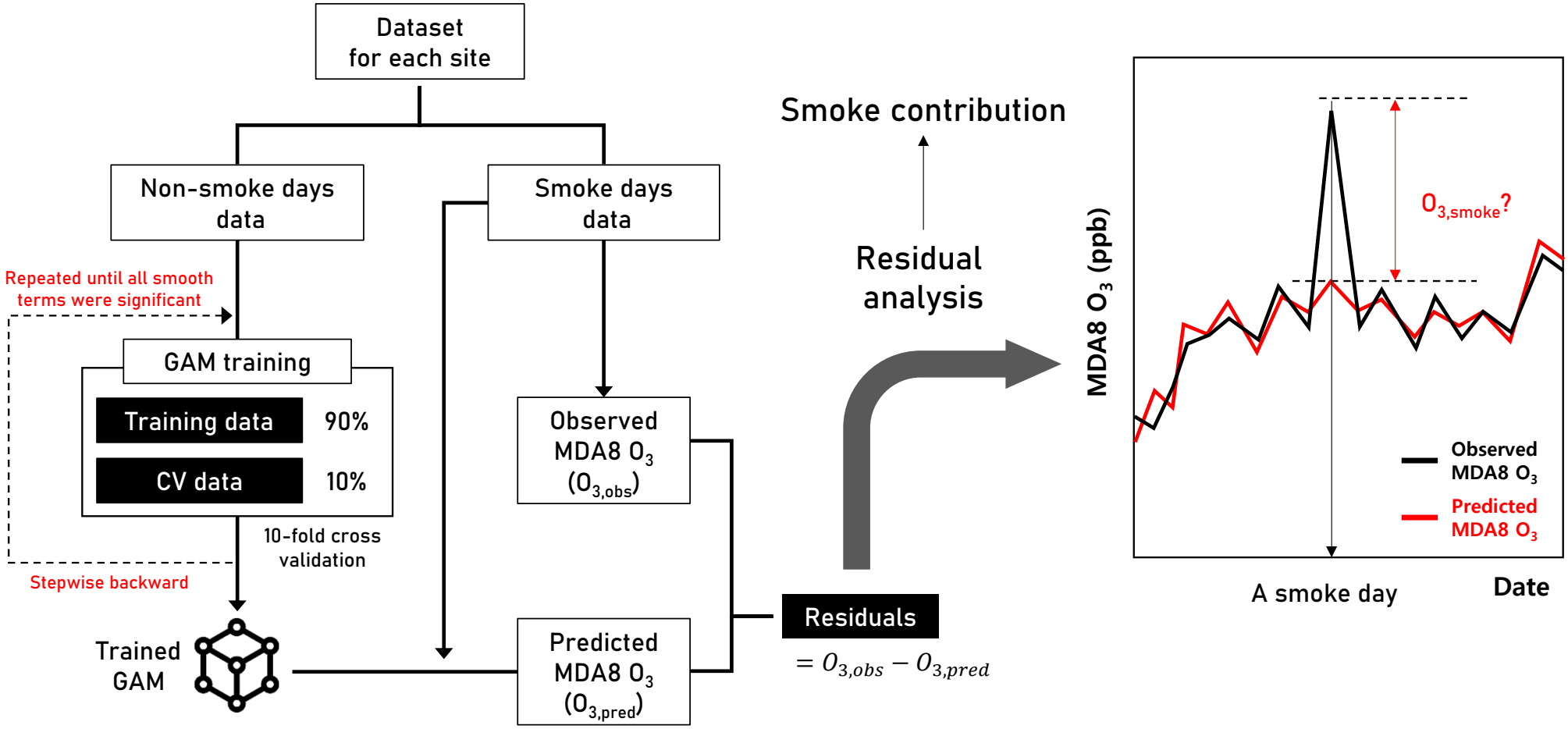
(Camalier et al 2007; CARB 2011; Sun et al 2015; Gong et al 2017; 2018; McClure and Jaffe 2018; Jaffe et al 2018; 2021; Lee et al 2023)



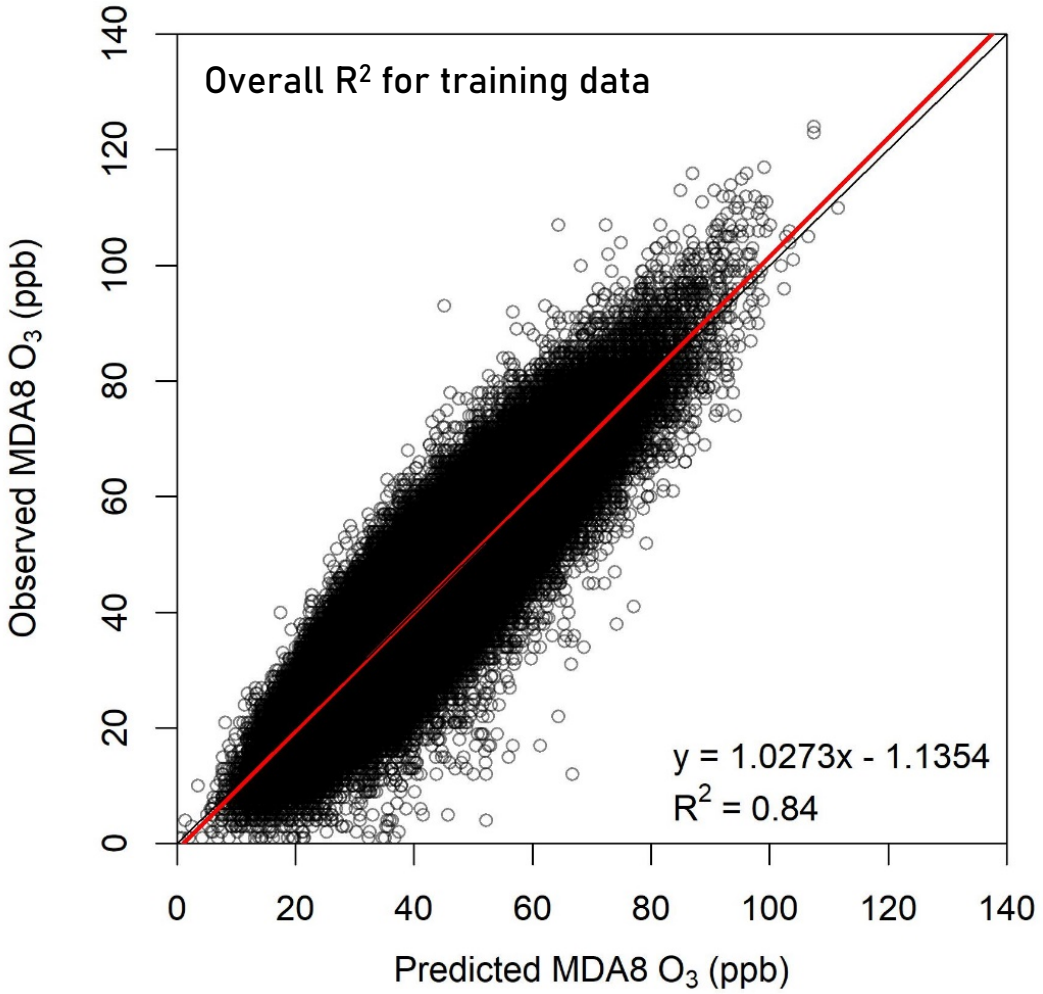


- Smoke is not include in the model, so we attribute difference between modeled and predicted to smoke influence

GAM procedure



Observed vs GAM predicted MDA8, 802 sites, training data (non-smoke), Apr-Oct, 2019-2023



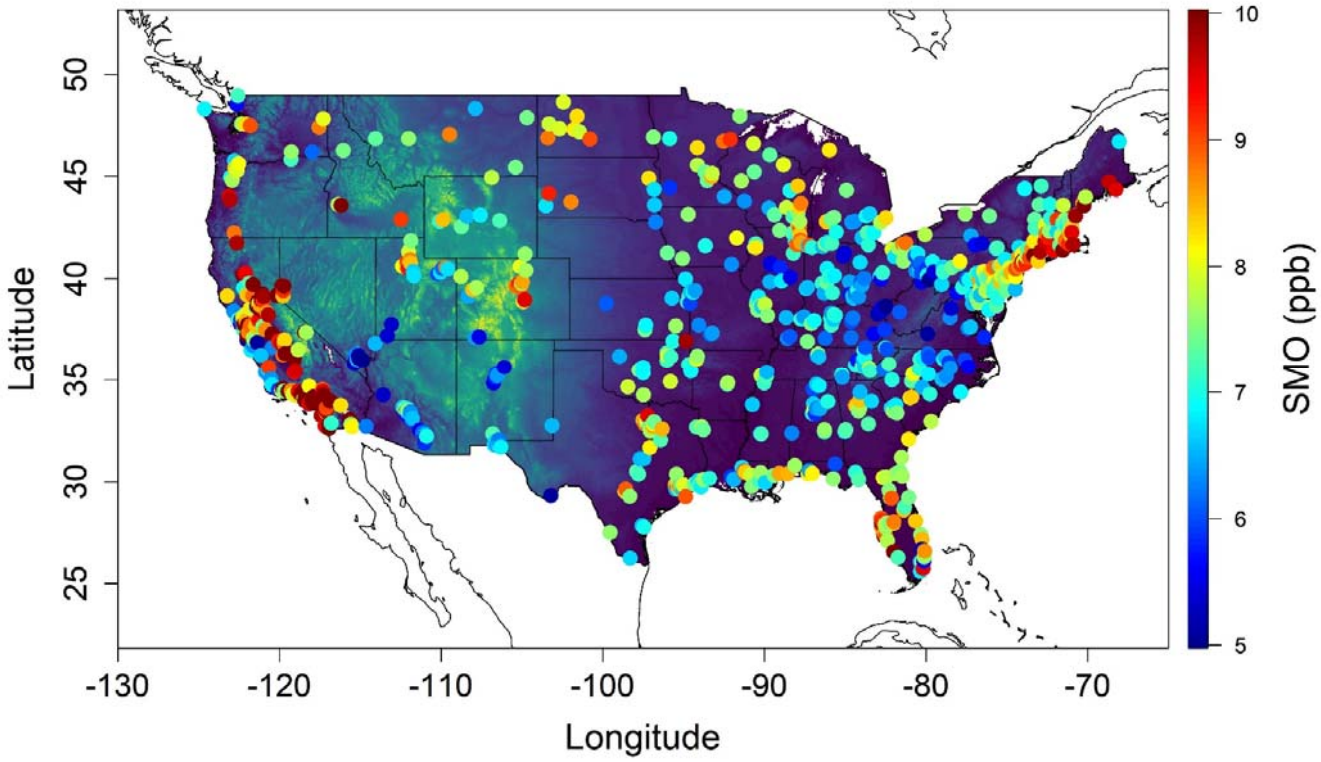
	Training	CV
No. of data	633760	633760
R^2 (all data)	0.84	0.74
Residual (Mean ± SD)	0.00 ± 4.78	0.00 ± 6.16

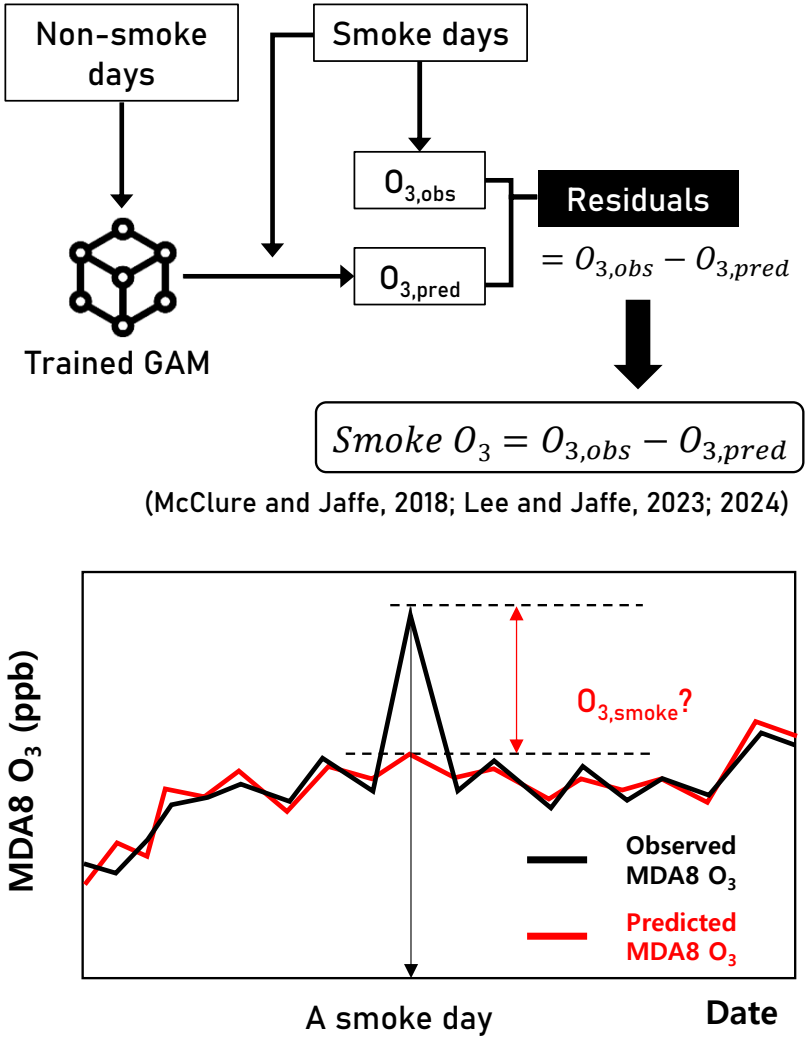
If looking at the average R^2 for the entire site…

	Training	CV
No. of sites	802	802
R^2 (mean all sites)	0.77 ± 0.05	0.62 ± 0.07

Smoke contributions to the O_3 MDA8
Smoke day's residual = Smoke O_3 = SMO

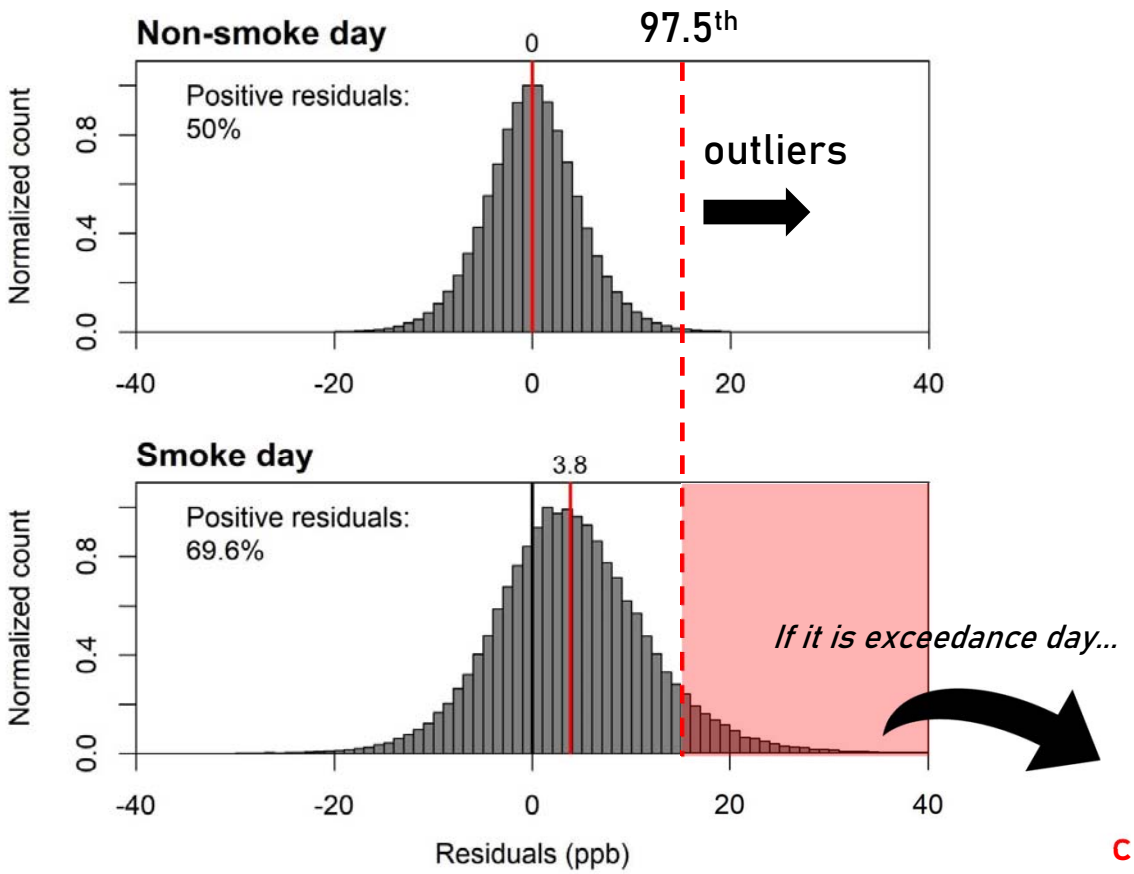
5-year average positive SMO for each site

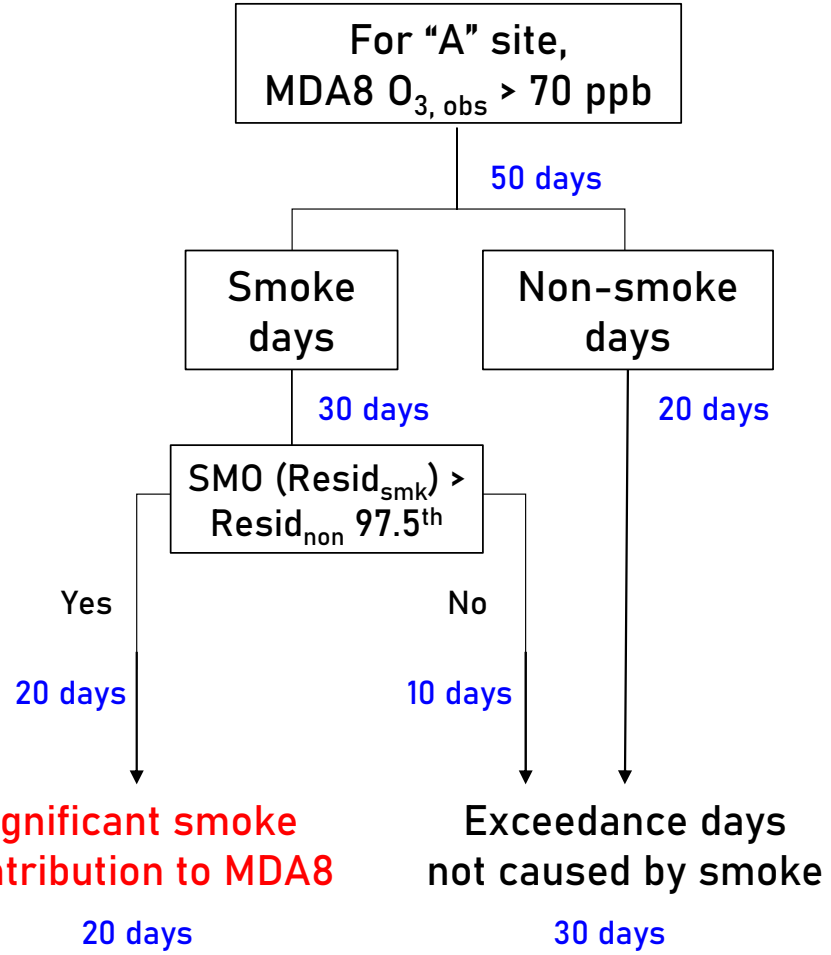




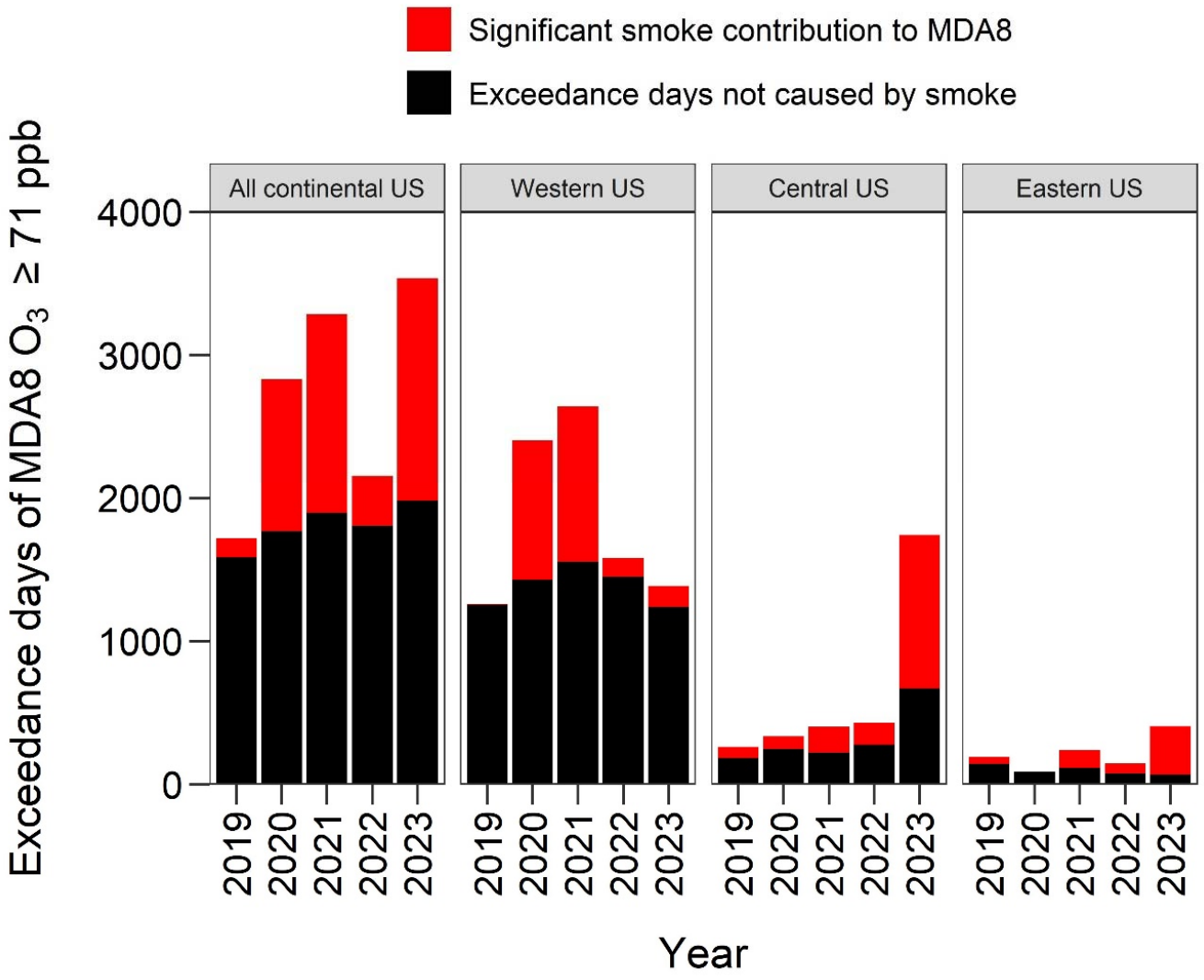
- A positive enhancement to the MDA8 O_3 on 70% of the days with smoke
- The mean positive SMO was 7.6 ± 6.0 ppb for Apr-Oct, 2019–2023

GAM residual and Exceedance days (> 70 ppb)





Number of exceedance days with/without smoke contribution?



- Overall exceedance days due to smoke was ~33% of total number of exceedance days
- Our results demonstrate the importance of wildfires as contributors to exceedances of the health based national air quality standards for O_3

Explore our results using R shiny - GAM Tool

GAM tool

- The purpose of this app is to demonstrate the application of General Additive Modeling (GAM) for predicting MDA8 O_3 using meteorological data under both smoke and non-smoke conditions. In this app, a GAM is derived using the 'mgcv' package in R. We hope this tool will be useful for state agencies and others in understanding the factors that control ozone production.

GAM manual

Data collection

- EPA O3 (44101)
- EPA PM2.5 (88101, 88502)
- Met data (from IEM)

HMS & $PM_{2.5}$-criteria

- HMS data
- Calculation $PM_{2.5}$-criteria
- Identify “smoke day”

Load dataset

- Load data from collected or your own dataset

GAM setup

- Build your GAM: response (y) and predictors (x)
- Set GAM parameters

GAM results

- Explore your GAM results

GAM previous

Overview

- Model summary
  - Metadata
  - Obs. vs Pred.
  - Time-series
  - Smoke frequency
  - Exceedance day
  - Export data

Layer map

- Spatial temporal data
  - MDA8
  - Residuals
  - SMO
  - $PM_{2.5}$
  - $PM_{2.5}$-criteria
  - HMS-smoke
  - HMS-fire
  - Export data

Data comparison

- Comparison between two datasets
  - Observed MDA8
  - Predicted MDA8
  - Residual
  - SMO
  - Positive SMO
  - PM2.5

- Current available dataset
  - GAM v1 (Lee and Jaffe, 2024)
  - GAM v2 (this study)
  - GAM v2-EDM (this study)
  - EPA EMBER (Simon et al., 2024)
- To be added later
  - Health Impact Assessments for $PM_{2.5}$ and O_3
  - Update with 2024

https://smoke.shinyapps.io/rsGAM/