10.1071/WF23148
International Journal of Wildland Fire

Supplementary Material

Is the smoke aloft? Caveats regarding the use of the Hazard Mapping System (HMS) smoke product as a proxy for surface smoke presence across the United States

Tianjia Liu[A,G,*], *Frances Marie Panday*[B], *Miah C. Caine*[C], *Makoto Kelp*[A], *Drew C. Pendergrass*[D], *Loretta J. Mickley*[D], *Evan A. Ellicott*[B], *Miriam E. Marlier*[E], *Ravan Ahmadov*[F] *and Eric P. James*[F]

[A]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

[B]Department of Geographical Sciences, University of Maryland, College Park, MD, USA

[C]Department of Computer Science, Harvard University, Cambridge, MA, USA

[D]John A. Paulson School of Engineering, Harvard University, Cambridge, MA, USA

[E]Department of Environmental Health Sciences, University of California, Los Angeles, Los Angeles, CA, USA

[F]Global Systems Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, USA

[G]Department of Geography, University of British Columbia, Vancouver, BC, Canada

*Correspondence to: Email: tianjia.liu@ubc.ca

ISD and HMS smoke days and trends at airport locations

Table S1. Number of ISD airports with statistically significant (p-value < 0.05) trends in smoke days per year from 2010-2011.

Region	ISD	HMS all	HMS medium/heavy	HMS heavy	Total
CONUS	493	386	639	1017	1598
Western U.S.	295	255	288	389	614
Alaska	16	14	9	0	108

Table S2. Average smoke days (± 1D) per year from 2010-2021 by region.

Region	ISD	HMS all	HMS medium/heavy	HMS heavy
Western U.S.	7.1 ± 4.5	36.2 ± 18.5	10.7 ± 9.8	3.7 ± 4.7
Pacific	11.6 ± 8.2	27.5 ± 18	11.1 ± 11.2	5 ± 6.6
Mountain	3.7 ± 2.8	29.4 ± 19.8	10.8 ± 11.3	4.1 ± 5.6
West North Central	3.8 ± 2.5	56.1 ± 20.6	17.8 ± 13.1	5.8 ± 7.2
East North Central	3.6 ± 2.5	44.8 ± 17.1	12.9 ± 10.2	3.9 ± 4.4
Northeast	2.5 ± 1.8	25.8 ± 15	6.1 ± 5.6	1.4 ± 1.8
West South Central	7 ± 4.3	41.4 ± 23	7.3 ± 7.4	1.4 ± 1.8
East South Central	3.6 ± 2.3	27.9 ± 19.9	4.7 ± 5.4	0.8 ± 1.2
South Atlantic	5.1 ± 2.6	21.6 ± 17.7	3.1 ± 3.8	0.7 ± 1
Alaska	2.4 ± 2.2	10.7 ± 8.8	2.7 ± 3.4	0.8 ± 1.6

Table S3. Linear trend in smoke days per year from 2010-2021 by region. The slope is shown with the standard error in parenthesis.

Region	ISD	HMS all	HMS medium/heavy	HMS heavy
Western U.S.	1.1 (0.2) *	2.9 (1.3)	1.7 (0.7) *	0.9 (0.3) *
Pacific	2 (0.4) *	4 (0.9) *	2.3 (0.7) *	1.4 (0.4) *
Mountain	0.7 (0.1) *	3.8 (1.2) *	2.1 (0.7) *	1.1 (0.4) *
West North Central	0.5 (0.2) *	2.1 (1.7)	2 (1)	1.3 (0.5) *
East North Central	0.3 (0.2)	1.9 (1.4)	1.7 (0.7) *	0.9 (0.3) *
Northeast	0.3 (0.1)	2.6 (1) *	1.1 (0.4) *	0.4 (0.1) *
West South Central	1 (0.2) *	1.4 (2)	0.7 (0.6)	0.3 (0.1)
East South Central	0.4 (0.2) *	2.6 (1.5)	0.7 (0.4)	0.2 (0.1) *
South Atlantic	0.5 (0.2) *	1.8 (1.4)	0.4 (0.3)	0.1 (0.1)
Alaska	0.4 (0.1) *	1 (0.7)	0.4 (0.3)	0.1 (0.1)

* p-value < 0.05

EPA $PM_{2.5}$ monitors

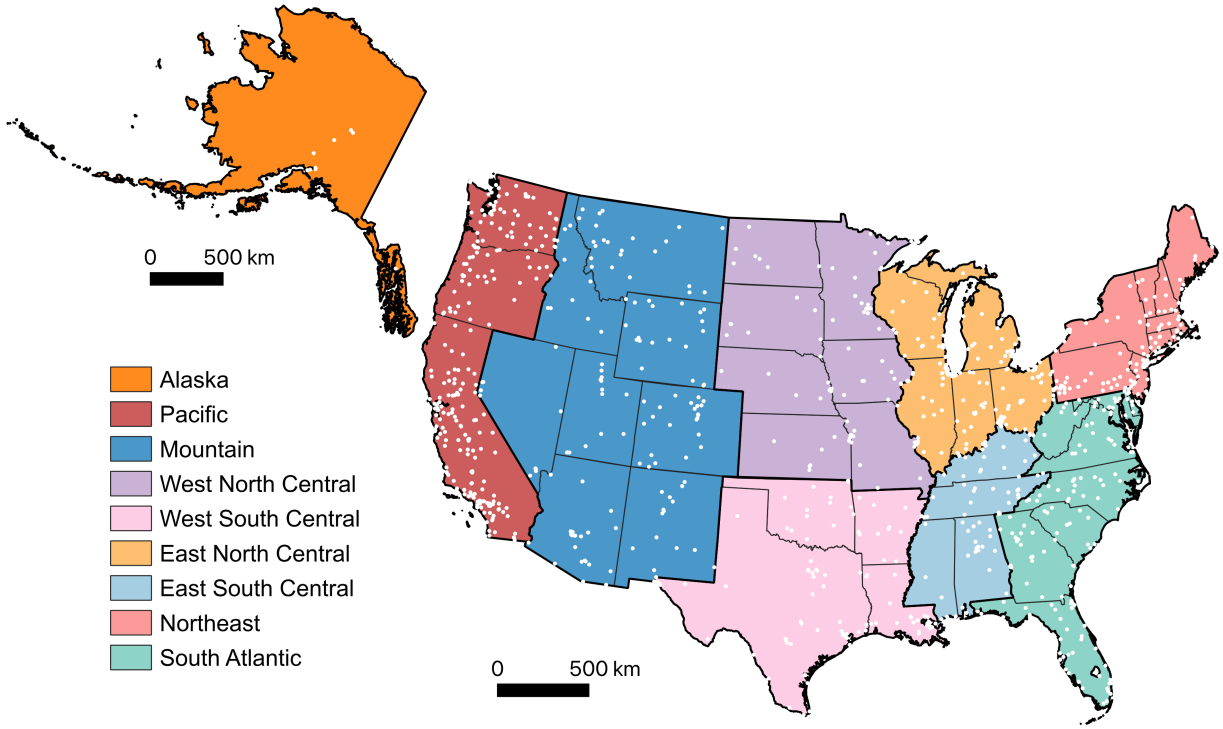

Fig. S1. Map of CONUS regions and Alaska with EPA $PM_{2.5}$ monitor locations. Each white dot represents the location of EPA $PM_{2.5}$ monitors used in this study. (Note that Alaska is not shown on the same scale as CONUS.)

Assessing uncertainty in HMS-based smoke $PM_{2.5}$ estimates

Table S4. Confidence categories for defining lower and upper bounds in smoke $PM_{2.5}$ estimation based on HMS smoke density categories and $PM_{2.5}$ anomalies as percentiles relative to the distribution of $PM_{2.5}$ anomalies on non-smoke days

Confidence	**HMS** (smoke density categories)	**EPA** (percentiles of non-smoke distribution of $PM_{2.5}$ anomalies)
High (lower bound)	Heavy-only plumes	> 85%
Medium	Medium and heavy plumes	> 70%
Low (upper bound)	All plumes	> 50%

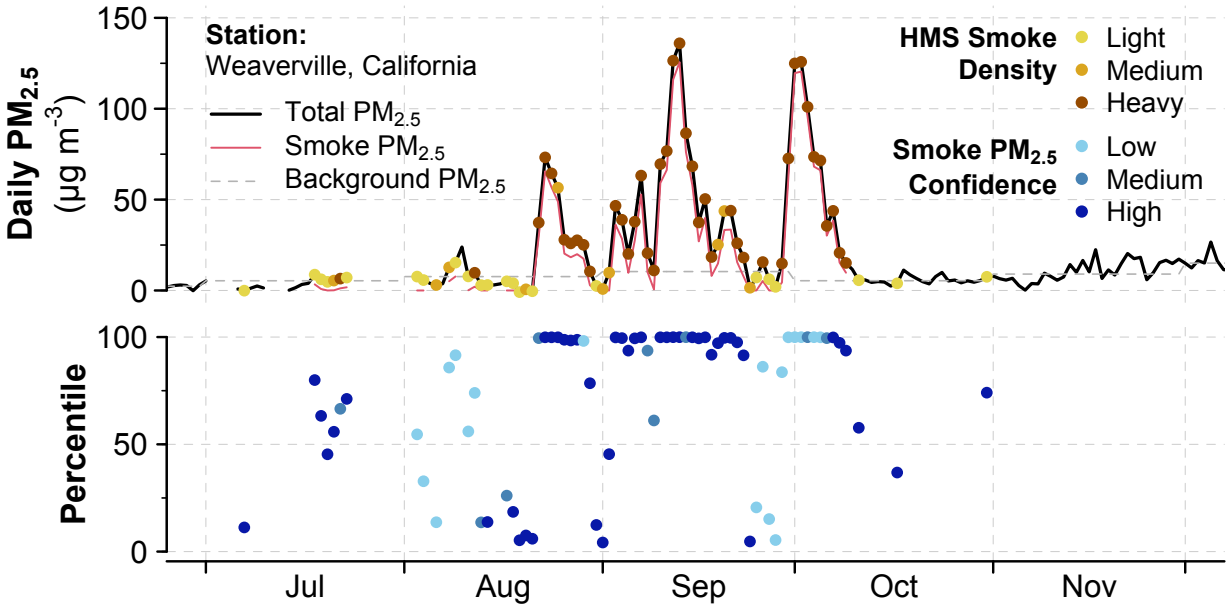

Fig. S2. Example of $PM_{2.5}$ observations with confidence levels for an EPA monitor in Weaverville, California from July to November 2020. The maximum HMS smoke plume density (*top*), percentile of $PM_{2.5}$ anomalies relative to the distribution of $PM_{2.5}$ anomalies on non-smoke days (*bottom*), and maximum confidence level category (*bottom*) are shown alongside the total, smoke, and background $PM_{2.5}$. Confidence level categories (low, medium, high) associated with each smoke $PM_{2.5}$ estimates are defined in Table S4. The average smoke $PM_{2.5}$ from July-November ranges from 13.1 µg m^{-3} for low confidence (upper bound) to 13 µg m^{-3} for medium confidence to 12.2 µg m^{-3} for high confidence (lower bound).

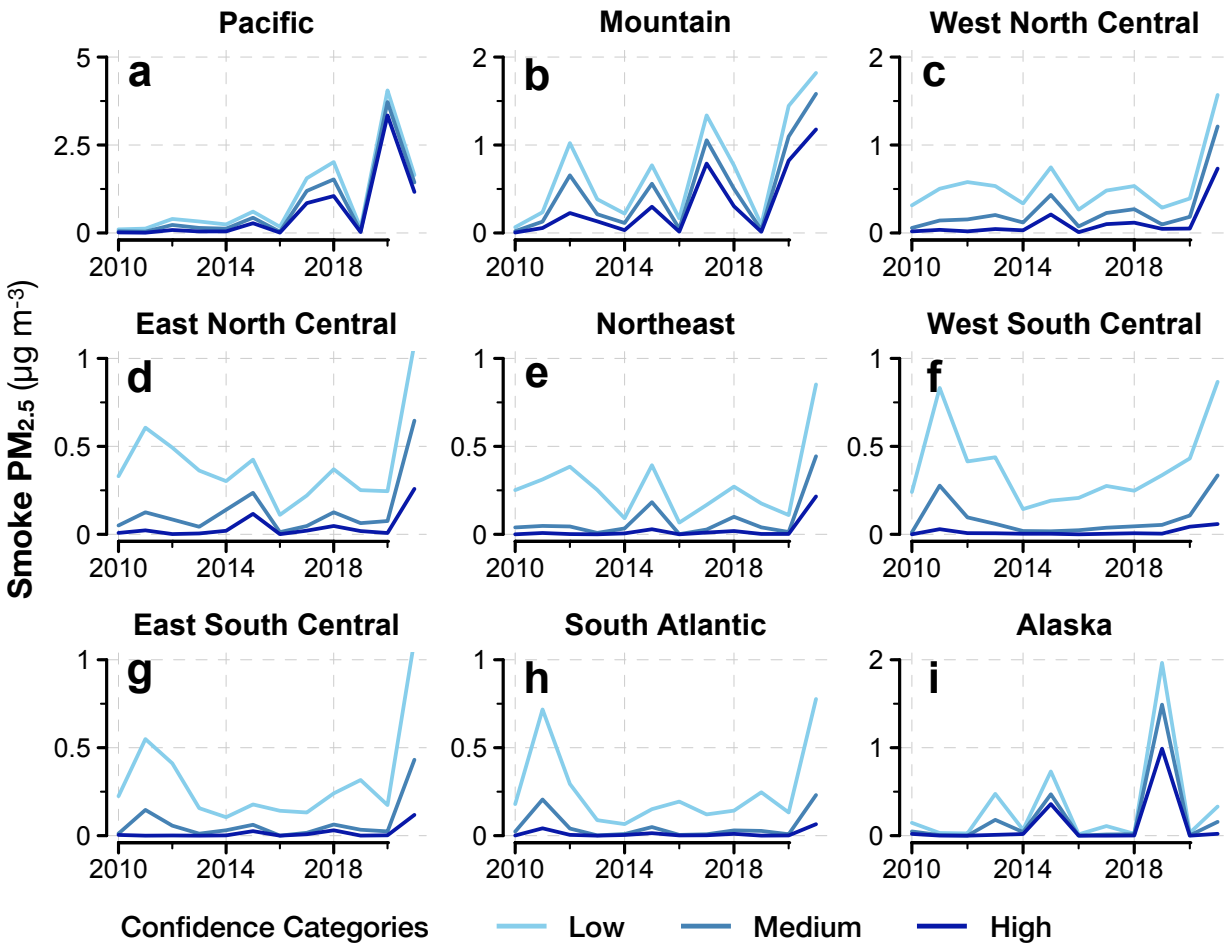

Fig. S3. Annual average smoke $PM_{2.5}$ from EPA monitor data from 2010-2021 at different confidence levels and by region. Confidence level categories (low, medium, and high) are defined based on HMS smoke densities and percentiles of $PM_{2.5}$ anomalies relative to the distribution of $PM_{2.5}$ anomalies on non-smoke days, as defined in Table S4.

Gap-filling missing HMS smoke densities using a random forest model

Starting from 2008, each polygon in the HMS dataset is consistently assigned a smoke density category, but there is a data gap from late 2008 to early 2010 when the density for 35,828 polygons is unspecified, possibly due to an error in the data archiving process. To fill this data gap, we train a random forest model on the density labels of smoke polygons from 2008-2021. For classification, the random forest algorithm is based on the majority vote of an uncorrelated ensemble of decision trees (Breiman 2001). Each decision tree is individually fit to a random bootstrap sample of the training data and features, or input variables. Decision tree training is recursive, splitting data into branches via an optimal split point determined from the features. Individual decision trees have high error variance but no inherent bias, so averaging many individual and uncorrelated trees yields a low variance, low bias prediction.

We use the following independent variables derived from HMS metadata and satellite data to model the density category: month, time of day of the first and last GOES image used to draw the polygon ("start" and "end"), duration of the animated set of images used to draw the polygon ("duration"), area of polygon ("area"), average Aerosol Optical Depth (AOD) within the polygon ("AOD"), and fraction of overlap with other polygons on the same day ("overlap") (Table S5). For AOD, we use the MODIS Multi-angle Implementation of Atmospheric Correction (MAIAC) product (MCD19A2, Collection 6) at 0.55 μm (Lyapustin *et al* 2018). MAIAC operates on a fixed 1-km grid and combines the advantages of the MODIS Dark Target and Deep Blue algorithms that specialize on dark vegetative and bright desert surfaces, respectively. The "overlap" variable takes advantage of the nested nature of the smoke polygons; that is, heavy smoke plumes are located within medium smoke extent, and medium smoke plumes are located within light smoke extent (Brey *et al* 2018). We calculate the fractional area of each smoke polygon that overlaps with other polygons from the same day. Medium and heavy smoke polygons have relatively high overlap, and light smoke polygons low overlap.

We train two random forest models with and without AOD. Some HMS polygons (n = 525) had missing AOD values due to cloud coverage preventing successful AOD retrievals. We use the model trained with AOD to gap-fill over 98% (n = 35303) of the unspecified densities, while we use the model trained without AOD to gap-fill the remaining unspecified densities. For 1000 bootstrap iterations, we undersample the light and medium categories so that all three densities are equally represented in the random forest model; we then split 2/3 of the dataset for training data and for 1/3 for test data. Without undersampling, the random forest model would prioritize the classification accuracy of light smoke, as light smoke plumes (75%) occur much more frequently than medium (18%) and heavy (8%) smoke.

The primary model, which includes all independent variables listed in Table S5, is used to gap-fill 35,303 polygons, while the secondary model, which excludes AOD, is used to gap-fill 525 polygons that have missing input AOD data. For the primary model, the test accuracy is 85% for light smoke, 58% for medium smoke, and 66% for heavy smoke (Fig. S4a). For the secondary model, the test accuracy is 83% for light smoke, 51% for medium smoke, and 67% for heavy smoke (Fig. S4b). The "overlap" variable, which specifies the fraction of overlap in one polygon with other polygons on the same day, is by far the most important variable, leading to a high mean decrease in model accuracy if that variable were excluded. The fractional overlap of a given HMS polygon with other polygons drawn at the same time is an innate property of HMS smoke product – i.e., heavy density polygons are nested within medium and light density

polygons. The lower accuracy for medium smoke relates to the weaker separation of medium smoke with light and heavy smoke by the overlap variable, which cannot distinguish between medium and heavy density polygons well if both are totally nested within a light density polygon. The mean AOD within the smoke polygon is the second most important variable; medium smoke density polygons tend to be associated with high AOD. However, clouds can obstruct AOD retrievals, and AOD values can highly vary within a polygon and throughout the day and year. MAIAC AOD relies on MODIS observations from the Terra and Aqua satellites, each of which overpass a location only once per day during daytime. Other variables, such as the start and time end of the satellite images used and polygon area, do not improve model performance much.

Table S5. Inputs and outputs of the random forest models used to gap-fill HMS smoke density labels

	Description	Format
Inputs		
Overlap	Fraction of overlap between a given polygon and other polygons in the same day	Numeric, [0-1]
AOD	Average MODIS MAIAC C6 aerosol optical depth within the smoke polygon	Numeric, [≥0] *
Start	Start time of the set of images used to delineate smoke polygon outline	Numeric, HHMM, UTC
End	End time of the set of images used to delineate smoke polygon outline	Numeric, HHMM, UTC
Duration	Duration of the set of images used to delineate smoke polygon outline, difference between start and end time	Numeric, hours
Month	Month that the smoke polygon is detected	Numeric, [1-12]
Area	Area of smoke polygon	Numeric, km^2
Outputs		
Density	HMS smoke density	Categorical, [light, medium, heavy]

* AOD values are generally ≥ 0, but small negative values are permitted in the retrievals

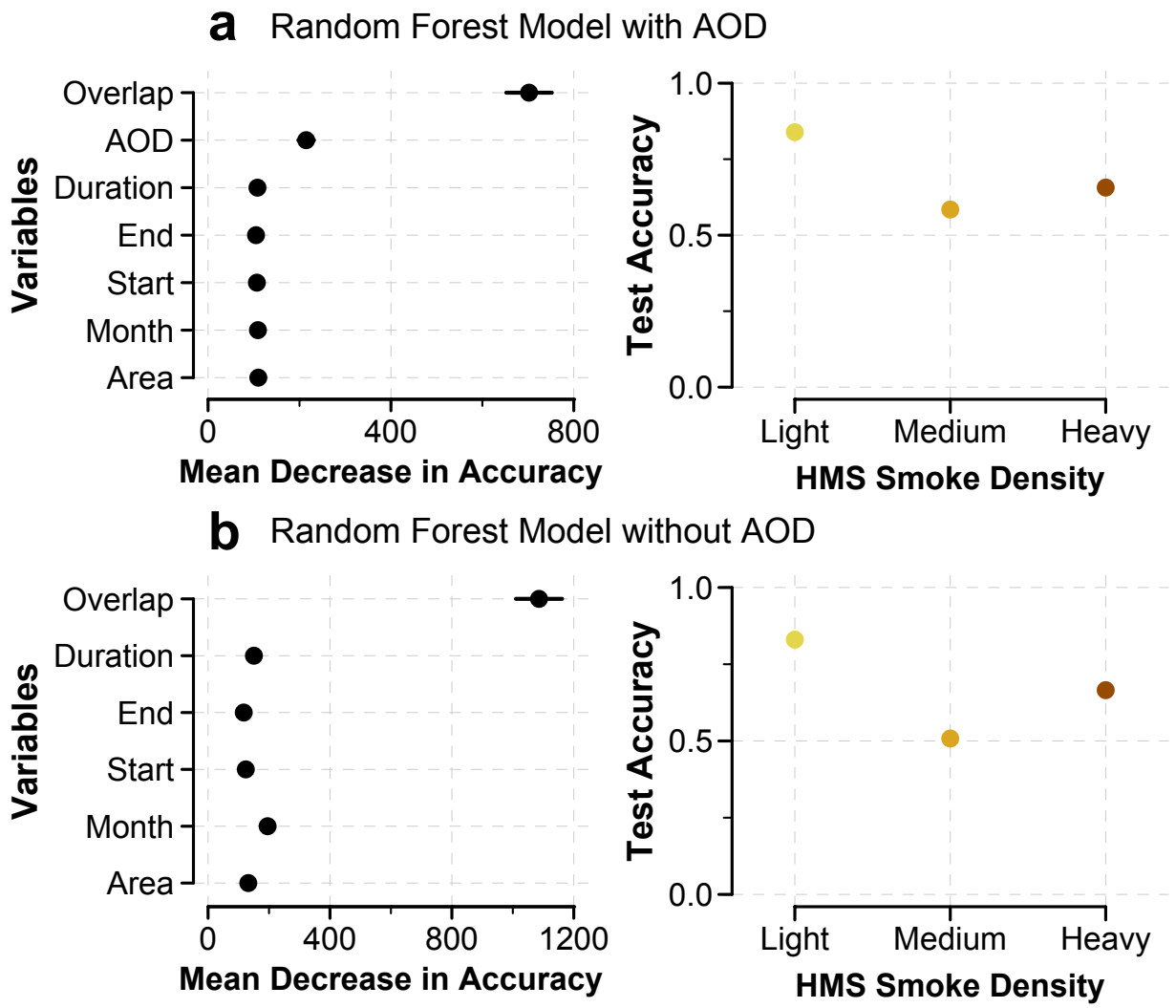

Fig. S4. Performance of random forest models for gap-filling HMS polygons with "unspecified" smoke density. Variable importance (*left*) and accuracy of the test set (*right*) for random forest models (**a**) with AOD as a predictor and (**b**) without AOD as a predictor. The plots show the average ± 1SD for variable importance and test set accuracy over 500 bootstrap iterations. Variable importance is indicated by the mean decrease in accuracy, where higher values represent more important variables.

References

Breiman L 2001 Random Forests *Machine Learning* **45** 5–32

Brey S J, Ruminski M, Atwood S A and Fischer E V. 2018 Connecting smoke plumes to sources using Hazard Mapping System (HMS) smoke and fire location data over North America *Atmospheric Chemistry and Physics* **18** 1745–61

Lyapustin A, Wang Y, Korkin S and Huang D 2018 MODIS Collection 6 MAIAC algorithm *Atmospheric Measurement Techniques* **11** 5741–65