

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

4 Tianjia Liu^{1*}, Frances Marie Panday², Miah C. Caine³, Makoto Kelp¹, Drew C.
5 Pendergrass⁴, and Loretta J. Mickley⁴

6 ¹Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

7 ²Department of Geographical Sciences, University of Maryland, College Park, MD, USA

8 ³Department of Computer Science, Harvard University, Cambridge, MA, USA

9 ⁴John A. Paulson School of Engineering, Harvard University, Cambridge, MA, USA

10

11 *Corresponding author: tliu@ucar.edu

12

13 **Keywords:** smoke, emissions, remote sensing, pollutants: air, scale: regional

14 **Abstract**

15 **Background:** NOAA’s Hazard Mapping System (HMS) smoke product comprises smoke
16 plumes digitized from satellite imagery. Recent studies have used HMS as a proxy for surface
17 smoke presence.

18 **Aims:** We quantify how well HMS agrees with airport observations, air quality station
19 measurements, and model estimates of near-surface smoke.

20 **Methods:** We quantify the agreement in smoke days and trends, regional discrepancies in levels
21 of near-surface smoke fine particulate matter (PM_{2.5}) within HMS polygons, and separation of
22 total PM_{2.5} on smoke and non-smoke days across the contiguous U.S. and Alaska from 2008-
23 2021.

24 **Key Results:** We find large overestimates in HMS-derived smoke days and trends if we include
25 light smoke plumes in the HMS smoke day definition. Outside of the western U.S. and Alaska,
26 near-surface smoke PM_{2.5} within areas of HMS smoke plumes are low and almost
27 indistinguishable across density categories, likely indicating frequent smoke aloft.

28 **Conclusions:** Compared to airport, EPA, and model data, HMS most closely reflects surface
29 smoke in the Pacific and Mountain regions and Alaska when smoke days are defined using only
30 heavy plumes or both medium and heavy plumes.

31 **Implications:** We recommend careful consideration of biases in the HMS smoke product for air
32 quality and public health assessments of fires.

33 1. Introduction

34 Smoke pollution from wildfires in the western United States is increasingly a major
35 public health concern with recent record-breaking fire seasons in 2018, 2020, and 2021 (Zhou *et al*
36 *et al* 2021, Burke *et al* 2021). Decades of fire suppression in the 1900s and droughts in a warming
37 climate together lead to longer and more severe fire seasons, punctuated by megafires that spiral
38 out of control (Williams *et al* 2019, Juang *et al* 2022, Xie *et al* 2022, Syphard *et al* 2017). The
39 growing human population living in the wildland-urban interface is vulnerable to fires and in
40 turn may cause more accidental ignitions. There is an increasing effort to attribute wildfire
41 smoke pollution to public health impacts, but the caveats of underlying datasets used to quantify
42 smoke are not yet fully explored (Zhou *et al* 2021, O’Dell *et al* 2021).

43 Recent public health studies have relied on the NOAA Hazard Mapping System (HMS)
44 smoke product to quantify the smoke fraction in surface fine particulate matter (PM_{2.5}) in the
45 U.S. (Zhou *et al* 2021, O’Dell *et al* 2021). This statistical approach diagnoses smoke PM_{2.5} in
46 surface PM_{2.5} observations on days when PM_{2.5} anomalies align with digitized HMS smoke
47 plume polygons. “Background” PM_{2.5} from other pollution sources in these studies is often
48 calculated as the median PM_{2.5} observed during non-smoke days (Burke *et al* 2021, Childs *et al*
49 2022). More advanced methods interpolate station measurements onto a grid (O’Dell *et al* 2021)
50 or fill in the cloud-induced gaps in HMS data by tracking the trajectory of smoke transport from
51 active fires (Childs *et al* 2022). Traditional air quality and public health assessments of fires on
52 air quality have relied on 3D chemical transport models with input emissions inventories to
53 estimate smoke PM_{2.5} by comparing model runs with and without fire (e.g., Wiggins *et al* 2018,
54 Carter *et al* 2020) or calculating the sensitivity footprint of a receptor to nearby emissions (e.g.,
55 Koplitz *et al* 2016, Marlier *et al* 2019, Kelp *et al* 2023); however, this process is computationally
56 expensive. The HMS statistical approach circumvents having to grapple with model biases
57 stemming from uncertainty in the meteorology driving the smoke transport and in the fire
58 emissions estimates, which are calculated from fire activity, fuel load, and combustion efficiency
59 and depend on poorly-constrained emissions factors (Liu *et al* 2020). Without prior knowledge
60 of emissions levels from different sectors, uncertainty arises from the reliance on the HMS
61 smoke product to distinguish smoke PM_{2.5} from other types of PM_{2.5}. Thus, here we seek to
62 understand: how well does the HMS smoke product reflect surface smoke conditions?

63 The HMS smoke product relies on NOAA analysts to digitize smoke plumes using
64 satellite imagery primarily from the Geostationary Operational Environmental Satellites (GOES).
65 However, the ability of the HMS smoke product to represent surface smoke conditions with high
66 spatial accuracy is uncertain as the product has not yet been fully validated against surface
67 observations. First, HMS smoke polygons represent column smoke presence and do not contain
68 information about the vertical location of smoke – i.e., whether the smoke is aloft or near the
69 surface. HMS may be a poor indicator of surface smoke where smoke is expected to be mostly
70 aloft, such as over states in the Midwest and Northeast that do not generate large amounts of
71 smoke from wildfires and prescribed fires but instead receive smoke transported from other
72 regions. Second, the spatial accuracy of HMS, particularly at the edges of smoke polygons, is
73 affected by the coarse spatial resolution of GOES imagery. The GOES imagery from which
74 HMS smoke is derived has 2-km spatial resolution at the equator, but the resolution over
75 CONUS and Alaska is lower depending on the pixel’s latitude and proximity to the edge of the

76 viewing disk – i.e., the satellite viewing angle. If a region is prone to high-altitude cloud cover,
77 GOES satellites have an advantage over polar-orbiting satellites (e.g., Terra, Aqua, S-NPP,
78 NOAA-20) as they can potentially wait until the clouds move away from the smoke layers.
79 Additionally, HMS does not account for the parallax effect, in which objects observed by GOES
80 are displaced from their actual location. This displacement is dependent on its location and
81 altitude and can affect spatial accuracy of HMS plume edges. Third, HMS does not fully capture
82 the dynamic nature of smoke dispersion. While HMS labels the apparent density of individual
83 plumes as light, medium, or heavy, there may still be high variation in smoke levels within
84 polygons. Because HMS analysts must cover North America every day with only two major
85 updates, the spatial and temporal information HMS provides is coarse. The potential spatial
86 heterogeneity in accuracy suggests that caution should be exercised in public health analyses
87 dependent on the HMS smoke product.

88 In this study, we quantify how accurately HMS represents surface smoke across the U.S.
89 For this evaluation, we use airport observations, U.S. Environmental Protection Agency (EPA)
90 air quality station (AQS) measurements, and model estimates of smoke. First, we compare the
91 magnitude and trends in HMS smoke days with a network of airport observations in the NOAA
92 Integrated Surface Database (ISD). Second, we use EPA AQS measurements and modeled
93 estimates to quantify variation in surface smoke PM_{2.5} concentrations within HMS density
94 categories and by region.

95 **2. Data and Methods**

96 *2.1 NOAA's Hazard Mapping System (HMS) smoke product*

97 To produce NOAA's Hazard Mapping System (HMS) smoke product, analysts use
98 visible satellite imagery to draw polygons of the extent of wildfire smoke (Rolph *et al* 2009).
99 The HMS smoke product is available from August 2005 and produced daily, in near-real-time
100 (<https://www.ospo.noaa.gov/Products/land/hms.html>). HMS analysts use true-color images
101 primarily from GOES geostationary satellites for smoke digitization and currently rely on
102 GOES-16 and GOES-18 imagery. The longitudinal position of GOES-16/East is 75°W and that
103 of GOES-18/West is 137°W. The GOES full disk view of North and South America is 2-km in
104 spatial resolution at the equator and recorded every 10 minutes, while the CONUS-specific view
105 is recorded every 5 minutes. Due to favorable optics at high solar zenith angles, analysts
106 typically update smoke plume polygons for large areas of smoke just twice per day – early
107 morning after sunrise and late afternoon before sunset – while smaller smoke plumes can be
108 updated anytime during daytime hours. Analysts use an animated sequence of satellite images to
109 identify smoke-affected areas and digitize the maximum extent of smoke visible. Each plume's
110 density is further qualitatively classified as light/thin, medium, or heavy/thick smoke based on
111 the apparent opacity of the plume in satellite imagery. HMS smoke plumes are categorically
112 labeled as 5, 16, and 27, which roughly correspond to PM_{2.5} equivalents based on the now
113 discontinued GOES Aerosol Smoke Product (GASP): 5 [0-10] µg/m³ (light/thin), 16 [10-21]
114 µg/m³ (medium), and 27 [21-32] µg/m³ (heavy/thick). However, an update to the HMS smoke
115 product in 2022 removed this connection to the PM_{2.5} equivalents, instead opting for the text
116 labels of “light,” “medium,” and “heavy.” For quality control, we remove malformed HMS
117 polygons with edges crossing, unclosed rings, out-of-bounds coordinates, and insufficient

118 number of vertices, i.e., drawn as lines; these excluded polygons comprise < 0.1% of all
119 polygons.

120 2.1.1 Gap-filling unspecified HMS smoke densities

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

131 We use the following independent variables derived from HMS metadata and satellite
132 data to model the density category: month, time of day of the first and last GOES image used to
133 draw the polygon (“start” and “end”), duration of the animated set of images used to draw the
134 polygon (“duration”), area of polygon (“area”), average Aerosol Optical Depth (AOD) within the
135 polygon (“AOD”), and fraction of overlap with other polygons on the same day (“overlap”)
136 (Table S1). For AOD, we use the MODIS Multi-angle Implementation of Atmospheric
137 Correction (MAIAC) product (MCD19A2, Collection 6) at 0.55 μm (Lyapustin *et al* 2018).
138 MAIAC operates on a fixed 1-km grid and combines the advantages of the MODIS Dark Target
139 and Deep Blue algorithms that specialize on dark vegetative and bright desert surfaces,
140 respectively. The “overlap” variable takes advantage of the nested nature of the smoke polygons;
141 that is, heavy smoke plumes are located within medium smoke extent, and medium smoke
142 plumes are located within light smoke extent (Brey *et al* 2018). We calculate the fractional area
143 of each smoke polygon that overlaps with other polygons from the same day. Medium and heavy
144 smoke polygons have relatively high overlap, and light smoke polygons low overlap.

145 We train two random forest models with and without AOD. Some HMS polygons ($n =$
146 525) had missing AOD values due to cloud coverage preventing successful AOD retrievals. We
147 use the model trained with AOD to gap-fill over 98% ($n = 35303$) of the unspecified densities,
148 while we use the model trained without AOD to gap-fill the remaining unspecified densities. For
149 1000 bootstrap iterations, we undersample the light and medium categories so that all three
150 densities are equally represented in the random forest model; we then split 2/3 of the dataset for
151 training data and for 1/3 for test data. Without undersampling, the random forest model would
152 prioritize the classification accuracy of light smoke, as light smoke plumes (75%) occur much
153 more frequently than medium (18%) and heavy (8%) smoke.

154 Evaluation of the gap-filling method for HMS smoke densities is discussed in
155 Supplemental Section S1. In brief, for the random forest model that considers AOD, the test
156 accuracy is 85% for light smoke, 58% for medium smoke, and 66% for heavy smoke. Accuracies
157 are similar for the model trained without AOD. The lower accuracy for medium smoke relates to
158 the weaker separation of medium smoke with light and heavy smoke by the most important input
159 variable, “overlap” (Figure S1), which takes advantage of the nested nature of the smoke

160 polygons (e.g., heavy smoke polygons nested within medium smoke polygons) and calculates
161 how much each polygon overlaps with other polygons of the same day.

162 2.2 NOAA’s Integrated Surface Database airport observations

163 NOAA’s Integrated Surface Database (ISD) collates observations of meteorological
164 parameters at airports at varying temporal frequencies (Smith *et al* 2011). Meteorological
165 observations include air temperature, surface pressure, visibility, as well as indicators of low
166 visibility due to haze, clouds/mist, dust, and smoke. We use the atmospheric condition codes
167 from the automated weather (AW) reports in the ISD dataset. To define a smoke observation, we
168 use the “smoke” (AW=5) code. We filter out airports that do not have any smoke observations or
169 do not on average have more than one observation per day from 2008-2021. We use a total of
170 1513 airports across CONUS and 104 airports in Alaska (Figure 1). To filter out spurious ISD
171 observations of smoke, we designate a day as a smoke day if > 5% of all observations during that
172 day are labeled as smoke.

173 2.3 Evaluating HMS smoke days with ISD airport observations

174 For HMS, we test three definitions of smoke days based on presence of the light,
175 medium, and heavy smoke density categories: 1) all (light, medium, or heavy), 2)
176 medium/heavy, and 3) heavy only. In the heavy-only definition, for example, we designate a day
177 as a smoke day only if a heavy smoke plume overlaps with a particular location; otherwise, days
178 are considered non-smoke days. At each airport, we compare the average smoke days and linear
179 trend in smoke days as derived from smoke observations from ISD airport and HMS data during
180 smoky-heavy months, or months with > 5% of annual HMS smoke days. This constraint limits
181 our analysis to months when fire-related smoke is likely a dominant pollution source.

182 For each airport location, we quantify the difference in HMS and airport average smoke
183 days per year and trend in smoke days from 2008-2021. We compare statistics and accuracy
184 metrics for nine sub-regions: Alaska, Pacific, Mountain, West North Central, West South
185 Central, East North Central, East South Central, Northeast, and South Atlantic (Figure 1). We
186 use two accuracy metrics, the Cohen’s kappa (κ) and Matthews correlation coefficient (MCC), to
187 evaluate the agreement between HMS and airport smoke day classifications. The Cohen’s kappa
188 is a widely used metric for validation in remote sensing studies that involve classification, such
189 as mapping land cover types and change (Cohen 1960). The MCC is a proposed alternative for
190 the Cohen’s kappa; although both metrics are derived from confusion matrices, the MCC
191 performs better on imbalanced datasets and overall is a more informative and reliable metric to
192 evaluate binary classification (Matthews 1975, Chicco *et al* 2021). For two-class comparisons,
193 the Cohen’s kappa and MCC metrics are calculated as follows:

$$194 \quad \kappa = \frac{2(TP \times TN - FP \times FN)}{(TP + FP) \times (TN + FP) + (TP + FN) \times (TN + FN)} \quad \text{Eq. 1}$$

$$195 \quad MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad \text{Eq. 2}$$

196 where TP is number the true positives (i.e., both airport and HMS = smoke day), TN is
197 the number of true negatives (i.e., both airport and HMS = non-smoke day), FP is the number of

198 false positives (i.e., airport = non-smoke day, HMS = smoke day), and FN is the number of false
199 negatives (i.e., airport = smoke day, HMS = non-smoke day).

200 Additionally, we calculate the true positive rate (TPR, recall), positive predictive value
201 (PPV, precision), false positive rate (FPR), and negative predictive value (NPV) to complement
202 our analysis:

$$203 \quad TPR = \frac{TP}{TP+FN} \quad \text{Eq. 3}$$

$$204 \quad PPV = \frac{TP}{TP+FP} \quad \text{Eq. 4}$$

$$205 \quad FPR = \frac{FP}{FP+TN} \quad \text{Eq. 5}$$

$$206 \quad NPV = \frac{TN}{TN+FN} \quad \text{Eq. 6}$$

207 2.4 Evaluating elevated $PM_{2.5}$ at EPA stations during HMS smoke days

208 As an additional way to evaluate the HMS smoke density categories, we use daily $PM_{2.5}$
209 measurements at EPA stations across CONUS and Alaska. We obtain daily average EPA $PM_{2.5}$
210 data under parameter codes 88801 and 88502, which refer to the designation of federal reference
211 method (FRM) and federal equivalent method (FEM) for quality control
212 (https://aqs.epa.gov/aqsweb/airdata/download_files.html). We use a total of 1025 EPA stations
213 that have at least a decade of measurements from 2008-2021 and over an average of 100
214 measurements per year (Figure S2). To approximate smoke $PM_{2.5}$, we subtract the total $PM_{2.5}$
215 from the background $PM_{2.5}$, or the median $PM_{2.5}$ on days designated as non-smoke by HMS
216 during that year. We then classify the total $PM_{2.5}$ on HMS smoke days by the maximum HMS
217 smoke density category of each day and compare across regions. Large variation exists in the
218 background $PM_{2.5}$, but we would expect the total $PM_{2.5}$ on the HMS smoke days to fall at the
219 higher end of the distribution of background values. To test this, we also report the percentile at
220 which the total $PM_{2.5}$ on smoke days lies on the cumulative probability distribution of
221 background $PM_{2.5}$ values. The percentile measures the separation between the $PM_{2.5}$ on smoke
222 and non-smoke days; higher percentiles imply that we have greater confidence in attributing
223 elevated $PM_{2.5}$ to smoke.

224 2.5 Evaluating the spatial consistency of modeled near-surface smoke $PM_{2.5}$ within HMS 225 polygons

226 We use the High-Resolution Rapid Refresh (HRRR)-Smoke modeling system to track the
227 spatial consistency in near-surface smoke $PM_{2.5}$ across CONUS
228 (<https://rapidrefresh.noaa.gov/hrrr/HRRRsmoke/>). HRRR-Smoke is based on the Weather and
229 Research Forecasting model coupled with Chemistry (WRF-Chem) and input fire emissions
230 calculated from fire radiative power (FRP), a proxy for fire intensity that is directly proportional
231 to emissions (Ahmadov *et al* 2017). The FRP is derived from observations by the Visible
232 Infrared Imaging Radiometer Suite (VIIRS) sensor aboard the Suomi-NPP satellite. HRRR-
233 Smoke provides near-real-time hourly surface smoke $PM_{2.5}$ at 3-km spatial resolution that we
234 then average to daily scale. We use the HRRR-Smoke 2D outputs ('wrfscf') at forecast hour 0 in
235 2021, a high fire year and the first year that the near-surface smoke $PM_{2.5}$ variable

236 ('MASSDEN') became available in the operational product (accessed from: <https://noaa-hrrr->
237 [bdp-pds.s3.amazonaws.com/index.html](https://noaa-hrrr-bdp-pds.s3.amazonaws.com/index.html)). We track how HRRR-Smoke PM_{2.5} concentrations vary
238 across smoke polygons with the same density category. For example, the occurrence of low
239 smoke PM_{2.5} values from HRRR located within heavy HMS smoke polygons signals that the
240 smoke is lofted, and that HMS does not accurately reflect surface smoke levels in those areas.

241 **3. Results and Discussion**

242 *3.1 Evaluating HMS and ISD average smoke days and trends in smoke days by airport*

243 We compare HMS and ISD average smoke days (Figure 2) and trends in smoke days
244 (Figure 3) from 2008-2021 across airport locations in CONUS ($n = 1513$) and Alaska ($n = 104$).
245 In general, HMS shows large-scale changes in smoke presence with high spatial autocorrelation,
246 while ISD shows more localized patterns in smoke days and their trends. Sporadic hotspots
247 evident in ISD smoke days across the East and Midwest may be attributed to inconsistencies in
248 the automated system for smoke detection or contamination from nearby local pollution sources.
249 Despite this caveat in ISD data, we can still examine differences between HMS and ISD on a
250 broad regional scale (Figure 1).

251 The dominant source of smoke varies by region. Wildfires dominate the West and
252 Alaska, while the Southeast mainly sees agricultural fires and prescribed burns; the Midwest and
253 Northeast typically experience smoke transported from western states or Canada (Brey *et al*
254 2018, Cottle *et al* 2014). HMS identifies the highest smoke pollution in Pacific and Midwest
255 states. Consistent across HMS and ISD-derived smoke days, Pacific states (CA, WA, and OR)
256 comprise the most smoke-polluted region (Figures 2-3). This finding is underscored by a cluster
257 of airport locations observing over 10 smoke days per year within California's Central Valley,
258 which is close in proximity to large wildfires and experiences frequent temperature inversions
259 that trap smoke near the surface. In contrast, a large discrepancy between HMS and ISD is
260 evident in the Midwest, or the East North Central and West North Central states. The high smoke
261 pollution derived from HMS in the Midwest – on par or exceeding that in Pacific states in some
262 cases – is largely absent in ISD data. This result suggests that the smoke over the Midwest is
263 often aloft and may not affect surface air quality.

264 The contrast between Pacific and Midwest states is supported by the spatial variation in
265 Cohen's kappa and MCC values calculated from the HMS-ISD agreement in smoke days (Figure
266 4). We observe the highest HMS-airport agreement in Pacific states (median $\kappa = 0.36$, MCC =
267 0.37), weak agreement in Mountain states and Alaska (median $\kappa = 0.15$ to 0.19, MCC = 0.18 to
268 0.20), and low agreement elsewhere (median $\kappa < 0.1$, MCC < 0.1) for the heavy-only HMS
269 smoke day definition (Figure 5). Across almost all regions, using heavy-only HMS smoke leads
270 to lower recall (TPR) but higher precision (PPV) and lower false positive rates. This results in
271 higher Cohen's kappa and MCC values for the heavy-only HMS smoke day definition compared
272 to those using both medium and heavy plumes or all HMS plumes. Exceptions where the
273 medium/heavy smoke definition slightly outperforms the heavy-only smoke definition are in
274 West South Central, East South Central, and South Atlantic, where the accuracy for all HMS
275 smoke definitions is among the lowest across all regions (median $\kappa \leq 0.03$, MCC ≤ 0.03). The

276 negative predictive value is close to 1 in all regions and for all HMS smoke definitions,
277 indicating low misclassification of non-smoke days.

278 The overestimation of smoke days and their trends by HMS compared to ISD is evident
279 when including medium smoke with heavy smoke, and even more pronounced when all smoke
280 types are considered (Figures 2-3, 6-7). In the western U.S., we estimate 5.8 average airport-
281 observed smoke days from 2008-2021 at 581 airport location. In contrast, the number of average
282 HMS-observed smoke days is highly variable depending on the definition, ranging from 3.5 days
283 for heavy smoke to 9.9 days for medium/heavy smoke to 33.6 days for all smoke categories
284 combined (Figure 6). This pattern extends across all CONUS regions and Alaska, where the
285 inclusion of light smoke plumes leads to 2.7 to 16 times the number of airport smoke days
286 (Figure 7). Our results suggest that light smoke plumes should generally be excluded for a binary
287 classification of smoke and non-smoke days at the surface.

288 3.2 Spatial variability in observed and modeled near-surface smoke $PM_{2.5}$ levels within HMS 289 smoke polygons

290 In general, we find that the EPA $PM_{2.5}$ – particularly on days with a heavy HMS plume
291 overhead – is more easily separated from the $PM_{2.5}$ on non-smoke days in the Pacific and
292 Mountain regions and Alaska (Figure 8). On HMS smoke days, surface concentrations of total
293 $PM_{2.5}$ in these regions fall in the range of 87 to 93% on the cumulative probability distribution of
294 background $PM_{2.5}$ values, while those in other regions range from 68 to 79%. Because the 50th
295 percentile, or the median, is often used as the upper limit for background $PM_{2.5}$ (e.g., Koplitz *et*
296 *al* 2016, Childs *et al* 2022). $PM_{2.5}$ on HMS smoke days falling in low percentiles may be
297 misclassified as smoke-affected. The percentiles are generally lowest for light smoke days, and
298 highest for heavy smoke days, which indicates greater confidence in attributing elevated $PM_{2.5}$ to
299 smoke during the latter.

300 We also find that the $PM_{2.5}$ equivalents of the HMS light (5 [0-10] $\mu\text{g}/\text{m}^3$), medium (16
301 [10-21] $\mu\text{g}/\text{m}^3$), and heavy (27 [21-32] $\mu\text{g}/\text{m}^3$) density categories correspond better to the EPA
302 and HRRR-Smoke near-surface smoke $PM_{2.5}$ concentrations in the Pacific and Mountain regions
303 and Alaska than elsewhere across CONUS (Figure 9a-b). Modeled smoke concentrations in 2021
304 for the Pacific region are close to the HMS equivalent values for those plumes, with averages of
305 9 $\mu\text{g}/\text{m}^3$, 17 $\mu\text{g}/\text{m}^3$, and 36 $\mu\text{g}/\text{m}^3$ in the three categories in order of increasing density (Figure
306 9b). For the Mountain region, the distinctions between near-surface modeled $PM_{2.5}$ within the
307 three categories of HMS plumes are much less, with averages of 5 $\mu\text{g}/\text{m}^3$, 9 $\mu\text{g}/\text{m}^3$, and 16 $\mu\text{g}/\text{m}^3$;
308 these modeled values also deviate from the HMS $PM_{2.5}$ equivalent ranges. For all other regions,
309 the average near-surface $PM_{2.5}$ within medium and heavy plumes all fall within the light smoke
310 $PM_{2.5}$ equivalent range ($< 10 \mu\text{g}/\text{m}^3$), which suggests that most smoke is actually aloft over these
311 regions. We find similar patterns in the EPA AQS-derived smoke $PM_{2.5}$ from 2008-2021, with
312 Alaska also seeing similar smoke $PM_{2.5}$ distributions in each HMS category as Pacific and
313 Mountain states (Figure 9a). Reasons for the lower smoke $PM_{2.5}$ from EPA relative to HRRR-
314 Smoke may include the imperfect assumption of the background $PM_{2.5}$ as the median $PM_{2.5}$ on
315 non-smoke days, missing data, and spatial bias of EPA stations in urban centers and overall
316 sparsity in spatial coverage. Previous studies have found nighttime overestimates in HRRR-
317 Smoke and underestimates when FRP is biased low compared to observations (Ye *et al* 2021,
318 Chow *et al* 2022).

319 Even within HMS plumes of the same category, we find regional biases in the magnitude
320 of the surface smoke $PM_{2.5}$ concentration and the separation of the $PM_{2.5}$ from the background
321 $PM_{2.5}$. While a smoke plume may have uniform opacity and thickness as seen from satellite
322 imagery — thereby allowing an analyst to justify labeling it with a single HMS density category
323 — the underlying surface smoke $PM_{2.5}$ may differ substantially depending on location. The re-
324 processing of the HMS smoke product in 2022 removed the link between the smoke density
325 categories and $PM_{2.5}$ equivalents, which discouraged the data user from incorrectly deriving
326 surface smoke $PM_{2.5}$ from HMS. We recommend that data users interpret the HMS smoke
327 density categories with caution and carefully assess potential regional biases.

328 When using a statistical method to calculate smoke $PM_{2.5}$ — that is, by using total $PM_{2.5}$
329 observations with HMS to partition smoke and non-smoke days — overestimates in smoke days
330 will result in overestimates of smoke-related air pollution and public health impacts. This is
331 because the calculation of the background $PM_{2.5}$ using median or mean values is imperfect, and
332 elevated $PM_{2.5}$ may be incorrectly attributed to smoke. We recommend that studies calculate the
333 uncertainty in smoke $PM_{2.5}$ estimates due to variance in background $PM_{2.5}$ and confidence in
334 smoke attribution.

335 *3.3 Comparison of strengths and caveats of HMS, airport, and model estimates of surface smoke* 336 *presence*

337 Here we outline the strengths and caveats of using HMS, airport observations, EPA AQS
338 measurements, and model estimates as indicators of surface smoke presence. Understanding the
339 strengths and caveats of these different datasets is an important step in designing a study on
340 quantifying the impacts of fire-induced smoke exposure.

341 *HMS smoke product.* The HMS smoke product is available in near-real-time and provides a
342 simple classification of smoke density (light, medium, heavy) for digitized smoke plumes.
343 However, the smoke plumes are mapped based on an analyst’s interpretation of true-color
344 satellite imagery. Human error, the coarse resolution and parallax displacement of GOES
345 imagery, as well as potential cloud cover, can all lead to biases and inconsistencies in the dataset.
346 Additionally, the HMS smoke product represents column observations of smoke. When used as
347 an indicator of surface smoke, regional biases arise, caused by variance in the altitude of smoke
348 plumes. Using HMS leads to inflated surface smoke estimates in regions with mostly aloft
349 smoke. This regional bias propagates to using the smoke density categories to differentiate
350 surface smoke levels.

351 *Airport observations.* Airport observations are available in near-real-time and provide a ground-
352 level view of smoke presence and levels of visibility reduction. However, the density of
353 observations is sparse given the available airport locations (Figure 1). Caveats include station-to-
354 station differences in observations, potential contamination by local sources, or misdiagnosis
355 smoke as some other air pollutant, which could lead to errors in reporting smoke influence. As
356 airport data is underused, these caveats are currently not well understood.

357 *EPA stations.* EPA stations offer high-quality, ground-based observations of air pollution levels,
358 often in near-real-time. Like the network of ISD airports, the EPA stations are sparsely
359 distributed across the U.S. with a bias toward urban centers (Figure S2). A main caveat is that
360 EPA stations often only report the total $PM_{2.5}$. The task to separate smoke $PM_{2.5}$ from the

361 background PM_{2.5} is non-trivial, with many studies relying on statistical methods. Station
362 measurements from the Interagency Monitoring of Protected Visual Environments (IMPROVE)
363 network offer some insights into the PM_{2.5} composition — e.g., organic and black carbon (OC
364 and BC) — but only report every three days. It is possible to infer smoke contribution to total
365 PM_{2.5} during days dominated by OC+BC, but direct attribution is difficult due to co-varying
366 sources, such as traffic, industrial facilities, dust, and secondary organic aerosol formation.
367 Additional data from low-cost sensors, such as the PurpleAir network, may supplement the EPA
368 and IMPROVE data and decrease the spatial sparsity of station locations.

369 *Model estimates.* Model estimates of smoke concentrations can be generated in near-real-time
370 but are generally used for historical analysis as emissions inventories are updated with some lag
371 time. Chemical transport models rely on fire and anthropogenic emissions inventories and
372 transport (i.e., meteorology) to be accurate, but substantial differences may exist among the
373 available input datasets. Such models are also computationally expensive. However, partitioning
374 fire from non-fire PM_{2.5} is an easy task. The model outputs provide spatially cohesive smoke
375 PM_{2.5} estimates and are important where there are little to no ground monitors.

376 Although airport observations, EPA AQS measurements, and model estimates have their
377 own biases and uncertainties, we can broadly pinpoint where HMS may not accurately reflect
378 surface smoke presence, namely outside of Alaska and the Pacific and Mountain regions. Future
379 studies can take advantage of the agreement and disagreement between ground, satellite, and
380 model estimates to make more robust conclusions.

381 **4. Conclusion**

382 In summary, we present three lines of evidence from airport observations, EPA AQS
383 measurements, and model estimates that across much of CONUS and Alaska, the HMS smoke
384 product conflates surface smoke presence with smoke aloft. Only in western U.S. and Alaska
385 does the HMS smoke product appear to agree consistently with other measures of surface smoke.
386 For example, compared to the airport-observed average of 5.8 smoke days per year in the
387 western U.S. from 2008-2021, HMS severely overestimates the number of smoke days if all
388 smoke density categories (light, medium, and heavy) are included (33.6 days). Using only
389 medium and heavy plumes (9.9 days) or only heavy plumes (3.5 days) leads to better agreement
390 with airport observations in this region. Outside of western U.S. and Alaska, observed and
391 modeled surface smoke PM_{2.5} concentrations occurring within medium and heavy HMS plumes
392 are similar to those of light plumes (< 10 µg/m³). This finding suggests that the impact of smoke
393 on surface air quality is relatively minimal in areas where smoke is often aloft, though the
394 corresponding plumes may be categorized as medium or heavy density by HMS. Exceptions to
395 this, however, can be seen from Canada's recent record-breaking fire season in 2023, when
396 smoke from these fires degraded surface air quality to unhealthy levels in northeastern and
397 midwestern states. For future studies, we urge caution in using the HMS smoke product as a
398 broad indicator of surface smoke, as its performance varies widely by region, and inclusion of
399 light smoke – and sometimes, even medium smoke – inflates both the number of and trend in
400 smoke days. For defining smoke days, using only heavy or both medium and heavy smoke
401 plumes can serve as lower and upper bound estimates, respectively.

402

403 **Acknowledgements**

404 This study was supported by the NOAA Climate Program Office's Modeling, Analysis,
405 Predictions, and Projections Program (MAPP), Grant NA22OAR4310140. T. Liu and D.C.
406 Pendergrass were funded by NSF Graduate Research Fellowships (NSF grant DGE1745303).
407 F.M. Panday was funded by the NSF program for Research Experiences for Undergraduates,
408 Grant 2150058. M.C. Caine was funded by the Harvard University Center for the Environment
409 (HUCE) Summer Undergraduate Research Fund and Harvard College Research Program
410 (HCRP). We thank John Simko and Wilfrid Schroeder for their help and comments on this work.

411

412 **Data Availability**

413 The Hazard Mapping System (HMS) smoke product
414 (https://satepsanone.nesdis.noaa.gov/pub/FIRE/web/HMS/Smoke_Polygons/Shapefile/),
415 Integrated Surface Database (ISD) of airport observations
416 (<https://www.ncei.noaa.gov/data/global-hourly/archive/csv/>), and HRRR-Smoke model outputs
417 (<https://rapidrefresh.noaa.gov/hrrr/HRRRsmoke/>) are distributed by NOAA. The
418 MODIS MAIAC aerosol product is distributed by NASA
419 (<https://doi.org/10.5067/MODIS/MCD19A2.006>) and available from the Google Earth Engine
420 public data catalog.

421

422

423 **Conflicts of Interest**

424 The authors declare no conflicts of interest.

425 **References**

- 426 Ahmadov R, Grell G, James E, Csiszar I, Tsidulko M, Pierce B, McKeen S, Benjamin S,
427 Alexander C, Pereira G, Freitas S and Goldberg M 2017 Using VIIRS fire radiative power
428 data to simulate biomass burning emissions, plume rise and smoke transport in a real-time
429 air quality modeling system 2017 *IEEE International Geoscience and Remote Sensing*
430 *Symposium (IGARSS)* pp 2806–8 Online: <https://doi.org/10.1109/IGARSS.2017.8127581>
- 431 Breiman L 2001 Random Forests *Mach. Learn.* **45** 5–32 Online:
432 <https://doi.org/10.1023/A:1010933404324>
- 433 Brey S J, Ruminski M, Atwood S A and Fischer E V. 2018 Connecting smoke plumes to sources
434 using Hazard Mapping System (HMS) smoke and fire location data over North America
435 *Atmos. Chem. Phys.* **18** 1745–61 Online: <https://doi.org/10.5194/acp-18-1745-2018>
- 436 Burke M, Driscoll A, Heft-Neal S, Xue J, Burney J and Wara M 2021 The changing risk and
437 burden of wildfire in the United States *Proc. Natl. Acad. Sci.* **118** e2011048118 Online:
438 <https://doi.org/10.1073/pnas.2011048118>
- 439 Carter T, Heald C, Jimenez J, Campuzano-Jost P, Kondo Y, Moteki N, Schwarz J, Wiedinmyer
440 C, Darmenov A and Kaiser J 2020 How emissions uncertainty influences the distribution
441 and radiative impacts of smoke from fires in North America *Atmos. Chem. Phys.* **20** 2073–
442 97 Online: <https://doi.org/10.5194/acp-20-2073-2020>
- 443 Chicco D, Warrens M J and Jurman G 2021 The Matthews Correlation Coefficient (MCC) is
444 More Informative Than Cohen’s Kappa and Brier Score in Binary Classification
445 Assessment *IEEE Access* **9** 78368–81 Online:
446 <https://doi.org/10.1109/ACCESS.2021.3084050>
- 447 Childs M L, Li J, Wen J, Heft-neal S, Driscoll A, Wang S, Gould C F, Qiu M, Burney J and
448 Burke M 2022 Daily Local-Level Estimates of Ambient Wildfire Smoke PM_{2.5} for the
449 Contiguous US *Environ. Sci. Technol.* **56** 13607–13621 Online:
450 <https://doi.org/10.1021/acs.est.2c02934>
- 451 Chow F K, Yu K A, Young A, James E, Grell G A, Csiszar I, Tsidulko M, Freitas S, Pereira G,
452 Giglio L, Friberg M D and Ahmadov R 2022 High-Resolution Smoke Forecasting for the
453 2018 Camp Fire in California *Bull. Am. Meteorol. Soc.* **103** E1531–52 Online:
454 <https://doi.org/10.1175/BAMS-D-20-0329.1>
- 455 Cohen J 1960 A Coefficient of Agreement for Nominal Scales *Educ. Psychol. Meas.* **20** 37–46
456 Online: <https://doi.org/10.1177/001316446002000104>
- 457 Cottle P, Strawbridge K and McKendry I 2014 Long-range transport of Siberian wildfire smoke
458 to British Columbia: Lidar observations and air quality impacts *Atmos. Environ.* **90** 71–7
459 Online: <https://doi.org/10.1016/j.atmosenv.2014.03.005>
- 460 Cusworth D H, Mickley L J, Sulprizio M P, Liu T, Marlier M E, Defries R S, Guttikunda S K
461 and Gupta P 2018 Quantifying the influence of agricultural fires in northwest India on urban
462 air pollution in Delhi, India *Environ. Res. Lett.* **13** 044018 Online:
463 <https://doi.org/10.1088/1748-9326/aab303>
- 464 Juang C S, Williams A P, Abatzoglou J T, Balch J K, Hurteau M D and Moritz M A 2022 Rapid
465 Growth of Large Forest Fires Drives the Exponential Response of Annual Forest-Fire Area

466 to Aridity in the Western United States *Geophys. Res. Lett.* **49** e2021GL097131 Online:
467 <https://doi.org/10.1029/2021gl097131>

468 Kelp M M, Carroll M C, Liu T, Yantosca R M, Hockenberry H E and Mickley L J 2023
469 Prescribed Burns as a Tool to Mitigate Future Wildfire Smoke Exposure: Lessons for States
470 and Rural Environmental Justice Communities *Earth's Futur.* **11** e2022EF003468 Online:
471 <https://doi.org/10.1029/2022EF003468>

472 Koplitz S N, Mickley L J, Marlier M E, Buonocore J J, Kim P S, Liu T, Sulprizio M P, DeFries
473 R S, Jacob D J, Schwartz J, Pongsiri M and Myers S S 2016 Public health impacts of the
474 severe haze in Equatorial Asia in September–October 2015: demonstration of a new
475 framework for informing fire management strategies to reduce downwind smoke exposure
476 *Environ. Res. Lett.* **11** 94023 Online: <https://doi.org/10.1088/1748-9326/11/9/094023>

477 Liu T, Mickley L J, Marlier M E, DeFries R S, Khan M F, Latif M T and Karambelas A 2020
478 Diagnosing spatial biases and uncertainties in global fire emissions inventories: Indonesia
479 as regional case study *Remote Sens. Environ.* **237** 111557 Online:
480 <https://doi.org/10.1016/j.rse.2019.111557>

481 Lyapustin A, Wang Y, Korkin S and Huang D 2018 MODIS Collection 6 MAIAC algorithm
482 *Atmos. Meas. Tech.* **11** 5741–65 Online: <https://doi.org/10.5194/amt-11-5741-2018>

483 Marlier M E, Liu T, Yu K, Buonocore J J, Koplitz S N, DeFries R S, Mickley L J, Jacob D J,
484 Schwartz J, Wardhana B S and Myers S S 2019 Fires, Smoke Exposure, and Public Health:
485 An Integrative Framework to Maximize Health Benefits From Peatland Restoration
486 *GeoHealth* **3** 178–89 Online: <https://doi.org/10.1029/2019GH000191>

487 Matthews B W 1975 Comparison of the predicted and observed secondary structure of T4 phage
488 lysozyme *Biochim. Biophys. Acta* **405** 442–51 Online: [https://doi.org/10.1016/0005-2795\(75\)90109-9](https://doi.org/10.1016/0005-2795(75)90109-9)

490 O'Dell K, Bilsback K, Ford B, Martenies S E, Magzamen S, Fischer E V. and Pierce J R 2021
491 Estimated Mortality and Morbidity Attributable to Smoke Plumes in the United States: Not
492 Just a Western US Problem *GeoHealth* **5** e2021GH000457 Online:
493 <https://doi.org/10.1029/2021GH000457>

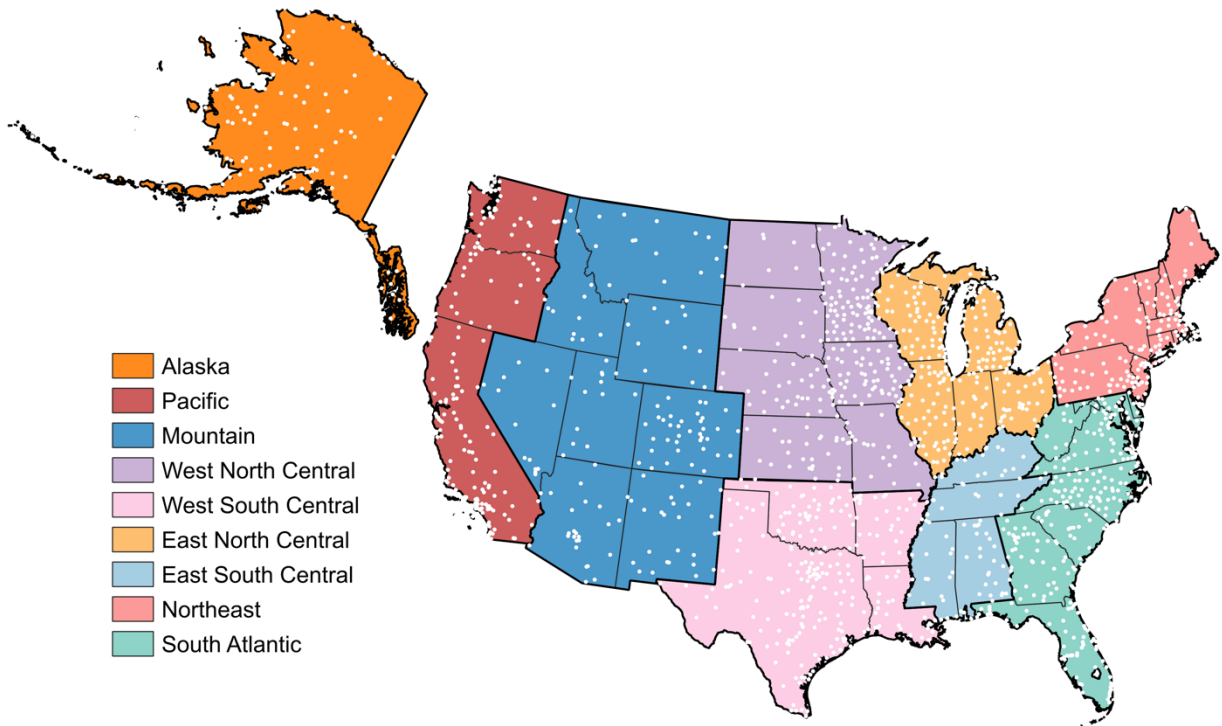
494 Rolph G D, Draxler R R, Stein A F, Taylor A, Ruminski M G, Kondragunta S, Zeng J, Huang H
495 C, Manikin G, McQueen J T and Davidson P M 2009 Description and verification of the
496 NOAA smoke forecasting system: The 2007 fire season *Weather Forecast.* **24** 361–78
497 Online: <https://doi.org/10.1175/2008WAF2222165.1>

498 Smith A, Lott N and Vose R 2011 The Integrated Surface Database: Recent Developments and
499 Partnerships *Bull. Am. Meteorol. Soc.* 704–8 Online:
500 <https://doi.org/10.1109/IGARSS.2017.8127581>

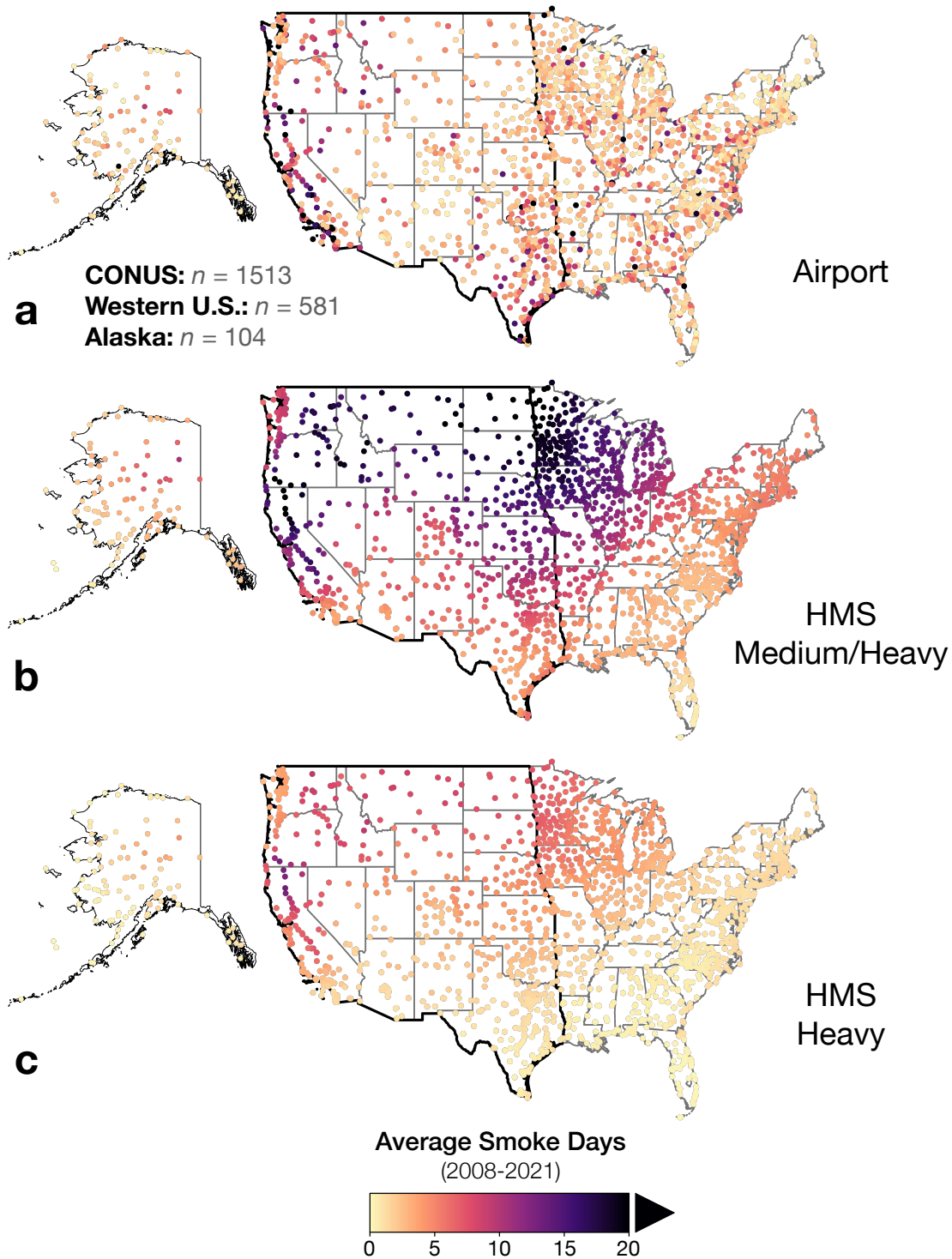
501 Syphard A D, Keeley J E, Pfaff A H and Ferschweiler K 2017 Human presence diminishes the
502 importance of climate in driving fire activity across the United States *Proc. Natl. Acad. Sci.*
503 **114** 13750–5 Online: <https://doi.org/10.1073/pnas.1713885114>

504 Wiggins E B, Yu L E, Holden S R, Chen Y, Kai F M, Czimczik C I, Harvey C F, Santos G M,
505 Xu X and Randerson J T 2018 Smoke radiocarbon measurements from Indonesian fires
506 provide evidence for burning of millennia-aged peat *Proc. Natl. Acad. Sci.* **115** 12419–24
507 Online: <https://doi.org/10.1073/pnas.1806003115>

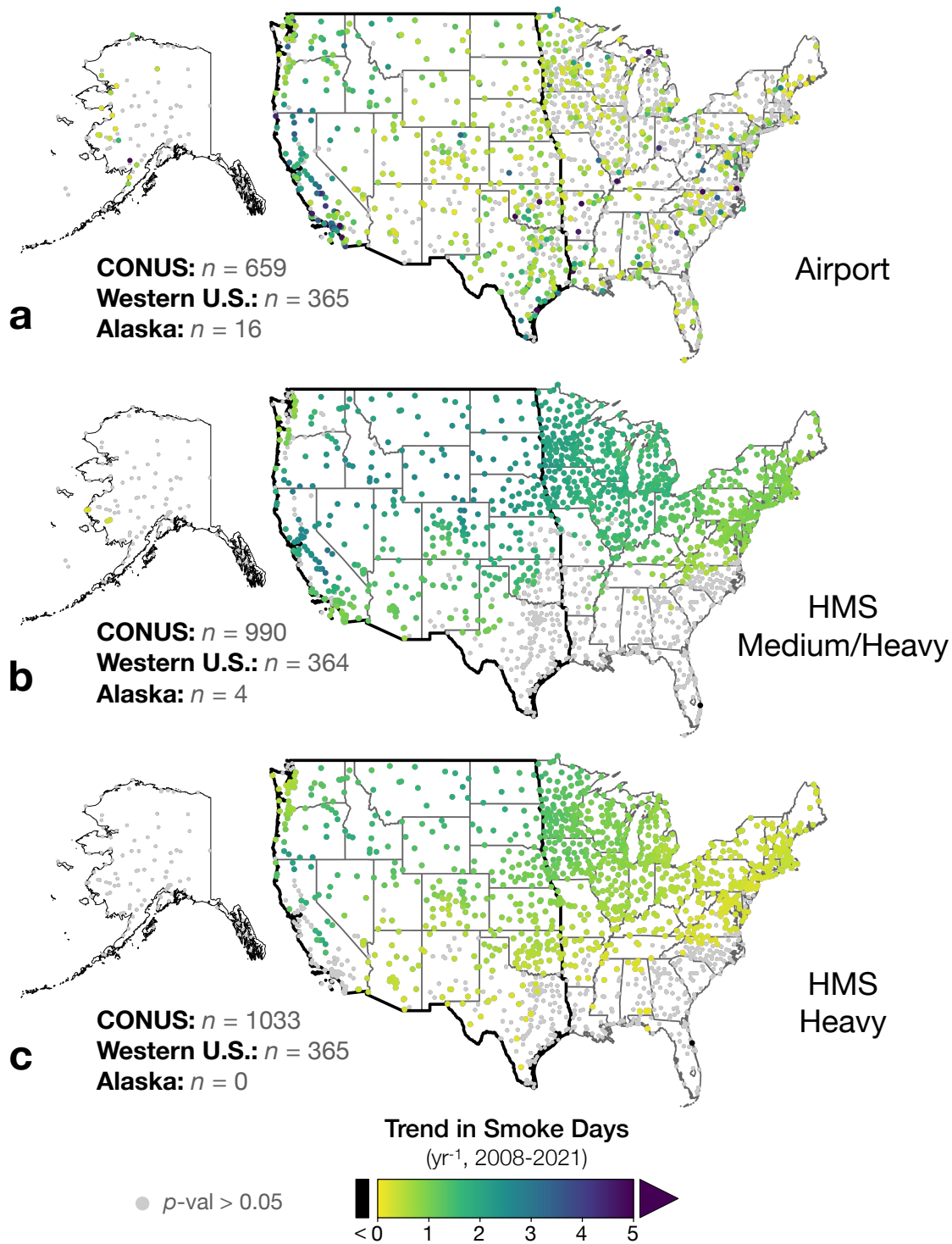
- 508 Williams A P, Abatzoglou J T, Gershunov A, Guzman-Morales J, Bishop D A, Balch J K and
509 Lettenmaier D P 2019 Observed impacts of anthropogenic climate change on wildfire in
510 California *Earth's Futur.* **7** Online: <https://doi.org/10.1029/2019EF001210>
- 511 Xie Y, Lin M, Decharme B, Delire C, Horowitz L W, Lawrence D M, Li F and Séférian R 2022
512 Tripling of western US particulate pollution from wildfires in a warming climate *Proc. Natl.*
513 *Acad. Sci. U. S. A.* **119** e2111372119 Online: <https://doi.org/10.1073/pnas.2111372119>
- 514 Ye X, Arab P, Ahmadov R, James E, Grell G A, Pierce B, Kumar A, Makar P, Chen J, Davignon
515 D, Carmichael G R, Ferrada G, Mcqueen J, Huang J, Kumar R, Emmons L, Herron-Thorpe
516 F L, Parrington M, Engelen R, Peuch V H, Da Silva A, Soja A, Gargulinski E, Wiggins E,
517 Hair J W, Fenn M, Shingler T, Kondragunta S, Lyapustin A, Wang Y, Holben B, Giles D M
518 and Saide P E 2021 Evaluation and intercomparison of wildfire smoke forecasts from
519 multiple modeling systems for the 2019 Williams Flats fire *Atmos. Chem. Phys.* **21** 14427–
520 69 Online: <https://doi.org/10.5194/acp-21-14427-2021>
- 521 Zhou X, Josey K, Kamareddine L, Caine M C, Liu T, Mickley L J, Cooper M and Dominici F
522 2021 Excess of COVID-19 cases and deaths due to fine particulate matter exposure during
523 the 2020 wildfires in the United States *Sci. Adv.* **7** eabi8789 Online:
524 <https://doi.org/10.1126/sciadv.abi878>
- 525



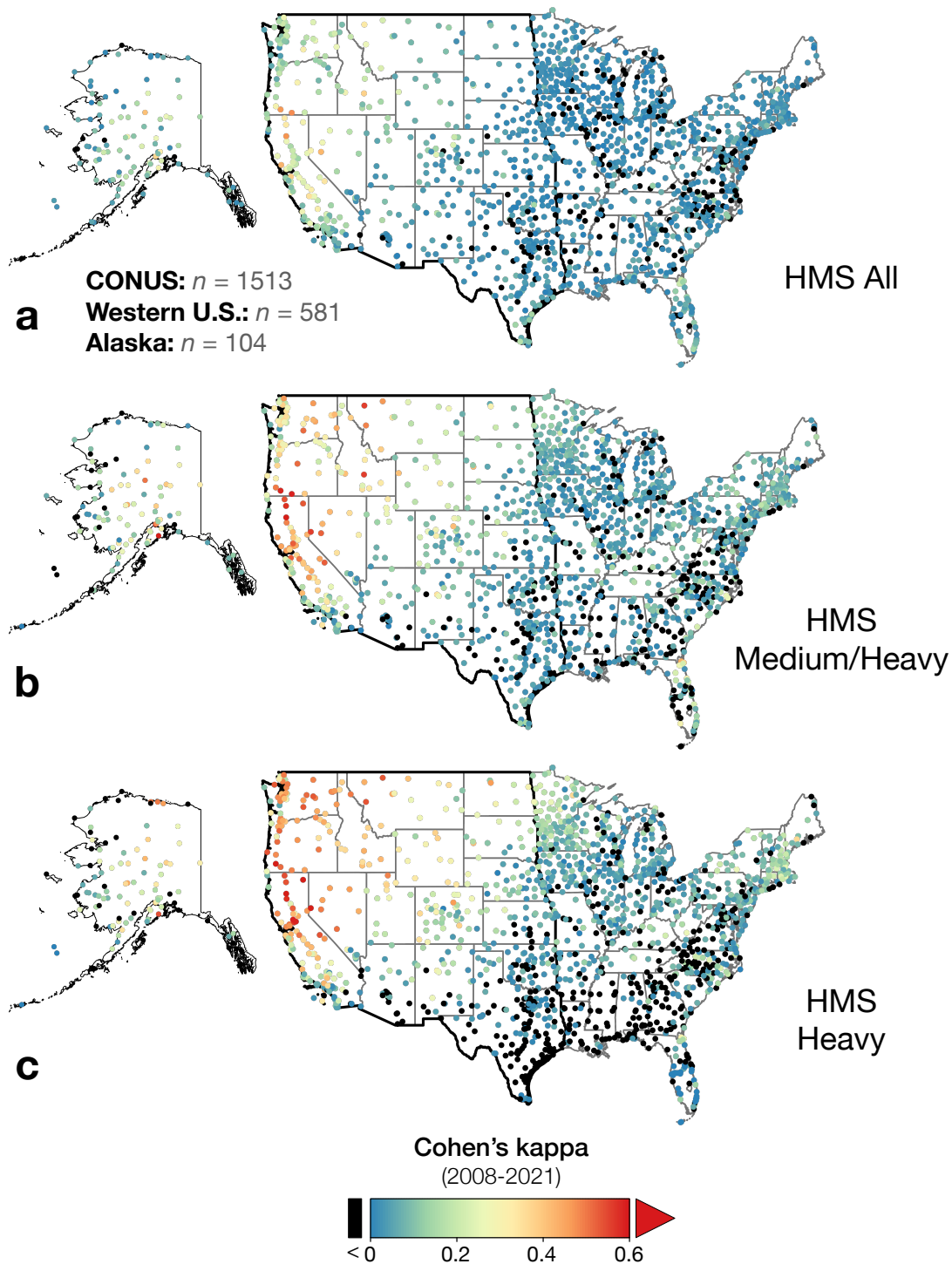
526
 527 **Figure 1. Map of CONUS regions and Alaska with ISD airport locations.** Each white dot
 528 represents the location of an airport in the Integrate Surface Database (ISD) used in this study.
 529 (Note that Alaska is not shown on the same scale as CONUS.)



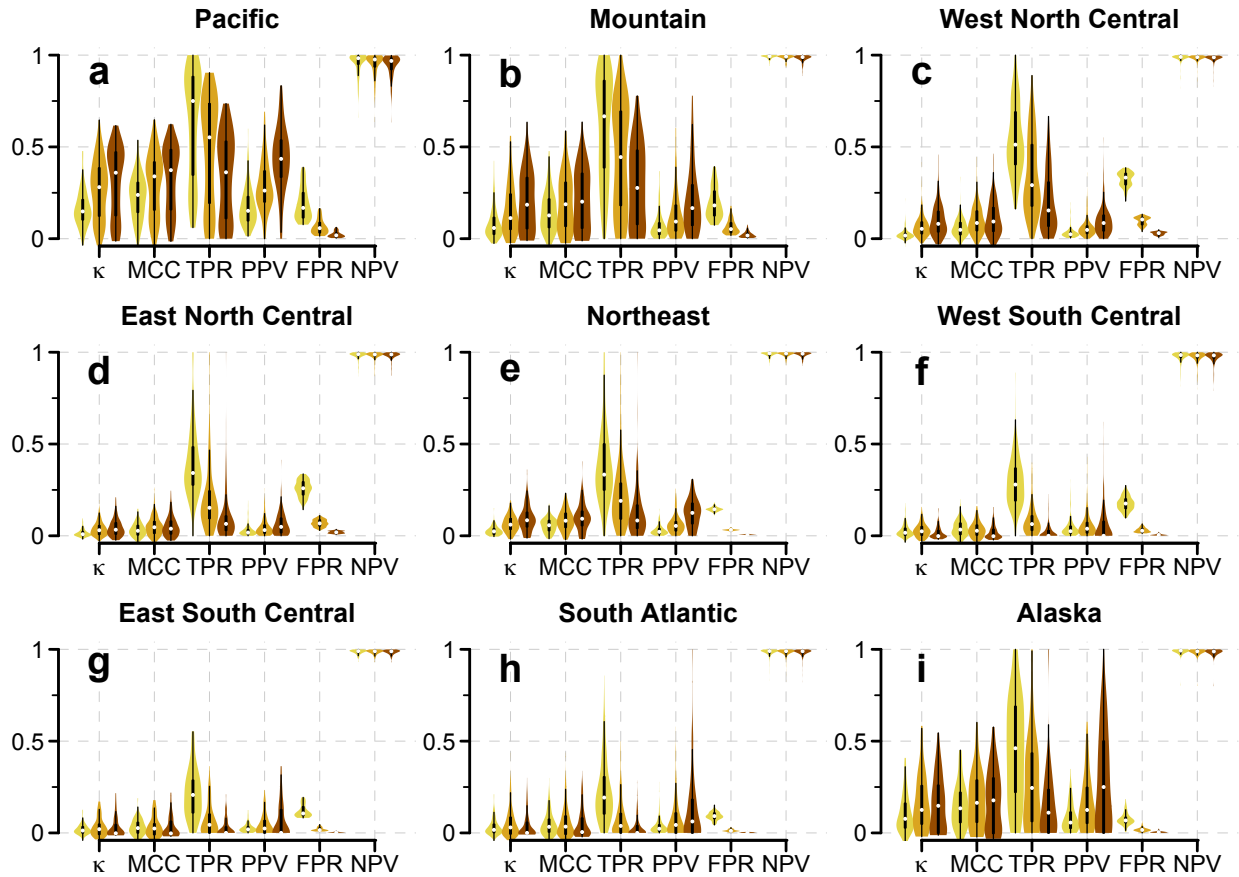
530
531 **Figure 2. Average smoke days across the contiguous United States (CONUS) and Alaska**
532 **from 2008-2021.** Smoke days for each year are derived from: (a) ISD airport smoke
533 observations, (b) HMS medium and heavy smoke plumes, and (c) HMS heavy smoke plumes.
534 HMS smoke days are shown at airport locations, and states in the western U.S. are outlined by
535 the thick border. Values inset indicate the number of total airport locations in CONUS, western
536 U.S., and Alaska.



537
 538 **Figure 3. Trend in smoke days across the contiguous United States (CONUS) and Alaska**
 539 **from 2008-2021.** Trends are derived from: (a) ISD airport smoke observations, (b) HMS
 540 medium and heavy smoke plumes, and (c) HMS heavy smoke plumes. HMS trends are shown at
 541 airport locations, and states in the western U.S. are outlined by the thick border. Values inset
 542 indicate the number of locations in CONUS, western U.S., and Alaska with trends statistically
 543 significant at $p < 0.05$. Trends that are not statistically significant are denoted by small gray dots.

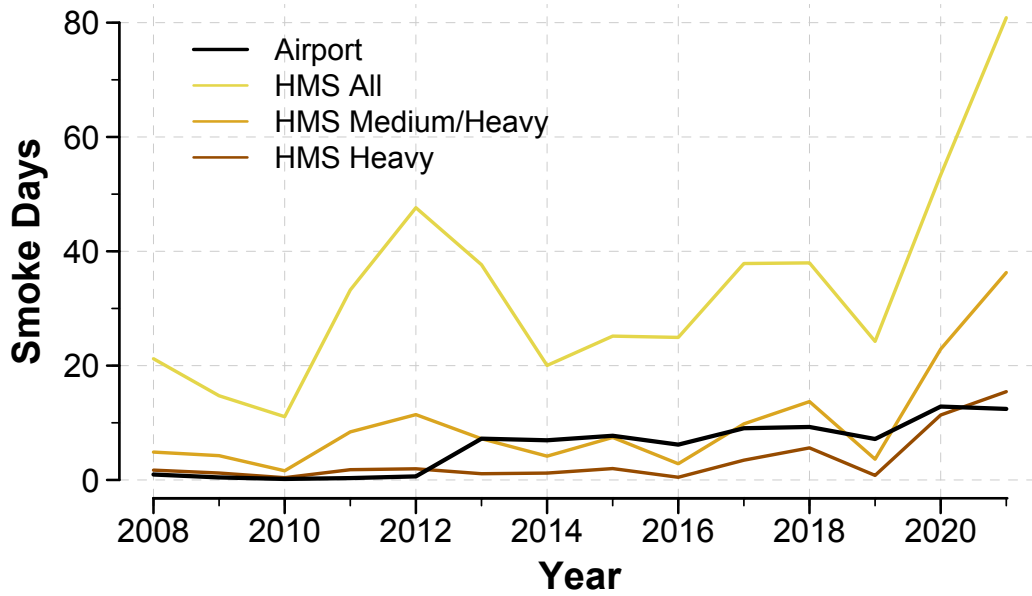


544
 545 **Figure 4. Agreement between airport and HMS smoke days across the contiguous United**
 546 **States (CONUS) and Alaska from 2008-2021.** For HMS, smoke days for each year are derived
 547 from: (a) all smoke plumes, (b) medium and heavy smoke plumes, and (c) heavy smoke plumes.
 548 Agreement is shown at airport locations, and states in the western U.S. are outlined by the thick
 549 border. Inset values denote the number of total airport locations in CONUS, western U.S., and
 550 Alaska. Agreement is shown as Cohen's kappa, where higher values (warmer colors) indicate
 551 greater agreement. Negative Cohen's kappa, or no agreement, are indicated by black dots.

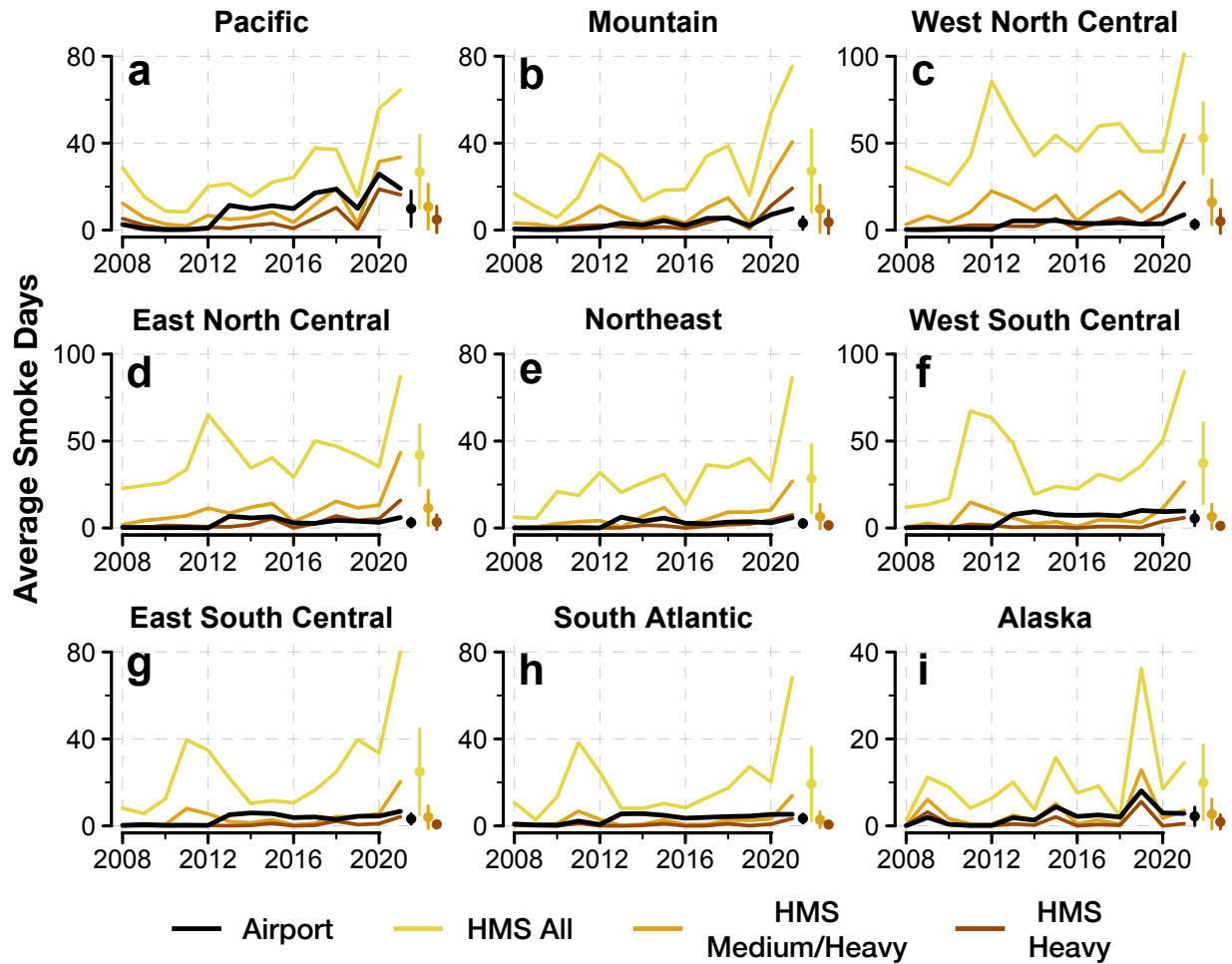


552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563

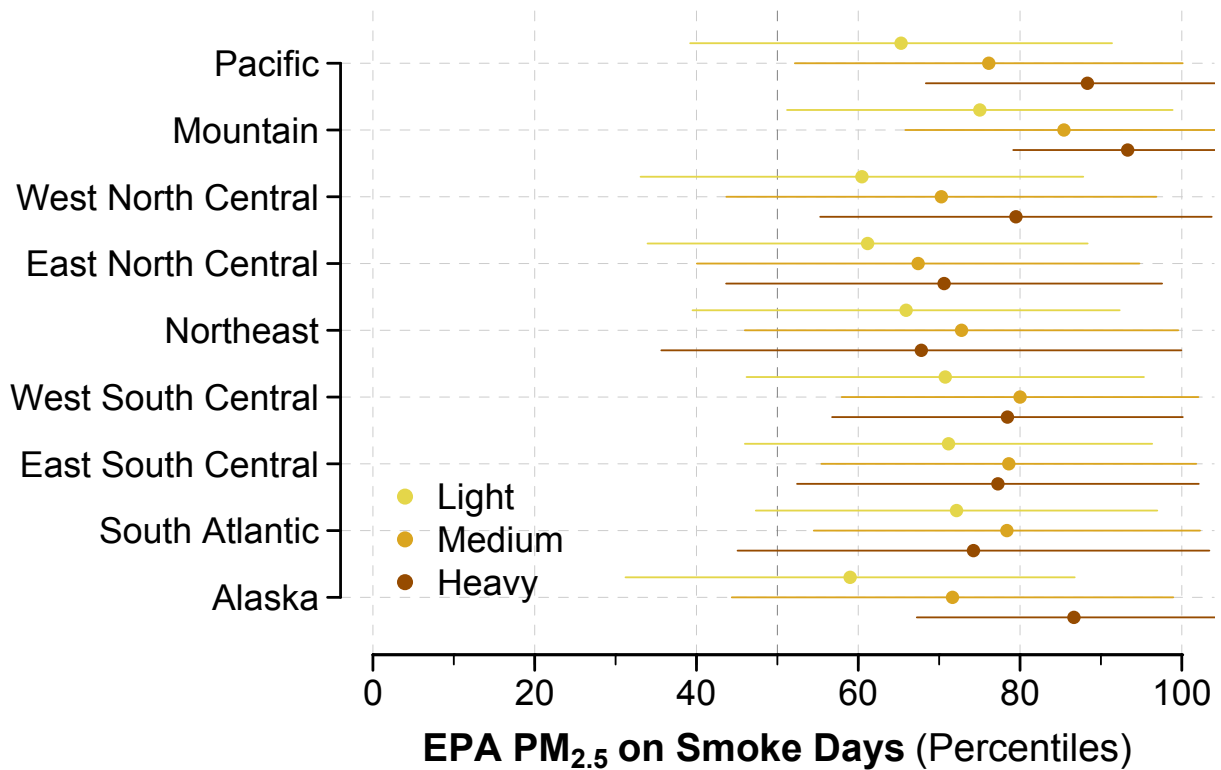
Figure 5. Violin plots of the agreement between HMS and airport smoke days in the United States and Alaska by region from 2008-2021. The violin plot is a hybrid of a box plot and a kernel density plot (as shown by the shape. Smoke days are derived from ISD airport smoke observations and compared to those derived from all HMS smoke plumes (yellow), HMS medium and heavy smoke plumes (goldenrod), and HMS heavy smoke plumes (brown). The agreement metrics – Cohen’s kappa (κ), Matthews correlation coefficient (MCC), true positive rate (TPR), positive predictive value (PPV), false positive rate (FPR), and negative predictive value (NPV) – are spatially averaged across airport locations in each region. A value of 1 for κ , MCC, TPR, PPV, or NPV would indicate perfect agreement, as would a value of 0 for FPR. The plots show that the best agreement between HMS and airport smoke days – e.g., the greatest κ and MCC – occurs in Pacific and Mountain states and Alaska.



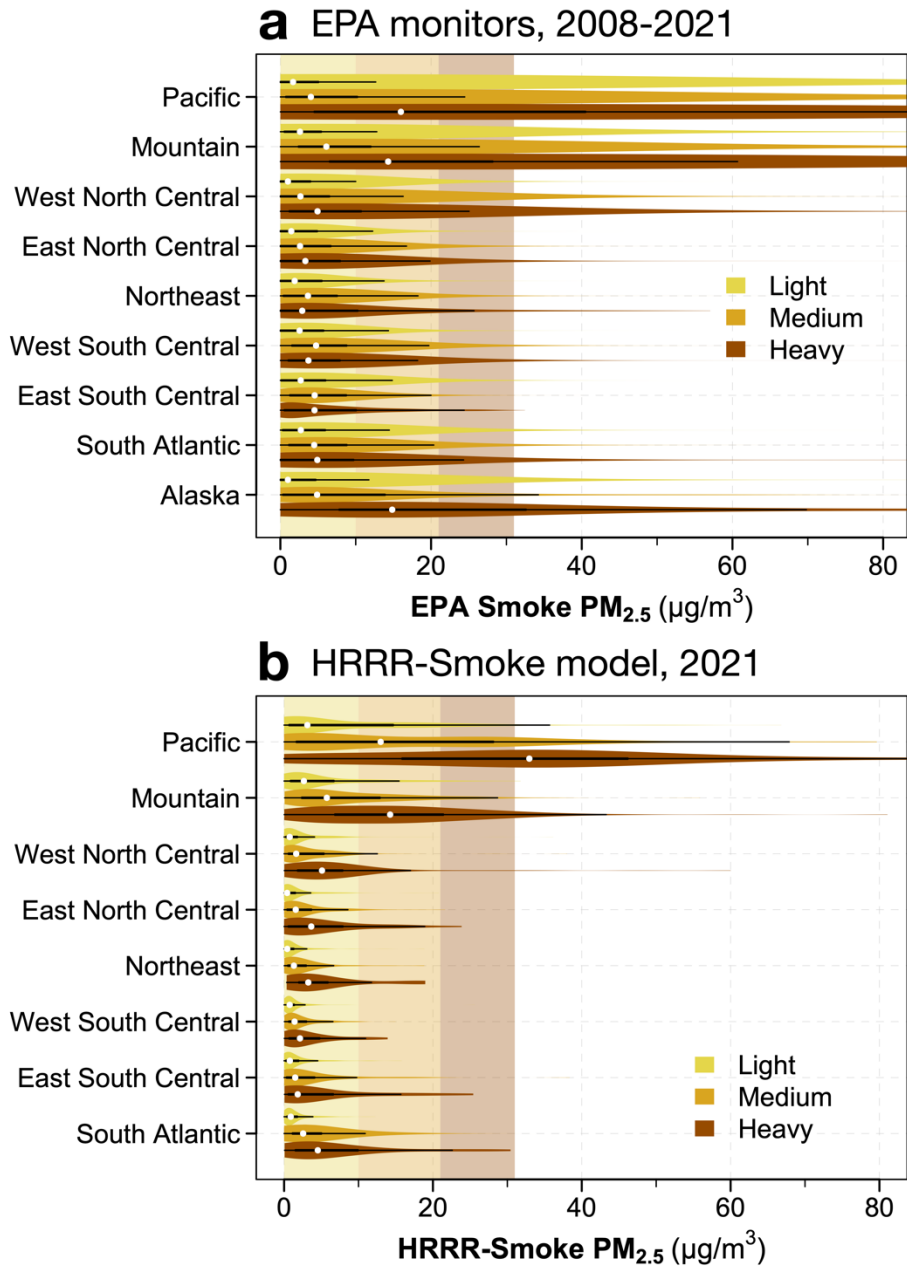
564
 565 **Figure 6. Smoke days in the western United States from 2008-2021.** Smoke days are spatially
 566 averaged across airport locations in the western U.S, as defined by Figure 2, and are derived
 567 from ISD airport smoke observations (black line), all HMS smoke plumes (yellow line), HMS
 568 medium and heavy smoke plumes (goldenrod line), and HMS heavy smoke plumes (brown line).



569
 570 **Figure 7. Smoke days in the United States and Alaska by region from 2008-2021.** Smoke
 571 days are spatially averaged across airport locations in each region, as defined in Figure 1, and are
 572 derived from ISD airport smoke observations (black line), all HMS smoke plumes (yellow line),
 573 HMS medium and heavy smoke plumes (goldenrod line), and HMS heavy smoke plumes (brown
 574 line). Dots to the right of each panel denote annually averaged smoke day number across all
 575 years for the four conditions, with error bars representing one standard deviation.



576
 577 **Figure 8. Separation of PM_{2.5} on smoke and non-smoke days by region at EPA stations.** The
 578 percentile of the PM_{2.5} on an HMS smoke day is calculated relative to the empirical cumulative
 579 distribution of PM_{2.5} on non-smoke days. Smoke days are classified as light, medium, and heavy
 580 according to the designation of HMS plume density on that day; if there are multiple plumes, we
 581 use the maximum HMS density. The dots show the mean percentile, and the horizontal bars
 582 show ± 1 standard deviation across EPA stations in each region. The 50th percentile, denoted by
 583 the vertical gray dotted line, represents the typical value used as the background PM_{2.5}. Higher
 584 percentiles denote more separation between the PM_{2.5} on smoke and non-smoke days and imply
 585 greater confidence in attribution of elevated PM_{2.5} to smoke.



586
 587 **Figure 9. Violin plots of daily smoke $PM_{2.5}$ from EPA monitors and the HRRR-Smoke by**
 588 **region and HMS smoke density category.** The violin plot is a hybrid of a box plot and a kernel
 589 density plot (as shown by the shape). The violin plots show the distribution of daily $PM_{2.5}$ within
 590 light (yellow), medium (goldenrod), and heavy (brown) HMS smoke polygons (a) at EPA
 591 monitors from 2008-2021 and (b) from the HRRR-Smoke model in 2021. The vertically shaded
 592 areas show the equivalent $PM_{2.5}$ ranges for the HMS smoke density categories. For example, the
 593 brown violin for the Northeast U.S. shows the range of EPA and HRRR-Smoke $PM_{2.5}$
 594 concentrations occurring within HMS polygons designated as heavy. The median of this subset
 595 in both the HRRR and EPA datasets in the Northeast (white dots) is $< 10 \mu g m^{-3}$, while the
 596 approximate range of values for heavy HMS smoke is designated as $21-32 \mu g m^{-3}$. This large
 597 mismatch suggests that much of the heavy smoke detected by HMS in this region is likely aloft.