





## Heat flux assumptions contribute to overestimation of wildfire smoke injection into the free troposphere

Laura H. Thapa<sup>1</sup><sup>✉</sup>, Xinxin Ye<sup>1</sup>, Johnathan W. Hair<sup>2</sup>, Marta A. Fenn<sup>2,3</sup>, Taylor Shingler<sup>2</sup>, Shobha Kondragunta<sup>4</sup>, Charles Ichoku<sup>5</sup>, RoseAnne Dominguez<sup>6</sup>, Luke Ellison<sup>7</sup><sup>3,7</sup>, Amber J. Soja<sup>2,8</sup>, Emily Gargulinski<sup>8</sup>, Ravan Ahmadov<sup>9,10</sup>, Eric James<sup>9,10</sup>, Georg A. Grell<sup>10</sup>, Saulo R. Freitas<sup>11</sup><sup>11</sup>, Gabriel Pereira<sup>12</sup><sup>12</sup> & Pablo E. Saide<sup>1,13</sup><sup>✉</sup>

Injections of wildfire smoke plumes into the free troposphere impact air quality, yet model forecasts of injections are poor. Here, we use aircraft observations obtained during the 2019 western US wildfires (FIREX-AQ) to evaluate a commonly used smoke plume rise parameterization in two atmospheric chemistry-transport models (WRF-Chem and HRRR-Smoke). Observations show that smoke injections into the free troposphere occur in 35% of plumes, whereas the models forecast 59–95% indicating false injections in the simulations. False injections were associated with both models overestimating fire heat flux and terrain height, and with WRF-Chem underestimating planetary boundary layer height. We estimate that the radiant fraction of heat flux is 0.5 to 25 times larger in models than in observations, depending on fuel type. Model performance was substantially improved by using observed heat flux and boundary layer heights, confirming that models need accurate heat fluxes and boundary layer heights to correctly forecast plume injections.

<sup>1</sup>Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, Los Angeles, CA, USA. <sup>2</sup>NASA Langley Research Center, Hampton, VA, USA. <sup>3</sup>Science Systems and Applications Inc, Hampton, VA, USA. <sup>4</sup>NOAA NESDIS Center for Satellite Applications and Research, College Park, MD, USA. <sup>5</sup>Department of Interdisciplinary Studies, Howard University, Washington, D.C., USA. <sup>6</sup>Bay Area Environmental Research Institute (BAERI), Moffett Field, CA, USA. <sup>7</sup>NASA Goddard Space Flight Center, Greenbelt, MD, USA. <sup>8</sup>National Institute of Aerospace, Hampton, VA, USA. <sup>9</sup>CIRES, University of Colorado Boulder, Boulder, CO, USA. <sup>10</sup>NOAA Global Systems Laboratory, Boulder, CO, USA. <sup>11</sup>National Institute for Space Research - INPE, São Paulo, Brazil. <sup>12</sup>Federal University of São João del-Rei, São João del-Rei, MG, Brazil. <sup>13</sup>Institute of the Environment and Sustainability, University of California, Los Angeles, Los Angeles, CA, USA. ✉email: [lthapa@ucla.edu](mailto:lthapa@ucla.edu); [saide@atmos.ucla.edu](mailto:saide@atmos.ucla.edu)

Wildfire smoke impacts Earth's radiative budget<sup>1</sup>, air quality<sup>2</sup>, visibility<sup>3</sup>, and human health<sup>4</sup> by increasing the atmospheric load of pollutants. Harsher wildfires and longer fire seasons have been observed and appear to be correlated with anthropogenic climate change<sup>5,6</sup>. Further degradation of air quality and visibility due to smoke are projected in future climate regimes<sup>7</sup>.

Plume rise refers to the vertical transport of hot smoke gasses and particles through the atmosphere. This process may result in plume injections, where smoke reaches the free troposphere, or in plume non-injections, where smoke remains confined within the planetary boundary layer (PBL). Previous studies have shown that injected plumes tend to occupy layers of ambient stability and that plume injections are associated with high fire radiative power (FRP), and favorable fire weather<sup>8–10</sup>. Observations from the Multi-angle Imaging Spectroradiometer (MISR) indicate that injection into the free troposphere occurs in 4–12% of North American smoke plume tops<sup>8,9,11,12</sup>, and the Cloud-Aerosol Lidar with Orthogonal Polarization data indicate that injection of the total smoke column into the free troposphere occurs in 78% of North American smoke plumes<sup>10</sup>. However, the MISR percentage may be biased low due to satellite overpasses occurring in the morning (~10:30 AM LT for MISR), when plumes have not fully developed<sup>11</sup>. Injected plumes can be advected downwind with little dilution<sup>13</sup> and thus tend to have more regional impacts. On the other hand, non-injected plumes tend to have more local impacts due to more efficient downward mixing by ambient PBL turbulence.

The Freitas plume rise parameterization is a 1-dimensional cloud-resolving model. It is typically embedded in a 3-dimensional host model that specifies the ambient environment<sup>14</sup>. This model represents fires as surface buoyancy fluxes that depend on instantaneous fire size, convective heat flux, and fuel type. This model tends to outperform models that assign emissions to a single level or a fixed vertical distribution<sup>15–17</sup>. However, the model depends on estimates of fire size and heat flux, overpredicts free tropospheric injection frequency, and tends to underestimate the range of observed plume heights<sup>18,19</sup>. Proposed causes of the Freitas model's shortcomings include uncertainties in lateral entrainment and in input parameters. Entrainment is dependent on fire size, so versions of the Freitas model allowing variable fire sizes can improve the modeled range

of plume heights<sup>15,18</sup>. In addition, explicitly adding entrainment to a later version of the model improved performance<sup>20,21</sup>. Uncertainty in the input parameters, notably fire size and FRP, may be caused by the plume masking FRP retrievals, incorrect fire shape or size, or variability of burning with vegetation type<sup>18,22</sup>.

Here, we use aircraft and model data from the NASA-NOAA Fire Influence on Regional to Global Environments and Air Quality (FIREX-AQ) field campaign to evaluate the Freitas model in developed plumes. We show that the Freitas model, implemented in the Weather Research and Forecasting Coupled with Chemistry (WRF-Chem) model and the High Resolution Rapid Refresh Smoke (HRRR-Smoke) model, overpredicts injection in comparison with airborne lidar measurements. Injection occurs in 35% of observed plumes and in 80% and 95% of WRF-Chem and HRRR-Smoke plumes, respectively. When the observed boundary layer height is used to evaluate injection in the both models, injection occurs in 59% of the WRF-Chem plumes and 72% of the HRRR-Smoke plumes. Comparing the fire radiant energy flux (FRE flux, "Methods") observed by the Moderate Resolution Imaging Spectroradiometer (MODIS)-Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) airborne simulator (MASTER) instrument with those assumed by WRF-Chem and HRRR-Smoke shows that the models overestimate FRE fluxes by up to a factor of 25 (0.55 kWm<sup>-2</sup> in the observations and 0.29–13.86 kWm<sup>-2</sup> in the models). When WRF-Chem total heat fluxes were reduced so radiant fractions matched observed FRE fluxes, the occurrence of false positives became less common, and further improvement was achieved when the WRF-Chem sensitivity simulations were evaluated using the observed PBL as the injection boundary. This effect was strongest for forest fires, which are assumed by the Freitas model to have the highest heat fluxes<sup>14</sup>.

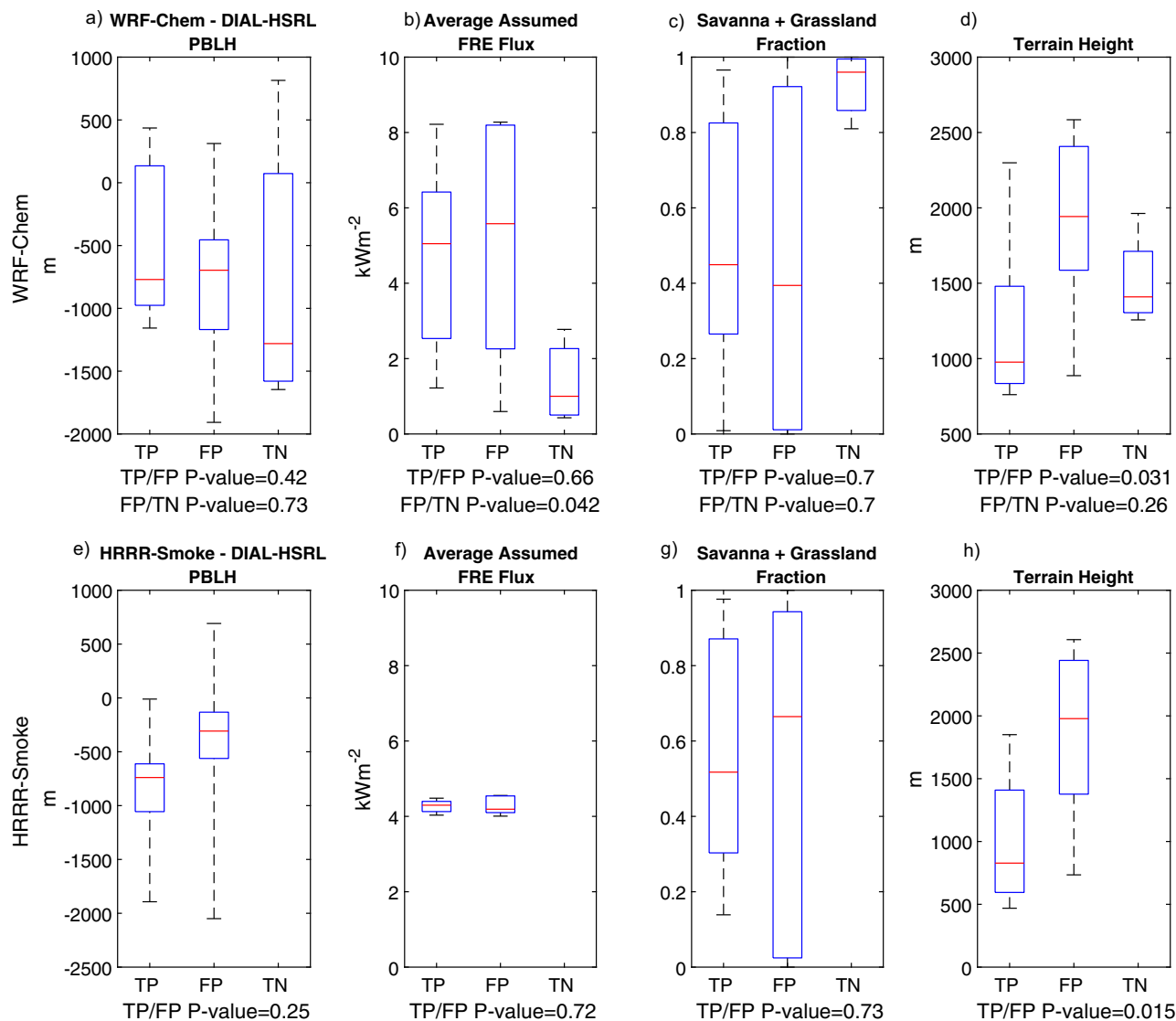
## Results

**Model performance on smoke injection.** Table 1 shows the injection behavior of modeled and observed smoke plume profiles using the observed planetary boundary layer height (PBLH) to classify the observations and the modeled PBLH to classify the models (see "Methods", Supplementary Figs. 1–51, Supplementary Tables 1–3). Injections occur in 35% of observed cases, 80% of WRF-Chem cases, and 95% of HRRR-Smoke cases. WRF-

**Table 1 Confusion matrices for comparing observed and modeled injection behavior counts.**

HRRR-Smoke against HRRR-Smoke PBL	Non-injection forecasted	Injection forecasted	Total
	Non-injection observed	FP = 24	26
	Injection observed	TP = 17	17
	TN = 2	41	43
HRRR-Smoke against DIAL-HSRL PBL	Non-injection forecasted	Injection forecasted	Total
	Non-injection observed	FP = 17	26
	Injection observed	TP = 14	17
	TN = 9	31	43
WRF-Chem against WRF-Chem PBL	Non-injection forecasted	Injection forecasted	Total
	Non-injection observed	FP = 24	33
	Injection observed	TP = 17	18
	TN = 9	41	51
WRF-Chem against DIAL-HSRL PBL	Non-injection forecasted	Injection forecasted	Total
	Non-injection observed	FP = 13	33
	Injection observed	TP = 16	18
	TN = 20	30	51

Counts of injection behavior are shown for HRRR-Smoke using the HRRR-Smoke PBL height to evaluate injection, HRRR-Smoke using the DIAL-HSRL PBL height to evaluate injection, WRF-Chem using the WRF-Chem PBL height to evaluate injection, and WRF-Chem when the DIAL-HSRL PBL height is used to evaluate injection. Columns are labeled with forecasted plume behavior and rows are labeled with observed plume behavior. TN stands for true negative, where non-injection was forecasted and observed. Similarly, TP stands for true positive. FN stands for false negative, where injection was not forecast but did occur. FP stands for false positive, where injection was forecast but did not occur. HRRR-Smoke has fewer total cases than WRF-Chem due to certain fires missing in the model (see "Methods", Sampling Method).



**Fig. 1** Box and whisker plots of model parameters and observations grouped by WRF-Chem and HRRR-Smoke injection behavior. **a, e** Comparison of modeled transect median PBLH with observed transect median PBLH, **b, f** average FRE flux, **c, g** savanna and grassland fraction, and **d, h** terrain height. Red horizontal line denotes medians, blue boxes denote the interquartile ranges, and whiskers denote the 1.5x interquartile range. TP, FP, and FN are defined as in Table 1 caption. **a–d** Refer to the WRF-Chem model and **e–h** refer to the HRRR-Smoke model.

Chem and HRRR-Smoke have accuracies of 0.51 and 0.44, respectively. Both models tend to capture observed injections well, with WRF-Chem and HRRR-Smoke having true positive rates of 0.94 and 1.0, respectively. However, WRF-Chem and HRRR-Smoke also exhibit false positive rates of 0.73 and 0.92, respectively. For WRF-Chem, in cases where the plumes are within the uncertainty range of the modeled PBLH (i.e., cases which exhibit different injection behavior when the modeled and observed PBLHs are used to evaluate injection, see next section), are excluded from the analysis, then accuracy is 0.67, the true positive rate is 1.0, and the false positive rate is 0.61. This only slightly improves on the error achieved by WRF-Chem when all cases are used. These statistics indicate that overprediction of injection leads to the high true positive rate, low accuracy error type present in both models, consistent with previous work which quantified uncertainty in the Freitas model<sup>15,18</sup>.

**Assessing modeled planetary boundary layer height.** Underestimation of the PBLH is a known problem with the WRF-Chem

model<sup>23,24</sup> that may cause plumes placed at the correct height to be misidentified as injections into the free troposphere. Figure 1a and e shows the bias on the modeled PBLH for the WRF-Chem and HRRR-Smoke models, respectively (Supplementary Data 1). For the false positive cases, WRF-Chem tends to underpredict the PBLH by 500–1000 m, and HRRR-Smoke tends to underpredict the PBLH by <500 m. Both models show large spread and even contain cases where the models overpredict PBLH and false injection still occurs. In addition, the WRF-Chem model correctly captures non-injections for generally the most underpredicted PBLHs.

We ran a standalone version of the Freitas plume rise scheme based on simulated meteorology for cases where WRF-Chem underpredicted the PBLH and HRRR-Smoke accurately captured it as determined by overlap in the interquartile ranges of the observed and modeled PBLH distributions (Table 2, Supplementary Fig. 52). It was found that plume height and plume injection behavior have little sensitivity to variations in PBLH present in the meteorology used to drive the model (Supplementary Fig. 53,

**Table 2 Summary of model details.**

Model	WRF-Chem v3.6.1	HRRR-Smoke	1D Freitas
Resolution	4 km horizontal, 53 vertical levels	3 km horizontal, 40 vertical levels	70 vertical levels, 100 m spacing
Initializaiton/Spin-Up	OOUTC initialization, 12 h spin up, 76 h analysis, 84 h forecast, model meteorology passed to Freitas model <sup>34</sup>	00, 06, 12, 18 UTC initializations, 48 h forecasts, model meteorology passed to Freitas model <sup>34</sup>	Uses model profiles of T, P, geopotential, RH, wind, potential temperature, specific humidity.
FRE Flux Assumptions	2.69–13.86 kWm <sup>-2</sup> for forest, 0.4–3.10 kWm <sup>-2</sup> for woody savanna/cerrado, and 0.57 kWm <sup>-2</sup> for grassland/pasture/ cropland.	3.4 kWm <sup>-2</sup> for grassland to 4.8 kWm <sup>-2</sup> for forest for 14 vegetation categories as defined by the IGBP <sup>28,43,44</sup>	Choose the Freitas maximum heat flux for the dominant fuel category
Fire Size Assumptions	0.25 km <sup>2</sup> per burning grid cell and per fuel category <sup>14,45</sup>	Converts MODIS/VIIRS FRP into area using fuel-varying heat fluxes	0.25 km <sup>2</sup> per burning grid cell and per fuel category <sup>14,45</sup>
Flaming Fraction of Emissions Assumption	50% flaming, undergo plume rise	75% flaming, undergo plume rise	N/A
Key Parameterizations	Rapid Radiative Transfer Model for GCMs <sup>45</sup> , Noah Land Surface Model <sup>46</sup> , aerosol-aware microphysics <sup>47</sup> , Grell-Freitas cumulus <sup>34</sup> , Mellor-Yamada-Janjić PBL <sup>48</sup> , MEE = 3.5 <sup>49</sup> , MOSAIC Aerosols <sup>50</sup>	Aerosol-aware microphysics <sup>47,51</sup> Mellor-Yamada-Nakanishi-Niino PBL scheme <sup>52</sup>	N/A
Data Assimilation	Near-real-time inversion of MAIAC AOD V6 <sup>53,54</sup> based on 6 largest fires in the region <sup>19</sup>	None	N/A
Emissions Source	QFED v2.4, 0.1 degrees resolution, regridded <sup>38</sup>	Top-down from MODIS/VIIRS FRP <sup>39</sup>	N/A
Variables Extracted	PM2.5 curtains, model PBL and theta-derived mixed layer heights <sup>19</sup> , land use fractions, Richardson number and Brunt-Vaisala frequency via centered-difference at twice and half the PBLH.	PM2.5 curtains, model PBL height, Richardson number and Brunt-Vaisala frequency calculated via centered-difference at twice and half the PBLH.	Maximum plume top height

The first column lists model details for WRF-Chem, the second feature lists model details for HRRR-Smoke, and the third column lists model features for the 1D Freitas plume rise model.

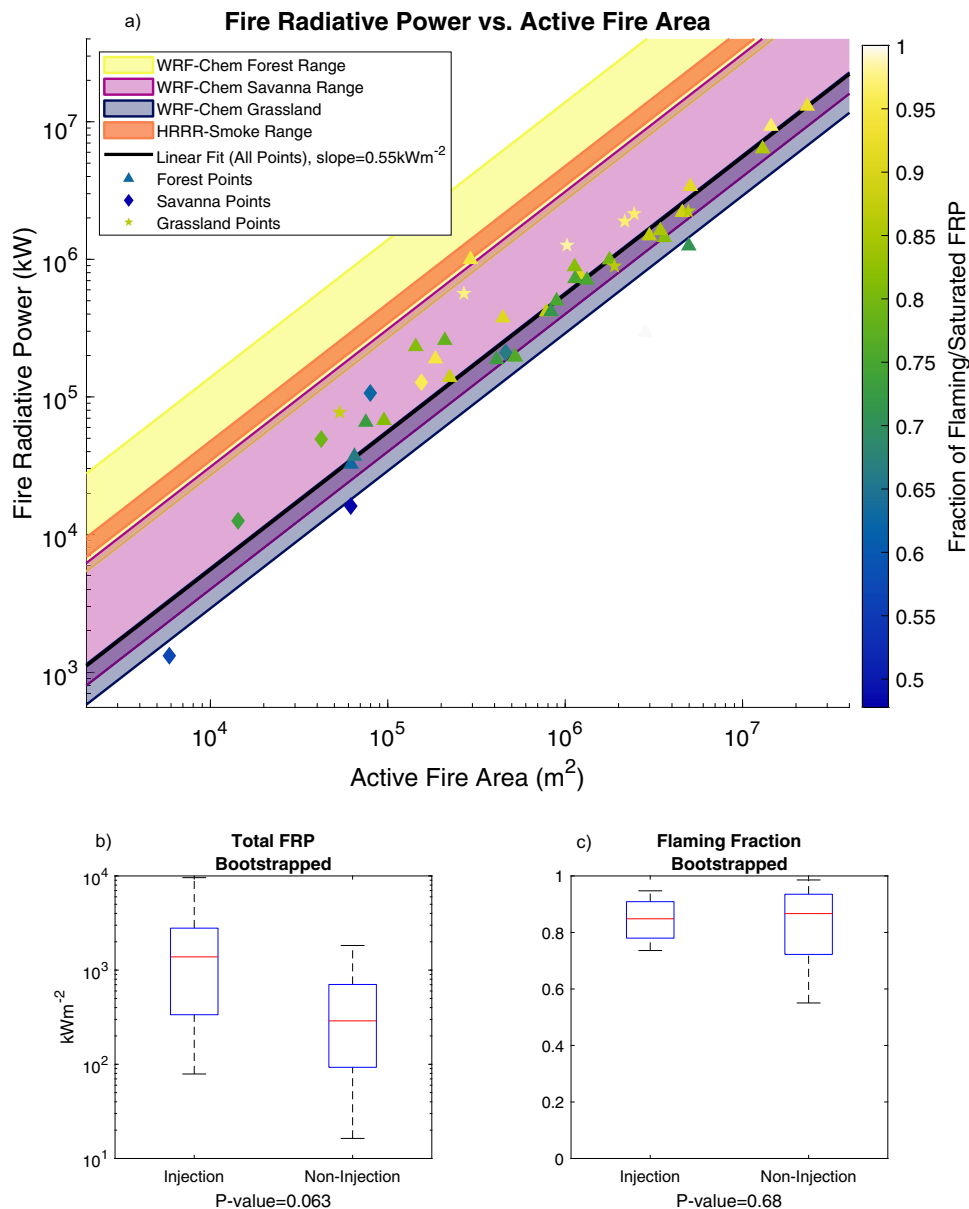
note percent error on plume heights do not exceed 8.8%). This allows us to re-classify injection behavior for WRF-Chem using the observed PBLH (Table 1) since plume top height would likely be similar if WRF-Chem was reproducing the observed PBLH correctly. There are 12 cases that get reassigned to the true negative category, two that get reassigned to false positive, and one that gets reassigned to false negative. For this evaluation, injection occurs in 59% of cases, resulting in an accuracy of 0.70, a true positive rate of 0.89, and a false positive rate of 0.39, indicating better model performance. Therefore, we conclude that for WRF-Chem, uncertainties in the boundary layer likely do contribute to the overprediction of injection. When the observed PBLH was used to evaluate the injection behavior of HRRR-Smoke, the injection rate was 59%, the accuracy was 0.53, the true positive rate was 0.82, and the false positive rate was 0.65. For HRRR-Smoke, which runs an newer PBL scheme (Table 2), uncertainties in the boundary layer are likely smaller and play less of a role in generating false injections.

**Associating model performance with additional fire characteristics.** Figure 1 shows distributions of three additional variables (Supplementary Data 1) that can help explain the injection behavior of the WRF-Chem and HRRR-Smoke models: average of the model's assumed FRE fluxes (see "Methods", 1b, 1f), model fuel type (1c, 1g), and terrain height (1d, 1h). Since FRE flux is assumed to depend on fuel type in the model, assumed FRE flux and fuel type are strongly correlated which each other. Other variables showing strong correlations with each

other include free tropospheric and boundary layer stability, and terrain complexity and heat flux in the HRR-Smoke model (Supplementary Figs. 54–55). The rest of the variables are at most moderately correlated with each other ( $r^2 < 0.46$ ). Box plots were generated using bootstrapping to avoid sampling bias towards any single fire ("Methods").

Figure 1b and f imply that modeled FRE fluxes are likely a strong contributing factor to the overprediction of injection by both models. WRF-Chem captures non-injections correctly when a lower FRE flux is prescribed and either overpredicts or captures injections for larger FRE fluxes. In fact, the difference in FRE flux between correct non-injection and false injection cases (which in the observations are showing the same non-injection behavior) is statistically significant ( $p$ -value = 0.042). While HRRR-Smoke did not contain enough correct non-injection cases for a meaningful comparison, HRRR-Smoke assumes similar FRE fluxes for all fuel types, and these values are comparable with what WRF-Chem assumes for cases where false injection is forecasted. Therefore, we may infer that these high heat fluxes contribute to HRRR-Smoke overpredicting injections as well.

Capturing non-injections correctly tends to occur in WRF-Chem for larger fractions of savanna and grasslands (Fig. 1c), which is consistent as WRF-Chem prescribes lower heat fluxes for savanna and grassland than for forests ("Methods"). Similarly, there is no clear relationship for HRRR-Smoke regarding fuel type as FRE fluxes assumed for all fuels are similar and within the range of that WRF-Chem assumes for forest. This indicates that the fuel-dependent heat fluxes assumed by the models ("Methods") specifically those corresponding to forest fuels in WRF-



**Fig. 2 Comparison of observed FRE flux and assumed FRE fluxes for WRF-Chem and HRRR-Smoke.** **a** Scatter plot of total FRP and flaming and saturated fire area from MASTER for each fire overpass (colored diamonds, stars, triangles). Black line is a linear fit to the MASTER points of the form  $y = mx$ . Colored bars represent the ranges of FRE flux based on models' assumptions for convective heat flux and uncertainty in converting between convective and radiant heat flux (see "Methods"). FRE flux increases toward the upper left. Shapes represent the campaign fuel category. Colors of diamonds, stars, and triangles, represent the fractional contribution of flaming combustion (including saturated pixels) to total FRP. Other panels show box and whisker plots of **b** observed FRP and **c** observed fraction of FRP due to flaming combustion, grouped by observed injection behavior. Box and whisker elements are as in Fig. 1 caption.

Chem and all fuels in HRRR-Smoke, could be high and lead to an overprediction of injection. Other variables such as time of day, static stability, and terrain complexity were investigated on their role on modeled injection behavior but none resulted in being a clear explanatory variable (Supplementary Figs. 54–55).

Figure 1d and h shows that false injections occur at higher terrain elevation than correct injections and that this difference is statistically significant across both models ( $p$ -values = 0.031 and 0.015). In addition, Fig. 1d shows that false injections occur at higher elevations than correct non-injections. Likely, this means that the plume rise scheme's input heat fluxes or meteorology are more uncertain in areas of higher (Fig. 1d, h) and more complex (Supplementary Figs. 54b, 55b) terrain. The Fretias model is also

limited in that it is 1-dimensional and thus cannot simulate interactions between neighboring model columns.

**Assessing heat flux assumptions.** Figure 2a provides an observational constraint on the FRE flux. Scatter points show observations of total FRP versus flaming fire area from the MASTER instrument for each transect and shapes denote the campaign fuel category for each MASTER overpass ("Methods", Supplementary Data 1). Colored regions denote the FRE flux assumptions made by the models. Active fire area is computed for flaming and saturated pixels only, assuming that these fire phases primarily contribute to plume injection<sup>14</sup>. FRP is computed for all categories, as this is

what satellites observe. Plots showing fire size and FRP for different combustion phases show similar qualitative results (Supplementary Fig. 56). MASTER, WRF-Chem, and HRRR-Smoke all track heat fluxes slightly differently, so for ease of comparison all have been converted into FRE flux (see “Methods”).

Figure 2a shows that the best fit for the observed FRE flux ( $0.55 \text{ kWm}^{-2}$ ) is within the range of the WRF-Chem savanna ( $0.40\text{--}3.09 \text{ kWm}^{-2}$ ) and grassland ( $0.29\text{--}0.57 \text{ kWm}^{-2}$ ) FRE fluxes and is not within the range the WRF-Chem forest FRE flux ( $2.69\text{--}13.86 \text{ kWm}^{-2}$ ) or any assumed HRRR-Smoke FRE fluxes ( $3.44\text{--}4.76 \text{ kWm}^{-2}$ ). Put another way, WRF-Chem and HRRR-Smoke assume FRE fluxes equal to 0.5–25 and 6–9 times the average observed FRE flux, respectively. This is consistent with the general lack of overlap between the scatter points and the yellow and orange colored bars in Fig. 2a. Thus, this provides observational evidence for the hypothesis that the assumed heat fluxes that get used in the Freitas model, specially for forest fuels, are likely too large. Since these heat fluxes are directly used to calculate initial plume buoyancy, this leads to simulated plumes which are more easily lofted into the free troposphere.

A caveat of this finding is that a fraction of the flaming pixels retrieved by MASTER are flagged as saturated, and thus FRP could be underpredicted, affecting our conclusions. Comparing MASTER to satellite observations (which are less prone to saturation due to larger pixel footprint) shows that MASTER tends to provide FRP estimates in between both MODIS (slope = 1.65) and the Visible Infrared Imaging Radiometer Suite (VIIRS, slope = 0.79) (Supplementary Fig. 57). If we assume that VIIRS retrievals are closer to reality, we can use the slope of the scatter plot to correct MASTER retrievals, yielding an average observed FRE flux of  $0.7 \text{ kWm}^{-2}$  which does not change our findings. Therefore, we do not expect saturation of MASTER to influence our final result to a large degree.

To investigate drivers of variability of FRE flux, FRE flux was calculated for MASTER points grouped by the dominant campaign fuel category (see “Methods”) and by the flaming fraction of FRP. This resulted in FRE flux of  $0.55 \text{ kWm}^{-2}$  for forest fires,  $0.51 \text{ kWm}^{-2}$  for savanna fires,  $0.59 \text{ kWm}^{-2}$  for grassland fires,  $0.56 \text{ kWm}^{-2}$  for >85% flaming fraction, and  $0.36 \text{ kWm}^{-2}$  for <85% flaming fraction. FRE flux is generally constant across fuel type and appears to vary more with fire combustion phase. In other words, given the right spread and fuel consumption conditions, forest and grass fires may burn with the same intensity, which is consistent with previous studies<sup>25</sup>. Thus, heat fluxes used to drive the Freitas model should be small, relatively constant across fuel types, and perhaps more closely associated with fire weather.

Figure 2a supports the idea that model heat flux assumptions assign too much buoyant energy per unit area, which is likely the primary contributor to injection overprediction. The overprediction occurs even when a small fire size is assumed in WRF-Chem ( $0.25 \text{ km}^2$  per grid cell), further supporting the idea that the models assume too much fire energy per unit area. In addition, WRF-Chem tends to correctly identify non-injection events in grass and savanna-dominated fires (Fig. 1b), as these fires have the lowest modeled heat flux. FRE flux used in HRRR-Smoke for all fuel types are generally in the range of FRE fluxes for forest specified in ref. 14, consistent with HRRR-Smoke showing a high rate of false injection regardless of fuel type (Table 1). HRRR-Smoke’s internal calculation of total heat flux from FRE flux (see “Methods”), likely keeps plume top heights from being unphysically high. Thus, heat fluxes used to drive the Freitas model should be updated to be consistent with the smaller observed FRE fluxes. The linear relationship between FRP and burned area (black solid line) is consistent across several orders of

magnitude which can be used to update the HRRR-Smoke scheme where burned area is a function of FRP.

Figure 2b shows observed injection behavior as a function of observed FRP showing a near-significant trend ( $p$ -value = 0.063) of injections happening at higher FRP. Thus, even though the heat fluxes need to be updated in HRRR-Smoke, there is value in driving plume injection with observed FRP, as is already implemented in HRRR-Smoke (see “Methods”) and other systems (e.g., ref. 26). MASTER FRP is highly correlated to burned area ( $R^2 = 0.96$ ), and thus approaches using observed burned area to drive plume rise would also be appropriate. On the other hand, Fig. 2c shows that observed injection behavior has little association with flaming fraction of MASTER pixels. However, flaming fraction does play an indirect role, as higher flaming fractions are associated with higher FRE fluxes, and thus heat flux could be implemented as a function of flaming fraction if this is available from observations.

**Sensitivity analysis.** The final step in this analysis was to re-run the WRF-Chem model for selected cases with an adjusted heat flux based on the observed average FRE flux of  $0.55 \text{ kWm}^{-2}$  (total heat flux of  $3.4\text{--}5.3 \text{ kWm}^{-2}$ ). The sensitivity test was run on the WRF-Chem false injections, and a total of 24 of the 51 cases were re-evaluated. These cases are shown as a fourth panel in the relevant transects in Supplementary Figs. 1–51. WRF-Chem model setup for the sensitivity simulations is as shown in Table 2, and HRRR-Smoke is not included in the sensitivity analysis as it was being run in only as a real-time forecast product for the FIREX-AQ period.

Of the 24 cases which were re-run with updated heat fluxes, 18 remained false injections and 6 were re-classified correctly as non-injections. This yields an improved false positive rate of 0.75, given that the initial false positive rate of this subset was 1. If injection behavior is evaluated using the observed PBLH as the injection boundary, then 7 cases remain false positive and 17 cases become correctly re-classified as non-injections for a false positive rate of the subset equal to 0.29. The sensitivity analysis shows that both an accurate boundary layer and an accurate heat flux are needed to correctly model injection behavior.

On average across these 24 cases, in-plume median mixed layer height (a proxy for smoke top height<sup>19</sup>) decreased by 455 m with the inclusion of more realistic heat fluxes. Grouped by model fuel category, this corresponds to a 687 m decrease for extratropical forest fires (14 cases), a 141 m decrease for savanna fires (7 cases), a 20 m decrease for grass fires (3 cases). Of the six cases that were re-classified as non-injections with respect to the WRF-Chem PBLH, 5 had extratropical forest as their dominant model fuel category and 1 was a savanna fire. All but one saw a reduction in the median in-plume smoke top height (average decrease of 506 m across the 6 cases). Of the 18 cases which remained injections, 9 were extratropical forest-dominated, 6 were savanna-dominated, and 3 were grassland-dominated.

Using heat fluxes and boundary layers that are in line with observed values reduces the occurrence of false positives and reduces mean smoke top height for the WRF-Chem model. This effect is consistent with other studies which have shown that modeled plume top heights decrease in proportion to decreasing heat flux values<sup>18</sup>, with the fact that the Freitas model assumes the highest heat flux values for forest fires<sup>14</sup>, and with studies that have looked at the underprediction of the PBLH by WRF-Chem<sup>23,24</sup>. While we note that classifying injection behavior with respect to the observed boundary layer does improve model performance, further improvements on WRF-Chem PBLH are expected by using more recent versions of the model (the one

used here dates to 2014) which are more consistent with the HRRR-Smoke model's reduced biases in PBLH.

## Discussion

We compared smoke injection behavior in the WRF-Chem and HRRR-Smoke models against airborne lidar observations and found that both models tend to overpredict injection behavior. Unrealistically high heat fluxes used to drive buoyancy of plumes is identified as a likely contributor to the high rate of false injections in both models, and underprediction of PBLH is identified as a second contributor to the high rate of false injections in the WRF-Chem model. Observations of FRP and fire area from the MASTER instrument provide an observational constraint on heat flux, which we used to determine that the assumed FRE fluxes in WRF-Chem (HRRR-Smoke) are generally higher than observed by a factor of 0.5–25 (6–9). The observations mainly fell within the ranges of FRE flux assumed in WRF-Chem for grassland and savanna fires (average FRE flux =  $0.55 \text{ kWm}^{-2}$ ), consistent with WRF-Chem correctly predicting non-injections for high fractions of grassland and with HRRR-Smoke predicting a high rate of false positives. FRE flux was shown to vary weakly with fuel type and more strongly with fraction of FRP due to flaming combustion. Adjusting the heat fluxes used in WRF-Chem based on field data partially improves the overprediction behavior that dominated the error in this model, and further improvement is achieved when the model injection behavior is evaluated using the observed PBLH as the injection boundary. Sensitivity tests for HRRR-Smoke were not performed, but we hypothesize that reducing heat fluxes will improve injection representation, although this effect may be somewhat damped by HRRR-Smoke's internal calculation of total heat flux.

The FIREX-AQ dataset we used is limited in that it targets only Western United States wildfires during the relatively normal 2019 fire season. In addition, due to the nature of observing wildfire smoke with an aircraft, more intense and lofted plumes were preferred by the flight planners. In spite of these limitations, the western FIREX-AQ dataset contains a range of fire sizes, injection behaviors, and fuel types (Fig. 2a). In addition, these results are consistent with studies which used satellites to evaluate plume rise<sup>15,18</sup>.

Our results also show that heat fluxes may not be dependent on fuel in the way the Freitas model assumes. The Freitas plume rise parameterization was originally developed for tropical forest fires<sup>14,18</sup>, so it is possible that these constants are tuned to this region and thus not wholly applicable to Western US wildfires.

Past studies have identified shortcomings with the Freitas model, including underpredicting the injection height range and misidentifying injections<sup>18</sup>. The data available from FIREX-AQ have allowed us to pinpoint the assumed heat fluxes and boundary layer uncertainty contributors to the known issues with the performance of the Freitas model. We expect our results will improve smoke modeling across a variety of disciplines, including air quality and visibility forecasting, wildfire-climate interactions, and health impact studies.

## Methods

**Campaign data.** The FIREX-AQ field campaign targeted smoke plumes using the DC-8 airborne laboratory. This dataset contains high-resolution measurements of plume backscatter, FRP, fire area, and fuels burned for each fire. In addition, WRF-Chem and HRRR-Smoke simulations implement slightly different versions of the Freitas parameterization. WRF-Chem was run in near-real time to support flight planning efforts and HRRR-Smoke is a NOAA operational system.

An airborne Differential Absorption Lidar-High Spectral Resolution Lidar (DIAL-HSRL<sup>27</sup>) provided aerosol backscatter coefficient at 532 nm (vertical resolution = 30 m, horizontal resolution = 10 s) to evaluate the plume injection height. FRP, area burned, and fire phase measurements (flaming, smoldering, or

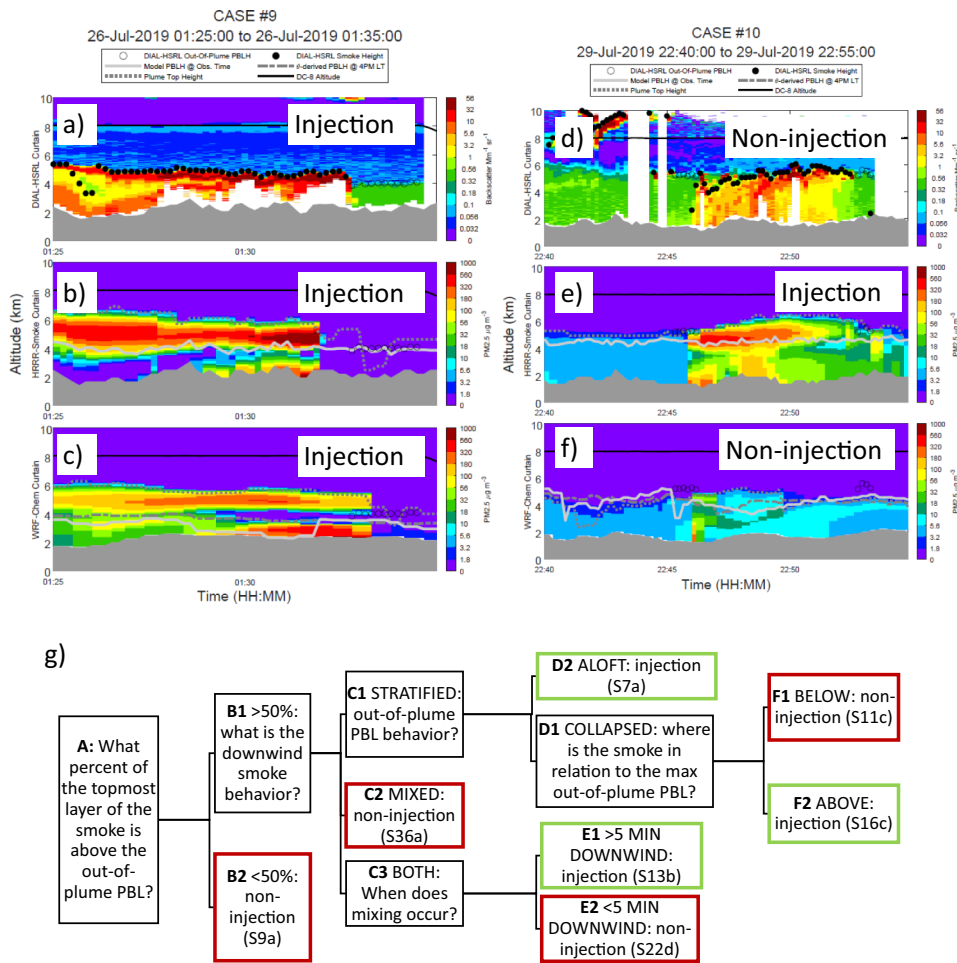
saturated) were taken using the MASTER multispectral imager<sup>28</sup>. For each fire, the campaign fuel category was determined as the dominant fuel category according to the Fuels Characteristic Classification System (FCCS), a high-resolution dataset built using remote sensing and in-situ measurements of land cover and fuel loading. In this study, FCCS fuel categories were mapped to the fuel categories assumed by the MODIS-International Geosphere Biosphere Programme (IGBP)<sup>29,30</sup>. A comparison of modeled and campaign fuels (Supplementary Fig. 58), shows that the model uses the correct fuel type in 47% of the cases. Since the fuel type directly influences the assumed heat flux, errors in model fuel can lead to the heatflux being over or under estimated with respect to heat flux that would be assumed had the campaign fuel been used to determine the heat flux. A majority of the errors (present in 39% of the cases) work to reduce the heat flux, and in spite of this heat flux estimates are still too high.

**MASTER algorithm and pixel saturation.** In this study, FRP retrievals and combustion phase classes from the MASTER instrument are used to evaluate the models' assumptions of fire heat flux. Fire detection from MASTER is based on the contextual algorithm concept originally developed for the MODIS instruments aboard the Terra and Aqua satellites<sup>31</sup>. Regarding the retrieval of FRP, previous studies indicate that 98% of the variance in fire radiative energy is explained by variations in the emitted mid-infrared (MIR) spectral radiance at  $\sim 4 \mu\text{m}$ <sup>32</sup>. Thus, retrieval of FRP from MASTER observations is based on the MIR radiance approach<sup>33</sup> using the MASTER 3.91  $\mu\text{m}$  spectral band (hereafter 4  $\mu\text{m}$ ). To classify MASTER-detected fire pixels by combustion phase (i.e., flaming and smoldering), efforts to derive and apply constant thresholds were unsuccessful, as the range of values representing these phases varied widely from scene to scene, partly because of differences in the fire regimes and variations of pixel fractions with different fire phases, and partly because of complex instrument radiometric calibration dynamics beyond the scope of this discussion<sup>34</sup>. Thus, an approach similar to the contextual fire detection algorithm<sup>31</sup> is used to accomplish this classification, with thresholds empirically derived from each scene of the MASTER imagery. Flaming pixels were determined as fire pixels for which the 4  $\mu\text{m}$  brightness temperature is higher than 3 standard deviations above the background mean 4  $\mu\text{m}$  brightness temperature and for which the difference between the 4 and 11  $\mu\text{m}$  brightness temperatures exceeds 100 K. Smoldering pixels were determined as fire pixels for which the 4  $\mu\text{m}$  brightness temperature is higher than 2 standard deviations above the background mean 4  $\mu\text{m}$  brightness temperature and for which the 11  $\mu\text{m}$  brightness temperature is higher than 1 standard deviation above the mean background 11  $\mu\text{m}$  brightness temperature. Since some of the more intense fire pixels were found to be saturated due to the limited dynamic range of the MASTER 4  $\mu\text{m}$  fire channel relative to the high temperatures that can be attained by pixels which are fully covered by flames (up to 1500 K), saturated pixels were determined using the 4  $\mu\text{m}$  radiance values. Given the scene with 20% valid background pixels, these pixels with 4  $\mu\text{m}$  radiance values  $>99.5\%$  of max value are flagged as saturated. These maximum values may vary from scene to scene and even scan line to scan line, but in a given scan line, they are generally equal (within 0.005%). When saturation occurs, clusters of pixels around saturated pixels tend to have the same 4  $\mu\text{m}$  radiance value. These pixels are also flagged as saturated, which is a limitation of this method.

Preliminary analysis (Supplementary Fig. 57) showed that MASTER FRP is comparable to FRP from MODIS, VIIRS, and the Geostationary Operational Environmental Satellite-Advanced Baseline Imager (GOES-ABI). MASTER tends to underpredict VIIRS FRP measurements (bias =  $-66.11 \text{ MW}$ ), and overpredict MODIS (bias =  $379.6 \text{ MW}$ ). In the MODIS plot, the strong association ( $R^2 = 0.97$ ) is driven in large part by the outlier, but removing the outlier still results in MASTER overpredicting MODIS ( $R^2 = 0.72$ , bias =  $190,215.3 \text{ MW}$ , slope = 1.39). See Supplementary Fig. 55 for the full comparison. Note that not all cases where we have plumes have associated MASTER data.

**Modeling systems.** In this study, we compared the Freitas plume rise parameterization in WRF-Chem and HRRR-Smoke, the details of each are listed in Table 2. The relevant differences for this research are in the way each model deals with heat flux and fire size. WRF-Chem assumes a constant fire size of  $0.25 \text{ km}^2$  per model grid cell and fuel-dependent total heat fluxes that range from  $4.4 \text{ kWm}^{-2}$  for grass fires to  $80 \text{ kWm}^{-2}$  for forest fires. HRRR-Smoke, on the other hand, assumes fuel-dependent FRE fluxes which range from  $3.44\text{--}4.76 \text{ kWm}^{-2}$  for all fuel types and uses these constants along with measurements of FRP to estimate fire size. Since HRRR-Smoke assumes FRE fluxes, it does an internal conversion to total heat flux using a factor of 1.6<sup>35,36</sup>. In both models, convective energy is computed using prescribed or calculated total heat flux based on the fuel category and assuming that convective energy is 55% of the total heat flux<sup>36</sup>. Both models assume that only flaming emissions undergo plume rise and assume a fixed fraction of emissions are flaming. The plume rise module is run every 30 min of model time.

The Quick Fire Emissions Dataset (QFED) emissions used in WRF-Chem were regridded to the model resolution by assigning emissions to the closest grid cell. The QFED daily emissions were regridded to hourly resolution by assigning a percentage of the daily emissions to each hour, with the peak of the diurnal cycle occurring at 16:00 LT and containing 17% of the daily emissions. In addition, this diurnal cycle assumes that a majority of the emissions are released during daytime



**Fig. 3 Selected transects with their corresponding injection behavior identified, and decision tree for classifying injections.** Vertical profiles of **a** DIAL-HSRL backscatter, **b** HRRR-Smoke PM<sub>2.5</sub>, and **c** WRF-Chem PM<sub>2.5</sub> for the Shady Fire on 2019-07-26 01:25-01:35 UTC. Vertical profiles of **d** DIAL-HSRL backscatter, **e** HRRR-Smoke PM<sub>2.5</sub>, and **f** WRF-Chem PM<sub>2.5</sub> for the North Hills Fire on 2019-07-29 22:40-22:55 UTC. **g** The decision tree that is used to classify each case. In the terminal nodes (red and green boxes) representative cases for each path ending at that terminal node are given in parentheses. Open circles in **(a)** and **(d)** denote out of plume boundary layer height, and filled circles denote plume top heights. Light gray solid lines and short dashed gray lines in **(b, c)** and **(e, f)** denote the modeled PBLH and mixed layer heights derived from vertical PM<sub>2.5</sub> gradients, respectively. For WRF-Chem, the mixed layer height at 4PM LT, the time when the mixed layer is thickest, is shown the long dashed dark gray line. 5 min of flight time corresponds to 50–70 km.

hours (around 10:00–19:00 LT) and less than 1% of the daily emissions are released per hour during nighttime hours<sup>37</sup>.

**Sampling method.** Supplementary Figs. 1–51 show the 51 transects that were selected for analysis in this work. Observations were taken during transects wherein the aircraft flew over the fire source and upwind or downwind along the longest axis of the plume. Modeled columns of particulate matter less than 2.5 microns in diameter (PM<sub>2.5</sub>) concentration were extracted along the flight path, considering fires that were slightly shifted in space and time in WRF-Chem relative to observations (Supplementary Figs. 59–60). All cross sections are shown in Supplementary Figs. 1–51 and Supplementary Tables 1–2 indicate injection behavior of all samples. Terminal nodes of the classification (Fig. 3g) are used to determine which PBL to evaluate injection against and to compare with the DIAL-HSRL PBLH.

In a few cases, the aircraft observed fires which are not present in the HRRR-Smoke model. This is due to data gaps in the satellite fire detections, which means that the FRP was not ingested into the model and thus fire emissions and plume height are never derived for those fires. This issue impacted the Shady fire (observed July 24, 2019) and the Tucker Fire (observed July 29, 2019). These missing detections are likely due to clouds in the area, which could lead to missed satellite FRP retrievals (Tucker), or persistent clouds dampening fire behavior and leading to missed detections even after the clouds cleared (Shady). These missing HRRR-Smoke cases are present in the WRF-Chem dataset, because the QFED emissions inventory has the capability to estimate emissions for cloud-covered regions<sup>38</sup>. Because these cases have WRF-Chem curtain plots available, they are present in our analysis, but for the HRRR-Smoke samples, the injection behavior has been labeled as not a number (nan, Supplementary Tables 1–2).

In about half of the WRF-Chem cases, modeled fires appear shifted with respect to observations. When sampling the model in the location of the flight, this appears as missing portions of the plume (Supplementary Fig. 59a) and only capturing the plume edge (Supplementary Fig. 59d). This apparent shift is a result of the resolution differences between the WRF-Chem model (4 km) and the QFED emissions inventory (0.1-degree, ~12 km). As each QFED sample is regridded to a single WRF-Chem grid cell, a shift in fire location can result. HRRR-Smoke is not subject to this issue because it is run at 3 km resolution and ingests 375 m VIIRS FRP to calculate and spatially allocate fire emissions, area, and plume rise<sup>39</sup>.

Rather than exclude these cases and lose half our sample size in the analysis, we implemented spatial shifts in the sampling algorithm to assure the model was sampled over the densest part of the plume. The shift distance is defined as the distance between the flight’s closest approach to the modeled fire location and the modeled fire location itself, which is defined the center of mass of the emissions in a 9 × 9 set of grid boxes around the observed fire location (FIREX-AQ data repository). The shift distance is then used to translate the flight track in space, and the model is sampled at these new locations. This assumes that the shifts are small enough that model meteorology and plume rise drivers do not change over the relatively small spatial range of the shifts (shifts are often a fraction of a degree). This substantially increased the number of comparisons between the model and the observations. However, it is acknowledged that shifting may result in higher error when the fuel type changes over the distance shifted, as this impacts the plume rise calculation<sup>14</sup>. Supplementary Fig. 59c, d (red line) shows how shifting allows us to sample the model in the location where plume rise occurs and how this impacts model performance.

In two cases, poor model-observation agreement was found for a different reason. In Supplementary Fig. 60b, the fire plume is not evident in the WRF-Chem



vertical profile of PM<sub>2.5</sub>, and Supplementary Fig. 60a shows that this is due to QFED emissions not being present in the model at the sampled time. As noted above, QFED daily emissions are assigned hourly values based on UTC time and each day (at 01 UTC) a new set of emissions is imported to QFED which tends to be around 5–6 pm local time for the FIREX-AQ domain. This leads to discontinuities in the emissions diurnal cycle which typically coincided with the DC-8 observing the fire plume. Therefore, a time shift was also implemented for the two cases where this occurred. In these cases, the model was sampled at a timestep one hour later than the time the DC-8 overpassed the plume. Similar to the spatial shift, the time shift assumes no change in meteorology (i.e., wind direction) from one timestep to the next. Although in reality meteorology can change drastically over the course of an hour, the timing of the overpasses (within 20 or 30 min of the next hour) we are able to neglect some of this variability. Supplementary Fig. 60c–d shows how sampling the WRF-Chem model one hour later brings us to a time where model emissions are present, and plume rise behavior can be evaluated.

**Classifying injection by visual inspection.** We define injection as a stratified layer of smoke above the PBL or residual layer above the fire whose stratified shape is maintained for at least five minutes (50–70 km based on aircraft groundspeed) of downwind transport. This definition is motivated by the differences between injected and non-injected smoke<sup>13</sup> and the tendency of injected smoke to aggregate in free tropospheric stable layers<sup>8</sup>. With this definition, injections are independent of the presence of PBL smoke. Examples of injection are shown in Fig. 3a–c, e, and examples of non-injection are shown in Fig. 3d, f.

In the DIAL-HSRL curtains (Fig. 3a, d), backscatter values of 1–56 Mm<sup>-1</sup> sr<sup>-1</sup> (yellow-dark red, black filled circles) represent smoke, and values of 0.18–0.56 Mm<sup>-1</sup> sr<sup>-1</sup> (dark green-light green, black open circles) represent the out-of-plume PBL. Since the presence of fire influences the mixed layer height observed by the DIAL-HSRL, upwind and downwind locations where the PBLH remained relatively stable were chosen as the out of plume region (black open circles in 3a–3f). In both models (Fig. 3b–c, e–f), PM<sub>2.5</sub> values of 5.6–1000 µg m<sup>-3</sup> (light blue-dark red) are assumed to be smoke. Determining the modeled out-of-plume PBL height is more complex due to the fact that WRF-Chem simulates background aerosols and HRRR-Smoke does not. Therefore, across all model cases, we use the out-of-plume modeled PBL height or the potential temperature-derived mixed layer height at 4PM local time (light gray solid or dark gray dashed lines in Fig. 3b–c, e–f in the location of black open circles) to classify injection. This methodology evaluates modeled injection with respect to the observed and modeled PBL and observed injection with respect to the observed PBL. This is different from previous studies that use modeled PBL heights only as the standard to judge injection<sup>11</sup>. Comparing ambient backscatter with PM<sub>2.5</sub> based on dry extinction is sufficient for our purposes because smoky regions and non-smoky regions are visually distinct from each other and have a satisfactory degree of spatial coincidence in both the backscatter and the PM<sub>2.5</sub> curtains.

Figure 3g shows the decision tree-style process that was applied to all cases to determine whether or not injection occurred. The first split in the tree (node A) asks what fraction of the topmost layer of smoke is above the out-of-plume PBLH. Node A identifies cases where smoke remains well-mixed throughout the local PBL (Fig. 3d) and is especially useful for evaluating the models (i.e., Fig. 3f) as the Freitas model injects flaming emissions at the height where the simulated plume stops rising but for a range of heat fluxes<sup>14</sup>, allowing the possibility for plumes to be injected into multiple vertical levels that could be below the PBLH.

For cases where a majority of the lofted plume resides in the free troposphere, node B1 indicates that downwind behavior will play a role in whether or not plumes will be classified as injections. If plumes maintain their stratified shape (i.e., in all panels of Fig. 3a), then they are considered injections. This comes with the caveat (node F1) that smoke in the residual layer, although it may appear stratified in shape, is not considered to be an injection in this work. These cases occur when fires were observed after dark, so the daytime convective PBL had been replaced with a less turbulent residual layer, leading to the stratified structure that likely will have more localized air quality impacts the day after if similar PBLH is reached. Cases where the plume mixes with the PBL within 5 min flight time are then considered to be non-injections (nodes C2, E2). If mixing occurs after 5 min of flight time, the case is considered an injection (node E1).

**Bootstrapping.** Some fires were observed multiple times in a single day, which can bias the data towards those fires. To mitigate such a bias, a bootstrapping approach was used that generates box and whisker plots by subsampling the dataset such that only one aircraft overpass per fire was selected per day. Final boxplots (Fig. 1 and Supplementary Figs. 50–52) were generated using the average statistics of all box plots from 2000 such iterations. The false non-injection category did not contain enough cases to include meaningful statistics for either model. HRRR-Smoke did not have enough true non-injection cases to generate statistics, but the true and false injection distributions are similar to the corresponding WRF-Chem distributions (i.e., Fig. 1).

**FRE flux computation.** FRE flux was chosen as the standard quantity with which to compare heat fluxes across the models and observations, and it is derived from the measurements as follows. First, the FRP is summed over all the MASTER pixels

present in a single scene, yielding the total FRP. Next, the sizes of the MASTER pixels are derived from the altitude of the aircraft above the terrain using a linear relationship (MASTER flight planning). FRE flux is then the slope of the linear fit  $FRP = m \cdot AREA$ , which in the main text is calculated on the entire dataset as well as subsets. This analysis can be done using smoldering pixels, flaming and saturated pixels, or all pixels.

HRRR-Smoke reports heat fluxes as area-to-FRP ratios, so we take their inverse to translate them into FRE fluxes. WRF-Chem uses the default convective heat fluxes assumed by the Freitas model, which are assumed to be 0.55 of the total heat fluxes<sup>14,36</sup>. The resulting convective heat flux is turned into a FRE flux using measured radiant and convective fractions<sup>40</sup>. These radiant and convective fractions were measured at ~3.9 µm, which is similar to wavelengths used by MASTER to measure FRP. The conversion factor is as follows: radiant fraction/convective fraction = 0.12/0.51 = 0.24<sup>40</sup>. Radiant and convective fraction measurements contain considerable uncertainty, and two instruments were used to measure the radiant fraction<sup>40</sup>. In order to produce the WRF-Chem ranges in Fig. 2, error propagation calculations were performed and the range including the errors from both radiant fraction measurements are plotted. We can also use error propagation to get a range of the radiant fraction/convective fraction ratio, which yields 0.239 ± 0.076 and 0.234 ± 0.07 when the different instruments are used to estimate the radiant fraction<sup>40</sup>. The inverse of these calculations can be done to obtain total heat fluxes from measured FRE fluxes. This produces analogous results to the ones described above (“Results”, Assessing heat flux assumptions).

## Data availability

Flight observational data from FIREX-AQ, along with the DIAL-HSRL data and campaign fuels data are archived by NASA/LARC/SD/ASDC<sup>41</sup> (<https://www-air.larc.nasa.gov/missions/firex-aq/>). MASTER Fire Radiative Power data products may be found here ([https://asapdata.arc.nasa.gov/share/Ichoku\\_RFP/](https://asapdata.arc.nasa.gov/share/Ichoku_RFP/)). WRF-Chem and HRRR-Smoke curtain plot data may be found here<sup>42</sup> (<https://doi.org/10.5281/zenodo.7033257>).

## Code availability

Custom code for this manuscript was written to generate all figures presented and to calculate FRE fluxes. Code was written in MATLAB R2017a and R2019b. Code available at the following <https://doi.org/10.5281/zenodo.7033257>.

Received: 19 August 2021; Accepted: 23 September 2022;

Published online: 10 October 2022

## References

- Ward, D. S. et al. The changing radiative forcing of fires: global model estimates for past, present and future. *Atmos. Chem. Phys.* **12**, 10857–10886 (2012).
- Munoz-Alpizar, R. et al. Multi-year (2013–2016) PM<sub>2.5</sub> wildfire pollution exposure over North America as determined from operational air quality forecasts. *Atmosphere* **8**, 179 (2017).
- Spracklen, D. V. et al. Wildfires drive interannual variability of organic carbon aerosol in the western U.S. in summer. *Geophys. Res. Lett.* **34**, L16816 (2007).
- Reid, C. E. et al. Critical review of health impacts of wildfire smoke exposure. *Environ. Health Perspect.* **124**, 1334–1343 (2016).
- Westerling, A. L. Warming and earlier spring increase western U.S. forest wildfire activity. *Science* **313**, 940–943 (2006).
- Zhuang, Y. et al. Quantifying the contributions of natural variability and anthropogenic forcings on increased fire weather risk over the western United States. *PNAS* **118**, e2111875118 (2021).
- Ford, B. et al. Future fire impacts on smoke concentrations, visibility, and health in the contiguous United States. *GeoHealth* **2**, 229–247 (2018).
- Kahn, R. A. et al. Aerosol source plume physical characteristics from space-based multiangle imaging. *J. Geophys. Res.* **112**, D11205 (2007).
- Val Martin, M. et al. Smoke injection heights from fires in North America: analysis of 5 years of satellite observations. *Atmos. Chem. Phys.* **10**, 1491–1510 (2010).
- Soja, A. J. et al. *Defining the Vertical Structure of Smoke Aerosols: Caliop-based Smoke Plume Detrainment Height* (American Geophysical Union, 2021).
- Kahn, R. A. et al. Wildfire smoke injection heights: two perspectives from space. *Geophys. Res. Lett.* **35**, L04809 (2008).
- Val Martin, M., Kahn, R. & Tosca, M. A global analysis of wildfire smoke injection heights derived from space-based multi-angle imaging. *Remote Sensing* **10**, 1609 (2018).
- Schum, S. K. et al. Molecular and physical characteristics of aerosol at a remote free troposphere site: implications for atmospheric aging. *Atmos. Chem. Phys.* **18**, 14017–14036 (2018).

14. Freitas, S. R. et al. Including the sub-grid scale plume rise of vegetation fires in low resolution atmospheric transport models. *Atmos. Chem. Phys.* **7**, 3385–3398 (2007).
15. Sessions, W. R., Fuelberg, H. E., Kahn, R. A. & Winker, D. M. An investigation of methods for injecting emissions from boreal wildfires using WRF-Chem during ARCTAS. *Atmos. Chem. Phys.* **11**, 5719–5744 (2011).
16. Roy, A. et al. in *Energy, Environment, and Sustainability* 245–278 (Springer, 2017).
17. Mallia, D., Kochanski, A., Urbanski, S. & Lin, J. Optimizing smoke and plume rise modeling approaches at local scales. *Atmosphere* **9**, 166 (2018).
18. Val Martin, M. et al. Space-based observational constraints for 1-D fire smoke plume-rise models. *J. Geophys. Res.* **117**, D22204 (2012).
19. Ye, X. et al. Evaluation and intercomparison of wildfire smoke forecasts from multiple modeling systems for the 2019 Williams Flats fire. *Atmos. Chem. Phys.* **21**, 14427–14469 (2021).
20. Freitas, S. R., Longo, K. M., Trentmann, J. & Latham, D. Technical Note: Sensitivity of 1-D smoke plume rise models to the inclusion of environmental wind drag. *Atmos. Chem. Phys.* **10**, 585–594 (2010).
21. Walter, C. et al. The importance of plume rise on the concentrations and atmospheric impacts of biomass burning aerosol. *Atmos. Chem. Phys.* **16**, 9201–9219 (2016).
22. Paugam, R., Wooster, M., Freitas, S. & Val Martin, M. A review of approaches to estimate wildfire plume injection height within large-scale atmospheric chemical transport models. *Atmos. Chem. Phys.* **16**, 907–925 (2016).
23. Hu, X. et al. Evaluation of three planetary boundary layer schemes in the WRF model. *J. Appl. Meteorol. Climatol.* **49**, 1831–1844 (2010).
24. Shin, H. H. et al. Impacts of the lowest model level height on performance of planetary boundary layer parameterizations. *J. Appl. Meteorol. Climatol.* **140**, 664–682 (2012).
25. Stocks, B. J. & Kauffman, J. B. in *Sediment Records of Biomass Burning and Global Change* 169–188 (Springer Berlin Heidelberg, 1997).
26. Sofiev, M., Ermakova, T. & Vankevich, R. Evaluation of the smoke-injection height from wild-land fires using remote-sensing data. *Atmos. Chem. Phys.* **12**, 1995–2006 (2012).
27. Hair, J. W. et al. Airborne High Spectral Resolution Lidar for profiling aerosol optical properties. *Appl. Opt.* **47**, 6734 (2008).
28. Hook, S. J., Myers, J. J., Thome, K. J., Fitzgerald, M. & Kahle, A. B. The MODIS/ASTER airborne simulator (MASTER)—a new instrument for earth science studies. *Remote Sens. Environ.* **76**, 93–102 (2001).
29. Strahler, A. et al. *MODIS Land Cover Product Algorithm Theoretical Basis Document (ATBD) Version 5.0 MODIS Land Cover and Land-Cover Change Principal Investigator* (1999).
30. Prichard, S. J. et al. *Fuel Characteristic Classification System version 3.0: technical documentation* (U.S. Department of Agriculture, Forest Service, 2013).
31. Giglio, L., Descloitres, J., Justice, C. O. & Kaufman, Y. J. An enhanced contextual fire detection algorithm for MODIS. *Remote Sens. Environ.* **87**, 273–282 (2003).
32. Wooster, M. Fire radiative energy for quantitative study of biomass burning: derivation from the BIRD experimental satellite and comparison to MODIS fire products. *Remote Sens. Environ.* **86**, 83–107 (2003).
33. Wooster, M. J., Roberts, G., Perry, G. L. W. & Kaufman, Y. J. Retrieval of biomass combustion rates and totals from fire radiative power observations: FRP derivation and calibration relationships between biomass consumption and fire radiative energy release. *J. Geophys. Res.* **110**, D24311 (2005).
34. Grell, G. A. & Freitas, S. R. A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.* **14**, 5233–5250 (2014).
35. Paugam, R. et al. Development and optimization of a wildfire plume rise model based on remote sensing data inputs—Part 2. <https://doi.org/10.5194/acpd-15-9815-2015> (2015).
36. McCarter, R. J. & Broido, A. Radiative and convective energy from wood crib fires. *Pyrodynamics* **2**, 65–85 (1965).
37. Western Regional Air Partnership. *2002 Fire Emission Inventory for the WRAP Region—Phase II Project No. 178–6* (Air Sciences, INC, 2005)
38. Darmenov, A. S., Da Silva, A. M. & Koster, R. D. *The Quick Fire Emissions Dataset (QFED): Documentation of Versions 2.1, 2.2 and 2.4* Vol. 38 (2015).
39. Ahmadov, R. et al. Using VIIRS fire radiative power data to simulate biomass burning emissions, plume rise and smoke transport in a real-time air quality modeling system. in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (IEEE, 2017).
40. Freeborn, P. H. et al. Relationships between energy release, fuel mass loss, and trace gas and aerosol emissions during laboratory biomass fires. *J. Geophys. Res.* **113**, D01301 (2008).
41. FIREX-AQ Science Team. Fire influence on regional to global environments and air quality. <https://doi.org/10.5067/SUBORBITAL/FIREXAQ2019/DATA001> (2019).
42. Thapa, L. H. Heat flux assumptions contribute to overestimation of wildfire smoke injection into the free troposphere for 2019 Western US wildfires. Zenodo. <https://doi.org/10.5281/zenodo.7033257> (2022).
43. Archer-Nicholls, S. et al. Characterising Brazilian biomass burning emissions using WRF-Chem with MOSAIC sectional aerosol. *Geosci. Model Dev.* **8**, 549–577 (2015).
44. Wiedinmyer, C. et al. The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning. *Geosci. Model Dev.* **4**, 625–641 (2011).
45. Iacono, M. J. et al. Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. *J. Geophys. Res.* **113**, D13103 (2008).
46. Chen, F. & Dudhia, J. Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.* **129**, 569–585 (2001).
47. Thompson, G. & Eidhammer, T. A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *J. Atmos. Sci.* **71**, 3636–3658 (2014).
48. Janjić, Z. I. Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. NCEP Office note 437, 61pp (2002).
49. Kleinman, L. I. et al. Rapid evolution of aerosol particles and their optical properties downwind of wildfires in the western US. *Atmos. Chem. Phys.* **20**, 13319–13341 (2020).
50. Zaveri, R. A., Easter, R. C., Fast, J. D. & Peters, L. K. Model for simulating aerosol interactions and chemistry (MOSAIC). *J. Geophys. Res.* **113**, D13204 (2008).
51. Thompson, G., Rasmussen, R. M. & Manning, K. Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.* **132**, 519–542 (2004).
52. Nakanishi, M. & Niino, H. An improved Mellor–Yamada level-3 model with condensation physics: its design and verification. *Boundary-Layer Meteorol.* **112**, 1–31 (2004).
53. Lyapustin, A. et al. Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *J. Geophys. Res.* **116**, D03211 (2011).
54. Lyapustin, A., Wang, Y., Korokin, S. & Huang, D. MODIS Collection 6 MAIAC algorithm. *Atmos. Meas. Tech.* **11**, 5741–5765 (2018).

## Acknowledgements

This work has been supported by the following grants: NASA 80NSSC18K0629, NASA 80NSSC20K1650, NASA 80NSSC18K0685, NOAA NA18OAR4310107, NSF 2013461. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center (GSFC). The FIREX-AQ data are archived by National Aeronautics and Space Administration, U.S. Government, NASA/LARC/SD/ASDC ([https://doi.org/10.5067/ASDC/FIREXAQ\\_Ground\\_InSitu\\_Data\\_1](https://doi.org/10.5067/ASDC/FIREXAQ_Ground_InSitu_Data_1)). <https://espo.nasa.gov/firex-aq>). R.A. and E.J. thank NOAA's JPSS PGR program for funding and the rest of the HRRR-Smoke team and collaborators for helping with the model development. C.I. acknowledges support from NOAA Educational Partnership Program for Minority Serving Institutions (NOAA/EPP/MSI) under agreement no. NA16SEC4810006 as well as Jeffrey Myers and the MASTER instrument team for the airborne instrument calibration and fire observation. S.K. thanks the help of Chuanyu Xu in processing MASTER data. Views expressed in this manuscript are those of the authors and do not reflect the views of NOAA or the Department of Commerce.

## Author contributions

Each author contributed as follows: L.H.T. designed the study, wrote MATLAB scripts to generate all figures, analyzed the data, and wrote the paper with contributions from all co-authors. X.Y. provided Matlab scripts to generate curtain plots of DIAL-HSRL, WRF-Chem, and HRRR-Smoke and to extract relevant variables from the model. J.W.H., M.A.F., and T.S. conducted DIAL-HSRL observations. S.K. processed the MASTER L1B data and contributed MASTER and satellite comparison, including data used to generate Fig. S55. C.I. and L.E. processed MASTER observations and generated the MASTER FRP L2 data files using their original algorithm. R.D. conducted and provided MASTER observations. A.J.S. and E.G. assisted with interpretation of results and provided the fuels information used to generate Fig. 2. R.A., E.J., G.A.G., and G.P. provided HRRR-Smoke forecast data and emissions. S.R.F. provided the 1D standalone plume rise code and aided in the interpretation of results. P.E.S. designed the study, provided MATLAB scripts to generate curtain plots, edited the manuscript, and advised L.H.T. during the project.

## Competing interests

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43247-022-00563-x>.

**Correspondence** and requests for materials should be addressed to Laura H. Thapa or Pablo E. Saide.

**Peer review information** *Communications Earth & Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Kerstin Schepanski and Clare Davis. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022