

### 1.General Comments:

Jin et al. clearly define the issue at hand and how they have contributed to the solution and what the resulting benefits are to science and society; namely, that common biomass burning emission inventories often omit diurnal information and vertical injection heights of fires, and that by incorporating their ideas, they have an emissions product for Southeast and East Asia (SEAF) that includes these two important pieces of information that ultimately improves the accuracy of models used for reporting and assessment of air quality, climate and public policy. They have done careful work to compare their results with measurements from TROPOMI, MISR and CALIPSO, and they present a comparison of their emissions dataset with common established BB emission datasets.

Of particular note is that they have been able to generate SEAF in such a way as to closely match not only the 2D structure of emissions observed from satellite, but the 3D structure as well, with the help of machine learning, albeit difficult for them to capture short-lived fires and emissions from low-lofted plumes. This is an encouraging contribution to the biomass burning emissions research community, and furthers the community's desire to see uncertainty in regional and global emissions decrease significantly. The main concern of this reviewer involves their method for generating and applying the diurnal FRP cycle. The authors need to include more discussion and review of the community's efforts regarding this very needed step in emission inventory development, and to show how their efforts are either similar to these previously published methods or are an improvement upon them.

Overall, I recommend this article for publication with minor revisions.

Response: We sincerely thank the reviewer for the comprehensive summary and the encouraging assessment of our work. We fully agree that the manuscript should better position our diurnal FRP-cycle method within the broader community efforts and more explicitly state what is inherited from prior approaches and what is new in our framework. We have revised the manuscript accordingly. **Please note that text in “italics and underlining” represents the revised sentences in the modified manuscript.**

**1. Enhanced review of community efforts on diurnal FRP modeling:** In the revised Introduction, we expanded the background discussion to provide a more complete and more balanced overview of the development of diurnal FRP modeling (Introduction, Lines 112–122: *A critical prerequisite for reducing uncertainty in emission estimates is the accurate reconstruction of the FRP diurnal cycle. Vermote et al., (2009) and Ellicott et al., (2009) established the foundational theoretical basis for using Gaussian functions to fit diurnal fire cycles from geostationary data. Subsequently, Kaiser et al., (2012) incorporated Kalman filtering in the Global Fire Assimilation System to address cloud gaps, while Andela et al., (2015) demonstrated the strong dependence of diurnal cycles on ecosystem types. More recently, Li et al., (2019, 2022) reconstructed sub-daily FRP variability by combining polar-orbiting and geostationary observations, with temporal gaps filled by integrating both available observations and ecosystem-specific diurnal climatologies, whereas Zheng et al., (2021) utilized Himawari-8 observations*

to implement an event-based Gaussian representation of the FRP diurnal cycle, establishing a critical regional reference for geostationary-based fire monitoring in East Asia.).

**2. Clarification of similarities and improvements:** In response to your request to show how our efforts are either similar to, or an improvement upon, previously published methods, we now clearly distinguish between methodological heritage and innovation.

**(1) Similarity (heritage):** As clarified in the revised Introduction and Methodology, our approach builds upon the community-established Gaussian representation of diurnal FRP cycles, thereby maintaining physical consistency with prior work. Historical statistics are used to determine the shape parameters of the diurnal curve, consistent with earlier Gaussian-based frameworks (Eq. (8)).

$$FRP(t) = ae^{-\frac{(t-b)^2}{2c^2}} + d \quad (8)$$

Where  $a, b, c$  are shape parameters controlling the amplitude, peak time, and width of the diurnal FRP cycle, respectively, derived from historical statistics for specific land cover types and regions.

**(2) Improvement (innovation):** We explicitly articulate the limitations of commonly used static climatological diurnal profiles, which are often applied when observations are missing and may introduce biases during non-active or extreme fire periods (Introduction, Lines 122–131: Despite these advances, when observations are partially missing due to cloud occlusion or limited temporal sampling, many FRP-based emission frameworks still rely on static diurnal representations to fill gaps, such as superimposing predefined Gaussian-shaped curves or adopting climatological diurnal profiles that are invariant in time (Wooster et al., 2021). Such static treatments can lead to biases in emissions during non-active periods or dampen peak fire activity during extreme events, thereby introducing additional uncertainty into emission estimates. These limitations indicate that, although Gaussian-based representations of diurnal FRP cycles are widely adopted and physically grounded, their application in regions with frequent cloud cover and episodic extreme fires requires dynamic, event-specific adjustment rather than reliance on climatological averages.). **To overcome these limitations, we introduce a two-stage adaptive fitting strategy (Eq. (8)).** Where  $a, b, c$  are shape parameters controlling the amplitude, peak time, and width of the diurnal FRP cycle, respectively, derived from historical statistics for specific land cover types and regions. The parameter  $d$  is an event-specific additive adjustment that shifts the diurnal curve vertically to match the observed FRP level of a given day and can either increase or decrease the overall magnitude relative to the climatological baseline. Distinct from traditional static climatological Gaussian schemes,

*we implement a two-stage adaptive fitting strategy. When sufficient instantaneous observations are available for a fire event ( $\geq 4$  detections) and the reconstruction is stable ( $R^2 \geq 0.8$ , indicating that the diurnal shape is sufficiently constrained by observations), all parameters ( $a, b, c, d$ ) are jointly optimized to capture event-specific diurnal variability. However, when observations are sparse ( $< 4$ ) or the reconstruction is unstable ( $R^2 < 0.8$ ), the shape parameters ( $a, b, c$ ) are fixed to their climatological values and only  $d$  is solved from the available observations. This constrained formulation preserves a physically realistic diurnal shape while anchoring the overall magnitude to satellite observations, thereby reducing systematic mean-level offsets and, in a least-squares sense, providing a closer fit to the available observations than the pure climatological Gaussian baseline. **Importantly, gap filling is not applied to all missing hours. Instead, the reconstruction is constrained to the high probability sustained burning windows determined in Table 1 for each region and vegetation type, and missing FRP is filled only within these statistically supported windows using Eq. (8) and the two-stage adaptive fitting strategy. This constraint prevents spurious reconstruction across low probability hours when sustained burning is unlikely, thereby preserving the physical realism of the reconstructed diurnal evolution.***

**(3) Evidence of improvement:** The effectiveness of the proposed dynamic adjustment is demonstrated by its dual capability to mitigate systematic underestimation while preventing artificial emission inflation.

- At the event scale, the method selectively recovers missing fire signals during critical windows. As stated in the revised manuscript: (Section 3.2.2: *Across all regions and vegetation types, reconstructed FRP profiles showed marked enhancements, particularly during key burning periods underrepresented in the original observations... In Region 1, cropland fires exhibited pronounced morning peaks (08:00–12:00 local time), with FRP increased by 67.7% after reconstruction... In Region 4, shrubland FRP increased by 88.7% during 16:00–00:00 local time, revealing active evening-to-night burning that was systematically underestimated due to twilight detection gaps.*). These localized adjustments ensure that missing fire activity is recovered based on physical evidence and agricultural practices.
- To ensure robustness when data are sparse, we implement a constrained logic: (Section 3.2.2: *In contrast, under observation-limited or unstable fitting conditions, the method employs a physically-constrained regularization strategy. By intentionally anchoring the diurnal shape to the established climatology and adjusting only the magnitude parameter, the framework effectively prevents the*

introduction of artificial peak shifts or structural artifacts...). This ensures we do not "over-fit" the data when observations are insufficient.

- At the regional scale, Fig. 7 demonstrates that the improvement is primarily a correction of systematic bias. We have added the following explanation: (Section 3.2.2: The original observations systematically underestimate FRP due to cloud contamination, low temporal sampling frequency, and twilight detection blind zones... It should be noted that the relatively small changes in mean FRP compared to the standard deviation reflect systematic bias adjustment rather than random variability... The dynamically fitted FRP showed improved regional means (e.g., 23.85 MW in Region 2 and 17.49 MW in Region 4), representing relative increases of 3.52% and 1.62% compared to the original data.).
- Overall, the dynamic adjustment framework improves the temporal representativeness of FRP by recovering underdetected fire activity during under-sampled periods, while avoiding artificial amplification associated with conventional climatological Gaussian fitting (Section 3.2.2: Across all regions and vegetation types, reconstructed FRP profiles showed marked enhancements, particularly during key burning periods underrepresented in the original observations; Overall, the dynamically adjusted FRP profiles exhibited markedly improved temporal continuity compared with the original observations, capturing the primary peaks of fire activity and recovering missing signals during under-sampled periods; and the Gaussian fitting method ... resulting in systematic overestimation in certain regions, primarily due to artificial amplification during inactive periods). As a result, the dynamically adjusted FRP yields more realistic regional mean values and improved statistical representativeness at the regional scale (Section 3.2.2: The dynamically fitted FRP showed improved regional means ... representing relative increases of 3.52% and 1.62% compared to the original data; and the relatively small changes in mean FRP compared to the standard deviation reflect systematic bias adjustment rather than random variability).

#### Specific Comments:

2.Lines 74-76: I believe this statement is a bit misleading. The authors are claiming here that the “integrated inventories” (i.e. the multi-source emissions datasets that incorporate diurnal cycles) generally have a large uncertainty when compared to the standard emission inventories. There may be two issues: 1) the authors are making the case that the standard inventories have uncertainty themselves and thus the need for integrated inventories, so although comparing against these standard inventories is important, we need to know if the areas where these standard inventories are weak and if the integrated inventories improve upon

them, and 2) Xu et al. 2023b is given as a paper that has assessed these integrated inventories and found large uncertainty (the numbers given in this sentence), but it seems as if the paper is only assessing their CHOBE inventory, which is not mentioned in their list of examples previously, and moreover, it does not assess any of the more well-known integrated inventories that the authors here present to us. It would be good to better present the issues with the integrated inventories that already exist, and to show how they improve upon the standard inventories but still underperform in areas that a new inventory can accomplish.

Response: Thank you for this careful and insightful comment. We fully agree that the statement in the original version of the manuscript regarding the uncertainty of “integrated inventories” was potentially misleading and did not clearly distinguish between the limitations of standard inventories and the advances achieved by multi-source integrated approaches. In response, we have substantially revised and restructured the Introduction to address this issue more clearly and transparently.

(1) We removed the generalized claim that integrated inventories exhibit “large uncertainty” and revised the associated citation strategy, avoiding reliance on Xu et al. (2023b), which evaluates a specific inventory rather than the broader class of integrated products. This revision ensures that the discussion of previous work remains contextually accurate and avoids inappropriate generalization.

(2) The revised Introduction now explicitly acknowledges the well-established improvements provided by existing integrated inventories. As stated in the revised text, recent studies combining geostationary and polar-orbiting FRP observations (e.g., Li et al., 2019,2022; Xu et al., 2023) have successfully produced hourly, kilometer-scale emission products that significantly improve spatiotemporal characterization compared with conventional daily or monthly inventories. This revision makes clear that integrated inventories represent a substantial methodological advancement over standard products.

(3) Rather than attributing remaining discrepancies to vague “uncertainty,” the revised text now identifies a specific and common methodological limitation shared by many existing frameworks. We explain that when observations are partially missing due to cloud occlusion or limited temporal sampling, many FRP-based emission frameworks still rely on static or climatological diurnal representations for gap filling. Such static treatments can introduce systematic biases by inflating emissions during inactive periods or damping peak fire activity

during extreme events, particularly in regions with frequent cloud cover and episodic fires such as Southeast Asia.

(4) We have emphasized that while previous integrated inventories improved surface (2D) temporal characterization, they often lack the explicit, observation-driven vertical injection heights required for comprehensive 3D atmospheric modeling. This justifies the necessity of the new SEAF inventory, which provides an observation-driven, hourly 3D BB emission dataset for the broader SEA region by integrating a dynamic diurnal adjustment with a machine-learning-based vertical injection model.

By restructuring the Introduction in this way, we clarify that the motivation for SEAF is not a general deficiency of integrated inventories, but rather a remaining methodological gap in dynamic, event-specific reconstruction of diurnal FRP cycles and their consistent extension to three-dimensional emission representations.

3.Lines 76-79: It is mentioned that using a Gaussian scheme to reconstruct the diurnal cycle is a problem, but the authors use a Gaussian scheme themselves, so it should be specified better how the authors use the Gaussian scheme differently than the other integrated inventories. These works should be mentioned and discussed: Ellicott et al. 2009, Vermote et al. 2009, Andela et al. 2015, Kaiser et al. 2012, Li et al. 2019/2022, and Zheng 2021. Ellicott and Kaiser are not mentioned in the paper at all. Kaiser in particular provides an alternative method for cloud correction using a Kalman filter temporal prediction.

Response: We sincerely thank the reviewer for highlighting these foundational studies and for pointing out that our original manuscript did not sufficiently clarify how our Gaussian-based approach differs from previous integrated inventories. We agree that this distinction needed to be made explicit, and we have revised both the Introduction and Methodology accordingly.

**(1) Expanded and updated literature review of Gaussian-based approaches:** In the revised Introduction, we have substantially expanded the discussion of prior work to trace the development of diurnal FRP modeling within the community (see our response to General Comments (Item 1, above): **Enhanced review of community efforts on diurnal FRP modeling**)).

**(2) Clarification of how our Gaussian implementation differs from previous approaches:**

As noted by the reviewer, our method also adopts a Gaussian representation of the diurnal FRP

cycle. However, the key distinction lies not in the functional form itself, but in how it is applied under observation-limited conditions. To avoid repetition, we refer the reader to our response to the previous Specific Comment (see our response to General Comments (Item 2, above): **Clarification of similarities and improvements**), where we detail the adaptive two-stage fitting strategy implemented in SEAF.

4.Lines 208-209: Using a beta value of 0.95 seems really high to me. Does this mean that if you have only 5% non-cloud, you still estimate the undetected FRP for the 95% cloud-filled portion of a grid cell from the 5%? Have you tried a lower threshold to see how the results change, or not?

Response: We thank the reviewer for raising this important concern regarding the choice of the cloud-fraction threshold. When the cloud fraction  $\beta$  approaches 0.95, the cloud-correction scheme estimates grid-cell-aggregated FRP using the available clear-sky observations; however, this procedure is not a simple linear extrapolation from a small non-cloud fraction, nor does it assume spatial homogeneity within the grid cell.

#### **(1) Mathematical constraint by damping term**

The formulation in Eq. (2) explicitly includes a quadratic damping term ( $\alpha = 0.25$ ) to constrain amplification under high cloud-cover conditions. This term suppresses the rapid growth that would otherwise occur under a simple reciprocal correction. For example, at  $\beta = 0.95$ , the effective correction factor is limited to approximately 3–4 times the observed clear-sky FRP, rather than the ~20-fold amplification implied by a purely linear extrapolation ( $1/(1-\beta)$ ). This design ensures that the cloud correction remains physically bounded even under heavily clouded conditions.

#### **(2) Sensitivity analysis across cloud-fraction thresholds**

To evaluate whether a high cloud-fraction threshold introduces instability or systematic overestimation, we conducted a sensitivity analysis in which  $\beta$  was varied from 0.70 to 0.95 for both NOAA-20 and Suomi-NPP VIIRS observations. The results show that total annual FRP estimates are highly stable across this range, with variations remaining below 0.5% (Fig. S2). This indicates that the cloud-correction scheme does not introduce numerical artifacts or bias regional-scale FRP budgets, even when a high threshold is applied.

#### **(3) Rationale for retaining $\beta = 0.95$**

Based on the combined mathematical constraint and the demonstrated empirical stability, a threshold of  $\beta = 0.95$  was retained to avoid systematic underestimation of FRP in persistently cloudy regions of Southeast Asia, while maintaining numerical stability. This choice represents a balance between recovering cloud-obscured fire signals and preventing excessive amplification under extreme cloud-cover conditions.

A summary of this sensitivity analysis has been added to Section 2.3.1 of the revised manuscript (Lines 314–316), and the full results are provided in the Supplementary Material (Fig. S2).

5.Eqs. 3 & 4: If you combine the equations, the 1 cancels out and you are simply left with multiplying the AHI FRP by the ratio of VIIRS to AHI FRP for the coincident data. Perhaps it would be simpler to define  $r$  as simply the fraction, unless it is beneficial to center the ratios around zero.

Response: We thank the reviewer for this sharp observation regarding the algebraic formulation. We fully agree that for a single instantaneous calibration, substituting Eq. (4) into Eq. (3) is algebraically equivalent to simply scaling AHI FRP by the VIIRS-to-AHI ratio. However, as the reviewer correctly surmised, we deliberately defined  $r_i$  in the form of a zero-centered relative bias  $(1 + r_i)$  rather than a simple ratio, because this formulation offers significant advantages for statistical aggregation and numerical stability when calibration factors must be averaged across time and land-cover classes. We have revised Section 2.3.2 to explicitly clarify the statistical rationale behind this choice.

**(1) Statistical Robustness in Aggregation:** Unlike a simple instantaneous correction, our method requires aggregating these factors to generate daily ( $r_{d,j}$ , Eq. 5) and monthly calibration factors when simultaneous observations are unavailable. Averaging zero-centered relative bias terms is generally more robust and statistically consistent than averaging direct ratios, which can be asymmetric and sensitive to outliers.

**(2) Interpretability:** Defining  $r_i$  as a relative bias allows  $r_i = 0$  to serve as a direct, intuitive baseline for perfect consistency between sensors. Positive and negative values immediately indicate the direction of the bias (underestimation or overestimation).

**(3) Unified Calibration Framework:** This formulation provides a unified mathematical structure that connects the instantaneous bias (Eq. 4) with the dynamic daily aggregation (Eq. 5) and monthly land-cover-specific adjustments (Eq. 6). This hierarchy ensures that the



calibration logic remains consistent whether we are using a direct match or a temporal/spatial fallback, maintaining the integrity of the emission inventory across all scales.

- 6.Eq. 6: I would be very interested to know if using  $r_i$  instead of  $r_{m,l}$  makes much of a difference in general. My idea would have been to prioritize the general calibration ( $r_{m,l}$ ) over of instantaneous calibrations ( $r_i$ ) in developing the diurnal cycle, but here you prioritize the instantaneous calibrations.

Response: We thank the reviewer for raising this important question regarding the calibration hierarchy in Eq. (6). Below we clarify the distinct roles of the instantaneous calibration factor  $r_i$  and the monthly land-cover-specific factor  $r_{m,l}$  and explain why instantaneous calibration is prioritized when available.

**(1) Separation between magnitude calibration and diurnal shape reconstruction:** Both  $r_i$  and  $r_{m,l}$  are applied solely to correct the magnitude (amplitude) of the AHI FRP to match the polar-orbiting reference. They do not determine the shape (e.g., peak time and width) of the diurnal cycle. As described in the Methodology, the shape parameters are derived from long-term AHI statistics (2016–2023) using Gaussian fitting, ensuring the structural stability of the diurnal curve regardless of which calibration factor is used.

**(2) Rationale for Prioritizing Instantaneous Calibration ( $r_i$ ):** We prioritize the instantaneous calibration factor ( $r_i$ ) whenever a reliable simultaneous VIIRS-AHI pair ( $\pm 5$  min) is available. The primary reason is event-specific fidelity.

- Capturing Real Intensity: Using  $r_i$  ensures that for a specific fire event, the AHI FRP magnitude is directly constrained by the high-quality VIIRS observation at that moment.
- Avoiding Over-smoothing: Prioritizing the general monthly average ( $r_{m,l}$ ) would artificially smooth out the variations of individual fire events. For example, during extreme fire episodes where intensities significantly exceed the climatological mean, relying on  $r_{m,l}$  would systematically underestimate the peak energy.

**(3) The Adaptive Hierarchy:** Our approach functions as an adaptive hierarchy: instantaneous calibration ( $r_i$ ) is used to anchor event-specific magnitudes when available, while statistically robust daily or monthly averages ( $r_{m,l}$ ) are used as a fallback under data-sparse conditions. This design maximizes physical fidelity at the event scale without sacrificing stability at longer temporal scales. While differences between using  $r_i$  and  $r_{m,l}$  are most evident for individual

high-intensity events, their influence is largely smoothed at monthly or regional scales, supporting the use of this hierarchical strategy.

We have clarified this physical and statistical distinction in the revised Methodology to avoid any ambiguity regarding the impact of  $r_i$  on the diurnal cycle structure.

- 7.Eq. 7: Did you mean to say “VIIRS/AHI FRP = 0” instead of “... < 0”?

Response: Yes. We thank the reviewer for pointing out this ambiguity. FRP is a physical quantity and cannot be negative. In the revised manuscript, we have reformulated Eq. (7) to remove the use of negative values and to explicitly distinguish between valid observations and missing data.

$$FRP_i^{fuse} = \begin{cases} FRP_i^{VIIRS} & i \in \Omega_{VIIRS} \\ \overline{FRP_i^{AHI}} & i \notin \Omega_{VIIRS} \wedge i \in \Omega_{AHI} \\ 0 & otherwise \end{cases} \quad (7)$$

- 8.Line 263: I may not be understanding correctly, but if I do, I think the term “T\_gap” conveys the opposite of what is being described. “T\_gap” makes me think it is the period between high intensity burn periods, not how long the high intensity burn lasts.

Response: We thank the reviewer for this helpful clarification. We agree that the notation “T\_gap” was potentially misleading, as it could be interpreted as the interval between burning periods, whereas our analysis focuses on the duration of persistent burning.

To remove this ambiguity, we have revised the manuscript by renaming “T\_gap” to “T<sub>SB</sub>” (sustained burning), which explicitly denotes the maximum duration of sustained burning derived from burning frequency statistics. We also revised the accompanying text and Table 1 to clearly distinguish between the duration of sustained burning and the separation interval between adjacent burning windows. These changes clarify the physical meaning of the parameter and ensure consistent interpretation throughout the manuscript.

- 9.Lines 268-269: Please explain how these Gaussian curves were constructed, e.g. were the peaks adjusted before averaging between days to account for any daily differences?

Response: Thank you for this important clarification request. We agree that the construction of the Gaussian diurnal curves requires a clearer description, particularly regarding how daily variability is treated.

(1) As clarified in the revised manuscript, the climatological Gaussian diurnal curves are derived from long-term Himawari-8/9 FRP observations (2016–2023), grouped by region and vegetation type. For each local-time bin, FRP observations from the full 2016–2023 record are

aggregated to compute a multi-year mean FRP, yielding a 24-hour climatological mean diurnal profile. A Gaussian function is then fitted to this mean profile to estimate the characteristic diurnal shape (peak timing and width). No explicit peak shifting, temporal alignment, or normalization at the day level is applied prior to this fitting. The resulting curves are intended to represent statistically dominant diurnal fire behavior rather than day-specific variability (Lines 281-285: Long-term AHI observations (2016–2023) are used to derive climatological Gaussian representations of FRP diurnal cycles (constructed by computing multi-year mean FRP at each local-time bin to form a 24-hour mean diurnal profile).

(2) Treatment of daily variability: Daily differences in fire intensity or peak strength are not handled through pre-alignment or averaging of daily peaks. Instead, as detailed in our response to the General Comments (see our response to General Comments (Item 2, above): **Clarification of similarities and improvements**), daily variability is captured through a subsequent dynamic adjustment step. When sufficient same-day observations are available, an event-specific magnitude parameter is solved while the diurnal shape remains physically constrained by the climatological Gaussian form (Sections 3.2.2–3.2.3).

This clarification has been added to the revised manuscript to explicitly state that peak positions are not adjusted prior to averaging and that daily variability is accounted for through event-specific dynamic adjustment rather than temporal alignment of peaks.

- 10.Lines 340-342: citation needed

Response: Thank you for this comment. The relevant text (*Furthermore, comparisons of thermal anomaly observations from drones with both VIIRS and Himawari-8 data indicate that VIIRS measurements are more reliable.*) has been fully revised in the updated manuscript (Section 3.1). The previous qualitative statement has been removed and replaced with a physically grounded, quantitative discussion of cloud-induced FRP underestimation and cross-sensor differences, explicitly supported by established FRP-based emission and uncertainty studies (e.g., Kaiser et al., 2012; Andela et al., 2015b; Freeborn et al., 2014; Deng et al., 2024; Wickramasinghe et al., 2018; Hall et al., 2023). The revised section now includes region-specific analyses showing 2–9% increases in VIIRS FRP due to cloud correction and explains the physical mechanisms behind observed sensor differences. This ensures that all statements are now properly referenced and methodologically justified.

- 11.Lines 342-355: There doesn't seem to be any citations or analysis done to show how the increased estimates of FRP in cloud-filled regions compare to reality. Clearly, there must be an increase, but without any general idea of how much is missing due to cloud cover, it doesn't seem to decrease the uncertainty in the results.

Response: We thank the reviewer for this insightful comment. We agree that it is critical to demonstrate that the cloud-correction scheme provides a physically constrained improvement rather than an arbitrary increase in FRP estimates. Accordingly, we have revised Section 3.1 to clarify the robustness of our results from three complementary aspects.

(1) Physical basis and literature support: We explicitly document that persistent cloud cover leads to systematic underestimation of satellite-derived FRP, a limitation that has been widely recognized in FRP-based emission inventories and FRP field reconstruction studies (Kaiser et al., 2012; Andela et al., 2015b). The adopted coverage-based cloud-correction approach follows established practices in the literature and is therefore physically motivated rather than ad hoc.

(2) Conservative magnitude constrained by known uncertainties: We quantify that cloud correction leads to a modest increase of approximately 2%–9% in VIIRS-derived FRP across all regions. Importantly, this adjustment range is comparable to, or smaller than, inherent sensor-level FRP uncertainties (approximately 15–30%) and atmospheric attenuation effects reported in previous validation studies (Freeborn et al., 2014; Deng et al., 2024). This indicates that the correction is conservative and primarily reduces cloud-induced systematic low bias, rather than inflating FRP estimates beyond plausible physical bounds. The fact that the applied corrections fall well within the known sensor uncertainty envelope provides indirect validation of their physical reasonableness.

(3) Cross-sensor consistency as internal validation: Crucially, the framework does not uniformly increase FRP values. While cloud correction increases VIIRS FRP, the subsequent cross-sensor calibration results in an overall decrease in AHI FRP (approximately –3% to –24%) by mitigating sensor-specific overestimations related to mixed-pixel effects and false detections (Hall et al., 2023; Wickramasinghe et al., 2018). This bidirectional adjustment—with VIIRS increasing and AHI decreasing—demonstrates that the method targets systematic biases in both directions and improves the overall physical consistency of the resulting FRP fields. This bidirectional behavior provides an internal plausibility check on the correction magnitude and

demonstrates that the framework is not a uniform inflation scheme but a bias-correction system anchored to the cloud-corrected VIIRS reference.

These clarifications and the corresponding quantitative analyses have been incorporated into the revised Results section (Section 3.1). While direct ground-truth validation in persistently cloudy regions remains challenging, as acknowledged in the broader FRP validation literature, the combination of a physically grounded methodology, conservative adjustment magnitudes within known uncertainty bounds, and improved cross-sensor consistency indicates that the applied corrections are unlikely to amplify uncertainty and instead act to reduce systematic low bias in the FRP estimates.

12.Line 370 / Fig. S4: Please list the units for time in the caption – I can't seem to make the "9:00-16:00 local time" for NE China (region 4) cropland (column 1) correspond to the figure. The x-axis shows this to peak at ~6:00, which if in UTC would correspond to a local CST time of 14:00, which is not the center of 9:00-16:00. Also, it seems as if for Region 1 the Gaussian doesn't work too well because of the prolonged right tail – perhaps a skew term should be introduced.

Response: We thank the reviewer for pointing out the inconsistency between the reported peak times and the time axis in Fig. S5 (the pre-revision Fig. S4). We agree that this required clarification.

(1) Time reference and unit clarification: In the revised manuscript, we have clarified that the x-axis in Fig. S5 represents hours in UTC. Correspondingly, all diurnal peak times reported in Section 3.2.1 have been revised to UTC, with local time provided only as a reference where necessary. For example, cropland fires in Northeast China (Region 4) peak between approximately 01:00 and 08:00 UTC, corresponding to 09:00–16:00 local time (UTC+8) (Section 3.2.1, Lines 591–592). These revisions ensure consistency between the figure and the text and remove ambiguity regarding the time reference.

(2) We agree that this asymmetry indicates a departure from an ideal symmetric Gaussian shape and merits explicit discussion. In the revised Discussion, we now clarify that the pronounced right tail in Region 1 is likely driven by region-specific fire and observational characteristics rather than by a deficiency of the Gaussian framework itself (*The prolonged right-tailed diurnal patterns observed in Region 1, which are potentially driven by the high thermal inertia of*

*smoldering peatlands or persistent late-afternoon cloud interference...*). Such conditions are common in tropical peatland-dominated regions and can lead to sustained or delayed fire radiative signals. We further acknowledge that, under these conditions, a symmetric Gaussian assumption may lead to a temporal shift or partial underestimation of evening emissions (*...a symmetric Gaussian assumption may lead to a temporal shift or underestimation of evening emissions*). Similar deviations from unimodal symmetric diurnal behavior have been reported in previous studies for tropical fire regions characterized by frequent cloud cover and long-lasting combustion processes (Page et al., 2009; Rein, 2009).

While a skewed fitting function could potentially better represent such asymmetric structures, we retain the Gaussian formulation in this study as a deliberate trade-off between computational robustness and the primary objective of climatological diurnal fitting, namely, the stable characterization of dominant fire peak timing for gap filling (*Although this simplification represents a necessary trade-off between computational efficiency and capturing primary fire peaks...*). This methodological choice is therefore intentional rather than an oversight, and we explicitly acknowledge this limitation and its associated uncertainty in the Discussion.

- 13.Lines 378-380 / Fig. 5: Would you please briefly discuss what is going on in panels g and l. How are the filled data points so far off the GLS in g, and why was the peak not shifted right in l? Of all the panels, only a, m, p, q and t seem convincing and substantial; the rest of the panels have peaks that are difficult to corroborate with the presented data.

Response: Thank you for the reviewer's insightful observation regarding Fig. 5. We acknowledge that the previous presentation, which included unstable GLS fits with  $R^2 < 0.8$  for demonstration, could be confusing. In the revised manuscript, we regenerated Fig. 5 to strictly follow our two-stage gap filling strategy. Each panel now displays only the fitting mode that is actually used for that event.

In the revised Fig. 5, panels a, p, q, s, and t illustrate cases where the within-day sampling is sufficient to constrain the diurnal curve shape ( $R^2 \geq 0.8$ , and  $\geq 4$  detections), therefore the diurnal cycle is reconstructed using a least-squares GLS fitting. In contrast, the remaining panels, including g and l, represent cases with sparse or uneven observations where the diurnal shape parameters cannot be reliably constrained. In these cases, the framework switches to the GVM mode, in which the diurnal shape is fixed to region- and land-cover-specific

climatological parameters and only an event-specific additive term  $d$  is solved from the available detections.

This design directly explains the two reviewer questions. First, in panel g, the filled points are not expected to lie on the GLS curve because GLS is not applied in this information-limited case. Instead, the GVM reconstruction prioritizes physical plausibility and robustness by avoiding underdetermined shape optimization. Second, in panel l, the peak timing is not shifted to the right because the available observations do not provide sufficient constraints to justify a data-supported peak-time adjustment. Allowing a peak shift under such sampling would improve the visual fit locally but would introduce spurious timing changes that cannot be verified.

Although GVM retains the climatological diurnal shape, it is still mathematically and practically superior to using a pure climatological Gaussian curve without event-level adjustment. Specifically, the additive term  $d$  is solved in a least-squares sense to minimize residuals to the available observations, thereby correcting systematic mean-level offsets between the climatological baseline and the observations for that specific day. As a result, the GVM mode anchors the reconstructed magnitude to the observed FRP level while avoiding artificial peak shifts or peak broadening under sparse sampling conditions.

Overall, Fig. 5 is intended to demonstrate that the proposed two-stage strategy does not apply a single fitting approach to all cases. Instead, it uses GLS when the data support event-resolved fitting and switches to GVM when the sampling is insufficient, which is a conservative choice to prevent overfitting while maintaining physical consistency. This design choice is further supported by independent evaluations presented later in the manuscript, including improved peak timing statistics (Fig. 6), enhanced recovery of daily FRP totals (Fig. 7), and strong agreement between emission-derived CO and satellite observations.

- 14.Lines 419-440 / Fig. 7: The differences between the original and filled mean FRP values are so low compared with their standard deviations that it's hard to argue for the significance of these changes. Please address this. Also, do I understand correctly that the conventional Gaussian fits result in lower FRP for most of the regions? How did you define the conventional fit? If simply by not using GLS and GVM, then would you mind mentioning how e.g. Ellicott et al. 2009, Zheng et al. 2021, and Andela et al. 2015 all have dynamic diurnal Gaussian fits that change amplitude and/or duration with each day's data, and how you are similar or different in your approach.

Response: We thank the reviewer for this insightful comment. We address below the interpretation of the mean FRP differences in Fig. 7, the definition of the conventional Gaussian fit, and the relationship between our approach and previously published dynamic diurnal Gaussian methods.

(1) Interpretation of small mean differences relative to standard deviations: The relatively small changes in mean FRP compared with the large standard deviations shown in Fig. 7 do not indicate weak correction. Rather, they reflect the conservative nature of the reconstruction and the fact that the standard deviation primarily represents intrinsic spatiotemporal fire variability across heterogeneous fire regimes, rather than reconstruction uncertainty. In other words, the standard deviation shown in Fig. 7 is dominated by real fire intermittency and regional heterogeneity, not by the magnitude of the applied correction. Fig. 7 illustrates two opposing systematic effects: missing observations during sustained high-burning periods lead to systematic underestimation in the unfilled product, while unconditional application of diurnal Gaussian curves over a full 24-hour cycle tends to introduce artificial FRP during low- or non-burning periods. The modest net change in the mean therefore reflects bias redistribution, whereby underestimation during active burning windows is reduced while overestimation during inactive hours is suppressed. As a result, statistical significance should not be assessed solely by the ratio of mean change to standard deviation, because the correction targets systematic temporal bias rather than overall variance reduction.

(2) Definition and behavior of the conventional Gaussian fit: In this study, the “conventional Gaussian fit” refers to a static climatological diurnal curve derived from long-term averages and applied uniformly across the entire 24-hour cycle without conditioning on event duration, high-burning probability windows, or observation availability. This definition intentionally mirrors the most common implementation of climatological Gaussian gap filling used in many FRP-based emission frameworks. Under this definition, the conventional fit can yield either higher or lower regional mean FRP depending on the balance between recovered emissions during active periods and artificially assigned emissions during inactive periods. In several regions, the latter effect dominates, resulting in lower or even negative net corrections relative to the dynamically adjusted reconstruction shown in Fig. 7. This explains why the conventional



Gaussian approach produces lower FRP estimates for most regions in Fig. 7, despite partially filling missing observations.

(3) Relation to previous dynamic Gaussian approaches: We acknowledge that previous studies, including Ellicott et al. (2009), Andela et al. (2015), and Zheng et al. (2021), implemented dynamic diurnal Gaussian reconstructions that allow amplitude and/or duration to vary with daily observations. Our approach is conceptually aligned with these studies in recognizing the need for event-specific adjustment. However, it differs in that dynamic parameter optimization is explicitly conditioned on observational sufficiency and reconstruction stability. When event-day observations are insufficient or the fit is unstable, the diurnal shape is intentionally constrained to climatology and only the magnitude is adjusted. This strategy avoids unsupported peak shifts and prevents artificial inflation of FRP outside physically meaningful burning periods.

Overall, the small net differences in mean FRP shown in Fig. 7 should be interpreted as evidence of conservative, physically constrained bias correction rather than weak adjustment. The proposed method improves the realism of daily FRP distributions by redistributing bias while preserving the intrinsic variability of fire activity. This design choice prioritizes physical plausibility and stability over maximizing mean differences, consistent with the objectives of emission gap filling.

15.Lines 556-557: Can you remind the reader to reference Figure 12a for this claim about the comparisons to MISR?

Response: We thank the reviewer for this suggestion. We have revised Line 863 to explicitly reference Fig. 12 (a), which presents the MISR-based plume height comparison supporting this conclusion.

- 16.Fig. 4: Please explain how the cloud-corrected AHI FRP data can be less than the non-cloud-corrected data.

Response: Thank you for raising this important point. We clarify that the “correction” applied to AHI FRP in this study is not a pure cloud-recovery procedure and therefore does not necessarily lead to a monotonic increase in FRP. To avoid any further ambiguity, we have updated the caption of Fig. 4 to explicitly distinguish between the cloud-gap correction applied to VIIRS and the cross-sensor calibration applied to AHI.

In the revised manuscript, cloud correction is applied exclusively to VIIRS FRP to compensate for cloud-induced under-detection, whereas AHI FRP is subsequently adjusted **through cross-sensor calibration** using the cloud-corrected VIIRS product as a radiometrically more reliable reference. This calibration is designed to improve physical consistency in FRP magnitude between sensors rather than to inflate AHI-derived values. The results shown in Fig. 4 and Section 3.1 demonstrate that after cross-sensor calibration using cloud-corrected VIIRS FRP as a reference, calibrated AHI FRP exhibits region-dependent responses. In most regions, AHI FRP decreases due to suppression of sampling-related overestimation, whereas in regions with substantial BB emissions (e.g., northern Laos), AHI FRP still increases as cloud correction reveals additional fire activity previously underestimated. Therefore, calibrated AHI FRP may either decrease or increase depending on the regional balance between cloud-induced underestimation and sampling-related overestimation already documented in the manuscript.

- 17.Figs. 4, 7 & S7b: The resolution of the images seems to be too coarse. There appears to be a gridded/stripped pattern in the images – it appears to be an artifact in the data (particularly in Fig. 7), but it could just be the poor image resolution. Please update the resolution – I would like to be able to see more detail when I zoom in. Please also confirm what the stripped pattern is if it is indeed an artifact.

Response: Thank you for your careful review and the opportunity to clarify our visualization approach. We have carefully examined Figs. 4, 7, and S7b regarding the "grid/stripping" patterns. We would like to clarify that these patterns are not instrumental artifacts (such as VIIRS detector striping) or algorithmic errors. Instead, they represent the native resolution of our gridded product. The "blocky" appearance is a deliberate choice for the following reasons:

(1) Data Fidelity and Transparency: We used a pixel-wise rendering method (pcolormesh with no interpolation). This ensures that each pixel's color accurately represents the underlying aggregated FRP value. Spatial interpolation (smoothing) often creates a false sense of continuity, which can misrepresent the magnitude of fire intensity and artificially "smear" fire signals into non-fire areas.

(2) Product Characteristics: Since the data is aggregated into a regular  $0.03^{\circ} \times 0.03^{\circ}$  grid, the boundaries between grid cells are inherently discrete. The "striping" the reviewer observed is the alignment of these discrete grid cells at the boundaries of fire plumes.

To facilitate detailed inspection during the review process, we have additionally provided  $\geq 300$  dpi versions of Figs. 4, 7, and S7b in the response to reviewers, allowing the grid structure and spatial gradients to be examined clearly when zoomed in. The figures included in the manuscript follow the journal’s standard publication resolution, and no spatial smoothing or interpolation has been introduced in either case.

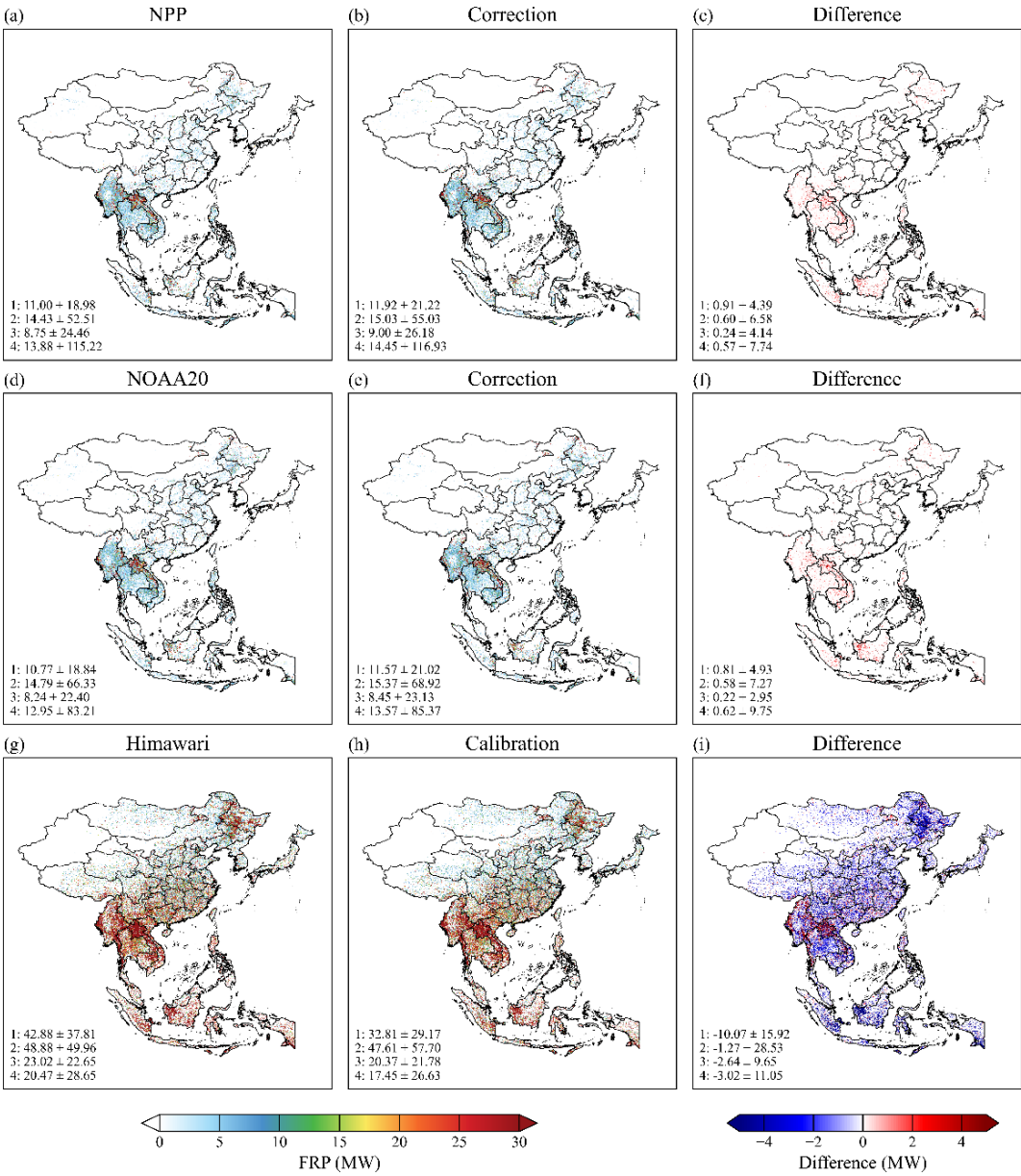
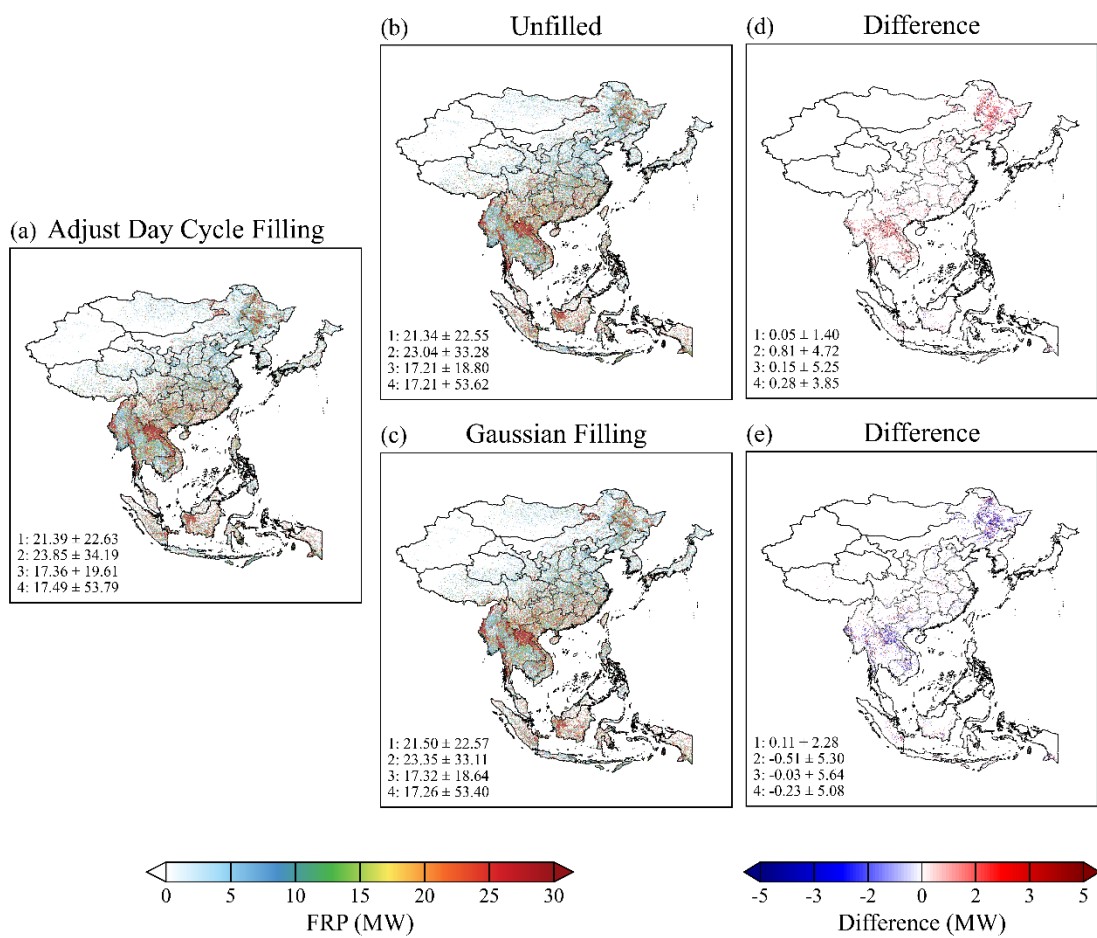
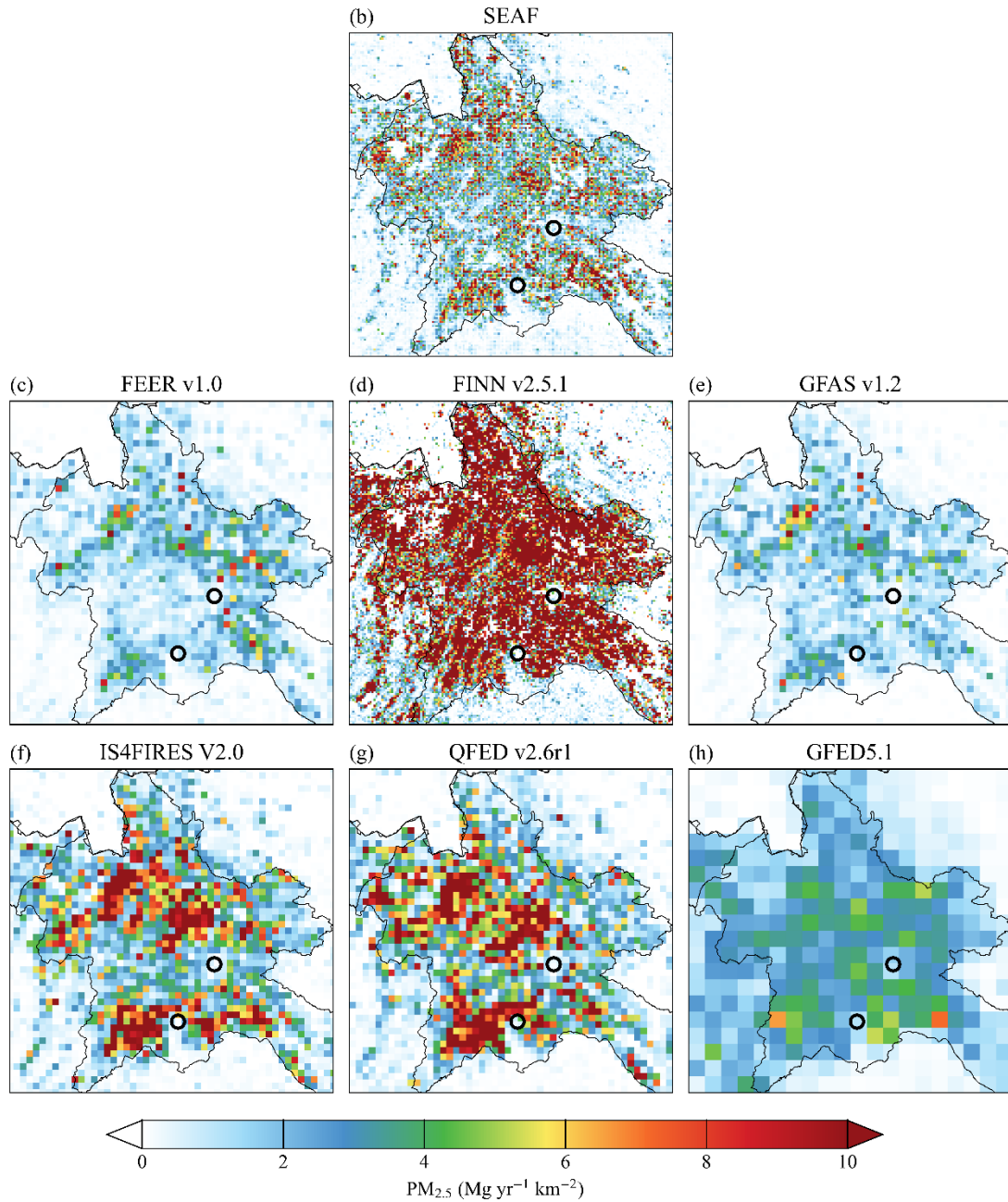


Figure 4



**Figure 7**



**Panel from Figure S8 (formerly Figure S7)**

- 18.Fig. 11: I think the images would be easier to interpret if you kept the SEAF panel as is and converted the rest of the panels to difference maps.

Response: We thank the reviewer for this helpful suggestion. Fig. 10 (formerly Fig. 11) has been revised accordingly. The figure caption and the corresponding text in the manuscript have been updated to reflect this change (Section 3.3.2).

- 19.Fig. 14: It appears to me that the SEAF emissions in panel d is greater than that of panel e, but in panel f, it is reported as lower. Is this a mistake, or is there a dynamic that is not visually observable with panels d and e?

Response: Thank you for raising this point. We apologize that the distinction between panels (a–e) and panel (f) was not sufficiently clear in the original manuscript. The apparent inconsistency does not indicate an error, but rather reflects that these panels represent different quantities.

Panels (a–e) show the layer-resolved  $\text{PM}_{2.5}$  emission mass after vertical allocation to five altitude layers (0.025 km, 0.275 km, 1.0 km, 2.75 km, and 5.5 km). In SEAF, smoke plume heights (SPH) are first predicted using our machine-learning model, and the predicted SPH is then used to guide the vertical allocation of emissions. Importantly, the allocation scheme distributes the emission mass across the layers that a plume is assumed to traverse, rather than assigning all mass exclusively to the highest layer. Consequently, the 2.75 km layer (panel d) can integrate contributions from both moderate injection cases and the lower portions of high-reaching plumes, whereas the 5.5 km layer (panel e) only contains the fraction assigned to the uppermost layer. This mechanism can lead to larger allocated emission mass at 2.75 km than at 5.5 km, even when high SPH occurrences are present. Panel (f), in contrast, compares the relative frequency distribution of SPH, that is, the frequency with which plume tops fall within each height interval. Therefore, the frequency at 5.5 km being higher than at 2.75 km does not imply larger emission mass allocated to the 5.5 km layer. As discussed in the revised text, the SEAF SPH distribution exhibits a relative dip near 2.75 km, which is consistent with a regime in which plumes either remain largely within the boundary layer or are lofted to higher altitudes under strong convection, resulting in fewer plume tops occurring at intermediate heights. Notably, SEAF retains a non-negligible occurrence at 5.5 km (approximately 0.12), which is broadly consistent with the MISR and CALIPSO distributions, while conventional inventories such as GFAS v1.2 tend to under-represent these high-altitude injection signals.

To prevent misinterpretation, we have revised panel (f) by changing the y-axis label from “Height” to “Smoke Plume Heights (SPH)”, and we have updated the Figure 14 caption to explicitly distinguish layer-allocated emission mass in panels (a–e) from the SPH frequency distribution in panel (f).

- 20.Fig. S7: The color scale is a bit unhelpful since the values saturate too quickly to be able to do any useful visual comparisons. Either stretch the scale, or as suggested with Fig. 11, convert the images to difference maps. Also, the units need to be changed to do proper comparisons

since the spatial resolutions are different between the panels. Please convert them to  $\text{Mg/yr/km}^2$ .

Response: We thank the reviewer for this helpful suggestion. Fig. S7 has been revised by rescaling the color bar to avoid saturation and by converting all emission fields to area-normalized units ( $\text{Mg yr}^{-1} \text{ km}^{-2}$ ), thereby enabling consistent spatial comparison across inventories with different native resolutions.

#### Technical Corrections:

- 21.Line 72: missing period

Response: Thank you for pointing out this minor typographical issue. We note that the sentence referred to in the original manuscript has been removed in the revised version as part of a substantial reorganization and condensation of the Introduction section. The revised Introduction now presents a more concise and focused discussion of prior work on diurnal FRP reconstruction and multi-source fire emission inventories, and the specific punctuation issue at Line 72 is therefore no longer applicable.

- 22.Line 77: “understate” might not be the best word, maybe underestimate, limit, etc.

Response: Thank you for this helpful suggestion regarding word choice. We note that the sentence containing the term “understate” in the original manuscript has been removed in the revised version as part of a substantial reorganization of the Introduction.

- 23.Line 120: missing space

Response: Thank you for pointing this out. The missing space has been corrected in the revised manuscript.

- 24.Line 332: delete first comma?

Response: Thank you for pointing this out. The formatting has been corrected to “Texas Commission on Environmental Quality (2022)” in the revised manuscript.