General Comments:

This paper generates three-dimensional biomass-burning emissions for Southeast and East Asia by developing new fire diurnal cycles and vertical injection profiles. The proposed diurnal cycle is derived by integrating fire radiative power data from both geostationary and polar-orbiting satellite observations. The vertical injection profile is produced using a machine-learning model trained on satellite-retrieved smoke plume heights and meteorological variables. However, the manuscript is not well-structured, and the methodology lacks clarity and robustness. Substantial revisions are needed before the work can be considered for publication.

Response: We thank the reviewer for this comprehensive assessment. We agree that the previous version of the manuscript did not sufficiently convey the methodological coherence and robustness of the proposed framework. We have conducted a substantial revision of the manuscript, focusing on restructuring, clarification, and strengthening of the methodological presentation, as summarized below. **Please note that text in "*italics and underlining*" represents the revised sentences in the modified manuscript.**

(1) Structural reorganization and methodological integration.

The manuscript has been reorganized to explicitly present the SEAF product as a single end-to-end workflow, rather than as independent derivations of diurnal emissions and plume heights. The revised structure emphasizes the logical sequence from FRP fusion and diurnal reconstruction to smoke plume height prediction and final 3D emission construction. This integrated framework is now explicitly summarized in Eq. (12) and visually illustrated in the revised Figure 2, clarifying that diurnal variability and vertical injection are coupled at consistent spatial and temporal resolutions.

(2) Improved methodological clarity and robustness.

To enhance transparency and reproducibility, we have expanded the description of key methodological components, particularly the machine-learning framework for SPH prediction. Detailed information on dataset composition, sample selection, training–testing strategy, hyperparameter tuning, and performance metrics has been added. In addition, a dedicated validation subsection (Section 2.3.6) has been introduced to systematically describe the evaluation of diurnal emissions, plume heights, and the resulting 3D emission structure using multiple independent observational datasets and statistical indicators.

(3) Strengthened regional context and evaluation framework.

The Introduction has been expanded to better reflect prior studies conducted in SEA, including work on agricultural burning, peatland fires, combined VIIRS–AHI emission estimation, and extreme regional fire events. These additions clarify how the present study builds upon existing regional research while addressing key gaps, particularly the lack of a consistent hourly 3D emission inventory for this region. Furthermore, the evaluation framework has been strengthened by incorporating additional widely used emission inventories (including GFED v5.1), providing clearer interpretation of inter-inventory differences, and explicitly justifying the use of satellite-based CO as an observation-based tracer for BB emission evaluation.

We are confident that these comprehensive revisions effectively address the reviewer's concerns regarding structure, clarity, and robustness, significantly enhancing the overall scientific quality of the manuscript.

Major comments:

1. **Weak integration between the fire diurnal cycle and vertical injection profile components**

Although the overarching objective is to develop a three-dimensional biomass-burning emission dataset, the manuscript presents the derivation of the fire diurnal cycle and the vertical injection profile as largely independent processes. The authors first generate and validate 2-D fire emissions, then separately develop and validate smoke plume heights. However, the connection between these two components—and how they integrate to form the final 3-D emissions—is not clearly articulated. I recommend extensively restructuring the manuscript to better emphasize the methodological coherence and the interdependencies between these two parts.

Response: We thank the referee for pointing out that the previous manuscript structure may have given the impression that the diurnal emission reconstruction and the vertical injection profile were treated as independent components. We agree that this was an issue of presentation rather than methodology.

To address this, we revised the manuscript to explicitly present the construction of the SEAF product as a single end-to-end workflow. The integrated generation of hourly 3D emissions is now summarized by **Eq. (12)**, which explicitly links fused hourly FRP, column-integrated emissions, Random Forest–predicted SPH, and layer-wise vertical allocation. This formulation clarifies that FRP simultaneously controls the temporal evolution of column emissions and serves as an input to the plume height prediction, which constrains the vertical distribution. In addition, **Figure 2** has been reorganized to highlight the unified workflow from hourly FRP to final 3D emissions, emphasizing that diurnal variability and vertical injection are coupled at the same spatial and temporal resolution. The evaluation of diurnal emissions and plume heights is conducted separately as a modular validation strategy to isolate uncertainties from different process components, rather than implying methodological independence. In practice, both the diurnal reconstruction and SPH prediction are applied to the same hourly FRP fields, ensuring that temporal variability and vertical allocation are driven by a consistent set of event-scale fire intensities.

2. **Insufficient background and regional context**

The literature review does not adequately cover relevant studies conducted in Asia. Much of the discussion focuses on work from the United States or other regions, while several closely related studies in Asia are overlooked. These omissions weaken the contextual grounding of the study and obscure how this work builds upon or differs from existing regional research. The authors should expand the background section to include key Asian studies and clearly articulate their linkages to the current work.

References:

'Dynamics of fire plumes and smoke clouds associated with peat and deforestation fires in Indonesia'

'Fire Particulate Emissions from Combined VIIRS and AHI Data for Indonesia, 2015-2020' 'Improved estimation of fire particulate emissions using a combination of VIIRS and AHI data for Indonesia during 2015–2020'

'Highly anomalous fire emissions from the 2019–2020 Australian bushfires'

Response: We sincerely thank the reviewer for this constructive comment regarding the insufficient background and regional context. We acknowledge that the previous version of the manuscript did not adequately integrate key Asian studies, which are essential for properly grounding this work in the specific fire regimes of SEA. To address this concern, we have substantially revised and expanded the Introduction to explicitly incorporate the suggested regional literature and to clearly articulate how the present study builds upon, and extends beyond, existing Asian research. The main revisions are summarized as follows.

(1) Strengthening the description of regional fire regimes in SEA.

We have expanded the background discussion to explicitly highlight the unique fire characteristics of SEA, with particular emphasis on peatland and deforestation fires in Indonesia. These fires exhibit distinct smoke plume dynamics and emission potentials compared to other regions, with strong implications for regional radiative balance (Tosca et al., 2011). This revision is reflected in the Introduction, where we now state that (*Notably, the SEA region exhibits unique fire regimes that require dedicated regional focus. For instance, peatland and deforestation fires in Indonesia possess distinct smoke plume dynamics and emission potentials compared to other regions, exerting a strong influence on the regional radiative balance (Tosca et al., 2011).*)

(2) Incorporating extreme climate-driven fire events and associated uncertainties.

We have added discussion of how climate forcing under warming and drying conditions can amplify fire emissions and increase uncertainties in conventional emission inventories, with explicit reference to El Niño–related extreme fire events in SEA. In particular, the severe Indonesian fires in 2015 are now cited as a representative example (Field et al., 2016; Huijnen et al., 2016). This is reflected in the revised Introduction, where we note that (*Furthermore, climate forcing under warming and drying conditions can substantially amplify fire emissions and drive strong deviations from climatological means, thereby increasing uncertainties in conventional emission inventories (Li et al., 2021). This effect is particularly relevant in SEA, where extreme fire activity frequently occurs during El Niño-related droughts, such as the severe Indonesian fires in 2015 (Field et al., 2016; Huijnen et al., 2016).*).

(3) Clarifying the linkage to existing VIIRS–AHI–based Asian studies.

We have explicitly incorporated recent regional studies that combine VIIRS and AHI observations to improve high-frequency fire emission estimates in Asia, including work focused on SEA (Lu et al., 2022; Li et al., (2019, 2022); Zheng et al., (2021)). These studies are now discussed in the Introduction to acknowledge prior regional efforts in multi-sensor FRP fusion ((1) *Recent studies have attempted to combine VIIRS and AHI data to improve high-frequency particulate emission estimates in specific areas such as Indonesia (Lu et al., 2022);*(2) *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.*).

(4) Explicitly identifying the remaining regional gap addressed by this study.

Building on the expanded regional background, we now explicitly clarify that several key methodological limitations remain in existing Asian biomass-burning studies, which are directly addressed in the present work. These limitations include (i) the reliance on static diurnal representations to reconstruct FRP under conditions of cloud occlusion or limited temporal sampling, which can bias emissions during non-active periods or dampen peak fire activity during extreme events, and (ii) the lack of plume-height-resolved emissions and the absence of a consistent hourly three-dimensional emission framework across SEA.

These points are now explicitly stated in the revised Introduction ((1) *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*; (2) *these studies are generally limited to two-dimensional (2D) emission estimates and specific fire episodes or subregions, without explicitly resolving plume injection heights or providing a consistent hourly 3D emission framework across SEA*). By explicitly identifying these gaps, the revised Introduction now clearly motivates the development of the SEAF inventory as a systematic, high-resolution, observation-driven hourly three-dimensional BB emission dataset for SEA.

3. Methodological issues and lack of robustness

• Choice of VIIRS 375-m product:

The manuscript uses 375-m VIIRS FRP, but this product is known to saturate under highintensity fires. Why was the 375-m product chosen instead of the 750-m VIIRS FRP? This choice needs stronger justification beyond the higher spatial resolution. Have the authors tested or compared the 750-m FRP?

Response: Thank you for the reviewer's professional inquiry regarding the applicability of the VIIRS 375 m active fire product. We agree with the reviewer that the VIIRS 375 m I-band, particularly the I4 channel, may be subject to detector dynamic range limitations under extremely high-intensity fire conditions, potentially leading to brightness temperature saturation or folding and thus affecting I-band-based fire detection and classification (Schroeder et al., 2020; Zhang et al., 2017).

It is important to clarify that the VIIRS 375 m active fire product used in this study (e.g., JPSS VIIRS Products-VIIRS Active Fires I-Band EDR) does not estimate FRP directly from the I4 channel. According to the NOAA VIIRS I-band active fire algorithm, the I-band observations are primarily used for precise fire pixel detection and classification, while the quantitative FRP retrieval is based on the co-located 750 m M13 dual-gain mid-infrared radiance data. In practice, the FRP retrieved for a single 750 m M13 pixel is distributed among the coincident 375 m I-band fire sub-pixels, resulting in an FRP product reported at 375 m spatial resolution (Schroeder et al., 2020).

Therefore, although the FRP is reported at 375 m resolution, its physical retrieval already incorporates information from the M-band, and it is not equivalent to an FRP estimate derived solely from I-band radiances. Previous studies have further demonstrated that, owing to its larger pixel size and substantially enhanced dynamic range, the M13 band is rarely affected by saturation under active fire conditions. For example, Zhang et al., (2017) showed that in agriculturally dominated regions, while some intense fires may saturate the I4 channel, the M13 band remained largely unsaturated and was able to provide stable and reliable FRP estimates at the locations of I-band-detected fire pixels.

Based on these product characteristics, we selected the VIIRS 375 m FRP product as the primary fire radiative input for this study. Fire activity in SEA is strongly influenced by agricultural burning and is characterized by highly fragmented spatial patterns, with a large number of small and edge fires, making fire detection particularly sensitive to spatial resolution and omission errors (Huang et al., 2024; Vadrevu et al., 2022; Yin, 2020). Compared with the 750 m active fire product, the VIIRS 375 m product provides more accurate fire localization and a better representation of spatial heterogeneity, which is critical for the construction of high-resolution emission inventories and for the analysis of transboundary pollutant transport (VIIRS I-Band 375 m Active Fire Data | NASA Earthdata, 2025; Schroeder et al., 2014).

To avoid any potential ambiguity, we have also added a brief clarification in Section 2.1.1 (*Although the product is reported at 375 m spatial resolution, the I-band observations are mainly used for fire detection and localization, whereas FRP is retrieved based on the co-located 750 m M13 dual-gain mid-infrared radiance data and then allocated to the detected 375 m fire pixels…*) of the revised manuscript, explicitly stating that although the VIIRS product is reported at 375 m resolution, FRP is retrieved using the co-located 750 m M13 dual-gain mid-infrared radiance data following the standard VIIRS I-band active fire algorithm.

4• AHI FRP correction:

The reference (Li et al., 2022) cited for the AHI FRP correction does not actually use AHI FRP, making this citation inappropriate.

Response: We agree that Li et al. (2022) does not involve Himawari-8/9 AHI FRP and have therefore removed this citation. The manuscript now cites Xu et al. (2022, 2023), which explicitly apply VIIRS-referenced calibration to correct AHI FRP (*Following (Xu et al., 2022, 2023), cloud-corrected VIIRS FRP was therefore adopted as an external reference to perform cross-sensor calibration of Himawari-8/9 AHI FRP using collocated observations, with the aim of reducing systematic biases associated with sensor characteristics and spatial resolution differences*), and clarifies that the specific cloud-corrected, multi-level fallback calibration framework is developed in this study.

5Unclear machine-learning description:

The section describing the machine-learning method lacks clarity. Important details such as the training–testing strategy, sample selection, and dataset composition are not provided.

Response: Thank you for this constructive comment. We agree that providing more technical details of the machine-learning workflow is essential for reproducibility. Following the reviewer's suggestion, we have significantly expanded Section 2.3.5 to explicitly document the modeling procedure. The specific additions are summarized below.

(1) Dataset composition and sample selection:

We clarified the data cleaning and filtering procedure used to construct the training dataset. (*Raw satellite observations were first spatially filtered to match the study domain, and records with non-physical values or missing FRP information were removed. This procedure resulted in a finalized dataset comprising 2,127 samples.*)

(2) Training–testing strategy:

A standard random split was implemented to ensure model generalizability and independent evaluation. (*For model development, the dataset was randomly divided into a training set (80%) and an independent testing set (20%).*)

(3) Hyperparameter tuning:

To ensure model robustness and avoid overfitting, we applied a grid-search strategy combined with cross-validation. (*A grid search combined with 5-fold cross-validation was employed to optimize model hyperparameters, yielding an optimal configuration of 200 trees (n_estimators) with a maximum tree depth of 10.*)

(4) Model performance:

Quantitative performance metrics were added to demonstrate the predictive reliability of the final model. (*The finalized RF model demonstrated strong predictive skill, with a root mean squared error (RMSE) of 334.68 m and an R^2 of 0.90 on the test set.*)

6• **Missing validation subsection in the Methods:**

A dedicated validation subsection should be added to the Methods section to clearly explain the evaluation workflow.

Response: Thank you for this constructive suggestion. We agree with the reviewer and have added a dedicated validation subsection to the Methods section. Specifically, a new subsection entitled "**2.3.6 Validation and evaluation strategy**" has been included.

This subsection systematically describes the evaluation workflow, including the validation of Random Forest–predicted smoke plume heights using independent MISR observations with statistical metrics (RMSE, R^2, R, and bias), the assessment of two-dimensional emissions through comparisons with TROPOMI CO columns and multiple BB emission inventories, and the evaluation of the three-dimensional emission structure using MISR, CALIPSO, and existing injection-height schemes. Detailed validation results are presented in Section 3. The validation strategy follows the same hierarchical structure as the emission construction, progressing from diurnal FRP evaluation (2D), to plume height prediction, and finally to the integrated 3D emission structure.

7• Lack of GFED products in evaluation:

It is unclear why the widely used GFED-related products are omitted from the evaluation. Including them would provide a more comprehensive comparison.

Response: Thank you for this constructive suggestion. We agree that including GFED-related products provides a more comprehensive benchmark and strengthens the inter-inventory comparison. We have therefore incorporated GFED v5.1 into our evaluation framework and updated the corresponding figures accordingly.

(1) Reason for the initial omission and clarification of product availability:

In the initial version, GFED v5.1 was not included because it was not publicly available at the time of manuscript preparation, and the latest officially released emissions only extended to 2022, which did not cover the 2023 study period analyzed in this work. To ensure temporal consistency across all inter-comparisons, GFED was therefore not included in the initial evaluation. We have now utilized the recently released extended version of GFED v5.1 and incorporated it into the revised manuscript, avoiding any ambiguity regarding the product release cycle.

(2) Revisions implemented in the manuscript:

We have updated Figures 10–11 and Figures S6–7 to include comparisons with GFED v5.1. This addition enables a more complete comparison against a widely used global burned-area based inventory and complements the FRP-constrained products discussed in this study.

(3) Value added to the evaluation framework:

Including GFED v5.1 strengthens the evaluation by placing SEAF in the context of both FRP-based and burned-area-based global emission frameworks, thereby improving the interpretability of inter-inventory differences and further supporting the robustness assessment of the SEAF product.

8• **Section 3.3.2 — Rationale for comparison with inventories:**

The manuscript evaluates the new emissions against existing emission inventories but mainly describes the differences without explaining the underlying causes. For example, why are the results lower than FINN but closer to FEER and QFED? Additional interpretation is needed in the Results or Discussion sections.

Response: Thank you for this constructive comment. We agree that providing a physical and methodological rationale for the inter-inventory discrepancies is essential for demonstrating the robustness of the SEAF inventory. Following your suggestion, we have added detailed interpretations in both the Results (Section 3.3.2) and the Discussion (Section 4) to clarify why SEAF estimates are lower than FINN but closer to FEER and QFED. The revisions explain the underlying causes from the following aspects.

(1) Methodological framework (FRP-based vs. burned-area-based): We clarify that SEAF, FEER, QFED, and IS4FIRES are all constructed within a top-down framework constrained by fire radiative power (FRP), whereas FINN relies on a burned-area-based approach. (*By resolving heterogeneous emission structures and sharp spatial gradients, SEAF captures small-scale fire clusters and localized hotspots that are often omitted or attenuated in burned-area-based products like GFED v5.1, which tend to display comparatively smooth and spatially diffuse emission patterns.*)

(2) Dynamic temporal representation of fire activity: A key driver of the inter-inventory differences lies in how fire activity is characterized over time. We emphasize that SEAF explicitly reconstructs sub-daily fire variability through a dynamic diurnal adjustment. (*Supported by the dynamically reconstructed diurnal FRP patterns, SEAF reproduces a pronounced seasonal peak of approximately 500 Gg month^{-1} in Region 2, whereas inventories relying on infrequent sampling from polar-orbiting sensors, such as GFAS v1.2 and FEER v1.0, tend to underestimate this seasonal maximum.*)

(3) Treatment of smoldering-dominated peatland fires: We further discuss peatland-dominated fires, which are prevalent in parts of Southeast Asia and remain challenging for FRP-derived top-down approaches because deep smoldering combustion may be weakly expressed in FRP observations. (*We acknowledge that FRP-constrained approaches, including SEAF, are generally more conservative in regions where smoldering combustion is prevalent because cool fires are more difficult to detect and quantify using thermal infrared sensors compared to burned-area algorithms, which remain highly sensitive to assumptions on fuel consumption and burning depth.*)

(4) Quantitative and spatial consistency with observations: These methodological differences are reflected in the quantitative results and spatial patterns. (*For 2023, SEAF's annual PM2.5 estimate of 2362 Gg yr-1 lies within the central range of the inter-inventory spread, showing close agreement with FEER v1.0 and QFED v2.6r1 while corresponding to a reduction of approximately 67% relative to the burned-area-driven FINN v2.5.1.*)

Note on emission factors: We emphasize that while emission factors are a known source of uncertainty in BB inventories, all inventories compared in this study apply emission factors derived from similar literature-based compilations. Accordingly, the observed spread among inventories over Southeast Asia arises primarily from differences in fire activity characterization, temporal representation, and spatial resolution, rather than from emission factor selection alone.

9• **Lines 321–324 — Transferability of methods:**

These lines describe a core component of the method, yet the supporting references are based on studies from the U.S. and Europe. The authors should justify whether such methodologies are appropriate for Asian fire regimes.

Response: Thank you for raising this important comment. We agree that some of the references supporting this methodological component were originally developed and validated under fire regimes in the United States and Europe, and that the transferability of these approaches to Asian fire conditions therefore requires clarification.

(1) We have revised the Introduction to explicitly clarify why methodologies developed in other regions remain applicable to SEA (**see our response to 2. Insufficient background and regional context**).

(2) In the revised manuscript, we now explicitly clarify that the core of the proposed methodology is grounded in general physical relationships between FRP, energy release, and biomass combustion (Eq. 12). These relationships are not region-specific and have been demonstrated to be broadly applicable across different fire types and fire regimes.

(3) More importantly, the present study does not directly transplant existing methods to the Asian region. Instead, the approach is specifically adapted through region-dependent calibration and constraints to better represent Asian fire conditions. These adaptations include the use of VIIRS and AHI observations over SEA, regionally optimized diurnal reconstruction schemes, and evaluation against independent observations (e.g., TROPOMI CO) as well as multiple existing emission inventories over the SEAF domain. The spatial patterns, seasonal evolution, and regional consistency shown in Figures 10–11 and Figure S8 provide empirical evidence for the robustness of the method under Asian fire regimes characterized by fragmented land cover, agricultural burning, and peatland fires.

10• **Line 465 — Limited validation case:**

Validating the method using only a single biomass-burning episode is insufficient. More cases are needed to demonstrate robustness.

Response: Thank you for this valuable comment. We agree that validation based on a single BB episode is insufficient to establish full robustness at the event scale. In the revised manuscript, we have clarified the role of the event-scale analysis, explicitly acknowledged this limitation, and explained how the overall robustness of the SEAF inventory is supported by broader evaluations. The revisions are summarized as follows:

(1) Clarification of scope and acknowledgment of limitations: We explicitly distinguish between physical plausibility at the event scale and methodological robustness at regional and longer temporal scales. (**4. Discussion** *In addition, validation based on a single BB episode is insufficient to establish full robustness at the event scale (Figure 9). In this study, the event-scale analysis serves to demonstrate the physical plausibility of the fused fire emission product, whereas methodological robustness is primarily supported by regional- and long-term statistical consistency across multiple emission inventories.*)

(2) Observational constraints in SEA: We provide a practical explanation for the limited availability of suitable event-scale validation cases in the study region. (**4. Discussion** *Furthermore, systematic inspection of operational satellite imagery (e.g., NOAA STAR, https://www.star.nesdis.noaa.gov/mapper/) reveals that the identification of isolated, cloud-free plumes in SEA is severely constrained by persistent cloud cover and plume interference. Therefore, additional observational evidence is required to further validate the method at the event scale*)

(3) Robustness supported by multi-scale evaluation:

We emphasize that the main conclusions of this study do not rely on the single-event analysis alone, but are supported by extensive regional and multi-temporal evaluations presented throughout Section 3, including comparisons with independent satellite observations and multiple emission inventories.

(4) Outlook for future validation: We note that additional event-scale validation will be pursued as more suitable observational constraints become available. (**4. Discussion** *Therefore, additional observational evidence is required to further validate the method at the event scale as data availability improves*.)

We believe that this transparent discussion of the limitations of event-scale validation, together with the comprehensive regional and statistical evaluations presented in the manuscript, adequately addresses the reviewer's concern.

11.Errors in basic information

There are several factual inaccuracies that need correction. For example:

• Mixing up sensor and satellite names (line 97).

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

12• Stating that MODIS fails to capture nighttime events (lines 131–132).

Response: Thank you for pointing this out. We have revised the text.

13• Citing a reference for AHI FRP correction that does not actually use AHI FRP.

Response: Thank you for pointing this out. Please see our response to **4• AHI FRP correction**.

Minor to moderate comments:

14• **Lines 33–34:** Please specify the versions of all datasets used—at minimum in the Data section—since different versions may produce substantially different values.

Response: Thank you for this comment. We agree that specifying dataset versions is essential for reproducibility and comparability. We have revised both the Abstract and the Data section accordingly. In the revised Abstract, all emission inventories included in the inter-comparison are now explicitly identified with their corresponding version numbers, as reflected in the following sentence: (*with estimates lower than FINN v2.5.1 (67%) and GFED v5.1 (25%), but closely aligned with FEER v1.0, QFED v2.6r1, and IS4FIRES v2.0*.) In addition, the Data section has been updated, and Table S2 now comprehensively documents the versions of all datasets used in this study, ensuring transparency and consistency throughout the manuscript.

15• **Lines 33–34:** Why not provide comparison results for all emission inventories included in the study?

Response: We appreciate the reviewer's suggestion. We agree that the scope of the emission inventory comparison should be clearly communicated in the Abstract. In the revised Abstract, we now explicitly state that SEAF is compared against six widely used global BB emission inventories, and we summarize the key quantitative relationships in a concise and balanced manner, as shown in the following revised sentence: (*Annual $PM_{2.5}$ emissions in SEAF are approximately 2362 Gg y^{-1}, placing it within the central range of six widely used global BB inventories, with estimates lower than FINN v2.5.1 (67%) and GFED v5.1 (25%), but closely aligned with FEER v1.0, QFED v2.6r1, and IS4FIRES v2.0.)*Detailed comparisons for all inventories are presented and discussed in the main text and supplementary material.

16• **Line 35:** Please use statistical metrics to demonstrate the performance of the SPH estimation rather than presenting a single numerical value.

Response: Thank you for this comment. We agree that a single numerical value is insufficient to characterize the performance of smoke plume height (SPH) estimation. In the revised Abstract, we now report multiple statistical metrics to quantitatively evaluate SPH prediction performance, as shown in the following revised sentence: *The RF-SHAP framework successfully predicts SPH (R^2 = 0.90, RMSE = 335 m) with over 90% of estimates within ± 500 m.*

17• **Line 38:** "Satellite observations" should be specified. If you mean MISR SPH, please explicitly name the product here.

Response: Thank you for pointing this out. We agree that the original wording referring to "satellite observations" was too general. In the revised Abstract, we now explicitly specify the satellite products used at different stages of the analysis. The relevant revised sentences are: (..., *yielding vertical profiles that are more consistent with MISR and CALIPSO observations.*).

18• **Lines 38–41:** It seems inconsistent that the study's final goal is to generate 3-D fire emissions, yet this section focuses only on analyzing drivers of SPH.

Response: We appreciate this comment and agree that the original wording did not sufficiently emphasize the role of SPH analysis in achieving the final objective of constructing a 3D emission inventory. In the revised Abstract, we clarify that SPH prediction and interpretation are not standalone objectives, but are explicitly used to constrain the vertical allocation of emissions. This is reflected in the following revised sentence: (*The fused FRP, together with ERA5 meteorology, drives a random forest (RF) model trained on MISR smoke plume heights (SPH) observations to predict SPH, which are then used to guide a multi-layer vertical allocation of emissions to construct the 3D emission inventory.; Compared with the widely used IS4FIRES v2.0 inventory, the resulting 3D SEAF dataset effectively mitigates near-surface–biased emission allocation and improves the representation of elevated smoke injection during peak burning periods, yielding vertical profiles that are more consistent with MISR and CALIPSO observations.*).

19• **Line 41:** The phrase "are anticipated to" is not appropriate, since you have already generated an observation-driven, hourly 3-D biomass-burning emission dataset for SEA.

Response: We thank the reviewer for this helpful comment. We agree that the phrase "are anticipated to" was overly tentative and did not accurately reflect the completed nature of the SEAF dataset. In the revised Abstract, this wording has been replaced with a definitive statement. The revised sentence now reads: (*By jointly mitigating systematic underestimation during key burning periods and alleviating low-altitude allocation bias while preserving elevated smoke occurrences, the SEAF inventory provides an observation-driven hourly 3 km 3D BB emission dataset for SEA, with improved temporal and vertical realism, supporting air quality and climate assessment applications.*)

20•Lines 47–51: If the manuscript focuses on validating CO, why not center the discussion on CO emissions in this section?

Response: Thank you for this suggestion. We agree that the role of CO should be made explicit in the Introduction. To address this point without expanding the scope of the section, we have revised the Introduction to explicitly include carbon monoxide (CO) among the major trace gases emitted from BB and to note that CO is commonly used as a tracer for BB (*Black carbon (BC) and primary organic aerosols (POA) derived from BB account for approximately 40% and 65% of global BC and POA emissions, respectively, while non-methane organic gases, carbon monoxide (CO, commonly used as a tracer for BB), and greenhouse gases such as methane (CH4), carbon dioxide (CO2), and nitrous oxide (N2O) contribute significantly to atmospheric chemistry and radiative forcing (Bond et al., 2013; Gkatzelis et al., 2024; Griffin et al., 2024).*). This minor revision clarifies the motivation for using CO as an observation-based tracer in the emission evaluation presented later in the manuscript.

21• **Lines 56–58:** The justification for selecting the SEA study area is not strong or straightforward.

Response: Thank you for this comment. We have revised the Introduction to clarify and strengthen the rationale for selecting SEA as the study region (**see our response to 2. Insufficient background and regional context and 9• Lines 321–324 — Transferability of methods**).

22• **Lines 56–81:** The background research for Asia is insufficient. Much prior work outside Asia is discussed, yet several important regional studies are missing.

Response: Thank you for this comment. We agree that the background research for Asia was insufficiently represented in the original version. In the revised Introduction, we have substantially expanded and restructured the regional background to better reflect prior work focused on SEA (**see our response to 2. Insufficient background and regional context and 9• Lines 321–324 — Transferability of methods**).

23• **Line 95:** Please add “three-dimensional” here for accuracy.

Response: Thank you for this helpful suggestion. We have added “3D” at Line 144 to improve accuracy.

24• **Line 97:** VIIRS and NOAA-20 should not be presented as parallel entities; one is a sensor, the other is a satellite platform.

Response: Thank you for pointing this out. We have revised the text (*SEAF was generated by fusing FRP data from the Advanced Himawari Imager (AHI), the VIIRS instruments onboard both the Suomi-NPP and NOAA-20 satellites.*).

25• **Lines 102–103:** Please remove the final sentence—its style is more appropriate for a proposal rather than a manuscript.

Response: Thank you for this suggestion. The final sentence has been removed.

26• **Lines 129–130:** The description of Himawari-8's spatial and temporal resolution is unnecessary here, as these specifics have already been provided.

Response: Thank you for the comment. We have revised the sentence to focus on the role of Himawari-8 observations over the SEA study region and removed redundant descriptions.

27• **Lines 131–132:** MODIS can detect nighttime events due to its ~1:30 a.m. overpass, although it may not capture all nighttime fires. This statement should be corrected.

Response: Thank you for this clarification. We have corrected the statement accordingly.

28• **Lines 163–165:** Please clarify what constitutes the testing dataset.

Response: Thank you for this comment. We have clarified the construction of the testing dataset in the revised manuscript by explicitly describing the quality control procedure, the train–test split (80% training and 20% independent testing), and the use of cross-validation and test-set performance metrics.

29• **Lines 171–174:** Why not directly include the common fire weather index variables (temperature, relative humidity, wind, and precipitation)?

Response: Thank you for this comment. We clarify that the ERA5 predictors used in this study already include near surface temperature, wind components, and precipitation. For atmospheric moisture, we used 2 m dew point temperature rather than relative humidity, since relative humidity can be directly derived from temperature and dew point and provides largely redundant information. This choice also helps reduce multicollinearity among predictors while retaining the physically relevant moisture constraint for plume development.

30• **Lines 177–180:** Why are widely used GFED products not included?

Response: Thank you for this constructive suggestion. **see our response to 7• Lack of GFED products in evaluation**.

31• **Section 2.3.1:** If this section refers to VIIRS FRP data calibration, please revise the title to reflect that.

Response: Thank you for this suggestion. We have revised the section title to explicitly reflect that this section describes the cloud correction of VIIRS FRP data.

32• **Line 5.7:** Please indicate the source of the test data.

Response: Thank you for pointing this out. We have clarified the source and partitioning of the test data in the revised manuscript. The test data is derived from the 2,127 MISR smoke plume height samples described in Section 2.1.3. To ensure a robust evaluation, the total dataset was randomly divided into a training set (80%) and an independent testing set (20%).

33• **Figure 2:** Please add a legend explaining the colors.

Response: Thank you for this suggestion. We have revised the caption of Figure 2 to include a clear explanation of the color-coded framework. Specifically, we have added a description clarifying that:

- Green boxes represent input datasets/reference frameworks.
- Blue boxes denote intermediate data processing, calibration, and calculation steps.
- The pink circle represents the machine learning core (RF-SHAP framework).
- Orange boxes represent the final output products, including the predicted SPH and 3D emission inventory.

34• **Figure 4:** Use either "correction" or "calibration" consistently throughout the text and figures.

Response: Thank you for pointing this out. We agree that the terminology should be used consistently and precisely. In this study, two different adjustment processes are applied to different sensors. Specifically, **VIIRS FRP is subject to cloud correction**, which aims to compensate for missing detections caused by cloud coverage without altering sensor-related biases. In contrast, Himawari FRP is adjusted through **cross-sensor calibration**, in which cloud-corrected VIIRS FRP serves as an external reference to correct systematic bias in Himawari observations. To avoid ambiguity, we have revised the figure caption and relevant text.

Reference

VIIRS I-Band 375 m Active Fire Data | NASA Earthdata: https://www.earthdata.nasa.gov/data/instruments/viirs/viirs-i-band-375-m-active-fire-data, last access: 22 December 2025.

Huang, H., Jin, Y., Sun, W., Gao, Y., Sun, P., and Ding, W.: Biomass burning in northeast China over two decades: Temporal trends and geographic patterns, Remote Sens., 16, 1911, https://doi.org/10.3390/rs16111911, 2024.

Schroeder, W., Oliva, P., Giglio, L., and Csiszar, I. A.: The new VIIRS 375m active fire detection data product: algorithm description and initial assessment, Remote Sens. Environ., 143, 85–96, https://doi.org/10.1016/j.rse.2013.12.008, 2014.

Schroeder, W., Giglio, L., Csiszar, I., and Tsidulko, M.: Algorithm theoretical basis document for NOAA NDE VIIRS I-band (375m) active fire, National Oceanic and Atmospheric Administration: Washington, DC, USA, 2020.

Vadrevu, K., Eaturu, A., Casadaban, E., Lasko, K., Schroeder, W., Biswas, S., Giglio, L., and Justice, C.: Spatial variations in vegetation fires and emissions in South and Southeast Asia during COVID-19 and pre-pandemic, Sci Rep, 12, 18233, https://doi.org/10.1038/s41598-022-22834-5, 2022.

Yin, S.: Biomass burning spatiotemporal variations over South and Southeast Asia, Environment International, 145, 106153, https://doi.org/10.1016/j.envint.2020.106153, 2020.

Zhang, T., Wooster, M. J., and Xu, W.: Approaches for synergistically exploiting VIIRS I- and M-Band data in regional active fire detection and FRP assessment: A demonstration with respect to agricultural residue burning in Eastern China, Remote Sensing of Environment, 198, 407–424, https://doi.org/10.1016/j.rse.2017.06.028, 2017.