

General comments

The new version of the paper has solved most of my concerns, although some issues persist. Even though the authors claim to have answered the concerns of two reviewers over the uncertainty estimation, the new analysis they provide does not seem to be very convincing, as they focused on comparing with the reference product GFED5, instead of providing a full uncertainty characterization by taking into account the different factors of uncertainty relevant for their model, related to the input data, including the static hotspots mask, the estimation of BA from AF, and the emission model: fuel loads and CC. To me this is more an inter-comparison analysis than an uncertainty characterization.

In fact, in the text the section where this comparison is made is entitled “Evaluation of the GFED5NRT”, including also the comparison with HR data. The aim of this section is to show that the product is consistent with GFED5, rather than the analysis of the uncertainty factors. In fact, the references to the uncertainty of the product are mostly introduced in the supplementary material and a brief comment in the conclusion section (lines 447-49). I wonder why the uncertainty (expressed as the dispersion around the estimated mean) is not considered in the time series (figure 4 or 7, for instance).

We thank the reviewer for their careful assessment and agree that understanding dataset uncertainty is critical. To partially address this comment, we have moved a figure and a table describing the performance of GFED5NRT relative to GFED5 from the Supplementary Material to the main text. The new Figure 5 (formerly Fig. S3) shows the global spatial distribution of correlations between GFED5NRT and GFED5. The new Table 1 (formerly Table S5) reports the normalized root mean square error and Taylor skill scores for different types of comparisons between the two datasets across different regions.

We acknowledge that our evaluation is fundamentally an inter-comparison and benchmarking exercise rather than a formal, bottom-up uncertainty characterization, and that this is a deliberate scope choice rather than an oversight, for the reasons detailed below.

Standard Practices in Global Fire Datasets: Providing a full, formal uncertainty characterization is not standard practice for global operational fire emissions products. Existing global remote sensing and process-based datasets—such as GFAS, QFED, FINN, and the GBBEPx system—rely primarily on inter-comparisons with higher-resolution data, cross-evaluations with other inventories rather than attempting formal error propagation, and comparisons with atmospheric observations. By benchmarking GFED5NRT against other global datasets and high-resolution regional data, we align with accepted community standards for operational fire data validation. In other work underway, we make systematic comparisons of GFED5 with atmospheric observations.

Challenges in Formal Error Propagation: Executing a formal error propagation analysis for global burned area and emissions is highly challenging and can yield artificial confidence intervals. The uncertainties in our framework originate from highly varied sources that defy straightforward mathematical quantification, including satellite retrieval errors, the static hotspot masking process, and the complex representation of carbon flows within the CASA biogeochemical model.

Furthermore, the numerous empirical scaling choices and assumptions regarding fuel loads and combustion completeness make classical error propagation deeply unreliable.

Scope and Objectives of this Data Descriptor: This manuscript serves strictly as a data descriptor for GFED5NRT, which was designed to provide continuity following the orbital drift of the MODIS sensors. Because the primary objective is to maintain operational and statistical consistency with the existing GFED5 historical record, the core validation metric must be conformity with the standard GFED5 time series. Attempting a full uncertainty analysis of the baseline GFED5 time series would require foundational re-evaluations of the standard product that are strictly outside the scope of this extension paper.

Comprehensive Dataset Evaluation Performed: In our previous revisions, we added extensive evaluations to quantify the uncertainties specific to our new methodology. We assessed the spatial and temporal uncertainties introduced by the scaling algorithm using multiple metrics including Taylor Skill Scores (TSS), normalized root-mean-square error (nRMSE), correlation coefficients, and linear regression slopes. Additionally, we systematically evaluated GFED5NRT against independent references, including the VIIRS VNP64A1 burned area product, daily emissions inventories (FINN, GFAS, GBBEPx), and high-resolution regional datasets across the globe (e.g., INPE, MapBiomass, NBAC, MTBS, EFFIS, ONFIRE).

Ultimately, while we agree that quantifying uncertainty is vital, a comprehensive bottom-up error propagation is neither standard practice for operational fire datasets nor readily feasible given the complex and interdependent modeling steps involved. The primary goal of this data descriptor is to establish a reliable near-real-time transition from MODIS to VIIRS observations. In the revised manuscript, we have added a new section 4 ‘*Uncertainty*’, which describes the uncertainty assessments conducted in our previous studies and in the present study, and explains why we adopted an intercomparison-based approach rather than attempting formal error propagation. The new section reads as follows:

“

4. *Uncertainty*

Uncertainty in GFED5NRT stems from two primary sources: baseline uncertainties inherited from the GFED5 framework, and methodological uncertainties introduced by the NRT scheme developed here. We distinguish these components because GFED5NRT is designed to provide a methodologically consistent temporal extension of the established GFED5 record.

4.1 Uncertainty inherited from GFED5

GFED5 burned area and emissions are subject to uncertainties in satellite-derived burned area and active fire observations, as well as in the subsequent modeling of fuel consumption and emissions, including assumptions regarding combustion completeness and emission factors and the representation of specific fire types such as agricultural burning, peat fires, and deforestation

fires. These uncertainties are inherited by GFED5NRT and are not re-estimated as part of the NRT extension.

The baseline uncertainty of the GFED5 product has been evaluated in our prior assessments against independent observational datasets (Chen et al., 2023; Hall et al., 2024; van der Werf et al., 2025). These evaluations provide an empirical basis for assessing the magnitude and implications of uncertainty in the underlying GFED5 estimates. In contrast, a comprehensive analytical, bottom-up propagation of GFED5 uncertainty is not currently feasible, because it would require well-constrained empirical probability distributions and full error covariance matrices for numerous interdependent satellite, ecological, and physical parameters. Accordingly, we rely on these prior evaluations rather than deriving a new formal uncertainty budget for the baseline GFED5 product.

4.2 Additional uncertainty introduced by GFED5NRT

The principal additional uncertainties in GFED5NRT arise from the use of static scaling factors and land-cover information to translate NRT VIIRS active fire observations into the GFED5 framework. Effective fire area, fuel consumption scalars, and land cover classifications are held fixed during the NRT extension, maintaining consistency with the historical GFED5 record but potentially limiting representation of changes in fire size, fuel availability, fuel moisture, and plant functional type. In addition, active fire observations provide instantaneous rather than cumulative measures of fire activity and are subject to limitations such as cloud and smoke obscuration and variations in detection conditions. These factors introduce uncertainty beyond that inherited from GFED5.

We quantified the additional uncertainty by comparing GFED5NRT with GFED5 over their 2012–2022 overlap period, using region-aggregated monthly time series and grid-cell-level spatial patterns of burned area and emissions (Table 1 and Figure 5). We report the Pearson correlation coefficient (r), linear regression slope (m), normalized root-mean-square error (nRMSE), and Taylor Skill Score (TSS), which combines the correlation and the ratio of standard deviations into a single measure of agreement (Taylor, 2001). Spatial TSS ranged from 0.62 to 0.95 for burned area and from 0.64 to 0.89 for emissions across regions, with global values of 0.94 and 0.88, respectively. Temporal TSS ranged from 0.62 to 0.98 for burned area and from 0.80 to 0.99 for emissions, with global values of 0.88 and 0.97, respectively.

These metrics demonstrate that GFED5NRT reliably reproduces continental-scale spatial gradients, seasonal cycles, interannual variability, and major fire anomalies captured in GFED5. However, regional divergences highlight the limitations of uniform regional scaling factors across landscapes with pronounced sub-regional heterogeneity in fuel types and fire behavior. Consequently, GFED5NRT is more robust when aggregated over broad spatial domains and long timeframes, whereas fine-scale assessments (e.g., grid-cell, national or subnational) carry substantially higher uncertainty, particularly regarding quantitative trends in burned area and emissions. Importantly, these evaluation metrics quantify inter-product consistency and transfer errors rather than formal parametric confidence bounds.

4.3 Independent benchmarking of GFED5NRT

To contextualize GFED5NRT within the broader observational landscape, we conducted systematic multi-dataset benchmarking (Sect. 3.2). Monthly GFED5NRT burned area was evaluated against independent global satellite products, including VIIRS VNP64A1, as well as high-resolution regional inventories (e.g., INPE, MapBiomass, NBAC, MTBS, EFFIS, and ONFIRE). In parallel, daily emissions were compared against operational fire-emission inventories, including FINN, GFAS, and GBBEPx, confirming the spatial and temporal coherence of GFED5NRT relative to existing services.

Future versions of GFED5NRT will target these uncertainties through several methodological refinements, including improved dynamic estimates of effective fire area and fuel consumption, incorporation of updated land-cover information, and refinement of algorithms for constraining fire-emissions estimates using fire radiative power observations. An important next step is establishing an integrated atmospheric validation pipeline to provide objective top-down constraints on fire emissions. Following the architecture of the International Land Model Benchmarking system (ILAMB; Collier et al., 2018), a chemical transport model driven by fire emissions inventories could be used to evaluate simulated trace gas (CO, NO₂) and aerosol fields (aerosol optical depth, surface particulate concentrations) against surface networks and satellite retrievals. Such a benchmarking framework would enable standardized, quantitative assessments of individual inventories and track performance across model iterations.

”

Minor issues:

Line 102-3: “generating two independent time series of burned area and emissions estimates with minimal latency”. Please indicate the value.

We have indicated the latency in Section 2.5 *Generating the GFED5NRT products*:

“Using this automated framework, daily GFED5NRT products are routinely generated with latencies of only a few hours.”

Line 116-18: “To reduce commission errors associated with the coarse spatial resolution of MODIS (500 m), burned area estimates were calibrated using higher-resolution (20-30 m) observations from Landsat and Sentinel-2.” I indicated in the previous version of my review that MODIS BA errors are mostly omission, and this should be the main motivation to use higher resolution BA products, but commission is still emphasized here, why?

While omission errors are indeed a major challenge for coarse-resolution burned area products, commission errors caused by mixed pixels within the 500-m MODIS footprint are also substantial and require explicit correction to ensure accurate burned area estimates.

As described in Chen et al. (2023), our calibration approach addresses both omission and commission errors. Specifically, commission errors occur because the MODIS algorithm marks a 500-m pixel as entirely burned if any portion of it is detected as burned, ignoring unburned patches within that footprint. To fix this, we used high-resolution (20–30 m) Landsat and Sentinel-2 observations to quantify the actual burned fraction within MODIS burned pixels and derived commission scaling coefficients to adjust down the burned area estimates.

Simultaneously, a parallel set of scaling coefficients was derived using MODIS active fire counts and high-resolution sampling outside MODIS pixels to remedy omission errors (i.e., small fires missed due to coarse resolution).

To avoid any ambiguity in the manuscript, we have revised the text to explicitly state that both error types are addressed by the calibration framework:

“To address both commission and omission errors associated with the coarse spatial resolution of MODIS (500 m), burned area estimates were calibrated using higher-resolution (20–30 m) observations from Landsat and Sentinel-2 alongside MODIS active fire data. Specifically, commission errors caused by unburned patches within burned MODIS pixels were corrected by scaling MODIS burned area using coefficients derived from high-resolution observations, while omission errors were corrected by relating MODIS active fire counts to high-resolution burned area sampled outside MODIS burned pixels.”

Lines 231–33: “In the first step, biome-specific scaling factors, termed effective fire area (EFA) scalars, were derived for multiple geographical regions by relating science-quality VIIRS active fire detections to GFED5 burned area data during the MODIS-VIIRS overlap period (2012–2022).” Be more specific with the regions, are those used for the analysis of BA trends? If so, please refer here to the table-figure describing them. It is odd that you used GFED BA as a reference instead of the original HR BA datasets, as this implies increasing the uncertainty.

We added the following sentence to clarify the definition of the GFED regions:

“We used the same 14 continental-scale regions (hereafter GFED regions) defined in the standard GFED5 product (Fig. 2 in van der Werf et al., 2025).”

We selected the GFED5 burned area dataset as a reference because global, full-coverage, high-resolution burned area data were still unavailable at the time of writing.

Line 242–43: In principle, the Regftp scheme could yield up to 224 EFA scalars globally, from the product of 14 continental regions and 16 burning types” in line 179 you mentioned 3 burning types, so I assume you mean here the combination of burning types and land covers. Please clarify.

We clarified this by changing the sentence to:

“In this scheme, active fires are aggregated within each GFED region by land cover or burning type.”

Line 268 “each grid cell are calculated as the arithmetic mean of the SNPP- and NOAA-20-based estimates when both”. Since NPP is close to end its lifetime, I assume you have adapted the system to include NOAA-21 AF. It would be good to mention somewhere, and comment on the potential implications.

We added the following sentence to address the potential adaptation to NOAA-21.

“The pipeline is also adaptable to active fire detections from VIIRS sensors on other platforms such as NOAA-21, pending systematic evaluations of data quality and consistency.”

Line 308: “Australia, we relied on the regional data compiled within the ONFIRE database”, please quote the original paper describing the dataset.

We have added the citation and reference for the original paper (Gincheva et al., 2024).

*Gincheva, A., Pausas, J. G., Edwards, A., Provenzale, A., Cerdà, A., Hanes, C., Royé, D., Chuvieco, E., Mouillot, F., Vissio, G., Rodrigo, J., Bedía, J., Abatzoglou, J. T., Senciales González, J. M., Short, K. C., Baudena, M., Llasat, M. C., Magnani, M., Boer, M. M., González, M. E., Torres-Vázquez, M. Á., Fiorucci, P., Jacklyn, P., Libonati, R., Trigo, R. M., Herrera, S., Jerez, S., Wang, X., and Turco, M.: A monthly gridded burned area database of national wildland fire data, *Sci. Data*, 11, 352, <https://doi.org/10.1038/s41597-024-03141-2>, 2024.*

Line 352. Here we compare the GFED5NRT burned area estimates with VNP64A1. Overall, the spatial distribution of GFED5NRT burned area in 2023 shows moderate to good agreement with the VIIRS-based burned area product”. This statement is vague, and the figure S6 shows strong differences between the two datasets with much higher estimations in GFED5NRT. A more adequate figure should show the differences between them in terms of total BA in each grid cell. Ideally, the authors should show differences with other global BA products, such as those derived from FireCCI (the MRBA60 product covers 2003-2025). Another figure with annual trends in total BA estimates from different products would be useful.

We thank the reviewer for their suggestions. We agree the original phrasing "moderate to good agreement" was ambiguous. We have revised the sentence to make this distinction explicit:

“GFED5NRT and VNP64A1 identify burning in broadly similar regions in 2023, but GFED5NRT detects substantially more burned area overall, including in areas where VNP64A1 shows little or no burning (Fig. S6), consistent with its greater sensitivity to small and low-intensity fires through the active fire-based scaling.”

Regarding the request for a figure showing grid-cell differences in total BA, we respectfully believe this is not necessary, as the discrepancy between the two products is already clearly visible in the side-by-side maps in Fig. S6. We believe an explicit difference panel would convey the same conclusion already apparent from the two panels, and have instead revised the main text to explicitly describe this systematic pattern (see response to the wording of "moderate to good agreement" above).

Regarding comparison with additional global BA products, we opted not to include MRBA60 (Torres-Vázquez et al., 2026) at this stage. Our manuscript already provides extensive cross-validation against ten established datasets (FINN, GBBEPx, GFAS, standard GFED5, VNP64, MTBS, INPE, NBAC, EFFIS, and ONFIRE), and we believe this is sufficient to establish the reliability of GFED5NRT. We note that MRBA60 is still an early-access version of record at the time of writing; we will consider including it in future updates to GFED5NRT documentation as the product matures.