

Responses to RC1

General comments

1. This manuscript presents GlobMap FFP v1.0, a global 30 m dataset of forest burned area patches derived from the Landsat archive for the period 1984–2022. Although the technical effort required to process the full Landsat archive at global scale is substantial, the study is built on a fundamentally flawed premise: Landsat’s temporal resolution is insufficient for reliable global burned area mapping. Burned area is a phenomenon characterized by short-lived spectral signals, often detectable for only a few weeks to a few months, and strongly dependent on observation timing and cloud-free conditions. A sensor with a nominal 16-day revisit cycle, further reduced by cloud cover and data gaps, cannot consistently observe this phenomenon at global scale, particularly in tropical and temperate forests.

The manuscript implicitly acknowledges this limitation by adopting multi-year temporal compositing, typically using five-year windows (and even longer periods prior to 2000). However, this strategy does not solve the underlying problem. Aggregating observations over five-year intervals inevitably suppresses short-lived burn signals and biases detection toward high-severity or slowly recovering fires. As a result, large fractions of real burned area are missed, especially in regions with frequent low-intensity fires or rapid vegetation recovery. An additional unresolved issue concerns areas that burn multiple times within the same compositing window. If I understand the methodology correctly, how are repeated fire occurrences affecting the same pixel over a five-year period handled, and how is fire recurrence represented when only a single observation per pixel is retained in the composite?

Response:

We thank the reviewer for these important comments. We agree that the relatively low temporal sampling frequency of Landsat, together with cloud contamination and uneven observation availability, presents significant challenges for global fire reconstruction, particularly in ecosystems where burn signals are short-lived and vegetation recovery is rapid.

The approximately five-year compositing interval was adopted as a compromise between temporal specificity and observation completeness, rather than to fully eliminate the constraints. Shorter intervals would provide greater temporal precision but would substantially reduce observation availability in many parts of the globe, leading to fragmented spatial coverage and reduced consistency across the time series. Longer intervals would increase observation availability but increase temporal aggregation and the likelihood of merging distinct fire events. The selected interval was therefore designed to balance these competing considerations while maintaining computational feasibility at the global scale.

Regarding repeated burning, as only one observation is retained for each pixel within a compositing interval, multiple fire occurrences affecting the same location during the same interval are be independently reconstructed. In practice, the composite preserves the observation associated with the strongest retained burn signal, and a single burned year is assigned to that pixel. Consequently, fire recurrence within a compositing interval is not explicitly represented. We have clarified this limitation in **Lines 509 – 513 of the Discussion**.

We acknowledge that multi-year compositing inevitably suppresses short-lived burn signals. This is an inherent trade-off when using medium-resolution optical sensors to achieve global, gap-reduced time series at decadal scales. We do not consider this a methodological flaw of compositing, but rather a constraint that defines the appropriate interpretation of the dataset, which we have clarified as a spatially explicit fire patch product rather than a complete burned area inventory.

To facilitate appropriate interpretation, we emphasize throughout the revised manuscript that GlobMap FFP should be regarded as a spatially explicit fire patch dataset rather than a complete inventory of all burned area occurrences. Its primary value lies in preserving patch geometry, spatial organization, and long-term fire patch dynamics at 30 m resolution across nearly four decades, while the limitations associated with temporal aggregation and repeated burning are now more explicitly acknowledged.

2. A critical consequence of this approach is evident in the reported burned area estimates. The dataset reports mean global forest burned area values of approximately 7.3 Mha yr^{-1} for the period 2001–2021, whereas established global burned area products report values close to 19–20 Mha yr^{-1} over the same period. This represents an underestimation by a factor of roughly three.

This strong underestimation is particularly pronounced in tropical regions, where the manuscript itself reports the largest divergences relative to existing products. These are precisely the regions where burned signals are short-lived, cloud cover is persistent, and Landsat’s sparse clear-sky observations are least capable of capturing fire effects. Therefore, I find it very difficult to reconcile the large disparity in burned area between the GlobMap product and MODIS with the validation metrics reported in the manuscript. According to Table 2, omission errors (24%) and commission errors (13%) are relatively low, the Dice coefficient is high (0.82), and the reported relative bias is relatively small (−11%). These values would normally indicate a product with only moderate underestimation. However, the intercomparison shows that GlobMap detects substantially less burned area (Fig. 7; MODIS $\approx 19.6 \text{ Mha yr}^{-1}$ versus GlobMap $\approx 7.3 \text{ Mha yr}^{-1}$). If such a large and systematic discrepancy truly exists, it is difficult to explain how it could coexist with a low relative bias and only moderate omission errors. This inconsistency is not resolved in the manuscript and fundamentally undermines the

credibility of the validation results.

Response:

We thank the reviewer for this insightful comment. We agree that the apparent inconsistency between the relatively good agreement metrics reported in Table 2 and the substantially lower burned area estimates relative to MCD64A1 requires clarification.

First, we acknowledge that the omission errors, commission errors, Dice coefficients, and relative bias reported in Table 2 should not be interpreted as measures of burned area completeness. These metrics were derived from sampled Landsat scenes and quantify agreement only for burned scars remained detectable in the available Landsat observations. Consequently, they characterize classification agreement conditional on burn detectability rather than the completeness of burned area reconstruction. Fires missed because of limited observation availability, persistent cloud cover, rapid post-fire vegetation recovery, or compositing effects are not represented in these statistics. Therefore, relatively good agreement within the evaluated samples does not necessarily imply close correspondence in cumulative burned area estimates when products are aggregated across regions and decades. The relative bias reported in Table 2 reflects agreement within the sampled locations and is not directly comparable to differences in total burned area between GlobMap FFP and MCD64A1. To clarify, we have added the following sentences in **Lines 260 – 262, Lines 321 – 324**:

“Because both datasets were derived from Landsat imagery and relied on related burned area detection procedures, these metrics should be interpreted as measures of internal consistency rather than fully independent estimates of product accuracy.”

“Yet, it is worth noting that the reported omission and commission errors characterize agreement between GlobMap FFP and the Landsat-based reference samples within evaluated locations where burned scars remained detectable in the available observations. These metrics quantify classification agreement conditional on burn detectability rather than the completeness of burned area reconstruction at regional or global scales.”

Second, we agree that the lower burned area estimates of GlobMap FFP relative to MCD64A1 warrant further discussion. In the original manuscript, we primarily attributed these differences to the coarser spatial resolution of MODIS. We acknowledge that this explanation alone is insufficient. In the revised manuscript, we now emphasize that the discrepancy likely reflects the combined effects of multiple factors, including incomplete detection of short-lived fire signals under heterogeneous Landsat observation conditions, particularly in cloud-prone and rapidly recovering tropical forests; differences in image compositing and fire patch delineation procedures among products; and limited representation of repeated burning within compositing intervals. We have expanded the **Discussion (Lines 501–514)** to explicitly address these

sources of discrepancy and no longer attribute the differences solely to spatial-resolution effects.

Finally, we have revised the positioning of both the manuscript and the product to avoid misunderstanding. The revised manuscript no longer presents GlobMap FFP as a complete global burned area inventory. Instead, the primary objective of the dataset is to provide a consistent, spatially explicit characterization of forest fire patches at 30 m resolution over nearly four decades. Its principal value lies in enabling analyses of fire patch geometry, spatial organization, patch size distributions, and long-term changes in fire patch structure, while the limitations and uncertainty sources associated with burned area completeness are now discussed more explicitly throughout the manuscript.

3. The validation framework itself further limits the interpretability of the reported accuracy metrics. Although the authors state that they follow the spatial sampling framework of the Burned Area Reference Database (BARD) by selecting Landsat TSAs, they do not use the BARD reference perimeters. Instead, reference burned area within each TSA is generated independently using the same algorithmic approach employed for training sample generation and product development. As a result, the validation does not rely on independent reference data, but rather evaluates internal methodological consistency.

Moreover, while BARD TSAs are part of a global, stratified sampling design explicitly constructed to support statistically rigorous accuracy assessment and uncertainty estimation, this sampling design is not adopted in the present study. The authors do not implement a probability-based sampling scheme adapted to their analysis, nor do they report uncertainty measures (e.g. confidence intervals or standard errors) for the accuracy metrics. Consequently, the reported omission and commission errors, Dice coefficients, and relative bias cannot be interpreted as statistically robust estimates of real-world burned area detection performance.

In addition, the validation methodology lacks essential temporal detail. The manuscript does not clearly specify which years or portions of the 1984–2022 period are actually covered by the validation, nor whether the reported accuracy metrics are representative of the entire time series. It remains unclear whether the validation is dominated by periods with higher observation density (e.g. the Landsat 7 and Landsat 8 eras) or whether earlier periods, characterized by sparser data availability and longer compositing windows, are adequately represented. Furthermore, the additional filtering of validation imagery to scenes with less than 40% cloud cover further reduces the number of usable observations, particularly in cloud-prone regions, compounding uncertainties regarding the representativeness and robustness of the reported metrics.

Response:

We thank the reviewer for this thoughtful and detailed comment. We agree that the assessment

does not constitute a fully independent validation. Although the reference dataset was generated separately from the final product and based on independently selected sample locations, both datasets were derived from the Landsat archive and relied on related burned area interpretation procedures. Thus, they share common sources of uncertainty, and the reported metrics should not be interpreted as fully independent estimates of burned area detection accuracy. To avoid overstating the level of independence, we have revised the terminology throughout the manuscript and now describe the assessment as a Landsat-based evaluation of product agreement and internal consistency rather than an independent validation.

We also agree that the use of BARD TSA locations should not be interpreted as adoption of the full BARD validation framework. Here we used the TSA locations to achieve broad geographic coverage across major forest regions. However, we did not use the BARD reference perimeters themselves, or adopt the probability-based sampling design developed for statistically rigorous global accuracy estimation, so that formal uncertainty estimation was not implemented. To avoid misunderstanding, we have clarified this in **Lines 248 – 249** of the **Methods** as below:

“To evaluate the performance of GlobMap FFP, we constructed a Landsat-derived reference dataset using the TSA units selected to achieve broad geographic coverage following the spatial distribution of the burned area reference database (BARD) (Franquesa et al., 2020).”

Regarding the temporal representativeness of the evaluation dataset, the sampled scenes were selected across the full study period (1984-2022) and include observations from the Landsat 5, Landsat 7, and Landsat 8 eras. Therefore, the assessment is not restricted to periods with higher observation density. Nevertheless, we acknowledge that the temporal distribution of available Landsat observations is inherently uneven. Earlier portions of the Landsat archive generally contain fewer cloud-free observations and are associated with greater uncertainty than more recent periods. The requirement for cloud cover below 40% was introduced to improve the interpretability and consistency of the reference data, but may reduce the representation of highly cloud-prone regions in the evaluation dataset. We have clarified the temporal coverage of the reference samples in the **Methods (Lines 252 – 255)** as below:

“In total, we sampled 74 Landsat TSA units across major forest biomes (Fig. 3). Then Landsat scenes with cloud cover below 40% in these units were sampled, spanning the full study period and encompassing observations from the Landsat 5, 7, and 8 archives, resulting in a total of 945 Landsat scenes. The cloud cover threshold was introduced to improve the interpretability and consistency of the reference data, but may reduce the representation of highly cloud-prone regions in the evaluation dataset.”

4. An additional conceptual limitation concerns the interpretation of the mapped units as 'individual' fire events. In the proposed dataset, burned pixels are aggregated into patches based

on spatial proximity and assignment to a single burned year, without explicit information on ignition timing, fire duration, or intra-annual separation. Under this framework, independent fires occurring at different moments within the same year may be merged into a single patch, while the same fire spreading over extended periods may be inconsistently represented depending on observation availability. As a result, the mapped patches cannot be unambiguously interpreted as fire events. This has direct implications for the analysis of fire size distributions, fire frequency, and fire regime characteristics, and further limits the comparability of the dataset with products that explicitly track fire events using finer temporal information. The manuscript does not sufficiently clarify these limitations or their consequences for downstream analyses.

Response:

We agree that the mapped units should not be interpreted as individual fire events in the strict sense. Because fire patches are delineated using annual temporal attribution and spatial clustering, temporally distinct fires occurring within the same year may be merged into a single patch, while detailed information on ignition timing and fire progression is not retained. To avoid ambiguity, we have modified the writing throughout the manuscript to consistently describe the product as a forest fire patch dataset rather than an event-based fire inventory. We have also expanded the discussion of the implications of annual aggregation for fire frequency estimates, fire patch size, and downstream analyses as below in **Lines 513 – 516**:

“Furthermore, the annual aggregation may merge temporally adjacent fires into a single fire patch, likely resulting in overestimated patch sizes and underestimated fire frequencies. Consequently, GlobMap FFP should be interpreted as a spatially explicit fire patch dataset rather than a complete inventory of all forest burned area, particularly in frequently burned tropical forests.”

5. Overall, while Landsat-based burned area mapping can be highly effective at regional scales under appropriate conditions, this manuscript does not demonstrate that such approaches can be straightforwardly generalized to a global product without substantial loss of information. The strong underestimation of burned area, the internal inconsistency between area estimates and validation metrics, the lack of independent validation, and the inability to represent repeated burning indicate that the proposed dataset does not provide a reliable or improved representation of global forest burned area. In this context, it remains unclear what scientific or practical value a new dataset based on higher spatial resolution sensors offers if it does not demonstrably improve the representation of burned area relative to existing products.

Response:

We thank the reviewer for this important comment. We agree that, if GlobMap FFP were

interpreted as a complete global burned area inventory, the limitations identified by the reviewer, including burned area underestimation, uncertainties in the evaluation framework, and the inability to fully represent repeated burning, would substantially constrain its utility. We also agree that the original manuscript did not sufficiently articulate the scientific objectives and intended applications of the dataset.

In response, we have revised the positioning of the product throughout the manuscript. The revised version no longer presents GlobMap FFP as a complete reconstruction of global forest burned area or as a replacement for existing burned area products. Instead, the primary objective of GlobMap FFP is to provide a consistent, spatially explicit characterization of forest fire patches at 30 m resolution across nearly four decades of Landsat observations.

The principal value of the dataset lies in its ability to preserve fine-scale fire patch geometry and spatial organization. Many existing global burned area products were developed primarily to quantify burned area extent and temporal dynamics. While highly valuable for monitoring burned area, their spatial resolution limits the representation of fire patch boundaries, internal heterogeneity, patch connectivity, and landscape-scale fire patterns. By contrast, GlobMap FFP was specifically designed to retain individual fire patch structure and associated attributes, enabling analyses of fire patch morphology, patch size distributions, spatial aggregation, landscape fragmentation, and long-term changes in fire regime characteristics across global forests.

We therefore emphasize that the scientific contribution of GlobMap FFP is not solely determined by its ability to reproduce burned area totals reported by existing products. Rather, it provides a complementary perspective on forest fire dynamics by explicitly representing the spatial structure of fire patches over a nearly four-decade period. This type of information is difficult to derive consistently from coarser resolution global products and is particularly relevant for studies of fire regime shifts, landscape ecological impacts of fire, spatial fire patterns, and fire patch scaling relationships.

To better reflect these objectives, we revised the **Introduction, Discussion, and Conclusions** sections to more clearly distinguish between burned area completeness and fire patch characterization. We also expanded the discussion of limitations and uncertainty sources associated with burned area reconstruction, while highlighting the specific applications for which the dataset was designed. We hope these revisions clarify the intended scope, limitations, and scientific value of GlobMap FFP.

Responses to RC2

General comments

1. The study entitled “A Global 30 m Landsat-based Dataset of Forest Fire Patches (GlobMap FFP v1.0) from 1984 to 2022” by He et al. reports a global forest fire area/patch dataset at 30 m for nearly four decades since 1984. By leveraging the computation power of GEE platform, the authors applied a BVI approach to reduce Landsat noises due clouds and shadows to generate pixel-level image composites, and used ANN models to map forest burned areas, which were further clustered into fire patches (or individual fire events), producing the final GlobMap product. The product was evaluated using independently derived burned area (BA) and by comparing with existing Landsat- and MODIS-based BA datasets.

A consistent, high-quality global fire area dataset at fine resolution would be very meaning for science communities and various applications. Aiming at this goal, there is no doubt that substantial efforts were put by the authors to produce a 30 m global fire area dataset across nearly four decades. As a data science paper, scientifically rigorous methods and high-quality data are critical. These two aspects are my main concerns for this manuscript.

Response:

We thank the reviewer for highlighting these important concerns. We appreciate the reviewer’s recognition of the substantial effort required to generate a global multi-decadal Landsat-derived forest fire patch dataset. We agree that both methodological rigor and transparent characterization of dataset quality are critical for a global fire dataset.

In revising the manuscript, we substantially clarified the design rationale, methodological assumptions, and limitations of the product. We added a new subsection describing the conceptual framework and product design rationale, expanded the discussion of Landsat observational constraints, clarified the rationale behind key methodological choices (including image compositing and fire patch delineation), and strengthened the discussion of uncertainty sources and dataset limitations.

We also revised the overall framing of the manuscript. Rather than presenting GlobMap FFP as a complete global burned area inventory, we now emphasize its primary objective as providing a long-term, spatially explicit characterization of forest fire patches under heterogeneous Landsat observation conditions. We believe these revisions improve the transparency of the dataset, clarify its intended applications, and facilitate more appropriate interpretation of its strengths and limitations.

2. On one hand, the INTRO has done a good job describing the challenges to map global fires using Landsat data, yet the Methods part is mostly descriptive and fails to provide solid,

quantitative evidence to demonstrate its superior performance in handling these challenges. The flow chart does help the reader to understand what steps have been taken to process the data. However, I don't see much demonstration and evidence to show any preferable advantages of the adopted approaches relative to existing approaches in literature.

Response:

We agree that the original manuscript focused primarily on describing the workflow and processing steps, while providing insufficient explanation of the rationale underlying several key methodological choices. We also recognize that the manuscript could be interpreted as implying superior performance relative to existing approaches without presenting direct quantitative comparisons to support such a conclusion. Our intention, however, was not to demonstrate that the adopted workflow universally outperforms existing burned area mapping approaches. Rather, the objective of GlobMap FFP was to develop a globally scalable framework capable of reconstructing forest fire patches from the heterogeneous Landsat archive while balancing several competing constraints, including observation availability, cloud contamination, temporal representativeness, preservation of spatial detail, and computational feasibility.

To address this concern, we substantially revised both the **Introduction** and **Methods** sections to more clearly explain the motivations and trade-offs associated with the adopted methodological choices. Specifically, we now emphasize that the development of GlobMap FFP was guided by the practical challenges of global multi-decadal Landsat processing rather than by the objective of maximizing burned area detection completeness under all conditions.

In the revised manuscript, we clarify the rationale for several key methodological decisions. For example, the use of multi-year compositing is now explicitly described as a compromise between observation availability and temporal specificity, particularly in cloud-prone regions and during the early Landsat era when clear-sky observations were sparse (in **Lines 172-175**). Similarly, the minimum-BVI compositing strategy is motivated by the need to preserve burn-related spectral signals over extended compositing intervals while reducing the influence of cloud, shadow, and atmospheric contamination (in **Lines 177-188**). We also clarify that annual-scale fire patch reconstruction was adopted because globally consistent sub-annual fire chronology is often difficult to recover from heterogeneous Landsat observations (in **Lines 224-232**). These methodological choices are therefore presented as practical solutions to known limitations of the Landsat archive rather than as universally superior alternatives to existing approaches.

We further revised the **Introduction** to better frame the methodological context. The revised text now explicitly discusses the trade-offs among disturbance sensitivity, contamination robustness, temporal representativeness, and computational efficiency associated with

different Landsat-based burned area mapping strategies (in **Lines 52-84**). We emphasize that globally scalable fire reconstruction requires balancing these competing objectives and that no single approach is optimal under all environmental and observation conditions.

Additionally, we expanded the **Discussion** section to acknowledge the limitations associated with the adopted workflow (in **Lines 501-534**), such as potential under-detection of short-lived fire signals, uncertainties related to multi-year compositing, and possible differences relative to alternative approaches. We believe these revisions provide a more balanced presentation of the methodology and clarify the practical considerations that motivated the design of GlobMap FFP.

3. More importantly, the underestimation of the proposed Globmap is surprisingly high. First, I would not call the comparisons of 30m Globmap with other 30m landsat-based BA datasets and even coarser-resolution MODIS BA a rigorous “validation” practice because higher spatial-resolution datasets are usually required as a reference for the validation purpose. Second, the relatively low omission errors (OE) and commission errors (CE) in all five forest types (Table 2) and the substantial BA underestimation compared with much coarser-resolution MODIS BA (MCD64A1; Figs 6&7) are striking. The validation of the MODIS MCD64A1 BA product using 30m Landsat-based BA by the science team has shown that MCD64A1’s OE is higher than 90% in temperate and tropical forests (Boschetti et al., 2019). In other words, MCD64A1 misses majority of fire areas in these two types of forests. Nevertheless, the 30m GlobMap BA is substantially lower than MCD64A1 in most global forests, except boreal North America (Figs 6&7), although this manuscript reports low OE and CE (Table 2). The authors attribute this underestimation in GlobMap (relative to MODIS BA) to MODIS’ coarser footprint size (Fig.12). As other regional Landsat-based BA datasets have shown much higher BA estimates than MCD64A1, at least in U.S. (e.g., Hawbaker et al., 2020), I would not expect that the finer footprint size in Landsat relative to MODIS would be able to account for the substantial underestimation in the GlobMap. This eventually casts great doubts on the robustness of the proposed BA mapping methods and the quality of the final GlobMap BA dataset.

Response:

We thank the reviewer for this insightful comment. We agree that comparisons with existing burned area products cannot substitute for independent validation based on higher-resolution reference data. We also agree that the apparent inconsistency between the relatively low omission and commission errors reported in Table 2 and the substantially lower burned area estimates relative to MCD64A1 requires careful explanation.

First, we acknowledge that the evaluation presented in the original manuscript could be interpreted as a conventional accuracy assessment. In the revised manuscript, we have

clarified that the reported metrics represent agreement between GlobMap FFP and separately generated Landsat-based reference samples rather than independent estimates of absolute burned area detection accuracy in **Lines 260 – 262, Lines 321 – 324** as below:

“Because both datasets were derived from Landsat imagery and relied on related burned area detection procedures, these metrics should be interpreted as measures of internal consistency rather than fully independent estimates of product accuracy.”

“Yet, it is worth noting that the reported omission and commission errors characterize agreement between GlobMap FFP and the Landsat-based reference samples within evaluated locations where burned scars remained detectable in the available observations. These metrics quantify classification agreement conditional on burn detectability rather than the completeness of burned area reconstruction at regional or global scales.”

We have also clarified the distinction between Landsat-based evaluation using reference samples and intercomparison with existing burned area products, and revised the terminology throughout the manuscript accordingly.

Second, we have clarified that the omission and commission errors reported in Table 2 should not be interpreted as a measure of global burned area completeness. These metrics are derived from sampled Landsat scenes and quantify agreement only within locations where burned scars remained detectable in the available Landsat observations. They therefore characterize classification agreement under Landsat observation conditions rather than the completeness of burned area reconstruction at regional or global scales. Consequently, relatively good agreement within sampled locations does not necessarily imply close correspondence in cumulative burned area estimates when products are aggregated across regions and decades. We have revised the manuscript to explicitly discuss this distinction and to avoid interpreting the reported metrics as independent estimates of global product accuracy in **Lines 324 – 327** of the manuscript:

“Fires that were not observable because of limited observation availability, persistent cloud cover, rapid post-fire vegetation recovery, or compositing effects are not represented in these statistics. Therefore, relatively low omission and commission errors within the evaluated samples do not necessarily imply complete recovery of burned area when estimates are aggregated across regions and decades.”

Third, we agree that the lower burned area estimates of GlobMap FFP relative to MCD64A1 warrant further discussion. In the original manuscript, we primarily attributed these differences to the coarser spatial resolution of MODIS. We acknowledge that this explanation alone is insufficient. In the revised manuscript, we now emphasize that the discrepancy likely reflects the combined effects of multiple factors, including incomplete detection of short-lived

fire signals under heterogeneous Landsat observation conditions, particularly in cloud-prone and rapidly recovering ecosystems; differences in image compositing and fire patch delineation procedures among products; and limited representation of repeated burning within compositing intervals. We therefore no longer attribute the observed discrepancies solely to spatial-resolution effects. We have added the following paragraph in the Discussion in **Lines 501 – 518**:

“Several factors contribute to the differences between GlobMap FFP and existing burned area products, particularly MCD64A1. Spatial discrepancies partly reflect differences in sensor resolution. The coarser MODIS pixels tend to produce larger and more spatially continuous burned perimeters, whereas Landsat’s 30 m observations better preserve small fires, patch boundaries, and within-fire heterogeneity (Robinson, 1991). Temporal discrepancies are mostly evident in tropical forests, where persistent cloud cover and rapid vegetation regrowth can cause burned signals to be missed by the relatively infrequent cloud-free Landsat observations. Since Landsat preferentially preserves persistent burn signals, this product is less effective at capturing short-lived fire effects, particularly in ecosystems characterized by rapid vegetation recovery or frequent low-severity burning, compared to the near-daily MODIS observations. Additional differences arise from the methodological choices adopted to achieve globally consistent fire patch reconstruction. The multi-year compositing strategy reduces the influence of cloud contamination, data gaps, and uneven observation availability while maintaining computational feasibility. Yet, because only a single observation is retained for each pixel within a compositing interval, repeated burning occurring at the same location during the same interval are not explicitly reconstructed. Fire recurrence may be underrepresented in frequently burned regions, and fire occurrence frequencies derived from the dataset should be interpreted with caution. Furthermore, the annual aggregation may merge temporally adjacent fires into a single fire patch, likely resulting in overestimated patch sizes and underestimated fire frequencies. Consequently, GlobMap FFP should be interpreted as a spatially explicit fire patch dataset rather than a complete inventory of all forest burned area, particularly in frequently burned tropical forests. Future work should evaluate how different compositing strategies influence fire patch reconstruction across contrasting fire regimes and compare the performance of BVI-based and NBR-based approaches under varying environmental conditions.”

We also agree that the reported agreement metrics should not be interpreted as evidence that GlobMap FFP provides a complete reconstruction of global forest burned area. Rather, they indicate the degree of agreement between the mapped fire patches and independently generated Landsat-based reference samples within the evaluated locations. The revised manuscript explicitly acknowledges that substantial underestimation of burned area may occur in some regions where observation availability is limited and post-fire recovery is rapid

in Lines 520 – 527:

“First, the irregular availability of cloud- and snow-free Landsat observations constrains the temporal precision and completeness of burned area detection, particularly in moist tropical forests where short-lived fires may disappear before the next clear-sky acquisition. Frequent cloud cover in these regions reduces the effective temporal sampling frequency and limits the ability to identify burning dates accurately. Fast-recovering surface fires may also be missed if post-fire spectral signals disappear before the next cloud-free overpass (Hislop et al., 2018). Second, incomplete spatial coverage prior to the 2000s may contribute to regional underestimation of burned area. In parts of western and central Africa and boreal Eurasia, Landsat acquisitions prior to the 2000s were sparse because of historical limitations in data storage, ground-station reception, and data archiving (Feng and Wang, 2024).”

To facilitate appropriate interpretation of the dataset, we have revised the positioning of both the manuscript and the product. The revised manuscript no longer presents GlobMap FFP as a complete global burned area inventory. Instead, the primary objective of the dataset is to provide a consistent, spatially explicit characterization of forest fire patches at 30 m resolution over nearly four decades. Its principal value lies in enabling analyses of fire patch geometry, spatial organization, patch-size distributions, and long-term changes in forest fire patch structure, while the limitations and uncertainty sources associated with burned area completeness are now discussed more explicitly throughout the manuscript.

References

- Boschetti, L., Roy, D. P., Giglio, L., Huang, H., Zubkova, M., & Humber, M. L. (2019). Global validation of the collection 6 MODIS burned area product. *Remote sensing of environment*, 235, 111490.
- Hawbaker, T. J., Vanderhoof, M. K., Schmidt, G. L., Beal, Y. J., Picotte, J. J., Takacs, J. D., ... & Dwyer, J. L. (2020). The Landsat Burned Area algorithm and products for the conterminous United States. *Remote Sensing of Environment*, 244, 111801.

Responses to CC1

General comments

1. This paper describes a data product for patches of fire occurrence across the world from 1984 to near present using Landsat images. It's an ambitious effort given the challenges of data volume and the heterogeneity in challenges that must be overcome to map fires across the entire earth. After some thought, and after reading the first reviewer's comments, I am not as down on the data product as reviewer number 1, but I do think it is perhaps better characterized as a beta version 0.5, rather than a full-on 1.0. I'll try not to be redundant with the first reviewer here, although I agree with much of what they say.

Responses:

We sincerely thank the reviewer for the thoughtful and constructive comments, which helped us substantially improve both the framing and transparency of the manuscript. We appreciate the reviewer's recognition of both the ambition and the technical challenges associated with generating a global multi-decadal Landsat-based fire patch dataset.

We agree that the original manuscript did not sufficiently clarify the scope, intended applications, and limitations of the current product, which may have led to an overly strong impression regarding its maturity and completeness. In the revised manuscript, we have substantially revised the positioning and presentation of the dataset to better reflect its current strengths and limitations.

Specifically, we now emphasize that the product is designed primarily as a spatially explicit characterization of forest fire patch patterns, rather than as a complete global burned area inventory or a fully resolved fire event reconstruction product. We have also expanded the discussion of methodological limitations, uncertainty sources, forest mask assumptions, temporal aggregation choices, and data usability considerations.

In addition, throughout the manuscript, we have clarified the intended application scenarios of the dataset, particularly for regional ecological analyses, fire patch morphology studies, and landscape-scale spatial fire pattern characterization in regions where Landsat observations provide reliable coverage.

2. Reviewer 1 rightly pointed out the lack of agreement in tropical regions as their reasoning behind their opinion that Landsat is simply not an appropriate product for mapping burned area globally. I'm not sure I would go that far. The thing that actually gave me both hope and concern was lack of agreement between the GlobMap FFP and the BAECV in the US. From my own

experience I've found the BEACV to be quite good, at least in the study areas in which I have used it. So why is this product not matching the BAECV? To me, if this algorithm were calibrated so that it can at least match the performance of well-validated reference data (BAECV or even to something human-verified like MTBS perimeters in the US), maybe we could at least feel confident that this product is capturing what Landsat can legitimately detect in the tropical areas where there is generally much less human-verified reference data. So I guess my takeaway is not that Landsat cannot capture burned area well in the tropics, rather that this particular product is doing a bad job in most places, so maybe a better product could.

Responses:

We thank the reviewer for this insightful comment, which highlights important differences in agreement among burned area products across well-studied regions.

We agree that discrepancies observed in the United States are particularly informative because several well-established reference datasets are available in this region. We also agree that evaluating agreement against products such as BAECV and MTBS provides valuable context for understanding the capabilities and limitations of Landsat-based fire mapping approaches.

Our additional examination suggests that differences between GlobMap FFP, BAECV, and MTBS should not necessarily be interpreted as evidence that one product is correct and another is incorrect. Rather, they largely reflect differences in mapping objectives, temporal attribution strategies, and operational definitions of burned area. Specifically, BAECV is based on pixel-level spectral change detection and may exhibit sensitivity to persistent post-fire spectral anomalies following high-severity fires. In such cases, persistent post-fire spectral signals may be represented differently depending on the temporal attribution and compositing strategy adopted by each product, contributing to differences in mapped burned area (**Figure R1 and Figure R2**). Meanwhile, MTBS provides fire perimeters derived from manual interpretation of Landsat imagery and auxiliary information. While MTBS is widely used as a reference dataset, its polygon-based representation necessarily generalizes fine-scale heterogeneity within fire perimeters, potentially including unburned or partially burned regions within mapped boundaries (**Figure R3**). This characteristic may lead to systematic differences in estimated burned area compared to pixel-based products, particularly in heterogeneous fire events.

Our product, in comparison, is designed to provide a globally consistent characterization of forest fire patches based on Landsat spectral trajectories, with annual temporal attribution intended to reduce repeated representation of persistent post-fire signals across multiple years. lies in its spatially explicit representation of long-term fire patch patterns and geometry, rather than in reproducing any single existing burned area product. We have revised the manuscript to clarify these conceptual differences and to emphasize that agreement with reference datasets

should be interpreted in the context of their respective mapping definitions and methodological assumptions.

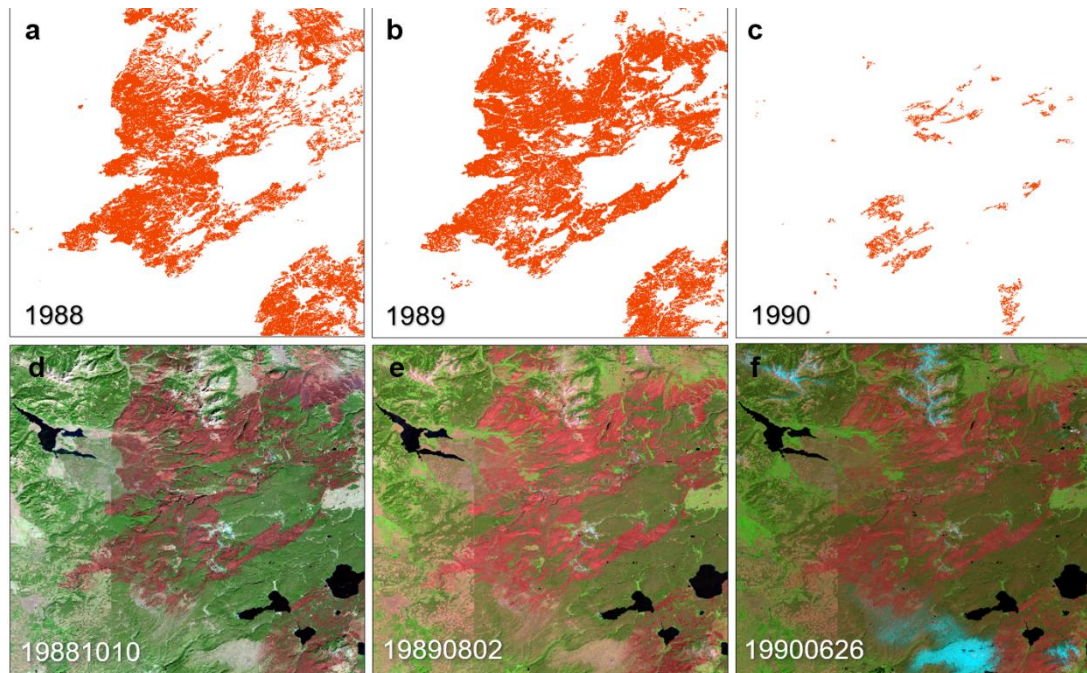


Figure R1. BAECV example of the 1988 fire at the Yellowstone National Park. a-c show the detected burned area in BAECV in 1988, 1989 (one year after burning), and 1990 (two years after burning), respectively. d-f show the false-color composites of Landsat imagery in the corresponding years.

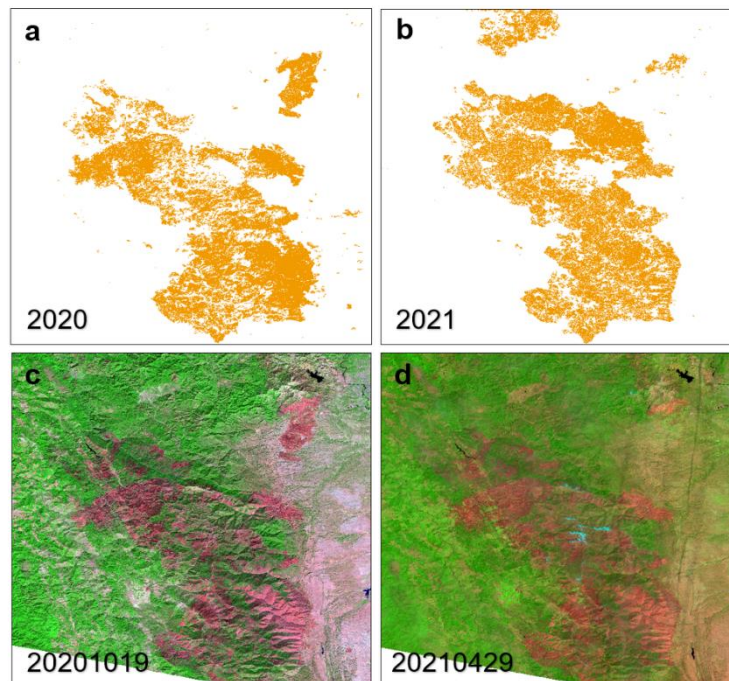


Figure R2. BAECV example of the 2020 fire in Northern California. a and b show the

detected burned area in BAECV in 2020 and 2021 (one year after burning), respectively. **c** and **d** show the false-color composites of Landsat imagery in the corresponding years.

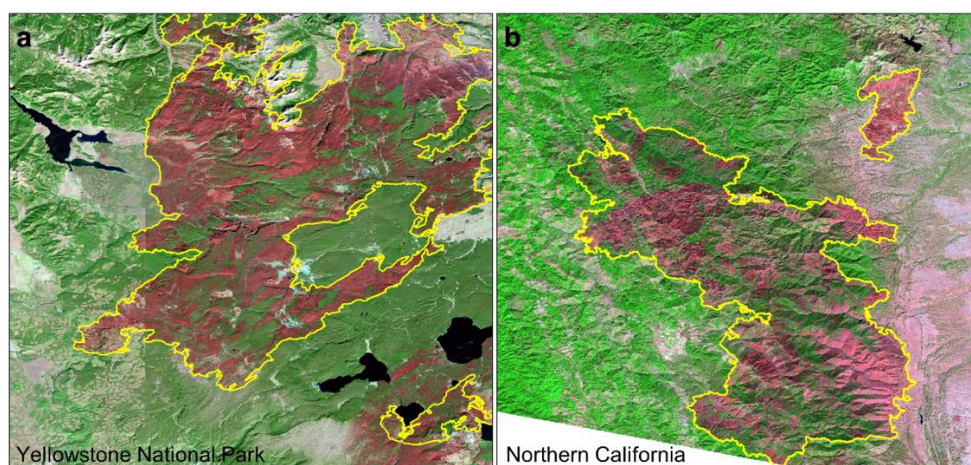


Figure R3. MTBS examples of the (a) 1988 fires in Yellowstone National Park and the (b) 2020 fires in Northern California.

3. My other major concern is the packaging of the data product. It is provided as each individual Landsat tile zipped up as a geotiff (666 zipped files for each 5 year time period), which has a fairly high barrier to convenient use. it would make a lot of sense to polygonize the individual fire events and provide regional shapefiles they way other products do. The rasters do not provide heterogeneous values within each event (for example, burn severity), so having each pixel is of limited utility compared to polygons with scar ID and date as attributes. One first step might be to polygonize the whole time series for each landsat tile and go from there. Even if there are concerns with the errors of omission, if it were actually convenient to use, it might still be useful to many people. Especially in ecological studies at regional scales in places where the algorithm did a good job.

Response:

We thank the reviewer for this valuable suggestion regarding the usability and packaging of the dataset. We agree that vectorized fire patch products can substantially improve convenience for ecological and regional-scale applications, particularly for users interested in fire geometry and patch-level analyses.

In the current version, we chose to distribute the dataset in raster format primarily to preserve the original 30 m spatial detail and within-patch heterogeneity represented in the Landsat-derived fire maps. Raster representation also avoids potential geometric simplification introduced during large-scale polygonization and provides greater flexibility for users

interested in pixel-level spatial analyses and customized post-processing workflows. In addition, polygonization at the global scale over the full multi-decadal archive would generate extremely large and computationally complex vector datasets, particularly for fragmented and spatially heterogeneous fire patches. Therefore, the current release prioritizes scalable raster-based storage and analysis.

We acknowledge, however, that vectorized products improve accessibility and usability for many regional ecological applications. To facilitate this, the dataset includes unique fire patch identifiers that allow users to perform customized polygonization according to their specific study objectives and regional requirements.

We have clarified this design rationale and the potential for future vectorized products in **Lines 550 – 552** of the **Data availability** as below:

“The dataset is distributed in raster format to preserve pixel-level spatial heterogeneity; future versions may explore polygon-based representations for improved usability.”

We have also modified the related sentences in **Lines 93 – 96** of the **Introduction** as:

“The final product is distributed in raster format, with each fire patch assigned a unique identifier together with associated attributes including burned year and quality assurance (QA) information. This raster-based representation preserves fine-scale spatial heterogeneity while supporting analyses of fire patch morphology, spatial organization, and long-term forest fire dynamics.”

Specific comments:

1. 97, Figure 5: The 25% cutoff for forest cover might be a little high. If I am interpreting figure 5 correctly, this cutoff value is excluding essentially all forests in the western US, for example. I think there needs to be a figure right up front that shows the land area that meets the filtering criteria and is considered 'forested'. In Figure 5b, for example, it is not clear whether 0 fire occurrences is a different color than unanalyzed non-forested regions. I'd suggest grey for 0 fires and white for unanalyzed.

Response:

We thank the reviewer for this helpful suggestion. We agree that the forest-cover threshold and the visualization of analyzed versus non-analyzed regions should be clarified more explicitly in the manuscript.

The 25% tree-cover threshold was adopted as a commonly adopted operational definition (Hansen et al., 2013; Harris et al., 2012; Myroniuk et al., 2020) to provide a globally consistent

forest mask and to reduce contamination from sparsely vegetated or non-forest fire-prone areas. However, we acknowledge that this threshold may exclude some open-canopy forest and woodland systems, particularly in regions such as the western United States.

Following the reviewer's suggestion, we have added a new map in **Figure 5** (showing as **Figure R4** below) showing the global spatial extent of regions that meet the forest filtering criteria and were included in the analysis to improve clarity. We have also revised **Figure 5** to clearly distinguish between areas with 0 detected fire occurrences and non-analyzed non-forest regions. non-analyzed regions are now shown in white, whereas analyzed regions with zero fire occurrences are displayed in grey.

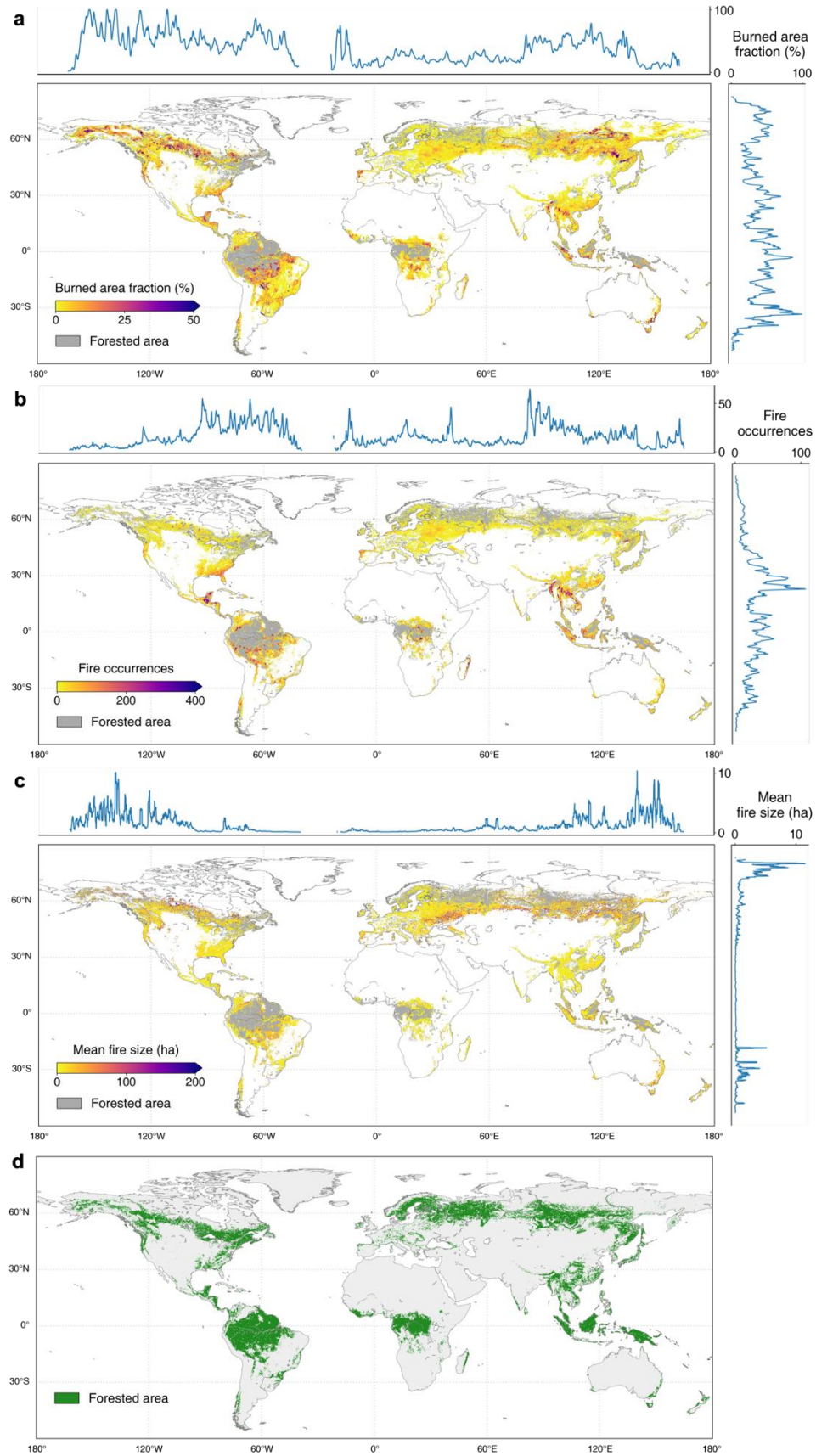


Figure R4. Spatial distributions of total global forest fires over 1984 – 2022 based on the

dataset developed in this study. (a) Burned area fraction (%), (b) fire occurrences, (c) mean fire size (ha). (d) shows the identified forested area. The variables were aggregated at a 0.1° resolution for data visualization.

We have also clarified the forest-mask definition and its implications in **Lines 117 – 120** of the **Datasets**:

“This operational threshold is commonly used in global forest mapping studies to distinguish forest from non-forest land cover types on satellite data (Myroniuk et al., 2020; Hansen et al., 2013; Harris et al., 2012). While it improves global consistency in forest delineation, it may exclude some sparsely wooded systems, particularly in dry forest–savanna transition regions such as parts of western North America.”

References:

Harris, N. L., Brown, S., Hagen, S. C., Saatchi, S. S., Petrova, S., Salas, W., Hansen, M. C., Potapov, P. V. & Lutsch, A. (2012). Baseline Map of Carbon Emissions from Deforestation in Tropical Regions. Science 336, 1573-1576.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R. et al. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. Science 342, 850-854.

Myroniuk, V., Kutia, M., J. Sarkissian, A., Bilous, A. & Liu, S. (2020). Regional-Scale Forest Mapping over Fragmented Landscapes Using Global Forest Products and Landsat Time Series Classification. Remote Sensing 12, 187.

2. Please do some kind of formatting with the reference section to make them easier to look through (number them, have some indentation on the left side, whatever is in accordance with the publication's style)

Response:

We have increased the spacing after each reference and added a hanging indent to improve readability.

3. 142-151: Please provide the brown vegetation index as an equation as this is not as commonly used for fire as the normalized burn ratio (NBR). Why was this used instead of NBR and how is it different?

Response:

We thank the reviewer for this important comment. We agree that the NBR is one of the most widely used indices for burned area mapping. The objective of adopting BVI in this study was not to replace NBR as a general burned area index, but to improve the burn-signal consistency within our multi-year Landsat compositing framework. Compared with NBR, which relies on the contrast between NIR and SWIR2 reflectance and is highly sensitive to vegetation structural changes, BVI emphasizes the contrast between green and SWIR2 reflectance. Because NIR generally exhibits stronger and more rapid post-fire recovery responses than visible wavelengths (Pérez-Cabello et al., 2021; McKenna et al., 2018), NBR-based signals may attenuate more rapidly in composited imagery, particularly in regions with sparse observations or rapid vegetation regrowth. In comparison, the green-SWIR2 contrast used in BVI was considered advantageous for preserving persistent burn darkening signals in composited imagery under heterogeneous observation conditions. Additionally, BVI was selected partly because cloud, snow, and aerosol contamination generally increase BVI values, making them less likely to be retained in minimum-value composites.

We have revised the manuscript to clarify that BVI was selected to improve the consistency and persistence of burn-signal characterization in the compositing step, rather than to imply universal superiority over NBR. We have added the equation of BVI and modified the following sentences in **Lines 177 – 188** of the **Methods** as below:

“Previous compositing approaches commonly relied on minimum NBR, NDVI, or NIR values (Chuvieco et al., 2005; Miettinen and Liew, 2008; Barbosa et al., 1998; Alencar et al., 2022). However, NIR-based burn signals often recover rapidly after fire, which can reduce their persistence in long-term composites, particularly in regions with sparse observations or rapid vegetation regrowth (McKenna et al., 2018; Pérez-Cabello et al., 2021). By contrast, the green-SWIR2 contrast captured by BVI tends to preserve burn-darkening signals for longer periods and is less sensitive to cloud, shadow, and aerosol contamination during minimum-value compositing (Liu, 2017). BVI is calculated as:

$$BVI = \frac{\rho_{Green} - \rho_{SWIR2}}{\rho_{Green} + \rho_{SWIR2}},$$

where ρ_{Green} and ρ_{SWIR2} represent surface reflectance in the green and SWIR at 2.1 μm (SWIR2) wavelengths, respectively. BVI exploits the contrasting spectral responses of burned surface in the green and SWIR2 bands. Following fire disturbance, SWIR2 reflectance typically increases because of vegetation moisture reduction and charcoal deposition, whereas green reflectance decreases with vegetation damage (Chuvieco et al., 2019; Liu, 2017), producing characteristically low BVI values over burned area.”

We have also added the following sentence in **Lines 516 – 518** of the **Discussion**:

“Future work should evaluate how different compositing strategies influence fire patch reconstruction across contrasting fire regimes and compare the performance of BVI-based and NBR-based approaches under varying environmental conditions.”

References:

Pérez-Cabello, F., Montorio, R., & Alves, D. B. (2021). Remote sensing techniques to assess post-fire vegetation recovery. Current Opinion in Environmental Science & Health, 21, 100251.

McKenna, P., Phinn, S., & Erskine, P. D. (2018). Fire Severity and Vegetation Recovery on Mine Site Rehabilitation Using WorldView-3 Imagery. Fire, 1(2), 22.

4. 186: maybe give this distance in meters (600m).

Response:

We have added the distance in meters in the manuscript.

5. 186: Same year is almost certainly too broad of a time range. Many other products have used same month, or even 5-15 days. Products using those tighter temporal windows also included grasslands, but still, considering the global scope, many forested areas with year-round growing seasons (e.g. SE United States) have been observed to burn multiple times in the same year.

Response:

We appreciate the reviewer’s important comment regarding the temporal aggregation window used to define fire patches. We agree that annual aggregation is not appropriate for studies focused on fire-event chronology, ignition timing, or short-interval reburning. The adopted temporal framework reflects a deliberate trade-off aimed at reconstructing long-term fire patch structure from heterogeneous Landsat observations at the global scale.

Compared with shorter temporal windows, annual compositing can help preserve the spatial coherence and complete burned scar extent of large fires, particularly for fires that span multiple months, which is common in boreal and temperate forests). This enables analyses of fire patch geometry, size distribution, and landscape-scale fire heterogeneity that are difficult to achieve using temporally fragmented monthly fire products.

Another consideration is the availability of cloud- and snow-free Landsat observations. In many regions globally, particularly in high latitudes, humid tropical regions, and mountainous areas, clear observations are often only available at intervals of several months (Feng and Wang, 2024; Flores-Anderson et al., 2023). As a result, consistent monthly-scale compositing of fire patches

with Landsat remains challenging at a global scale.

Given these constraints, a yearly temporal window was adopted as a pragmatic compromise to ensure global consistency in mapping burned scar extent while maximizing spatial coverage and completeness. This design prioritizes the reconstruction of spatially continuous burn scars rather than resolving within-season fire chronology.

We acknowledge that multiple fire occurrences within a single year may be merged in regions with repeated burning. However, because the primary objective of the dataset is to characterize the spatial structure and geometry of fire patches, the annual aggregation framework remains suitable for the intended applications of GlobMap FFP. This issue is likely to be most relevant in regions characterized by frequent reburning within a year, such as parts of the southeastern United States and some tropical fire-prone landscapes.

We have modified the writing in **Lines 223 – 232** of the **Methods** to clarify this rationale:

“Burned area pixels detected within the same burned year and separated by less than 20 Landsat pixels (600 m) were grouped as a single fire patch and assigned a unique fire identifier. The objective of this procedure was to generate spatially coherent fire patches suitable for fire regime analyses at a global scale rather than to reconstruct individual ignition events. This distance threshold was selected to bridge small unburned gaps commonly caused by heterogeneous fire spread, cloud contamination, or omission errors in burned area detection while avoiding excessive merging of spatially independent fires. Temporal segmentation was based on annual burned year assignments rather than sub-annual fire chronology, because the heterogeneous availability of cloud- and snow-free Landsat observations limits reliable global-scale reconstruction of fire timing (Feng and Wang, 2024; Flores-Anderson et al., 2023). As a consequence, multiple fires occurring within the same year and in close spatial proximity may be represented as a single fire patch, potentially overestimating patch sizes and underestimating fire occurrences.”

We have added the following sentences in **Lines 473 – 478** of the **Discussion**:

“Because the availability of cloud- and snow-free Landsat observations varies substantially across regions and time periods, GlobMap FFP focuses on reconstructing the spatial characteristics of fire patches rather than resolving detailed within-season fire chronology to achieve global consistency. The adopted multi-year compositing strategy therefore represents a pragmatic trade-off between global consistency, temporal precision, burned scar completeness, and computational efficiency. Although this design inevitably sacrifices some temporal precision, it enables a globally consistent reconstruction of long-term fire patch patterns from the heterogeneous Landsat archive.”

We have also discussed the limitation of our methods in **Lines 508 – 516**:

“Additional differences arise from the methodological choices adopted to achieve globally consistent fire patch reconstruction. The multi-year compositing strategy reduces the influence of cloud contamination, data gaps, and uneven observation availability while maintaining computational feasibility. Yet, because only a single observation is retained for each pixel within a compositing interval, repeated burning occurring at the same location during the same interval are not explicitly reconstructed. Fire recurrence may be underrepresented in frequently burned regions, and fire occurrence frequencies derived from the dataset should be interpreted with caution. Furthermore, the annual aggregation may merge temporally adjacent fires into a single fire patch, likely resulting in overestimated patch sizes and underestimated fire frequencies. Consequently, GlobMap FFP should be interpreted as a spatially explicit fire patch dataset rather than a complete inventory of all forest burned area, particularly in frequently burned tropical forests.”

References:

Feng, L., & Wang, X. (2024). Quantifying Cloud-Free Observations from Landsat Missions: Implications for Water Environment Analysis. Journal of Remote Sensing, 4, 0110.

Flores-Anderson, A. I., Cardille, J., Azad, K., Cherrington, E., Zhang, Y., & Wilson, S. (2023). Spatial and Temporal Availability of Cloud-free Optical Observations in the Tropics to Monitor Deforestation. Scientific Data, 10(1), 550.