

Response to topic editor:

[Comment 1] We have received an additional review from reviewer 1, and they generally find the technical aspects of REA to appear robust. However, they note several major clarity issues. The most major is that a new sensitivity analysis was added without any methodological explanation. I agree with this assessment, and agree that such an analysis needs to be fully described in the main text in the methods. They also felt that there are several locations throughout that are unclear yet, and that the messaging needs improvement. There appear to be instances where they are requesting an edit that had been requested in the previous round. Therefore, while the reviewer concerns about technical aspects appear addressed, clarity needs to be improved before the manuscript can be accepted for publication.

Response: We sincerely appreciate the editor's time and effort. In response to your and the reviewer's suggestions, we have added the detailed methodological explanation of the sensitivity analysis, and revised the manuscript throughout to enhance clarity and strengthen the messaging. Please find our detailed replies to each of the reviewers' comments below.

Response to Reviewer #1:

[Comment 1] Thanks to the authors for their edits and thorough response to my previous comments. The discussion of the REA method limitations and the methods for the "trend" analyses are more robust. The discussion section also has more connection to the literature now. However, it appears that a new analysis has been added without thorough explanation. In the discussion section (lines 458-473), the authors mention that they conducted a sensitivity analysis. This is the first time this analysis is mentioned in the paper. It should have been included much earlier in the methods section and the process thoroughly explained (e.g., "groups" are mentioned in lines 468-469, but there's no explanation of what a "group" is). I think I pieced together what the authors are trying to do from the figures, but there needs to be far more explanation in the text about what this is and why/how it was done. Given more context, I think this can be important piece of information to support the REA dataset credibility/robustness.

Response: We sincerely thank the reviewer for the positive and thoughtful suggestions, which have significantly improved the clarity and rigor of our manuscript. We appreciated the reviewer's help in the acknowledgments in the revised manuscript. Specifically, we have added detailed descriptions of the sensitivity analysis earlier in the Methods section and improved the text according to the reviewer's suggestions, please refer to the text in Lines 241-247, and lines 472-474, lines 476-477 in the revised manuscript.

[Comment 2] Also, as stated before, please also be careful with the use of "phenocam" vs "PhenoCam" throughout. Below are additional line edits.

Response: Following the reviewer's suggestion, we have revised the use of "phenocam" and "PhenoCam", please refer to line 33, 37, 178, 405 in the revised manuscript.

Line edits:

[Comment 3] Line 18: "largest correlation" compared to what? The original datasets?

Response: We have revised it as "*The start and end of the growing season in the newly merged dataset showed the highest correlation with ground-based phenocam observations, compared to the original datasets.*" (lines 17-18 in the revised manuscript).

[Comment 4] Line 20: Is this the average rate for the entire Northern Hemisphere? I would add that in.

45 **Response:** It is the average rate for the regions north of 30° N. Following the reviewer’s suggestion, we have clarified this information in the revised manuscript in lines 21-22.

[Comment 5] Line 32-33: When you say PhenoCam here, do you mean “the PhenoCam Network” or phenocams in general? The PhenoCam Network started in 2008, so it is NOT over 20 years old yet. However, the use of phenocams more generally has occurred for over 20

50 years. I think you mean “phenocams” in this context.
Response: We are sorry for the unclear description, in the revised manuscript, we have revised this sentence as “*For example, phenocams, a near-surface remote sensing tool, has been operational for more than 20 years*” (lines 33-34).

[Comment 6] Line 26: change to “phenocam”

55 **Response:** changed following the reviewer’s suggestion (line 37 in the revision).

[Comment 7] Line 48-49: Define acronyms AVHRR and VIPPHEN

Response: Defined following the reviewer’s suggestion (line 49-51 in the revision).

[Comment 8] Line 51: This is the first time NDVI is mentioned, so please provide a brief explanation of what it measures.

60 **Response:** Following the reviewer’s suggestion, we have provided a brief explanation of NDVI in the revised manuscript, please refer to lines 54-57.

[Comment 9] Line 68-69: This wording is awkward: “method shows its advantages by simplicity and efficiency

65 **Response:** We have revised it as “Compared to traditional data fusion methods, the REA method shows clear advantages in simplicity and computational efficiency” (lines 86-87 in the revised manuscript).

[Comment 10] Line 73: This is the first time the “voting principle” is mentioned. Please explain what it is here.

70 **Response:** Following the reviewer’s suggestion, we have provided an explanation of voting principle in the revised manuscript, please refer to lines 91-92.

[Comment 11] Line 75-88: I suggest switching the order so that this paragraph comes before the previous paragraph. That way you describe remote sensing phenology methods before you discuss combining them using REA.

75 **Response:** We thank the reviewer for this thoughtful comment, and we have switched the order of these two paragraphs following the reviewer’s suggestion, please refer to the lines 62-93 in the revised manuscript.

[Comment 12] Line 118-119: What is the “specific segment of the growing season” ? Does that just mean the entire growing season? If so, please state that.

80 **Response:** The amplitude is calculated as the difference between the maximum and minimum values of the EVI2 time series during the entire growing season. To clarify, we updated the text in the revised manuscript, please refer to lines 122-123.

[Comment 13] Line 121-123: This is a repeated sentence. I suggest deleting it in this location.

Response: Deleted following the reviewer's suggestion.

85 **[Comment 14]** Line 148-149: You already defined amplitude above. I suggest deleting this sentence.

Response: Deleted following the reviewer's suggestion.

[Comment 15] Line 150: How were they produced? Were all 5 averaged? Or combined in a different way?

90 **Response:** yes, it was averaged across the five methods. In the revised manuscript we revised it as: "*The average spring (SOS) and autumn (EOS) phenological dates were produced by averaging the results obtained from the five fitting methods*" (lines 150-151).

[Comment 16] Line 162: Briefly explain what GCC is and provide a source (e.g., a measure of greenness intensity calculated from digital imagery)

95 **Response:** Following the reviewer's suggestion, we updated the explanation, please refer to lines 162-165 in the revised manuscript.

[Comment 17] Line 165: add "the" before ROI

Response: Added following the reviewer's suggestion (line 167).

100 **[Comment 18]** Line 174: I'm confused by this- isn't the point that some remote sensing products are better for certain ecosystems than others? Isn't this biasing your comparison if you're removing these challenging sites?

105 **Response:** Our intention was to ensure the validity of inter-dataset comparisons by focusing on sites where all products provided usable data. Including sites with invalid phenology estimates in remote sensing data would introduce noise and potentially distort the evaluation of consistency and performance. To clarify, in the revised manuscript, we updated the sentence as follows: "*For this study, we excluded 26 sites that only provided one type of transition date (either SOS or EOS) and removed 90 sites where none of the four remote sensing datasets provided valid phenology estimates. These excluded sites were primarily located in cropland-dominated areas or regions with sparse vegetation, where the low spatial resolution of remote sensing data limits the reliable detection of phenological transitions.*" (lines 174-179).

110 **[Comment 19]** Line 176: I think you mean "phenocam" here.

Response: We have revised it into "phenocam" (line 178).

[Comment 20] Line 260: Delete the word "hypothesis"

Response: Deleted following the reviewer's suggestion.

115 **[Comment 21]** Line 261: Please explain what the SOS/EOS “trend” is (e.g., the change in SOS/EOS date through time).

Response: Following the reviewer’s suggestion, we updated the explanation, please refer to lines 270-273 in the revised manuscript.

[Comment 22] Line 271: Didn’ t you also use a Mann-Kendall trend test to analyze the growing season greenness trend? If so, please state that.

120 **Response:** yes, we also used the MK trend test. Following the reviewer’s suggestion, we updated the statement in the revised manuscript please refer to lines 271-273.

[Comment 23] Line 337: change to: “North America and on the Qinghai – Tibet Plateau”

Response: Changed following the reviewer’s suggestion (line 347-348).

[Comment 24] Line 343: end of season shouldn’ t be capitalized

125 **Response:** Changed following the reviewer’s suggestion (line 353).

[Comment 25] Line 351: Missing parentheses around “Fig. 4(b, e)”

Response: Changed following the reviewer’s suggestion (line 360).

[Comment 26] Line 353-354: Don’ t need to write out SOS and EOS again here.

Response: Deleted following the reviewer’s suggestion.

130 **[Comment 27]** Line 369: “VIP dataset has a lower estimation in the SOS range of DOY 100 – 140” - what does this mean? That VIP is bad a predicting SOS dates when they occur early in the season? I guess I don’ t understand why the DOY 100-104 range is given here.

135 **Response:** Thank you for pointing this out. The VIP dataset show an earlier SOS dates compared to other datasets, particularly in regions where the actual SOS falls between day of year (DOY) 100 – 140, which corresponds to early spring. To clarify, in the revised manuscript, we have revised the sentence as follows: “*The VIP dataset tends to estimate earlier SOS dates than the other datasets, especially in areas where SOS occurs early in the season (DOY 100 – 140).*” (lines 378-379).

[Comment 28] Line 377: Don’ t need to define EOS again

140 **Response:** Changed following the reviewer’s suggestion (line 387).

[Comment 29] Line 379: Better compared to what? A simple average of the datasets?

Response: yes. in the revised manuscript, we clarify this statements as: “The REA-based EOS also shows better performance compared to the simple average of the original datasets” (lines 389-390).

145 **[Comment 30]** Line 394: from “the PhenoCam Network”

Response: Changed following the reviewer’s suggestion (line 405).

[Comment 31] Line 395: Remove period after US

Response: Changed following the reviewer’s suggestion (line 405).

[Comment 32] Line 405: Should this say Fig 6 instead of 7?

150 **Response:** Thank you for pointing this mistake and it has been corrected (line 416).

[Comment 33] Line 428: Replace “across” with “between”

Response: Changed following the reviewer’s suggestion (line 439).

[Comment 34] Line 442: Start new paragraph here - this doesn't relate to differences in ecological regions (topic sentence of this paragraph). Also, define "greening trend" (e.g., how average greenness over the growing season changes through time). Put this in context- what does it mean?

155

Response: Changed following the reviewer’s suggestion (line 454-459).

[Comment 35] Line 458-473: See my notes above for this section. Also, line 460” “average deviation” of what? SOS and EOS?; Line 463: What is a “multi-time span analysis” ?

160

Response: Changed following the reviewer’s suggestion (line 241-247, 472-474, 476-477).

[Comment 36] Line 488-489: What was the difference in rates they found due to this difference?

Response: In the revised manuscript, we updated this statements as: “*For instance, Jeong et al. (2011) analyzed temperate vegetation between 30° N – 80° N (SOS advance in 0.2 days per year), and Wang et al. (2015) focused on 30° N-75° N(SOS advance in 0.14 ± 0.6 days per year), excluding evergreen forests and managed landscapes.*” (line 500-501).

165

Figures

[Comment 37] Figure 2 caption: It’ s confusing to relist the panel letters again- I suggest changing to “(a and c) Weights of the four phenology datasets during 1982 – 2020 and (b and d) latitudinal differences for (upper) the SOS and (lower) the EOS”

170

Response: Changed following the reviewer’s suggestion (line 330).

[Comment 38] Figure 3 caption: Recommended re-wording “Spatial distribution of the mean contribution of the four datasets to the merged SOS and EOS results. The mean SOS (a-d) and EOS (e-h)weight derived from the GIM_4g, MCD12Q2, VIP, and GIM_3g datasets.” And delete everything after.

175

Response: Changed following the reviewer’s suggestion (line 344).

[Comment 39] Figure 4: Would it be possible to use the same color scales for panels a & d? It would make it easier to compare and interpret.

Response: Thank you for the suggestion, but we intentionally used different color scales to visually distinguish the spring (SOS) and autumn (EOS) phenophases, reflecting their seasonal characteristics—green tones for spring onset and brown tones for autumn senescence. We therefore prefer using the different colors.

180

[Comment 40] Figure 5 m & n: I don’ t see the average group on the radar charts. Is it

complexly covered by another group? Perhaps use a color that's more different from the REA group?

185 **Response:** Changed following the reviewer's suggestion (line 395).

[Comment 41] Figure S1 caption: I'm a little confused- what's the difference between a "given" threshold and the "dynamic vegetation threshold" ?

190 **Response:** The "given" threshold refers to a fixed NDVI or EVI2 value, which is used to identify phenological events. The "dynamic" threshold is calculated by multiplying the given threshold by the range of the vegetation index, and then adding the minimum value of the vegetation index. In the revised manuscript, we have clarified the difference in the caption. (line 5)

[Comment 42] Figure S3 caption: Please state what a positive vs negative Sens slope means in context

195 **Response:** Changed following the reviewer's suggestion (line 20-25).

[Comment 43] Figure S6 caption: Include the original REA timeframe to remind readers.

Response: Changed following the reviewer's suggestion (line 43).

[Comment 44] Figure S7: How were the groups selected? It doesn't include all the possible dataset combinations.

200 **Response:** Because including all possible dataset combinations would take up too much space here, we therefore selected three representative combinations that each include all four datasets used in the complete fusion. These combinations were chosen to illustrate the influence of different data source selections on the final fusion results.