Responses and Revisions to Comments

Dear reviewer

Thanks for your reviewing and valuable comments of our manuscript entitled "***A global estimate of monthly vegetation and soil fractions from spatio-temporally adaptive spectral mixture analysis during 2001–2022***". We also appreciate you for providing insightful feedback and comments to strengthen our manuscript.

We have revised our manuscript with considering each detailed suggestion that you and reviewers have graciously provided. These major revisions include:

1 Added validation and comparison of the data.

2 Clearly defined the methodological workflow involving endmember selection, hierarchical clustering, validation, etc.

3 Analyzed and discussed the necessity and advantages of the data.

Besides these revisions, all authors checked the manuscript carefully and several minor revisions have been done to finalize the manuscript.

The following is a point-by-point response to the questions and comments delivered in your letter. For your convenience, revisions made by the authors have been highlighted in red color in the both response and revised manuscript, which could be easily checked. We hope that our revisions and responses can satisfactorily address all the issues and concerns.

Sun et al. generated global monthly vegetation and soil fraction maps during 2001-2022 using a spectral mixture analysis method. Application of this method in global scale is interesting and deserves a scientific publication. However, the proposed product at current stage is not enough to publish on ESSD and the accuracy of this product is not very well demonstrated, from the perspective of a dataset and users. There are some main concerns:

Thank you for your valuable feedback on our manuscript. We acknowledge your concerns regarding the reliability of the vegetation and soil fractions products described. We understand these concerns may stem from insufficiently validation in our manuscript. We have consequently enhanced the description of the validation process and added more validation/comparisons data.

On the one hand, we further authenticate our product through incorporating comprehensive global land cover and land use reference data, which were obtained from the Geo-Wiki crowdsourcing platform across four campaigns (page 12, line 265-276).

“Besides, we also authenticate our product through incorporating comprehensive global land cover and land use reference data (Fritz et al. 2017), which were obtained from the Geo-Wiki crowdsourcing platform across four campaigns: Human impact, wilderness, reference and disagreement. Over 150000 samples of land cover and land use were acquired in this reference data. To effectively validate our product, we need to filter the reference data, considering aspects such as data acquisition time, measurement methods, and credibility. We select first three campaigns, which have a good match with MODIS pixels (size 1×1km) and were observed during 2001 to 2022. High feasibility reference data is then selected through the confidence information of land cover estimates and the status of use of high spatial resolution imagery provided by the metadata. Similarly to the procedural description used for fractional vegetation-soil compared to GLCVRD, we reclassify ten classes of this dataset into our four groups of endmembers, including (1) tree cover, shrub cover, herbaceous vegetation/grassland, cultivated and managed, and mosaic of cultivated and managed/natural vegetation to PV and NPV; (2) flooded/wetland and open water to DA; (3) urban and barren to BS; (4) snow and ice to IS. This involve comparing the measured percent of land cover with the mean of endmember fractions within the corresponding 1×1km pixels.”

Moreover, we strengthened the comparisons between our generated data and four existing datasets: NDVI, MOD44B Vegetation Continuous Fields product, GLASS FVC dataset, and GEOV Fcover dataset (page 12-13, line 277-301).

“**2.4 Comparisons and limitations analysis**

To verify the consistency and merits of our dataset against existing ones, we conducted comparisons with four distinct pre-existing datasets: NDVI, MOD44B Vegetation Continuous Fields product, GLASS fractional vegetation cover dataset, and GEOV Fcover dataset. NDVI is derived from monthly synthesized MCD43A4 images. Both mean values of NDVI and our estimated fractional PV across all years and months are considered for comparison. The MOD44B Vegetation Continuous Fields product provides annual information about the percent tree cover, percent non-tree cover, and percent non-vegetated within each 250-meter pixel globally (DiMiceli et al., 2015). Consequently, we compare vegetation cover proportions—sum of percent tree cover and percent non-tree cover—to the sum of fractional PV and NPV. To align spatial and temporal resolutions, we aggregated the sum of percent tree cover and percent non-tree cover to a 500-meter scale. Simultaneously, we computed monthly Fractional PV and NPV as annual averages. The GLASS fractional vegetation cover dataset, offering an 8-day temporal frequency and dual spatial resolutions of 0.05° and 500 meters, was generated using a machine learning approach correlating MODIS reflectance with fractional vegetation cover (Jia et al., 2015). In our study, the 500-meter GLASS data was utilized to validate our estimated fractions. We computed annual averages from all the CLASS fractional vegetation cover data within a year and compared it with the annual averages of Fractional PV and NPV. GEOV FCover is a 10-day product estimated through the neural network using visible, near-infrared and shortwave infrared at 1km resolution (Baret et al. 2013). We aggregate our product to a 1km spatial resolution, and compare their annual averages with the annual averages of GEOV FCover.

Moreover, we also carry out a comparison with traditional linear spectral mixture analysis to demonstrate the advantages of our spatio-temporally adaptive spectral mixture analysis. Such comparison is performed using average of monthly $RMSE_{sma}$ of fully-constrained framework based on two fixed endmember spectral curves: (1) average of all spectral spectra for each endmember and (2) existing spectral spectra from Small and Sousa (2019).

Furthermore, to validate the uncertainties of the hierarchical clustering, we select a spectral spectrum from selected endmember spectra that exhibit the largest mean squared error from the mean of cluster for each cluster. These selected spectral spectra were then used to reconstruct an extreme library of endmember spectra and used to estimate fractional vegetation and soil using MESMA."

1 The selection of endmember spectra is critical in this method. The authors establish a library of endmember spectral using a nested framework, combined with MODIS derived endmember spectra used in previous studies. However, if the proposed method is valid enough, why is it necessary to use those from previous studies? What are the differences between spectra decided in this study compared to existing ones? How about the proportion of endmember from your method and previous studies in the final library? Furthermore, the authors used a hierarchical clustering method to generate sub-groups of endmembers. This clustering method is quite dependent on the input parameters. Among a lot of clustering algorithm, why did you select this one? How about the input parameters, as well as the performance of its accuracy?

First, we selected endmember spectra in ten globally representative regions using nested framework, and utilizing these previous spectra aimed to complement and enrich the diversity of the spectral library. In fact, we gather 7 PV, 5 NPV, 5 BS, and 1 DA endmember spectra through such literature search method, accounting for less than 15% of all endmember spectral curves.

We have revised this unclear description (page 8, line 205-207).

"Besides, we collect MODIS derived endmember spectra used in previous study to complement and enrich the diversity of the spectral library. (Okin et al. 2013; Daldegan et al. 2019; Meyer and Okin 2015; Sousa and Small 2019). We gather 7 PV, 5 NPV, 5 BS, and 1 DA endmember spectra through such literature search method."

Second, hierarchical clustering boasts strong interpretability and adaptability for clustering at diverse scales within data analysis[1]. Our use of hierarchical clustering aims to streamline the number of endmembers, thereby optimizing the efficiency of MESMA. This is crucial as MESMA requires intricate per-pixel adjustments, a challenge with

[1] Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. John Wiley & Sons.

extensive models due to current computational constraints.

In our approach, input parameters are all spectral curves per endmember to group similar curves to compute their mean—a representative typical spectral curve for each cluster. We analyze the uncertainties of such clustering in estimates of global vegetation and soil fractions, the results indicate the relative stability of the unmixed results as well as the effectiveness of the clustering.

We thus improved the descriptions of hierarchical clustering (page 8, line 210-214).

"To ensure feasibility of pixel-by-pixel operations in GEE, we also consider the similarity between the spectral curves, the hierarchical clustering method is selected to aggregate these spectra of each endmember as sub-groups, we input all spectral curves per endmember, grouping similar curves to compute their mean—a representative typical spectral curve for each cluster. Such hierarchical clustering boasts strong interpretability and adaptability for clustering at diverse scales within data analysis."

2 Only five endmembers are select to represent the surface. More evidences are necessary to demonstrate the selection of five endmembers, e.g., group of current land classifications.

Thank you for highlighting the importance of providing further evidence regarding the selection of five endmembers to represent the surface. Indeed, additional evidence, such as a comparison with current land classifications or other relevant groups, would bolster the rationale behind selecting these specific five endmembers.

(1) In fact, current research has demonstrated that the global land surface can essentially be characterized by four components—substrate, vegetation, dark and ice/snow—across diverse climate patterns and land cover types[2]. Moreover, to enrich the diversity of surface elements on a global scale, we also expanded the variety of endmember types based on endmember types commonly used in drylands[3]. These elements can characterize the fundamental composition of the Earth surface. Thus, we embody five endmembers to represent surface units.

We have improved the descriptions of five endmembers (page 6, line 163-171).

"(1) Recent studies have proposed various compositional endmember frameworks in different application contexts. For example, a framework including substrate, vegetation, dark and ice/snow was proposed and verified globally for both Landsat and MODIS to allow estimated fractions, this framework ensures consistent comparison of

[2] Sousa, D., Small, C.: Globally standardized MODIS spectral mixture models. Remote Sens Lett, 10, 1018-1027, https://doi.org/10.1080/2150704X.2019.1634299, 2019

[3] Guerschman, J. P., Scarth, P. F., McVicar, T. R., Renzullo, L. J., Malthus, T. J., Stewart, J. B., Trevithick, R.: Assessing the effects of site heterogeneity and soil properties when unmixing photosynthetic vegetation, non-photosynthetic vegetation and bare soil fractions from Landsat and MODIS data. Remote Sens. Environ., 161, 12-26, https://doi.org/10.1016/j.rse.2015.01.021, 2015

estimated fractions across diverse climate patterns and land cover types (Small and Milesi 2013; Sousa and Small 2019). Another framework includes photosynthetic vegetation, non-photosynthetic vegetation, soil, and shade (Roberts et al. 1993), this framework was widely adopted for presentation surface structure worldwide, particularly in tropical rainforest and dryland ecosystems (Guerschman et al., 2015). These elements can characterize the fundamental composition of the Earth surface. Thus, we embody five endmembers to represent surface units, these five endmembers include photosynthetic vegetation (PV), non-photosynthetic vegetation (NPV), bare soil (BS), dark (DA), ice/snow (IS)"

(2) Besides, Considering the scalability of hierarchical clustering and MESMA, we can analyze land cover types within the 15 subclasses under the five endmembers. We discover that these 15 subclasses encompassed tropical rainforest, temperate forest, cropland, grasslands, temperate deciduous forest in winter, crop residues, shrubs in dryland, moving sands, sand dunes, bare ground, moving sands. waters, bare rock, polar glaciers, and alpine glaciers. They cover major land cover types globally, indicating the representativeness of our selected endmembers (Fig. S2).

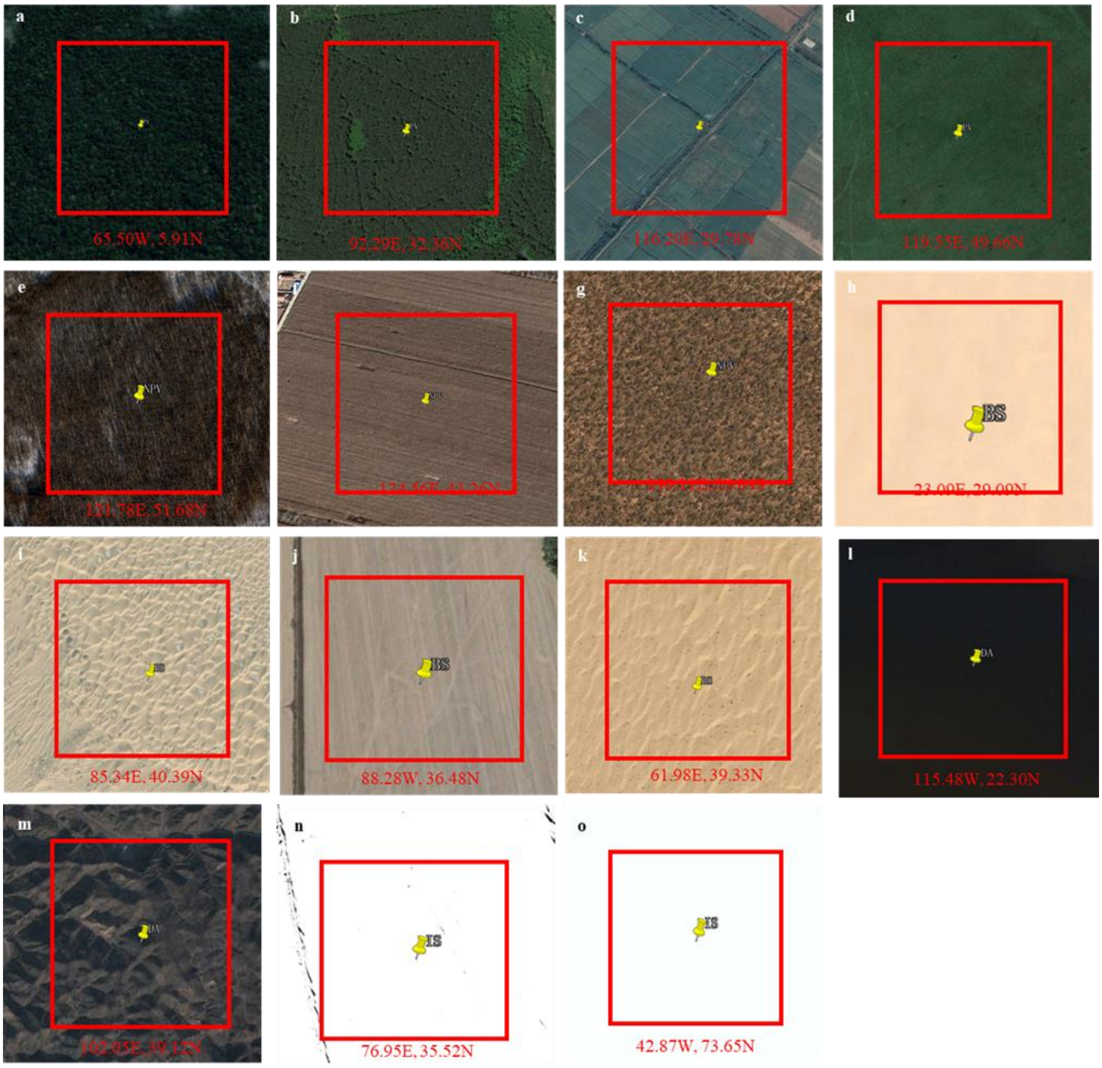

Figure S2: Typical images representing selected pure pixel of each endmember. a-o are tropical rainforest, temperate forest, cropland, grasslands, temperate deciduous forest in winter, crop residues, shrubs in dryland, moving sands, sand dunes, bare ground, moving sands. waters, bare rock, polar glaciers, and alpine glaciers.

(3) Furthermore, given sufficient computational resources, we can conduct more in-depth analyses and utilize MESMA for these 15 subclasses (Discussions, page 25, line 461-469).

"The spatio-temporally adaptive framework employed helps to increase the representativeness of endmember selection, and MESMA also considers the suitability of each combination of these endmembers within each pixel. However, considering the limitations of computational resources, our solution on hierarchical clusters of the endmember spectra can improve considerably cost-effective unmixing of long time-series satellite records over globe under the neglect of certain accuracy requirements (Fig. 3). With the assumption of increased computational power in the future, we believe that utilization of combination models from selected endmember spectra (35 GV spectra, 40 BS spectra, 25 NPV spectra, 16 DA spectra, and 15 IS spectra) or expanded endmember spectra may further improve the accuracy and stability of estimates of gradations of five surface vegetation and soil components at global scale."

3 The validation step is very important to convince the users to use your datasets, instead of other existing products. However, the validation is not enough. For example, the authors used GLCVRD reference dataset which was produced in 2010 to validate the new products during the period during 2001-2022. How did you consider the inconsistency between time period and land changes during this long period? I would suggest the authors to improve the validation section, either by cross-validation with other existing products or by independent datasets.

Due to challenges in conducting fraction estimation validation through field surveys, we employed validation methods based on high-resolution remote sensing imagery. The CLCVRD dataset is an example of such high-resolution reference data, hence we utilized it for corresponding validation. This dataset has demonstrated effectiveness in global-scale land cover verification. It's important to note that this dataset spans 2003-2012 (mainly 2010), and I apologize for any misleading information in previous descriptions. Thus, this data can effectively serve for fraction validation across different years and months.

(1) We have improved our description on CLCVRD and corresponding processes (page 11, line 236-251).

"Moreover, due to challenges in conducting fraction estimation validation through field surveys, we employ reference data obtained from high spatial resolution images as validation set. We thus select for two sets of reference data that their land cover classification systems are closely related to our five endmembers. Global Land Cover Validation Reference Dataset (GLCVRD) is provided with a 2m reference dataset from very high resolution commercial remote sensing data within 5 × 5 km blocks from 2003 to 2012 (Olofsson et al. 2012; Pengra et al. 2015; Stehman et al. 2012). These datasets support global estimates of classification accuracy for four major land cover classes: tree, water, barren, other vegetation, cloud, shadow, ice & snow. Various recent studies have selected this dataset to evaluate the continuous fields of land cover types (Baumann et al. 2018; Qin et al. 2019; Song et al. 2018). We use all GLCVRD reference dataset (Fig. 3a) to assess the accuracy of globally fractional vegetation and soil estimates from MESMA. Firstly, we filter the estimated fractions based on the corresponding year and month obtained from the reference data. Simultaneously, aligning the interpretations of land cover types with our endmembers, we pair them accordingly, that is, tree and other vegetation represent PV and NPV, barren stands for BS, water and shadow correspond to DA, and ice & snow denote IS. Subsequently, we reclassify these paired land cover types and calculated their percentage within 5×5 km blocks, in which we exclude cloud coverage (named no data). Additionally, utilizing these cloud-free pixels in each block, we compute the mean of fractional values for each endmember, and then compare these estimated fractions with the measured percentage of paired the reclassified land cover types to validate the reliability of our product (Fig. S4)."

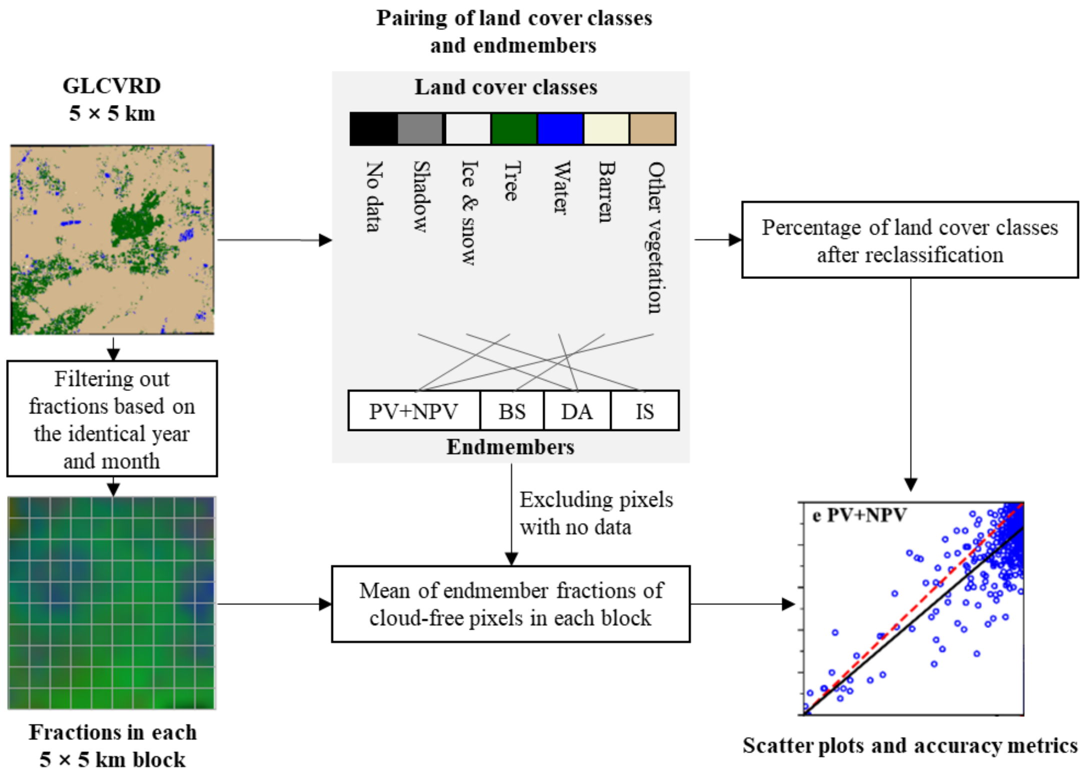

Figure S4: Procedural description for estimating fractional vegetation-soil compared to GLCVRD.

(2) Additionally, we conducted in-depth validation using new datasets[4] (page 12, line 265-276).

"Besides, we also authenticate our product through incorporating comprehensive global land cover and land use reference data (Fritz et al. 2017), which were obtained from the Geo-Wiki crowdsourcing platform across four campaigns: Human impact, wilderness, reference and disagreement. Over 150000 samples of land cover and land use were acquired in this reference data. To effectively validate our product, we need to filter the reference data, considering aspects such as data acquisition time, measurement methods, and credibility. We select first three campaigns, which have a good match with MODIS pixels (size 1×1km) and were observed during 2001 to 2022. High feasibility reference data is then selected through the confidence information of land cover estimates and the status of use of high spatial resolution imagery provided by the metadata. Similarly to the procedural description used for fractional vegetation-soil compared to GLCVRD, we reclassify ten classes of this dataset into our four groups of endmembers, including (1) tree cover, shrub cover, herbaceous vegetation/grassland, cultivated and managed, and mosaic of cultivated and managed/natural vegetation to PV and NPV; (2) flooded/wetland and open water to DA; (3) urban and barren to BS; (4) snow and ice to IS. This involve comparing the measured percent of land cover with the mean of endmember fractions within the corresponding 1×1km pixels."

[4] Fritz, S., See, L., Perger, C., McCallum, I., Schill, C., Schepaschenko, D., Duerauer, M., Karner, M., Dresel, C.,r Laso-Bayas, J. C., Lesiv, M., Moorthy, I., Salk, C. F., Danylo, O., Sturn, T., Albrecht, F., You, L., Kraxner F., Obersteiner, M.: A global dataset of crowdsourced land cover and land use reference data. Scientific data, 4, 1-8, https://doi.org/10.1038/sdata.2017.75, 2017.

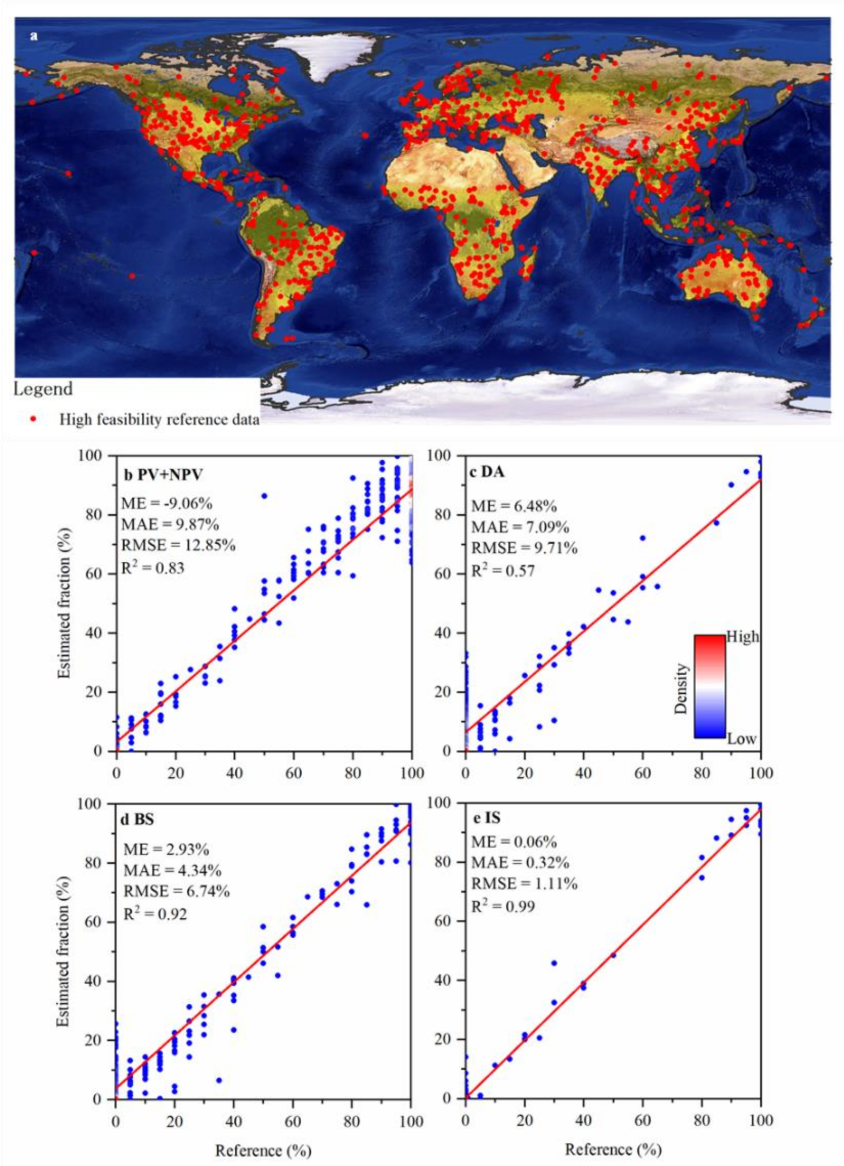

Fig. S5 Evaluation of global fractional endmember estimates based on land cover reference data. a, the location of high-feasibility land cover reference data. b-d, Scatter plots of PV+NPV, BS, DA, IS fractions against land cover reference data

(3) Moreover, we strengthened the comparisons between our generated data and four existing datasets: NDVI, MOD44B Vegetation Continuous Fields product, GLASS FVC dataset, and GEOV Fcover dataset (page 12, line 277-301).

"**2.4 Comparisons and limitations analysis**

To verify the consistency and merits of our dataset against existing ones, we conducted comparisons with four distinct pre-existing datasets: NDVI, MOD44B Vegetation Continuous Fields product, GLASS fractional vegetation cover dataset, and GEOV Fcover dataset. NDVI is derived from monthly synthesized MCD43A4 images. Both mean values of NDVI and our estimated fractional PV across all years and months are considered for comparison. The MOD44B Vegetation Continuous Fields product provides annual information about the percent tree cover, percent non-tree cover, and percent non-vegetated within each 250-meter pixel globally (DiMiceli et al., 2015). Consequently, we compare vegetation cover proportions—sum of percent tree cover and percent non-tree cover—to the sum of fractional PV and NPV. To align spatial and temporal resolutions, we aggregated the sum of percent tree cover and percent non-tree cover to a 500-meter scale. Simultaneously, we computed monthly Fractional PV and NPV as annual averages. The GLASS fractional vegetation cover dataset, offering an 8-day temporal frequency and dual spatial resolutions of 0.05° and 500 meters, was generated using a machine learning approach correlating MODIS reflectance with fractional vegetation cover (Jia et al., 2015). In our study, the 500-meter GLASS data was utilized to validate our estimated fractions. We computed annual averages from all the CLASS fractional vegetation cover data within a year and compared it with the annual averages of Fractional PV and NPV. GEOV FCover is a 10-day product estimated through the neural network using visible, near-infrared and shortwave infrared at 1km resolution (Baret et al. 2013). We aggregate our product to a 1km spatial resolution, and compare their annual averages with the annual averages of GEOV FCover.

Moreover, we also carry out a comparison with traditional linear spectral mixture analysis to demonstrate the advantages of our spatio-temporally adaptive spectral mixture analysis. Such comparison is performed using average of monthly $RMSE_{sma}$ of fully-constrained framework based on two fixed endmember spectral curves: (1) average of all spectral spectra for each endmember and (2) existing spectral spectra from Small and Sousa (2019).

Furthermore, to validate the uncertainties of the hierarchical clustering, we select a spectral spectrum from selected endmember spectra that exhibit the largest mean squared error from the mean of cluster for each cluster. These selected spectral spectra were then used to reconstruct an extreme library of endmember spectra and used to estimate fractional vegetation and soil using MESMA.”

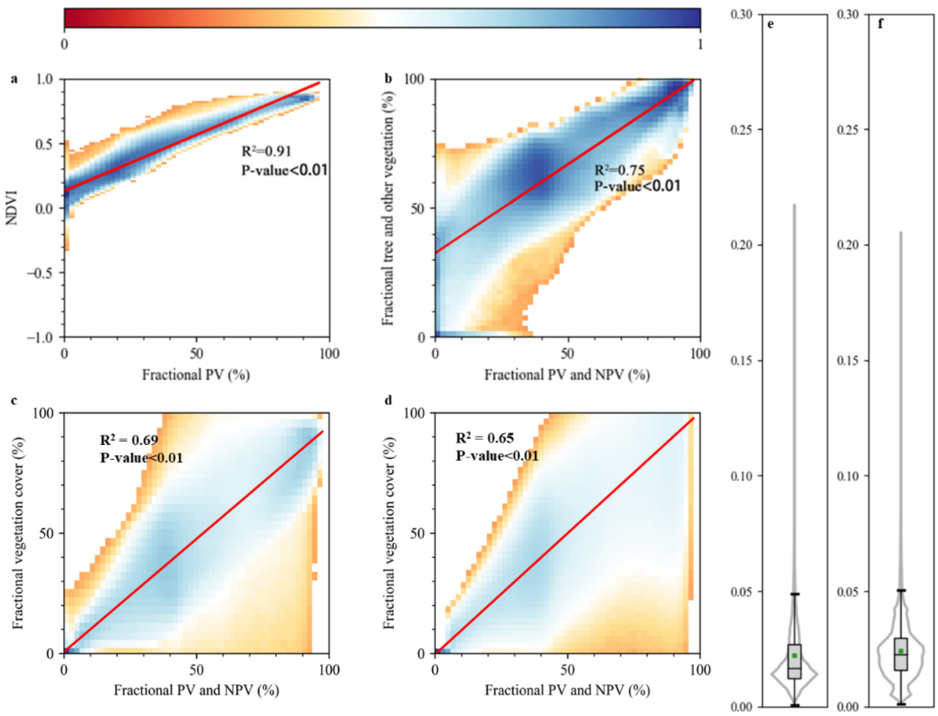

Figure 4: Comparisons with other datasets and traditional spectral mixture analysis models. a, b, **c**, **d** the bi-dimensional histogram of fractional endmembers and other dataset with bin size of 2%, including fractional PV against NDVI (a), fractional PV and NPV against fractional tree and non-tree vegetation of MOD44B vegetation continuous fields product (b), fractional PV and NPV against GLASS fractional vegetation cover product (c), fractional PV and NPV against fractional vegetation cover of GEOV Fcover product; **e**, **f**, the boxplot and violin plot for average of monthly $RMSE_{sma}$ for two fixed endmember spectral curves using fully constrained linear spectral mixture models, including (e) average of all spectral spectra for each endmember and (f) existing spectral spectra from Small and Sousa (2019).

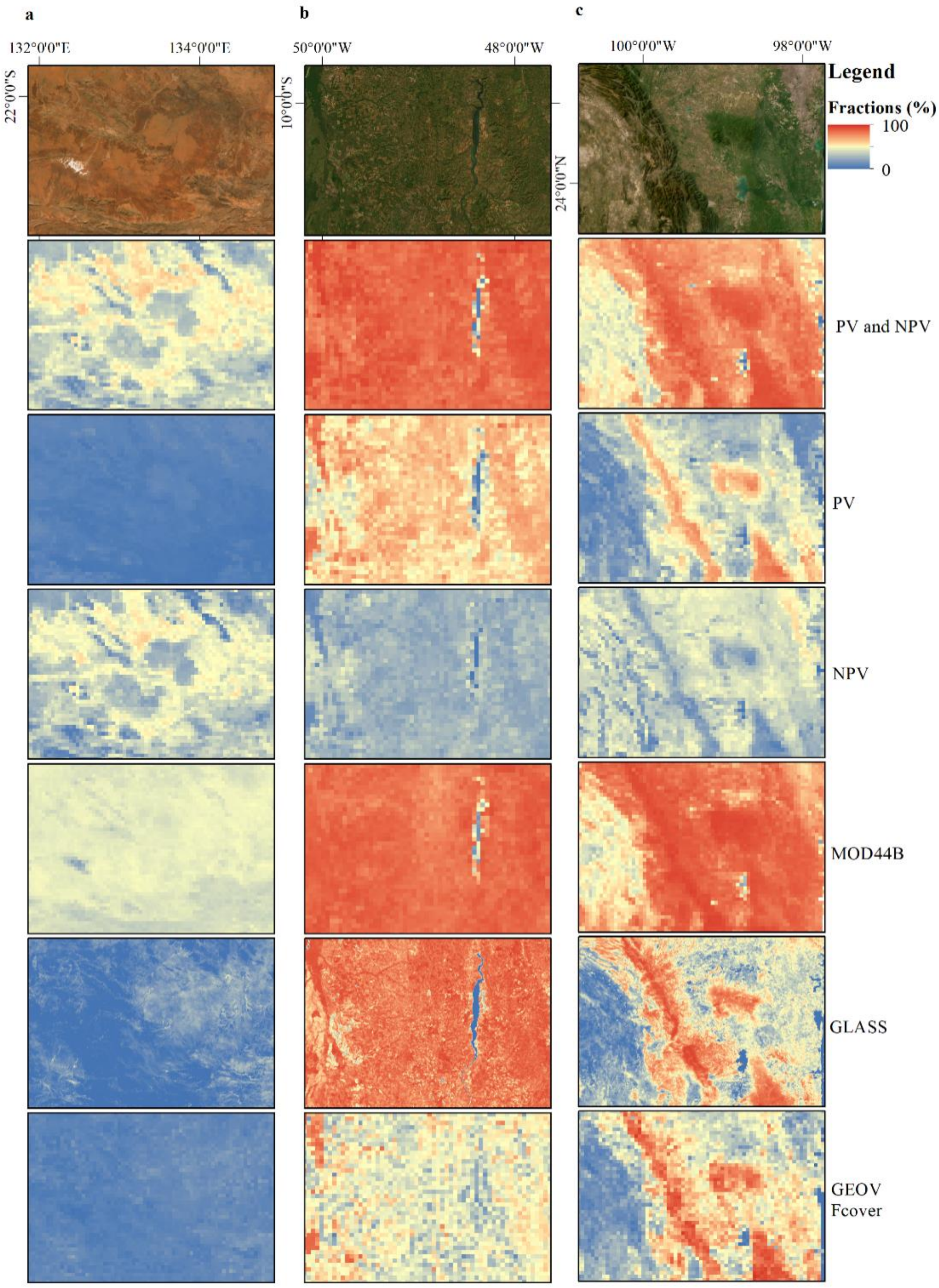

Fig. S6 The detailed graphs for comparing different datasets. a, b, and c represent comparisons of vegetation abundance products in different scenarios, specifically, regions with low vegetation cover in arid areas, high vegetation cover in tropical rainforests, and transitional zones from low to high values. The compared products include our produced PV and NPV, MOD44B, GLASS, and GEOV.

4 The structure of this manuscript should be re-organized, especially the results and discussion. Put all figures in results section. In discussion section, I expect more deep analysis and discussions on the intercomparison with existing products, the advantages and disadvantages of the new products, and uncertainties.

Thank you for your valuable feedback on the manuscript's structure, particularly emphasizing the need for reorganization, especially in the Results and Discussion sections. We will consolidate some figures within the Results section for improved coherence. In the Discussion section, we aim to delve deeper into comprehensive analyses and discussions. We'll focus on intricate intercomparisons with existing products, elucidating the advantages, disadvantages, and uncertainties inherent in the new products. This restructured discussion will offer a more comprehensive and insightful exploration of our findings, ensuring a clearer and more informative narrative.

We have moved “Compared with other datasets and traditional SMA model” and “Uncertainties of estimates of global vegetation and soil fractions” to Section 3.2 and Section 3.3. And we have improved the discussions of Advances and limitations of estimates of global vegetation and soil fractions in Section 4.1.

5 Some minor comments:

Thank you for your detailed feedback regarding the technical problems, we have corrected these errors and have also carefully examined and revised the manuscript to ensure the accuracy of the presentation.

Line 85-86: this is not the key scientific question of this study. More emphasis should be started from issues of existing global vegetation/soil fraction products, instead of application of spectral mixture analysis method in global scale.

- We have improved descriptions of existing global vegetation/soil fraction products (page 3, line 75-87).

  “In recent years, there have been significant advancements in fractional vegetation cover within the fields of remote sensing and environmental science. This progress has led to the development of various products at multiple resolutions, such as long-term global land surface satellite (GLASS), GEOV Fcover, multi-source data synergized quantitative remote sensing production system (MuSyQ) fractional vegetation cover (Baret et al. 2013; Jia et al., 2015; Mu et al., 2017; Zhao et al., 2023). These products primarily integrate and utilize data from different spectral bands and sensors, employing methods including machine learning and radiative transfer model. However, these data primarily focus on green vegetation, posing significant limitations in capturing information regarding non-photosynthetic vegetation and bare soil. This constraint also restricts the applicability of this data in arid regions. Although some initiatives and products focused on multi-element fractions, such as MOD44B and the Global Vegetation Fractional Cover Product (DiMiceli et al., 2015; Guerschman et al., 2015). For instance, the Global Vegetation Fractional Cover Product primarily targets arid regions, particularly Australia, focusing on photosynthetic vegetation, non-photosynthetic vegetation, and bare soil. Meanwhile, MOD44B achieves global-scale acquisition of trees, non-trees, and non-vegetative cover. There is a lack of unified classification systems among these products across global scale”

Line 135: Spectral mixture analysis, 's' should be capitalized.

- We have improved as "Spectral mixture analysis"

Line 166: what is the resolution of this land cover product? There are multiple land cover products, why do you select this one? How did you overlap the climate classification zones and the land cover products when the resolution is quite different?

- MCD12Q1 is a 500m land use cover data from Terra and Aqua Version 6 product. We selected the MCD12Q1 land cover product due to its widely acknowledged accuracy, high spatial resolution, and global coverage. This product is generated by a robust algorithm using multiple remote sensing datasets, providing comprehensive and detailed land cover information across various land cover classes. Its consistency, reliability, and compatibility with our study's objectives make it an ideal choice for comparison and validation against our newly developed products.

  Moreover, we do not overlap the climate classification zones and the land cover products, we used MODIS sinusoidal grid (10°× 10°intervals) to count the land cover diversity (D) and the full coverage capacity of climate types.

  "Finally, we selected the top 10 grids (i.e., h08v05, h12v12, h13v09, h16v01, h21v03, h22v02, h22v08, h24v06, h26v05, h27v06, h29v12) in terms of Simpson's Diversity Index (D) among all MODIS grids (Fig. S1a, b), and containing all Köppen-Geiger climate types" (page 7, line 182-185)

Line 180-181: While ***, especially in vegetation growing seasons. This sentence seems not complete.

- We have revised that as "PC eigenvector with relatively high contrast between the near-infrared band and other bands primarily captures information related to photosynthetic vegetation (PV), particularly during vegetation growing seasons. (page 7, line 193-195)"

Line 192: 35 GV spectra, should be PV spectra?

- We establish a library of endmember spectra considering spatio-temporal variability, this library includes 35 PV spectra, 40 BS spectra, 25 NPV spectra, 16 DA spectra, and 15 IS spectra (page 8, line 208).

Lien 212-213: how long it took to generate one global map on GEE?

- Exporting the data will take approximately 30-40 minutes per month per our generated grid, according to the grid of longitude 60° and Latitude 50°.

Line 260: It is strange that Sahara Desert and polar regions had higher RMSE. These regions were well-known for unique land cover type.

- Even though the two regions have a single land cover type of desert and snow/ice, the internal spectral profile is more complex and extreme, for example, deserts have both high and low reflectance sands. However, to realize the global-scale dominant element information extraction, we input all spectral curves per endmember, grouping similar curves to compute their mean, resulting in larger $RMSE_{sma}$ for these extreme regions, but basically, they are also in line with the requirements of model fitting accuracy.

  Such inaccurately pixels have been discussed in Section 4.1 Advances and limitations of estimates of global vegetation and soil fractions (page 25, line 481-485).

Figure 3.e-h are not enough to support your conclusion.

- We conducted in-depth validation using new datasets and discussed the limitations of validation using current land cover reference dataset (see details in response of comment 3).

Section 3.1: add more analysis for pixels that were estimated inaccurately.

- We have added more discissions for pixels that were estimated inaccurately (page 25, line 481-485).

  "We observed higher $RMSE_{sma}$ values in seemingly homogeneous areas like the Sahara Desert and Arctic regions. However, within these regions, there often exist extremely diverse land cover types, such as high and low reflectance sands and ice. When selecting endmembers and hierarchical clustering models, we might not have adequately considered these extreme spectral curves. As a result, these extreme areas exhibit a higher uncertainty."

Lie 358: indicate the resolution of MODIS pixel.

- each MODIS pixel (500 m) (page 24, line 434)