Response to Reviewer Comments

A global daily seamless 9-km Vegetation Optical Depth (VOD) product from 2010 to 2021

Die Hu [a], Yuan Wang [b], Han Jing [a], Linwei Yue [c], Qiang Zhang [d], Lei Fan [e*], Qiangqiang Yuan [a, f, g, †], Huanfeng Shen [h], and Liangpei Zhang [i]

[a] School of Geodesy and Geomatics, Wuhan University, Wuhan, China.

[b] School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing, China.

[c] School of Geography and Information Engineering, China University of Geosciences, Wuhan, China.

[d] Center of Hyperspectral Imaging in Remote Sensing (CHIRS), Information Science and Technology College, Dalian Maritime University, Dalian, China.

[e] Chongqing Jinfo Mountain Karst Ecosystem National Observation and Research Station, School of Geographical Sciences, Southwest University, Chongqing, China.

[f] Key Laboratory of Geospace Environment and Geodesy, Ministry of Education, Wuhan University, Wuhan, China.

[g] Key Laboratory of PolarEnvironment Monitoring and Public Governance, Ministry of Education, Wuhan University, Wuhan, China.

[h] School of Resource and Environmental Science, Wuhan University, Wuhan, China.

[i] State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China.

* Corresponding authors. E-mail address: leifan33@swu.edu.cn

† Corresponding authors. E-mail address: qqyuan@sgg.whu.edu.cn

Response to Comments of Reviewer #1:

General comments: This study introduces a novel approach to reconstructing and merging global daily L-band Vegetation Optical Depth (L-VOD) data at a 9-km resolution from 2010 to 2021. The authors integrate datasets from the SMOS and SMAP satellites. They use a three-dimensional discrete cosine transform-based penalized least square regression (DCT-PLS) method to fill gaps in the data, followed by a spatiotemporal fusion model (STFM) based on non-local filtering to enhance the spatial and temporal continuity of the dataset. This material is important for long-term canopy and vegetation change studies. The manuscript is well-prepared, and the VOD data are valuable and should be published in ESSD. However, I have some questions about the data and their seasonal pattern, hope the authors can clarify them and confirm the data quality.

Response: We sincerely thank the reviewer for his/her positive comments. According to the concerns, we have tried our best to improve the manuscript. The point-by-point responses are listed below. Thanks for your time.

Response to Specific Comments:

Q: 1) In Fig. 6, the description indicates that the rectangles emphasize some extreme value "reconstruction" results, but why does the first panel with a 30% missing proportion have a rectangle in the "Original" part line (magenta color)?

Response: This is a good question. In Figure 6, the red line represents the original values and is overlaid on the blue line representing the reconstructed values. In other words, the original values are missing, while the reconstructed values are continuous. The results show that the DCT - PLS model does not alter the original pixel values themselves. It preserves the original characteristics of the data and maintains the continuity of the reconstructed results. In the first panel with a 30% missing proportion, the rectangle is marked on the original values, indicating that the reconstruction process can not only maintain the original non - missing parts but also capture extreme values well. Similar studies [1] also mark extreme values on the original values.

References:

[1] Wang, G., Garcia, D., Liu, Y., De Jeu, R., and Dolman, A. J.: A three-dimensional gap filling method for large geophysical datasets: Application to global satellite soil moisture observations, Environmental Modelling & Software, 30, 139–142, 2012.

Q: 2) For Fig. 6 and Fig. 7, the colors for "Reconstructed" and "Original" are too similar. Please consider using more distinct colors, such as red and blue or black and blue. Additionally, the yaxis uses different scales in these figures. I suggest using the same y-axis scale range to better compare the relative size differences.

Response: Thanks for the reviewer's suggestion regarding the detail of our mapping. According to the advice, we have changed the colors of the reconstructed values and the original values to distinct blue and red. The extreme values are marked with black boxes.

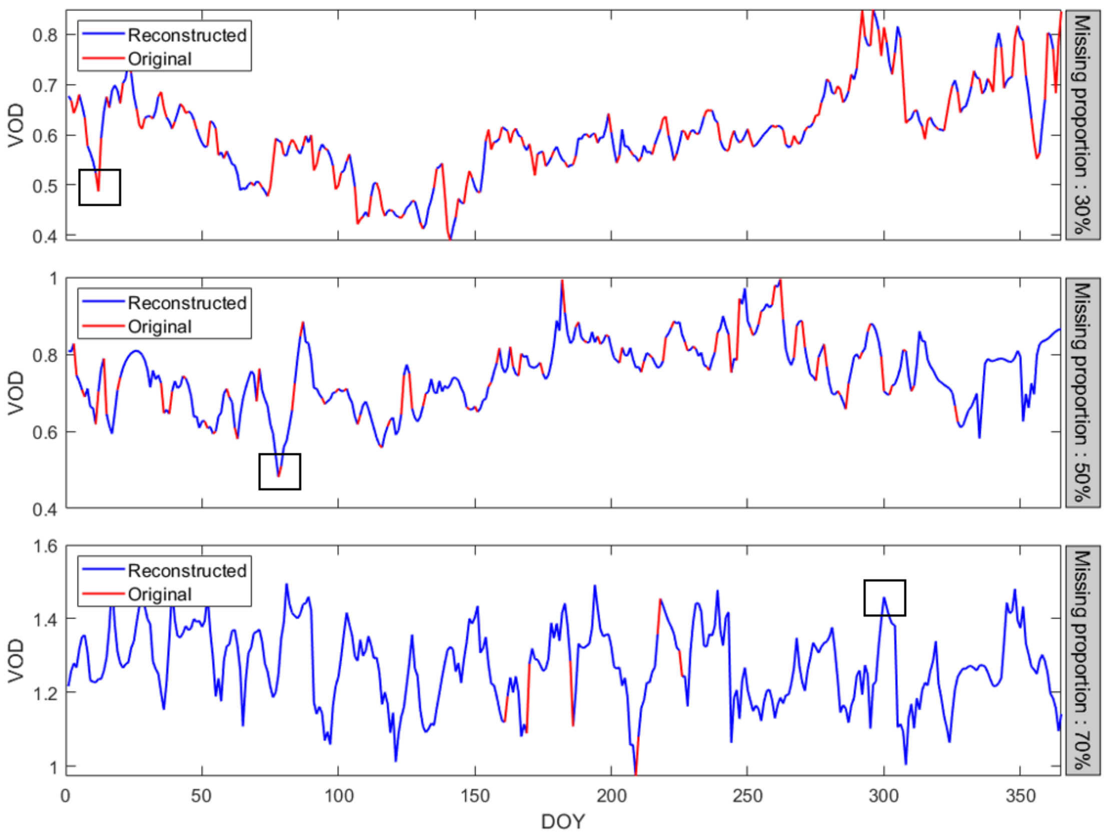

Fig. 1. Results of temporal variation in selected pixel at different missing data ratios in 2018, with red representing original values, blue representing model reconstructed values, and rectangles emphasizing some extreme value reconstruction results.

In the comparison of the temporal variation results of different land cover types, to ensure consistency, we select pixels with a data - missing proportion of 52% throughout the year for analysis.

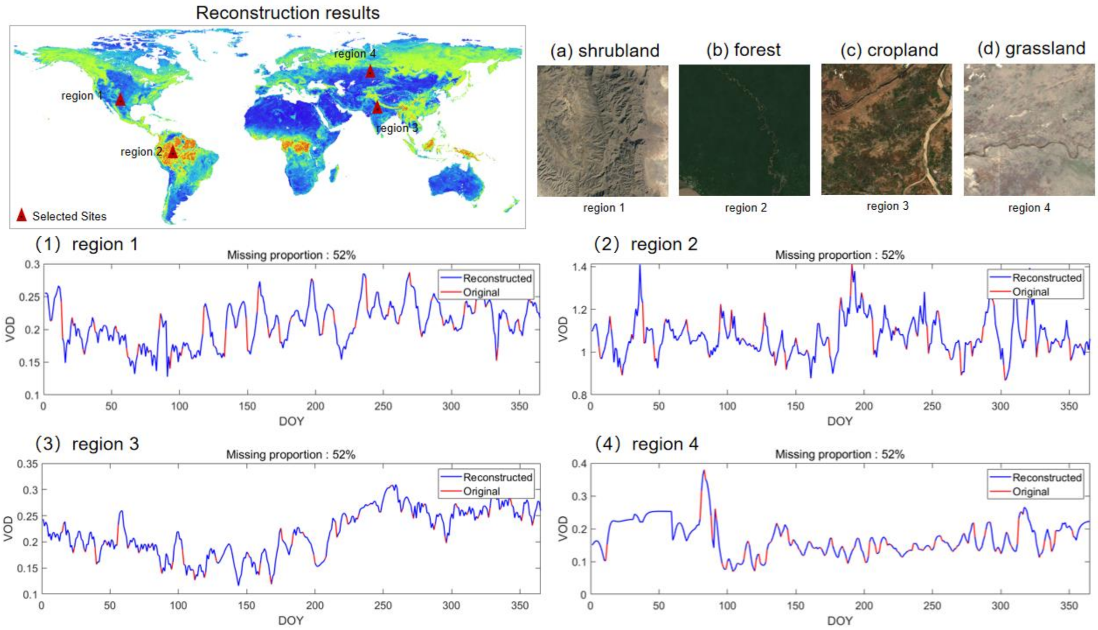

Fig. 2. The red dots in the figure indicate the pixel points selected to characterise the temporal variation of L-VOD under different vegetation conditions. Four different surface types are selected here, namely (a) scrub, (b) forest, (c) cropland, and (d) grassland; (1)-(4) represent the time-series variation maps of the corresponding pixels under the above surface types, respectively.

Q: 3) In Fig. 10, the x- and y-axes represent VOD. What do the colors and the color bar indicate? Please clarify their meaning.

Response: Thanks for the comment. The colors and the color bar indicate the density of data points in the scatter plot. The color bar indicates the range of density, with the colors transitioning from blue to red, where blue represents lower density and red indicates higher density. This helps visualize the distribution of the original and reconstructed VOD values, with more frequent data points represented by warmer colors (red) and less frequent points represented by cooler colors (blue).

Q: 4) In Fig. 18, VOD_st has more blank data compared to VOD_resmap. Can you provide a clearer explanation of why VOD_st is considered close to VOD_resmap

(line 490)?

Response: Thanks for the comment. During the spatiotemporal fusion process, VOD_st learns the temporal characteristics of VOD_resmos (from 2010 to 2015) and the spatial characteristics of VOD_resmap (with a 9-km spatial resolution). During the period from 2010 to 2015, only the SMOS satellite provides L-band VOD products. Therefore, the spatial coverage of VOD_st from 2010 to 2015 is completely dependent on VOD_resmos. In comparison, the spatial coverage of the SMOS satellite products is not as extensive as that of the SMAP satellite. In our paper, VOD_st is considered closer to VOD_resmap (line 490) in terms of numerical accuracy. Specifically, it is closer to the high-resolution product in terms of numerical performance of the valid data (9 km) rather than in terms of spatial coverage (VOD_resmos). This numerical proximity further demonstrates that the fusion product (VOD_st) has a certain degree of reliability in reflecting relevant features.

Q: 5) In Fig. 19, the red boxes and zoomed-in plots show that the reconstructed data appear smoother and free of striping. However, it is difficult to conclude that the reconstructed data are necessarily better. Why was the Black Sea region chosen as an example instead of another region with a higher VOD signal?

Response: Thanks for the comment. We select the Black Sea region as an example due to its representative ecosystem, which primarily consists of grasslands and croplands. Moreover, the proportion of missing data in this area is moderate, mostly ranging from 40% to 50%. Since the percentage of missing data is not very high, the data distribution in the region is relatively uniform. The data characteristics are generally consistent, reducing the impact of extreme values or unusual data clusters. Therefore, the difference between the monthly averages before and after reconstruction is not significant.

In order to better compare the results before and after the reconstruction, we have re-selected a relatively more representative area, the Kalimantan Island (5° S - 8° N, 108° E - 120° E). The VOD signals on Kalimantan Island are higher, and the missing - data proportion mainly ranges from 50% to 80%, which can better reflect the reconstruction ability. Kalimantan Island is characterized by its large - area and diverse - type tropical rainforests. The dynamic changes in vegetation are significantly affected by human activities. Located in the tropical climate zone, it has complex climatic conditions, abundant precipitation, and extreme weather events that can impact vegetation. With diverse landforms and a special geographical location, as well as social and economic activities such as agricultural development and eco - tourism, this island becomes a typical area for testing the effectiveness and reliability of the reconstruction method in complex environments.

In the selected local area, the original data presents blocky patterns. There are significant differences in VOD values between different patches, and the edges are rather abrupt. Meanwhile, there may be some noises or outliers in the original data, resulting in a non - smooth spatial distribution of the data. The reconstructed data shows a smoother spatial transition. It indicates that the reconstruction algorithm not only fills in the missing data values but also processes the noises and outliers in the original data to a certain extent. It is important to note that VOD products exhibit insignificant variations at the daily scale. So the difference between the monthly average data before and after the reconstruction is not significant.

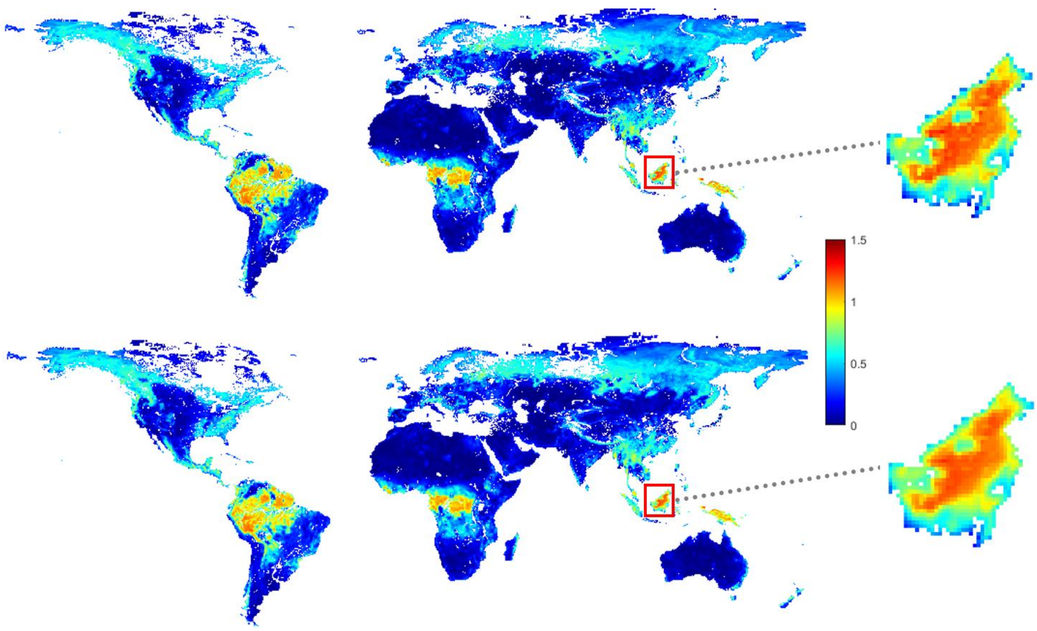

Fig. 3. Original (top) and reconstructed (bottom) results for July 2015 SMOS VOD monthly average. The red boxes highlight local areas.

Response to Data part Comments:

Q: 1) Each data file includes only one time step and has a file size of 3.7 MB. I suggest merging them by year into a single NetCDF file.

Response: We sincerely thank the reviewer for the suggestion. We have followed this suggestion and merged the daily data within each year into one NetCDF file. The variable names are named as VOD_xxxxyydd, where xxxx represents the year, yy represents the month, and dd represents the day. The longitude variable is named "lon" with a dimension of 4000×1, and the latitude variable is named "lat" with a dimension of 2000×1. It should be noted that these NetCDF files are saved using the netCDF4 library in Python, with the dimension order being (lat, lon). When reading these NetCDF files in MATLAB, the default data dimension order is (lon, lat). Therefore, it is necessary to transpose the variables to match the correct dimension order.

Q: 2) I selected VOD data from two days: summer (2021-07-07, left panel) and winter (2021-01-07, right panel). In the Eastern US and many other regions, the seasonal pattern appears reasonable, with higher VOD in summer and lower VOD in winter in the Northern Hemisphere. However, in Sweden (red rectangle) in Europe, the summer VOD (~0.6) is lower than the winter VOD (~0.8). Additionally, the black boxes indicate that the western US, Alaska and western Canada have higher VOD in winter (almost as high as Amazon rainforest, VOD >1.0). Please confirm whether this seasonal pattern is reasonable in those regions.

Response: Thank you for your careful review and the valuable comment. We appreciate your attention to the seemingly unusual seasonal patterns of VOD in specific regions. Regarding the situation in Sweden where the summer VOD (~0.6) is lower than the winter VOD (~0.8), this phenomenon can be reasonably explained by several factors. In Sweden, during the winter, although the vegetation may be less physiologically active, the presence of snow cover can have a significant impact on the VOD measurement. Snow has unique electromagnetic properties, and its high dielectric constant can lead to increased microwave backscattering, which in turn can elevate the VOD value [1]. In contrast, during the summer, although the vegetation is growing, the relatively sparse forest cover, combined with possible effects of soil moisture and vegetation structure changes, might result in a lower VOD compared to the winter with snow cover.

As for the regions in the western US, Alaska, and western Canada where the winter VOD is higher (almost as high as in the Amazon rainforest, VOD >1.0), this can be attributed to the local environmental conditions. A large part of these areas is covered by coniferous forests. These forests are characterized by evergreen trees such as spruce, pine, and fir. The evergreen nature of these trees means that they retain their foliage throughout the winter [2]. Compared to deciduous trees that shed their leaves in winter, the continuous presence of foliage in coniferous forests leads to a relatively stable and high VOD [3]. In addition, the density of coniferous forests in these regions is relatively high in many areas. This high - density vegetation further contributes to the elevated VOD values, making them comparable to those in the Amazon rainforest in terms of magnitude.

In conclusion, based on the local environmental characteristics and the influence mechanisms of various factors on VOD, the observed seasonal patterns in these regions are reasonable and consistent with our understanding of the interaction between the environment and microwave remote sensing signals.

References:

[1] Mätzler C. Applications of the interaction of microwaves with the natural snow cover[J]. Remote sensing reviews, 1987, 2(2): 259-387.

[2] Tian F, Wigneron J P, Ciais P, et al. Coupling of ecosystem-scale plant water storage and leaf phenology observed by satellite[J]. Nature ecology & evolution, 2018, 2(9): 1428-1435.

[3] Jones H G, Vaughan R A. Remote sensing of vegetation: principles, techniques, and applications[M]. OUP Oxford, 2010.

Q: 3) There are large areas of missing data in Russia, China, and Japan during winter, but a specific spot in Russia appears red. Please confirm whether this

pattern is reasonable.

Response: We sincerely thank the reviewer for pointing out this interesting observation. In the final L-VOD product, the data from April 1, 2015, to July 31, 2021, is seamless data obtained through gap - filling model. The VOD product from January 1, 2010, to March 31, 2015, is further processed using a spatiotemporal fusion algorithm based on the seamless data. Both of these datasets maintain spatial consistency with the original satellite products.

As shown in Figure 4, which presents the monthly average of VOD_resmap in January 2021, we can observe that the SMAP satellite product has missing data during this period. Our reconstruction model constructs three - dimensional data on a monthly basis to learn spatiotemporal features. Since the input data (the original SMAP satellite data) has missing values, the reconstruction results also exhibit missing data. This is the reason for the large - scale missing data in Russia, China, and Japan during winter in our product.

The presence of the red spot in Russia can be attributed to the availability of original satellite data in that area. Unlike the surrounding regions in Russia, China, and Japan where there are large - scale data gaps during winter, this particular location has accessible original satellite data. The sensors are able to capture signals in this area, enabling reconstruction model to generate a valid VOD value.

In summary, these spatial patterns are reasonable given the characteristics and spatial consistency of the original data.

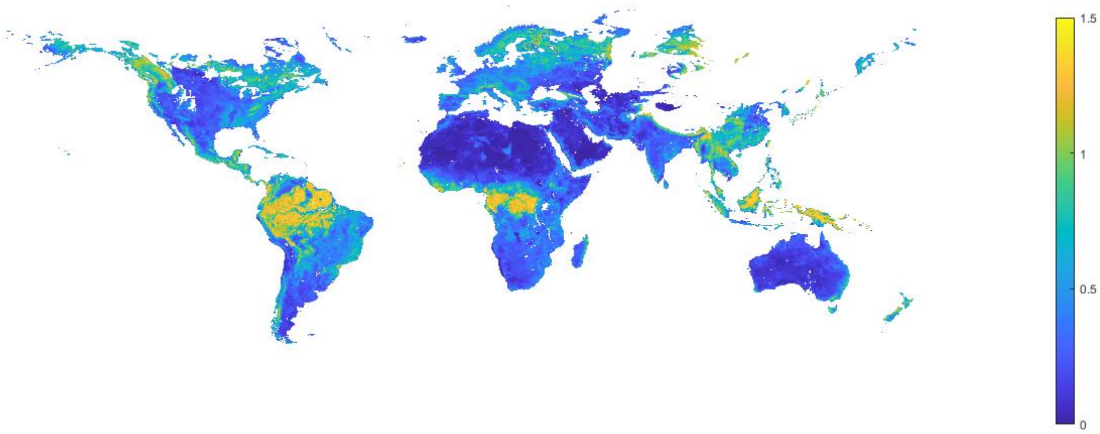

Fig. 4. This chart shows the monthly average of VOD_resmap in January 2021.