

## Comment on **essd-2021-326**

Anonymous Referee #2

---

Referee comment on "A global land aerosol fine-mode fraction dataset (2001–2020) retrieved from MODIS using hybrid physical and deep learning approaches" by Xing Yan et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-326-RC2>, 2021

---

The authors need to address the following major points.

The motivation of this manuscript is to address the following issues raised by the authors, the first issue is the Chen et al (2020) method is not applied on a global scale, the second issue is that 'Zhang et al (2016) noted that satellite –measured multi-spectral reflectance of ground-based data alone was not sufficient to retrieve FMFs with high accuracy' (the meaning of this sentence is not quite clear to me), my interpretation is that the use of (only) spectral information from satellite measurements is not enough for the retrieval of FMF, this two reasons are not really solid enough to continue the work proposed in this manuscript. Both manuscripts mentioned above focus on the original level 2 spatial scale (with a very quick look at these two papers), while this work focuses on the level 3 data, the average of spatial resolution from kilometre to degree can make all problems significantly easier, both technically and scientifically. Even later, the authors point to their own previous publication (Yan et al., 2021b) and claimed 'As shown by Yan et al. (2021b), the global land Phy-based FMF is still unreliable.', in which the authors even started 'seasonal FMF characteristics and trends' analysis using the 'unreliable' dataset. This is quite misleading for the understanding of the motivation for the developments in this manuscript, even from the very first step. I think the work should start from level 2 dataset rather than level 3.

The second major point the authors highlighted is that the method is a combination of physical and deep learning approach, this is also quite confusion. As presented in section 2.4, this dataset is created by the new algorithm described in this manuscript for the first time, however, my feeling is that it is a mixture of previous publications without a clear description of the method itself. Meanwhile, the authors mark this paper as a 'Data description paper' in the submission. I never saw a 'Data description paper' without a clear and solid 'Method description paper' before. The key physical part is the LUT-SDA, described in equation (1). Firstly, I am confused why we have two of here? Second, I can not understand how the combination between the physical method and deep learning is achieved. To my personal view, it is still a deep learning method with certain parameters from some physical derivation, however, your major input, aerosol optical thickness, is derived from a physical model (MODIS retrieval algorithm) as well, you cannot claim it is a

combination of physical and deep learning method because some inputs for the deep learning is from a product derived from a physical method.

The third major point is the comparison between different satellite products, the authors need to be aware that it is really the same parameter in the comparison or not, 'bad' agreements between satellite FMF products and the AERONET FMF product do not reveal anything because these FMF are not the same FMF due to different assumptions in particle size distribution and the 'cutting criteria' in the level 2 retrieval process, the FMF derived from this paper is somehow with knowledge (maybe certain inputs as well, not sure about it) from AERONET, it is not surprise at all to have a better agreement with AERONET measurements later. It really makes no sense to include the MODIS comparison since it is already removed in the new version of dataset. The remove of FMF product in MODIS dataset also indicates how uncertain such a parameter can be.

The fourth major point is the application of this dataset, my personal view is that it is too early to sell it as a dataset which is mutual enough for a trend analysis, especially since the application of (both fine and coarse as total) aerosol optical thickness is still quite questionable. Even at the last part of the manuscript, the authors claim some significant trends from the new satellite product are not revealed by AERONET measurements due to the scale issue, this simply reveals the limited representative of these AERONET sites in your regions, rather than anything with respect to the satellite data quality.