RC2: 'Comment on essd-2024-266', Anonymous Referee #2, 07 Nov 2024

Major concern

The paper proposes a new dataset of spectral indices for Europe, based on Landsat Analysis Ready Data version 2 (ARD V2) for the period 2000-2022. The dataset is fully available to the public on Zenodo, and the code is accessible on GitHub, greatly facilitating the review process. The paper is well-written. The article's strengths are the strategies followed for building the bi-monthly composites (what authors call Tier 1) and the extensive validation performed. The dataset validation includes ground truth data from several sources, such as the Land Use and Cover Area Frame Survey (LUCAS), European cropland surveys (Edlinger papers), and Eurostat's tillage area statistics. These ground truth datasets are primarily used to assess the utility of the temporal-spectral feature (i.e., Tier 3, Figures 5, 6, 7, 8, and 9). Additionally, the paper evaluates annual trends in long-term temporal-spectral indices (i.e., Tier 4, Figures 10, 11, 12, 13, 14, and 15). Finally, a classification (land cover) and regression models (soil organ carbon) are trained using the dataset presented in this paper as input data, with LUCAS data as the target variable (Figures 18, 19, 20, and Tables 5, 6, 7). A statistical comparison of classes of land cover with other Landsat-derived products (i.e., EcoDataCube, Witjes et al. (2023)) was also carried out (Figure 16). I have some concerns about the high correlation between this paper and the Consoli et al. (2024) paper, which is available only as a preprint and published just one month ago (https://www.researchsquare.com/article/rs-4465582/v1). However, if the editor thinks this similarity is irrelevant, I would recommend a minor revision (see comments below).

This paper, "Bi-monthly and Annual Landsat Spectral Indices for Europe 2000-2022," is very similar to the dataset presented in Consoli et al.'s (2024) "Global Bi-monthly Landsat Aggregated Product 1997-2022. " Both papers have not been published yet but are available as preprints. This journal states: "The editors encourage submissions on original data or data collections of sufficient quality that have potential to contribute to these aims." The Consoli et al. (2024) paper focuses on Tier 1, i.e., generating plausible aggregation and performing gap-filling in a single step by simply adjusting a convolution kernel. It introduces a new algorithm, TSIRF, which appears to outperform traditional methods like Savitzky-Golay in terms of both computation time and accuracy. The editor should consider that many authors of the Consoli et al. (2024) paper are also authors of this paper under review. The Consoli et al. (2024) paper is available as a preprint here: https://www.researchsquare.com/article/rs-4465582/v1. The primary difference between these two papers lies in the estimation of spectral indices, specifically Tier 2 (Vanilla Spectral Index), Tier 3 (Temporal-Spectral Index), and Tier 4 (Long-term Temporal-Spectral Index). Essentially, this paper builds on the data presented by Consoli et al. (2024). For instance, on Page 3, Line 24 the Bimonthly aggregated cloud-optimized bands must be the same that Consoli et al. (2024) paper. I believe that an ESSD brief communication, as an extension of the Consoli et al. (2024) paper—published after its

acceptance—would be more appropriate for this paper, as the principal novelty is the estimation of complex temporal-spectral indices and their validation.

RE: We acknowledge the connection between our work and Consoli et al. (2024) and the overlap in authorship due to our shared institute. However, despite this overlap, the two works provide distinct contributions, each addressing different scientific objectives.

Consoli et al. (2024) focuses on developing the TSIRF computational framework for efficient global aggregation and gap-filling of Landsat data, demonstrating this approach with a global bi-monthly aggregation of raw spectral bands (i.e., the global version of Tier 1 products referenced in our paper). In contrast, our study emphasizes Europe, with a focus on the preparation of a comprehensive data cube of biophysical indices. This includes bi-monthly time series (Tier 2), aggregated annual series (Tier 3), and long-term temporal signatures (Tier 4). Additionally, we conduct extensive plausibility checks and machine learning applicability tests to enhance the utility and relevance of the dataset for targeted applications within Europe.

Leveraging existing datasets is a common and essential practice in advancing data processing. While this work builds on the data produced in Consoli et al. (2024), we believe this does not compromise its originality. Furthermore, all text in this manuscript, aside from standard abbreviations and input data descriptions, is original. The work of Consoli et al. (2024) has been consistently cited to ensure transparency and proper acknowledgment. We have also revised the manuscript to further clarify and highlight these distinctions (see Section 2.3.1 on Page 6).

Experiments demonstrating the time-series reconstruction performance in Tier-1 would be highly relevant for readers.

RE: We have incorporated performance metrics for time-series reconstruction in continental Europe, detailed in Section 3.1 and summarized in Table 4 (Page 15), with the corresponding methodology described in Section 2.4.1 (Page 10). Additional details on the testing materials and performance across various land cover classes are provided in the supplementary material to ensure the manuscript remains concise. For a broader global analysis of time-series reconstruction performance, we refer to Consoli et al. (2024). The code for this experiment has also been updated in the corresponding GitHub repository.

Line 10, Page 2: "which cover approximately 67% of the Earth's surface" – Please either remove this phrase or provide an accurate citation.

RE: This phrase has been removed as it inaccurately quantifies the data (see Page 3 Line 16-17). In our experiments, we observed that valid pixels account for approximately 67% of the total for several years between 1997 and 2022. However, this figure does not specifically refer to cloudy pixels—it represents valid pixels, not invalid ones. Additionally, not all invalid pixels are due to cloud cover; other factors, such as atmospheric interference, snow cover and polar night in northern areas also contribute. We have removed this phrase to avoid any misunderstanding.

Line 20, Page 5: What is SIRCLE? This term is not mentioned in either Consoli et al. (2024) or the scikit-map documentation. Please clarify the difference between SIRCLE and TSIRF.

RE: SIRCLE was the initial name of the framework at the time of submission but has since been renamed to TSIRF. We have now updated the manuscript to reflect this change. Please refer to Section 2.2 on Page 5 for details.

Line 5, Page 6: SIRCLE is cited as part of Consoli et al. (2024), but the acronym only appears in Figure 3 without further reference. Please ensure that SIRCLE is introduced properly if it is central to the methodology.

RE: We sincerely thank the reviewer for spotting this. The use of the acronym SIRCLE in Figure 3 of Consoli et al. (2024) was a typographical error. This will be corrected in Consoli et al. (2024)'s final published version to align with the updated terminology.

Line 20, Page 3 and Line 5, Page 36: Referring to Witjes et al. (2023) as "limited" for presenting quarterly rather than bi-monthly maps may be a bit misleading. Especially considering the high amount of missing data at high latitudes (see Figure 1 in this paper), a more nuanced wording may be more appropriate.

RE: We acknowledge that referring to EcoDataCube.eu as "limited" may not fully capture its strengths and nuances. While our work provides a finer temporal resolution and additional data products, we recognize the challenges posed by high latitudes in achieving higher resolutions. We revised the wording to more accurately reflect the contributions of EcoDataCube.eu while highlighting the distinctions of our dataset. Please refer to Page 3 Line 21-23 for details.

Line 5, Page 37: What is HLS? Please provide a brief explanation of the acronym.

RE: We have now clarified the acronym for HLS in the revised manuscript. The updated text provides a brief explanation of HLS as the Harmonized Landsat Sentinel product, detailing its data sources and relevance. Please refer to Page 38 Line 34-35 for details.

Line 5, Page 37: The limitations section should address the known limitations of ARD V2, such as cloud detection accuracy over Europe and the challenges in harmonizing Landsat products.

RE**:** Added. The limitations section now includes the known challenges of ARD V2, such as the surface reflectance normalization used for harmonizing ARD V2, and its struggles in processing winter time images. Please refer to Section 4.4 Paragraph 1-2 on Page 41-42 for details.