

Interactive comment on “A remote sensing-based dataset to characterize the ecosystem functioning and functional diversity of a Biosphere Reserve: Sierra Nevada (SE Spain)” by Beatriz P. Cazorla et al.

Anonymous Referee #2

Received and published: 26 March 2020

Are the data and methods presented new? -An interesting approach is presented for inter-annual heterogeneity; it is left open why for assessing the spatial variability a certain kernel size had been chosen

Is there any potential of the data being useful in the future? -In principle yes, however, there are details missing, see next

Are methods and materials described in sufficient detail? - No. Why is the kernel size 4x4? How have borderline pixels be processed with the kernel? (kernel processed

C1

raster layer have same extension) How variable are the quartile boundaries (could you name a standard deviation?)

Are any references/citations to other data sets or articles missing or inappropriate? -reference/URL to the database REDIAM is missing, also, which particular datasets have been employed from it; by what data got the MODIS data clipped/masked?

Is the article itself appropriate to support the publication of a data set? - yes with respect to gain an understanding of the data. The article does not provide necessary information to re-use the data: the legend for EFTs is part of Fig 2; the values of the EFTs do not correspond to the values in the TIFs (there they are 1-64 encoded)

Check the data quality: Is the data set accessible via the given identifier? -yes Is the data set complete? -yes Are error estimates and sources of errors given (and discussed in the article)? - well, not error but there is no reference to variability eg the means of internal quartiles given Are the accuracy, calibration, processing, etc. state of the art? - The article employes community-"standard" pre-processed data; however, it does not provide accuracy information of intermediate processing steps. Also, the derivation of spatial heterogeneity, the chosen size of the kernel and how this affects the results is not discussed

Are common standards used for comparison? - the resulting data are not compared Is the data set significant – unique, useful, and complete? -The data set is useful

Consider article and data set: Are there any inconsistencies within these, implausible assertions or data, or noticeable problems which would suggest the data are erroneous (or worse). - using a kernel to derive values I would have expected that the resulting layer is smaller in size than the input layer, unless some "mirroring" is done to extend the input layer in size. The article does not provide any information on how this was handled

If possible, apply tests (e.g. statistics). - looking up the TIFs with standard GIS software

C2

(QGIS) did not reveal any problems. The histograms of values seem ok, although because of missing legend they could not be really interpreted

Is the data set itself of high quality? Check the presentation quality: Is the data set usable in its current format and size? -yes, the GeoTIFF is a well accepted and documented file format

Are the formal metadata appropriate? - No, I am unable to discover any formal metadata. The GeoTIFF come with some metadata in their header, but do require specialized software for extraction, eg. of the bounding box or employed projection. additional TFW file would be readable with common editors. Additional formal metadata is missing.

Check the publication: Is the length of the article appropriate? - given, that it is a data publication, the article dwells much on discussion of the application/biodiversity/structure but is much shorter when it comes to describing data and methodology

Is the overall structure of the article well structured and clear? -yes

Is the language consistent and precise? -there are a few language errors but the article is language wise in good shape

Are mathematical formulae, symbols, abbreviations, and units correctly defined and used? - Eq.3 uses X any Y without explicit definition; this equation does not provide additional information content Are figures and tables correct and of high quality? Quality is mostly acceptable, in Fig.2, part 3 the legend is hardly readable

Finally: By reading the article and downloading the data set, would you be able to understand and (re-)use the data set in the future? -No, eg. the EFT type as encoded in the TIFs cannot be interpreted

Uniqueness: It should not be possible to replicate the experiment or observation on a routine basis. - all resulting data can be reproduced as the primary source is generally

C3

available However, the derivation needs expertise with GIS/remote sensing software, and a target audience of ecologists is usually easier reached with data products which are deemed useful for such clientele

The introduced methods are not trivial nor obvious, however, would benefit from a discussion why certain approaches had been taken (kernel size eg.) The data seem complete. All derived data sets are provided (annual data), also the summary data. In theory one could re-calculate all results (if eg. interval boundaries were to be now known, EVI_max).

I would request information on hardware and software used to derive products (algorithmic deviations) Also, to reproduce the data information on masking/clipping the covered regions is necessary but absent. (which dataset, which method)

Technical details:

line24: imagery do not provide a continuous characterization as reflectance is integrated per pixel

line 26: from 2001 to 2018

line 79 not the EFT approach has exp. grown but the application of EFT approaches

line 137 EFT seasonal curve: the term has not been introduced properly; I presume it refers to the 23 measurements taken per year, please clarify

line 146: one cannot understand the present derivation as the methodology is referred to another article; worse, the authors write of a "similar" approach without making clear how/where they differ

line 147 EVI_DMAX: unclear, whether you chose the intervals according to the definition of the seasons or you derived them and they turned out to coincide with the seasons; please clarify

line 149-150: the derivation of quartile borders was understandable only after con-

C4

sulting the reference. How stable are the boundaries, that is, provide a standard deviation for each mean Table 1: values cannot be reproduced nor checked, e.g. EVI_Mean_2001_C006_MOD13Q1_Pixel232.tif shows values between 11.5-4471.9 (QGIS), table 1 reports 75% values are less than 0.241 EVI_mean: problem with "sealed" class boundaries: derivation relies on mean of a 18y period. If say, you want to show the time series of 2001-2020, would you need to do the derivation of the boundaries or "extrapolate" from 2018? Table 1, EVI_Max: values of 1-4 do not correspond to values found in TIFs (1-12)

line 159: justification for a 4x4 kernel? Why not 3x3 or 5x5? Could the kernel be dependend on the question being asked? How have borderline pixels be processed/why eg share richness and inter-annual mode the same borders?

line 359: database is maintained

line 360: please include a reference/URL to the database REDIAM, also, indicate which datasets of REDIAM have been included in your work

Fig 2.1; <https://lpdaac.usgs.gov/products/mod13q1v006/> states 250m GSD, not 230m. Fig 2.2: the mean is not the area under the curve, but the area normalized by the range; there is no curve at all but 23 discrete values/year Fig 2.4: the legend is crucial for re-using data but is not provided as individual data (eg. numerical values corresponding to a class, or pseudo color code for GoogleEarth); at present, the TIF files for eg EFTs show values between 1-64; how to map to your classes?

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2019-198>, 2020.