Comment on essd-2021-146
Anonymous Referee #1

Referee comment on "Fish and cephalopods monitoring on the Bay of Biscay and Celtic Sea continental shelves" by Pascal Laffargue et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-146-RC1, 2021

First of all I have to acknowledge that I am not familiar with this type of data-papers. Reading it as a regular review, I missed the scientific questions and findings in the abstract and only then understood that the only purpose of the paper was to describe the data set.

I find making only the data available an interesting concept, and I find it very good that the authors are putting the effort in making the whole EVHOE-set available. As a whole set this is definitely an interesting data set, that can be used in numerous fisheries and ecological studies.

However, a large part of the dataset is (of course) already for many years available via the ICES DATRAS database and most of the survey design is already (better) described in the ICES manual (ICES, 2017). Some of the sentences/sections in the MS are straight from that manual. The ICES database is (more or less) restricted to uploading a fixed set of fish, cephalopods and commercial invertebrate species, which means a large part of the invertebrate data collected on the EVHOE survey is not available via Datras. So reading the MS I expected that the dataset described in the MS would include the information that (currently) can’t be upload to DATRAS. However, on line 61-62, it is stated that the non-commercial invertebrate are actually not yet included. Also the spring data, which is not available in Datras, is not included in the current data set. The only part currently different from the Datras database seems to be the first 7 years (1987-1995) which as is stated in the document are complicated to be combined owing to the vessel change. Which is the reason that these older data are not available in Datras as they are in the assessments not used in combination with the recent data. The authors further state on line 266 that a strength of the data set is the additional data. I agree that that type of data would make the use of the dataset more interesting. However, neither that data is currently provided. While, I expect that the hydrological data is already available via ICES, as is the litter data (mentioned in line 57). Concluding, at the moment I do not see the added value of this data set over the data set already available at ICES. So maybe better to publish this dataset when these additional data become available instead of promising these data
without a planned date for making them available. If the editors think it is worth publishing the current dataset, I advise to make clearer in the text which data is currently part of the dataset and what the difference is with the ICES dataset.

A worry I have related to the dataset is how the authors plan to maintain their local dataset, the one at ices and the current one provided. From experience, I know that on a regular basis things within the ICES database are corrected/updated. Which is likely also the case when people start using this dataset. So how will the authors deal with that, as it undesirable that the content of the similar datasets start to deviate.

I made small corrections/additions in the pdf and added some comments in it as well. I only made small corrections, however I constantly had the urge to rewrite large sections or reorder sections in the MS. I believe rewriting large parts of the text, including putting some of the things in tables, taking out duplications and also taking out information not necessary to understanding the data set and its use, could make the text of the MS shorter and much better readable. As its purpose is to describe the dataset, there is no need for a lengthy document, a short and comprehensive text would according to me better fit the purpose.

I am not convinced that the figures currently used to showcase parts of the content of the data set are the most relevant or the most useful choices. However I can't come up with better solutions at the moment.

Overall: the concept of publishing the full survey data set seems interesting and relevant. The currently provided dataset is not really unique and the MS would in my opinion require a major rewrite to make it a better reference document than the current ICES manual. So I advise a large revision, and maybe even extending the publication until a larger data set with unique data currently not available yet can be provided.

Please also note the supplement to this comment: https://essd.copernicus.org/preprints/essd-2021-146/essd-2021-146-RC1-supplement.pdf