

Interactive comment on “A dataset of distributed global water withdrawal from 1960 to 2017” by Denghua Yan et al.

Denghua Yan et al.

lixn0555@163.com

Received and published: 1 April 2020

Reply to the reviewer comments Responses to Reviewer #3: Dear Reviewer: Thank you for your comments concerning our manuscript entitled “A dataset of distributed global water withdrawal from 1960 to 2017” (ID: ESSD-2019-224). Those comments are all valuable and very helpful for revising and improving our paper, as well as the important guiding significance to our researches. We have studied comments carefully and modified the dataset which we hope meet with approval. The main corrections in the dataset and the responds to your comments are as follows:

Line-by-line comments: Referee comment: 1. The Figshare link allows users easy access but the presented data remain almost useless. THE FILES CONTAIN NO META-

C1

DATA!! A user finds no column headings, no introductory lines (rows) of explanation or attribution, nothing to help the user know what he or she looks at. More information comes from the filename (e.g. Data_W_Sources) than from the files themselves. Most folders include only maps, not referenced or sorted except by date. Data availability section in the manuscript (lines 221 to 235) provides only minimal guidance. A large set of guidelines and taxonomies exist for surface water hydrology - the authors have neither used nor referenced any of that. Authors should look at almost any other ESSD product to find very good examples of how to present and record metadata. Not acceptable in present form. Entire text very weak, disorganized, confusing. This review highlights dozens of issues below. Basically, authors show some skill in mapping and interpolation, but have not shown understanding or skill with data or description to help users. Does not meet standards expected for a top data journal. Author's response: We introduced the attribute information of the dataset in the manuscript, which may not be detailed enough. Based on your comments, we will add metadata and provide more detailed data guidance in the manuscript. Author's changes to the manuscript: We will add the metadata, column headings, lines of explanation and we will provide more detailed data guidance in the data availability section.

Referee comment: 2. Line 29 - “United Nations World Water Development Report 2018”: after authors have cited this report once (line 24) they should use a standard acronym. The report itself suggests a proper citation, which these authors should use. Author's response: Thank you. We will include a standard acronym and reference. Author's changes to the manuscript: In the revision, we will use a standard acronym and format of reference.

Referee comment: 3. Line 33 - “regularity of water space distribution” - What does this mean? Regularity in time or space? Regularity in classification? This entire introductory paragraph basically repeats the same issues two or three times. Whatever motivation might exist remains hidden by or obscured by random language. Author's response: What we want to express is that by studying changes in global water with-

C2

drawal, accurately analyzing the structure of water withdrawal intensity, revealing the regional characteristics and spatial distribution of water withdrawal. Author's changes to the manuscript: That sentence will be changed to "regularity of the spatial and temporal distribution of water withdrawal".

Referee comment: 4. Line 35 - "withdrawal intensity is the main forms of water consumption" - Singular / plural confusion, happens frequently throughout the manuscript, often confuses the reader about what authors intend. Author's response: We will send the manuscript to a professional linguist for editing to ensure the readability of the manuscript. Author's changes to the manuscript: That sentence will be changed to "withdrawal intensity is the main form of water consumption".

Referee comment: 5. Lines 35 to 48: this entire paragraph reminds user that authors have produced a geographic data product, not a hydrologic data product. But, from the title, authors implied a hydrological product? Authors should make clear their intent, their tools and their skills. This is a GIS / mapping exercise, not an effort to produce a valid global hydrologic data product. Author's response: I think the hydrological data must contain the corresponding geographic information to be meaningful. There is no clear boundary between geographic data product and hydrologic data product, and all hydrologic data products need the support of geographic data. Author's changes to the manuscript: We will make clear the intent of manuscript in introduction.

Referee comment: 6. Line 54 - "errors in the statistics and collection of the original water withdrawal data": most users will know that accurate data on water supply or use remains highly restricted and highly distorted by most countries. Here the hydrology / geography confusion arises clearly: what use does a higher spatial resolution product have if the underlying data remain almost completely unreliable. These authors with their spatial GIS skills never address the fundamental issue of the quality or availability of the basic hydrological data? 7. Line 56 - "With the reference of the official data available, the accuracy of the data set is sufficient to meet the current research": yes, the authors can assemble and provide nice maps but the fundamental data remain almost

C3

useless. 8. Line 58 - "products that can reflect the spatial and temporal changes of water withdrawal in the world": No, strongly disagree. New maps from bad data, not an overall improvement. Author's response: First, all data cannot be guaranteed to be 100% accurate, but it is considered reliable as long as the data is within the allowed error range. There is almost no continuous long series of water withdrawal data globally, which leads to underestimation or overestimation in the analysis of global water consumption. Without global water withdrawal data, biases can arise in assessing social, economic and ecological water requirement and in developing water management strategies. Moreover, water withdrawal is not just a physical concept, and it contains a large number of socio-economic implications, so it is impossible for any research to calculate the water withdrawal of the world or even a region with 100% accuracy. The significance of this study is not to provide a set of datasets with 100% accuracy, but to provide a complete and credible data product based on the fact that there is currently no global water withdrawal data set. Of course, with the gradual increase of data collection, the accuracy of this data product can continue to be improved in the future. Author's changes to the manuscript: None

Referee comment: 9. Line 60 - "improvement of the accuracy of the original data": improvement of spatial resolution does not equal improvement of accuracy. Author's response: I acknowledge that the reviewer's view that "improvement of spatial resolution does not equal improvement of accuracy" is correct. It is clear, however, that the reviewer did not carefully review the manuscript. What we said on line 60 of the manuscript is that in the future, when the government releases data with better accuracy (e.g., water withdrawal data by region), we can further improve the accuracy of the dataset. Author's changes to the manuscript: None

Referee comment: 10. Line 62, Table 1: The FAO data products are notoriously unreliable. Likewise for other UN sources. Chinese sources seem perhaps interesting and useful but authors provide no publications, validations, or reliability analyses? Author's response: Although the data of FAO and other international organizations do

C4

have some deviations from the real data, the data is to a certain extent credible. A large amount of current water research worldwide is also based on existing FAO and other data. In addition, for more than 200 countries, we have collected a large amount of data released by national government departments. The credibility of these data is often higher than that of FAO. The introduction of these data sources and download links have been uploaded to the database of this study (lines 207 to 208). According to your opinion, we will supplement references in the manuscript with Chinese sources, validations, and reliability analyses. Author's changes to the manuscript: We will supplement references in the manuscript with Chinese sources, validations, and reliability analyses

Referee comment: 11. Lines 175 to 182: Here the authors provide a weak description of GlobeLand30, the Chinese remote sensing product (based on USGS Landsat). No details, no references, no uncertainties, nothing accessible or open access that can convince other users of the quality of this product. Author's response: The classified images used for the development of GlobeLand30-2010 data are mainly 30-meter multispectral images, including the United States Landsat TM5, ETM+ multispectral images, and the China Environmental Disaster Reduction Satellite (HJ-1). In the verification process, a total of more than 150,000 test samples in 9 categories were selected for accuracy evaluation. The overall accuracy of GlobeLand30-2010 data was 83.51%, and the Kappa coefficient was 0.78. The method of obtaining data and listing it in the manuscript, reviewers can download the data through this website. Since this set of data is not original, it was not uploaded to the database. Author's changes to the manuscript: We will add details, references and uncertainties about GlobeLand30 in the manuscript.

Referee comment: 12. Line 182 - "water is only used on artificial surface and cultivated land": Reader / user never learns what the authors mean by the term 'artificial surface' (reservoir?, impervious urban pavements?, compacted drought-impacted land with high run-off?) but the assumption stated in this phrase is almost certainly false.

C5

The authors demonstrate no understanding whatsoever of surface hydrology! Author's response: Artificial surface refers to the surface formed by artificial construction activities, including various residential areas such as towns, industries, mines, and transportation facilities. The artificial surface mainly includes industrial and domestic water, and the cultivated land mainly includes agricultural water, so we are confident that the assumptions we have listed are correct. Author's changes to the manuscript: We will add the definition of artificial surface.

Referee comment: 13. Line 206 and following - Technical validation: Validation consists of comparing their interpolated products to two countries, India and China, for which they have data of higher spatial resolution. But from exactly the same FAO/UN national report source data! How does that represent independent data? Many other data products, including those derived from satellites (e.g LandSat or others), exist, for which these authors could at least construct some intercomparison maps. Not one quantitative assessment (run-off, storage, retention, etc.) in useful hydrological terms anywhere in the entire manuscript. Validation, uncertainty, reliability completely absent. Author's response: When we verified the data, we mainly verified our interpolation and extrapolation methods. For China and India, their data sources are from the data released by their governments, not the same data source (lines 211 to 213). We first assume that there are no data for some years, then use our method to obtain the data for these years and compare them with real data. Of course, the comparison in the manuscript is indeed insufficient, and we will add more data verification in the future. Author's changes to the manuscript: We will add more detailed data validation.

Referee comment: 14. Lines 64 to 205, Methods: This is basically a GIS, data-filling, data interpolation exercise, to fill in missing national reports and then to fill in spatial gaps. The authors, who may have skill in using these products and tools, give no indication that they understand, much less question, the underlying data. I repeat: better maps of unreliable data do not result in better data. 15. Line 216 - "dataset derived from the existing water withdrawal data is accurate" Authors may believe this,

C6

but due to complete ignorance of surface hydrology, reader /user can only accept that maps might prove faithful to national reports but that underlying data has neither validity nor accuracy. Authors can neither make nor verify any statement about accuracy. Large community of hydrologists sharing data through ESSD will not find this product in any way useful. 16. Line 237 - "fill the blanks of complete water withdrawal sequence data, enhance the accuracy and spatial variability of water withdrawal data, and can reflect the spaceapace changes of water withdrawal." Complete = no. Accuracy = no. Spatial variability = a product only of their interpolation tools, never verified. Space-apace = ???? Author's response: As we mentioned earlier, by collecting and compiling a large amount of data in more than 200 countries, we can trust that the water withdrawal dataset based on government data / FAO data and other sources is accurate. We will add more detailed verification and instructions in subsequent revisions. Author's changes to the manuscript: We will add more detailed verification and instructions in subsequent revisions.

Referee comment: 17. Line 270 and following: 13 references, 4 of which represent unreviewed UN or private technical reports. If well and carefully done, a product like this could have reference to dozens of primary and validation sources and to literally hundreds of hydrological and social applications. The authors seem aware of none of that, not even the vast remote sensing or interpolation literature. Author's changes to the manuscript: We will add references in the manuscript.

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2019-224>, 2020.