Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-521-AC1, 2021

Hydrology and Earth System Sciences
Discussions

Interactive comment on "CABra: a novel large-sample dataset for Brazilian catchments" *by* André Almagro et al.

André Almagro et al.

andre.almagro@gmail.com

Received and published: 18 February 2021

Response to Anonymous Referee #2

Comment #1: This study has collected, synthesized, organized, and made available more than 100 topography, climate, streamflow, groundwater, soil, geology, land-use/land cover, and hydrological disturbance attributes for 735 catchments in Brazil. The dataset is valuable for many hydrological and other relevant scientific studies. However, this paper does not provide any in-depth/innovative scientific analysis based on the established dataset. It looks more like a dataset paper rather than a scientific research paper. I would like to recommend a rejection in HESS but would like to suggest a transfer to the journals focusing on data such as Earth System Science Data,

Scientific Data - Nature, etc.

Author's response #1: We would like to thank the referee for the insightful comments, suggestions, and kind words in support of our manuscript. The manuscript has been revised in accordance with your comments, which were highly insightful and enabled us to improve the quality of our manuscript. With the updated manuscript, we took care to answer the reviewer's questions, comments, and requests. We hope that the revisions in the manuscript and our accompanying responses will now meet the requirements for publication. Thank you again for your consideration.

Regarding to the transfer to another journal focusing on data, we disagree with referee. Our dataset is focused on hydrology, providing a large variety of catchment attributes that may contribute to the advancement of hydrological modelling, process concepts, besides the fact that all provided data is analysed in the study. The main inspiration for the CABra dataset is the MOPEX (Duan et al., 2006) and CAMELS-US datasets (Addor et al., 2017), which were published in journals focused in hydrology. The last one also collected, organized, synthetised, and analysed a wide range of catchment attributes for 671 catchments in continental US, being universally used and cited since its publication. We also have seen many other datasets being published in Hydrology and Earth System Sciences journal. Oubeidillah et al. (2014) made available a dataset of post-calibrated model parameters for hydroclimate impact assessment, where authors collected and organized data including meteorological forcings, soil, land-cover, vegetation and elevation from the best-available source at that time. Siebert et al. (2015) provided a global dataset of historical evolution of irrigated areas, collected from sub-national irrigations statistics from a variety of sources. For German territory, Zink et al. (2017) provided a high-resolution dataset of land surface variables. A new global land-based product of precipitation was made available by Contractor et al. (2020), by merging multiple archives of in situ data. The well-known ERA-Interim/Land reanalysis dataset (Balsamo et al., 2015) was also published in HESS journal, providing one of the biggest archives on land surface variables in the world. Finally, the CAMELS-CL

(Alvarez-Garreton et al., 2018), which is a catchment-based dataset of geophysical attributes, similar to CABra, for Chile was also published in HESS journal. Given the above we believe that CABra dataset is within the scope of HESS journal for publication, since we have not only organized existent data, but also presented novel products (e.g., hydrological disturbance index, gridded climate ensemble, and potential evapotranspiration) and conducted deep analysis of the data.

Comment #2: More specific comments are as follows: 1) Some attributes are from first-hand investigation data (e.g., streamflow, meteorology, geology, catchment delineation, etc.), and a lot of attributes are extracted from global re-analysis dataset. The former attributes owe higher accuracy and are really valuable for the community to do a variety of hydrometeorological modeling and assessment. The latter attributes are also useful to reduce other uses' time consumption to re-prepare them. But I would like to suggest providing the delineated catchment-based digital map and/or tabular dataset. The spatially distributed dataset is important for distributed modeling.

Author's response #2: Thank you for the comment. We agree with referee that the spatially distributed information about the attributes provided by CABra is of great importance for distributed modelling. But for such large dataset as CABra is quite unworkable to provide the catchment-based digital map for 735 catchments due to its large amount of data. To attend referee's requirement, we will add a table (attached to the dataset files in Zenodo and attached in this response) indicating the link to download each of the digital maps used for CABra's development. Since we already provide the catchment boundary, it is easy to the user to download and clip the digital map using a GIS application.

Comment #3: 2) For groundwater attribute: spatial variation of groundwater table is subject to complicated driving factors. Fan et al.'s (2013) can be useful for a regional scale analysis for pattern recognition but would have a lot of uncertainties at catchment scale. The observed well observation dataset would be more useful for such a large-sample catchment dataset effort.

Author's response #3: Thank you for your comment and suggestion. We will add to the dataset the static and dynamic levels observed in 2010, which is the last year from the daily timeseries of climate and streamflow in CABra dataset (Figure R1). We have obtained the wells observations at the Geological Survey of Brazil – CPRM database for groundwater, the Groundwater Information System – SIAGAS (http://siagasweb.cprm.gov.br/).

Comment #4: 3) Typos to be checked: L44, than? L301 the all the year? L303 be showed seen?

Author's response #4: Thanks for noting! We removed "than" at L44 (L45 of track change revised manuscript). We corrected "through the all the year" to "through the year" at L301 (L338 of track change revised manuscript). We removed "showed" at L303 (L340 of track change revised manuscript).

REFERENCES

Addor, N., Newman, A. J., Mizukami, N. and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, Hydrol. Earth Syst. Sci., 21, 5293–5313, doi:10.5194/hess-21-5293-2017, 2017.

Alvarez-Garreton, C., Mendoza, P. A., Pablo Boisier, J., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J. and Ayala, A.: The CAMELS-CL dataset: Catchment attributes and meteorology for large sample studies-Chile dataset, Hydrol. Earth Syst. Sci., 22(11), 5817–5846, doi:10.5194/hess-22-5817-2018, 2018.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Munõz-Sabater, J., Pappenberger, F., De Rosnay, P., Stockdale, T. and Vitart, F.: ERA-Interim/Land: A global land surface reanalysis data set, Hydrol. Earth Syst. Sci., 19(1), 389–407, doi:10.5194/hess-19-389-2015, 2015.

Contractor, S., Donat, M. G., Alexander, L. V., Ziese, M., Meyer-Christoffer, A., Schnei-

der, U., Rustemeier, E., Becker, A., Durre, I. and Vose, R. S.: Rainfall Estimates on a Gridded Network (REGEN) - A global land-based gridded dataset of daily precipitation from 1950 to 2016, Hydrol. Earth Syst. Sci., 24(2), 919–943, doi:10.5194/hess-24-919-2020, 2020.

Duan, Q., Schaake, J., Andréassian, V., Franks, S., Goteti, G., Gupta, H. V., Gusev, Y. M., Habets, F., Hall, a., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O. N., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T. and Wood, E. F.: Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, J. Hydrol., 320(1–2), 3–17, doi:10.1016/j.jhydrol.2005.07.031, 2006.

Oubeidillah, A. A., Kao, S. C., Ashfaq, M., Naz, B. S. and Tootle, G.: A large-scale, high-resolution hydrological model parameter data set for climate change impact assessment for the conterminous US, Hydrol. Earth Syst. Sci., 18(1), 67–84, doi:10.5194/hess-18-67-2014, 2014.

Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N. and Scanlon, B. R.: A global data set of the extent of irrigated land from 1900 to 2005, Hydrol. Earth Syst. Sci., 19(3), 1521–1545, doi:10.5194/hess-19-1521-2015, 2015.

Zink, M., Kumar, R., Cuntz, M. and Samaniego, L.: A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty, Hydrol. Earth Syst. Sci., 21(3), 1769–1790, doi:10.5194/hess-21-1769-2017, 2017.

Please also note the supplement to this comment:
https://hess.copernicus.org/preprints/hess-2020-521/hess-2020-521-AC1-supplement.pdf

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-521, 2020.

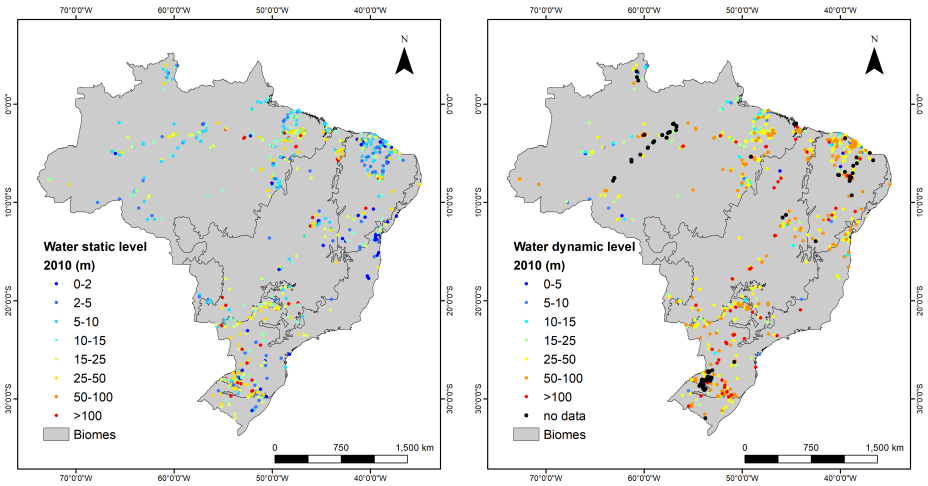

Fig. 1. Figure R1