



Comment on **essd-2021-214**

Anonymous Referee #1

Referee comment on "Global distribution of wastewater treatment plants and their released effluents into rivers and streams" by Heloisa Ehalt Macedo et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-214-RC1>, 2021

General comments

The study combines multiple existing wastewater treatment plant (WWTP) datasets, supplementing available information with additional (inferred) plant characteristics, to produce a consistent and coherent global dataset of wastewater treatment plants. Information captured within the dataset includes both geographical location and important plant characteristics such as flow rates and treatment level. The authors calculate an average wastewater treatment plant dilution factor, and link WWTP outfall locations to the river network, to calculate the length of the stream network downstream of WWTPs as a basic indicator for identifying areas of potential environmental concern.

Overall, I believe this work to be both of high quality and interest for the scientific community. The manuscript clearly describes the dataset and is mostly well written, although I would recommend some edits to improve readability (particularly shortening sentences!). Therefore, after some revisions, I would recommend publication in ESSD. Please see my detailed comments below for my recommendations on how the manuscript could be further improved.

Specific comments

- I would revise the title to be more specific to the contents of the paper and dataset. "Their released effluents" is quite generic – for example this could suggest there is directly data on effluent quality (i.e. pollutant loads or concentrations) which is not contained within the dataset. Keep the focus on the spatial distribution and the core aspects of the dataset: volumes, populations served and treatment level.
- It is stated in the abstract that WWTP information is particularly lacking in developing countries (which is true). This suggested that the dataset would improve this. However, the presented dataset is still vastly dominated by information from developed countries (particularly the USA and Europe with around 2/3 of the WWTP). Additionally, looking to the comparison to Jones et al., 2021; there is still a significant volume of wastewater treatment in the "remaining countries" that is not contained within the dataset. I

appreciate the lack of data availability in many countries, however I think this should further be discussed in the manuscript.

- "Wastewater discharge" is often used in the manuscript to refer to the outflow from wastewater treatment plants. But please note the term "wastewater discharge" could also be interpreted akin to "wastewater production" – the total amount of wastewater produced (regardless of if there is subsequent collection and treatment). Please consider renaming this throughout the manuscript (e.g. "discharge from WWTP" or "treated wastewater discharge") to make it clearer.
- Global wastewater in Table 1 vs. Table 3. Firstly, please make clear that this is **treated** wastewater discharge. Also, while the JMP-WASH and Jones et al., totals are consistent across both data tables, why are the "dataset" values different (e.g. global wastewater discharge in Table 1 = 5 million m³/day; while in Table 3 = 521 million m³/day)? Looking at the dataset, 521 million m³/day looks the correct value.
- The authors do a very good job of inferring the outfall location (coordinates) of each WWTP, developing a rule based method in GIS. More as a comment, many (particularly global) water quality models are grid-based and typically at 30 arc-min (50 by 50km at the equator) or 5 arc-min (10 by 10km at the equator). So while the additional spatial resolution is good, it may be required to aggregated this back up to a gridded level for global water quality model runs, which may in turn cause a loss of usefulness of this data! Nevertheless, interesting data to provide and good that you also acknowledge these do not necessarily represent precise locations.
- For estimating the missing attributes for "population served"; I suggest the authors conduct and reflect upon the sensitivity of their results to their assumptions. The authors select the minimum value calculated from the three methods, but how do the results look if one method is used consistently instead? Also important to reflect upon the uncertainty of these derived attributes in general.
- The manuscript consistently refers to the treatment levels as "primary", "secondary" and "advanced". While "primary" and "secondary" are very commonly described in the literature, the authors should confirm exactly what is meant by the description "advanced" (i.e. what specific treatment practices). More commonly used in literature is the term "tertiary" (describing a third level of treatment); with "advanced" or "quaternary" used to describe a fourth level of treatment (see work from van Puijenbroek¹). Clarification here is very important for water quality modelers who need to define removal efficiencies of pollutants.
- Please acknowledge that the treatment capacities reported in the individual (source) datasets likely refer to maximum or operational design capacities, and thus might not actually represent volumes of wastewater being treated. For example, wastewater treatment plants may be shut or not operating at full capacity, leading to an overestimation of wastewater treatment². This is very difficult to overcome, but should be acknowledged nevertheless.
- It would be very interesting to see how the dilution ratio varies intra-annually, due to large seasonality in river discharge. However I do appreciate that the analysis presented in this paper is more with the purpose of showcasing the HydroWASTE dataset. If possible, analysis of the seasonality in dilution ratios would further improve the manuscript, but is not essential.
- I find the analysis in 3.3 to be OK, but quite rudimentary. The focus is very much strongly on rivers (potentially) impacted by treated wastewater effluent as a result of the dilution factor, thus ignoring the treatment level. A 'low' dilution factor may not always be bad if the wastewater is adequately treated. Therefore, I would suggest to alter Table 7 to also include additional information, such as treatment level.

Technical corrections

Line 9: Revise and shorten this sentence to improve readability.

Line 20: I would use "spatially-explicit" instead of "location-explicit".

Line 38: Revise this sentence. E.g. "...reduce the pollutant loads reaching downstream water bodies, typically with a focus on organic matter and macro-pollutants"

Line 44 - 48: Shorten this sentence.

Line 49: I would recommend citation of UNEP, 2016³ and van Vliet et al., 2021⁴ here.

Line 55 – 58: Shorten this sentence.

Line 108: "calculations are **typically** based..."

Line 246 and 285: There is unnecessary repetition here.

Line 327: The average distance between WWTP location and effluent outfall location seems quite large here? (I thought that this search radius was constrained to a maximum of 10km in most cases?). Please briefly elaborate on this.

Line 424 and 430: Some repetition here.

Figure 3: I appreciate the density of WWTPs in Europe and east-coast USA is very high and thus difficult to present visually, but perhaps slightly reduce the point size is possible..! Alternatively, you could perhaps consider only mapping wastewater treatment plants with a treatment capacity greater than a pre-defined threshold.

Line 364: "but" instead of "and"

References

1. Jones, E. R., van Vliet, M. T. H., Qadir, M., and Bierkens, M. F. P.: Country-level and gridded estimates of wastewater production, collection, treatment and reuse, *Earth Syst. Sci. Data*, 13, 237-254, 10.5194/essd-13-237-2021, 2021.
2. UNEP: A Snapshot of the World's Water Quality: Towards a global assessment, United Nations Environment Programme, Nairobi, Kenya, 162pp, 2016.
3. van Puijenbroek, P. J. T. M., Beusen, A. H. W., and Bouwman, A. F.: Global nitrogen and phosphorus in urban waste water based on the Shared Socio-economic pathways, *Journal of Environmental Management*, 231, 446-456, <https://doi.org/10.1016/j.jenvman.2018.10.048>, 2019.
4. van Vliet, M. T. H., Jones, E. R., Flörke, M., Franssen, W. H. P., Hanasaki, N., Wada, Y., and Yearsley, J. R.: Global water scarcity including surface water quality and expansions of clean water technologies, *Environmental Research Letters*, 16, 024020, 10.1088/1748-9326/abbfc3, 2021.