

This manuscript presents the Global Aquatic N<sub>2</sub>O Emission Database (GANED), compiling global observations of dissolved N<sub>2</sub>O concentrations and air–water fluxes from various aquatic systems, including rivers, streams, lakes, reservoirs, ponds, estuaries, coastal waters, and seas. The topic clearly fits within ESSD’s scope. A harmonized database integrating concentration and flux data, linking observations to specific sources and sites, and incorporating related environmental variables is valuable for supporting synthesis studies, model evaluations, data-gap analyses, and future upscaling research on aquatic greenhouse gas emissions.

The dataset has several strong points. First, it integrates both dissolved N<sub>2</sub>O concentrations and fluxes, significantly improving over databases focusing on only one parameter. Second, the relational structure using Source-ID and Site-ID enhances data traceability and usability. Additionally, including ancillary water-quality parameters along with other greenhouse gases (CO<sub>2</sub> and CH<sub>4</sub>) further boosts the dataset’s applicability. The manuscript effectively acknowledges existing spatial biases in data collection, clearly highlighting the concentration of observations in regions like Asia and North America, and gaps in Africa, South America, high-latitude areas, and less-studied aquatic ecosystems.

However, several major issues should be addressed before publication. These involve data documentation and transparency, methodological clarity, analytical results presentation, and aspects of data quality and metadata completeness.

### **Comments:**

1. The manuscript states that 52% of observations estimated emissions using dissolved gas concentration and gas exchange coefficient ( $k$ ), while 48% used chamber methods. This distinction is critical for interpretation and database reuse. Derived fluxes depend significantly on dissolved concentration, equilibrium concentration ( $C_{eq}$ ), and gas transfer velocity ( $k$ ), each introducing uncertainty. Specifically,  $k$  values depend on hydrodynamics, wind speed, flow velocity, channel slope, turbulence, water-body morphology, and parameterization method. The manuscript acknowledges substantial discrepancies due to differing parameterizations and methods.

Combining the two flux types is understandable given the heterogeneity of the source literature. The existing *Flux-Method* field is useful, but it is not sufficient to document the provenance of concentration-derived fluxes; documenting the  $k$  and  $C_{eq}$  assumptions would help users interpret correlations involving temperature, wind, and flow, which partly enter the derived fluxes by construction. Therefore, I strongly recommend adding fields such as `flux_origin`, `k_value`, `k_method`, `Ceq_method`, `solubility_formula`, and `temperature_used_for_conversion` to the archived data. Additionally, the authors should explicitly document how  $C_{eq}$  was calculated, including assumptions about temperature, salinity, and atmospheric N<sub>2</sub>O concentration. If feasible, a comparative analysis between chamber-based and concentration-gradient flux estimates from studies employing both methods would significantly enhance understanding of method-related uncertainties in the database (I recognize that studies applying both methods at the same sites may be scarce, so this is only a suggestion).

2. The correlation analysis supporting Table 2, Fig. 10, and the main driver conclusion should be reproducible. The current correlation analysis is critical because it supports the conclusion regarding substrate control of aquatic N<sub>2</sub>O dynamics. However, the unusually high correlation values reported, such as  $R = 0.943$  for NH<sub>4</sub><sup>+</sup> and  $R = -0.977$  for DOC, necessitate explicit and reproducible methodologies.

The archived script that appears to correspond to this analysis seems to contain pre-specified correlation values rather than computing them directly from the archived data. If another script or data subset was used to generate Table 2 and Fig. 10, the authors should identify and archive it. Otherwise, the current code should be revised so that these results can be reproduced directly from the deposited database.

Moreover, the description of Table 2 requires detailed clarification. Currently, the table caption broadly mentions ~190 filtered rows but lacks specificity. Authors should explicitly document the input file used, exact filtering criteria, handling of NA / zero / negative / extreme / non-log-transformable values, any data transformations applied, and the aggregation level used (e.g., site, study, or aquatic-type means).

When I attempted to recompute the correlations from the archived data, the NH<sub>4</sub><sup>+</sup>-N<sub>2</sub>O flux correlation did not match the reported  $R = 0.943$  and appeared to be much weaker. The DOC relationship also needs to be checked because the manuscript reports a very strong negative correlation, while direct analysis from the database may give a positive relationship. The manuscript also contains internal inconsistencies: DOC is described in the text as a positive relationship (L453 and L615), whereas Table 2 reports a very strong negative correlation. I recognize that my data subset may differ from the authors' "approximately 190 rows," but this is precisely the problem, since the manuscript does not define that subset clearly enough for readers to reproduce the result.

3. The overall Random Forest model has an  $R^2$  of 0.13. This means that the model explains only about 13% of the variance in N<sub>2</sub>O flux, leaving about 87% unexplained. This level of performance is not sufficient to support strong statements such as substrate availability being the primary global control on aquatic N<sub>2</sub>O production.

The Random Forest analysis can be useful as an example of how GANED may be used, but it should be presented as exploratory. The current text interprets variable importance too strongly. Feature importance from a low-skill model should not be treated as proof of a global mechanism.

The relatively low model performance may also indicate that water-quality variables alone are not enough to explain global N<sub>2</sub>O flux variability. N<sub>2</sub>O flux may also depend on watershed properties, water-body geometry, hydrology, land use, climate, channel morphology, nitrogen loading, etc. In a preliminary exploratory test using the archived data, adding variables such as coordinates, depth, surface area, and elevation appeared to improve predictive performance. I do not suggest that this test should replace the

authors' analysis, but it indicates that spatial, geomorphic, and system-level factors may deserve consideration.

The authors should also clarify how repeated sites, multiple depths, and closely related sampling points were handled. For example, if records such as Min River Estuary MP1 (D1–D13) represent different depths or sub-sites at the same location, the authors should explain whether these values were averaged or treated as independent. A site- or study-blocked cross-validation would help avoid overestimating model performance.

4. The manuscript states that rivers contribute the largest share (41.04%) of total emissions (Line 314). However, Fig. 3b appears to show the proportion of observational sites or water-body types, not an area-weighted or flux-weighted contribution to global emissions. Emission contributions should be derived from  $\text{emission} = \text{flux} \times \text{wetted area} \times \text{time}$ , accounting for flux magnitude, water surface area, temporal representativeness, and associated uncertainties. Without an explicit upscaling framework, it is inappropriate to report observational shares as emission contributions. Suggested wording revisions include stating rivers accounted for the largest share of observational sites or were among the most frequently represented water-body types, rather than implying their emission contributions.

5. Several unit and data-quality issues should be checked before the dataset can be safely reused. First, some ancillary water-quality units appear inconsistent. TP spans a very wide range, from about 0.04 to 4941, which seems more plausible as  $\mu\text{g L}^{-1}$  than as  $\text{mg L}^{-1}$ . Some DO values are greater than 20  $\text{mg L}^{-1}$  and cluster around 85–95, suggesting that percent saturation may have been mixed into a  $\text{mg L}^{-1}$  column. The authors should check these variables against the original sources and document the harmonized units.

Second, the database contains negative nutrient concentrations, including negative  $\text{NH}_4^+$  and  $\text{NO}_2^-$  values. If these values represent below-detection-limit handling, blank correction, or another source-specific convention, they should be flagged. Otherwise, they should be corrected or set to NA.

Third, Site\_ID 2061, “Brisbane River Estuary 11,” has an impossible latitude of  $-277.5509$ . Given the location and longitude, this is likely a misplaced digit and should be approximately  $-27.5509$ . This error indicates that checks should be conducted for latitude, longitude, concentrations, pH, salinity, and other key fields.

Fourth, the authors should clarify the scope of the  $>3$  SD filtering rule. Did this rule apply only to  $\text{N}_2\text{O}$  concentration and flux, or also to ancillary water-quality variables? If it only applied to  $\text{N}_2\text{O}$  variables, the text should say so. If it was intended to apply to all variables, the remaining extreme ancillary values need to be explained.

6. The manuscript states that nearby sampling points were merged and that data were averaged to produce representative records in some cases. This is a reasonable step, but the rule needs to be more transparent.

The archived site table appears to contain records such as Min River Estuary MP1 (D1–D13), which may represent multiple depths or sub-sites at the same station. The authors should clarify whether such records were averaged, retained as separate observations, or treated as separate sites. If retained separately, the reason should be given. If averaged, the averaging rule and the number of affected records should be reported.

7. The manuscript states that the database spans “8 continents.”, which is misleading. Under the conventional classification there are seven continents, and Central America is a sub-region of the Americas rather than a continent. If Central America is represented by only one site or observation, treating it as an eighth continent overstates the geographic coverage. The authors should use “regions” or “geographic regions” if they want to include Central America as a separate label. The terms Oceania and Australia should also be used consistently across the manuscript, appendix, and data files.

8. Mean and median values should both be reported for highly skewed variables. N<sub>2</sub>O concentrations and fluxes are strongly right-skewed, especially in rivers, streams, and ponds, where the mean can be strongly influenced by a few extreme observations. The manuscript should therefore report medians and interquartile ranges alongside means and standard deviations. This is especially important when comparing water-body types.