Author Comment (AC1)

Reply to Referee comments #1 (RC1)

Manuscript: essd-2026-109

Title: A Comprehensive Global Aquatic N_2O Emission Database (GANED): Unravelling N_2O Emission Patterns from Different Water Bodies

Response to: Anonymous Referee #1 (RC1)

Original comment date: 07th June 2026

Dear Referee:

We sincerely thank the Referee for the thorough and constructive evaluation of our manuscript. The comments provided have highlighted several important aspects of the database structure, the analytical workflow, and interpretation that require clarification and substantial revision. This Author Comment (AC) is submitted by the contact author on behalf of all co-authors, who have reviewed and endorsed the revisions described below. In this AC, we will outline the specific changes that will be implemented in the revised manuscript and the accompanying updated data. A complete point-by-point description of the changes, together with the revised manuscript and updated dataset, will be provided upon formal resubmission following the ESSD final-response and finalize workflow.

In particular, we acknowledge your recommendation to cite existing marine database (MEMENTO) and related inland and freshwater N_2O compilations such as PANGAEA, GO-SHIP databases to strengthen the description of our data-processing and quality control procedures, and to be more cautious in the interpretation of temporal trends and Random-Forest (RF) performance. We plan to address every comment in the revised submission as summarized below.

RC1 Major comments:

Comment 1

Lines 120-125: The description of literature search across Web of Science, Scopus, and Google Scholar lacks screening of duplicate records, as the same article may be recorded in different databases. The authors should make sure one record is counted only once. The number of unique *sources retained from each database after duplicate removal and screening should also be briefly summarized.*

Response: We agree that the literature-search workflow requires a transparent description of duplicate identification and removal. We will revise Section 2.1 to explicitly report the number of records at each PRISMA stage. A total of 2,847 records were initially identified, including 1,412 from Web of Science, 1,038 from Scopus, and 397 from Google Scholar. We will report that 954 duplicate records were removed in Zotero using DOI, title, author, and publication-year matching, followed by manual verification. The remaining 1,893 unique records were screened by title and abstract, 740 records were assessed at the full-text stage, and 426 sources were retained. A PRISMA flow diagram summarizing the complete search, de-duplication, screening, and inclusion process will be added as Figure S1 in the Supplementary Materials (SI).

Comment 2:

Lines 226-252: It seems that the authors combined flux data derived from concentration-gradient calculations and chamber-based measurements. These two methods differ not only in parameterization but also in sampling scale and measurement principle. It remains controversy whether these datasets, having different inherent biases, are comparable, or can be combined as a whole for subsequent analyses. Lacking these information makes it difficult to evaluate whether the main conclusions are robust.

Response: We agree that chamber-based and concentration-gradient flux estimates differ in measurement principle, parameterization, spatial footprint, and temporal representativeness, and therefore should not be treated as methodologically equivalent without further evaluation. In the updated data files, we will add a categorical field, *"Flux_Origin"*, to the archived Fluxes table. This field will distinguish static chambers, floating chambers, concentration-gradient estimates based on k600, concentration-gradient estimates using directly reported gas-transfer velocities, and mixed-method records. We will also add a new subsection, Section 3.2.4, entitled "*Methodological Comparability"*, which will present a paired comparison for studies that reported chamber and concentration-gradient measurements at overlapping sites and sampling periods. Correlation and RF analyses will be repeated for the pooled dataset and separately for each "*Flux_Origin"* category as a robustness assessment. We will clearly state that pooled analyses provide a broad synthesis and do not imply equivalence among measurement methods, and users will be encouraged to filter the database by '*Flux_Origin'* for method-specific applications.

Comment 3:

Lines 270-275: The removal procedure of outlier datapoints needs justification. Given the size of dataset, and the skewed distribution nature of N_2O concentrations and fluxes, excluding values greater than three standard deviations from the mean within each aquatic type may remove real "hotspot" observations rather than erroneous records.

Response: We accept this identification which is not appropriate for highly skewed environmental data and may limit accurate emission hotspots. In the revised workflow, concentrations and fluxes will first be $\log_{10}$-transformed, where appropriate, and potential outliers will be identified within each aquatic-system category using the Tukey 1.5 times interquartile-range (IQR) criterion on the transformed scale. Flagged observations will remain in the database and will be identified through a new *'Outlier_Flag'* field. Records will be removed only when verification against the original source confirms a transcription, unit-conversion, or physically implausible value. We will repeat the main statistical analyses and provide sensitivity analyses with and without flagged observations to determine whether the principal conclusions are affected.

Comment 4:

Line 315: Calculating the proportional contribution of different aquatic systems to total N_2O emissions is unclear. I am confused whether these percentages are based on simple aggregation of observations, spatially weighted estimates, or upscaled global fluxes.

Response: We agree that the original wording was ambiguous. The revised Abstract, Section 3.1, and Conclusions will explicitly state that the reported percentages represent the relative distribution of observational flux records among aquatic-system categories in GANED. They do not represent area-weighted contributions or globally upscaled shares of aquatic N_2O emissions. We will replace language referring to "contributions to total emissions" with "relative representation within the database". A new paragraph in Section 4.5 will distinguish observational coverage from spatially upscaled emission budgets and will direct readers to relevant riverine, inland-water, freshwater, and global assessments, including Tian et al. (2024). GANED will therefore be presented as an observational database that can support future upscaling.

Comment 5:

Temporal trends in Figures 6 and 7 may be affected by changes in sampling density and sampling locations across years, reflecting varying proportions of headwater streams, agricultural drainage channels, or urban-impacted rivers rather than realistic temporal changes in N_2O levels.

Response: We agree that changes in sampling density, site distribution, and system composition may strongly influence the apparent temporal patterns. In the revised manuscript, Figures 6 and 7 will be presented as two-panel figures. Panel (a) will show annual means based on the complete observational dataset, retained for transparency, while panel (b) will show a spatially balanced subset restricted to sites with at least three years of repeated sampling. A new paragraph in Section 3.3 will explain that opposing concentration and flux patterns in some systems may arise from changes in sampling composition, particularly the increasing representation of small or impacted water bodies with high gas-transfer velocities, and from changes in flux methodology, including the greater use of chamber-based measurements after 2015. We will avoid interpreting these patterns as unbiased global temporal trends and will state explicitly that robust temporal attribution requires repeated observations at fixed sites.

Comment 6:

Figure 9 shows large within-system variability (N_2O concentrations spanning nearly six orders of magnitude in rivers). Within-type heterogeneity may be driven by finer sub-categories such as polluted rivers, background rivers, agricultural drainage channels, or urban-impacted systems.

Response: We agree that broad system categories can conceal substantial ecological and methodological heterogeneity. We will use the existing *'Channel_type'* field, including the DIT, UIT, DRC, and related categories described in Table 1 of original submitted manuscript, to examine variation among river and channel subtypes. A new Figure will present N_2O distributions across these subcategories, and Section 4.4 of revised manuscript will report the proportion of variance associated with channel type using a one-way analysis of variance applied to log-transformed flux values. We will explicitly acknowledge that the river category includes background rivers, polluted rivers, agricultural drainage channels, urban-impacted systems, and other distinct environments. The revised manuscript and data documentation will encourage users to apply *'Channel_type'* filters when conducting subsystem-specific analyses.

Comment 7:

Line 305, Fig. 10 and Fig. 2: The data-processing procedure used for the correlation analysis is insufficiently described. It is unclear how missing values, repeated observations, aggregated records, different flux methods, and potential non-independence among records from the same site were handled is unclear. The correlation should distinguish more clearly between R and R^2. R^2 cannot indicate positive or negative associations; yet Fig. 10 and explanatory texts use R^2 together with correlation. This statistical notation should be made consistent throughout the main text and figure.

Response: We agree that the correlation analysis requires substantially greater transparency and statistical consistency. We will rerun the complete analysis in R, revise the code where necessary, and deposit a fully reproducible script with the archived CSV files so that the values in Table 2 and Figure 10 can be regenerated. The revised Methods will specify the input tables, treatment of missing pairs, $\log_{10}$ transformation of concentration and flux variables, handling of aggregated records, stratification by flux method, and aggregation to the site-mean level. Spearman rank correlation will be used as the primary statistic, Pearson correlation will be provided as a sensitivity analysis, and multiple comparisons will be adjusted. Site-level correlations will be reported as the primary results, while record-level correlations will be included in Table S4 as a sensitivity assessment. We will also conduct systematic hyperparameter tuning for the RF models using an automated search procedure combined with cross-validation (CV) to identify an appropriate combination of model parameters and ensure robust performance assessment. Model evaluation will continue to use site-blocked CV to minimize spatial dependence between the training and validation datasets. Further, we will also review the entire manuscript to ensure that R is used for signed correlation coefficients and R^2 is reserved exclusively for goodness-of-fit measures in regression and RF analyses.

Comment 8:

In fig. 10, the non-significant relationship between N_2O flux and TN needs to be clarified, as N_2O flux is reported to be strongly positively correlated with NH_4^+, NO_3^-, and NO_2^-, which are components of TN.

Response: We agree that the weak relationship between TN and N_2O flux requires a clearer mechanistic explanation. A dedicated paragraph will be added to Section 4.2 to clarify that TN combines the bioavailable dissolved inorganic nitrogen (DIN) pool with dissolved and particulate organic nitrogen fractions, many of which are relatively refractory and do not directly support short-timescale N_2O production. Interpretation is further complicated by differences among studies in filtration, digestion, and the inclusion of particulate nitrogen. In contrast, NH_4^+, NO_2^-, and NO_3^-, serve as direct substrates or intermediates in nitrification and denitrification and may therefore show stronger and more consistent relationships with N_2O dynamics. The revised discussion will present this interpretation cautiously and will cite the relevant literature.

Comment 9:

Based on the dataset, the Random Forest model may be limited in the predictive performance, having an overall R^2 = 0.13 (line 481 - 482) and low or negative R^2 values for some aquatic systems (stream, reservoir and ocean). The statement about the RF model having "robust predictive capability" appears to be overstated (line 481). The RF model may be useful in identifying the most important predictor, e.g. DIN. In table 3, the "Feature Importance" should be interpreted more cautiously, as they indicate relative importance within a weak predictive model rather than strong evidence for dominant controls on N_2O flux.

Response: We agree that the current RF performance does not support the description of robust predictive capability. Because the reported R^2 values require further verification, we will repeat the complete statistical run and carefully look into the R coding for potential problems in data preprocessing, and performance calculation. We will conduct systematic hyperparameter tuning using an automated search procedure integrated with CV to identify an appropriate parameter combination while preventing information leakage among repeated observations from the same site. We will evaluate both the water-quality model and an extended model that includes latitude, longitude, depth, surface area, and elevation. The final CV R^2 values will be reported transparently, regardless of whether model performance improves after code validation and tuning. Unless the revised models demonstrate consistently robust performance, RF will be described as an exploratory multivariate variable-ranking approach rather than a predictive model. The Table 3 caption and related discussion will also clarify that feature importance reflects relative ranking within the fitted model and should not be interpreted as direct evidence of dominant mechanistic controls.

Comment 10:

Section 4.3 provides a mechanistic discussion of temperature effects across aquatic systems, but the relationship between water temperature and N_2O flux in a global setting is weak and non-significant (as in Figure 10). System-specific temperature effects may exist, but the current analysis does not provide sufficient explanations.

Response: We agree that temperature interpretation require substantial explanation. Accordingly, we will substantially revise Section 4.3 and the temperature-related discussion in Section 4.2. In the revised manuscript, the temperature-flux relationship will be reported as weakly non-significant (rather than negative); a new Figure S8 will further show temperature-flux relationships stratified by water type, revealing the expected weakly positive association in rivers and streams and the negative signal in marine system. The weak association will be interpreted as counteracting effects of temperature on microbial kinetics, gas solubility, and dissolved oxygen (DO)-mediated redox conditions.

Comment 11:

The authors attribute the strong heterogeneity in GANED mainly to real differences in N_2O production and emission processes (Section 4.4). However, it should be noted that the database inherits the data variabilities from sampling design, observation frequency, and measurement approaches. Without separating these sources of variability, the interpretation of heterogeneity as an inherent process-level feature may not be justified.

Response: This is an important point that need to be sufficiently addressed. The heterogeneity in GANED is a composite of (a) environmental variability (different N_2O production rates, pathways, drivers across sites), (b) methodological variability (chamber vs. *k*-based methods, different k parameterizations, different sampling frequencies), and (c) sampling design variability. We will revise Section 4.4 to explicitly acknowledge that observed heterogeneity reflects a mixture of these process-level and methodological variability. We will add: The heterogeneity preserved in GANED reflects both environmental variation and methodological differences across source studies. Users should not attribute all observed variability to process-level differences without accounting for the methodological factors reported in the flux-method and aggregation metadata fields. We will also discuss the potential for future analyses to partition these variance components.

RC1 Minor comments:

1. *The abstract should contain a brief explanation about how these heterogeneous datasets were harmonized, standardized, and quality-controlled before cross-system comparisons and correlation analyses.*

Response: We agree and will add a concise statement to the Abstract describing the harmonization and quality-control procedures. The revised text will state that records were converted to common units, standardized using consistent variable definitions, checked for coordinate and unit consistency, screened using log-scale outlier flagging, and independently reviewed by co-authors before cross-system comparison and statistical analysis.

2. *Line 39-40: There appears to conflate atmospheric N_2O concentration with N_2O emissions; ppb is not an emission unit.*

Response: We agree and will correct Lines 39 to 40 to refer to an atmospheric N_2O concentration of 336 ppb. The term "emissions" will be removed because ppb is a concentration unit.

3. *Line 128: As pH is not included in subsequent analysis of predictor, the authors need to explain here.*

Response: We will add an explanation for why pH was excluded from the correlation analysis: "pH was excluded from the regression analysis in Table 2 due to insufficient non-NA pairwise data with N_2O flux observations ($n < 10$ complete cases)."

4. *Lines 145-150: It is unclear whether concentration and flux measurements collected at the same site and date can be directly matched at the record level.*

Response: We will clarify that concentration and flux measurements collected at the same site and date can be matched through the shared *'Site-ID'* and *'Flux-Name/Conc-Name'* fields. When both concentration and flux exist for the same site-date combination, the *'Observation_Type'* field is marked as "Both" and the Flux-Name matches the corresponding Conc-Name. We will add a brief statement documenting this linkage.

5. *Lines 165-170: The details of the authors' own data need to be described, i.e. the size of dataset, and whether including these data could have changed the interpretation in certain aquatic systems.*

Response: The revised text will explicitly report the authors' own contribution to the database (i.e.; new observations from the Pearl River Estuary and coastal Guangdong, 2019-2023), together with a sensitivity analysis confirming that removing our own records does not change any qualitative conclusion.

6. *Table 1 lists several special site or channel types that may influence N_2O concentrations or fluxes, but the distinction between coded and uncoded sites is not clearly defined.*

Response: We agree that the distinction between coded and uncoded metadata should be clearer. Table 1 and the accompanying data dictionary will include a legend identifying coded categorical fields used in analysis and uncoded free-text fields retained for traceability and source documentation.

7. *The number of flux records is inconsistent across the Results section. The database is first described as containing 7,386 flux records, whereas Fig. 9 is a compilation of 6,456 flux measurements.*

Response: The difference reflects the removal of records with NA or missing N_2O flux values before generating the boxplots in Fig. 9. We will add an explicit note in the figure caption: "6,456 flux measurements with valid N_2O flux values are shown (from 7,386 total records; 930 records were excluded due to missing or invalid flux values).

8. *Lines 315-320: The authors indicate that only 2% of sites have more than 10 concentration observations and 4% have more than 10 flux records, suggesting very limited temporal coverage. The meaning of "total = 1629 observational counts" and "total = 3807 counts" should be clearly stated, and the manuscript should better justify how such sparse long-term observations support the subsequent temporal trend analyses.*

Response: Lines 315-320 will be clarified: 1,629 observational counts and 3,807 counts refer to unique Site × Year combinations for concentration and flux, respectively. The sparse long-term coverage will be acknowledged as a limitation of temporal-trend interpretations. We will add a discussion of the limitations of using sparse temporal data for trend analysis, acknowledging that the small number of sites with >10 observations limit the representativeness of temporal trend analyses. We will explicitly state that the temporal trends shown in Figs. 6 and 7 are primarily characterizations of the observational record rather than robust estimates of temporal change in N_2O emissions.

9. *Lines 450-465: The interpretations of correlation analyses appear to be inconsistent. DOC is described as positively associated with N_2O flux, whereas the abstract and Table 2 report a negative correlation; water temperature is described as having an inverse relationship, whereas Table 2 shows weakly positive and/or non-significant relationship. These inconsistencies need to be double checked throughout the text in the next round of revision.*

Response: We agree that the direction and significance of the reported correlations must be internally consistent. We will audit and rerun the full statistical analysis, review the R code, and cross-check every correlation reported in the Abstract, main text, Table 2, Figure 10, and Supplementary Materials. Particular attention will be given to DOC and water temperature. The revised wording will be based only on the validated results, and the sign, strength, and statistical significance of each relationship will be reported consistently.

We thank the Reviewer once again for the constructive and thoughtful comments. Following this final-response phase, we will finalize the interactive discussion via the ESSD portal and prepare a fully revised manuscript for peer-review completion, addressing all points documented above. A complete point-by-point description of the changes will accompany the revised manuscript upon resubmission.

On behalf of all co-authors,