# Review Comments

This manuscript investigates the spatial and temporal variability of stable water isotopes in precipitation and rivers across Romania and the Republic of Moldova, integrating monthly $\delta^{18}O$ and $\delta^{2}H$ observations from 22 precipitation stations and 19 river stations. From a scientific perspective, the study area lies in a transition zone influenced by Atlantic, Mediterranean, Black Sea–Caspian moisture sources and by the orographic barrier of the Carpathian Mountains. Because stable-isotope observations remain relatively sparse in this region, establishing a comparatively dense precipitation–river isotope database has clear value in filling an important data gap. From a practical perspective, the dataset could provide useful baseline information for regional water-source identification, investigations of surface water–groundwater interactions, catchment-scale water-cycle model evaluation, and future assessments of water-resource variability.

However, the present manuscript has several substantial shortcomings that affect the core reliability of a data paper. The temporal and spatial representativeness of the precipitation and river samples is not well matched, yet the two datasets are used directly for process attribution. Interpretations involving heat waves, Mediterranean cyclones, moisture recycling, snowmelt, and catchment "resilience" extend beyond what the analyses presented here can support. Data completeness, missing observations, quality control, and the provenance/originality of the data product are insufficiently documented. The statistical framework is incomplete, making several key quantitative results difficult to reproduce. In addition, there are numerous errors or inconsistencies in figures, coordinates, notation, and references. These issues cannot be resolved through local editing alone; they would require a substantial redefinition of the analytical scope, reconstruction of the data-quality framework, and additional analyses. In my view, the manuscript in its current form does not meet the standards of Earth System Science Data for data reliability, reproducibility, and documentation of the data product. I therefore recommend rejection (Reject).

## Major comments

1. The precipitation and river datasets are fundamentally mismatched in temporal scale and spatial representativeness, yet they are directly compared to infer hydrological processes. Page 5, Lines 120–146 indicate that precipitation samples represent monthly cumulative collections, whereas river water was sampled only once, at the end of each month. Pages 3–4, Lines 100–115 further indicate that most river sampling sites are located in the lower parts of their catchments, while the precipitation stations cannot represent the entire catchment, particularly the high-elevation source areas. Although the authors themselves acknowledge on Page 10, Lines 281–285 that the precipitation stations are generally located in the lower parts of the catchments and that this may affect the interpretation of river–precipitation isotope differences, the manuscript nevertheless interprets the approximately 0.8‰ depletion of river water relative to

precipitation as evidence for a greater winter-snow contribution and compares concurrent monthly values in Fig. 6 as "co-variability." Large catchments integrate precipitation through storage, mixing, damping, and time-lag effects. Without catchment-area/elevation-weighted precipitation isotope estimates, lag analysis, or a more defensible spatial matching scheme, the current precipitation–river process attribution is not sufficiently robust. This issue directly affects the central conclusions of the river-water section.

2. Several atmospheric-circulation and moisture-source mechanisms are presented as causal explanations, but the manuscript provides no independent analyses capable of testing them. Page 8, Lines 221–232 attribute the high and spatially homogeneous summer isotope values to regional heat waves and the high autumn d-excess values to Mediterranean cyclones. Pages 9–10, Lines 258–263 further state that "heat waves block moisture advection and enhance local recycling," while the Conclusions (Page 13, Lines 320–331) again identify the NAO, Siberian High, blocking patterns, and local moisture recycling as primary controls. However, the study includes no backward-trajectory analysis, moisture-flux diagnostics, circulation-type classification, cyclone-track analysis, moisture-recycling calculation, or event-scale composite analysis. The present evidence is essentially a combination of seasonal isotope patterns and interpretation based on previous literature, which is insufficient to establish these mechanisms. If the authors wish to retain such conclusions, independent atmospheric-circulation/moisture-source analyses are required; otherwise, the relevant statements throughout the manuscript should be systematically reframed as possible mechanisms or as patterns consistent with previous regional studies.

3. Reporting of data quality control and missing observations is insufficient for a data paper and does not adequately support traceability or reproducibility. Page 6, Lines 182–191 explicitly state that some precipitation samples had negative d-excess values, which the authors suspect resulted from post-precipitation evaporation in the collectors, yet all of these data were retained. Figure 4 on Page 9 also contains numerous blank cells, but the manuscript does not explain whether these represent missing samples, months without precipitation, sample loss, analytical failure, or quality-control exclusion. The manuscript does not report, by station, the expected number of observations, number collected, number considered valid, missing months, number of negative-d-excess samples, or quality flags; nor does it explain whether QC flags are included in the Zenodo dataset. This is particularly important because d-excess is subsequently used to infer Mediterranean moisture and evaporation processes. Sampling artifacts and atmospheric signals therefore need to be distinguished explicitly. At minimum, the authors should provide a complete station-by-month data-completeness and QC table and perform a sensitivity analysis comparing results with and without questionable samples.

4. The precipitation-sampling design may introduce evaporation, sublimation, and inter-station methodological differences, but the current methods do not demonstrate that these effects are negligible.

Page 5, Lines 120–141 indicate that the self-built collectors were used without oil, collector heights ranged from 20 to 120 cm, and open plastic containers were used in winter. Samples were transported to the laboratory only every 3–6 months and, in some cases, were analyzed after storage periods approaching nine months. Although the authors cite previous experience suggesting that the absence of oil does not systematically alter isotope values, no consistency test is presented for this study among collector types, collector heights, or the different summer and winter collection methods. Because the manuscript places considerable emphasis on d-excess and evaporation, this methodological uncertainty is particularly important. If different stations used different collector types or heights, these details should be reported station by station in the metadata, together with cross-validation results or at least a quantitative assessment of the potential bias.

5. The statistical analysis is insufficiently described, and several key quantitative results cannot be independently reproduced. Page 9, Lines 250–258 report $r^2$ values for relationships between $\delta^{18}O$, air temperature, and elevation, including values after excluding two high-elevation stations; Page 11, Lines 285–290 state that river and precipitation isotopes show a "very good correlation." However, there is no dedicated Statistical analysis subsection in the Methods. The manuscript does not specify the sample size (N), whether regressions use station means or all monthly observations, whether Pearson correlation or OLS regression was applied, whether temporal autocorrelation or repeated observations were considered, how seasons were defined, how missing data were handled, or what significance level and confidence intervals were used. With only three years of monthly data, pooling all station-month observations would also create substantial non-independence and potential pseudoreplication. The statistical framework should therefore be rebuilt and all correlations/regressions should report N, slope and/or correlation coefficient, $R^2$, p value, and 95% confidence intervals.

6. The provenance of the data and the boundary between newly released and previously available data are not sufficiently clear. Page 4, Lines 111–115 state that 17 precipitation stations belong to the authors' network, while five stations in the Republic of Moldova are obtained from GNIP and are "partly unpublished." The same section also states that data from four rivers (Ocoale, Deblei, Mniera, and Toaia) were previously published by Feher et al. (2024). The manuscript does not clearly distinguish data first released in this study from data already publicly available through GNIP or previously published river data, nor does it explain how source and licensing information are recorded for individual observations in the Zenodo dataset. More seriously, Feher et al. (2024) is cited in the text but no corresponding 2024 reference is included in the reference list. For a data paper, data provenance, publication status, licensing, and the specific new contribution of the present dataset should be documented at least at the station level, and preferably at the record level, so that the dataset's originality and reuse conditions can be evaluated.

## Specific comments

(a) Title, abstract, and wording

Page 1, title: The title uses “Eastern Europe,” whereas the dataset covers only Romania and the Republic of Moldova. The manuscript also alternates among “South Eastern Europe” and “East-Central Europe.” The geographical terminology should be narrowed and used consistently to avoid overstating the regional representativeness of the dataset.

Page 1, Lines 34–40: The Abstract states that regional isotope differences “are linked to” the NAO, Siberian High, Mediterranean cyclones, and heat waves. This wording is stronger than the evidence presented in the manuscript. Consider replacing it with language such as “may be associated with” or “are consistent with.”

Page 2, Lines 57–60: The phrase “the (NAO) Atlantic Oscillation (AO)” conflates two different names/abbreviations. Please clarify whether the intended term is the North Atlantic Oscillation (NAO) or the Arctic Oscillation (AO).

Page 8, Lines 221–225: The statement “Altitude and latitude do not seem to be playing a role” is too categorical, especially because Page 9 later reports an $r^2$ of 0.56 between $\delta^{18}O$ and elevation. These statements should be reconciled on the basis of the formal statistical results.

Page 10, Lines 260–263: “heat waves block moisture advection and enhance local recycling” is a strong mechanistic statement, yet moisture recycling is not quantified in this study. The phrase “enhance local recycling” should be removed or explicitly framed as an untested hypothesis.

Page 12, Lines 304–307: The manuscript describes lowland catchments as “the least resilient,” but resilience is neither defined nor quantified. Please remove this wording or replace it with a more defensible statement such as “may be more sensitive to …”.

Page 13, Lines 328–331: “bimodal distribution” is not an appropriate term for a seasonal pattern characterized by winter minima and summer maxima. Unless a true bimodal probability distribution is demonstrated, “pronounced seasonal/annual cycle” would be more accurate.

(b) Methods, data, and statistical details

Page 4, Table 1: The longitude of Chișinău is given as 28°53′60.00″E and the latitude of Cahul as 45°47′60.00″N. Seconds should not reach 60; these coordinates should be carried over to the next minute and checked against the original metadata.

Page 4, Table 1: Brașov uses commas as decimal separators in the coordinates, whereas other stations use decimal points; straight and curly prime symbols are also mixed. Please standardize the DMS format, number of decimal places, and coordinate notation throughout the table.

Page 4, Table 2: Length and/or drainage area are blank for rivers such as Săsar and Turia. Please

indicate whether these values are unavailable or not applicable, use a consistent code such as “NA,” and define it in the table note.

Page 5, Lines 121–126: Collector heights range from 20 to 120 cm, which is a substantial difference. Please report the collector type and installation height for each station and explain why this variation is not expected to materially affect collection representativeness.

Page 5, around Line 130: “walls of the collect” should read “walls of the collector.”

Page 5, Lines 151–155: “water was drawn from each vials” should be “water was drawn from each vial”; “high amount of particulate matter” would be better written as “a large amount of particulate matter.”

Pages 5–6, Lines 160–175: The preceding text states that measurements were normalized to VSMOW2 and SLAP2, whereas the δ definition subsequently refers to SMOW. Please use the VSMOW–SLAP scale consistently and clearly specify the reference standards and normalization procedure.

Page 6, Lines 173–175: “where R is the either …” should be corrected to “where R is either …”.

Page 5, Lines 157–160 and Page 6, Lines 174–175: The manuscript requires the standard deviation of the final four injections to be below 0.03‰ for $\delta^{18}O$ and 0.3‰ for $\delta^{2}H$, but later reports analytical precision of 0.16‰ and 0.7‰. These values may refer to different metrics (e.g., within-run repeatability versus long-term external precision), but this distinction is not explained and should be clarified.

Page 6, Lines 187–191: The statement that “no memory and linearity (drift) errors occurred” is too strong without reported memory tests, drift tests, or acceptance thresholds. Please provide the supporting QC results or revise the wording to something such as “no detectable memory or drift within the applied QC criteria.”

Page 6, Lines 193–196: River discharge is said to come from the gauging station nearest to each isotope sampling site, but the gauging-station names, distances to the isotope sites, and whether they are located on the same river reach are not reported. A station-matching table should be provided.

Pages 6–7, Lines 198–203: LMWLs are calculated using data from 2012–2017, whereas the main dataset covers 2015–2017. Please specify the source, station coverage, and publication status of the 2012–2014 observations, and explain why these earlier data are combined with the present dataset.

Page 10, Table 3: The LMWLs are reported only as slope and intercept with uncertainties. Please also report the sample size (n), $R^2$, p value and/or an appropriate measure of goodness of fit, and explain the precipitation-amount weighting and missing-data treatment used for PWLSR.

Page 9, Lines 250–255: Please clarify whether the elevation regression uses long-term station means or all monthly observations. If all monthly data are used, elevation is repeated for each station and the analysis would involve substantial pseudoreplication. If station means are used, the actual sample size (e.g., N = 22 or the number of stations with complete data) should be stated.

Page 9, Lines 250–257: It is not rigorous to describe a relationship as positive or negative using $r^2$ alone, because $R^2$ has no sign. Please report Pearson's r or the regression slope together with $R^2$, p value, and confidence intervals.

Page 9, Lines 250–252: The text states that $\delta^{18}O$ is negatively correlated with elevation but then gives an "altitude gradient" of 0.24‰/100 m. If the gradient is defined as the change in $\delta^{18}O$ with increasing elevation, the sign should be negative. Please verify the original regression result.

Page 10, Lines 276–280: The extremely low d-excess of the Olt River is attributed directly to "massive evaporation" from 19 upstream reservoirs, but no reservoir-isotope data or evaporation estimates are presented. Please provide quantitative evidence and a supporting source, or frame this interpretation as a hypothesis.

Page 10, Lines 281–285: The approximately 0.8‰ depletion of river water relative to precipitation is interpreted as evidence for a larger winter snowmelt contribution. High-elevation precipitation, groundwater/baseflow contributions, and differing water ages may all contribute to this offset. The difference should therefore not be attributed to snowmelt alone.

(c) Figures, formatting, notation, and references

Page 7, Figure 2: The station names on the x axis are crowded and difficult to read. Please use a consistent station-ordering scheme and explain in the caption the definitions of the box, whiskers, mean symbol, and outliers.

Page 8, Figure 3: The upper and lower panels use precipitation and river stations, respectively, with complex ordering and colors. The caption should report the numerical values represented by the red and black dashed lines and the sample sizes used to calculate them.

Page 9, Figure 4: The figure lacks a color legend/color bar, so the numerical d-excess ranges represented by green, yellow, and red cannot be interpreted. The station abbreviations (BM, BN, BZ, etc.) are not defined, and the meaning of the numerous blank cells is also unclear. In its current form, the figure is not self-contained.

Page 9, Lines 240–243: The text states that PWLSR slopes are higher at 13 of 19 stations and lower at 5 of 19 stations, which accounts for only 18 stations. Please explain the result for the remaining station.

Page 10, Lines 268–270: "The range of temporal variability (Fig. 5)" appears to cite the wrong figure. Figure 5 shows $\delta^{18}O$–$\delta^{2}H$ cross-plots rather than temporal variability.

Page 11, Figure 5 caption: "Gârda (precipitation) and Arieș (river)" is inconsistent with the station names in Table 1, where the relevant high-elevation station is listed as Ghețar. Please clarify whether Gârda is a different station, a locality name, or an error.

Pages 10–11: The text states, based on Figure 5, that river samples define lines with lower slopes and

intercepts than the local meteoric water lines, but Figure 5 does not display river regression lines, regression equations, $R^2$ values, LMWLs, or the GMWL. These elements should be added if this interpretation is to be supported directly by the figure.

Page 12, Figure 6: Precipitation and river $\delta^{18}O$, which have the same units, are plotted against different left and right y-axis ranges. This can visually exaggerate apparent synchrony. A common y-axis scale is preferable; if dual axes are retained, the different scales should be made explicit and the visual comparison should be accompanied by quantitative correlation/lag results.

Page 13: "5. Data availability" is followed directly by "Conclusions" without a section number. The numbering should be made consistent, for example "6. Conclusions."

Page 13, Data availability: Only the Zenodo DOI is provided. For an ESSD data paper, this section should also summarize the number and format of files, main variables, time fields, missing-value encoding, QC flags, station metadata, and dataset version.

Page 4, Line 115: Feher et al. (2024) is cited in the text, but no corresponding 2024 reference appears in the reference list on Pages 14–17; only Feher et al. (2021) is listed. The citation/reference entry must be added or corrected.

Page 8, Lines 228–230: Forray et al. (2024) is cited in support of Mediterranean cyclones and moisture sources, but the cited reference is titled "Past anthropogenic impacts revealed by trace elements in cave guano," which does not appear relevant to the claim. Please verify and replace this citation with the intended reference.

Page 15: The DOI URL for Forray et al. (2024) is written as "https://soi.org/…". This should be corrected to "https://doi.org/…".

Page 16: The DOI given for Persoiu et al. (2019), "10.5194/cp-2018-161," is the discussion/preprint DOI. Please verify the bibliographic record and use the final published reference consistently.

Page 2, around Line 53 uses IAEA/WMO (2025), whereas Page 6, Line 196 uses IAEA/WMO (2026), and the reference list presents the GNIP database with an access date. Please standardize the database citation year and access-date format throughout the manuscript.

Page 7, Lines 205–214: Several numerical expressions appear to contain editing errors, including the order in "maxima between -1.4 … -5.7," the extra full stop in "absolute range of 3.4. ‰," and missing δ/isotope notation. These values should be checked against the source data and the paragraph should be carefully proofread.

Throughout the manuscript, $\delta^{18}O$, $\delta^{2}H$, d-excess, δprec, and δriv are not fully consistent in font, subscript/superscript formatting, and spacing. Please standardize the isotope notation according to the journal style.

Author names in the citations and reference list are inconsistent in the use of diacritics and spelling (e.g., Persoiu/Perșoiu, Badaluta/Bădăluță, Ionita/Ioniță). Please verify each entry against the published source and use consistent spelling between the text and reference list.

DOI/URL formats are inconsistent in the reference list (e.g., https://doi.org versus http://dx.doi.org), and journal abbreviations and punctuation are not standardized. The reference list should be regenerated and then manually checked against the original publications.

All figures should be checked for self-contained readability. Each figure should clearly identify variables, units, colors/symbols, sample size, and missing-data information without requiring the reader to search the main text. Figures 4–6 currently lack several of these elements.

Overall, the regional stable-isotope observations compiled in this manuscript have potential value, but the current submission has major deficiencies in several areas that are central to a data paper: data provenance and the boundary of originality, quality control and missing-data documentation, matching of precipitation and river observations in space and time, statistical reproducibility, and consistency between mechanistic interpretation and the strength of the evidence. Addressing these issues would require substantial reconstruction rather than routine revision. I therefore recommend rejection (Reject) of the manuscript in its present form.