

The author has made careful and substantive revisions in response to the initial peer-review comments, resulting in a significant improvement in the manuscript's quality. Therefore, I recommend accepting this manuscript after minor revisions. To further enhance the rigor and usability of this data-driven paper, please address the following revision suggestions for the author's reference and refinement in the final version.

We sincerely thank the reviewer for the positive evaluation of our revised manuscript and for the constructive suggestions. We have addressed all comments as described below.

Please note that the line numbers referred to in the responses correspond to the tracked-changes version of the revised manuscript.

Comment 1:

It is recommended to provide a simple quantification or discussion of the typical error of digital data (e.g., the standard deviation or mean absolute error compared to the original data) in the Methods section (Section 2.1.2) or in the Supplementary Materials, to enhance the credibility of the data.

Response: We thank the reviewer for this helpful suggestion. We agree that the uncertainty associated with digitised data should be addressed more explicitly. Since original numerical values were not available for the digitised records, a direct calculation of the mean absolute error or standard deviation against the original data was not possible in all cases. We therefore treat digitisation uncertainty qualitatively, following the approach used in the WoKaS database, where digitised hydrographs were assessed through quality classes rather than fixed numerical error estimates. In WoKaS-Iso, all digitised output records are explicitly identified in the metadata and assigned a lower score than collected observations in the data quality attribution scheme. We have added the following discussion in Lines 260–266:

“Previous methodological studies have shown that carefully extracted graphical data can closely reproduce original numerical data and generally have limited influence on downstream statistical analyses, although extraction accuracy depends on graph quality, axis scaling, symbol clarity, and the size of plotted features (Van der Mierden et al., 2021; Turner et al., 2023). Following the approach used in the WoKaS database, digitization uncertainty is treated qualitatively through metadata documentation and data quality attribution rather than through fixed numerical error estimates (Olarinoye et al., 2020). All digitized records are

explicitly identified in the metadata, allowing users to distinguish them from directly collected observations and to account for this additional uncertainty in subsequent analyses.”

Comment 2:

Throughout the text, a clear distinction should be made between "number of caves" (74) and "number of drip sites" (365), with this hierarchical relationship explicitly stated upon its first mention. For example, the title of Table 1 could be revised to "Summary of observations for karst springs and cave drip sites".

Response: Thank you for this suggestion. We have clarified the hierarchical relationship between caves and drip sites in the Abstract (Lines 84–85) and revised the title of Table 1 accordingly (Lines 217–218).

“The database incorporates data sourced from 241 springs and 74 caves, of which 73 caves contain 365 individual drip sites and one cave contains input data only. In total, the database comprises 1007 time series, including 389 spring time series and 618 cave drip water time series.”

Comment 3:

Section 2.2.1 introduces various input data sources (local data, meteorological stations, GNIP, and global products), while Section 4.2 provides a quality scoring system; however, this section offers limited guidance on how users should select these data sources in actual research projects. It is recommended to include a section on data usage priority within the "Recommended Usage" section, for example: "Local observational data are recommended as the primary source; followed by data from nearby GNIP stations. Among global products, IsoGSM is recommended as the preferred choice due to its performance characteristics."

Response: Thank you for this helpful suggestion. We have added guidance on data usage priority in the Recommended Usage and Outlook section (Lines 780–791):

“The choice of input data source should depend on the specific research objective, required temporal coverage, spatial representativeness, and target variable. For site-specific process studies, local observations should be prioritised whenever available, as they provide the closest representation of the monitored spring or cave system. Where local precipitation or

precipitation isotope observations are unavailable or incomplete, nearby meteorological stations or GNIP records can provide useful alternatives, provided that the distance, elevation difference, and record continuity are suitable for the intended application. Global products are recommended for large scale comparative analyses, gap filling, or studies requiring spatially consistent and temporally continuous forcing data across multiple sites. They may also be useful alongside local observations to provide broader spatial context or to extend temporal coverage. Among the evaluated global precipitation isotope products, IsoGSM generally showed the most consistent performance and is therefore recommended as the preferred global substitute when local or nearby station-based precipitation isotope observations are not available. Nevertheless, users should consider site-specific uncertainties and compare alternative products where possible, particularly in regions with complex karst topography or sparse observational networks, where the spatial representativeness of a single station or grid cell may be limited.”

Comment 4:

The article provides a detailed description of the spatial distribution and data resolution, but does not specify the overall time span of the dataset (e.g., in which year did the earliest records begin? In which year did the latest records end?). At the beginning of the abstract, include additional information regarding the overall temporal coverage of the database, for example: "The compiled records span from XXXX to YYYY, with the majority of data collected after 2000."

Response: Thank you for this suggestion. We have added this information to the Abstract (Lines 86-87):

“The compiled isotope records span the period from 1967 to 2024, with the strongest temporal coverage during the 2010s and continuing coverage into the early 2020s.”

Comment 5:

In Section 3.3, which describes scripts, add an explanatory sentence—for example, by providing a small example of coordinates and expected output values for testing script functionality, or by explaining how to retrieve version information for the original NetCDF file that corresponds to the script.

Response: Thank you for this suggestion. The requested information is already provided in the README file accompanying the extraction scripts. Specifically, the section “Prepare the Site

Coordinates File” describes the required coordinate CSV format (name, latitude, and longitude) and provides example coordinates. The section “Prepare the Required Global Product Files” explains which external product files are needed before running the scripts, including NetCDF files for MSWEP, GLEAM, and ERA5, DAT files for IsoGSM2, and TIFF files for the Sine Curve model. The section “Set Paths in the MATLAB Scripts” specifies the input data folder path, site coordinates CSV file path, and output folder path that users need to update. The section “Run the Scripts” explains how to execute the scripts and select daily or monthly output where applicable. Finally, the section “Output CSV Format” describes the expected columns in the daily, monthly, and precipitation isotope output files. To make this clearer in the manuscript, we have expanded the description of the script folder in Sect. 3.3, including information on example coordinate input, expected output CSV structures, and NetCDF version checks. Please see Lines 582-593:

“The accompanying README.md file provides detailed instructions on obtaining and preparing the required input datasets, including links to download or access the relevant global product data, preparing coordinate files for user-specified extraction, setting file paths, running the scripts, selecting temporal resolution where applicable, and interpreting the resulting CSV outputs. It also provides example coordinate input and describes the expected output CSV structures for daily hydroclimatic data, monthly hydroclimatic data, and precipitation isotope data. The scripts were developed for the global product versions and file formats listed in Table 2 and described in the README.md file. For NetCDF-based products, users should verify the product version, temporal coverage, variable names, dimensions, and time metadata from the downloaded NetCDF files or the corresponding data provider documentation, as the extractable time period and version depend on the files provided by the user. If updated product versions are used and these properties differ from the supported formats, the script settings may need to be adjusted. In addition, detailed comments are also included within each script to guide users through extraction workflow.”

Comment 6:

In the bibliography, the journal titles for some references use abbreviations (e.g., Hydrogeol J), while for others they use full names (e.g., Hydrology and Earth System Sciences), resulting in inconsistent formatting. Please perform a final, unified proofreading of the journal titles,

volume numbers, and page ranges for all references in accordance with the ESSD journal formatting requirements.

Response: Thank you for this suggestion. We have proofread and revised the full reference list according to the ESSD reference style, including journal title abbreviations, volume numbers, page ranges, article numbers, DOIs, and general formatting throughout the bibliography (Lines 823–1061).