



OPEN

DATA DESCRIPTOR

A nationwide dataset of stable isotopes in meteoric and terrestrial water across Peru

Carol Romero¹ & James Apaéstegui^{1,2}

Water Stable Isotopes ($\delta^{18}\text{O}$, $\delta^2\text{H}$) are valuable tools for tracing sources and interactions in the water cycle, providing important information dedicated to understanding physical mechanisms related to global climate. Despite their significance, the topic of isotopic research in South America has been hindered by limited data. To address this gap, we launched a national-level water stable isotope dataset covering different water sources in Peru (WSI-PeruDB). The dataset contains curated in-house data and incorporates previously published records from various locations collected between 2000 and 2021. The WSI-PeruDB dataset is composed of 489 water collection sites and allows a comprehensive use of the dataset by implementing standardized metadata templates containing essential geographical information such as latitude, longitude, and altitude (from sea level to 5000 m a.s.l.), and sampling information such as sample type (e.g. groundwater, precipitation, river, spring, and others) and sampling frequency (e.g. biweekly, daily, monthly). The WSI-PeruDB dataset is publicly available on Zenodo, facilitating access and use for the scientific community.

Background & Summary

Understanding water cycle dynamics is critical in the context of ongoing climate variability and increasing human pressures on water resources^{1,2}. Stable isotopes of water ($\delta^{18}\text{O}$, $\delta^2\text{H}$) provide a valuable means for tracing the movement and sources of water within the hydrological cycle across a range of spatial and temporal scales. Isotopic fractionation in water elements during evaporation, condensation, and precipitation processes allows for insights related to atmospheric circulation, moisture sources^{3,4}, and hydrological processes^{5,6}.

In Peru, climate variability, glacier retreat, urbanization, and agricultural expansion⁷ have altered water availability and distribution, particularly in regions such as the Cordillera Blanca^{8,9} the coastal regions¹⁰ and Andean zones¹¹. Events like El Niño–Southern Oscillation (ENSO) episodes have further influenced regional hydrological patterns^{12–14}. Understanding these impacts requires detailed, spatially explicit data on water sources and flow paths. Stable isotope data, by serving as ideal geochemical tracers, support the investigation of processes not covered by conventional methods, such as precipitation sources, groundwater recharge, streamflow origin, and catchment storage. Several previous studies have contributed to the understanding of isotopic patterns in Peru and into the broader context of the Andean region^{15–26}. These works underscore the relevance of stable isotope applications for understanding climate variability, water sources, and the impacts of extreme events. The dataset presented here builds on this foundation by expanding geographic coverage and including a wider array of water types and measurement periods.

Building on these applications, stable isotopes have been employed in the evaluation and refinement of general circulation models (GCMs)²⁷, offering a means to compare modelled and observed isotope distributions. More recently, studies have applied stable isotopes to examine key hydrological processes associated with extreme weather events^{28–30}.

Despite the global availability of water isotope data through platforms like the Global Network of Isotopes in Precipitation³¹ (GNIP) and the Waterisotopes Database³² (wiDB), isotopic data coverage in Peru remains sparse and inconsistently updated. This represents a significant gap for researchers and decision-makers aiming to evaluate climate–water interactions, validate hydrological and climate models, or support environmental planning.

¹Subdirección de Ciencias de la Atmósfera e Hidrosfera, Instituto Geofísico del Perú (IGP), Lima, 15012, Peru.

²Programa de Maestría en Recursos Hídricos, Universidad Nacional Agraria La Molina (UNALM), Lima, 15024, Peru.

✉e-mail: romeroroldancarol@gmail.com

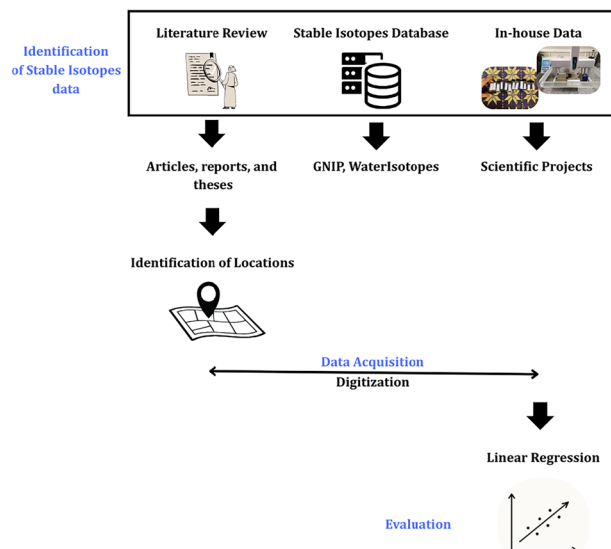


Fig. 1 Workflow of the water stable isotopes dataset in Peru (WSIPeruDDB) development.

Here, we present a comprehensive, curated dataset of water stable isotope measurements ($\delta^{18}\text{O}$, $\delta^2\text{H}$) from meteoric and terrestrial waters across Peru (WSIPeruDDB)³³, compiled from published studies, technical reports, theses, and global databases. This dataset aggregates observations from diverse sources, some of which are originally presented and have not been previously integrated into international repositories. By consolidating and standardizing these data, we aim to improve accessibility and enable broader reuse in hydrological, climatological, and ecological studies.

Methods

The development of the water stable isotopes dataset followed three steps: (i) collection of the isotopic information from published data, including scientific articles, reports and theses, project partners, and public databases; (ii) the identification of isotopic station's locations across Peru; (iii) evaluation of the stable isotopes from in-house data; (iv) creation of standardized metadata templates. The workflow is shown in Fig. 1.

Isotopic Dataset Compilation from scientific articles, reports, theses, project partners, and public databases.

We compiled the WSIPeruDDB dataset through extensive retrieval and transcription of information from a variety of sources such as scientific articles, technical reports, theses, and global databases. All content was standardized using a structured metadata template (described in Table 1) to ensure consistency and usability across the entire dataset. These sources include both in-house data and external datasets, each clearly referenced in the WSIPeruDDB_site_information.xlsx hosted at the WSIPeruDDB Zenodo repository³³ in accordance with their respective data-sharing policies.

Among the 489 water collection points, 17 are associated with the GNIP³¹, accessible via the International Atomic Energy Agency (IAEA) WISER platform (<https://nucleus.iaea.org/wiser>). While the metadata for these points is included in our dataset, users must register on the GNIP platform to access the full data, in accordance with IAEA's terms of use. An additional 34 collection points originate from datasets hosted in the Purdue University research repository. Twelve of these are linked to the study by Welp *et al.*³⁴ (<https://purr.purdue.edu/publications/4121/1>), and 22 to the work of Alvarez-Campos *et al.*³⁵ (<https://purr.purdue.edu/publications/3919/citations/1>). Both datasets are openly available for academic use and do not require registration. Furthermore, we incorporated metadata from 347 isotopic collection points listed in the WaterIsotopes database³² managed by the University of Utah (<https://wateriso.utah.edu/waterisotopes/index.html>), which is also publicly accessible.

In all cases, we followed the respective data contributor policies. For all databases mentioned above, only metadata and site identifiers were included in the WSI-Peru dataset to ensure compliance with reuse restrictions. Users are directed to the original repositories for full data access when required. Each data point included in the WSIPeruDDB dataset is accompanied by a clear source reference by adding the project ID and database location, as shown in Table 1, enabling full traceability.

Identifying isotopic station locations. To collect geographic information from the scientific article by Lambs *et al.*²⁶, we used the open-source Geographic Information System (GIS) tool QGIS (Quantum GIS)³⁶ to georeference the map included in the manuscript. We extracted the figure from the article and used the Layer → Georeferencer function in QGIS to georeference the image. This process generated a TIFF file, which was overlaid onto a base map of Peru. As a result, we were able to accurately determine the geographic coordinates of each isotopic station presented in the study as shown in Fig. 2b.

	Description
ID	Station coding with water type classification based on projects. We handle the coding
Station	Name of the station
Latitude	Latitude in decimal degrees
Longitude	Longitude in decimal degrees
Altitude	Elevation above sea level in meters
Department	Region of the study area
Start Date	Start Date Time for collection of the time-integrated sample, in local time, time in 24 hr format
End Date	End Date Time for collection of the time-integrated sample, in local time, time in 24 hr format
Sampling frequency	The temporal resolution of the sample: daily, biweekly, monthly, annual
Sample Type	precipitation, river, spring, lake, ice, tap, mine, wetland, groundwater, soil, snow, cave drip water
Number of $\delta^{18}\text{O}$ data	Number of $\delta^{18}\text{O}$ data
Number of $\delta^2\text{H}$ data	Number of $\delta^2\text{H}$ data
$\delta^{18}\text{O}$ analytical precision	analytical standard deviation
$\delta^2\text{H}$ analytical precision	analytical standard deviation
Contact	Contact name
Contact email	Contact email
References	Article or page citation where the database is located
Project_ID	It refers to the ID used if it is stored in another repository
Database	Name or link of the database related to the project ID

Table 1. Template with the categories for site information.

Evaluation of the stable isotopes from in-house data. In-house data refer to stable isotope measurements collected by our research institution^{15,16,20} and project collaborators^{17–19}. All datasets were collected following IAEA standard protocols for water sampling for isotopic analysis. Samples were stored in high-density polyethylene (HDPE) bottles to prevent evaporation. Isotopic compositions were analyzed using laser absorption spectrometers: a Picarro L2120i at the University of Brasília and a Picarro L2130i at the University of São Paulo. Post-processing was performed using the Laboratory Information Management System (LIMS) for laser-based and light-stable isotope data.

Calibration was carried out using international standards (VSMOW), and all isotopic values are reported relative to VSMOW. Analytical precision for each dataset is detailed in the supplementary file *WSIPeruDDB_site_information.xlsx*, available via the Zenodo repository.

To ensure data quality, we implemented a validation procedure based on the Local Meteoric Water Line (LMWL), which characterizes the linear relationship between $\delta^{18}\text{O}$ and $\delta^2\text{H}$ in precipitation for a given location or region. Data points falling outside the 3σ range of the LMWL regression were considered outliers and excluded. This quality control procedure follows the IAEA's recommended guidelines for isotope data processing³⁷. Additionally, the resulting LMWLs were compared to the Global Meteoric Water Line (GMWL), as described by Craig³⁸ ($\delta^2\text{H} = 8 \times \delta^{18}\text{O} + 10$). Only data that passed this validation step were included in the WSIPeruDDB dataset.

Data Records

The dataset is publicly available through the WSIPeruDDB Zenodo repository³³ and includes stable isotope data and associated metadata from 489 water collection points across 19 regions of Peru, covering the period from 2000 to 2021. The dataset is organized into four main files; all provided in open formats to support accessibility and reuse.

1. **WSIPeruDDB_site_information.xlsx.** This file contains metadata for each collection site, including station name, geographic coordinates, elevation, region, sample type, sampling period and frequency, and source project identifiers.
2. **WSIPeruDDB_dataset_information.xlsx.** This file includes in-house stable isotope measurements ($\delta^{18}\text{O}$, $\delta^2\text{H}$, and d-excess), along with the corresponding sampling date, location data and meteorological data.
3. **WSIPeruDDB_template.xlsx.** A standardized Excel template designed for future data contributions. It contains four worksheets: two for metadata descriptors (site and dataset information), and two editable sheets for users to input new collection data. Categories and field descriptions are summarized in Tables 1 and 2.
4. **site_information_numeral.geojson.** A geospatial file compatible with GIS platforms, showing the spatial distribution of collection points along with relevant site metadata.

The WSIPeruDDB dataset covers a wide diversity of environmental sample types. To facilitate their identification, we have assigned specific labels to the sample type: 01 for rainwater, 02 for river water, 03 for spring, 04 for lake, 05 for ice, 06 for tap, 07 for mine, 08 for wetland, 09 for groundwater, 10 for soil, 11 for snow, and 12 for cave drip water. The distribution of these sample types of our dataset can be shown in Fig. 2a. These codes are integrated into the ID using a numbering system based on project identifiers of the water collector points. This choice

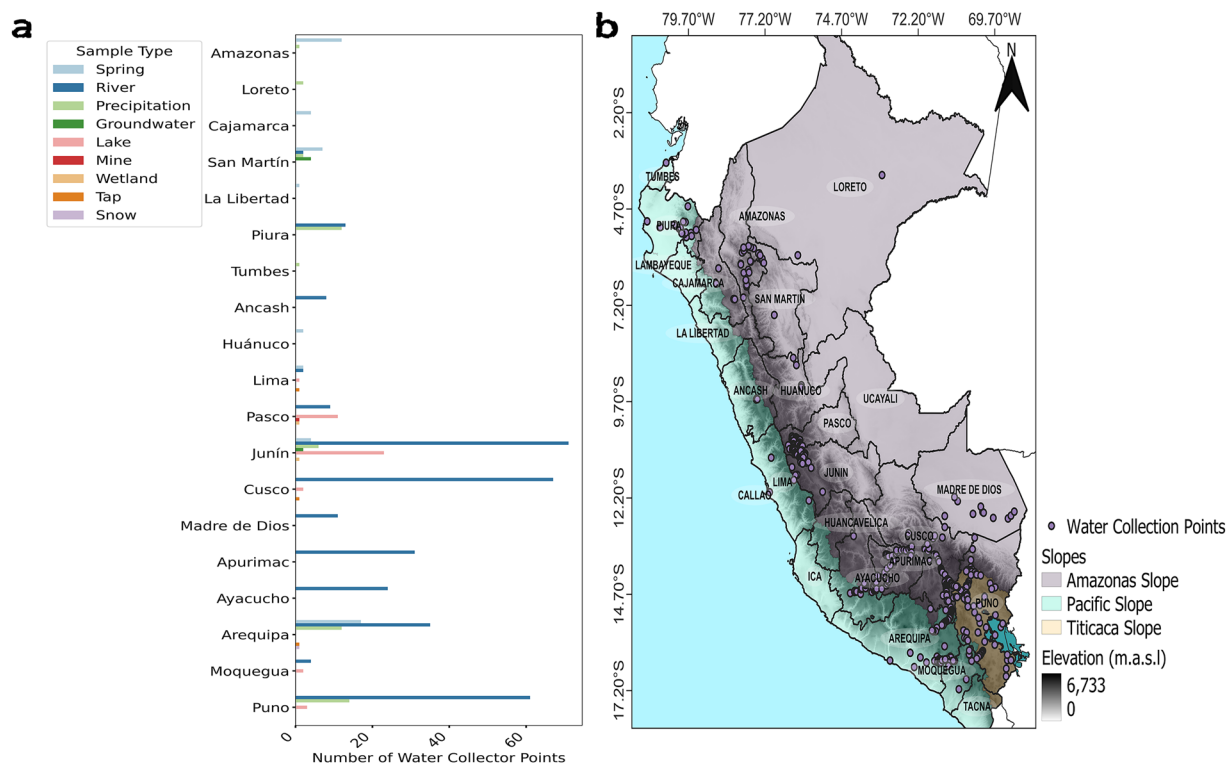


Fig. 2 (a) Water collection points in Peru according to the sample type across the WSIPeruDDB dataset. (b) Geospatial distribution of water collection points (purple circles) in Peru. Shaded areas represent the slopes: the Amazonas Slope (grey color), the Pacific Slope (cyan color), and the Titicaca Slope (brown color). The topography was derived from SRTM (Shuttle Radar Topography Mission) data with a resolution of 90 meters.

	Description
ID	Station coding with water type classification based on projects. We handle the coding
Sample Collection Date	Datetime of sample collection in local time
$\delta^{18}\text{O}$	$\delta^{18}\text{O}$ data
$\delta^2\text{H}$	$\delta^2\text{H}$ data
Dxs	Dxs data
Sample Type	precipitation, river, spring, lake, ice, tap, mine, wetland, groundwater, soil, snow, cave drip water
Department	Region of the study area
Latitude	Latitude in decimal degrees
Longitude	Longitude in decimal degrees
Altitude	Elevation above sea level in meters
Meteorological data	Precipitation, snowfall

Table 2. Template with the categories for isotopic data information.

ensures the uniqueness of the project numbers, reducing the possibilities of changes or confusion, and provides greater scalability, especially with the expectation of new water collector point contributions in the future.

Each external dataset included in WSIPeruDDB is referenced according to its original citation and detailed in the accompanying metadata file. Users interested in full datasets from external sources are directed to the original repositories, in compliance with each source's data-sharing policy.

It is important to highlight that researchers who are interested in contributing to new water collector points are welcome to upload their data to our WSIPeruDDB dataset. We encourage researchers to archive their data as it can be driven by various factors such as scientific journal requirements, grant funding obligations, or even a genuine commitment to supporting the open data initiative.

Technical Validation

To ensure the reliability of the dataset, all in-house stable isotope measurements were subjected to a quality control procedure based on the LMWL. A linear regression was performed, and data points falling outside three standard deviations (3σ) from the regression line were identified as outliers and excluded. This method follows the quality assessment guidelines provided by the IAEA.

For external sources such as GNIP and the WaterIsotopes database, only site information as shown in Table 1 were incorporated into the WSIPeruDDB dataset. No actual isotope measurement data from these sources were included. This approach is intended to facilitate users access to those databases while ensuring compliance with their respective data-sharing policies.

Code availability

The *WSIPeruDDB* Python package is an open-source and provides tools for visualizing the linear regression of the in-house stable isotope data. The code and documentation are available on GitHub at <https://github.com/karoru23/WSI-PeruDDB>. A public-facing visualization of the data is also accessible via the Instituto Geofísico del Perú (IGP) - institutional geospatial viewer at <https://ide.igp.gob.pe/geovisor/isotopicas/>. There are no restrictions on access or use.

Received: 7 October 2024; Accepted: 16 June 2025;

Published online: 12 July 2025

References

- Oki, T., Entekhabi, D. & Harrold, T. I. The global water cycle. in *Geophysical Monograph Series* vol. 150 225–237 (Blackwell Publishing Ltd, 2004).
- Intergovernmental Panel on Climate Change. Water Cycle Changes. in *Climate Change 2021 – The Physical Science Basis* 1055–1210 (Cambridge University Press, 2023).
- Aggarwal, P. K., Fröhlich, K., Kulkarni, K. M. & Gourcy, L. L. Stable isotope evidence for moisture sources in the Asian summer monsoon under present and past climate regimes. *Geophys Res Lett* **31**, 2–5 (2004).
- Fiorella, R. P. *et al.* Spatiotemporal variability of modern precipitation $\delta^{18}\text{O}$ in the central Andes and implications for paleoclimate and paleoaltimetry estimates. *Journal of Geophysical Research: Atmospheres* **120**, 4630–4656 (2015).
- Rodgers, P., Soulsby, C. & Waldron, S. Stable isotope tracers as diagnostic tools in upscaling flow path understanding and residence time estimates in a mountainous mesoscale catchment. *Hydrol Process* **19**, 2291–2307 (2005).
- Zhou, J. *et al.* Using stable isotopes as tracer to investigate hydrological condition and estimate water residence time in a plain region, Chengdu, China. *Sci Rep* **11** (2021).
- Wei, X. *et al.* Article hydrologic analysis of an intensively irrigated area in southern Peru using a crop-field scale framework. *Water (Switzerland)* **13**, 1–27 (2021).
- Somers, L. D. *et al.* Quantifying groundwater–surface water interactions in a proglacial valley, Cordillera Blanca, Peru. *Hydrol Process* **30**, 2915–2929 (2016).
- Glas, R. *et al.* A review of the current state of knowledge of proglacial hydrogeology in the Cordillera Blanca, Peru. *Wiley Interdisciplinary Reviews: Water* **5** (2018).
- Takahashi, K. *The atmospheric circulation associated with extreme rainfall events in Piura, Peru, during the 1997–1998 and 2002 El Niño events.* vol. 22 (2004).
- Drenkhan, F., Carey, M., Huggel, C., Seidel, J. & Oré, M. T. The changing water cycle: climatic and socioeconomic drivers of water-related changes in the Andes of Peru. *Wiley Interdisciplinary Reviews: Water* **2**, 715–733 (2015).
- Takahashi, K., Montecinos, A., Goubanova, K. & Dewitte, B. ENSO regimes: Reinterpreting the canonical and Modoki El Niño. *Geophys Res Lett* **38** (2011).
- Peng, Q., Xie, S. P., Wang, D., Zheng, X. T. & Zhang, H. Coupled ocean–atmosphere dynamics of the 2017 extreme coastal El Niño. *Nat Commun* **10** (2019).
- Garreaud, R. D. A plausible atmospheric trigger for the 2017 coastal El Niño. *International Journal of Climatology* **38**, e1296–e1302 (2018).
- Apaéstegui, J. *et al.* Hydroclimate variability of the northwestern Amazon Basin near the Andean foothills of Peru related to the South American Monsoon System during the last 1600 years. *Climate of the Past* **10**, 1967–1981 (2014).
- Apaéstegui, J., Romero, C., Vuille, M., Sulca, J. & Ampuero, A. Moisture Sources and Rainfall $\delta^{18}\text{O}$ Variability over the Central Andes of Peru—A Case Study from the Mantaro River Basin. *Water (Basel)* **15**, 1867 (2023).
- Ampuero, A. *et al.* The Forest Effects on the Isotopic Composition of Rainfall in the Northwestern Amazon Basin. *Journal of Geophysical Research: Atmospheres* **125** (2020).
- Hidalgo L. Rôle d'un karst andin tropical (Alto Mayo, Pérou) sur la dynamique de production de matériel dissous vers l'Amazone: analyse du fonctionnement hydrogéologique et des flux associés. PhD thesis, Sorbonne Université (2020).
- Ulloa F. Étude de l'hydrogéochimie des systèmes karstiques dans les Andes du Pérou. Contrôle de l'altération des carbonates en milieu orogénique tropical le long d'un gradient altitudinal et climatique. MSc thesis, Université Paris Sud, Orsay (2018).
- Apaéstegui, J. & Romero, C. Reconociendo fuentes de humedad durante los eventos de precipitación extrema en el norte del Perú. Boletín técnico: Generación de modelos climáticos para el pronóstico de la ocurrencia del Fenómeno El Niño. *Instituto Geofísico del Perú* **6**(2), 10–12 (2019).
- IAEA. Estudios de hidrología isotópica en América Latina 2006. Colección de Documentos Técnicos del OIEA. *Organismo Internacional de Energía Atómica* **04**, 153–173 (2009).
- Giovanni, M. K., Horton, B. K., Garzzone, C. N., McNulty, B. & Grove, M. Extensional basin evolution in the Cordillera Blanca, Peru: Stragigraphic and isotopic records of detachment faulting and orogenic collapse in the Andean hinterland. *Tectonics* **29**, TC6007 (2010).
- Katz, S. A. *et al.* Detecting hydrologic distinctions among Andean lakes using clumped and triple oxygen isotopes, Earth and Planetary Science Letters, 602 (2023).
- Bershaw, J., Saylor, J. E., Garzzone, C. N., Leier, A. & Sundell, K. E. Stable isotope variations ($\delta^{18}\text{O}$ and δD) in modern waters across the Andean Plateau. *Geochimica et Cosmochimica Acta* **194**, 310–334 (2016).
- Aron, P. G. *et al.* Triple oxygen isotopes in the water cycle. *Chemical Geology* **565**, ISSN 0009–2541 (2020).
- Lambs, L., Horwath, A., Otto, T., Julien, F. & Antoine, P. O. Isotopic values of the Amazon headwaters in Peru: comparison of the wet upper Río Madre de Dios watershed with the dry Urubamba–Apurímac River system. *Rapid Communications in Mass Spectrometry* **26**, 775–784 (2012).
- Risi, C. *et al.* Process-evaluation of tropospheric humidity simulated by general circulation models using water vapor isotopologues: 1. Comparison between models and observations. *Journal of Geophysical Research Atmospheres* **117**, 1–26 (2012).
- Nishimura, H. *et al.* A triggering mechanism of quasi-stationary convective bands in the vicinity of southwestern Japan during the summer season as deduced from moisture origins. *Atmospheric Res.* **308**, 107544 (2024).
- Li, X., Kawamura, R., Sugimoto, A. & Yoshimura, K. Isotopic composition and moisture sources of precipitation in midlatitude regions characterized by extratropical cyclones' route. *J. Hydrol.* **612**, 128047 (2022).
- Li, X., Kawamura, R., Ichiyana, K. & Yoshimura, K. Moisture sources and isotopic composition of a record-breaking heavy Meiyu-Baiu rainfall in southwestern Japan in early July 2020. *Atmospheric Res.* **286**, 106693 (2023).

31. IAEA/WMO. Global Network of Isotopes in Precipitation. *The GNIP Database* (2024).
32. Waterisotopes Database. <http://waterisotopesDB.org> (2024).
33. Romero, C. & Apaestegui, J. A Nationwide Dataset of Stable Isotopes in Meteoric and Terrestrial Water Across Peru (0.2.2) [Data set]. *Zenodo*. <https://doi.org/10.5281/zenodo.15625378> (2025).
34. Welp, L. R. *et al.* Reinterpreting Precipitation Stable Water Isotope Variability in the Andean Western Cordillera Due to Sub-Seasonal Moisture Source Changes and Sub-Cloud Evaporation. *Geophys Res Lett* **49** (2022).
35. Alvarez-Campos, O. *et al.* Evidence for high-elevation salar recharge and interbasin groundwater flow in the Western Cordillera of the Peruvian Andes. *Hydrol Earth Syst Sci* **26**, 483–503 (2022).
36. QGIS Development Team. QGIS Geographic Information System, version 3.22.1. *Open-Source Geospatial Foundation Project* <https://qgis.org/en/site/> (2024).
37. IAEA. *Statistical treatment of data on environmental isotopes in precipitation*. *Int Atomic Energy Agency Tech Rep Ser* (1992).
38. Craig, H. Isotopic Variations in Meteoric Waters. *Science* (1979) **133**, 1702–1703 (1961).

Acknowledgements

This research was partially supported by the CHARISMA Project (JE0ECCHARI, JEAI-IRD), and PROCENCIA about contract No. 124-2020 FONDECYT, Perú.

Author contributions

C.R. developed the WSI-PeruDB dataset. The dataset collection was the result of a synergistic collaboration between C.R. and J.A., with each bringing unique insights to the table. In the initial stages, C.R. took the lead in drafting the manuscript, with valuable input from J.A. Subsequently, both C.R. and J.A. reviewed the original manuscript, culminating in the final revision.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025