Journal of the Royal Society of New Zealand

RESEARCH ARTICLE OPEN ACCESS

Community Participation as a Pathway to Strengthen National Precipitation Isotope Products in New Zealand

Bruce D. Dudley[1,2] | Alice F. Hill[3] | Andy McKenzie[4] | Peter W. Holder[5] | W. Troy Baisden[6] | Stephen Collins[7] | James Dare[8] | Michael Darling[9] | Russell D. Frew[9] | Scott L. Graham[10] | Robert van Hale[9] | Travis W. Horton[2] | Elizabeth D. Keller[11,12] | Shelley MacDonell[2] | Jack McMecking[2] | Jeffrey J. McDonnell[13,14,15] | Magali F. Nehemy[16] | Ewen Rodway[17] | Zane Shadbolt[2] | Cerra Simmons[16] | Andrew Swales[18] | Michael K. Stewart[11] | Vanessa Trompetter[11] | Thomas Wöhling[19,20]

[1]Earth Sciences New Zealand, Christchurch, New Zealand | [2]University of Canterbury, Christchurch, New Zealand | [3]New Zealand Department of Conservation, Nelson, New Zealand | [4]New Zealand Qualifications Authority, Wellington, New Zealand | [5]Lincoln University, Lincoln, New Zealand | [6]Motu Economic and Public Policy Research, Wellington, New Zealand | [7]Horizons Regional Council, Palmerston North, New Zealand | [8]Bay of Plenty Regional Council, Tauranga, New Zealand | [9]Oritain Global Limited, Dunedin, New Zealand | [10]New Zealand Institute for Bioeconomy Science, Lincoln, New Zealand | [11]Earth Sciences New Zealand, Wellington, New Zealand | [12]Victoria University of Wellington, Wellington, New Zealand | [13]School of Environment and Sustainability, University of Saskatchewan, Saskatoon, Canada | [14]North China University of Water Resources and Electric Power, Zhengzhou, China | [15]School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK | [16]Department of Earth and Environmental Science, University of British Columbia Okanagan, Kelowna, Canada | [17]Environment Southland, Invercargill, New Zealand | [18]Earth Sciences New Zealand, Hamilton, New Zealand | [19]Dresden University of Technology, Dresden, Germany | [20]Lincoln Agritech Limited, Christchurch, New Zealand

Correspondence: Bruce D. Dudley (bruce.dudley@earthsciences.nz)

Received: 13 February 2026 | **Revised:** 9 June 2026 | **Accepted:** 11 June 2026

Keywords: environmental datasets | isoscapes | machine learning | New Zealand | precipitation isotopes | spatial prediction | stable isotope hydrology | XGBoost

ABSTRACT

Accurate environmental and climate modelling depends on the availability of large observational datasets, yet the generation of such data is often costly and logistically challenging for individual research groups. This limitation is particularly acute in New Zealand, where geographic isolation, complex topography, and strong climatic gradients require dense observational networks for robust national and regional scale modelling. Here we present New Zealand's national precipitation isotope database, compiled through a community data-collection effort. The database integrates historical and contemporary datasets with varied analytical methods, sampling strategies, and temporal coverage. To enable combined use of these data, we describe a standardisation framework, including weighting using gridded precipitation data, and alignment of sampling periods to calendar months. We then show how the database updates and improves the Precipitation Isotopes New Zealand (PINZ) machine-learning isoscapes. Incorporation of newly contributed data, particularly from inland and high-elevation regions, reduces overall root mean square error relative to earlier model versions and substantially expands the model's area of applicability in New Zealand. Our results highlight that further advances will benefit from the availability of submonthly precipitation isotope measurements and improved coverage of alpine environments. We show that community participation offers a practical pathway to strengthening national precipitation isotope products and their application across hydrology, climate science, and related environmental science disciplines.

1 | Introduction

The stable isotopes of hydrogen and oxygen are now crucial data in hydrological and climate research. They integrate information across scales, and their patterns can reveal processes—such as recharge sources, water pathways, and residence times—that control future changes to the local and regional water cycles (Bowen 2010; Bowen et al. 2019). Precipitation is the "input" to hydrological systems and therefore perhaps the most important and widely developed data to be mapped and shown as isoscapes (Rozanski et al. 1993) and modelling techniques are rapidly advancing that enable increasingly detailed representations of precipitation isotope signals in complex climate systems (Erdélyi et al. 2023; Eyring et al. 2024).

Yet accuracy of both machine-learning isoscape predictions and statistically downscaled global climate model outputs remains constrained by the availability of high-quality isotope data. Machine-learning models require large, diverse training datasets to perform reliably on new conditions (Reichstein et al. 2019; Meyer et al. 2021), while downscaling approaches benefit—either directly or indirectly—from dense networks of rainfall measurements for calibration and validation (Breil et al. 2021; Lima et al. 2021; Wootten et al. 2021). As a result, advances in precipitation isotope modelling largely hinge on access to extensive in situ datasets, and the data-generation process for these is often prohibitively expensive or cumbersome for a single research team or institution. At the same time, it is common for multiple teams to be, or have been, working on similar topics generating data with the potential to be combined. Leveraging existing data from within the research community or from citizen scientists—to increase both spatial coverage and length of time series—can be an efficient way to augment datasets and overcome resource constraints (Zhong et al. 2021; Fraisl et al. 2022).

However, while simple in concept, developing a collaborative community dataset requires management, curation, and standardization of data across projects to be of use to the broader community, and this can be a time consuming and complex task. Data management considerations include sampling and sample storage protocols, methods to ameliorate timestep mismatches, and quality-control of metadata. Also, credit and citation expectations for participating people and funding bodies must be addressed. These technical and logistical issues require time and effort within the community.

A collaborative database approach is particularly valuable for supporting resource-intensive environmental datasets in geographically isolated and environmentally unique countries such as New Zealand, especially given its relatively small research funding landscape. However, the benefits of this approach extend well beyond New Zealand and are equally applicable at a global scale. Other regions with complex climatic variation face similar challenges, and dense data networks such as those we present here can serve as effective testbeds for applications like downscaling isotope-enabled climate model outputs (Watson et al. 2024), or machine-learning predictions across alpine regions (Lei et al. 2025). In this study, we describe the process used to compile a community-sourced dataset and present the resulting national precipitation isotope database for New Zealand. We then apply this community database in environmental research to improve a machine-learning precipitation isoscape for New Zealand, which in turn provides a powerful, publicly available modelled dataset that can be used in fields spanning paleoclimate to hydrology to forensics.

2 | Methods

2.1 | Initial Data Collation

Following a presentation at the 2024 New Zealand Hydrological Society Conference, institutions and researchers holding precipitation isotope data were invited to contribute to the development of a national database. Those who responded were provided with a standardised data spreadsheet template (modified from that of Bowen 2025) and an accompanying methods table (following Hill et al. (2025)) to provide consistent reporting of site metadata, sample information, and analytical methods. Example entries were supplied to illustrate expected data formats, date conventions, and units of measurement. These template formats are largely retained in the metadata file provided as supplementary material to this manuscript.

2.2 | Data Processing

All data processing and analysis were performed in R version 4.5.1 (R Core Team 2025). The full workflow included importing and quality controlling measurements of precipitation isotope values and precipitation amounts, then integrating precipitation isotope measurements with gridded (interpolated) daily precipitation estimates to generate weighted δ^2H and δ^{18}O values at monthly timestep and 5 km resolution.

All submitted datasets were reviewed for completeness, adherence to template requirements, and consistency of units prior to integration into the database. Dates were standardised to ISO format and checked for potential errors (e.g., day–month inversion in U.S. date formats). Sequential consistency was verified visually by checking that collection dates increased monotonically within each site record. Any anomalies or reversed date sequences were inspected and corrected manually.

Geographic coordinates were verified by manual checking of mapped locations. Sites with isotope data that could not be converted to monthly values (e.g., one-off event samples, or snow samples with uncertain timing) were retained in the database but excluded from subsequent modelling.

2.3 | Community Database Demonstration in New Zealand Using PINZ

Precipitation Isotopes NZ (PINZ) is modelled δ^{18}O and δ^2H isoscapes generated using the Extreme Gradient Boosting (XGBoost) machine learning algorithm. PINZ is driven by a regional precipitation isotope database, uses a combination of geographic, temporal, and climate predictors, and produces a national monthly precipitation isotope model at 5 km resolution (Hill et al. 2025; McKenzie et al. 2025).

The dataset used for training PINZ described in Hill et al. (2025) includes long-term time series from two International Atomic Energy Agency (IAEA) Global Network of Isotopes in Precipitation (GNIP) sites, two cross-alpine transects to target orographic processes, a snow-sampling citizen science campaign, as well as datasets from previous publications (e.g., Baisden et al. 2016; Dudley et al. 2018). Nevertheless, that combined dataset contained clear spatial deficiencies—including gaps in the central North Island, and an underrepresentation of high-elevation areas—that restrict its appropriateness for national scale analyses (Dudley et al. 2025). Collecting precipitation isotope data is inherently laborious and is difficult in remote or high-elevation regions where coverage is most needed. The XGBoost modelling framework requires on the order of thousands of data points to develop robust models in real-world scenarios, although more data are required for cross-validation to ensure training and validation sets are representative. As a result, obtaining sufficient spatially and temporally representative data remains a major limitation for national-scale use of PINZ—one that would be difficult to overcome for any single research group.

We repeated the 'PINZ' modelling method described in Hill et al. (2025) using both the dataset described in that paper, and the expanded dataset resulting from this community outreach project for which the vast majority of the data required both virtual climate station network (VCSN) rainfall weighting and temporal alignment to calendar months, as described below. Model performance was evaluated using the 20% of test dataset aside from the original dataset within δ^2H and δ^{18}O quartile strata to ensure representative distribution of data in the test set. Predictor variable data and methods for model structure, and assessment of model performance and uncertainty were retained from Hill et al. (2025), while the code updates for this manuscript include spatial cross-validation in the hyperparameter tuning (i.e., optimisation of user-defined model settings controlling tree complexity and learning behaviour). Full methodological details are provided in Hill et al. (2025), which should also be consulted when using the updated code provided with this paper (McKenzie et al. 2025).

As a final examination of the individual and regional benefits of data contributions such as those provided by the co-authors of this work, we show changes in the spatial reliability of the PINZ model using an area of applicability (AOA) approach (Meyer et al. 2021), using national datasets available before and after this community outreach effort. The AOA is the area where the prediction data are similar to the training data and the estimated cross-validation prediction performance holds. AOA was calculated using the R package waywiser (Meyer et al. 2024).

2.4 | Rainfall Weighting and Temporal Aggregation

Many isotope values in the community database did not have accompanying precipitation amount data with which to generate amount-weighted values at monthly timestep. To address this, we obtained daily precipitation amount data from the VCSN, a national gridded dataset providing interpolated estimates of key climate fields at 5 km spatial resolution (Tait et al. 2006). Each isotope sampling site was assigned to the nearest VCSN grid point using the sf package in R (Pebesma 2018). Because multiple sites assigned to the same VCSN point caused problems with later PINZ training and predictions, sites within 2 km of each other were assigned a single site location (and VCSN point) at this step.

For each site, daily VCSN precipitation data were extracted for the full range of isotope sampling dates (first deployment to last collection). For all months, isotope values (δ^2H and δ^{18}O) were assigned to daily intervals between the 'Start Date' and 'Collection Date' of each isotope sample. Where exact collection times were unavailable, sampling periods were treated as inclusive of the start date and exclusive of the collection date. Each day within a collection interval inherited the same isotopic composition measured for its composite sample. These daily isotope estimates were then joined with the rainfall record from the nearest VCSN grid point to produce a continuous daily time series of precipitation and isotopic composition for each site, from which monthly values were calculated. Specifically, for each month m and site i, the weighted mean was calculated as

$$\delta_{\mathrm{m,i}} = \frac{\sum_{\mathrm{d}} \left(R_{\mathrm{d,i}} \cdot \delta_{\mathrm{d,i}} \right)}{\sum_{\mathrm{d}} R_{\mathrm{d,i}}} \tag{1}$$

where $R_{\mathrm{d,i}}$ is the daily rainfall amount (e.g., in mm) for site i on day d, and $\delta_{\mathrm{d,i}}$ is the isotope composition (δ^2H or δ^{18}O) for that day.

Months with incomplete isotope records were flagged as incomplete in the database used for modelling, and the number of missing days noted. We removed a small number of δ^2H and δ^{18}O values of exactly zero (caused by erroneous zero monthly rainfall values in VCSN records) and identical δ^2H and δ^{18}O values over 2 or more months (caused by collectors being left unattended across months) during data processing. We flagged data rows in the resulting dataset to indicate the steps involved in processing:

- 'composite' = TRUE if multiple samples contributed to that month's estimate,
- 'exact month' = TRUE if sampling periods exactly match the calendar month, including continuous submonthly samples (e.g. daily sampling covering the whole month).
- and 'incomplete month' or 'sparse month' flags indicate any days and ≥10 days of missing data, respectively.

2.5 | Testing for the Effects of Rainfall Weighting and Temporal Alignment

We note that the method above causes no change to isotope values representing a full calendar month (e.g., where a single composite sample is collected on the morning of the first day of each month). Thus, the majority of monthly δ^2H and δ^{18}O records available for New Zealand—those from GNIP stations and those of the database described in Baisden et al. (2016)—are unaltered in the database used for our model training. Our approach also allows the inclusion of a great number of monthly δ^2H and δ^{18}O estimates from collection time series across dates not matched exactly to calendar months (i.e., where 'exact month' = FALSE), and those where precipitation amounts measured on site were not available. However, both the correction of temporal alignment to calendar months, and the use of VCSN data for rainfall weighting introduce error into monthly δ^2H and δ^{18}O values. To address these issues, we use sites with higher-frequency measurements of precipitation amount and δ^2H and δ^{18}O (at event, daily or weekly time scale) to estimate the error added by these two steps. We selected five sites for this demonstration: Ashley Dene (weekly sampling), B3 Kaitaia (daily to weekly sampling), B3 Christchurch (daily to weekly sampling), Melrose Farm (daily sampling), and Maimai M8 catchment (subdaily sampling). At these sites we compared monthly values weighted with VCSN rainfall data against monthly values weighted with precipitation amounts measured on site. We then tested the sensitivity of monthly weighted isotope values to collection-date frame shifts, simulating offsets (±5, ±10, ±15 days) from ideal calendar-month sampling. This allows assessment of how deviations in sampling period alignment affect monthly rainfall-weighted isotope estimates. We calculated root mean square error

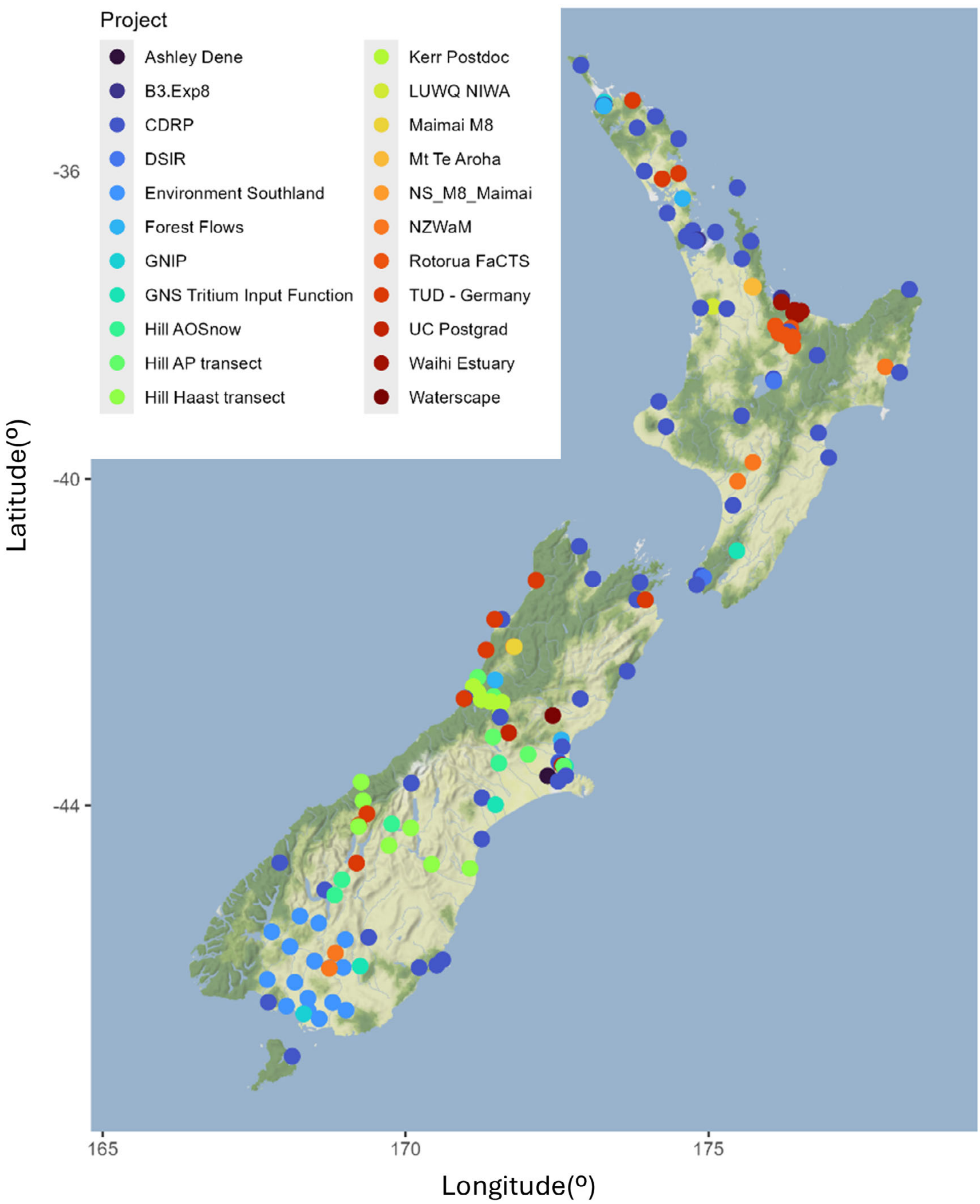

FIGURE 1 | Locations of monitoring sites, split by contributing project. Contributing projects vary in funding source (e.g., University research funding, or applied science projects from New Zealand regional government), and institutional origin (e.g., visiting scientists from non-New Zealand universities). Project codes are shown in abbreviated form, with both abbreviated and full project names available in metadata provided as Supporting Information.

(RMSE) resulting from our methods for each demonstration site, and for all months in this reduced dataset.

Code for the data processing steps described above is available in the supplementary code and data repository accompanying this work (Dudley et al. 2026).

3 | Results

3.1 | Composition and Coverage of the National Precipitation Isotope Database

Data from 245 sampling stations are included in the database, with 148 sampling stations reporting data from precipitation (rather than snowpack) samples. Of these 148 precipitation stations, 108 provide monthly time series, 18 provide isotope time series at intervals between daily and monthly, two provide daily data, and 20 report ad-hoc or one-off observations. The locations of all precipitation sites in the expanded database are shown in Figure 1, colour-coded by contributing project. The highest number of observations are available at the two long-running GNIP stations in Kaitaia (241) and Invercargill (282), with a median of 26 observations per site across the 148-site precipitation network.

Figure 2 shows the effect of new data additions on the distributions of available data with respect to year of collection (Figure 2A) and elevation (Figure 2B). Of note are increases in data available prior to the year 2000 through the incorporation of the data reported in Taylor (1990), which included measurements from Taupo and Christchurch Airport not previously included in the GNIP dataset used by Hill et al. (2025). Additions post-2000 (which are included in model training) add balance to a dataset previously dominated by collections from the Cross Departmental Research Project (CDRP) analysis of rainwater collected from weather stations across New Zealand between 2007 and 2010 (Baisden et al. 2016). Notably, Figure 2B shows that these additions also add considerable data density to that available from middle elevations (~300–800 m) across New Zealand (a notable shortcoming of the previous database), as much of the new data originates from the Taupo and Rotorua areas, and at higher elevations in the Canterbury region. Further data details are available in the metadata file included as supplementary material, which describes station locations, project affiliations, contact details for data holders, sampling frequency, collection date ranges, and the number of $\delta^{18}O$ and $\delta^{2}H$ precipitation measurements at each site.

3.2 | Effects of Rainfall Weighting and Temporal Alignment on Monthly Isotope Values

Errors resulting from using interpolated gridded precipitation estimates instead of on-site precipitation measurements are summarised in Table 1. Across the five test sites, the root mean squared errors of the monthly estimates generated using this approach were 5.53‰ for $\delta^{2}H$ and 0.71‰ for $\delta^{18}O$. Errors associated with shifting the sampling dates away from the first day of the month ('frame shifts' – again, shown in Table 1) were approximately double those associated with precipitation weighting. Values were relatively consistent across the simulated 5-, 10-, and 15-day offsets. The effects of these treatments—intended to mimic monthly values derived from 'imperfect' time-series data—are also shown in plots in the supplementary material (Figures S1 and S2).

3.3 | PINZ Model Trained Using the Expanded Database

Gridded predictions of $\delta^{2}H$ and $\delta^{18}O$ from the PINZ model trained using the expanded database are exemplarily shown for February and August 2024 (Figure S3), and time series of monthly predictions are provided as supplementary material to this manuscript. These predictions cover New Zealand in its entirety at 5 km resolution, from 1997 to 2025.

For $\delta^{2}H$, the RMSE for the overall model (national scale) evaluated using the test data is 13.3‰, RMSE evaluated on the training data is 10.1‰, and RMSE evaluated on all data is 10.8‰. Comparable $\delta^{2}H$ RMSE statistics using the dataset of Hill et al. (2025) were 13.4, 10.8, and 11.4‰, respectively. For $\delta^{18}O$, the RMSE for the overall model (national scale)

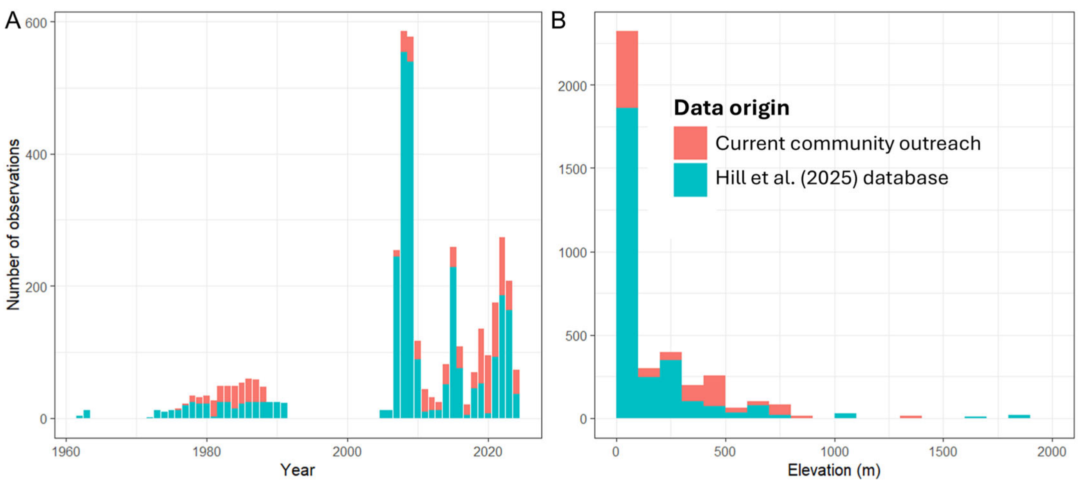

FIGURE 2 | Temporal and elevation origin of data added to New Zealand's national-scale precipitation isotope database from this community outreach project in addition to that used for training the model of Hill et al. (2025).

TABLE 1 | Root mean square errors (RMSE) associated with shifts in monthly sampling date, and use of interpolated (VCSN) rainfall data for weighting for monthly precipitation δ^2H and $\delta^{18}O$ values.

Site name	Isotope	RMSE 5-day date shift	RMSE 10-day date shift	RMSE 15-day date shift	RMSE VCSN rainfall weighting
Ashley Dene	δ^2H	5.09	6.62	6.87	2.42
Ashley Dene	$\delta^{18}O$	0.54	0.72	0.75	0.33
B3. Kaitaia	δ^2H	4.19	6.23	7.24	3.45
B3. Kaitaia	$\delta^{18}O$	0.57	0.85	0.95	0.44
B3.Christchurch	δ^2H	9.52	11.9	12	9.6
B3.Christchurch	$\delta^{18}O$	1.86	2.22	2.29	1.2
Melrose Farm	δ^2H	17.7	19.1	15.5	3.52
Melrose Farm	$\delta^{18}O$	1.81	2.01	1.62	0.47
Maimai (M8)	δ^2H	18.7	19.6	18.7	2.79
Overall (pooled)	δ^2H	10.4	12	11.1	5.53
Overall (pooled)	$\delta^{18}O$	1.32	1.56	1.52	0.71

Note: Shown are results for five example sites where rainfall was collected on an event-by-event basis. Overall (pooled) RMSE values give the overall error across all months in the 5-site dataset.

evaluated using the test data is 1.75‰, evaluated on the training data is 1.37‰, and evaluated on all data is 1.45‰ (1.76, 1.37, and 1.46‰, respectively, using the Hill et al. (2025) dataset). The overall R^2 of monthly δ^2H predictions was 0.55 for test data, and 0.77 for training data (0.55 and 0.75 using the Hill et al. (2025) dataset). For $\delta^{18}O$, the overall R^2 was 0.51 for test data, and 0.72 for training data (0.52 and 0.71, respectively, using the Hill et al. (2025) dataset). These results show a slight improvement in RMSE at test sites resulting from additional data, but a slight decrease in test R^2 for $\delta^{18}O$.

Plots of RMSE and median absolute deviation (MAD) at the site level are provided in Figure S4. In general, sites in alpine regions showed the greatest RMSE and MAD values; additional alpine sites in the Canterbury and Southland regions likely worsened overall model performance metrics.

3.4 | Changes to Model Spatial Reliability (Area of Applicability)

Changes to model reliability in space, resulting from the addition of data described in this study, are shown in Figure 3. Improvements were apparent at higher (but not the highest) elevations nationally, and also in areas where data density increased substantially. The inset to Figure 3 shows one such area, where data additions, largely from the Rotorua FaCTS sampling programme and sampling across an elevation range at Mount Te Aroha, added considerably to the AOA across the nearby Kaimai Range.

4 | Discussion

4.1 | On the Need to Share Climate Science Data in New Zealand

The need for data sharing in climate science is particularly acute in New Zealand. This is due first to the country's small economy and limited science-funding environment, and second to its geographic isolation, meaning environmental data collected in neighbouring regions are less applicable to New Zealand. In addition, the country's varied topography and climate require high-density observations for accurate modelling. New Zealand's island setting and long, narrow shape expose it to moisture sources from all directions, from cold Antarctic southerlies to warm, humid subtropical systems, resulting in highly variable weather patterns (Salinger 1980). Frequent atmospheric river events deliver high-intensity, multiday rainfall (Prince et al. 2021), while the Southern Alps, which run the length of the South Island, generate strong orographic precipitation gradients, with steep rises on the west (averaging 12% from sea to highest point) and pronounced rain shadows in the east. These interacting climatic and topographic controls result in strong spatial and temporal variability in the isotopic composition of precipitation across New Zealand (Baisden et al. 2016), placing particularly high demands on the density and coverage of observational isotope data.

Because of this variability, our community dataset has immense potential value as input into further local modelling and associated applications. For example, global-scale precipitation isotope models, while limited in spatial resolution, hold advantages over PINZ in terms of temporal resolution (Watson et al. 2024) and mechanistic representation of the climate processes that drive variation across New Zealand (Bong et al. 2024, 2026). As such, use of the database compiled in this work for high-spatial resolution correction of global models may provide a future pathway to accurate, event-scale precipitation isotope data for New Zealand, and other similarly climatologically complex countries. A key component of such advances is likely to be further additions of isotope data to our national database.

4.2 | Spatial and Temporal Representativeness of the National Precipitation Isotope Dataset

Early collections of precipitation isotopes in New Zealand were limited in site numbers, with initial sampling at Kaitaia, Taupo,

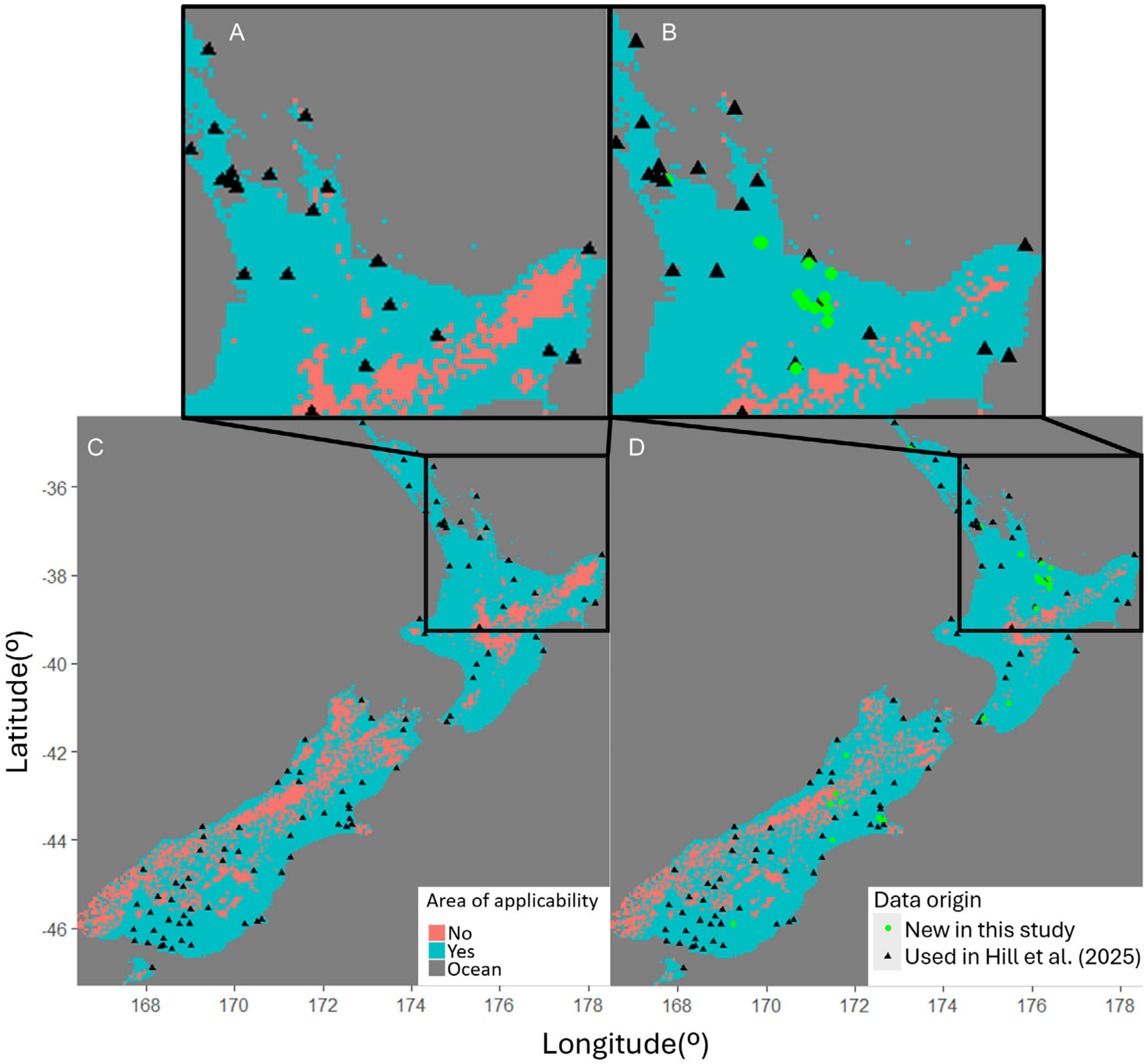

FIGURE 3 | Effects of data addition on Area of Applicability for the XGBoost spatial predictions model 'PINZ' without (Panels A and C) and with (Panels B and D) the data collated via community outreach described in this paper. Area of Applicability is the region in which the model has encountered similar predictor data characteristics to that during training and model prediction performance is similar (Meyer et al. 2021).

Christchurch and Invercargill contributing to the GNIP database (Taylor 1990). Of these early sites, sampling continued longest at Kaitaia and Invercargill, both near sea level, which biased long-term records toward coastal environments. As a result, global products built primarily from GNIP inputs provide weak representation of New Zealand's pronounced elevation-related variability (Dudley et al. 2024). The national-scale dataset described in Baisden et al. (2016) expanded national coverage considerably and remains foundational, but again, dictated by the practicalities of monthly sampling over several years, was also biased towards accessible, low-elevation areas. More recent additions of alpine sites (where rainfall volumes are disproportionately large (Caloiero 2014)) have improved representation of these hydrologically critical areas. It is notable that many of these recently added alpine sites feature among those with the poorest fit statistics to model outputs (as described in Hill et al. (2025) and shown in supplementary Figure S3). However, this variability should not be seen as environmental noise in this context, but instead a reflection of the true climatic, and isotopic, heterogeneity present at high elevations in New Zealand (Purdie et al. 2010; McMecking 2025). For this reason, these high-elevation sites are of immense value to understanding the isotopic composition of the precipitation that feeds much of New Zealand's river water and groundwater. Within the confines of the current PINZ model structure any reduction in statistical neatness in model results when including these sites is the cost of increased hydrological relevance. Improving model fits at these sites remains an important direction for future work.

4.3 | Benefits of Local Data Contributions for Model Applicability

A key outcome of this study is that new observations from previously under-represented regions directly improve the PINZ model's local predictive skill. AOA analysis demonstrates this effect: additional sites shrink extrapolation zones, thereby increasing confidence in predictions at and around the contributing locations. This provides a tangible return for data contributors; where short-term funding cycles provide resources for a limited period of data collection, this can be used to constrain model predictions for the region of collection over a much longer period. Across New Zealand, the incorporation of additional data from high-elevations and inland areas increased variance within the training pool and marginally reduced the R^2 fit to training data; however, overall RMSE fits improved relative to those of Hill et al. (2025), and R^2 performance in well-sampled lowland regions was not compromised. In practical terms, additional precipitation isotope datasets enhance the reliability of modelled precipitation isotopes where contributors most need it. These results suggest that community contributions of precipitation isotope data hold similar benefits at the local level to those of citizen scientist contributions of weather data, adding resolution at local-scale not resolved by large-scale networks (Muller et al. 2015; Medina et al. 2024). A key difference of our approach is that we have not relied on a single, centralised laboratory and/or standardised collection methods common to similar community initiatives for precipitation isoscapes (Terzer et al. 2013; Good et al. 2014; Cole et al. 2021).

While precipitation isoscape development for New Zealand does not necessarily require the continued development of the PINZ model, aspects of the PINZ approach that facilitate ongoing refinement include automated input of predictor variables from the VCSN dataset, the precoded capacity to track changes in 'area of applicability' using the method of Meyer et al. (2021) and consistent training architecture that enables the addition of new predictors and response data. The PINZ model operates at a monthly timestep because the vast majority of precipitation isotope data available for New Zealand has been taken from rainfall totalisers (see Michelsen et al. (2018)), at monthly timestep. However, almost none of the new data collated within this latest effort has been collected over exact calendar months. As described above, the estimation of monthly values from these data is not without error, and at most sites this cannot be quantified. For example, at high elevations, where VCSN rainfall estimates may be less accurate (Tait et al. 2012), error associated with using VCSN rainfall for rainfall weighting may be higher than the estimates in Table 1.

Nevertheless, the spatially continuous time series of monthly precipitation isoscapes provided by PINZ lay a foundation for a wide range of applications across fields such as plant physiology (Lin et al. 2022; Siegwolf et al. 2023), hydrology of groundwater (Schroeter et al. 2025) and rivers (Vystavna et al. 2024), nitrate provenance (Matiatos et al. 2023), and sediment transport (Hirave et al. 2023). PINZ and the community database described in this study are already supporting applications in New Zealand including tracer-based studies of river water age (Dudley et al. 2025), susceptibility of planted forests to climate change (Klein et al. 2026) and isotope-enabled hydrological modelling. Their continued refinement will expand opportunities for both applied water-resource management and fundamental research.

4.4 | Potential Improvements for New Zealand's Precipitation Isotope Network

Future, deliberate additions to New Zealand's precipitation isotope network would substantially strengthen the use of hydrogen and oxygen isotopes in environmental applications, particularly by improving our understanding of (1) submonthly

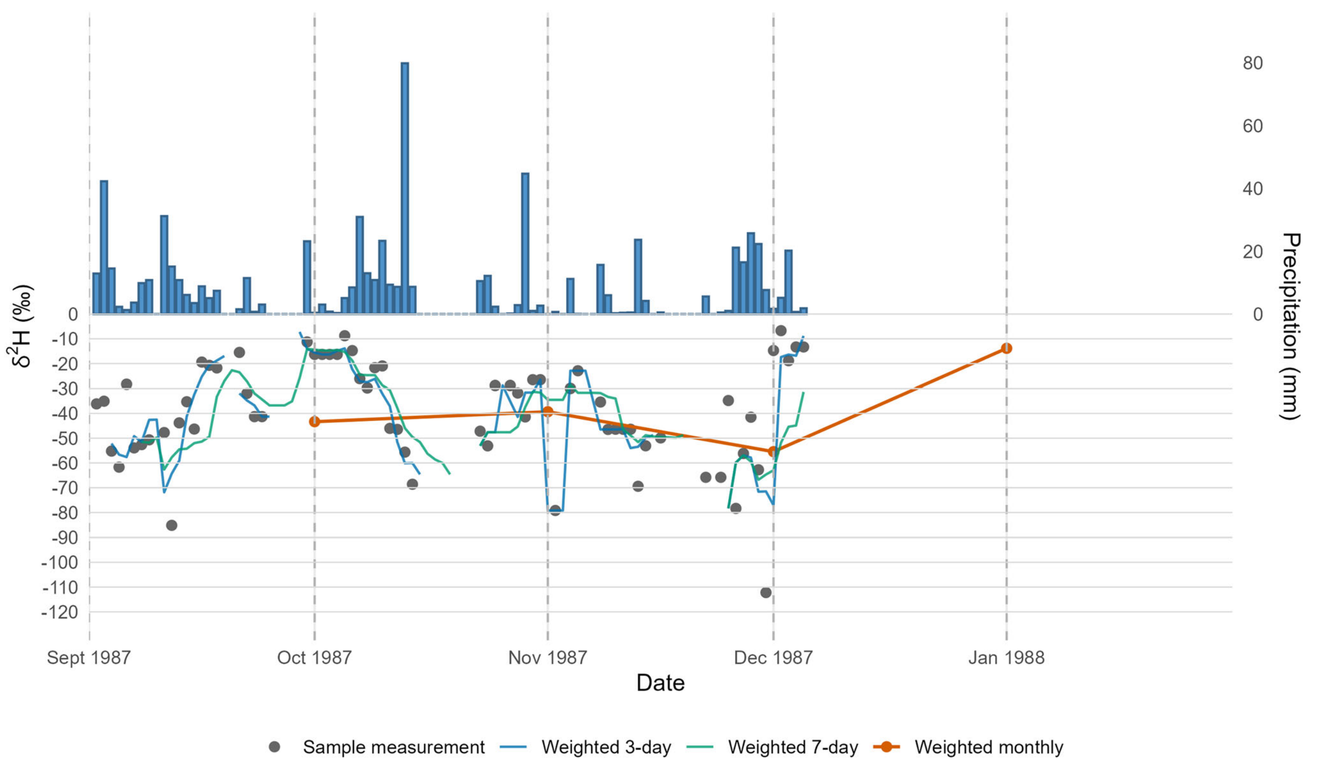

FIGURE 4 | Damping effect of monthly weighting, shown using daily precipitation δ^2H and amounts recorded at the Maimai experimental catchment near the southwest coast of New Zealand. Rolling 3-day and 7-day weighted averages are shown, in comparison to monthly values weighted for calendar months.

variability in precipitation isotope values and (2) spatial and temporal patterns in high-alpine environments.

Regarding submonthly variability, reliance on monthly composite samples dampens event-scale fluctuations and thus constrains our ability to model temporal variation. The effect of monthly sampling is evident from Figure 4, where strong departures above and below monthly means (likely driven by moisture-source variability (McDonnell 1988; Stewart et al. 2020)) would be largely smoothed at a monthly timestep but far better resolved with weekly or higher-frequency sampling. This is particularly pertinent to New Zealand, because submonthly variation dominates in marine climates with weak seasonal temperature cycles (Dudley et al. 2024). Improved tracer input data at submonthly timestep would thus benefit hydrological applications such as hydrograph separation, tracer-based transit-time estimation, and isotope-enabled hydrological modelling (Benettin et al. 2022). Increased sampling frequency would also support more direct use of daily isoGSM outputs, which at present require downscaling and correction using sparse station data (Watson et al. 2024).

High-elevation regions remain the largest spatial gap in the national network. Fit statistics from this study show that processes controlling isotopic variability at alpine sites are still imperfectly captured by the PINZ model. This likely reflects both the relatively limited number of high-elevation observations available for model training and the increased importance of interacting controls at these sites, such as elevation, aspect, and moisture-source variability, which may differ from those dominating lower elevations (Moran et al. 2007; Tian et al. 2007). Targeted improvements in observational coverage across alpine gradients would therefore help constrain these relationships and improve model performance. Moreover, elevation gradients in New Zealand occur over spatial scales far smaller than those resolved by global isotope-enabled models, meaning that topographic effects are effectively averaged out; alpine observations are therefore essential for maintaining model realism even at large-catchment scales (Watson et al. 2024).

Future precipitation isotope modelling would also benefit from the careful collection of precipitation amount data alongside isotope samples. This would reduce reliance on gridded rainfall products for rainfall weighting and the associated uncertainty demonstrated in Figure S1. Error associated with temporal alignment for monthly modelling (Figure S2) could be reduced or eliminated by inclusion of a precipitation isotope collection day on the first day of each month. Participation in interlaboratory comparison exercises such as the IAEA Water Isotope Interlaboratory Comparison (WICO) programme (Ahmad et al. 2012; Wassenaar et al. 2018, 2021), together with the use of standardised field protocols (Gröning et al. 2012), would help ensure data quality consistency and facilitate the integration of future datasets into national compilations and precipitation isotope models.

5 | Summary and Conclusions

This study delivers two key national-scale resources for New Zealand: an expanded precipitation isotope database compiled from historical and contemporary studies, and an updated spatially continuous time series of monthly precipitation isoscapes generated using the PINZ modelling framework. Together, these products improve national-scale estimates of precipitation δ^2H and δ^{18}O for New Zealand and provide a foundation for a wide range of hydrological and climatological applications. Although inclusion of data from multiple sources, alignment of data to calendar months, and weighting using gridded rainfall data introduce uncertainty, additional training data (especially that from high elevations) added valuable environmental heterogeneity to the training dataset of our national-scale machine learning model (PINZ) and expanded its AOA without degrading its performance in well-sampled lowland areas. The primary limitation on further improvement to modelling capability is not technological, but the availability of submonthly scale precipitation isotope measurements, especially in regions characterised by strong orographic effects, and at high elevations. Continued community contributions of alpine and higher-frequency data will further improve model realism and extend the utility of precipitation isotope products for hydrological and climatological applications in New Zealand.

Author Contributions

Bruce D. Dudley: conceptualization (lead), data curation (lead), formal analysis (co-lead), investigation (lead), methodology (co-lead), visualization (co-lead), writing – original draft (lead), writing – review and editing (lead), project administration (lead). **Alice F. Hill:** writing – original draft (supporting), methodology (co-lead), writing – review and editing (supporting). **Andy McKenzie:** formal analysis (co-lead), methodology (co-lead), visualization (co-lead), writing – review and editing (supporting). **Peter W. Holder:** data curation (supporting), investigation (supporting), writing – review and editing (supporting). **W. Troy Baisden, Stephen Collins, James Dare, Michael Darling, Russell D. Frew, Scott L. Graham, Robert van Hale, Travis W. Horton, Elizabeth D. Keller, Shelley MacDonell, Jack McMecking, Jeffrey J. McDonnell, Magali F. Nehemy, Ewen Rodway, Zane Shadbolt, Cerra Simmons, Andrew Swales, Michael K. Stewart, Vanessa Trompetter,** and **Thomas Wöhling:** resources; data curation (supporting); investigation; writing – review and editing. All authors contributed data and associated methodological information, reviewed the analyses and manuscript, and approved the final version for publication.

Acknowledgements

The authors wish to acknowledge the many individuals across New Zealand who have collected precipitation samples for isotope analysis over the past several decades. This includes the pioneering researchers who first established precipitation isotope monitoring in New Zealand, along with the technicians, students, citizen scientists, and research teams who have sustained and expanded these efforts. The national database described in this study is built on the cumulative work of this community.

Open access publishing facilitated by Earth Sciences New Zealand, as part of the Wiley - Earth Sciences New Zealand agreement via the Council of Australasian University Librarians.

Funding

This work was supported by the Ministry for Business Innovation and Employment (C04X1905, C09 X1610, C05X1711, C05X1803, Strategic Science Investment Funding); Better Border Biosecurity New Zealand (Cross Departmental Research Project (CDRP) database funding);

Biosecurity New Zealand (Cross Departmental Research Project (CDRP) database funding); Department of Conservation, New Zealand (Cross Departmental Research Project (CDRP) database funding); Environment Southland; Horizons Regional Council; Bay of Plenty Regional Council; Gisborne Regional Council; Brian Mason Scientific and Technical Trust; Rutherford Foundation; International Atomic Energy Association (CRP F31008).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Analysis code is available from the project GitHub repository (Dudley et al. 2026). Raw data are available via the repository's request process as described in the accompanying documentation at https://github.com/BruceDudley/Precipitation_isotopes_NZ. Model output (monthly δ^2H and δ^{18}O isoscapes, 1997–2025) are archived at Zenodo and available at https://doi.org/10.5281/zenodo.20501203.

References

Ahmad, M., P. Aggarwal, M. van Duren, L. Poltenstein. 2012. "Final Report on Fourth Interlaboratory Comparison Exercise for δ^2H and δ^{18}O Analysis of Water Samples (WICO2011)." INIS-XA--12S0138. Vienna: International Atomic Energy Agency, Isotope Hydrology Laboratory, Department of Nuclear Sciences and Applications.

Baisden, W. T., E. D. Keller, R. Van Hale, R. D. Frew, and L. I. Wassenaar. 2016. "Precipitation Isoscapes for New Zealand: Enhanced Temporal Detail Using Precipitation-Weighted Daily Climatology." *Isotopes in Environmental and Health Studies* 52: 343–352.

Benettin, P., N. B. Rodriguez, M. Sprenger, et al. 2022. "Transit Time Estimation in Catchments: Recent Developments and Future Directions." *Water Resources Research* 58: e2022WR033096.

Bong, H., A. Cauquoin, A. Okazaki, et al. 2024. "Process-Based Intercomparison of Water Isotope-Enabled Models and Reanalysis Nudging Effects." *Journal of Geophysical Research: Atmospheres* 129: e2023JD038719.

Bong, H., A. N. LeGrande, S. Dee, et al. 2026. "Water Isotope Model Intercomparison Project (WisoMIP): Present-Day Climate." *Journal of Geophysical Research: Atmospheres* 131: e2025JD044985.

Bowen, G. J. 2010. "Isoscapes: Spatial Pattern in Isotopic Biogeochemistry." *Annual Review of Earth and Planetary Sciences* 38: 161–187.

Bowen, G. J. 2025. "Waterisotopes Database." Accessed February 3, 2025. https://wateriso.utah.edu/waterisotopes/.

Bowen, G. J., Z. Cai, R. P. Fiorella, and A. L. Putman. 2019. "Isotopes in the Water Cycle: Regional- to Global-Scale Patterns and Applications." *Annual Review of Earth and Planetary Sciences* 47: 453–479.

Breil, M., E. Christner, A. Cauquoin, M. Werner, M. Karremann, and G. Schädler. 2021. "Applying an Isotope-Enabled Regional Climate Model over the Greenland Ice Sheet: Effect of Spatial Resolution on Model Bias." *Climate of the Past* 17: 1685–1699.

Caloiero, T. 2014. "Analysis of Daily Rainfall Concentration in New Zealand." *Natural Hazards* 72: 389–404.

Cole, A., and D. F. Boutt. 2021. "Spatially-Resolved Integrated Precipitation-Surface-Groundwater Water Isotope Mapping from Crowd Sourcing: Toward Understanding Water Cycling across a Post-Glacial Landscape." *Frontiers in Water* 3: 645634.

Dudley, B., A. Hill, and A. McKenzie. 2024. "A Seasonal Precipitation Isoscape for New Zealand." *Journal of Hydrology: Regional Studies* 52: 101711.

Dudley, B. D., A. McKenzie, and A. F. Hill. 2026. Precipitation Isotopes NZ: Supplementary Data and Analysis Code for the Manuscript *'Community Participation as a Pathway to Strengthen National Precipitation Isotope Products in New Zealand'* [Computer Software]. GitHub repository. https://github.com/BruceDudley/Precipitation_isotopes_NZ.

Dudley, B. D., A. McKenzie, J. S. Diamond, and A. F. Hill. 2025. "Complementary Precipitation Isotope Models Reveal Young and New Water Fractions in New Zealand Rivers." *Hydrological Processes* 39: e70265.

Dudley, B. D., H. Marttila, S. L. Graham, R. Evison, and M. Srinivasan. 2018. "Water Sources for Woody Shrubs on Hillslopes: An Investigation Using Isotopic and Sapflow Methods." *Ecohydrology* 11: e1926.

Erdélyi, D., I. G. Hatvani, H. Jeon, M. Jones, J. Tyler, and Z. Kern. 2023. "Predicting Spatial Distribution of Stable Isotopes in Precipitation by Classical Geostatistical- and Machine Learning Methods." *Journal of Hydrology* 617: 129129.

Eyring, V., W. D. Collins, P. Gentine, et al. 2024. "Pushing the Frontiers in Climate Modelling and Analysis with Machine Learning." *Nature Climate Change* 14: 916–928.

Fraisl, D., G. Hager, B. Bedessem, et al. 2022. "Citizen Science in Environmental and Ecological Sciences." *Nature Reviews Methods Primers* 2: 64.

Good, S. P., D. V. Mallia, J. C. Lin, and G. J. Bowen. 2014. "Stable Isotope Analysis of Precipitation Samples Obtained via Crowdsourcing Reveals the Spatiotemporal Evolution of Superstorm Sandy." *PloS One* 9: e91117.

Gröning, M., H. Lutz, Z. Roller-Lutz, M. Kralik, L. Gourcy, and L. Pöltenstein. 2012. "A Simple Rain Collector Preventing Water Re-Evaporation Dedicated for δ^{18}O and δ^2H Analysis of Cumulative Precipitation Samples." *Journal of Hydrology* 448: 195–200.

Hill, A. F., A. McKenzie, B. D. Dudley, and D. B. Nelson. 2025. "Precipitation Isotopes New Zealand (PINZ): Improvements in Precipitation Isoscapes with Machine Learning." *Environmental Chemistry* 22: EN24098.

Hirave, P., D. B. Nelson, M. Glendell, and C. Alewell. 2023. "Land-use-Based Freshwater Sediment Source Fingerprinting Using Hydrogen Isotope Compositions of Long-Chain Fatty Acids." *Science of the Total Environment* 875: 162638.

Klein, L., B. Dudley, and J. Klaus. 2026. "Soil and Plant Water Ages in Two *Pinus radiata* Forests with Contrasting Precipitation Regimes." *Ecohydrology* 19: e70225. https://doi.org/10.1002/eco.70225.

Lei, H., H. Li, and H. Zhao. 2025. "Refining Daily Precipitation Estimates Using Machine Learning and Multi-Source Data in Alpine Regions with Unevenly Distributed Gauges." *Journal of Hydrology: Regional Studies* 58: 102272.

Lima, C. H., H.-H. Kwon, and Y.-T. Kim. 2021. "A Bayesian Kriging Model Applied for Spatial Downscaling of Daily Rainfall from GCMs." *Journal of Hydrology* 597: 126095.

Lin, W., M. M. Barbour, and X. Song. 2022. "Do Changes in Tree-ring δ^{18}O Indicate Changes in Stomatal Conductance?" *New Phytologist* 236: 803–808.

Matiatos, I., C. Moeck, Y. Vystavna, et al. 2023. "Nitrate Isotopes in Catchment Hydrology: Insights, Ideas and Implications for Models." *Journal of Hydrology* 626: 130326.

McDonnell, J. J. 1988. "The Age, Origin and Pathway of Subsurface Stormflow in a Steep Humid Headwater Catchment." PhD thesis, University of Canterbury. http://dx.doi.org/10.26021/6481.

McKenzie, A., A. F. Hill, and B. D. Dudley. 2025. "PINZ: Precipitation Isotopes New Zealand; a Modelled Precipitation Isotope Surface (isoscape) Using the XGBoost Machine Learning Algorithm." [Computer software]. GitHub repository. https://github.com/niwa/PINZ.

McMecking, J. 2025. "Chasing Storms: Tracking Event Flow in a Mountain Catchment." Master of Science thesis, University of Canterbury. https://doi.org/10.26021/15985.

Medina, D. C., M. G. Delgado, J. S. Ramos, T. P. Amores, L. R. Rodríguez, and S. A. Dominguez. 2024. "Empowering Urban Climate Resilience and Adaptation: Crowdsourcing Weather Citizen Stations-Enhanced Temperature Prediction." *Sustainable Cities and Society* 101: 105208.

Meyer, H., and E. Pebesma. 2021. "Predicting into Unknown Space? Estimating the Area of Applicability of Spatial Prediction Models." *Methods in Ecology and Evolution* 12: 1620–1633.

Meyer, H., M. Ludwig, C. Milà, J. Linnenbrink, and F. Schumacher. 2024. "The CAST Package for Training and Assessment of Spatial Prediction Models in R." arXiv preprint arXiv:2404.06978.

Michelsen, N., R. van Geldern, Y. Roßmann, et al. 2018. "Comparison of Precipitation Collectors Used in Isotope Hydrology." *Chemical Geology* 488: 171–179.

Moran, T. A., S. J. Marshall, E. C. Evans, and K. E. Sinclair. 2007. "Altitudinal Gradients of Stable Isotopes in Lee-Slope Precipitation in the Canadian Rocky Mountains." *Arctic, Antarctic, and Alpine Research* 39: 455–467.

Muller, C. L., L. Chapman, S. Johnston, et al. 2015. "Crowdsourcing for Climate and Atmospheric Sciences: Current Status and Future Potential." *International Journal of Climatology* 35: 3185–3203.

Pebesma, E. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10: 439–446.

Prince, H. D., N. J. Cullen, P. B. Gibson, J. Conway, and D. G. Kingston. 2021. "A Climatology of Atmospheric Rivers in New Zealand." *Journal of Climate* 34: 4383–4402.

Purdie, H., N. Bertler, A. Mackintosh, J. Baker, and R. Rhodes. 2010. "Isotopic and Elemental Changes in Winter Snow Accumulation on Glaciers in the Southern Alps of New Zealand." *Journal of Climate* 23: 4737–4749.

R Core Team. 2025. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Reichstein, M., G. Camps-Valls, B. Stevens, et al. 2019. "Deep Learning and Process Understanding for Data-Driven Earth System Science." *Nature* 566: 195–204.

Rozanski, K., L. Araguás-Araguás, and R. Gonfiantini. 1993. "Isotopic Patterns in Modern Global Precipitation." *Climate Change in Continental Isotopic Records* 78: 1–36.

Salinger, M. 1980. "New Zealand Climate: I. Precipitation Patterns." *Monthly Weather Review* 108: 1892–1904.

Schroeter, S. A., A. M. Orme, K. Lehmann, et al. 2025. "Hydroclimatic Extremes Threaten Groundwater Quality and Stability." *Nature Communications* 16: 720.

Siegwolf, R. T. W., M. M. Lehmann, G. R. Goldsmith, et al. 2023. "Updating the Dual C and O Isotope—Gas-Exchange Model: A Concept to Understand Plant Responses to the Environment and Its Implications for Tree Rings." *Plant, Cell & Environment* 46: 2606–2627.

Stewart, M. K., U. Morgenstern, and R. W. van der Raaij. 2020. "An Effect of Australia on New Zealand Precipitation." NZ Hydrological Society, NZ Rivers Group & NZ Freshwater Sciences Society Joint Conference.

Tait, A., J. Sturman, and M. Clark. 2012. "An Assessment of the Accuracy of Interpolated Daily Rainfall for New Zealand." *Journal of Hydrology (New Zealand)* 25–44.

Tait, A., R. Henderson, R. Turner, and X. Zheng. 2006. "Thin Plate Smoothing Spline Interpolation of Daily Rainfall for New Zealand Using a Climatological Rainfall Surface." *International Journal of Climatology: A Journal of the Royal Meteorological Society* 26: 2097–2115.

Taylor, C. B. 1990. "Stable Isotope Compositions of Monthly Precipitation Samples Collected in New Zealand and Rarotonga." Physical Sciences Report No. 3. Lower Hutt, New Zealand: Nuclear Sciences Group, DSIR Physical Sciences, Department of Scientific and Industrial Research.

Terzer, S., L. Wassenaar, L. Araguás-Araguás, and P. Aggarwal. 2013. "Global Isoscapes for $\delta^{18}O$ and $\delta^{2}H$ in Precipitation: Improved Prediction Using Regionalized Climatic Regression Models." *Hydrology and Earth System Sciences* 17: 4713–4728.

Tian, L., T. Yao, K. MacClune, et al. 2007. "Stable Isotopic Variations in West China: A Consideration of Moisture Sources." *Journal of Geophysical Research: Atmospheres* 112: D10.

Vystavna, Y., L. Chavanne, A. Harjung, et al. 2024. "Predicting River Flow Dynamics Using Stable Isotopes for Better Adaptation to Climate and Land-use Changes." *Nature Water* 2: 741.

Wassenaar, L., S. Terzer-Wassmuth, and C. Douence. 2021. "Progress and Challenges in Dual-and Triple-Isotope ($\delta^{18}O$, $\delta^{2}H$, $\delta^{17}O$) Analyses of Environmental Waters: An International Assessment of Laboratory Performance." *Rapid Communications in Mass Spectrometry* 35: e9193.

Wassenaar, L. I., S. Terzer-Wassmuth, C. Douence, L. Araguas-Araguas, P. Aggarwal, and T. B. Coplen. 2018. "Seeking Excellence: An Evaluation of 235 International Laboratories Conducting Water Isotope Analyses by Isotope-Ratio and Laser-Absorption Spectrometry." *Rapid Communications in Mass Spectrometry* 32: 393–406.

Watson, A., C. Birkel, S. Arciniega-Esparza, et al. 2024. "Evaluating Input Data Sources for Isotope-Enabled Rainfall-Runoff Models." *Hydrological Processes* 38: e15276.

Wootten, A. M., K. W. Dixon, D. J. Adams-Smith, and R. A. McPherson. 2021. "Statistically Downscaled Precipitation Sensitivity to Gridded Observation Data and Downscaling Technique." *International Journal of Climatology* 41: 980–1001.

Zhong, S., K. Zhang, M. Bagheri, et al. 2021. "Machine Learning: New Ideas and Tools in Environmental Science and Engineering." *Environmental Science & Technology* 55: 12741–12754.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** Effect of using daily rainfall data from the interpolated Virtual Climate Station Network (VCSN) to amount-weight event-scale isotope values when generating monthly isotope values. Results are shown for five sites where isotopes were collected at event scale and rainfall was measured on site. Root mean-squared errors arising from this method are presented in Table 1. **Figure S2:** Effect of shifting monthly collection dates on estimates of amount-weighted monthly isotope values for the period ending on the first day of each month. Results are shown for five sites with event-scale isotope collections and on-site rainfall measurements. Errors associated with shifting the collection date to the 5th, 10th, and 15th of the month were calculated from these time series. Root mean-squared errors for each date shift are presented in Table 1. **Figure S3:** Representative PINZ precipitation isoscapes for a warm month (February) and cold month (August) for 2024. Upper panels show results for $\delta^{18}O$, lower panels show results for $\delta^{2}H$. **Figure S4:** Median Absolute Deviation (MAD) and Root Mean Squared Error (RMSE) for PINZ trained using the database described here ('All Sites') and the database described in Hill et al. (2025) ('Old Sites'). **Table S1:** Methods summary for laboratories contributing to the database.