

IsoRole improves response-blind precipitation-isotope reconstruction at new stations with a physical-reference residual ensemble

Shaofeng Guo

Water Conservancy and Civil Engineering College, Inner Mongolia Agricultural University

Debin Jia

jdb@imau.edu.cn

Water Conservancy and Civil Engineering College, Inner Mongolia Agricultural University

Yusheng Hao

Water Conservancy and Civil Engineering College, Inner Mongolia Agricultural University

Darifu Ba

Department of Chemical and Environmental Engineering, Hetao College

Shuang Li

Research Institute of System Engineering, PLA Academy of Military Science

Donghui Zhang

Institute of Remote Sensing Satellite, China Academy of Space Technology

Article

Keywords: IsoRole, precipitation isotopes, complete-station reconstruction, physical reference, residual ensemble, hydroclimate

Posted Date: October 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10859632/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

IsoRole improves response-blind precipitation-isotope reconstruction at new stations with a physical-reference residual ensemble

Shaofeng Guo^{1,2}, Debin Jia^{1*}, Yusheng Hao¹, Darifu Ba³, Shuang Li⁴, Donghui Zhang⁵

¹Water Conservancy and Civil Engineering College, Inner Mongolia Agricultural University, Hohhot 010019, China; gsfnd@163.com (S.G.); jdb@imau.edu.cn (D.J.); 2021202060008@emails.imau.edu.cn (Y.H.)

²Department of Water Conservancy and Civil Engineering, Hetao College, Bayannur 015000, China;

³Department of Chemical and Environmental Engineering, Hetao College, Bayannur 015000, China; bdrf2000@126.com (D.B.)

⁴Research Institute of System Engineering, PLA Academy of Military Science, Beijing 100091, China; 834588643@qq.com (S.L.)

⁵Institute of Remote Sensing Satellite, China Academy of Space Technology, Beijing 100094, China; zhangdonghui@alu.cdut.edu.cn (D.Z.)

* Correspondence:

Debin Jia

jdb@imau.edu.cn; Tel.: +86-135-147-12055

Abstract: Precipitation-isotope inputs support hydrological and ecohydrological inference, yet stations differ in sampling interval, temporal repetition, spatial support, and access to physical isotope fields. We developed IsoRole (Applicability-Gated Isotope Role Fusion), a response-blind method that learns record-scale displacement around an independent monthly isotope reference in $\delta^{18}\text{O}$ –d-excess geometry and returns one coupled isotope pair. Complete-station development partitioned 8,267 dated records from 67 stations across three public archives into five folds. The residual-geometry foundation reduced source-balanced joint error by 13.13% relative to meteorology-only learning and by 7.00% relative to physical-only learning. Independent confirmation evaluated 175 response-blind paired-isotope queries from Kathmandu and Lete. A fixed physical-reference residual ensemble (A-RGF) reduced joint loss from 1.775 for direct CatBoost to 1.403; complete IsoRole attained the lowest point loss, 1.360, and reduced measured-pair error by 23.37% relative to direct CatBoost and 50.39% relative to the raw physical field. Observation-centered analyses identified archive-specific annual, spatial, physical-background, and antecedent structures, while the out-of-fold structural analysis tracked selected phase, isotope-line, and 30-day precipitation relations. IsoRole thus delivered transferable response-blind reconstruction within the tested active support: A-RGF produced most of the Nepal improvement, and coordinate-role composition yielded the lowest point loss among the evaluated transfer configurations.

Keywords: IsoRole; precipitation isotopes; complete-station reconstruction; physical reference; residual ensemble; hydroclimate

35 1 Introduction

36 Stable oxygen and hydrogen isotope ratios encode the history of evaporation, condensation, transport,
37 and rainout that precedes a precipitation event. Their paired geometry supplies both the dominant
38 enrichment–depletion coordinate and a secondary departure from the slope-eight meteoric relation,
39 allowing precipitation to act as a dynamic input tracer for recharge, streamflow, soil water, and plant-
40 water studies^{1–4}. The scientific value of a precipitation end member consequently depends on matching
41 the location, season, and antecedent state represented by the receiving water pool. Recharge and
42 ecohydrological partitioning studies have shown that temporally aggregated annual precipitation can
43 obscure the seasonal inputs actually transmitted into groundwater and vegetation^{5–8}.

44 Precipitation-isotope observations remain heterogeneous by design. Long monthly station records
45 resolve repeated seasonal and interannual change; daily tropical observations resolve event sequences;
46 multi-station event compilations broaden geographic coverage but often contain fewer repeated
47 observations per site. The three public archives used here exemplify that trade-off and are
48 supplemented by a two-station Himalayan record for response-blind confirmation^{9–13}. Reconstruction
49 at an isotope-response-unobserved station is therefore a transfer problem in which every destination-
50 station isotope response remains outside model fitting.

51 Regionalized isoscapes and dynamic isotope products provide spatially coherent background
52 information, while machine-learning models can resolve nonlinear relationships with sampling context
53^{14–18}. Existing precipitation-isotope learners estimate the response directly from climatic, geographic,
54 and isotope-field predictors; direct concatenation can exploit all of these inputs but does not explicitly
55 constrain the different roles of a monthly physical state and a sampling-record displacement. More
56 broadly, hybrid Earth-system learning gains scientific leverage when known physical structure is
57 represented in the learning target rather than left for a flexible learner to rediscover^{19,20}. This motivates
58 a sharper construction for precipitation isotopes: place each dated record relative to a target-
59 independent monthly isotope reference, learn the displacement from that reference, and preserve the
60 paired-isotope identity during reconstruction.

61 The invertible $\delta^{18}\text{O}$ –d-excess representation expresses paired-isotope geometry as an enrichment–
62 depletion coordinate and a slope-eight departure while preserving exact recovery of both isotope ratios
63^{21,22}. This geometry allows A-RGF oxygen and direct-pair d-excess to form a coupled $\delta^{18}\text{O}$ – $\delta^2\text{H}$
64 estimate through exact inversion.

65 IsoRole, or Applicability-Gated Isotope Role Fusion, combines four elements: a complete-station
66 residual-geometry foundation (RGF), a one-third XGBoost/two-thirds CatBoost residual ensemble (A-
67 RGF), target-independent geographic applicability, and complementary A-RGF-oxygen/direct-pair-d-
68 excess coordinates. At an unseen station, these elements return one coupled isotope pair without using
69 destination-station isotope responses.

70 The central question is whether IsoRole can improve paired-isotope reconstruction across
71 heterogeneous observation regimes at stations absent from fitting. We evaluate the residual-geometry
72 foundation against meteorology-only, physical-only, raw-field, climatological, and same-information
73 direct learners across 67 held stations; quantify IsoRole at Kathmandu and Lete; and test whether out-
74 of-fold predictions retain annual-phase, isotope-line, and hydroclimatic relations. These analyses
75 separate the contribution of A-RGF from coordinate composition and connect predictive performance
76 to precipitation end-member structure.

2 Materials and Methods

2.1 Data, reconstruction task, and validation design

2.1.1 Measured archives and validation roles

The development analysis combined 2,849 monthly observations from 22 stations drawn from the Australian GNIP rainfall-isotope archive released with the 2015–2024 update, 4,729 tropical daily observations from 15 stations, and 689 event or collection-interval observations from 30 stations in the Mendeley compilation. The combined dataset contained 8,267 dated records from 67 stations, each with finite measured $\delta^{18}\text{O}$ and $\delta^2\text{H}$. The independently acquired uncorrected monthly physical-isotope product, referred to as ICM in comparator labels, covered 8,197 records; 70 records dated after its final month used the meteorology-only estimate. Figure 1 maps the resulting station support and station-mean isotope composition^{9,10,12,13,23}.

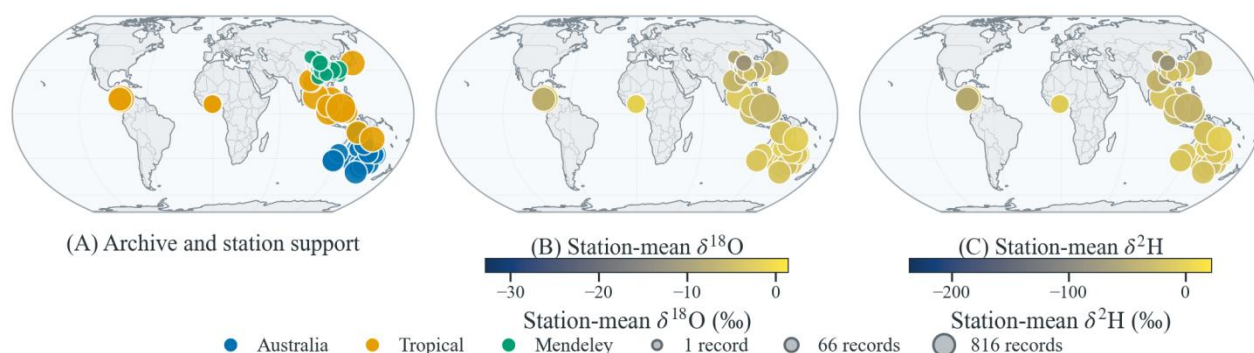


Figure 1. Global support and station-mean isotope composition across three measured precipitation archives. Symbol area denotes support; panels identify archive membership and station-mean oxygen and hydrogen isotope ratios for 67 stations and 8,267 records.

Data-source coverage is summarized in Supplementary Table S1.

Four non-overlapping station sets supported complete-station learning, geographic-applicability estimation, coordinate assignment, and independent evaluation. Complete-station analysis of 8,267 records selected mass-one residual geometry and two component learners. PAPIN and Indianapolis records (doi:10.5281/zenodo.4767053; doi:10.1594/PANGAEA.895300) informed the target-independent geographic-applicability rule, whose threshold τ was derived from development-station geometry. Yadong and Pali records (doi:10.5281/zenodo.8059700) defined the coordinate assignment. Kathmandu and Lete were reserved for independent response-blind evaluation.

Independent response-blind evaluation used the Kali Gandaki High Mountain Observatory archive. Records were required to contain a valid calendar date; one undated Lete record was excluded. At each station, the earliest 16 date-ordered records supplied calendar information for the support-gap descriptors, with same-day duplicates retained and isotope values withheld. The evaluation set contained 64 Kathmandu and 111 Lete queries (175 total), whose isotope responses were unavailable during model fitting, threshold construction, estimator assignment, and prediction. Neither station occurred in the 67-station development set¹¹. This protocol represents new-station reconstruction from a short date-only support history.

2.1.2 Response-blind task and isotope geometry

For record i , let $x_i \in \mathbb{R}^{34}$ denote the fixed-order target-independent predictor vector, comprising 24 meteorological/contextual fields followed by 10 physical-isotope fields. The admissible reconstruction map $\mathcal{F}$ returned the paired prediction $\hat{y}_i$ from this vector:

$$\hat{y}_i = \mathcal{F}(x_i) \quad (1)$$

Equation (1) excludes every isotope response from the destination station and defines the admissible information boundary for all estimators.

The measured response vector y_i fixed the order of the two isotope ratios; superscript T denotes transpose:

$$y_i = (\delta^{18}O_i, \delta^2H_i)^T \quad (2)$$

Both coordinates are expressed in per mille relative to VSMOW. The derived coordinate d_i was defined as

$$d_i = \delta^2H_i - 8\delta^{18}O_i \quad (3)$$

and the geometry transform mapped a measured or predicted isotope pair into oxygen and d-excess coordinates:

$$\mathcal{T}(y_i) = (\delta^{18}O_i, d_i)^T \quad (4)$$

For arbitrary oxygen and d-excess coordinates u_{18} and u_d , the inverse transform preserved isotope identity during every geometry-based reconstruction:

$$\mathcal{T}^{-1}(u_{18}, u_d) = (u_{18}, 8u_{18} + u_d)^T \quad (5)$$

Equations (2)–(5) fix the target order, the derived d-excess coordinate, and the unique inverse used for every geometry-based paired prediction.

The estimand is retrospective sampling-record reconstruction: measured isotope targets were unavailable, whereas record timing, collection-interval forcing, and the monthly physical product were allowed.

Figure 2 summarizes the complete-station design, the IsoRole reconstruction logic, and the parallel prediction and measured-observation analyses.

2.1.3 Predictor variables and complete-station folds

The ordered predictor matrix contained 24 meteorological/contextual fields followed by 10 physical fields. Meteorological/contextual fields were annual sine and cosine; $\ln(1 + \text{collection duration in days})$; sine and cosine transforms of latitude and longitude; collection-interval, prior 7-day, and prior 30-day precipitation; collection-interval temperature, relative humidity, shortwave radiation, and longwave radiation; seven corresponding observed-value masks; POWER forcing coverage; log-transformed median support gap; and support fraction. Meteorological forcing was obtained as NASA POWER Daily UTC data through the official point service or the AWS Analysis Ready Data source²⁴. For each Nepal query, the median support-gap field was the natural logarithm of one plus the median positive day gap from the query date to the 16 earlier calendar-only support-record date entries at the same station, with same-day duplicates retained; support fraction was 1.0. The physical block

comprised uncorrected monthly $\delta^{18}\text{O}$, $\delta^2\text{H}$, and publisher-supplied source d-excess; their gridded station-cell calendar-month climatology medians; their anomalies; and physical-coverage availability. The publisher-supplied d-excess field was used as an independent product variable. The physical-coverage indicator equaled one when the uncorrected physical isotope pair was finite and zero otherwise.

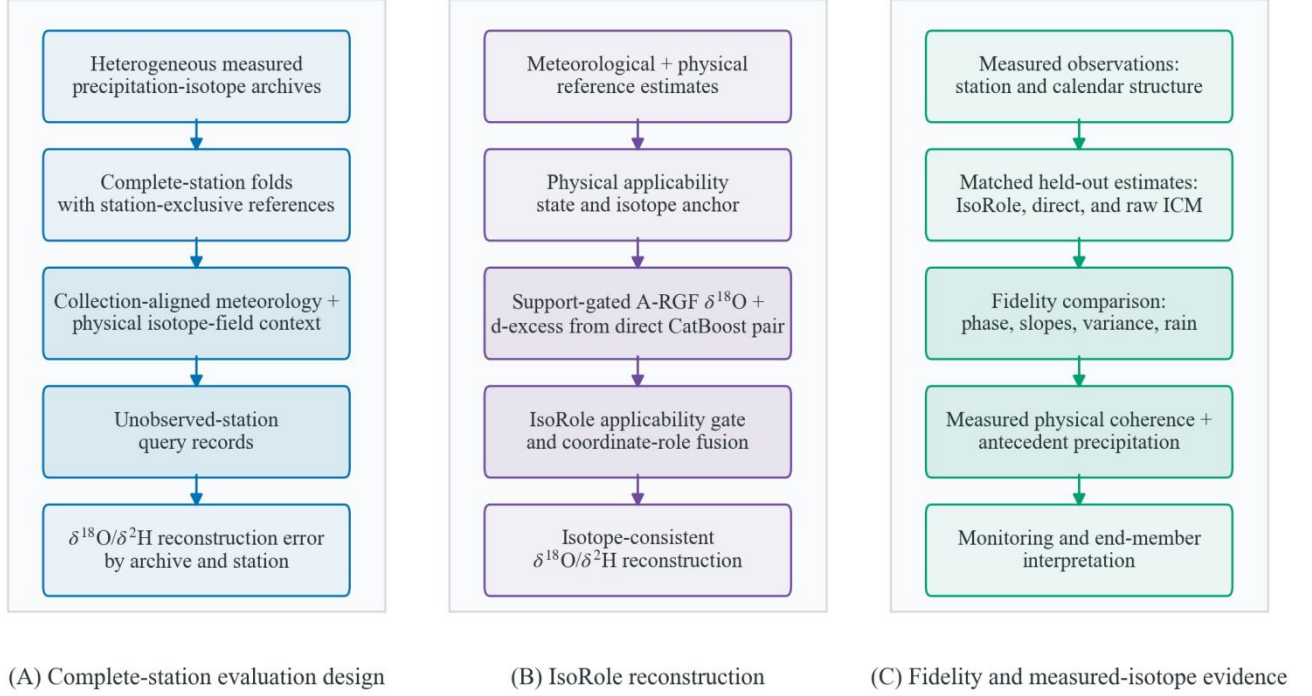


Figure 2. IsoRole combines complete-station residual learning with geographic applicability and paired-isotope reconstruction. Panels show development folds, A-RGF oxygen, direct-pair d-excess, out-of-fold structural fidelity, and observation-centered isotope relations.

Five deterministic complete-station folds were formed with station as the group and archive as the stratum. Let s_i be the station of record i , and let $K(s_i) \in \{1, \dots, 5\}$ be its fold. The validation records in fold k were

$$\mathcal{V}_k = \{i: K(s_i) = k\} \quad (6)$$

With $\mathcal{D}_{\text{dev}}$ denoting all 8,267 development records, the corresponding training set was $\mathcal{D}_k = \mathcal{D}_{\text{dev}} \setminus \mathcal{V}_k$. All three archives occurred in every fold. The seed-3407 split produced five station-held-out folds, each containing stations from all three archives. Grouped validation prevents record-level leakage from repeated station identity and aligns the evaluation with application at a new station^{25,26}.

Training weights were defined once on the full development universe so that each archive had equal total mass, each station within an archive had equal mass, and each record within a station had equal mass. Let $A = 3$ be the archive count, $\sigma(i)$ the archive of record i , J_r the number of stations in archive r , N_s the record count at station s , and $N_{\text{dev}} = 8267$ the development-record count. The mean-one weight was

$$w_i = \frac{(AJ_{\sigma(i)}N_{s_i})^{-1}}{N_{\text{dev}}^{-1} \sum_{\ell=1}^{N_{\text{dev}}} (AJ_{\sigma(\ell)}N_{s_\ell})^{-1}} \quad (7)$$

The summation index ℓ traverses the full development universe. Equation (7) was evaluated once. Each fold inherited its indexed weights without renormalization, so fold-training archive totals varied with the stations retained while full-universe balance remained fixed. Equations (6)–(7) therefore enforce complete-station exclusion and one weighting scheme across candidates. The same weights entered all control, direct, residual, and structural fits. Development reporting used source-balanced station aggregation; Nepal reporting used equal-station aggregation. The out-of-fold structural analysis compared reconstructed and measured isotope relations independently of the observation-centered models.

2.2 Residual-geometry foundation

2.2.1 Tree ensembles and coordinate objective

The reference XGBoost 3.2.0 coordinate learner used 400 boosted trees, learning rate 0.03, maximum depth 3, minimum child weight 20, row and column subsampling 0.85, L2 penalty 10, histogram construction with 256 maximum bins, seed 3407, and one CPU thread²⁷. The CatBoost 1.2.10 residual parent and direct-pair learner used MAE loss, 400 iterations, depth 3, learning rate 0.03, L2 leaf penalty 10, Bernoulli row subsampling 0.85, column sampling 0.85, zero random strength, seed 3407, and one CPU thread²⁸. LightGBM 4.6.0 was confined to the matched learner sensitivity analysis²⁹. All learner identities and hyperparameters were held fixed before independent confirmation.

The XGBoost structural-grid coordinate learners used Pseudo-Huber loss. For scalar coordinate residual e and slope parameter $\kappa = 1$ in the fitted coordinate's per-mille units, the loss was

$$\rho_\kappa(e) = \kappa^2 \left(\sqrt{1 + (e/\kappa)^2} - 1 \right) \quad (8)$$

This differentiable objective is the smooth Charbonnier or pseudo-Huber form within the Huber robust-loss family^{30,31}; Equation (8) is quadratic near zero and asymptotically linear for large $|e|$, retaining local curvature while bounding the marginal influence of large residuals. The structural candidate set was $\mathcal{C} = \{\text{DP, DG, RP, RG}\}$, denoting direct physical, direct geometry, residual physical, and residual geometry. Candidate $c \in \mathcal{C}$ determined both its coordinate domain— $q \in \{18, 2\}$ for physical coordinates or $q \in \{18, d\}$ for geometry coordinates—and its scalar target $t_{iq}^{(c)}$. Direct targets were measured coordinates; residual targets were measured coordinates minus the matching physical reference. Let $\mathbf{x}_i$ be the concatenated 34-predictor vector, $\mathcal{H}$ the prespecified tree-ensemble function class, $\mathcal{D}_{k,c}$ the fold-training records admissible for candidate c , and $\Omega(f)$ its tree-complexity penalty. For function $f \in \mathcal{H}$, the fitted coordinate learner minimized

$$\hat{f}_{q,k}^{(c)} = \operatorname{argmin}_{f \in \mathcal{H}} \left\{ \sum_{i \in \mathcal{D}_{k,c}} w_i \rho_\kappa(t_{iq}^{(c)} - f(\mathbf{x}_i)) + \Omega(f) \right\} \quad (9)$$

For $c \in \{\text{DP, DG}\}$, $\mathcal{D}_{k,c} = \mathcal{D}_k$; for $c \in \{\text{RP, RG}\}$, $\mathcal{D}_{k,c}$ retained only physically covered rows. Direct physical learning used the measured oxygen and hydrogen coordinates, whereas direct geometry used oxygen and d-excess. Residual physical learning used measured-minus-raw physical coordinates, and residual geometry used transformed-minus-anchor coordinates. This mapping fully defines the four parameterizations compared in the structural grid.

195 2.2.2 Physical reference and residual coordinates

196 For a physically covered record, the geometry anchor $\mathbf{a}_i$ used the product’s uncorrected monthly
 197 oxygen field $a_{18,i}$ and independently provided source d-excess $a_{d,i}^{\text{src}}$:

$$a_i = (a_{18,i}, a_{d,i}^{\text{src}})^T \quad (10)$$

198 For $q \in \{18, d\}$, let $\mathcal{T}_q$ denote coordinate q of the transform in Equation (4). The residual target r_{iq} was

$$r_{iq} = \mathcal{T}_q(y_i) - a_{iq} \quad (11)$$

199 For development record i , $K(i)$ denotes its excluded fold and superscript RG identifies the residual-
 200 geometry candidate in Equation (9). Its out-of-fold reference-centered coordinate estimate restored the
 201 anchor after learning the record-specific displacement:

$$\hat{g}_{iq}^{\text{RG, OOF}} = a_{iq} + \hat{f}_{q, K(i)}^{(\text{RG})}(x_i) \quad (12)$$

202 With $\hat{\mathbf{g}}_i^{\text{RG, OOF}} = (\hat{g}_{i,18}^{\text{RG, OOF}}, \hat{g}_{i,d}^{\text{RG, OOF}})^T$ and superscript dec denoting a decoded isotope pair, the paired RGF
 203 prediction was obtained by isotope-consistent inversion:

$$\hat{\mathbf{y}}_i^{\text{RGF, dec}} = \mathcal{T}^{-1}(\hat{\mathbf{g}}_{i,18}^{\text{RG, OOF}}, \hat{\mathbf{g}}_{i,d}^{\text{RG, OOF}}) \quad (13)$$

204 Equations (10)–(13) make the physical product a reference position rather than another exchangeable
 205 feature: the learners estimate record-specific displacement, restore each geometry coordinate, and
 206 decode one coupled isotope pair.

207 2.2.3 Residual parameterization and fusion-mass selection

208 For candidate $c \in \mathcal{C}$, let $\hat{\mathbf{y}}_i^{(c, \text{dec})}$ be its decoded oxygen–hydrogen prediction and $\hat{\mathbf{y}}_i^{\text{M}}$ the meteorology-
 209 only control pair. The development grid crossed all four parameterizations with fusion mass
 210 $m \in \{0.50, 0.75, 1.00\}$:

$$\hat{\mathbf{y}}_i^{(c, m)} = (1 - m)\hat{\mathbf{y}}_i^{\text{M}} + m\hat{\mathbf{y}}_i^{(c, \text{dec})} \quad (14)$$

211 Fusion was performed in the measured isotope-pair space. With $C_i \in \{0, 1\}$ denoting physical coverage
 212 for record i , absence of the physical reference invoked the exact meteorological pair:

$$\hat{\mathbf{y}}_i^{\text{RGF}} = \begin{cases} \hat{\mathbf{y}}_i^{\text{RGF, dec}}, & C_i = 1, \\ \hat{\mathbf{y}}_i^{\text{M}}, & C_i = 0. \end{cases} \quad (15)$$

213 Candidate selection used the same complete-station out-of-fold predictions as the development
 214 comparison. For each (c, m) , $L_{c,m}$, $E_{18,c,m}$, and $E_{2,c,m}$ denote the source-balanced joint loss, oxygen
 215 MAE, and hydrogen MAE formalized in Equations (26)–(28). The fixed identifier $\text{ID}_{c,m} \in \{0, \dots, 11\}$
 216 follows direct physical, direct geometry, residual physical, and residual geometry, with masses 0.50,
 217 0.75, and 1.00 nested in that order. The deterministic rule was

$$(\hat{c}, \hat{m}) = \text{lexmin}_{(c,m) \in \mathcal{C} \times \{0.50, 0.75, 1.00\}} (L_{c,m}, E_{18,c,m}, E_{2,c,m}, \text{ID}_{c,m}) \quad (16)$$

218 Equation (16) orders candidates lexicographically: joint loss is minimized first, followed only on an
 219 exact tie by $\delta^{18}\text{O}$ MAE, $\delta^2\text{H}$ MAE, and the fixed candidate identifier. Equations (14)–(16) define
 220 physical-space fusion, exact missing-coverage behavior, and deterministic selection. The selected

configuration was residual geometry at mass one (Supplementary Table S5). Residual geometry at mass one was therefore retained for the independent Nepal evaluation.

2.2.4 Model fitting for external evaluation

After structural selection, the retained meteorological XGBoost control and direct CatBoost pair were refitted with the fixed 34-field order, sample weights, targets, losses, and hyperparameters on all 8,267 development records. The residual-geometry XGBoost and CatBoost parents were refitted on the 8,197 physically covered records, retaining the development weights without covered-subset renormalization; Equation (15) supplied the meteorological pair for the other 70 records. For parent L , $\hat{\mathbf{g}}^L(\mathbf{x})$ denotes the full-development two-coordinate residual estimate at an independent query, with the physical anchor restored as in Equation (12). Development comparisons use the corresponding fold-excluded estimates. These fitted parents were applied to PAPIN/Indianapolis, Yadong/Pali, and Kathmandu/Lete, with estimator identities unchanged for the Kathmandu/Lete evaluation.

2.3 IsoRole: applicability-gated coordinate roles

2.3.1 A-RGF and geographic applicability

A-RGF combined the residual-geometry XGBoost and CatBoost parents in geometry coordinates. For learner $L \in \{\text{RGF} - \text{XGB}, \text{RGF} - \text{CAT}\}$, define $\hat{\mathbf{g}}^L(\mathbf{x}) = (\hat{g}_{18}^L(\mathbf{x}), \hat{g}_d^L(\mathbf{x}))^T$; $\hat{\mathbf{g}}^A(\mathbf{x})$ denotes their A-RGF blend. A-RGF used the same one-third/two-thirds blend in every external evaluation:

$$\hat{\mathbf{g}}^A(\mathbf{x}) = \frac{1}{3}\hat{\mathbf{g}}^{\text{RGF}-\text{XGB}}(\mathbf{x}) + \frac{2}{3}\hat{\mathbf{g}}^{\text{RGF}-\text{CAT}}(\mathbf{x}) \quad (17)$$

The blended geometry was decoded to a paired isotope estimate; when physical fields were unavailable, the meteorology-only estimate was used. The support rule used only station coordinates and the fixed development fold map. For stations a and b , let ϕ_a, ϕ_b denote latitude, λ_a, λ_b longitude, and $R = 6371.0088$ km Earth radius, with angles in radians. Great-circle distance was

$$D(a, b) = 2R \arcsin \sqrt{\sin^2\left(\frac{\phi_b - \phi_a}{2}\right) + \cos\phi_a \cos\phi_b \sin^2\left(\frac{\lambda_b - \lambda_a}{2}\right)} \quad (18)$$

Let $\ell^* = (\phi^*, \lambda^*)^T$ denote a query location and $\mathcal{S}_{\text{dev}}$ the 67 development stations. Its nearest-station support distance was

$$D_{\min}(\ell^*) = \min_{j \in \mathcal{S}_{\text{dev}}} D(\ell^*, j) \quad (19)$$

With j and ℓ indexing development stations, the envelope threshold τ was constructed entirely from the seed-3407 fold map as the largest leave-fold-out nearest-station distance:

$$\tau = \max_{j \in \mathcal{S}_{\text{dev}}} \min_{\ell \in \mathcal{S}_{\text{dev}}: K(\ell) \neq K(j)} D(j, \ell) = 1699.753317 \text{ km} \quad (20)$$

The indicator $\mathbb{1}\{\cdot\}$ equals one when its condition is true and zero otherwise. The target-free geographic-support state was

$$G(\ell^*) = \mathbb{1}\{D_{\min}(\ell^*) \leq \tau\} \quad (21)$$

Equations (17)–(21) define the mixed residual parent, query distance, threshold, and target-free support state. All 175 Nepal queries used the supported IsoRole formulation throughout the 0.75τ – 1.25τ sensitivity range.

251 2.3.2 Coordinate roles and complete inference

252 Yadong and Pali defined the coordinate-role assignment. Joint loss was 2.645 for support-gated A-
 253 RGF and 2.644 for direct CatBoost, while A-RGF had lower $\delta^{18}\text{O}$ MAE (4.5488 versus 4.5929‰) and
 254 direct CatBoost had lower d-excess MAE (7.3482 versus 9.2923‰). IsoRole therefore combined A-
 255 RGF oxygen with direct-pair d-excess for Kathmandu and Lete.

256 Let $\mathbf{x}^* \in \mathbb{R}^{34}$ denote the complete fixed-order query record, including the location ℓ^* , meteorological
 257 inputs, physical fields, and contextual descriptors. Direct CatBoost produced a paired isotope estimate
 258 from this record. Its d-excess coordinate was derived from that pair rather than trained as a separate
 259 model:

$$\hat{d}_{\text{dir}}(\mathbf{x}^*) = \widehat{\delta^2 H}_{\text{dir}}(\mathbf{x}^*) - 8\widehat{\delta^{18} O}_{\text{dir}}(\mathbf{x}^*) \quad (22)$$

260 For $C(\mathbf{x}^*) = G(\ell^*) = 1$, where $C(\mathbf{x}^*)$ is the query-level physical-coverage indicator, A-RGF supplied
 261 oxygen and the direct pair supplied d-excess. Superscript I denotes the supported IsoRole coordinate
 262 vector, and $[\cdot]_{18}$ extracts the oxygen coordinate from the geometry vector in Equation (17):

$$\hat{u}^I(\mathbf{x}^*) = ([\hat{\mathbf{g}}^A(\mathbf{x}^*)]_{18}, \hat{d}_{\text{dir}}(\mathbf{x}^*))^T \quad (23)$$

263 The supported pair was decoded exactly:

$$\hat{y}^I(\mathbf{x}^*) = \mathcal{T}^{-1}(\hat{u}^I(\mathbf{x}^*)) \quad (24)$$

264 Under all physical-coverage and geographic-support conditions, IsoRole returned one paired estimate.
 265 Let $\mathbf{a}^{\text{raw}}(\mathbf{x}^*)$ be the uncorrected monthly physical isotope pair, distinct from the oxygen–source-d-
 266 excess anchor in Equation (10), and let $\hat{\mathbf{y}}^M(\mathbf{x}^*)$ be the meteorological control pair:

$$\hat{\mathbf{y}}^{\text{IsoRole}}(\mathbf{x}^*) = \begin{cases} \hat{\mathbf{y}}^M(\mathbf{x}^*), & C(\mathbf{x}^*) = 0, \\ \mathbf{a}^{\text{raw}}(\mathbf{x}^*), & C(\mathbf{x}^*) = 1, G(\ell^*) = 0, \\ \hat{\mathbf{y}}^I(\mathbf{x}^*), & C(\mathbf{x}^*) = 1, G(\ell^*) = 1. \end{cases} \quad (25)$$

267 Equations (22)–(25) preserve the identity of the direct pair, assign the two coordinate roles, and return
 268 one isotope-consistent pair under every physical-coverage and geographic-support condition. Physical
 269 coverage is checked before geographic support because a supported location cannot activate an
 270 unavailable physical anchor. Algorithm 1 summarizes response-blind inference, including the date-
 271 only support history used in the 34-field predictor vector.

272 Algorithm 1. Response-blind IsoRole estimation under physical-coverage and geographic-support
 273 conditions. The procedure returns one isotope-consistent pair and identifies the governing estimator.

Input	Query rows; 16 earlier calendar-only date entries per confirmation station; ordered 34-variable predictor vector; fixed meteorological, residual-XGBoost, residual-CatBoost, direct-CatBoost, and raw-field estimators; 67 development coordinates and fold map; τ .
Output	One paired prediction per query, derived d-excess, physical-coverage and geographic-support indicators, and governing estimator.
1	Exclude measured isotope targets; derive date-only support descriptors and construct the ordered query vector $\mathbf{x}^* \in \mathbb{R}^{34}$.
2	For each query $\mathbf{x}^*$, evaluate physical coverage $C(\mathbf{x}^*)$.

Input	Query rows; 16 earlier calendar-only date entries per confirmation station; ordered 34-variable predictor vector; fixed meteorological, residual-XGBoost, residual-CatBoost, direct-CatBoost, and raw-field estimators; 67 development coordinates and fold map; τ .
Output	One paired prediction per query, derived d-excess, physical-coverage and geographic-support indicators, and governing estimator.
3	If $C(\mathbf{x}^*) = 0$, return the meteorology-only estimate; geographic distance is not required.
4	Otherwise , compute the nearest-station distance from the query coordinate to the 67 development stations using Equation (19).
5	If the nearest-station distance exceeds τ, return the uncorrected physical pair.
6	Otherwise , evaluate both residual parents, form A-RGF by Equation (17), and decode its physical isotope pair.
7	Evaluate the direct CatBoost isotope pair and derive its d-excess by Equation (22).
8	Assign A-RGF oxygen and direct-pair d-excess, and decode the supported pair by Equations (23)–(24).
9	Record the returned pair, derived d-excess, physical-coverage and geographic-support indicators, and governing estimator.
10	Return all paired estimates.

274 2.4 Evaluation and multiscale fidelity analysis

275 2.4.1 Comparator hierarchy and error estimands

276 Development placed RGF against six fixed baselines: fold-training climatology, raw ICM with
 277 meteorological fallback, meteorology-only XGBoost, physical-only XGBoost, direct XGBoost, and
 278 direct CatBoost. All models used the same complete-station folds and predictor specification. The 12-
 279 member grid isolated reference, coordinate, and mass effects; objective and learner analyses
 280 characterized the residual family. In Nepal, direct CatBoost and raw ICM with meteorological fallback
 281 served as the two prespecified primary comparators, and the A-RGF contrast evaluates the complete
 282 residual-ensemble configuration.

283 For station j , let $\mathcal{J}_j$ be its record set and $n_j = |\mathcal{J}_j|$. Coordinate $q \in \{1,2\}$ indexes oxygen or hydrogen;
 284 y_{iq} and $\hat{y}_{iq}$ are the measured and predicted values for record i , respectively. For $p \in \{1,2\}$, the station-
 285 level L_p error gives MAE at $p = 1$ and RMSE at $p = 2$:

$$E_{jq}^{(p)} = \left(\frac{1}{n_j} \sum_{i \in \mathcal{J}_j} |y_{iq} - \hat{y}_{iq}|^p \right)^{1/p}, p \in \{1,2\} \quad (26)$$

286 Let $\mathcal{S}_r$ be the station set of archive r and $J_r = |\mathcal{S}_r|$. The source-balanced coordinate error averaged
 287 station errors within each archive and then averaged archive means:

$$E_{q,SB}^{(p)} = \frac{1}{A} \sum_{r=1}^A \frac{1}{J_r} \sum_{j \in \mathcal{S}_r} E_{jq}^{(p)} \quad (27)$$

288 For Nepal, equal-station aggregation was $E_{q,ES}^{(p)} = \frac{1}{2} \sum_{j \in \{\text{Kathmandu, Lete}\}} E_{jq}^{(p)}$. Let aggregation label
 289 $\mathcal{A} \in \{\text{SB, ES}\}$ denote source-balanced development or equal-station Nepal. The primary endpoint

290 normalized coordinate MAEs by the prespecified contextual scales 1.8‰ and 14‰, fixed before
 291 structural candidate comparison and Nepal evaluation to balance the two isotope coordinates:

$$L_{\text{joint},\mathcal{A}} = \frac{1}{2} \left(\frac{E_{18,\mathcal{A}}^{(1)}}{1.8} + \frac{E_{2,\mathcal{A}}^{(1)}}{14} \right) \quad (28)$$

292 For evaluated method $\mathcal{M}$ and baseline $\mathcal{B}$, comparator-relative error reduction was

$$\text{Skill}_{\mathcal{A}}(\mathcal{M}; \mathcal{B}) = 100 \left[1 - \frac{L_{\text{joint},\mathcal{A}}(\mathcal{M})}{L_{\text{joint},\mathcal{A}}(\mathcal{B})} \right] \quad (29)$$

293 so positive values indicate lower error for $\mathcal{M}$ under the same aggregation $\mathcal{A}$ used for both method and
 294 baseline. Development intervals used 20,000 paired station-bootstrap replicates stratified by archive
 295 (seed 20260803). Nepal intervals used 20,000 paired circular moving-record-block replicates within
 296 station (seed 20260806), followed by equal-station aggregation. Records were ordered by event date
 297 and confirmation-row identifier. Replicates sampled circular 14-record blocks within station and
 298 truncated each concatenated sequence to the original station length. This block length was prespecified
 299 for primary IsoRole comparisons with direct CatBoost and raw ICM; 7-, 21-, and 28-record blocks
 300 provided sensitivity intervals. A paired analysis of IsoRole and A-RGF used the same interval
 301 construction across the 175 confirmation records. These intervals describe record-sequence uncertainty
 302 at Kathmandu and Lete; station-to-station uncertainty was not estimated. For target statistic θ ,
 303 bootstrap replicate θ_b^* , replicate index $b \in \{1, \dots, 20000\}$, and empirical quantile operator Q_α , the
 304 percentile interval was

$$\text{CI}_{95}(\theta) = [Q_{0.025}(\{\theta_b^*\}), Q_{0.975}(\{\theta_b^*\})] \quad (30)$$

305 Equations (26)–(30) fix the aggregation level, paired endpoint, comparator direction, and interval
 306 construction. For upper-tail performance, normalized record error was $\frac{1}{2}(|e_{18}|/1.8 + |e_2|/14)$, where
 307 $e_q = y_q - \hat{y}_q$; the 95th percentile (P95) was calculated within each station and then averaged equally
 308 across the two stations. Observation-model intervals used 20,000 station-bootstrap replicates (seed
 309 20260807).

310 2.4.2 Out-of-fold structural fidelity

311 The out-of-fold structural analysis applied the coordinate-role formulation to 8,267 complete-station
 312 predictions. The 8,197 physically covered rows used the active composition and the 70 uncovered rows
 313 the exact meteorological fallback. The reconstructed IsoRole series was compared with observations,
 314 direct CatBoost, and raw ICM for annual-harmonic peak day, within- and between-station isotope-line
 315 slopes, the adjusted Mendeley 30-day precipitation coefficient, and within-station variance. Candidate-
 316 minus-observation phase difference Δ_{phase} was wrapped so that $-182.625 \leq \Delta_{\text{phase}} \leq 182.625$ days.
 317 Variance bias was $100(V_{\text{cand}}/V_{\text{obs}} - 1)\%$, where V_{cand} and V_{obs} are equal-station means of within-
 318 station population variance for the reconstructed and measured series, respectively; signed error was
 319 $e = y - \hat{y}$. Each metric used 20,000 paired station-bootstrap replicates (master seed 20260810; offsets
 320 in Supplementary Table S12). This analysis quantified alignment with measured structures relevant to
 321 hydroclimate interpretation.

2.4.3 Observation-centered multiscale models

Observation-centered models were fitted to measured values and defined the reference structures used in the out-of-fold fidelity analysis. For coordinate $q \in \{18, 2, d\}$ across J stations, let $\mu_{jq} = n_j^{-1} \sum_{i \in \mathcal{I}_j} y_{iq}$ be the mean at station j . With Var denoting population variance, the between-station fraction B_q was

$$B_q = \frac{\text{Var}_j(\mu_{jq})}{\text{Var}_j(\mu_{jq}) + J^{-1} \sum_{j=1}^J \text{Var}_{i \in \mathcal{I}_j}(y_{iq} - \mu_{jq})} \quad (31)$$

Both variance operators in Equation (31) use the represented station or record count as divisor ($\text{ddof} = 0$). Isotope-line geometry was estimated at representation $v \in \{\text{raw}, \text{within}, \text{between}\}$. Observation index v denotes a record-station pair for raw and within-station fits or a station for the between-station fit; $X_v^{(v)}$ and $Y_v^{(v)}$ are the corresponding oxygen and hydrogen values. In Equation (32), α_v , β_v , and $\varepsilon_v^{(v)}$ are the intercept, isotope-line slope, and residual:

$$Y_v^{(v)} = \alpha_v + \beta_v X_v^{(v)} + \varepsilon_v^{(v)}, v \in \{\text{raw}, \text{within}, \text{between}\} \quad (32)$$

Raw-record and within-station fits assigned equal total mass to each station; between-station fits used one mean per station. Here i indexes a record within station j . For zero-based day of year t_{ij} , the annual basis was $s_{ij}^{\text{ann}} = \sin(2\pi t_{ij}/365.25)$ and $c_{ij}^{\text{ann}} = \cos(2\pi t_{ij}/365.25)$. Superscript ctr denotes station centering. For $q \in \{18, 2, d\}$, α_q is the intercept, $b_{s,q}$ and $b_{c,q}$ are the sine and cosine coefficients, and ε_{ijq} is the residual:

$$y_{ijq}^{(\text{ctr})} = \alpha_q + b_{s,q} s_{ij}^{(\text{ann}, \text{ctr})} + b_{c,q} c_{ij}^{(\text{ann}, \text{ctr})} + \varepsilon_{ijq} \quad (33)$$

Annual-cycle magnitude was summarized by $A_q^{\text{ann}} = (b_{s,q}^2 + b_{c,q}^2)^{1/2}$, whereas timing was represented by $t_q^{\text{peak}} = \text{mod}\left(\frac{365.25}{2\pi} \text{atan2}(b_{s,q}, b_{c,q}), 365.25\right)$. Here $\text{mod}(a, b)$ returns the least nonnegative remainder and maps phase to $[0, 365.25)$. For the pooled latitude relation, $h \in \{18, 2, d\}$ indexes isotope coordinate, $\mu_{jh} = n_j^{-1} \sum_{i \in \mathcal{I}_j} y_{ijh}$ is the station mean, ϕ_j is station latitude, $\sigma(j)$ its archive, $\mu_{\sigma(j)h}$ the archive response mean, and $\mu_{|\phi|, \sigma(j)}$ the archive mean absolute latitude. The latitude-model intercept, slope, and residual are $\alpha_h^{(\text{lat})}$, $\beta_{|\phi|, h}$, and $\varepsilon_{jh}^{(\text{lat})}$:

$$\mu_{jh} - \mu_{\sigma(j)h} = \alpha_h^{(\text{lat})} + \beta_{|\phi|, h} (|\phi_j| - \mu_{|\phi|, \sigma(j)}) + \varepsilon_{jh}^{(\text{lat})} \quad (34)$$

Source-specific latitude fits weighted stations equally; the pooled fit weighted archives equally and stations equally within archive. For physical coherence and hydroclimate associations, let $z_{\text{ES}}(a) = [a - \mu_{\text{ES}}(a)]/\sigma_{\text{ES}}(a)$ denote variable-specific equal-station standardization. Here $\mu_{\text{ES}}(a)$ and $\sigma_{\text{ES}}(a)$ are the analysis-weighted mean and population standard deviation of variable a ; this equal-station analysis weighting is distinct from the training weights w_i in Equation (7). Let $\xi_{ij}^{(\text{ctr})}$ be the centered focal predictor and $\mathbf{v}_{ij}^{(\text{ctr})} \in \mathbb{R}^{d_v}$ the centered adjustment vector. In Equation (35), $\alpha_h^{(\text{assoc})}$ is the association-model intercept, β_h the focal standardized coefficient, $\boldsymbol{\gamma}_h \in \mathbb{R}^{d_v}$ the adjustment coefficients, and $\varepsilon_{ijh}^{(\text{assoc})}$ the residual:

$$z_{\text{ES}}(y_{ijh}^{(\text{ctr})}) = \alpha_h^{(\text{assoc})} + \beta_h z_{\text{ES}}(\xi_{ij}^{(\text{ctr})}) + \boldsymbol{\gamma}_h^T z_{\text{ES}}(\mathbf{v}_{ij}^{(\text{ctr})}) + \varepsilon_{ijh}^{(\text{assoc})} \quad (35)$$

The vector operator $\mathbf{z}_{ES}$ standardizes each component with the same rule. For physical coherence, ξ was the matching monthly physical anomaly and $d_v = 0$. For the 689 Mendeley event/interval records, three models set ξ to $\ln(1 + P)$, where P was collection-interval, prior 7-day, or prior 30-day precipitation; $\mathbf{v}$ contained temperature, relative humidity, and station-centered annual sine and cosine, so $d_v = 4$. Equations (31)–(35) estimate the variance, isotope-line, annual-phase, latitude, physical-background, and antecedent-precipitation structures analyzed below.

3 Results

3.1 Physical-reference residualization improves over isolated meteorological and physical learners

3.1.1 Source-balanced accuracy and comparator structure

Across 8,267 held-station records, RGF achieved source-balanced station MAE of 2.750‰ for $\delta^{18}\text{O}$ and 21.431‰ for $\delta^2\text{H}$, with RMSE of 3.362‰ and 26.141‰ and normalized joint loss 1.529. Relative to meteorology-only XGBoost, RGF reduced joint error by 13.13% (95% station-bootstrap interval, 8.26–17.59%). Relative to physical-only XGBoost, it reduced joint error by 7.00% (1.14–12.26%). The same model reduced joint loss by 25.68% relative to raw ICM and by 28.89% relative to fold-training climatology (Figure 3; Table 1).

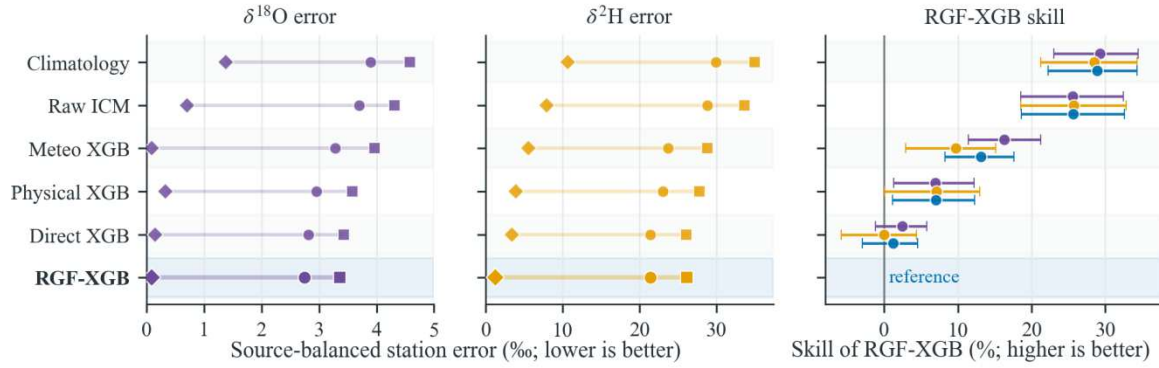
Table 1. Complete-station development accuracy and comparator-relative skill. Rows report source-balanced coordinate errors, joint loss, and paired intervals.

Method	Comparison role	$\delta^{18}\text{O}$ MAE (‰)	$\delta^2\text{H}$ MAE (‰)	Joint loss	Skill vs row [95% CI] (%)
Cross-fold climatology	traditional baseline	3.891	29.960	2.151	+28.89 [+22.28, +34.31]
Raw ICM + fallback	raw physical field	3.697	28.865	2.058	+25.68 [+18.63, +32.63]
Meteorology XGBoost	meteorology-only learner	3.286	23.742	1.761	+13.13 [+8.26, +17.59]
Physical-only XGBoost	physical-field learner	2.955	23.062	1.644	+7.00 [+1.14, +12.26]
Direct full-feature XGBoost	same-information direct control	2.819	21.432	1.548	+1.24 [−2.91, +4.58]
Direct full-feature CatBoost-MAE	strongest same-information control	2.738	21.460	1.527	−0.15 [−7.00, +5.31]
Residual-geometry fusion (XGBoost)	primary residual-geometry implementation	2.750	21.431	1.529	—

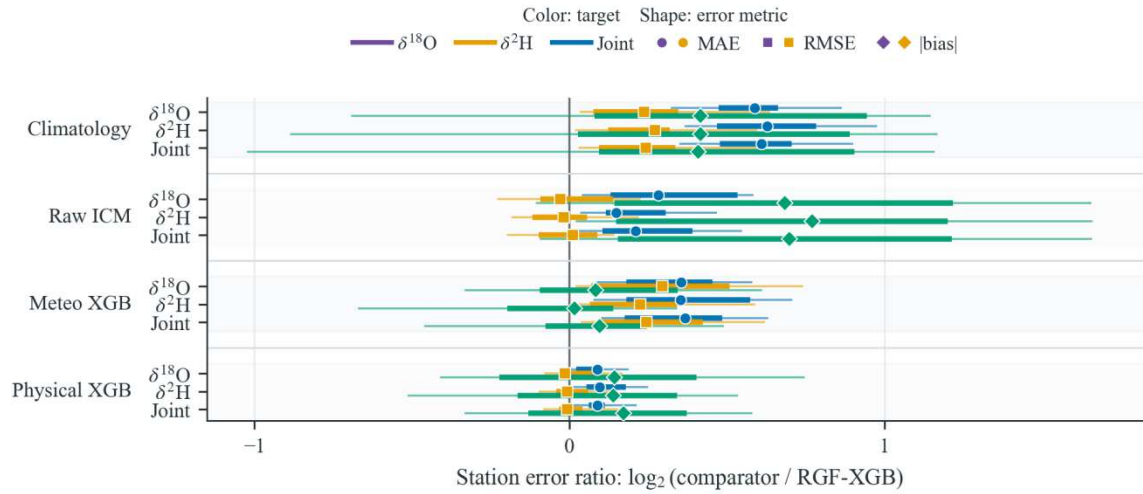
RGF and direct CatBoost had joint losses of 1.529 and 1.527, respectively (Table 1; Supplementary Table S9). RGF supplied the reference-centered residual model used in A-RGF and IsoRole.

3.1.2 Recurrence across archives and stations

RGF reduced station-level joint error relative to meteorology-only XGBoost at 56 of 67 sites (Figure 3; Supplementary Table S7). In Australian GNIP, oxygen and hydrogen MAE were 1.604‰ and 11.780‰, joint skill was 23.54%, and RGF yielded lower error at 21 of 22 stations. The tropical archive yielded MAE of 2.402‰ and 19.079‰, 18.28% joint skill, and lower error at 14 of 15 stations. The Mendeley collection showed 5.02% positive joint skill across its geographically broad 30-station set. These gains recurred across all three archives.



(A) Network benchmark and relative skill



(B) Held-station gains by archive

Figure 3. Residual-geometry gains persist from source-balanced network summaries to held stations across three isotope archives. Panel A aligns network error and bootstrap skill; panel B shows archive-specific station medians, interquartile ranges, and 10th–90th percentiles.

3.2 Physical–meteorological disagreement stratifies development gains

3.2.1 Prior disagreement stratifies realized development gains

Physical–meteorological disagreement provided a target-independent ranking of realized reconstruction gains. Across five equal-record disagreement strata, joint skill relative to meteorology increased from 3.36% to 8.05%, 11.70%, 20.68%, and 35.10% (Figure 4). In the highest stratum, oxygen and hydrogen skill reached 37.55% and 32.51%. The monotonic trajectory identified progressively larger RGF gains as the independently generated meteorological and physical states diverged.

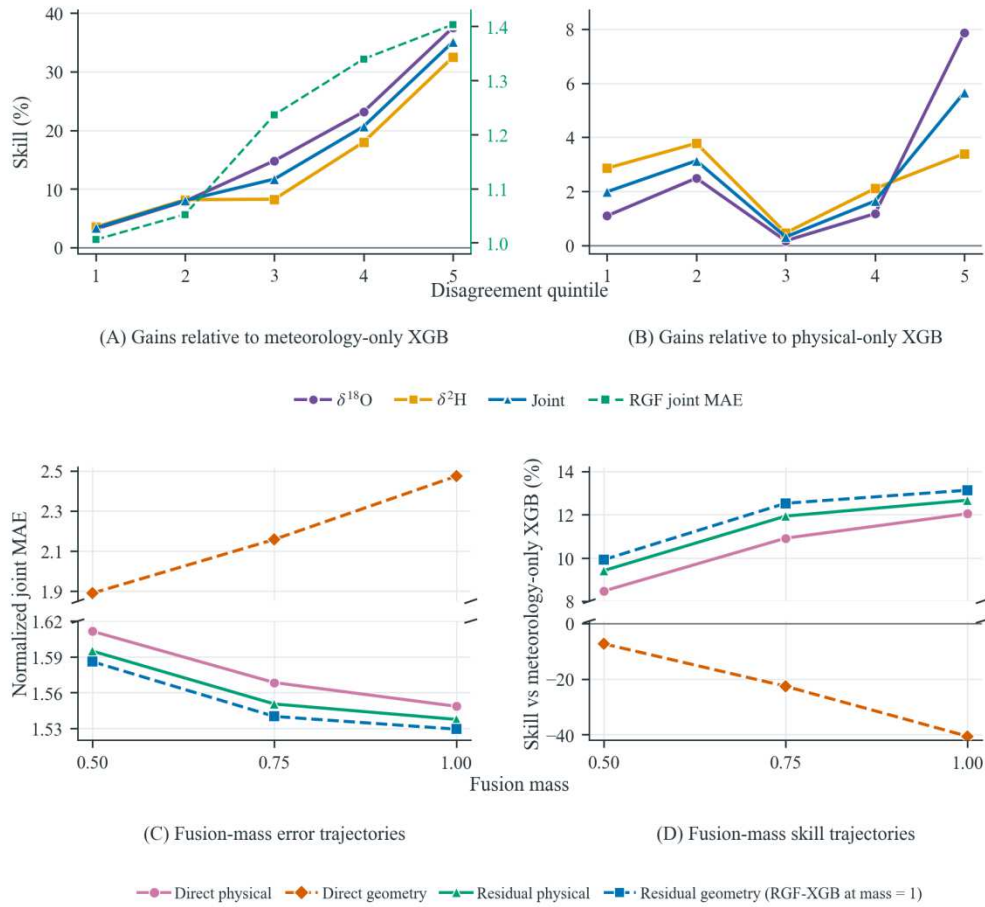


Figure 4. Physical–meteorological disagreement stratifies residual-geometry gains, while fusion mass traces blend sensitivity. Panels A–B show target-specific skill by disagreement quintile; panels C–D show four mass parameterizations, with diagonal breaks separating non-overlapping value bands.

3.2.2 Mass, objective, and learner evidence

Fusion-mass trajectories separated residual geometry from the other parameterizations (Figure 4; Supplementary Table S5). At mass one, residual geometry yielded the lowest joint loss among the four parameterizations (1.529), compared with 1.556 for direct physical and 1.538 for residual physical learning. Both residual formulations strengthened as the physical-residual component received more mass.

The matched sensitivity experiments characterize the RGF family (Figure 5; Supplementary Table S6). Pseudo-Huber reached joint loss 1.529, compared with 1.556 for L1 and 1.574 for L2. Under learner substitution, CatBoost-MAE reached 1.513, XGBoost/Pseudo-Huber 1.529, CatBoost-RMSE 1.540, and LightGBM/Huber 1.570. The rankings identified two strong residual parents with distinct tree-building implementations. Their one-third XGBoost/two-thirds CatBoost-MAE blend was used in every external evaluation.

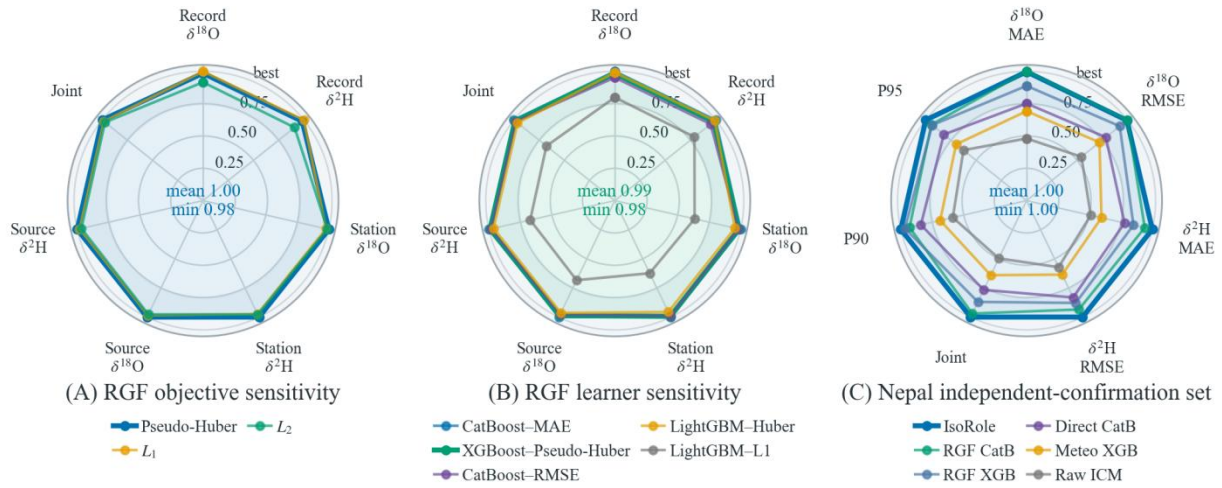


Figure 5. Objective and learner sensitivities identify the residual learners that form A-RGF and compare Nepal transfer configurations. Normalized radii show relative joint loss within each panel; larger values denote lower error.

3.3 IsoRole delivers the lowest paired point loss in response-blind Nepal confirmation

All 175 Nepal queries had physical fields and nearest-development distances of 750.829–944.400 km, below $\tau=1699.753$ km, so IsoRole used the supported residual-correction formulation; this assignment was unchanged at 0.75τ and 1.25τ . A-RGF reduced equal-station joint loss from 1.775 to 1.403 (20.95%), accounting for most of IsoRole's improvement. Coordinate composition further reduced loss to 1.360 (3.06% below A-RGF); the paired comparison across 175 confirmation records yielded a 95% interval of -2.62 to 7.69% . IsoRole reduced joint error by 23.37% relative to direct CatBoost (95% interval, 9.11–33.81%) and by 50.39% relative to raw ICM (38.74–60.02%; Figure 6; Table 2).

Table 2. IsoRole reconstruction accuracy in the Nepal independent-confirmation set. Brackets denote 95% intervals; A-RGF is a point estimate.

Method	$\delta^{18}\text{O}$ MAE (%)	$\delta^2\text{H}$ MAE (%)	Joint loss	Skill vs direct (%)	Skill vs raw (%)
IsoRole	2.404	19.382	1.360	+23.37 [+9.11, +33.81]	+50.39 [+38.74, +60.02]
A-RGF	2.404	20.584	1.403	+20.95	+48.83
Direct full-feature CatBoost-MAE	3.189	24.890	1.775	—	—
Raw ICM + fallback	4.985	37.994	2.742	—	—

IsoRole MAE was 2.404‰ for $\delta^{18}\text{O}$ and 19.382‰ for $\delta^2\text{H}$. Joint loss was 1.410 at Kathmandu and 1.311 at Lete, and both stations showed positive margins over direct CatBoost and raw ICM (Table 3). IsoRole also had the lowest P95 normalized record error among the evaluated transfer configurations. Because IsoRole inherits d-excess from the direct pair, IsoRole and direct CatBoost had identical d-excess MAE (5.785‰); full coordinate results are reported in Supplementary Figure S1 and Supplementary Table S8.

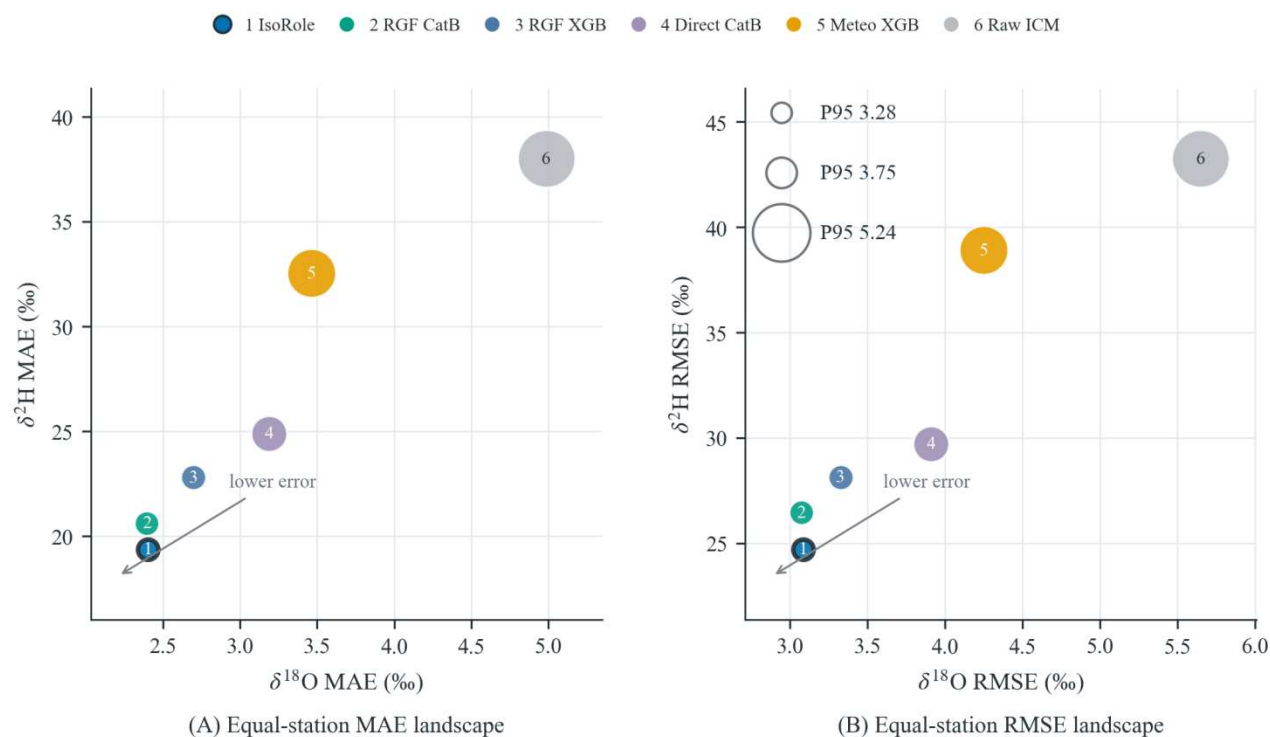


Figure 6. IsoRole attains the lowest paired error in the Nepal independent-confirmation set. Panels compare equal-station MAE and RMSE for 175 response-blind queries; symbol area denotes upper-tail normalized error.

Table 3. Station-specific Nepal confirmation separates the Kathmandu and Lete contributions. Rows report query support, coordinate MAE, and joint loss for IsoRole and its two primary comparators.

Method	Station	Queries	$\delta^{18}\text{O}$ MAE (%)	$\delta^2\text{H}$ MAE (%)	Joint loss
IsoRole	Kathmandu	64	2.484	20.147	1.410
IsoRole	Lete	111	2.325	18.616	1.311
Direct CatBoost-MAE	Kathmandu	64	2.472	20.840	1.431
Direct CatBoost-MAE	Lete	111	3.907	28.941	2.119
Raw ICM + fallback	Kathmandu	64	4.010	32.777	2.285
Raw ICM + fallback	Lete	111	5.961	43.211	3.199

Positive joint skill persisted across circular block lengths of 7, 14, 21, and 28 records. The 95% interval lower bound ranged from 6.86% to 12.32% relative to direct CatBoost and from 38.74% to 40.95% relative to raw ICM (Supplementary Table S11).

3.4 Reconstruction fidelity connects IsoRole with archive-specific isotope structure

3.4.1 Held-station reconstruction tracks selected isotope relations

In the out-of-fold structural analysis, IsoRole tracked selected measured structures more closely than raw ICM; Supplementary Table S12 reports the corresponding direct CatBoost comparison. In the Mendeley archive, predicted annual peak day differed from the measured peak by 4.47 days for $\delta^{18}\text{O}$ and -2.55 days for $\delta^2\text{H}$; raw ICM shifted the same peaks by -93.54 and -83.80 days. The tropical

between-station isotope-line slope was 7.890 in observations and 7.617 in IsoRole, whereas raw ICM yielded 6.879. For the Mendeley within-station line, the measured, IsoRole, and raw ICM slopes were 7.562, 7.826, and 8.040, respectively.

The adjusted 30-day precipitation coefficients were also retained in the reconstructed relation. Measured and IsoRole coefficients were -0.306 and -0.380 for $\delta^{18}\text{O}$ and -0.276 and -0.360 for $\delta^2\text{H}$. Standardized IsoRole error coefficients were -0.129 for $\delta^{18}\text{O}$ and -0.092 for $\delta^2\text{H}$, closer to zero than the corresponding raw-ICM values of -0.201 and -0.183 . Across the six archive–isotope combinations, IsoRole within-station variance remained 57.33–75.06% below the measured level. The reconstructed relations retained phase, slope, and antecedent-response structure, localizing further calibration to distributional amplitude.

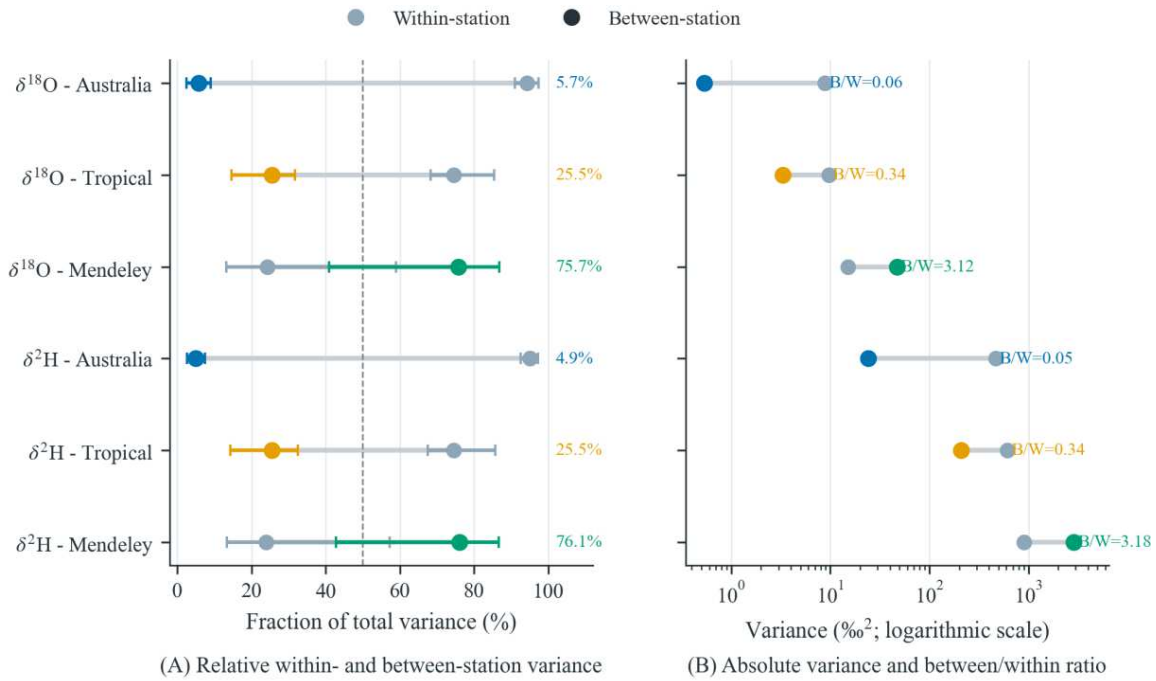


Figure 7. Precipitation-isotope variance shifts from within-station to between-station dominance across archives. Points show relative and absolute oxygen and hydrogen components; whiskers are station-bootstrap 95% intervals and labels give between/within ratios.

3.4.2 Archive-specific observation regimes separate temporal and spatial isotope structure

Between-station differences accounted for 5.7% of oxygen variance and 4.9% of hydrogen variance in the Australian monthly archive, 25.5% for both coordinates in the tropical daily archive, and 75.7% and 76.1%, respectively, in the Mendeley event/interval collection (Figure 7; Table 4A; Supplementary Table S2). The Australian archive concentrated information in repeated variation within stations, the tropical archive combined temporal and spatial structure, and the Mendeley collection concentrated information in geographic differences among station means. The corresponding d-excess components are shown in Supplementary Figure S2.

Isotope-line geometry changed across the archive-specific within- and between-station representations (Figure 8; Table 4A). Within-station slopes were 7.02 in Australia, 7.80 in the tropical archive, and 7.56 in Mendeley. Between-station slopes were 5.86, 7.89, and 7.78, respectively. The tropical and Mendeley records showed closely aligned temporal and spatial geometry, whereas Australia separated local trajectories from station-mean structure.

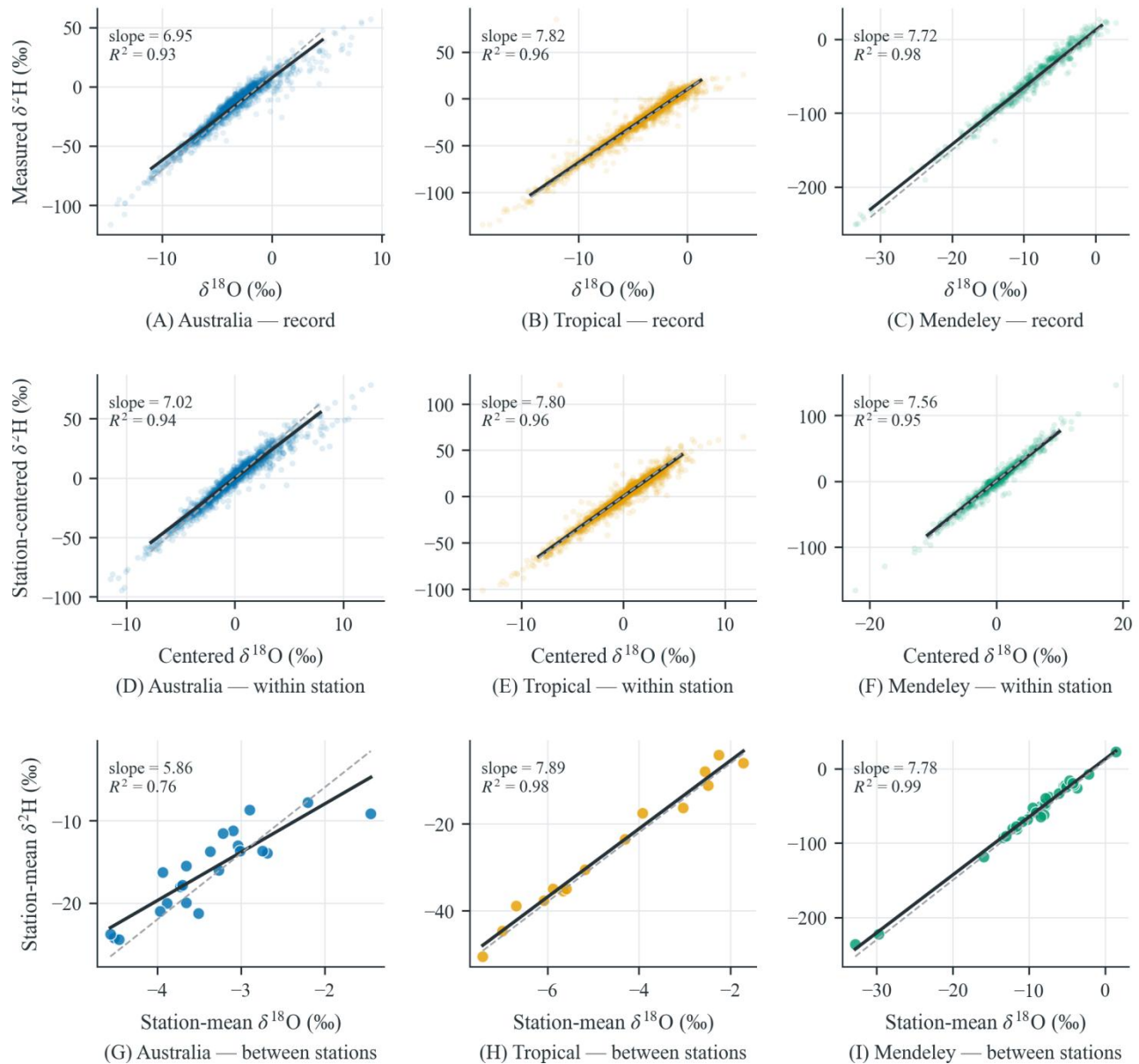


Figure 8. Isotope-line geometry separates record, within-station, and between-station representations across archives. Rows show raw records, station-centered observations, and station means; solid fits are compared with the slope-eight reference.

3.4.3 Annual phase, latitude, and physical background

Equal-station annual harmonics yielded oxygen and hydrogen amplitudes of 1.45‰ and 10.84‰ in Australia, 0.73‰ and 6.46‰ in the tropical archive, and 3.03‰ and 23.14‰ in Mendeley (Figure 9; Table 4B; Supplementary Table S3). Peak enrichment occurred near days 307 and 298 in Australia, 65 and 62 in the tropical archive, and 122 and 111 in Mendeley. These distinct cycles reinforce the need for season-matched precipitation inputs³²; d-excess harmonics are shown in Supplementary Figure S3.

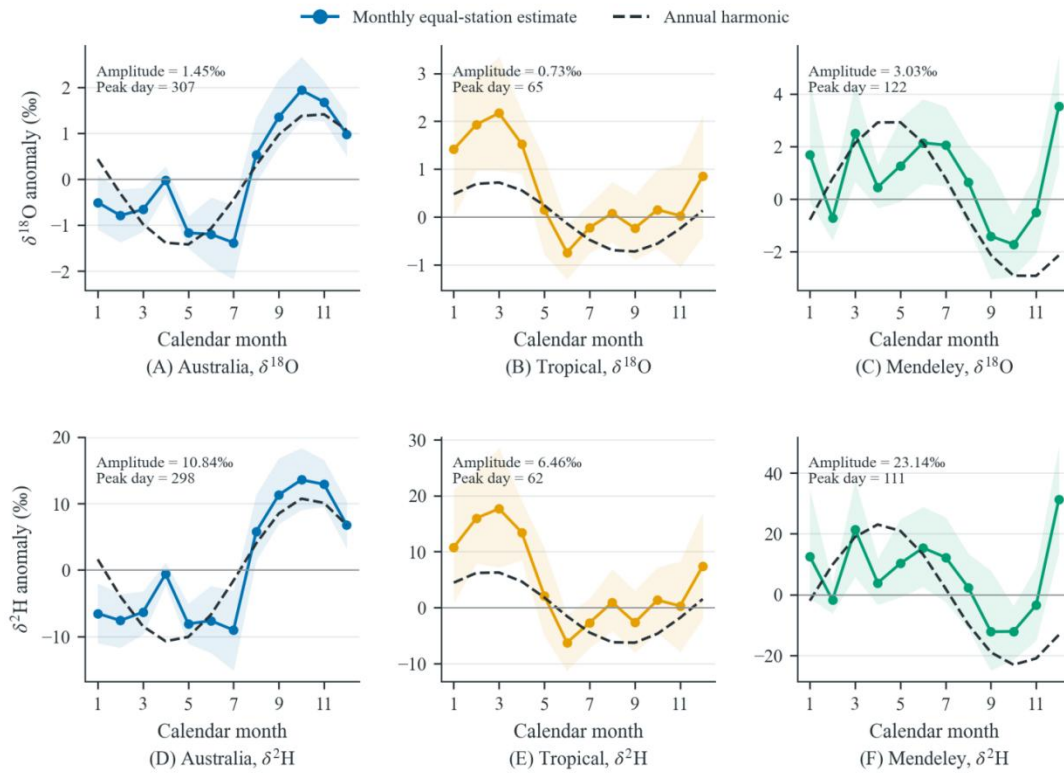


Figure 9. Archive-specific annual cycles organize oxygen and hydrogen isotope precipitation end members. Lines combine monthly equal-station anomalies, station-bootstrap 95% intervals, and annual harmonics for the Australian, tropical, and Mendeley archives.

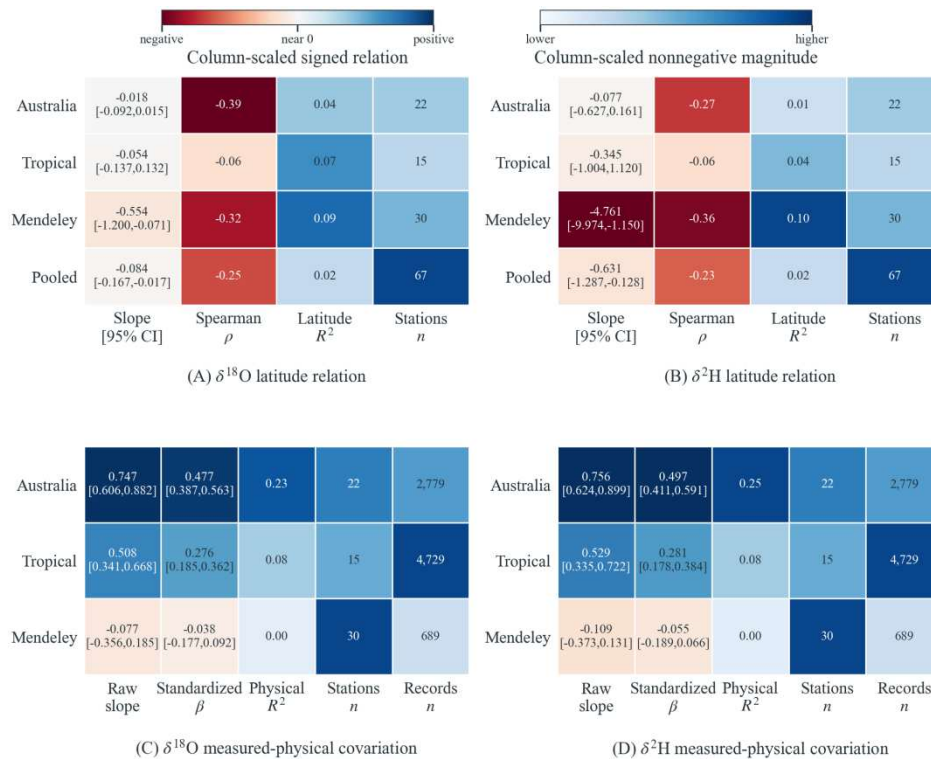


Figure 10. Latitude depletion and measured-physical coherence organize precipitation-isotope variation across archives. Cells report slopes, rank associations, explained variance, and observation support for measured oxygen and hydrogen isotope ratios.

After archive means were removed, station-mean isotope values became progressively depleted with absolute latitude: -0.084‰ per degree for oxygen and -0.631‰ per degree for hydrogen (Figure 10; Table 4B). This source-centered pooled direction was strongest in the geographically broad Mendeley archive. Supplementary Figure S4 reports the corresponding d-excess relations.

Measured monthly anomalies covaried with matching target-independent physical anomalies in the two long-sequence archives. Standardized oxygen and hydrogen coefficients were 0.477 and 0.497 in Australia and 0.276 and 0.281 in the tropical archive, with positive station-bootstrap intervals for all four estimates (Figure 10; Table 4B). This repeated coherence connects the physical reference used by RGF to measured temporal organization across two distinct network designs.

3.4.4 Thirty-day antecedent precipitation carries the strongest adjusted association

In the Mendeley event/interval records, the prior 30-day precipitation model had the strongest adjusted relationship with both measured isotope ratios (Figure 11; Table 4B; Supplementary Table S4). Standardized coefficients were -0.306 for oxygen and -0.276 for hydrogen, while collection-interval and 7-day coefficients were smaller. The 30-day association identifies antecedent precipitation as the most informative tested moisture-state index for event-scale isotope composition.

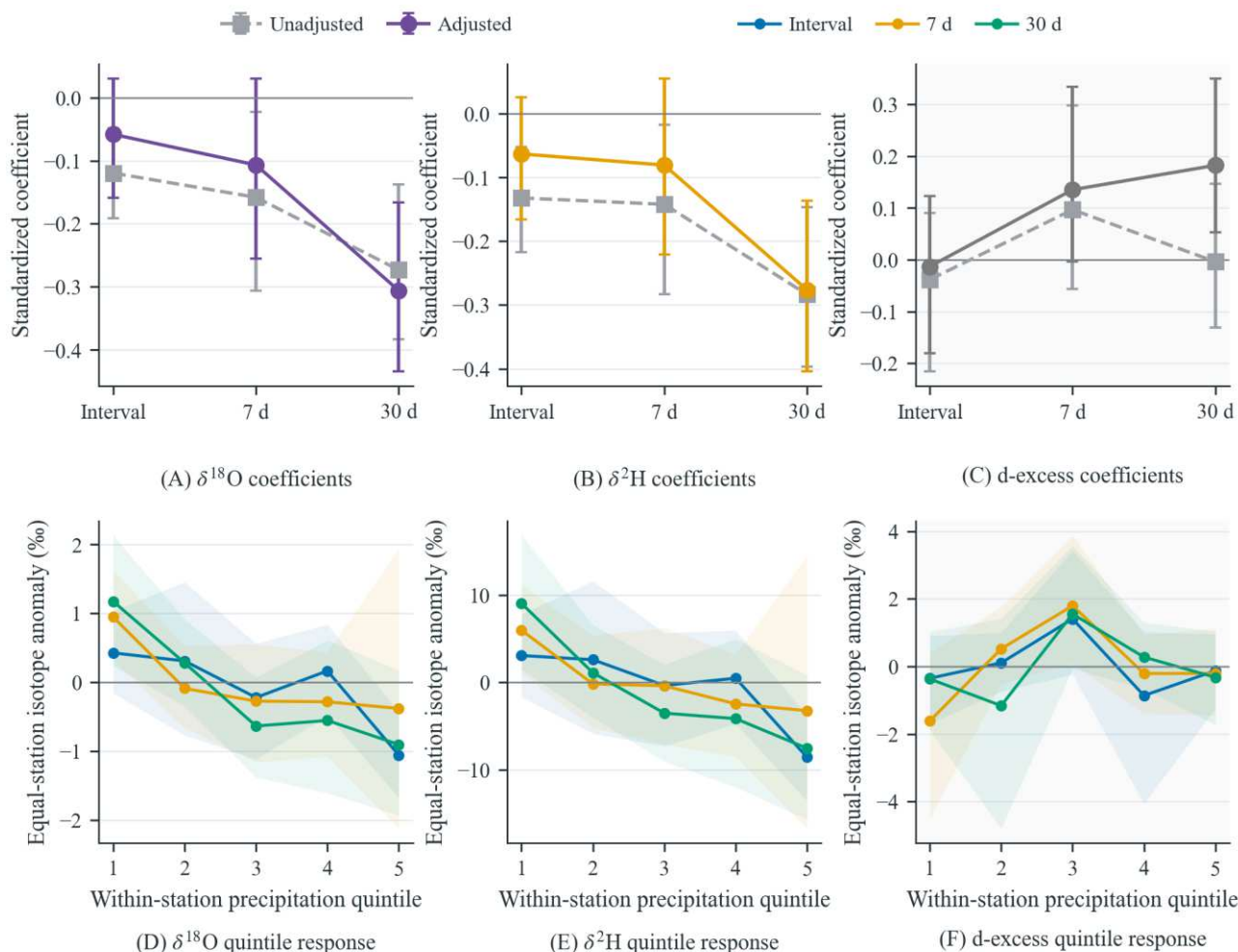


Figure 11. Thirty-day antecedent precipitation carries the strongest adjusted isotope association in the Mendeley archive. Upper panels compare standardized coefficients; lower panels show equal-station quintile responses for both measured isotope coordinates.

Table 4. Measured archive structure spans variance, isotope-line, seasonal, spatial, and antecedent regimes. Parts A and B report structural and environmental estimates; brackets denote bootstrap intervals, and em dashes denote unavailable estimates.

(A) Archive variance partitioning and isotope-line geometry

Source	Records, n	Stations	$\delta^{18}\text{O}$ between [95% CI] (%)	$\delta^2\text{H}$ between [95% CI] (%)	Within-station slope [95% CI]	Between-station slope [95% CI]
Australian GNIP monthly	2,849	22	+5.7 [+2.5, +9.0]	+4.9 [+2.7, +7.4]	7.016 [+6.650, +7.302]	5.857 [+4.494, +7.978]
IAEA tropical daily	4,729	15	+25.5 [+14.6, +31.8]	+25.5 [+14.3, +32.5]	7.796 [+7.547, +8.025]	7.890 [+7.211, +8.608]
Mendeley event/interval collection	689	30	+75.7 [+41.0, +86.7]	+76.1 [+42.8, +86.7]	7.562 [+7.205, +7.844]	7.775 [+7.543, +8.167]

(B) Seasonal, spatial, physical, and antecedent organization

Source / scope	Coordinate	Annual amplitude [95% CI] (‰)	Peak DOY	$ \phi $ slope [95% CI] (‰ degree ⁻¹)	Physical-anomaly β [95% CI]	30-day precipitation β [95% CI]
Australian GNIP monthly	$\delta^{18}\text{O}$	1.45 [+1.12, +1.89]	307	-0.018 [-0.092, +0.015]	+0.477 [+0.387, +0.563]	—
Australian GNIP monthly	$\delta^2\text{H}$	10.84 [+8.11, +14.49]	298	-0.077 [-0.627, +0.161]	+0.497 [+0.411, +0.591]	—
IAEA tropical daily	$\delta^{18}\text{O}$	0.73 [+0.20, +1.44]	65	-0.054 [-0.137, +0.132]	+0.276 [+0.185, +0.362]	—
IAEA tropical daily	$\delta^2\text{H}$	6.46 [+1.66, +12.26]	62	-0.345 [-1.004, +1.120]	+0.281 [+0.178, +0.384]	—
Mendeley event/interval collection	$\delta^{18}\text{O}$	3.03 [+2.07, +4.33]	122	-0.554 [-1.200, -0.071]	-0.038 [-0.177, +0.092]	-0.306 [-0.434, -0.166]
Mendeley event/interval collection	$\delta^2\text{H}$	23.14 [+16.02, +33.08]	111	-4.761 [-9.974, -1.150]	-0.055 [-0.189, +0.066]	-0.276 [-0.404, -0.136]
Source-adjusted pooled	$\delta^{18}\text{O}$	—	—	-0.084 [-0.167, -0.017]	—	—
Source-adjusted pooled	$\delta^2\text{H}$	—	—	-0.631 [-1.287, -0.128]	—	—

4 Discussion

4.1 IsoRole combines physical reference and learned displacement for transfer

IsoRole's principal contribution is a change in how the reconstruction target is represented. Regional isoscapes and dynamic isotope products encode a spatially and temporally organized background state, whereas direct machine-learning models estimate absolute isotope ratios from climate and geographic predictors. RGF instead places each dated observation around an independently generated monthly

515 isotope reference and assigns the learner the record-scale displacement from that reference. The
516 physical field therefore carries broad background organization, while sampling-aligned predictors
517 explain how the observed record departs from it. This division follows the wider logic of hybrid Earth-
518 system learning: retain available physical structure in the target representation and use statistical
519 learning for the unresolved component^{14,15,19,20}.

520 The process literature gives this division a physical interpretation. Moisture-source position, upstream
521 rainout, and large-scale evaporation–precipitation organization create coherent regional and monthly
522 isotope states^{33–35}. Convective organization, rain evaporation, entrainment, and meteorological
523 conditions along transport pathways add departures that can vary at the event or collection scale^{36–38}.
524 A reference-and-displacement formulation mirrors these nested controls without requiring the physical
525 field to reproduce every measured value directly. The anchor contributes its organized isotope state,
526 and the residual learner corrects the mismatch at the observation scale.

527 The development evidence is consistent with that mechanism. RGF improved on both meteorology-
528 only and physical-only learners, and its advantage increased as their target-independent predictions
529 diverged. Disagreement is therefore scientifically informative: it identifies records for which neither
530 an isolated meteorological mapping nor the physical field alone expresses the complete isotope state,
531 and where reference-centered correction has the largest empirical leverage. The gain recurred at 56 of
532 67 stations across all three archives.

533 A-RGF and coordinate composition contributed at different scales. The one-third XGBoost/two-thirds
534 CatBoost ensemble reduced Nepal joint loss by 20.95% relative to direct CatBoost. The
535 complementary Yadong/Pali ordering assigned A-RGF to $\delta^{18}\text{O}$ and the direct pair to d-excess for
536 Kathmandu/Lete confirmation. Because $\delta^2\text{H}$ is decoded from the selected $\delta^{18}\text{O}$ and d-excess
537 coordinates, the rule returns one algebraically coupled isotope pair rather than independently
538 combining two final isotope outputs. Coordinate composition further lowered Nepal loss to 1.360,
539 giving IsoRole its 23.37% advantage over direct CatBoost.

540 This positioning distinguishes IsoRole from both classical isoscape construction and direct
541 precipitation-isotope machine learning. Existing approaches have advanced regional interpolation,
542 climate-regression isoscapes, monthly time-series prediction, and spatial machine-learning maps^{14–17}.
543 IsoRole uses that same scientific arena for a more specific transfer question: whether an independent
544 physical isotope state can serve as the origin of a learned correction when the destination station
545 contributes no isotope response. Against direct CatBoost with the same predictors and complete-station
546 folds, the Nepal margin supports the combined value of reference-centered representation, residual
547 ensembling, and coordinate assignment under the same response-blind information set.

548 **4.2 Geographic support defines IsoRole applicability**

549 Response-blind evaluation changes the meaning of reconstruction accuracy. Repeated records from
550 one station share geography, climate, and temporal dependence; random record splits can therefore
551 reward recovery of a partly observed station rather than transfer to an unseen one. Complete-station
552 folds align training and evaluation with the intended scientific use by excluding the destination station
553 as a unit, consistent with recommendations for structured ecological and spatial data^{25,26}. Kathmandu
554 and Lete were absent from the 67-station development set, and their 175 isotope targets were
555 unavailable during fitting, threshold construction, routing, and prediction.

556 At a new station, IsoRole reconstructs a dated precipitation-isotope pair from target-independent
557 physical and meteorological information plus a short date-only support history. Direct CatBoost

558 controls the predictor set, and raw ICM represents the uncorrected physical alternative. Positive
559 margins at Kathmandu and Lete demonstrate transfer of the supported IsoRole formulation to stations
560 without fitted isotope responses.

561 Geographic support links each query to an estimator according to physical-field availability and
562 distance from the development network. Queries with physical coverage inside the envelope use
563 IsoRole; those outside the envelope use the uncorrected physical pair; queries without physical
564 coverage use the meteorological estimate. Because τ depends only on development-station coordinates,
565 each returned pair has a reproducible, target-independent applicability classification for hydrological
566 or ecological analysis.

567 All 175 Nepal queries used residual correction across thresholds from 0.75τ to 1.25τ . IsoRole
568 distinguished residual-corrected, physical-only, and meteorological-only estimates, and the Nepal
569 evaluation supports residual correction within the tested geographic support.

570 **4.3 Archive structure informs precipitation-isotope monitoring**

571 The three measured archives demonstrate that observation count and isotope information are not
572 interchangeable. In the Australian monthly archive, more than 94% of $\delta^{18}\text{O}$ and $\delta^2\text{H}$ variance occurred
573 within stations, so repeated observations resolved temporal distributions around comparatively stable
574 station means. In the Mendeley event/interval collection, approximately three quarters of the variance
575 occurred among station means, making geographic differentiation the dominant source of information.
576 The tropical daily archive occupied the middle of this continuum and contained substantial temporal
577 and spatial structure. These contrasts explain why archives with different network geometries
578 contribute different kinds of evidence even when all records enter one reconstruction task.

579 The variance partition yields a direct monitoring principle. A network intended to estimate seasonal
580 distributions, event sequences, or antecedent-state responses benefits from repeated reference stations,
581 because those estimands depend on within-station variation. A network intended to characterize
582 regional precipitation end members benefits from geographic spread, because its target is variation
583 among station means. Where both temporal transfer windows and regional end-member diversity
584 matter, a hybrid design combines intensively sampled sentinel stations with a broader set of spatially
585 distributed sites. Together, the three archives identify a network design that combines temporal depth
586 with spatial coverage for complementary isotope estimands.

587 Scale-specific isotope-line geometry sharpens the same point. Station-centered lines describe the
588 paired temporal trajectory around each local baseline, whereas lines fitted to station means describe
589 geographic differentiation among baselines. Tropical and Mendeley point estimates were closely
590 aligned across those representations, indicating similar paired-isotope geometry at their temporal and
591 spatial scales. Australia showed a larger separation between within- and between-station point
592 estimates, with overlapping intervals, indicating that a single pooled line would compress two distinct
593 structures. Reporting the representation together with the slope is therefore essential when isotope-line
594 geometry is used to evaluate reconstructed precipitation inputs.

595 These archive contrasts also clarify the value of the study's source-balanced evaluation. Equalizing
596 archive and station influence prevents the record-dense monthly network from defining performance
597 by volume alone, while complete-station exclusion requires the method to transfer across each
598 network's dominant structure. The heterogeneous archive set is thus not simply a larger training pool:
599 it combines repeated temporal trajectories, daily tropical event sequences, and broad geographic end-
600 member variation. A complementary network design combines intensively sampled reference stations

601 with spatially distributed sites, strengthening temporal-state estimation, regional end-member
602 coverage, and station-level evaluation.

603 **4.4 Seasonal, spatial, and antecedent states frame end-member relevance**

604 The observation-centered analyses organize precipitation-isotope inputs along three nested axes.
605 Annual harmonics locate enriched and depleted phases within the year; archive-centered latitude
606 relations describe a broad spatial direction among station means; and antecedent precipitation
607 characterizes the hydroclimatic state preceding an individual collection. Global precipitation isotopes
608 are strongly seasonal, but the phase and amplitude observed here differed among archives³². The
609 relevant precipitation end member is therefore indexed jointly by place, time of year, and recent
610 moisture history rather than by a regional annual mean alone.

611 The 30-day Mendeley association adds an event-scale interpretation to that hierarchy. Precipitation
612 isotopes integrate source-region rainout, transport, and convective processing over pathways that need
613 not coincide with the collection interval or local event total^{33,36,38}. The strongest adjusted association
614 at 30 days indicates that recent accumulated precipitation provides a more informative antecedent index
615 in this archive than shorter tested windows. The association identifies the temporal context that should
616 accompany a reconstructed event value used as a hydroclimatic tracer.

617 This context matters because receiving waters integrate precipitation over storage and travel-time
618 distributions, and those distributions differ among catchments and compartments^{7,39}. Groundwater
619 recharge and plant-water use are also seasonally selective, so annual amount-weighted precipitation
620 need not represent the water actually transmitted to an aquifer or accessed by vegetation^{5,8}. IsoRole
621 supplies dated, sampling-aligned paired estimates for aggregation over application-relevant transfer
622 windows. The archive analyses identify the temporal repetition, spatial coverage, and antecedent
623 window needed to align those estimates with a receiving-water process.

624 The out-of-fold structural analysis connects predictive performance with application-relevant isotope
625 relations. IsoRole tracked the highlighted annual-phase, isotope-line, and 30-day precipitation
626 coefficients more closely than raw ICM while retaining exact paired-isotope identity, and measured-
627 physical anomaly coherence supports use of the monthly field as an organized reference state. The
628 lower reconstructed within-station variance localizes additional calibration to distributional amplitude
629 while retaining the reference-centered mean, phase, and coupled geometry. IsoRole reconstructs the
630 sampling-aligned paired state, and distributional calibration can refine its spread for variance- or tail-
631 sensitive applications.

632 Together, these results connect method design to end-member interpretation. A physical isotope field
633 supplies the regional and monthly reference, learned displacement resolves the dated observation,
634 response-blind evaluation establishes transfer to an unseen station, and archive structure determines
635 how those reconstructed observations should be aggregated for hydrological use. Together, the
636 evidence defines an application principle: precipitation-isotope reconstruction is most informative
637 when the estimator, applicability state, and receiving-water transfer window are represented together.

638 **5 Conclusion**

639 IsoRole delivered response-blind paired precipitation-isotope reconstruction at new stations through a
640 fixed physical-reference residual ensemble and coupled-isotope completion rule. RGF established a
641 13.13% source-balanced joint gain over meteorology-only learning and a 7.00% gain over physical-
642 only learning across 67 held stations. At two stations in the Nepal independent-confirmation set, A-

RGF accounted for most of the point reduction, lowering joint loss by 20.95% relative to direct CatBoost; the complete IsoRole rule attained the lowest point loss and overall reductions of 23.37% relative to direct CatBoost and 50.39% relative to raw ICM.

The out-of-fold structural analysis retained selected annual-phase, isotope-line, and 30-day-response relations, while measured networks showed archive-specific balances between temporal and spatial information; distributional amplitude remains the principal calibration target. Within the evaluated station and geographic-support domains, the evidence supports a unified end-member principle: anchor reconstruction to the physical isotope state, learn sampling-record displacement, and match reconstructed precipitation to the receiving-water transfer window.

Data availability

All source datasets analysed in this study are publicly available. The Australian precipitation-isotope records are available from the Australian GNIP Rainfall Data 1960–2014 collection in the ANSTO Open Science Portal (<https://openscience.ansto.gov.au/collection/881>; associated publication: <https://doi.org/10.1016/j.scitotenv.2018.07.082>) and from the 2015–2024 update archived at <https://doi.org/10.5281/zenodo.17773309>. The tropical daily precipitation-isotope data are available from Figshare (<https://figshare.com/s/fdfabb43a844cad530a5>) and are described at <https://doi.org/10.1038/s41598-019-50973-9>. The global precipitation stable-isotope compilation is available from Mendeley Data at <https://doi.org/10.17632/75tscprd2.1>, and the Kathmandu and Lete records are available from the Kali Gandaki High Mountain Observatory database at <https://doi.org/10.5880/GFZ.4.6.2021.004>. The PAPIN, Indianapolis, and Yadong–Pali precipitation-isotope datasets are available at <https://doi.org/10.5281/zenodo.4767053>, <https://doi.org/10.1594/PANGAEA.895300>, and <https://doi.org/10.5281/zenodo.8059700>, respectively. Meteorological forcing data were obtained from the publicly accessible NASA POWER Daily API and Analysis Ready Datastore (<https://power.larc.nasa.gov/docs/services/>). The physical-isotope predictors were extracted from the uncorrected monthly NetCDF fields (UncorrPrec_Monthly_netCDF.zip) in version 3 of the global precipitation-isotope product archived at <https://doi.org/10.5281/zenodo.21377929>. The processed data, deterministic station-fold assignments, and model outputs underlying the figures and tables are available from the corresponding author upon reasonable request.

Code availability

The custom code used for data preprocessing, predictor construction, model training, complete-station cross-validation, independent evaluation, statistical analysis, and figure generation is available from the corresponding author upon reasonable request.

Author contributions

S.G.: Methodology, Software, Formal Analysis, Data Curation, and Writing—Original Draft. D.J.: Conceptualization, Supervision, Project Administration, Funding Acquisition, and Writing—Review and Editing. Y.H.: Investigation, Data Curation, and Validation. D.B.: Resources, Investigation, and Data Curation. S.L.: Validation and Writing—Review and Editing. D.Z.: Supervision, Project Administration, and Writing—Review and Editing. All authors reviewed and approved the final manuscript.

Funding

684 This research was supported by the Natural Science Foundation of Inner Mongolia Autonomous
685 Region (grant no. 2024ZD18).

686 Conflict of interest

687 The authors declare that the research was conducted in the absence of any commercial or financial
688 relationships that could be construed as a potential conflict of interest.

689 References

- 690 1. Craig, H. Isotopic Variations in Meteoric Waters. *Science* **133**, 1702–1703 (1961).
- 691 2. Dansgaard, W. Stable isotopes in precipitation. *Tellus* **16**, 436–468 (1964).
- 692 3. Gat, J. R. OXYGEN AND HYDROGEN ISOTOPES IN THE HYDROLOGIC CYCLE.
693 *Annual Review of Earth and Planetary Sciences* **24**, 225–262 (1996).
- 694 4. Pfahl, S. & Sodemann, H. What controls deuterium excess in global precipitation? *Climate of*
695 *the Past* **10**, 771–781 (2014).
- 696 5. Jasechko, S. *et al.* The pronounced seasonality of global groundwater recharge. *Water*
697 *Resources Research* **50**, 8845–8867 (2014).
- 698 6. Evaristo, J., Jasechko, S. & McDonnell, J. J. Global separation of plant transpiration from
699 groundwater and streamflow. *Nature* **525**, 91–94 (2015).
- 700 7. Ala-aho, P., Tetzlaff, D., McNamara, J. P., Laudon, H. & Soulsby, C. Using isotopes to
701 constrain water flux and age estimates in snow-influenced catchments using the STARR
702 (Spatially distributed Tracer-Aided Rainfall–Runoff) model. *Hydrology and Earth System*
703 *Sciences* **21**, 5089–5110 (2017).
- 704 8. Allen, S. T., Kirchner, J. W., Braun, S., Siegwolf, R. T. W. & Goldsmith, G. R. Seasonal origins
705 of soil water used by trees. *Hydrology and Earth System Sciences* **23**, 1199–1210 (2019).
- 706 9. Hollins, S. E., Hughes, C. E., Crawford, J., Cendón, D. I. & Meredith, K. T. Rainfall isotope
707 variations over the Australian continent – Implications for hydrology and isoscape applications.
708 *Science of The Total Environment* **645**, 630–645 (2018).
- 709 10. Munksgaard, N. C. *et al.* Data Descriptor: Daily observations of stable isotope ratios of rainfall
710 in the tropics. *Scientific Reports* **9**, 14419 (2019).
- 711 11. Andermann, C. *et al.* Kali Gandaki High Mountain Observatory, Stable Water Isotope database.
712 GFZ Data Services <https://doi.org/10.5880/gfz.4.6.2021.004> (2021).
- 713 12. Zhu, G. Global precipitation stable isotope data. Mendeley Data
714 <https://doi.org/10.17632/75tsccprd2.1> (2024).
- 715 13. Hughes, C. E., Hegarty, N. R., Farrawell, K.-A. L. & Chisari, R. Australian rainfall $\delta^{18}\text{O}$ and
716 $\delta^2\text{H}$ data 2015–2024 (GNIP). Zenodo <https://doi.org/10.5281/zenodo.17773309> (2025).
- 717 14. Bowen, G. J. & Revenaugh, J. Interpolating the isotopic composition of modern meteoric
718 precipitation. *Water Resources Research* **39**, 1299 (2003).
- 719 15. Terzer, S., Wassenaar, L. I., Araguás-Araguás, L. J. & Aggarwal, P. K. Global isoscapes for
720 $\delta^{18}\text{O}$ and $\delta^2\text{H}$ in precipitation: improved prediction using regionalized climatic regression
721 models. *Hydrology and Earth System Sciences* **17**, 4713–4728 (2013).

- 722 16. Nelson, D. B., Basler, D. & Kahmen, A. Precipitation isotope time series predictions from
723 machine learning applied in Europe. *Proceedings of the National Academy of Sciences* **118**,
724 e2024107118 (2021).
- 725 17. Erdélyi, D. *et al.* Predicting spatial distribution of stable isotopes in precipitation by classical
726 geostatistical- and machine learning methods. *Journal of Hydrology* **617**, 129129 (2023).
- 727 18. Adachi, I. & Yamanaka, T. Global dynamic precipitation isoscapes over three-quarters of a
728 century. Discussion paper at <https://doi.org/10.5194/essd-2026-52> (2026).
- 729 19. Reichstein, M. *et al.* Deep learning and process understanding for data-driven Earth system
730 science. *Nature* **566**, 195–204 (2019).
- 731 20. Karniadakis, G. E. *et al.* Physics-informed machine learning. *Nature Reviews Physics* **3**, 422–
732 440 (2021).
- 733 21. Risi, C. *et al.* What controls the isotopic composition of the African monsoon precipitation?
734 Insights from event-based precipitation collected during the 2006 AMMA field campaign.
735 *Geophysical Research Letters* **35**, L24808 (2008).
- 736 22. He, S. *et al.* Understanding Tropical Convection Through Triple Oxygen Isotopes of
737 Precipitation From the Maritime Continent. *Journal of Geophysical Research: Atmospheres*
738 **126**, e2020JD033418 (2021).
- 739 23. Adachi, I. & Yamanaka, T. Global climatology, monthly, daily precipitation isotope dataset,
740 1947–2024 (NetCDF and GeoTIFF). Zenodo <https://doi.org/10.5281/zenodo.21377929> (2026).
- 741 24. NASA Langley Research Center POWER Project. POWER Universal Time Coordinated (UTC)
742 Time Series. NASA Prediction Of Worldwide Energy Resources (POWER) Project.
- 743 25. Roberts, D. R. *et al.* Cross-validation strategies for data with temporal, spatial, hierarchical, or
744 phylogenetic structure. *Ecography* **40**, 913–929 (2017).
- 745 26. Valavi, R., Elith, J., Lahoz-Monfort, J. J. & Guillera-Aroita, G. blockCV: An R package for
746 generating spatially or environmentally separated folds for k-fold cross-validation of species
747 distribution models. *Methods in Ecology and Evolution* **10**, 225–232 (2019).
- 748 27. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *Proceedings of the*
749 *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–
750 794 (Association for Computing Machinery, 2016). doi:10.1145/2939672.2939785.
- 751 28. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V. & Gulin, A. CatBoost: unbiased
752 boosting with categorical features. Preprint at <https://doi.org/10.48550/arxiv.1706.09516> (2017).
- 753 29. Ke, G. *et al.* LightGBM: A Highly Efficient Gradient Boosting Decision Tree. in *Advances in*
754 *Neural Information Processing Systems* vol. 30 (2017).
- 755 30. Huber, P. J. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*
756 **35**, 73–101 (1964).
- 757 31. Charbonnier, P., Blanc-Féraud, L., Aubert, G. & Barlaud, M. Two deterministic half-quadratic
758 regularization algorithms for computed imaging. in *Proceedings of 1st International Conference*
759 *on Image Processing* vol. 2 168–172 (IEEE Computer Society Press, Austin, TX, USA, 1994).
- 760 32. Allen, S. T. *et al.* Global sinusoidal seasonality in precipitation isotopes. *Hydrology and Earth*
761 *System Sciences* **23**, 3423–3436 (2019).

33. Tang, Y. *et al.* Effects of changes in moisture source and the upstream rainout on stable isotopes in precipitation – a case study in Nanjing, eastern China. *Hydrology and Earth System Sciences* **19**, 4293–4306 (2015).
34. Galewsky, J. *et al.* Stable isotopes in atmospheric water vapor and applications to the hydrologic cycle. *Reviews of Geophysics* **54**, 809–865 (2016).
35. Bailey, A., Posmentier, E. & Feng, X. Patterns of Evaporation and Precipitation Drive Global Isotopic Changes in Atmospheric Moisture. *Geophysical Research Letters* **45**, 7093–7101 (2018).
36. Le Duy, N., Heidbüchel, I., Meyer, H., Merz, B. & Apel, H. What controls the stable isotope composition of precipitation in the Mekong Delta? A model-based statistical approach. *Hydrology and Earth System Sciences* **22**, 1239–1262 (2018).
37. Risi, C., Muller, C. & Blossey, P. Rain Evaporation, Snow Melt, and Entrainment at the Heart of Water Vapor Isotopic Variations in the Tropical Troposphere, According to Large-Eddy Simulations and a Two-Column Model. *Journal of Advances in Modeling Earth Systems* **13**, e2020MS002381 (2021).
38. Galewsky, J. *et al.* The Influence of Convective Aggregation on the Stable Isotopic Composition of Water Vapor. *AGU Advances* **4**, e2023AV000877 (2023).
39. McGuire, K. J. *et al.* The role of topography on catchment-scale water residence time. *Water Resources Research* **41**, W05002 (2005).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [GuoSFIsoRolesupplementarymaterialv2.pdf](#)