

Temporal eligibility rules alter country composition and ecoregion representation in global river water-quality monitoring networks

Andrea Mariel Malavé Gómez^{1*} , Luis Angel Bucheli Carpio¹ ,
Roger Freire Avilés¹ , Norma Valencia Castillo² ,
Sara Sofia Valencia Castillo³ , Fabiola Rosa Lopezdomínguez Rivas⁴ 

¹ Facultad de Ciencias de la Ingeniería, Universidad Estatal de Milagro, Milagro, Ecuador

² Facultad de Ciencias Agrarias, Universidad Agraria del Ecuador, Guayaquil, Ecuador

³ Carrera de gestión en talento humano, Tecnológico Universitario Argos, Guayaquil, Ecuador

⁴ Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, Guayaquil, Ecuador

* Corresponding author's e-mail: amalaveg@unemi.edu.ec

ABSTRACT

Temporal eligibility can reorganize a monitoring network in directions that station loss alone does not predict. This instability was characterized in Caravan-Qual Lite version 1.1 over 2000–2020 using six constituents, country labels, 426 freshwater ecoregions, seven realms, and Hill numbers $q=0, 1$, and 2 . The primary rule retained 14,387 electrical-conductivity stations (Cohort A) and a 4,373-station intersection of the independently eligible conductivity, phosphorus, and dissolved-oxygen networks (Cohort B); overlap in eligible years or sampling dates was not required. With minimum requirements of three, five, or ten nonconsecutive eligible calendar years, dissolved organic carbon was lost in 90.7% of stations, while ecoregion $q = 2$ increased from 3.909 to 6.594. Electrical conductivity and total phosphorus also gained $q = 2$ ecoregion representation at selected thresholds, whereas dissolved oxygen declined across countries, ecoregions, and realms. Nitrate and suspended solids realm $q = 2$ rose slightly from three to five years before falling to ten. Thus, comparable attrition did not imply a common geographic response. Equal-size references were treated as descriptive baselines rather than exchangeability tests. In Cohort B, ecoregion $q = 0$ was 10.8% below its equal-size median while $q = 2$ was 46.1% above it; the +0.5% within-country $q = 2$ residual showed that the evenness excess was largely reproduced after country counts were fixed, whereas the $q = 0$ deficit persisted. Requiring coverage in all four calendar quarters removed 40.1% of Cohort B stations and nine ecoregions and shifted country leverage. Temporal rules, therefore, changed country composition and FEOW representation in constituent- and order-specific ways that station retention alone could not summarize.

Keywords: freshwater ecology, monitoring-network design, temporal eligibility, country composition, Hill diversity, ecoregion representation.

INTRODUCTION

Water-quality monitoring networks balance information demand, station location, revisit frequency, constituent portfolio, and cost. Recent studies optimize sampling locations or evaluate network layouts against several decision criteria rather than maximizing station count alone (Alvarez-Vázquez et al., 2023; Li et al., 2024). Reviews of surface-water networks likewise treat

spatial coverage, temporal design, and information content as linked choices (Jiang et al., 2020; Strobl and Robillard, 2008).

Monitoring-network design also considers temporal feasibility, spatial information, redundancy, and cost (Alilou et al., 2019; Camara et al., 2020). Multiparameter systems make temporal coverage and constituent portfolios explicit choices (Syafudin et al., 2024). Temporal settings can change the evidence produced by a

network: coarser reporting can attenuate peaks, miss events, and reverse event rankings (Bastidas Sánchez, 2026). The present study addresses a different temporal decision – whether stations remain eligible after minimum-eligible-year and calendar-quarter-coverage rules are applied.

A harmonized archive introduces a second design stage after field deployment. Analysts define a retained monitoring network whenever they impose a minimum-eligible-year threshold, calendar-quarter-coverage rule, or constituent intersection. Station loss is automatic; its geographic consequence is not. The design question is whether different constituents and temporal thresholds reorganize country composition and FEOW representation in consistent or divergent directions.

This retrospective selection differs from preferential site placement, although the inferential risk is related. Preferential-sampling models evaluate whether monitoring locations depend on the environmental process, and changing-network methods address bias when sites enter or leave over time (Watson et al., 2019; Zidek et al., 2014). Here, the non-obvious outcomes are whether $q = 0$, $q = 1$, and $q = 2$ disagree, whether occupancy can fall while evenness rises, whether direction changes by constituent, and how much of the dominance pattern is reproduced after country counts are fixed. Treating deterministic archive filtering as a second network-design operation makes each outcome auditable before inference.

Global water-quality archives inherit unequal national capacity, monitoring mandates, and data-source histories after harmonization (Jones et al., 2024; Kirschke et al., 2020). Recent freshwater research has emphasized that large data volumes can coexist with incomplete or uneven representation of ecological settings (Harwell et al., 2024; Smits et al., 2025). Comparable diagnostics remain uncommon for post hoc eligibility rules, even though the retained observations may support restoration appraisal, pollution-control design, or model transfer.

River assessments depend on whether observations span the ecological settings across which stressor responses and management outcomes vary (Su et al., 2021). Freshwater biodiversity risk is both severe and unevenly observed (Reid et al., 2019; Sayer et al., 2025). Monitoring standards increasingly emphasize comparable biological and biogeographic coverage (Birk et al., 2020; Simaika et al., 2024). Restoration and

basin-management evaluations likewise require monitoring evidence that matches the intended spatial domain (Odnorih et al., 2020; Palmer et al., 2005). A numerically large but compositionally concentrated network can therefore create a false impression of broad ecoregion representation.

The contribution is not to show that stricter rules remove stations or that record abundance is geographically clustered. It is to identify directional instability: whether constituent, minimum eligible-year threshold, and Hill order produce responses that station loss alone cannot anticipate, and how much of the dominance pattern is reproduced after country counts are fixed. Equal-size departures are retained as descriptive context, not presented as proof of selection beyond monitoring effort.

Caravan-Qual integrates stream-water-quality observations with large-sample hydrology (Jones et al., 2026a; Kratzert et al., 2023). GRQA illustrates the value and inherited provenance of related global products, although no cross-archive station match is attempted here (Virro et al., 2021). Freshwater Ecoregions of the World (FEOW) provide the ecological geography (Abell et al., 2008). Hill numbers separate occupied categories ($q=0$), Shannon-weighted common and uncommon categories ($q=1$), and dominance-sensitive effective number ($q=2$) (Hill, 1973; Jost, 2006).

We asked four questions: whether temporal rules produced a common or directionally unstable response across six constituents, three diversity orders, and two observation windows; whether FEOW occupancy could fall while dominance-sensitive balance rose; how much FEOW variation remained after country composition was fixed; and what a stricter calendar-quarter rule implied for a worked monitoring-network decision. The estimand is the fixed retained network in a documented release, not a hypothetical population of stations.

The interpretation had explicit failure conditions. Directional instability would fail if constituent–hierarchy profiles moved monotonically in a common direction as eligibility tightened. The country-composition interpretation would fail if within-country $q=2$ deviations remained comparable to equal-size deviations despite informative within-country variation. The worked design consequence would fail if the stricter calendar-quarter rule changed station count but left FEOW occupancy, balance and country leverage unchanged.

MATERIALS AND METHODS

The primary source was Caravan-Qual Lite version 1.1, Zenodo record 21337532 (<https://doi.org/10.5281/zenodo.21337532>; accessed 20 August 2026) (Jones et al., 2026c). Six members were streamed directly from the official ZIP: electrical conductivity (EC), total phosphorus (TP), dissolved oxygen (DO), nitrate as nitrogen (NO₃-N), total suspended solids (TSS), and dissolved organic carbon (DOC). Full source and member SHA-256 hashes are reported in Table S3. Country metadata provenance was checked against the public construction code (Jones et al., 2026b).

The primary window was 1 January 2000 through 31 December 2020. It was selected as a common 21-year contemporary horizon with broad cross-constituent support, so that the minimum requirements of three, five, or ten eligible calendar years were evaluated against the same opportunity period rather than unequal historical availability. The analysis did not assume that the archive's most recent years were complete; source-specific reporting lags and right-edge truncation could affect 2021–2025. The 2000–2020 interval is therefore a worked design horizon, not a regulatory standard, and every primary result was compared with a 1980–2025 sensitivity window rather than a presumed complete release window.

Published numeric values flagged '<' were retained as released. Because eligibility depends on sampling dates rather than concentrations, the archive's censoring convention does not affect any result reported here; variable counts appear in Table S2. Values flagged '*' were excluded from eligibility because release-identified anomalies could create station-years, and an inclusion sensitivity was retained. The 19 finite rows carrying a different nonblank flag (17 EC, one DO, and one TSS) were retained because the release did not designate them as outliers; their dates and values entered the same eligibility and daily-median processing. Multiple finite station-date values were collapsed to their median.

The primary eligible year contained at least four distinct sampling days in at least three calendar quarters, and a station required a minimum of five eligible calendar years within the observation window. Eligible years did not have to be consecutive. This worked rule demands a calendar-quarter spread without requiring all four quarters and permits one within-quarter repeat; it

was fixed before geographic comparisons and is not proposed as universal. Sensitivity to the minimum number of eligible years used thresholds of 3, 5, and 10 for all six constituents. Within-year sensitivity held the five-year minimum fixed and evaluated 3d/3q, 4d/2q, 4d/3q, 4d/4q and 6d/3q (constituent counts in Table S13). Cohort A was the primary EC network. Cohort B was the station-level intersection of the independently eligible EC, TP, and DO networks; it established the availability of all three constituents at the same stations but did not require overlap among their eligible years or sampling dates. This intersection spans ionic condition, nutrients, and oxygen while retaining a usable sample, but it was not treated as an exhaustive or optimal constituent portfolio. NO₃-N, TSS and DOC remained in the complete minimum-eligible-year audit, and alternative intersections are reported in Table S11, so the EC–TP–DO station intersection does not define the general conclusion.

Country labels were taken from version 1.1 station metadata; they originate from a boundary overlay and are not agency or monitoring-program identifiers. All inferences therefore use 'country composition' rather than 'national program'. Ecoregions were assigned independently by point overlay against the FEOW first-edition boundary update aligned to HydroBASINS in January 2013 (GIS_hs_snapped; https://www.feow.org/files/downloads/GIS_hs_snapped.zip; accessed 20 August 2026). Station longitude/latitude and FEOW polygon coordinates were handled as WGS 84 geographic coordinates in decimal degrees (EPSG:4326), with no reprojection. The fixed reference contained 426 ecoregions in Afrotropic, Australasia, Indo-Malay, Nearctic, Neotropical, Oceania, and Palearctic realms. FEOW has no Antarctic freshwater realm in this geometry.

For countries, ecoregions, and realms, Hill $q=0$ richness, $q=1$ exponential-Shannon diversity and $q=2$ inverse-Simpson effective diversity were calculated. Unassigned stations remained in station and country totals but were excluded from FEOW profiles; their counts were reported for every network. These are exact descriptors of one release, window, and rule, so no station-bootstrap confidence intervals were attached. For station counts, $q=2$ is the reciprocal of the Herfindahl concentration index over category shares; the Hill formulation places $q=0$, $q=1$, and $q=2$ on a common effective-number scale. Throughout, FEOW representation denotes the distribution of retained

stations among FEOW ecoregions and realms; it does not measure ecological condition, upstream exposure, river-network density, biodiversity, or ecological representativeness in a causal sense.

The equal-size reference drew 10,000 subsets without replacement from the corresponding constituent or intersection pool, always matching the retained station count (master seed 20260810). Station pools were ordered canonically by `wqms_id`; analysis-by-hierarchy random substreams were derived from the master seed and canonical labels, and category vectors were lexically ordered. This makes results independent of set, merge, and analysis-call order. The observed value, median, 2.5th–97.5th percentile range and exact number of draws at or below the observation were reported. Endpoint values such as 0/10,000 and 10,000/10,000 mean that none or all of the simulated draws were at or below the observation; they are empirical ranks, not exact p-values of zero or one. The equal-size reference assumes exchangeable stations and is not conditioned on station record abundance. Because eligibility is generated by the minimum number of eligible years and calendar-quarter sampling, while monitoring effort is clustered by country, agency, and basin, departure from this reference is expected a priori. Its medians, ranges, and ranks are used only as descriptive context; they cannot establish selection beyond monitoring effort. Therefore, interpretation emphasizes magnitude, direction, and changes across constituents and q rather than rank alone.

The within-country reference preserved the observed station count in every country and resampled within those strata (10,000 draws; master seed 20260811) using the same canonical ordering and named-substream rule. It was applied to 28 analysis configurations, corresponding to 27 distinct networks, because EC at five years is identical to the primary Cohort A 4d/3q network and was simulated once. No fixed percentage threshold was applied. Each q profile was interpreted from its reference range and empirical rank.

A separate feasible-range audit was performed only for within-country $q=2$ in the two primary networks. At fixed country totals and fixed FEOW-assigned station count, maximizing $q=2$ is equivalent to minimizing $\sum_j n_j^2$. For ecological unit j , binary marginal variables y_{jk} represented the k th selected station, with $n_j = \sum_k y_{jk}$ and $n_j^2 = \sum_k (2k-1)y_{jk}$; this converts the quadratic concentration objective to a mixed-integer linear

program. SciPy `scipy.optimize.milp` (HiGHS) solved the exact maximum-evenness assignment with relative MIP gap 0.0. Deterministic coordinate ascent supplied a concentrated feasible assignment, not a certified global minimum. These calculations show what departures can occur in Cohorts A and B; they do not establish feasible ranges for the other 25 networks, particularly smaller networks with less freedom.

It was also measured how much ecological variation remained available within countries. The candidate pool was restricted to countries represented in the retained network. Using natural logarithms and station weights, $H(E|C) = \sum_c p_c H(E|C=c)$, conditional $q_1 = \exp[H(E|C)]$ and association $= 1 - H(E|C)/H(E)$. All represented strata entered the entropy calculation. Sampling slack means that a country's retained target is smaller than its candidate-pool count; a stratum is informative when it has slack and more than one assigned ecological category. Category degrees of freedom equal $\sum_c (K_c - 1)$ across informative strata. These quantities distinguish a small residual under limited within-country variation from one observed when substantial variation was available. This information-theoretic diagnostic is consistent with the use of entropy measures in water-monitoring network design (Keum et al., 2017).

Country influence was assessed systematically by removing every represented country from every distinct network. The United Kingdom and Mexico decompositions were then reported because the leave-one-country-out ranking identified exceptional influence in the worked EC and Cohort B cases. Area-proportional FEOW references, VPU 605 exposure, outlier inclusion, and proximity-based duplicate collapse were retained as scale and processing checks. The VPU 605 audit used the GEOGLOWS v2 package-metadata snapshot (<http://geoglows-v2.s3-website-us-west-2.amazonaws.com/tables/package-metadata-table.parquet>; accessed 20 August 2026) only for a tabular LINKNO–VPUCODE join; no additional coordinate reference system was involved. The complete protocol and workflow are given in Table S1 and Figure S1; the equal-size results for every network are in Table S4.

The primary and four-quarter Cohort B rules were compared using station retention, $q=0/1/2$ FEOW representation, country leverage, and within-country residual profiles. This worked comparison organizes design consequences; it does not optimize a new field network.

All analyses were run on Linux 6.18.35 with Python 3.12.13, NumPy 2.3.5, pandas 2.2.3, SciPy 1.17.0 (HiGHS 1.8.0 backend), Matplotlib 3.10.8, scikit-learn 1.8.0, and Pillow 12.3.0. FEOW geometry was read by the script's deterministic shapefile/DBF parser; GeoPandas was not used.

RESULTS

The six-file union contained 77,258 stations during 2000–2020. After excluding outlier-only records, the candidate pools were 56,816 EC, 38,888 TP, 48,817 DO, 45,403 NO₃-N, 32,627 TSS and 23,009 DOC stations. Cohort A retained 14,387 stations; the station-level EC–TP–DO intersection (Cohort B), a subset of Cohort A, retained 4,373. FEOW assignment was available for 14,353 Cohort A stations (99.76%) and 4,360 Cohort B stations (99.70%).

Cohort A occupied 48 countries and all seven realms, while Cohort B occupied 36 countries and four realms (Table 1; Figure 1). The unfiltered EC–TP–DO intersection had already been restricted to five realms, so most of the realm limitation was inherited from the underlying station-level constituent intersection; temporal eligibility removed one further realm. The retained four were Indo-Malay, Nearctic, Neotropic, and Palearctic. Within-realm FEOW representation is reported in Table S10, and station-level chemical summaries in Table S12.

All six constituents lost stations and effective countries as the minimum eligible-year requirement increased, but FEOW trajectories differed (Table 2; Figure 2). EC and TP gained $q=2$ ecoregion representation between three and ten years; EC realms peaked at five years. DO declined at all three $q=2$ hierarchies. NO₃-N and TSS declined for countries and ecoregions, but their realm $q=2$ rose marginally between three and five years before falling at ten. DOC lost 90.7% of stations but increased $q=2$ ecoregions from 3.909 to 6.594.

Therefore, station attrition did not predict either the sign or the scale of the geographic response.

For descriptive context, the primary equal-size comparison changed interpretation across diversity orders (Table 3; Figure 3). Cohort A ecoregions were below the equal-size median by 23.5%, 37.0%, and 48.3% for $q=0$, 1, and 2. Cohort B ecoregions had fewer occupied units (−10.8%; rank 0/10,000) but higher $q=1$ (+24.7%) and $q=2$ (+46.1%), meaning fewer ecoregions with a more even distribution among those retained. Cohort B realm $q=1$ and $q=2$ were 19.8% and 23.4% above their medians. Realm $q=0$ was four rather than the median of five, but its rank of 1.631/10,000 placed it within the equal-size reference; it was therefore a discrete, structurally limited contrast rather than evidence of unusual realm loss. These exchangeability-based contrasts are descriptive baselines, not tests of selection beyond monitoring effort. A $q=2$ -only summary would nevertheless have hidden the ecoregion occupancy deficit present in the retained network.

The $q=2$ equal-size magnitudes reinforced the constituent dependence (Table 4). EC ecoregions were 43.0–49.6% below the reference at all minimum eligible-year thresholds. TP changed from an ecoregion deficit at three years to an excess at ten. DO, NO₃-N, and TSS produced progressively concentrated country and ecoregion distributions, while DOC produced a late ecoregion excess.

Within-year rules produced a second source of divergence. Requiring coverage in all four calendar quarters retained 2,621 rather than 4,373 Cohort B stations, a loss of 1,752 stations (40.1%). It increased country $q=2$ from 4.460 to 4.902, but reduced ecoregion $q=2$ from 19.066 to 9.722 and realm $q=2$ from 2.904 to 2.334. It also removed nine occupied ecoregions while all four retained realms remained present, shifted the leading country from Mexico to the United States and increased the leading-country share by 6.6 percentage points (Table 5).

Removing the United Kingdom changed the complete EC minimum-eligible-year trajectory

Table 1. Hill profiles and country dominance in the primary networks

Cohort	Stations	Countries $q0 / q1 / q2$	Ecoregions $q0 / q1 / q2$	Realms $q0 / q1 / q2$	Leading country
Cohort A	14,387	48 / 8.318 / 4.839	143 / 24.217 / 5.541	7 / 3.541 / 2.706	United Kingdom (38.5%)
Cohort B	4,373	36 / 7.583 / 4.460	91 / 39.665 / 19.066	4 / 3.158 / 2.904	Mexico (34.4%)

Note: $q=0$ is richness; $q=1$ is exponential-Shannon diversity; $q=2$ is inverse-Simpson effective diversity. FEOW profiles use assigned stations.

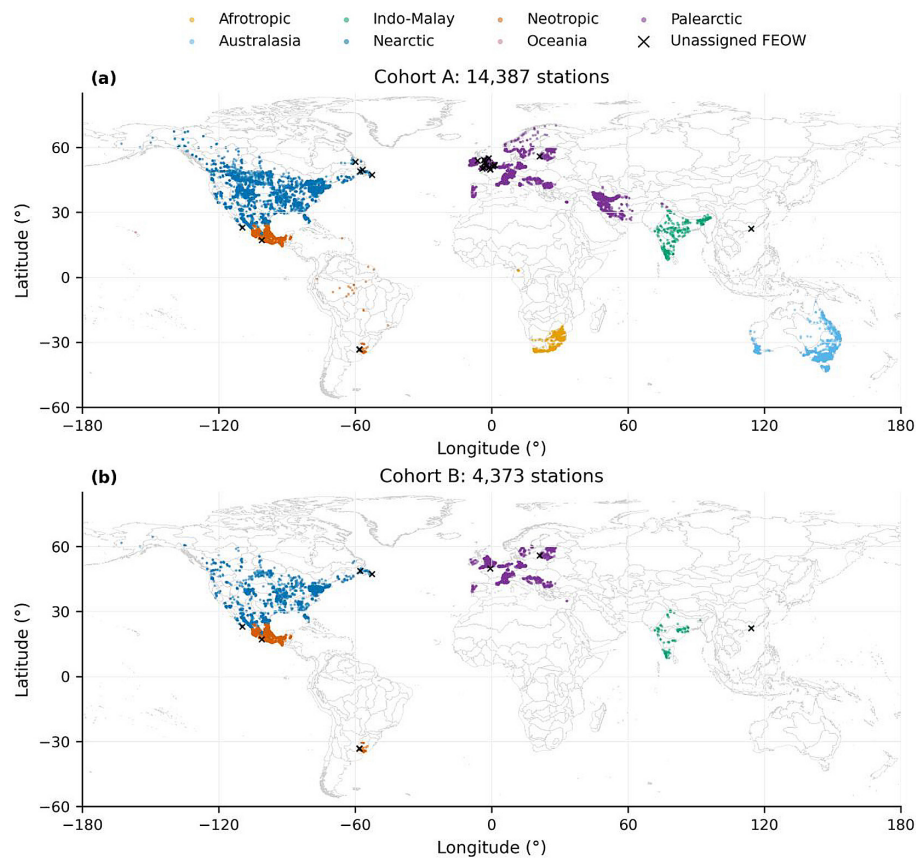


Figure 1. Geographic distribution of (a) Cohort A and (b) Cohort B. Points are colored by FEOW realm. Black crosses, identified in the legend as unassigned FEOW, mark the 34 Cohort A and 13 Cohort B stations without FEOW assignment

Table 2. Six-constituent sensitivity to minimum number of eligible calendar years

Variable	Min. years	Stations	Countries q0 / q1 / q2	Ecoregions q0 / q1 / q2	Realms q0 / q1 / q2	Leading country
EC	3	20,460	53 / 9.761 / 5.682	154 / 23.417 / 5.403	7 / 3.222 / 2.410	United Kingdom (34.7%)
EC	5	14,387	48 / 8.318 / 4.839	143 / 24.217 / 5.541	7 / 3.541 / 2.706	United Kingdom (38.5%)
EC	10	5,328	19 / 4.985 / 3.821	91 / 20.198 / 6.108	4 / 3.006 / 2.605	United Kingdom (38.9%)
TP	3	14,088	49 / 14.134 / 9.717	122 / 28.473 / 8.500	7 / 3.408 / 2.724	United States (17.8%)
TP	5	9,554	47 / 12.581 / 8.031	115 / 34.570 / 11.815	7 / 3.670 / 3.028	United States (20.0%)
TP	10	2,163	20 / 6.263 / 3.861	73 / 31.010 / 16.589	6 / 3.271 / 2.491	United States (45.9%)
DO	3	18,775	53 / 7.696 / 4.185	128 / 10.713 / 2.865	5 / 2.502 / 1.902	United Kingdom (44.9%)
DO	5	13,594	47 / 5.720 / 2.978	118 / 8.871 / 2.533	5 / 2.292 / 1.807	United Kingdom (54.9%)
DO	10	6,560	19 / 2.411 / 1.682	69 / 4.190 / 1.750	4 / 1.865 / 1.539	United Kingdom (75.2%)
NO3-N	3	18,635	52 / 7.871 / 4.198	122 / 8.530 / 2.554	7 / 2.151 / 1.639	United Kingdom (44.4%)
NO3-N	5	14,633	50 / 6.461 / 3.462	111 / 7.768 / 2.412	6 / 2.192 / 1.665	United Kingdom (49.9%)
NO3-N	10	6,144	22 / 2.401 / 1.571	66 / 3.430 / 1.583	5 / 1.722 / 1.399	United Kingdom (78.8%)
TSS	3	14,765	51 / 9.831 / 5.516	111 / 13.374 / 3.876	7 / 2.765 / 2.045	United Kingdom (37.0%)
TSS	5	11,488	46 / 8.399 / 4.948	98 / 12.313 / 3.605	6 / 2.792 / 2.054	United Kingdom (38.7%)
TSS	10	3,786	25 / 3.504 / 2.015	50 / 5.131 / 2.056	5 / 2.225 / 1.661	United Kingdom (68.9%)
DOC	3	8,735	40 / 7.689 / 5.599	96 / 12.385 / 3.909	6 / 2.482 / 2.078	France (27.2%)
DOC	5	6,410	35 / 7.051 / 5.156	92 / 12.750 / 4.212	6 / 2.617 / 2.234	France (29.7%)
DOC	10	812	12 / 5.259 / 4.065	46 / 15.033 / 6.594	3 / 2.173 / 2.016	United States (38.4%)

Note: Annual rule: ≥ 4 days in ≥ 3 calendar quarters. All six variables were analyzed under identical filters.

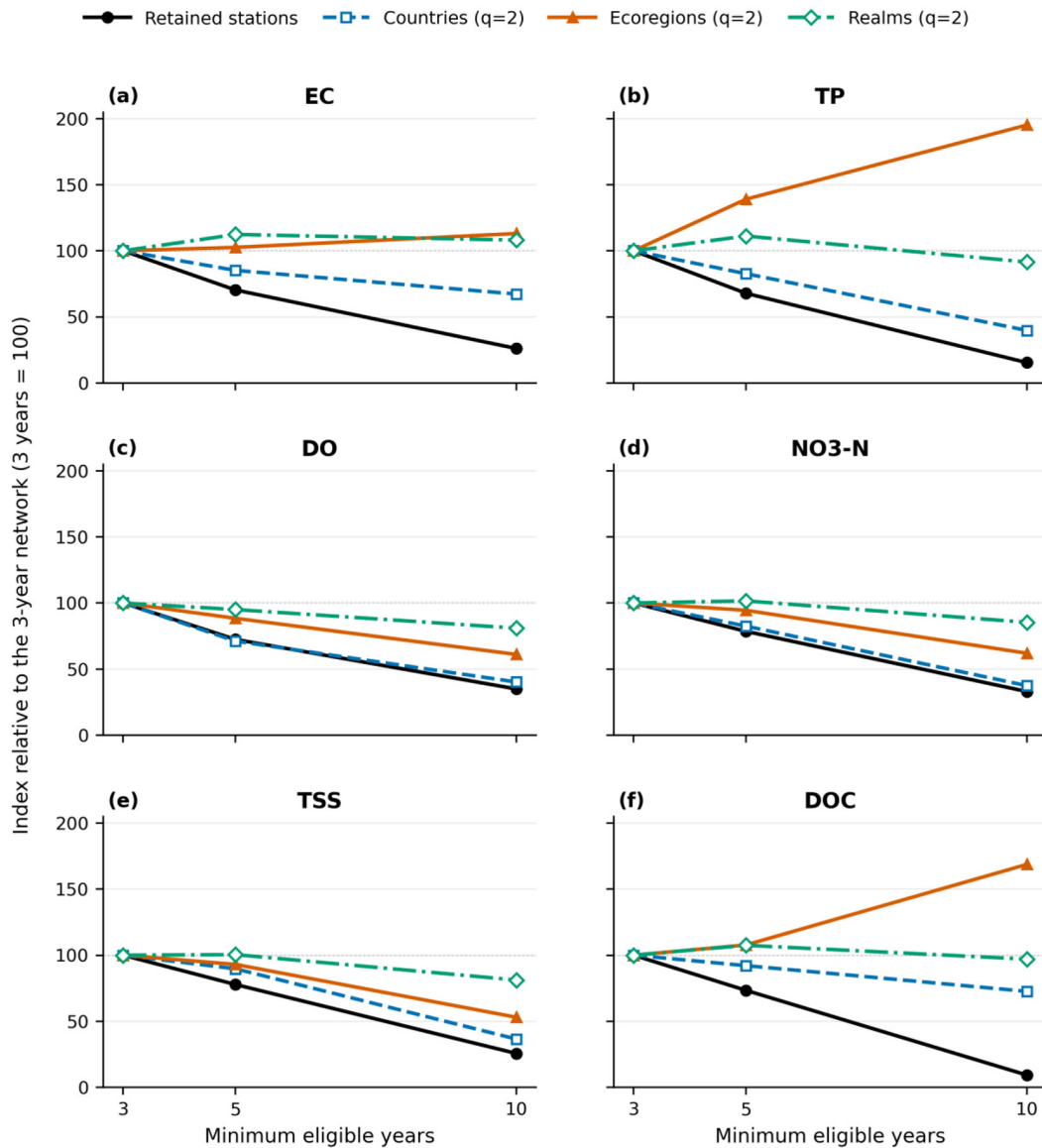


Figure 2. Trajectories across the minimum number of eligible years for all six constituents. Each line is indexed to the corresponding three-year value; all panels use the same vertical scale, separating station attrition from changes in $q=2$ effective diversity.

rather than eliminating all contrasts (Table 6). Without the United Kingdom, ecoregion $q=2$ followed $30.265 \rightarrow 46.499 \rightarrow 30.167$. From three to five years, country $q=2$ declined while ecoregion and realm $q=2$ increased; from five to ten years, all three declined. Thus, only the three-to-ten-year endpoint contrast disappeared.

The five-year peak was reproducible. Of 5,538 United Kingdom stations, 5,353 were in FEOW 404, 170 in FEOW 402, and 15 were unassigned. Removal reduced FEOW 404 from 41.7% to 7.1% of assigned stations and increased ecoregion $q=2$ 8.4-fold. The remaining network was most even at five years: its largest ecoregion share was 7.1%, compared with 13.2% at three

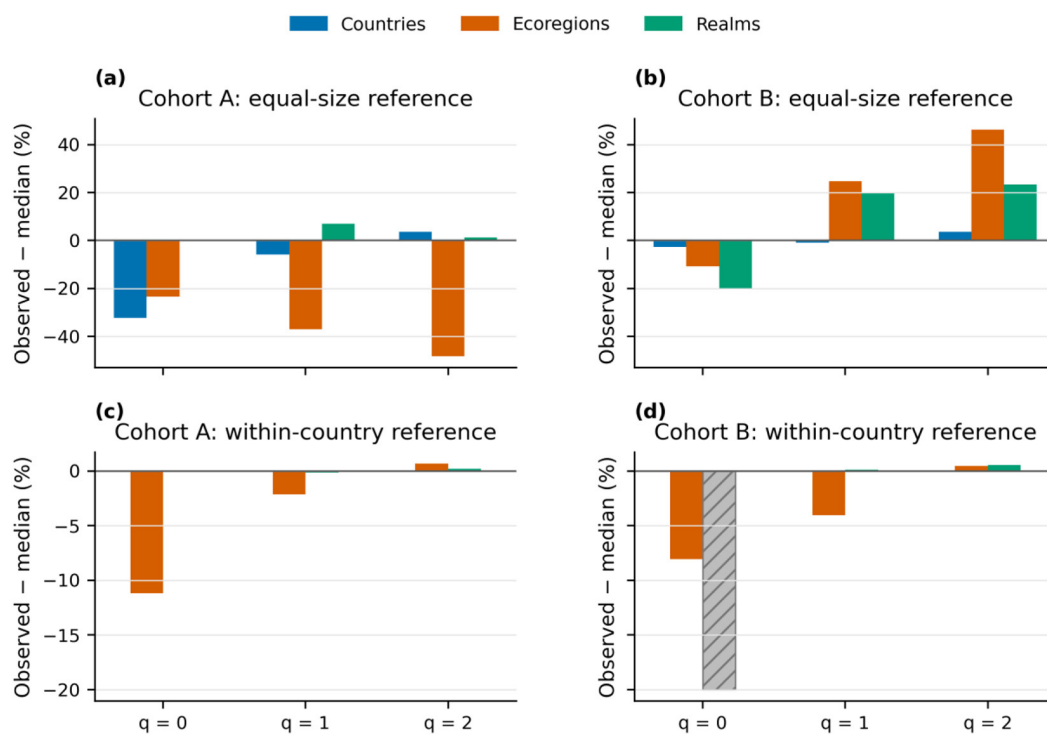
years and 10.2% at ten. All 143 five-year ecoregions remained occupied because FEOW 402 was shared with Ireland and FEOW 404 with ten continental labels; at ten years, removal did reduce richness from 91 to 90 (full FEOW decomposition in Table S19).

Mexico supplied 1,506 Cohort B stations under the primary rule across 25 ecoregions ($q_1=15.858$; $q_2=12.913$), but only 166 stations across 16 ecoregions under four-quarter coverage ($q_1=11.584$; $q_2=9.373$). Its network share fell from 34.4% to 6.3%; the United States became the leading country at 41.0%, so the leading share increased by 6.6 percentage points (Table 7; Figure 4; FEOW decomposition in Table S18). The

Table 3. Equal-size Hill-profile references for the primary networks

Cohort	Hierarchy	q	Pool / retained	Observed	Null median [95% range]	%Δ / contrast	r≤obs
Cohort A	Countries	0	56,816 / 14,387	48	71 [65–76]	-32.4%	0/10000
Cohort A	Countries	1	56,816 / 14,387	8.318	8.842 [8.662–9.028]	-5.9%	0/10000
Cohort A	Countries	2	56,816 / 14,387	4.839	4.675 [4.576–4.779]	+3.5%	9992/10000
Cohort A	Ecoregions	0	56,816 / 14,387	143	187 [181–192]	-23.5%	0/10000
Cohort A	Ecoregions	1	56,816 / 14,387	24.217	38.443 [37.468–39.430]	-37.0%	0/10000
Cohort A	Ecoregions	2	56,816 / 14,387	5.541	10.713 [10.329–11.118]	-48.3%	0/10000
Cohort A	Realms	0	56,816 / 14,387	7†	7 [7–7]	+0.0%	10000/10000
Cohort A	Realms	1	56,816 / 14,387	3.541	3.313 [3.271–3.353]	+6.9%	10000/10000
Cohort A	Realms	2	56,816 / 14,387	2.706	2.674 [2.646–2.701]	+1.2%	9910/10000
Cohort B	Countries	0	21,524 / 4.373	36	37 [34–41]	-2.7%	2949/10000
Cohort B	Countries	1	21,524 / 4.373	7.583	7.660 [7.385–7.941]	-1.0%	2982/10000
Cohort B	Countries	2	21,524 / 4.373	4.460	4.307 [4.133–4.495]	+3.5%	9471/10000
Cohort B	Ecoregions	0	21,524 / 4.373	91	102 [98–106]	-10.8%	0/10000
Cohort B	Ecoregions	1	21,524 / 4.373	39.665	31.818 [30.574–33.066]	+24.7%	10000/10000
Cohort B	Ecoregions	2	21,524 / 4.373	19.066	13.046 [12.260–13.905]	+46.1%	10000/10000
Cohort B	Realms	0	21,524 / 4.373	4†	5 [4–5]	4 vs 5†	1631/10000
Cohort B	Realms	1	21,524 / 4.373	3.158	2.635 [2.590–2.682]	+19.8%	10000/10000
Cohort B	Realms	2	21,524 / 4.373	2.904	2.353 [2.319–2.387]	+23.4%	10000/10000

Note: 10,000 draws. q=0 values are counts; q=1 and q=2 are shown to three decimals; r≤obs is the exact empirical rank among the 10,000 draws and is not an exact p-value; 0/10,000 and 10,000/10,000 do not mean probabilities of exactly zero or one. † Realm q=0 is structurally saturated: Cohort A is fixed at seven realms and Cohort B spans only four or five; the latter is shown as 4 vs 5 rather than a percentage.

**Figure 3.** Primary-network Hill profiles relative to equal-size and within-country medians.

Panels (a) and (b) share one vertical scale, as do panels (c) and (d). Country counts are fixed by design in the within-country reference and are therefore omitted from panels (c) and (d). In panel (d), the hatched Cohort B realm q=0 bar is the within-country one-unit contrast (rank 2.994/10,000); the corresponding equal-size contrast in panel (b) has rank 1.631/10,000

Table 4. q=2 departures from equal-size medians across minimum eligible-year thresholds

Variable	Countries q2 %Δ (3 / 5 / 10 years)	Ecoregions q2 %Δ (3 / 5 / 10 years)	Realms q2 %Δ (3 / 5 / 10 years)
EC	+21.5% / +3.5% / -18.3%	-49.6% / -48.3% / -43.0%	-9.9% / +1.2% / -2.5%
TP	+53.0% / +26.5% / -39.2%	-24.5% / +5.0% / +47.9%	+4.5% / +16.2% / -4.4%
DO	-19.6% / -42.8% / -67.7%	-48.4% / -54.4% / -68.5%	-15.6% / -19.8% / -31.7%
NO3-N	-36.8% / -47.8% / -76.3%	-44.6% / -47.7% / -65.7%	-21.4% / -20.2% / -32.9%
TSS	-12.2% / -21.2% / -67.9%	-15.2% / -21.1% / -55.0%	-3.4% / -3.0% / -21.6%
DOC	+13.3% / +4.4% / -17.4%	-37.5% / -32.7% / +5.8%	-6.5% / +0.5% / -9.3%

Note: Values are ordered as 3 / 5 / 10 eligible years. Full observed values, reference ranges and empirical ranks are reported in Table S4.

Table 5. Sensitivity to five within-year rules

Annual rule	Cohort	Stations	Countries n / q2	Ecoregions n / q2	Realms n / q2	Leading country
≥3 days in ≥3 quarters	Cohort A	14,773	48 / 4.869	149 / 5.597	7 / 2.711	United Kingdom (38.3%)
≥3 days in ≥3 quarters	Cohort B	4,506	36 / 4.530	92 / 19.512	4 / 2.903	Mexico (33.9%)
≥4 days in ≥2 quarters	Cohort A	15,096	48 / 4.862	144 / 5.655	7 / 2.719	United Kingdom (38.3%)
≥4 days in ≥2 quarters	Cohort B	4,551	36 / 4.427	91 / 19.272	4 / 2.911	Mexico (34.8%)
≥4 days in ≥3 quarters*	Cohort A	14,387	48 / 4.839	143 / 5.541	7 / 2.706	United Kingdom (38.5%)
≥4 days in ≥3 quarters*	Cohort B	4,373	36 / 4.460	91 / 19.066	4 / 2.904	Mexico (34.4%)
≥4 days in 4 quarters	Cohort A	11,055	43 / 4.036	130 / 4.106	7 / 2.231	United Kingdom (44.4%)
≥4 days in 4 quarters	Cohort B	2,621	33 / 4.902	82 / 9.722	4 / 2.334	United States (41.0%)
≥6 days in ≥3 quarters	Cohort A	11,550	46 / 5.194	136 / 5.878	7 / 2.816	United Kingdom (36.1%)
≥6 days in ≥3 quarters	Cohort B	3,218	33 / 4.904	85 / 13.658	4 / 2.858	United States (32.1%)

Note: All scenarios require ≥5 eligible years; the asterisk marks the primary rule. Quarter counts refer to calendar quarters.

Table 6. United Kingdom influence across EC minimum-eligible-year thresholds

Years	Configuration	Stations	Countries n / q2	Ecoregions n / q2	Realms n / q2	Leading country
3	Full	20,460	53 / 5.682	154 / 5.403	7 / 2.410	United Kingdom (34.7%)
3	Without United Kingdom	13,362	52 / 7.666	154 / 30.265	7 / 3.366	United States (27.4%)
5	Full	14,387	48 / 4.839	143 / 5.541	7 / 2.706	United Kingdom (38.5%)
5	Without United Kingdom	8,849	47 / 6.467	143 / 46.499	7 / 3.628	United States (30.7%)
10	Full	5,328	19 / 3.821	91 / 6.108	4 / 2.605	United Kingdom (38.9%)
10	Without United Kingdom	3,255	18 / 3.382	90 / 30.167	4 / 2.743	United States (47.3%)

Note: Full and country-removal rows are paired at each threshold.

leave-one-country-out analysis placed Mexico among the strongest Cohort B influences, but not alone: removing Mexico reduced ecoregion q=2 by 47.6%, removing the United States reduced it by 45.2%, and removing the United Kingdom increased it by 54.3% (Table S14).

Within-country resampling largely reproduced the q=2 profiles after country counts were fixed without eliminating FEOW occupancy and q1 deficits (Table 8). In Cohort A, ecoregion residuals were -11.2%, -2.1%, and +0.7% for q=0,

1, and 2. In primary Cohort B, they were -8.1%, -4.1%, and +0.5%; under four-quarter coverage, they were -13.7%, -8.5%, and +1.2%. Several q=0 and q=1 observations lie outside all 10,000 draws. The detailed medians, ranges, and ranks for all 27 networks are retained in Table S5 and the machine-readable CSV.

Available within-country variation depended on scale (Table 9). For the two primary networks, normalized country–ecoregion association was 0.45–0.48, and conditional q1 diversity was

Table 7. Mexico decomposition across Cohort B within-year filters

Rule	Configuration	Stations	Countries n / q2	Ecoregions n / q2	Realms n / q2	Leading country
3d/3q	Full	4.506	36 / 4.530	92 / 19.512	4 / 2.903	Mexico (33.9%)
3d/3q	Without Mexico	2.977	35 / 4.132	77 / 10.368	4 / 2.245	United States (46.0%)
4d/2q	Full	4.551	36 / 4.427	91 / 19.272	4 / 2.911	Mexico (34.8%)
4d/2q	Without Mexico	2.969	35 / 4.051	76 / 10.095	4 / 2.239	United States (46.3%)
4d/3q	Full	4.373	36 / 4.460	91 / 19.066	4 / 2.904	Mexico (34.4%)
4d/3q	Without Mexico	2.867	35 / 4.070	76 / 9.995	4 / 2.229	United States (46.3%)
4d/4q	Full	2.621	33 / 4.902	82 / 9.722	4 / 2.334	United States (41.0%)
4d/4q	Without Mexico	2.455	32 / 4.387	68 / 8.571	4 / 2.185	United States (43.7%)
6d/3q	Full	3.218	33 / 4.904	85 / 13.658	4 / 2.858	United States (32.1%)
6d/3q	Without Mexico	2.301	32 / 4.167	68 / 7.656	4 / 2.193	United States (44.9%)

Note: The systematic leave-one-country-out ranking is reported in Table S14.

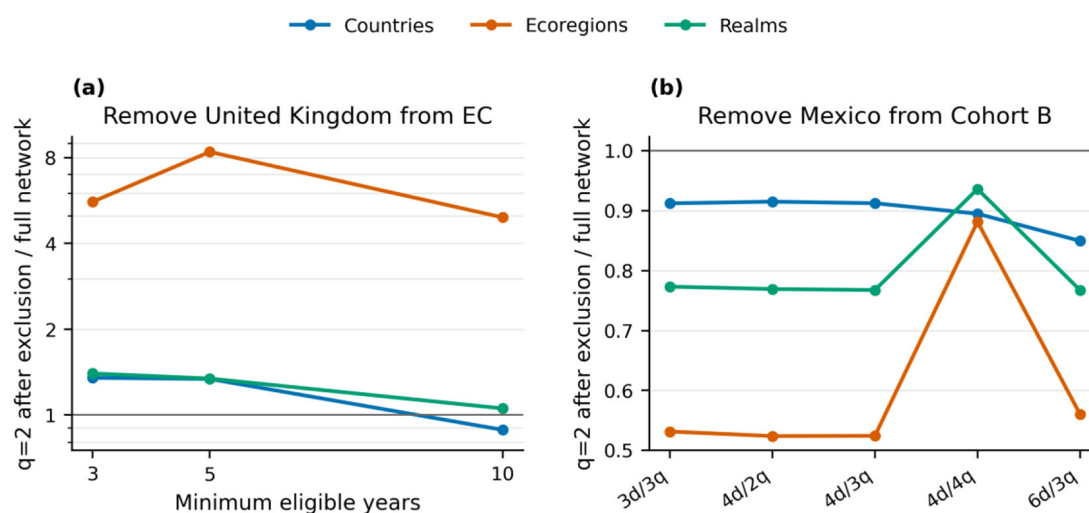


Figure 4. Country influence in the worked networks, expressed as the q=2 ratio after excluding the focal country to the full network. (a) United Kingdom removal from EC. (b) Mexico removal from Cohort B. Panel (a) uses a logarithmic vertical scale; panel (b) uses a linear scale. A ratio of one indicates no change

6.53–6.66 ecoregions; 94.8–96.7% of retained stations occurred in informative country strata. The ecoregion comparison, therefore, retained meaningful within-country freedom. Country–realm association was 0.93–0.96, and conditional q1 was only 1.04–1.07 realms, so country nearly determined realm membership, and the realm comparison had little freedom.

The within-country q=2 feasible-range audit showed broad departures in the two primary ecoregion networks (Table S20). For Cohort A, a feasible concentrated assignment reached –18.8% and the exact maximum-evenness assignment reached +14.5% relative to the within-country median; Cohort B reached –55.8% and +51.9%. Realm freedom was much narrower: the corresponding demonstrated ranges were –3.1%

to +3.3% and –6.4% to +7.5%. These bounds apply only to the two audited primary networks. No comparable claim is made for the other 25 networks, including smaller configurations such as DOC at ten years.

The 1980–2025 sensitivity window increased Cohort A to 20,681 and Cohort B to 6,540 stations. Equal-size q=2 magnitudes changed substantially – Cohort A ecoregions –42.5% and Cohort B ecoregions +81.4% – while within-country q=2 residuals remained +0.9% and –3.1%. Thus, the selected window affected magnitude but did not restore neutrality (Table S15). FEOW non-assignment never exceeded 0.37% across the 27 primary-window networks (Table S16).

Input-integrity checks confirmed that all six official archive members were streamed

Table 8. Selected within-country Hill profiles

Network	Hierarchy	q	Observed	Median [95% range]	Δ / contrast	Range width	r _{sobs}
Cohort A primary	Ecoregions	0	143	161 [157–165]	-11.2%	5.0%	0/10000
Cohort A primary	Ecoregions	1	24.217	24.742 [24.487–25.002]	-2.1%	2.1%	0/10000
Cohort A primary	Ecoregions	2	5.541	5.503 [5.461–5.546]	+0.7%	1.5%	9612/10000
Cohort A primary	Realms	0	7†	7 [7–7]	+0.0%	0.0%	10000/10000
Cohort A primary	Realms	1	3.541	3.547 [3.535–3.560]	-0.2%	0.7%	1672/10000
Cohort A primary	Realms	2	2.706	2.701 [2.696–2.707]	+0.2%	0.4%	9506/10000
Cohort B primary	Ecoregions	0	91	99 [96–103]	-8.1%	7.1%	0/10000
Cohort B primary	Ecoregions	1	39.665	41.341 [40.577–42.113]	-4.1%	3.7%	0/10000
Cohort B primary	Ecoregions	2	19.066	18.976 [18.508–19.456]	+0.5%	5.0%	6497/10000
Cohort B primary	Realms	0	4†	5 [4–5]	4 vs 5†	20.0%	2994/10000
Cohort B primary	Realms	1	3.158	3.155 [3.141–3.170]	+0.1%	0.9%	7005/10000
Cohort B primary	Realms	2	2.904	2.889 [2.872–2.905]	+0.5%	1.2%	9628/10000
Cohort B four-quarter	Ecoregions	0	82	95 [90–99]	-13.7%	9.5%	0/10000
Cohort B four-quarter	Ecoregions	1	26.284	28.728 [27.909–29.549]	-8.5%	5.7%	0/10000
Cohort B four-quarter	Ecoregions	2	9.722	9.609 [9.323–9.921]	+1.2%	6.2%	7681/10000
Cohort B four-quarter	Realms	0	4†	5 [4–5]	4 vs 5†	20.0%	3747/10000
Cohort B four-quarter	Realms	1	2.630	2.655 [2.624–2.685]	-0.9%	2.3%	531/10000
Cohort B four-quarter	Realms	2	2.334	2.350 [2.328–2.371]	-0.7%	1.9%	749/10000

Note: No percentage threshold is applied. † Realm q=0 is discrete and structurally constrained. The two Cohort B contrasts are shown as 4 vs 5; empirical ranks 2.994/10,000 and 3.747/10,000 place both within their references; these ranks are not exact p-values.

Table 9. Country–FEOW association and within-country variation in the primary networks

Cohort	Hierarchy	Pool q1	Association	Conditional q1	Informative strata	Category df	Retained in informative strata
Cohort A	Ecoregions	37.524	0.482	6.532	30 / 48	214	96.7%
Cohort A	Realms	3.305	0.963	1.045	6 / 48	6	33.4%
Cohort B	Ecoregions	31.525	0.450	6.664	23 / 36	131	94.8%
Cohort B	Realms	2.638	0.932	1.068	4 / 36	4	66.9%

Note: Pool $q1 = \exp[H(E)]$ and conditional $q1 = \exp[\sum_c p_c H(E|C=c)]$ use natural logs and station weights over the candidate pool restricted to represented countries. Association = $1 - H(E|C)/H(E)$. All represented strata enter the entropy terms. Sampling slack means retained country count < candidate-pool country count; informative strata also contain >1 ecological unit. Category df = $\sum_c (K_c - 1)$ over informative strata.

completely; member sizes, complete SHA-256 hashes, and verification status are reported in Tables S3 and S17.

Outlier inclusion, VPU 605 exposure, and duplicate collapse did not explain the main patterns. Reincorporating outliers changed q=2 by at most 0.220; VPU 605 contributed 13 Cohort A, and no Cohort B stations; and high-confidence duplicate collapse shifted q=2 by at most 0.076 (Tables S6, S7, and S8). Area-proportional comparisons remain a scale benchmark rather than a target distribution (Table S9).

DISCUSSION

The primary-to-four-quarter comparison demonstrates why a temporal rule cannot be evaluated by a single headline metric. Country q=2 increased by 0.442, but the same rule removed 1.752 stations, nine ecoregions, and 13.4 units of FEOW ecoregion q1, shifted the leading country, and increased the leading share. Its within-country ecoregion q2 residual remained small, yet q0 and q1 deficits deepened. These are simultaneous properties of one retained network, not contradictory statistical messages.

Table 10 integrates the archive audit into a monitoring-network decision record: temporal feasibility, country coverage, FEOW occupancy, biogeographic breadth, leverage, and within-country FEOW structure are evaluated together. The four-quarter rule is not labeled good or bad in isolation; it produces a different network whose suitability depends on the intended assessment.

The principal result is directional instability, not the unsurprising rejection of exchangeability. EC and TP produced selected increases in dominance-sensitive ecoregion representation, DO declined across all three hierarchies, NO₃-N and TSS showed short-lived realm increases, and DOC lost 90.7% of stations while gaining ecoregion $q=2$. The same filter can change occupied units, Shannon-weighted balance, and dominance in different directions. The Hill profile is therefore substantive: $q=0$, $q=1$ and $q=2$ answer different design questions.

Cohort B illustrates the danger of relying on $q=2$ alone. Relative to equal-size draws, it occupied fewer ecoregions but distributed stations more evenly among those retained. The +46.1% equal-size $q=2$ excess shrank to a +0.5% within-country residual, so the equal-size evenness excess was largely reproduced after country counts were fixed and was therefore consistent with country composition rather than an independent property of the retained FEOW profile. The $q=0$ deficit persisted. Realm $q=0$, by contrast, was four rather than a median of five but remained within its reference range. No single statement that representation increased or decreased is therefore adequate across Hill orders.

The within-country comparison also depended on the FEOW scale. Because resampling fixes country counts, small $q=2$ residuals mean that dominant-category concentration was largely

reproduced after country counts were fixed and was therefore consistent with country composition. They do not identify a monitoring agency, prove a causal national mechanism, or exclude within-country FEOW selection. Most countries occupy one realm, leaving little room for realm composition to change. Ecoregions vary more within countries: $q=0/q1$ deficits remained even when $q=2$ was near the within-country median, and the feasible-range audit confirmed substantial $q=2$ freedom only for the two primary networks.

The country label is therefore an explanatory stratum, not a program identifier. Caravan-Qual contains harmonized stations the country of which is assigned geographically. Agency provenance would be required to test whether one institution or national program caused the retention pattern. The present design can state that most of the $q=2$ pattern was reproduced after country counts were fixed, making it consistent with country composition, and that residual ecoregion richness loss remains; it cannot replace those statements with a causal mechanism.

Systematic leave-one-country-out analysis strengthened the influence diagnosis and exposed asymmetry. The United Kingdom's removal overwhelmingly changed EC because one country-category block was concentrated in FEOW 404. In primary Cohort B, Mexico and the United States changed ecoregion $q=2$ by -47.6% and -45.2% , whereas the United Kingdom's removal changed it by $+54.3\%$. A country can therefore be influential by contributing either concentration or breadth, and dominance alone does not determine the sign. The 8.4-fold ecoregion $q=2$ change after the United Kingdom's removal is a property of this release rather than a general consequence of temporal filtering; a different national composition could produce a different influence ranking.

Table 10. Worked Cohort B decision audit for stricter calendar-quarter coverage

Design criterion	Indicator	Primary 4d/3q	Four-quarter 4d/4q	Change / diagnosis
Temporal feasibility	Retained stations	4.373	2.621	-1.752 (-40.1%)
Country coverage	Countries $n / q2$	36 / 4.460	33 / 4.902	-3 / +0.442
FEOW coverage	Ecoregions $n / q2$	91 / 19.066	82 / 9.722	-9 / -9.344
Biogeographic breadth	Realms $n / q2$	4 / 2.904	4 / 2.334	0 / -0.570
Country leverage	Dominant country	Mexico (34.4%)	United States (41.0%)	Leader changed; leading share +6.6 pp
Within-country FEOW profile	$q0 / q1 / q2 \% \Delta$	-8.1% / -4.1% / +0.5%	-13.7% / -8.5% / +1.2%	Richness and $q1$ decline; $q2$ alone does not show it

Note: The within-country row reports ecoregion $q0/q1/q2$ percentage residuals.

The 1980–2025 sensitivity-window analysis provides a second caution. Extending the opportunity period increased retained stations but amplified some equal-size departures. The minimum number of eligible years is not interpretable without its observation window. The primary 2000–2020 results are a worked contemporary design, while the 1980–2025 analysis shows that magnitude is release-window specific.

An archive filter specifies the retained monitoring network used to assess stressors, restoration outcomes, or transferable models. The worked case has the same decision structure as field-network design: temporal regularity competes with station retention, FEOW representation, and concentration. It extends basin-scale monitoring logic (Odnorih et al., 2020) to the selection that occurs after a global archive has been assembled; it does not claim to design a new physical station network.

Station retention records temporal feasibility; $q=0$ records occupied FEOW units; $q=1$ records balance across common and uncommon units; $q=2$ records dominance; leading share exposes country leverage; and the two resampling references distinguish an exchangeability baseline from patterns conditional on country counts. Reporting these quantities together avoids reducing a network change to sample size or evenness alone.

For applied use, the eligibility rule can be reported as part of the network specification together with station retention, $q=0/q=1/q=2$ FEOW representation, country leverage, and observation-window sensitivity before the archive is used to evaluate restoration, stressors, or model transfer.

The study remains descriptive and release-specific. FEOW point assignment measures station distribution among mapped ecoregions and realms, not upstream-catchment exposure or ecological representativeness in a strong causal sense; polygon area also ignores river-network density. The equal-size reference is not conditioned on station record abundance and therefore cannot test selection beyond monitoring effort. The within-country reference does not preserve basin-, agency-, source-program- or spatial-clustering dependence, so remaining patterns may partly reflect those structures. Censored values were not re-estimated, and the analysis does not optimize new locations or infer ecological conditions. Predictive use would additionally require grouped validation and area-of-applicability assessment because representation descriptors do not establish transferability (Meyer and Pebesma, 2021; Roberts et al., 2017).

Together, the six-constituent analysis, complete-member input verification, systematic country-influence assessment, duplicate-network accounting, and observation-window sensitivity provide a reproducible characterization of the retained monitoring networks.

CONCLUSIONS

Temporal eligibility changed the composition of the retained monitoring network, but the direction depended on constituent, FEOW hierarchy, and Hill order. Richness, Shannon-weighted balance, and dominance-sensitive representation were not interchangeable: Cohort B retained fewer ecoregions while becoming more even among those that remained.

After country counts were fixed, most $q=2$ concentration patterns were reproduced, yet did not erase richness and $q1$ deficits. The country nearly determined realm membership, whereas several effective ecoregions remained within countries. The dominant-category pattern is therefore consistent with country composition, but the data neither identify national programs nor establish selection beyond monitoring effort.

The four-quarter rule increased country $q=2$ while removing 40.1% of stations and nine ecoregions, reducing FEOW balance and increasing the leading-country share. Temporal feasibility, ecoregion occupancy, balance, and country leverage thus formed one network change, and no single metric represented it completely.

Acknowledgements

The authors acknowledge the maintainers of Caravan-Qual, FEOW and the source monitoring systems for making the underlying data and provenance available.

REFERENCES

1. Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., Coad, B., Mandrak, N., Balderas, S. C., Bussing, W., Stiassny, M. L. J., Skelton, P., Allen, G. R., Unmack, P., Naseka, A., Ng, R., Sindorf, N., Robertson, J., Armijo, E., ... Petry, P. (2008). Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *BioScience*, 58(5), 403–414. <https://doi.org/10.1641/B580507>

2. Alilou, H., Moghaddam Nia, A., Saravi, M. M., Salajegheh, A., Han, D., Bakhtiari Enayat, B. (2019). A novel approach for selecting sampling points locations to river water quality monitoring in data-scarce regions. *Journal of Hydrology*, 573, 109–122. <https://doi.org/10.1016/j.jhydrol.2019.03.068>
3. Alvarez-Vázquez, L. J., Martínez, A., Rodríguez, C., Vázquez-Méndez, M. E. (2023). Optimizing the design of an estuarine water quality monitoring network by optimal control techniques. *Environmental Modeling & Assessment*, 28(4), 665–672. <https://doi.org/10.1007/s10666-023-09882-5>
4. Bastidas Sánchez, C. J. (2026). Temporal resolution requirements for event-oriented indoor air quality monitoring: Peak attenuation, missed events, and event-ranking reversals. *Ecological Engineering & Environmental Technology*, 27(9), 361–378. <https://doi.org/10.12912/27197050/226530>
5. Birk, S., Chapman, D., Carvalho, L., Spears, B. M., Andersen, H. E., Argillier, C., Auer, S., Baattrup-Pedersen, A., Banin, L., Beklioglu, M., Bondar-Kunze, E., Borja, A., Branco, P., Bucak, T., Buijse, A. D., Cardoso, A. C., Couture, R.-M., Cremona, F., De Zwart, D., ... Hering, D. (2020). Impacts of multiple stressors on freshwater biota across spatial scales and ecosystems. *Nature Ecology & Evolution*, 4(8), 1060–1068. <https://doi.org/10.1038/s41559-020-1216-4>
6. Camara, M., Jamil, N. R., Abdullah, A. F. B., Hashim, R. B., Aliyu, A. G. (2020). Economic and efficiency based optimisation of water quality monitoring network for land use impact assessment. *Science of The Total Environment*, 737, 139800. <https://doi.org/10.1016/j.scitotenv.2020.139800>
7. Harwell, L. C., McMillion, C. A., Lamper, A. M., Summers, J. K. (2024). Development of a generalized pseudo-probabilistic approach for characterizing ecological conditions in estuaries using secondary data. *Environmental Monitoring and Assessment*, 196(8), 753. <https://doi.org/10.1007/s10661-024-12877-8>
8. Hill, M. O. (1973). Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology*, 54(2), 427–432. <https://doi.org/10.2307/1934352>
9. Jiang, J., Tang, S., Han, D., Fu, G., Solomatine, D., Zheng, Y. (2020). A comprehensive review on the design and optimization of surface water quality monitoring networks. *Environmental Modelling & Software*, 132, 104792. <https://doi.org/10.1016/j.envsoft.2020.104792>
10. Jones, E. R., Graham, D. J., Van Griensven, A., Flörke, M., van Vliet, M. T. H. (2024). Blind spots in global water quality monitoring. *Environmental Research Letters*, 19(9), 091001. <https://doi.org/10.1088/1748-9326/ad6919>
11. Jones, E. R., Kratzert, F., van Vliet, M. T. H. (2026a). Caravan-Qual: A global scale integration of stream water quality observations into a large-sample hydrology dataset. *Scientific Data*, 13(1), 1015. <https://doi.org/10.1038/s41597-026-07352-7>
12. Jones, E. R., Kratzert, F., van Vliet, M. T. H. (2026b). *Caravan-Qual: Code for recreating or extending the Caravan-Qual dataset* (Version 0cc1527015449c4b4498d6854fd860af7ab0cde0) [Computer software]. GitHub. <https://github.com/SustainableWaterSystems/Caravan-Qual/tree/0cc1527015449c4b4498d6854fd860af7ab0cde0>
13. Jones, E. R., Kratzert, F., & van Vliet, M. T. H. (2026c). *Caravan-Qual (lite): A global scale integration of water quality observations into a large sample hydrology dataset* (Version 1.1) [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.21337532>
14. Jost, L. (2006). Entropy and diversity. *Oikos*, 113(2), 363–375. <https://doi.org/10.1111/j.2006.0030-1299.14714.x>
15. Keum, J., Kornelsen, K., Leach, J., Coulibaly, P. (2017). Entropy applications to water monitoring network design: A review. *Entropy*, 19(11), 613. <https://doi.org/10.3390/e19110613>
16. Kirschke, S., Avellán, T., Bärlund, I., Bogardi, J. J., Carvalho, L., Chapman, D., Dickens, C. W. S., Irvine, K., Lee, S., Mehner, T., Warner, S. (2020). Capacity challenges in water quality monitoring: Understanding the role of human development. *Environmental Monitoring and Assessment*, 192(5), 298. <https://doi.org/10.1007/s10661-020-8224-3>
17. Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D., Nevo, S., Shalev, G., Matias, Y. (2023). Caravan—A global community dataset for large-sample hydrology. *Scientific Data*, 10(1), 61. <https://doi.org/10.1038/s41597-023-01975-w>
18. Li, J., Yang, N., Shen, Z. (2024). Evaluation of the water quality monitoring network layout based on driving-pressure-state-response framework and entropy weight TOPSIS model: A case study of Liao River, China. *Journal of Environmental Management*, 361, 121267. <https://doi.org/10.1016/j.jenvman.2024.121267>
19. Meyer, H., Pebesma, E. (2021). Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods in Ecology and Evolution*, 12(9), 1620–1633. <https://doi.org/10.1111/2041-210X.13650>
20. Odnorih, Z., Manko, R., Malovanyy, M., Soloviy, K. (2020). Results of surface water quality monitoring of the Western Bug River Basin in Lviv Region. *Journal of Ecological Engineering*, 21(3), 18–26. <https://doi.org/10.12911/22998993/118303>
21. Palmer, M. A., Bernhardt, E. S., Allan, J. D., Lake, P. S., Alexander, G., Brooks, S., Carr, J., Clayton, S., Dahm, C. N., Follstad Shah, J., Galat, D. L., Loss, S. G., Goodwin, P., Hart, D. D., Hassett, B.,

- Jenkinson, R., Kondolf, G. M., Lave, R., Meyer, J. L., ... Sudduth, E. (2005). Standards for ecologically successful river restoration. *Journal of Applied Ecology*, 42(2), 208–217. <https://doi.org/10.1111/j.1365-2664.2005.01004.x>
22. Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., Smol, J. P., Taylor, W. W., Tockner, K., Vermaire, J. C., Dudgeon, D., Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, 94(3), 849–873. <https://doi.org/10.1111/brv.12480>
23. Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. <https://doi.org/10.1111/ecog.02881>
24. Sayer, C. A., Fernando, E., Jimenez, R. R., Macfarlane, N. B. W., Rapacciuolo, G., Böhm, M., Brooks, T. M., Contreras-MacBeath, T., Cox, N. A., Harrison, I., Hoffmann, M., Jenkins, R., Smith, K. G., Vié, J.-C., Abbott, J. C., Allen, D. J., Allen, G. R., Barrios, V., Boudot, J.-P., ... Darwall, W. R. T. (2025). One-quarter of freshwater fauna threatened with extinction. *Nature*, 638(8049), 138–145. <https://doi.org/10.1038/s41586-024-08375-z>
25. Simaika, J. P., Stribling, J., Lento, J., Bruder, A., Poikane, S., Moretti, M. S., Rivers-Moore, N., Meissner, K., Macadam, C. R. (2024). Towards harmonized standards for freshwater biodiversity monitoring and biological assessment using benthic macroinvertebrates. *Science of The Total Environment*, 918, 170360. <https://doi.org/10.1016/j.scitotenv.2024.170360>
26. Smits, A. P., Hall, E. K., Deemer, B. R., Scordo, F., Barbosa, C. C., Carlson, S. M., Cawley, K., Grossart, H., Kelly, P., Mammola, S., Pintar, M. R., Robbins, C. J., Ruhi, A., Saccò, M. (2025). Too much and not enough data: Challenges and solutions for generating information in freshwater research and monitoring. *Ecosphere*, 16(3), e70205. <https://doi.org/10.1002/ecs2.70205>
27. Strobl, R. O., Robillard, P. D. (2008). Network design for water quality monitoring of surface freshwaters: A review. *Journal of Environmental Management*, 87(4), 639–648. <https://doi.org/10.1016/j.jenvman.2007.03.001>
28. Su, G., Logez, M., Xu, J., Tao, S., Villéger, S., Brosse, S. (2021). Human impacts on global freshwater fish biodiversity. *Science*, 371(6531), 835–838. <https://doi.org/10.1126/science.abd3369>
29. Syafrudin, S., Sarminingsih, A., Juliani, H., Budi-hardjo, M. A., Puspita, A. S., Auliya Arlin Mirhan, S. (2024). Water quality monitoring system for temperature, pH, Turbidity, DO, BOD, and COD Parameters based on internet of things in Garang Watershed. *Ecological Engineering & Environmental Technology*, 25(2), 1–16. <https://doi.org/10.12912/27197050/174412>
30. Virro, H., Amatulli, G., Kmoch, A., Shen, L., Uuemaa, E. (2021). GRQA: Global river water quality archive. *Earth System Science Data*, 13(12), 5483–5507. <https://doi.org/10.5194/essd-13-5483-2021>
31. Watson, J., Zidek, J. V., Shaddick, G. (2019). A general theory for preferential sampling in environmental networks. *The Annals of Applied Statistics*, 13(4). <https://doi.org/10.1214/19-AOAS1288>
32. Zidek, J. V., Shaddick, G., Taylor, C. G. (2014). Reducing estimation bias in adaptively changing monitoring networks with preferential site selection. *The Annals of Applied Statistics*, 8(3). <https://doi.org/10.1214/14-AOAS745>