Supplementary Information

S1. Supplementary Data Structure and Analytical Framework

S1.1. Data Standardization and Analytical Input

All community-level analyses in this study were conducted using a unit-sum relative-abundance (RA) matrix in which the sum of taxa within each sample equals $\Sigma p = 1.0$ before rounding. Raw diatom valve counts were standardized on a per-sample basis to ensure comparability across sampling sites, riverbanks, and sampling periods.

This RA matrix served as the exclusive analytical input for all community-based statistical procedures, including: (i) diversity indices, namely Shannon diversity (H′), Pielou's evenness (J′), Margalef richness (e′), and dominance (D); (ii) β-diversity partitioning into turnover and nestedness components; (iii) ordination analyses, including NMDS, PCA, and dbRDA; (iv) multivariate diagnostic tests, including PERMANOVA, PERMDISP, and SIMPER; and (v) indicator species analysis (ISA).

Absolute densities, scaled counts, intermediate abundance metrics, and non-standardized values were not used in any statistical tests. Hellinger transformation was applied only where required by Euclidean-based ordination methods, whereas Bray–Curtis-based procedures were conducted on non-transformed RA data. Accordingly, the analytical workflow was designed to maintain consistency across community-level comparisons while minimizing ambiguity in data handling.

S1.2. Analytical Scope and Interpretation

The supplementary datasets and figures presented in this appendix are intended to complement the basin-scale summaries shown in the main text by providing additional spatial, temporal, and methodological detail. In particular, Appendix Tables S1–S3 and Figures S1–S3 extend the interpretation of the results by documenting the structure of the analytical dataset, the sampling framework, and the multivariate patterns underlying the principal findings.

All ordination, β-diversity, and machine-learning results should be interpreted as descriptive and exploratory, rather than deterministic or causal. The observed temporal groupings ("earlier monitoring period" and "later monitoring period") were defined as analytical categories used to summarize assemblage trajectories during the 2020–2024 study window and should not be interpreted as strict before–after experimental treatments or direct causal evidence of restoration success. Although bank position was evaluated separately as a preliminary design factor, bank-specific observations were retained as distinct sampling units in the final analytical matrix because bank position was not treated as a focal explanatory variable in the present study.

S2. β-Diversity Partitioning and Community Diagnostics

This table summarizes β-diversity partitioning following the Baselga framework and associated multivariate analyses derived exclusively from the final RA matrix. Turnover accounted for 72% of total dissimilarity, whereas nestedness accounted for 28%, indicating that temporal differentiation in benthic diatom assemblages was driven primarily by species replacement rather than by ordered species loss or gain. PERMANOVA detected a significant contrast in community composition between the defined monitoring periods, whereas PERMDISP was not significant, supporting the interpretation that between-

group differences reflected shifts in group centroids rather than increased within-group dispersion.

Table S1. Summary of β-diversity partitioning and multivariate diagnostics based on the unit-sum relative-abundance matrix.

Metric	Statistic	Result	Interpretation
β-diversity (Baselga, 2010)	Turnover component	72%	Turnover-dominated β-diversity, indicating that compositional differentiation among samples was primarily associated with species replacement rather than nested species loss or gain.
	Nestedness component	28%	Secondary contribution of nestedness, suggesting that systematic species loss or gain across sites played a subordinate role in explaining compositional differences.
PERMANOVA (adonis2)	$F = 4.86$, $p < 0.01$	Group-level compositional contrast detected	PERMANOVA identified a statistically supported contrast in multivariate community composition between the defined monitoring periods, reflecting shifts in group centroids.
PERMDISP (betadisper)	$F = 2.47$, $p = 0.11$	Not significant	No significant difference in multivariate dispersion, supporting the interpretation that PERMANOVA contrasts were associated with centroid shifts rather than heterogeneity changes.
SIMPER	Top five contributing taxa	Nitzschia gregaria, Stephanodiscus hantzschii, Achnanthidium gracillimum, Planothidium ellipticum, Cocconeis choctawhatcheeana	SIMPER results descriptively identify taxa contributing most strongly to Bray–Curtis dissimilarity, providing qualitative consistency across analytical approaches.
Site-level correlation	Spearman $\rho = 0.48$, $p = 0.04$	Moderate association	A weak-to-moderate descriptive association was detected between site-level Shannon diversity and compositional uniqueness.

Notes. Results represent descriptive summaries of community-level patterns and are not intended to imply mechanistic or causal relationships.

S3. Species Matrix and Data Availability

The complete species-by-sample relative-abundance (RA) matrix ($\Sigma p = 1.0$ per sample) is provided as a separate Excel file (Table S2). This dataset constitutes the final analytical community matrix used in all statistical analyses reported in the main text and appendices and is fully consistent with the preprocessing procedures described in Section 2.4.

Data structure of the Excel file

The matrix is organized in a species-by-sample format, with:

- Rows (sampling units): Individual observations defined as site × bank (L/R) × sampling event (date). Although bank position was evaluated separately as a preliminary design factor, bank-specific observations were retained as distinct sampling units in the final analytical matrix used for community-level analyses.

- Columns (variables): Epilithic diatom taxa identified to the lowest practicable taxonomic level
- Cell values: Relative abundance (proportion) of each taxon within each sample

Data preprocessing and standardization

Raw valve counts (≥400 valves per sample) were standardized independently for each sample such that:

$\sum p_i = 1.0$

before rounding, where p_i represents the proportional abundance of taxon i.

No additional preprocessing steps were applied, including:

- rarefaction
- zero-adjustment
- abundance scaling
- transformation for statistical input

Analytical use of the matrix

This matrix served as the exclusive analytical input for:

- diversity indices (H′, J′, e′, and dominance)
- β-diversity partitioning
- NMDS, PCA, and dbRDA
- PERMANOVA and PERMDISP
- SIMPER
- indicator species analysis (ISA)

For PCA and dbRDA, a Hellinger transformation was applied internally, whereas all Bray–Curtis-based analyses were conducted using non-transformed relative-abundance data. Accordingly, Table S2 represents the bank-specific final analytical matrix used to characterize temporal and spatial patterns across the monitoring network.

Data integrity and reproducibility

The Excel file contains a single, non-redundant analytical dataset, and no alternative matrices (e.g., absolute densities, transformed datasets, or scaled abundances) were used.

Accordingly, Table S2 represents the complete and reproducible analytical foundation for all results presented in this study.

Table S2. Species-level unit-sum relative-abundance matrix.

S4. Sampling Design and Temporal Classification

This table provides the sampling framework used in the study, including site and bank identifiers, exact sampling dates, and classification into the earlier monitoring period (2020–2021) and the later monitoring period (2022–2024). Sample identifiers were harmonized to match the site codes used consistently in the main text, figures, and appendices. These temporal groupings were used to summarize compositional trajectories of benthic diatom assemblages during the coordinated monitoring window and should not be interpreted as strict treatment phases. Bank-specific observations were retained in the final analytical matrix, even though bank position itself was not treated as a focal explanatory factor in the present study.

Table S3. Sampling calendar and analytical grouping.

Sample ID	Weir	Site	Bank	Date	Monitoring period
BJ1-L-2020-06-22	Baekje	BJ1	L	2020-06-22	Earlier monitoring period

BJ1-R-2020-06-22	Baekje	BJ1	R	2020-06-22	Earlier monitoring period
GJ1-L-2020-06-22	Gongju	GJ1	L	2020-06-22	Earlier monitoring period
GJ1-R-2020-06-22	Gongju	GJ1	R	2020-06-22	Earlier monitoring period
SJ1-L-2020-06-22	Sejong	SJ1	L	2020-06-22	Earlier monitoring period
SJ1-R-2020-06-22	Sejong	SJ1	R	2020-06-22	Earlier monitoring period
GJ2-L-2020-06-22	Gongju	GJ2	L	2020-06-22	Earlier monitoring period
GJ2-R-2020-06-22	Gongju	GJ2	R	2020-06-22	Earlier monitoring period
BJ1-L-2021-04-27	Baekje	BJ1	L	2021-04-27	Earlier monitoring period
BJ1-R-2021-04-27	Baekje	BJ1	R	2021-04-27	Earlier monitoring period
GJ1-L-2021-04-27	Gongju	GJ1	L	2021-04-27	Earlier monitoring period
GJ1-R-2021-04-27	Gongju	GJ1	R	2021-04-27	Earlier monitoring period
SJ1-L-2021-04-27	Sejong	SJ1	L	2021-04-27	Earlier monitoring period
SJ1-R-2021-04-27	Sejong	SJ1	R	2021-04-27	Earlier monitoring period
GJ2-L-2021-04-27	Gongju	GJ2	L	2021-04-27	Earlier monitoring period
GJ2-R-2021-04-27	Gongju	GJ2	R	2021-04-27	Earlier monitoring period
BJ1-L-2021-05-09	Baekje	BJ1	L	2021-05-09	Earlier monitoring period
BJ1-R-2021-05-09	Baekje	BJ1	R	2021-05-09	Earlier monitoring period
GJ1-L-2021-05-09	Gongju	GJ1	L	2021-05-09	Earlier monitoring period
GJ1-R-2021-05-09	Gongju	GJ1	R	2021-05-09	Earlier monitoring period
SJ1-L-2021-05-09	Sejong	SJ1	L	2021-05-09	Earlier monitoring period
SJ1-R-2021-05-09	Sejong	SJ1	R	2021-05-09	Earlier monitoring period
GJ2-L-2021-05-09	Gongju	GJ2	L	2021-05-09	Earlier monitoring period

GJ2-R-2021-05-09	Gongju	GJ2	R	2021-05-09	Earlier monitoring period
SJ1-L-2021-05-27	Sejong	SJ1	L	2021-05-27	Earlier monitoring period
BJ1-L-2022-09-29	Baekje	BJ1	L	2022-09-29	Later monitoring period
BJ1-R-2022-09-29	Baekje	BJ1	R	2022-09-29	Later monitoring period
GJ1-L-2022-09-29	Gongju	GJ1	L	2022-09-29	Later monitoring period
GJ1-R-2022-09-29	Gongju	GJ1	R	2022-09-29	Later monitoring period
GJ2-R-2022-09-29	Gongju	GJ2	R	2022-09-29	Later monitoring period
BJ1-L-2022-10-25	Baekje	BJ1	L	2022-10-25	Later monitoring period
BJ1-R-2022-10-25	Baekje	BJ1	R	2022-10-25	Later monitoring period
GJ1-L-2022-10-25	Gongju	GJ1	L	2022-10-25	Later monitoring period
GJ1-R-2022-10-25	Gongju	GJ1	R	2022-10-25	Later monitoring period
SJ1-L-2022-10-25	Sejong	SJ1	L	2022-10-25	Later monitoring period
SJ1-R-2022-10-25	Sejong	SJ1	R	2022-10-25	Later monitoring period
GJ2-L-2022-10-25	Gongju	GJ2	L	2022-10-25	Later monitoring period
GJ2-R-2022-10-25	Gongju	GJ2	R	2022-10-25	Later monitoring period
BJ1-L-2023-06-20	Baekje	BJ1	L	2023-06-20	Later monitoring period
BJ1-R-2023-06-20	Baekje	BJ1	R	2023-06-20	Later monitoring period
GJ1-L-2023-06-20	Gongju	GJ1	L	2023-06-20	Later monitoring period
GJ1-R-2023-06-20	Gongju	GJ1	R	2023-06-20	Later monitoring period
SJ1-L-2023-06-20	Sejong	SJ1	L	2023-06-20	Later monitoring period
SJ1-R-2023-06-20	Sejong	SJ1	R	2023-06-20	Later monitoring period
GJ2-L-2023-06-20	Gongju	GJ2	L	2023-06-20	Later monitoring period

GJ2-R-2023-06-20	Gongju	GJ2	R	2023-06-20	Later monitoring period
BJ1-L-2023-08-08	Baekje	BJ1	L	2023-08-08	Later monitoring period
BJ1-R-2023-08-08	Baekje	BJ1	R	2023-08-08	Later monitoring period
GJ1-L-2023-08-08	Gongju	GJ1	L	2023-08-08	Later monitoring period
GJ1-R-2023-08-08	Gongju	GJ1	R	2023-08-08	Later monitoring period
SJ1-L-2023-08-08	Sejong	SJ1	L	2023-08-08	Later monitoring period
SJ1-R-2023-08-08	Sejong	SJ1	R	2023-08-08	Later monitoring period
GJ2-L-2023-08-08	Gongju	GJ2	L	2023-08-08	Later monitoring period
GJ2-R-2023-08-08	Gongju	GJ2	R	2023-08-08	Later monitoring period
BJ1-L-2023-09-10	Baekje	BJ1	L	2023-09-10	Later monitoring period
BJ1-R-2023-09-10	Baekje	BJ1	R	2023-09-10	Later monitoring period
GJ1-L-2023-09-10	Gongju	GJ1	L	2023-09-10	Later monitoring period
GJ1-R-2023-09-10	Gongju	GJ1	R	2023-09-10	Later monitoring period
SJ1-L-2023-09-10	Sejong	SJ1	L	2023-09-10	Later monitoring period
SJ1-R-2023-09-10	Sejong	SJ1	R	2023-09-10	Later monitoring period
GJ2-L-2023-09-10	Gongju	GJ2	L	2023-09-10	Later monitoring period
GJ2-R-2023-09-10	Gongju	GJ2	R	2023-09-10	Later monitoring period
BJ1-L-2023-10-27	Baekje	BJ1	L	2023-10-27	Later monitoring period
BJ1-R-2023-10-27	Baekje	BJ1	R	2023-10-27	Later monitoring period
GJ1-L-2023-10-27	Gongju	GJ1	L	2023-10-27	Later monitoring period
GJ1-R-2023-10-27	Gongju	GJ1	R	2023-10-27	Later monitoring period
SJ1-L-2023-10-27	Sejong	SJ1	L	2023-10-27	Later monitoring period

SJ1-R-2023-10-27	Sejong	SJ1	R	2023-10-27	Later monitoring period
GJ2-L-2023-10-27	Gongju	GJ2	L	2023-10-27	Later monitoring period
GJ2-R-2023-10-27	Gongju	GJ2	R	2023-10-27	Later monitoring period
BJ1-L-2024-06-14	Baekje	BJ1	L	2024-06-14	Later monitoring period
BJ1-R-2024-06-14	Baekje	BJ1	R	2024-06-14	Later monitoring period
GJ1-L-2024-06-14	Gongju	GJ1	L	2024-06-14	Later monitoring period
GJ1-R-2024-06-14	Gongju	GJ1	R	2024-06-14	Later monitoring period
SJ1-L-2024-06-14	Sejong	SJ1	L	2024-06-14	Later monitoring period
SJ1-R-2024-06-14	Sejong	SJ1	R	2024-06-14	Later monitoring period
GJ2-L-2024-06-14	Gongju	GJ2	L	2024-06-14	Later monitoring period
GJ2-R-2024-06-14	Gongju	GJ2	R	2024-06-14	Later monitoring period

Notes. Monitoring periods represent analytical groupings rather than strict experimental phases. Sampling covered multiple seasons, and peak monsoon periods were mostly excluded to improve comparability among surveys, although selected summer observations were retained when field conditions allowed. Some survey dates are represented by incomplete site coverage in the final analytical matrix because only samples retained for analysis are listed here. Bank-specific observations were retained in the final analytical matrix, even though bank position itself was not treated as a focal explanatory factor in the present study.

S5. Spatial Configuration of Sampling Sites

This table summarizes the spatial and geomorphological configuration of the principal sampling sites relative to the Sejong, Gongju, and Baekje weirs. Distances represent along-channel measurements derived from GIS-based river centerline data following the main thalweg, and all sites were located within the expected hydromorphological influence zones of the respective weirs.

Table S4. Spatial characteristics of sampling sites relative to weirs.

Weir	Site	Position	Distance (km)	Description
Sejong	SJ1	Upstream	1.2	Gravel substrate
Gongju	GJ2	Downstream	0.8	Reduced velocity
Gongju	GJ1	Upstream	1.5	Transitional reach
Baekje	BJ1	Downstream	0.6	Fine sediment

Notes: Distances were measured along the main channel using GIS-based thalweg data.

S6. Supplementary Figures

This figure presents heatmaps of nine physicochemical variables, including temperature, pH, dissolved oxygen (DO), biochemical oxygen demand (BOD_5), electrical conductivity (EC), chlorophyll-a (Chl-a), turbidity (NTU), total nitrogen (TN), and total phosphorus (TP). Values represent annual means normalized to a 0–1 scale within each parameter to highlight relative temporal and spatial variability across bank-specific stations rather than absolute magnitudes. The figure complements the basin-scale summaries presented in the main text by revealing site-level heterogeneity and showing that nutrient-related variables and turbidity generally declined after 2021, whereas dissolved oxygen tended to increase.

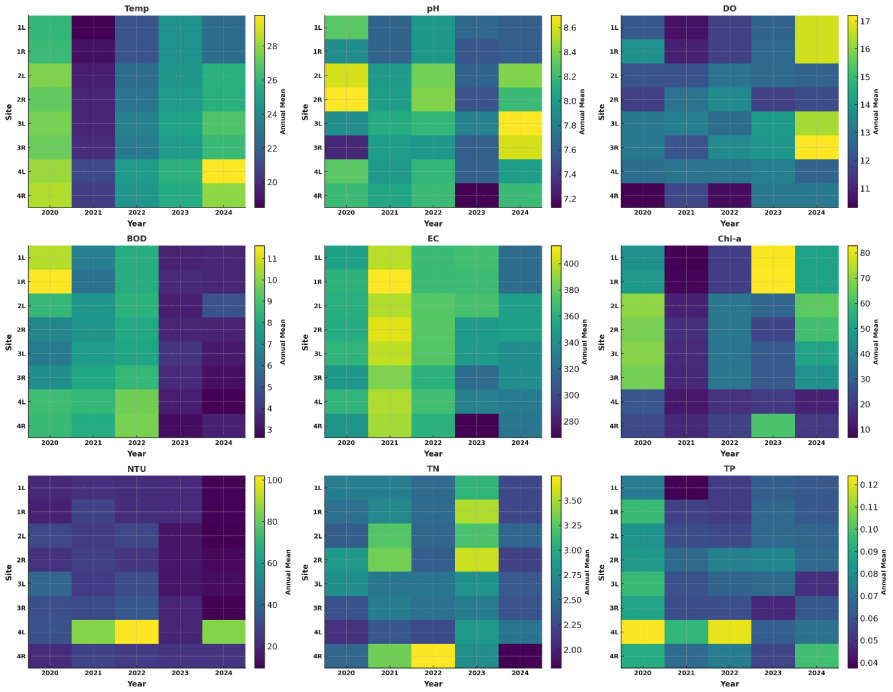


Figure S1. Site-level spatiotemporal heatmaps of water-quality variables across eight bank-specific biomonitoring stations (2020–2024).

This figure summarizes the temporal and spatial variation of benthic diatom ecological indices, including dominance (D), Shannon diversity (H′), Pielou's evenness (J′), species number, Margalef richness (e′), and total cell count. All diversity-based indices were calculated from the unit-sum RA matrix; however, total cell count is shown for descriptive reference only and was not used in the community-level statistical analyses. Compared with the earlier monitoring period (2020–2021), later monitoring-period patterns generally show lower dominance and relatively higher diversity, evenness, and richness at multiple bank-specific stations, supporting the interpretation of gradual community reorganization rather than abrupt or uniform recovery.

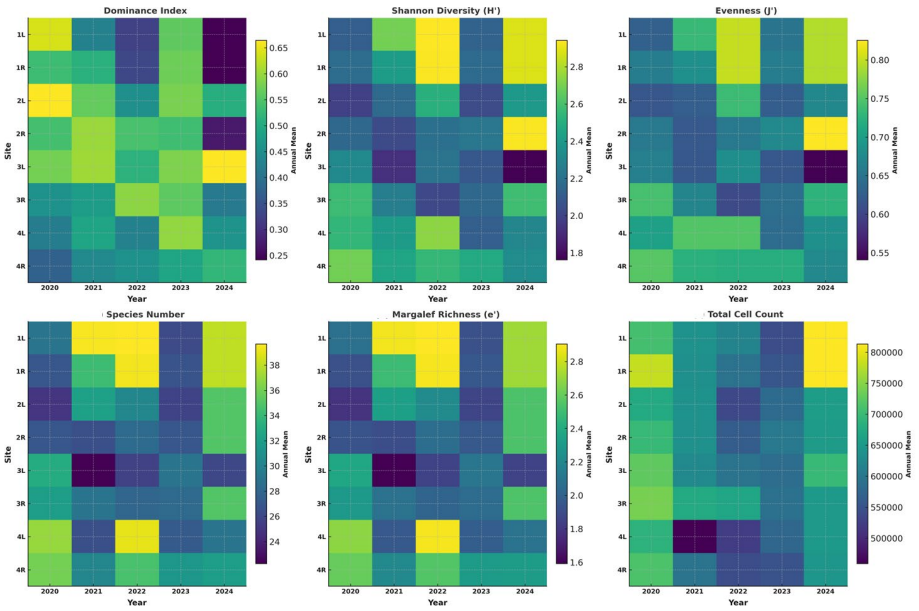


Figure S2. Site-level spatiotemporal heatmaps of benthic diatom ecological indices across eight bank-specific biomonitoring stations (2020–2024).

This figure presents the dbRDA ordination of benthic diatom assemblages constrained by measured environmental gradients. Red arrows indicate fitted vectors for TP, TN, chlorophyll-a, turbidity, BOD, DO, EC, pH, and temperature. The ordination summarizes multivariate associations between assemblage composition and environmental gradients across years, with later monitoring period samples occupying a broader region of ordination space than earlier samples. These patterns are interpreted as descriptive evidence of gradual temporal differentiation in assemblage structure under changing hydrological and physicochemical conditions, rather than as direct evidence of deterministic environmental control or restoration success.

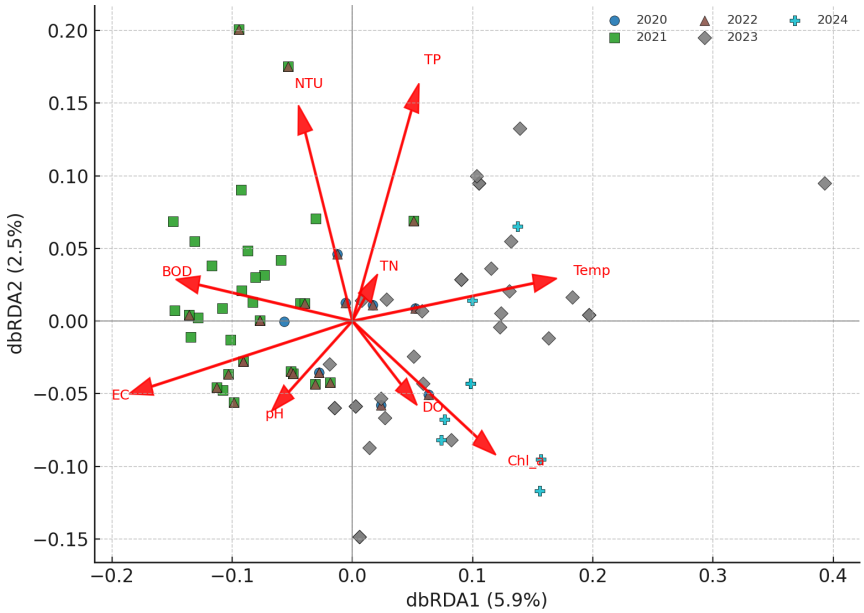


Figure S3. Distance-based redundancy analysis (dbRDA) of benthic diatom assemblages constrained by environmental variables (2020–2024).