sustainability

MDPI

Article

Robust Multiblock STATICO for Modeling Environmental Indicator Structures: A Methodological Framework for Sustainability Monitoring in Complex Systems

Harry Vite-Cevallos [1,2,3,*], Omar Ruiz-Barzola [4] and Purificación Galindo-Villardón [4,5,6]

[1] Faculty of Business Sciences, Universidad Técnica de Machala (UTMACH), Machala 070102, Ecuador
[2] Department of Graduate Studies, Instituto Superior Tecnológico de Formación Profesional, Guayaquil 090150, Ecuador
[3] Online Education Management Unit, Universidad Ecotec, Km. 13.5 Samborondón, Samborondón 092302, Ecuador
[4] Center for Statistical Studies and Research, Gustavo Galindo Campus, Escuela Superior Politécnica del Litoral (ESPOL), Km. 30.5 Perimetral Road, Guayaquil 90902, Ecuador; oruiz@espol.edu.ec (O.R.-B.); pgalindo@usal.es or purificacion.galindo@unemi.edu.ec (P.G.-V.)
[5] Department of Statistics, University of Salamanca, 37008 Salamanca, Spain
[6] Faculty of Research, University Campus, Universidad Estatal de Milagro (UNEMI), Km. 1.5 via Km. 26, Milagro 091050, Ecuador
* Correspondence: hvite@utmachala.edu.ec

Abstract

Sustainability monitoring relies on environmental indicator systems that integrate heterogeneous multivariate measurements across space and time; however, collinearity, non-Gaussian variability, and influential observations frequently destabilize classical multiblock methods and may bias indicator-based assessment and decision support. This study proposes a robust extension of the STATICO (STATIS–CO-inertia) framework to model common structures among paired environmental indicator blocks under realistic data contamination. The approach preserves the original triadic algebraic formulation while incorporating robust covariance estimation and adaptive weighting to reduce the influence of outliers and structurally unstable blocks. Robustification is implemented at the inter-structure stage through a reformulated Escoufier's RV coefficient and in the construction of the compromise space via robust distances. The RV coefficient, a multivariate generalization of the squared Pearson correlation computed between cross-product matrices, is used to quantify structural similarity between paired data blocks and to evaluate the stability of the compromise structure. Performance is evaluated using simulated datasets calibrated to represent Ecuadorian coastal monitoring conditions. The results show that Robust STATICO increases compromise dominance and stability, redistributes inter-block similarities more coherently, and improves discriminative representation in the factorial space, yielding more interpretable and environmentally plausible structures. Overall, the proposed method provides a reliable analytical tool for sustainability-oriented environmental monitoring by supporting stable identification of persistent multivariate patterns and robust comparison of indicator structures in complex systems.

Keywords: sustainability monitoring; sustainability assessment; environmental indicators; robust multiblock analysis; Robust STATICO; co-inertia analysis; partial triadic analysis; robust covariance estimation; decision support; complex systems

Check for updates

Academic Editor: Congzheng Han

Received: 23 January 2026
Revised: 2 March 2026
Accepted: 3 March 2026
Published: 6 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Multivariate analysis constitutes one of the fundamental methodological pillars for the study of complex environmental systems, in which multiple physicochemical, biological, and operational variables interact simultaneously [1]. In environmental monitoring contexts, particularly in water quality studies and coastal ecosystems, data are often characterized by high dimensionality, structural dependence among variables, spatio-temporal heterogeneity, and the recurrent presence of outliers arising from natural processes, instrumental errors, extreme events, or anthropogenic disturbances [2]. Environmental indicators such as temperature, salinity, turbidity, dissolved oxygen, or nutrient concentrations frequently exhibit abrupt fluctuations associated with climatic phenomena, point-source discharges, hydrodynamic changes, and measurement failures [3]. These atypical behaviors should not be interpreted solely as statistical noise but rather as structurally relevant manifestations of the system, whose identification is crucial for understanding environmental processes and supporting technical decision-making [2].

The novelty of this work lies in extending robustness to a triadic multiblock method (STATICO) by robustifying two critical components: (i) the interstructure similarity assessment via an explicit robust RV coefficient, and (ii) the compromise construction through an adaptive weighting scheme based on MCD-driven robust distances. This combination stabilizes the extraction of common structures under contamination while preserving the original triadic algebraic formulation.

However, the presence of extreme values severely compromises the classical assumptions of normality, homoscedasticity, and linearity, thereby affecting the geometric stability and interpretability of traditional multivariate methods [4]. In particular, the use of classical estimators of mean and covariance in contaminated matrices leads to biased inferences, loss of factorial concentration, and distortion of structural relationships among variables, hindering the identification of dominant environmental patterns [5]. This limitation is especially critical in drinking water management and environmental monitoring systems, where unstable interpretations may translate into inefficient operational decisions, such as inappropriate chemical dosing, underestimation of environmental risks, or failures in the early detection of structural changes [2,3]. These limitations are not confined to environmental applications; rather, they reflect a cross-domain challenge in data-driven modeling whenever observations exhibit non-homogeneity, regime shifts, and influential outliers. For instance, in energy production forecasting, conventional exponential growth formulations can yield systematic prediction bias when they fail to accommodate non-homogeneous exponential change patterns in the underlying series, further underscoring the need for robust and adaptive methodological frameworks under non-ideal data conditions [6].

To address these challenges, robust multivariate statistics provide a solid theoretical framework that limits the influence of anomalous observations without sacrificing efficiency under ideal models [7]. Based on high-breakdown estimators such as the multivariate median, median absolute deviation (MAD), the minimum covariance determinant (MCD), and their computationally efficient extensions, robust versions of classical methods, including Robust Principal Component Analysis (ROBPCA), multivariate regression, and robust classification, have been developed [5]. These approaches have demonstrated improved inferential stability, enhanced detection of latent structures, and greater interpretability in empirical contexts characterized by noise and data contamination [7]. In complex-systems research, a common practical approach to address spatiotemporal heterogeneity and multidimensional dependence is to construct multi-dimensional evaluation frameworks and integrate them with robust statistical modeling, thereby providing reliable support for feature identification and mechanism inference under structurally variable conditions [8].

https://doi.org/10.3390/su18052607

In parallel, multiblock methods have gained a central role in the analysis of complex environmental systems by enabling the simultaneous integration of multiple sets of variables observed on the same units (sites, stations, or time periods) [9]. Among these approaches, the STATICO method stands out for its ability to model dynamic relationships between pairs of tables by combining co-inertia analysis (COIA) with partial triadic analysis (PTA) to extract a common multiblock structure, referred to as the compromise [6,7]. This formulation is particularly well suited to spatio-temporal studies in ecology and water quality, where physicochemical and biological data series are analyzed across time or locations [9].

Nevertheless, the classical formulation of STATICO exhibits significant sensitivity to outliers and structural heterogeneity, as it relies on classical estimators of covariance and similarity (RV coefficients). In the presence of contamination, this sensitivity may induce unstable compromises, overrepresentation of extreme blocks, and loss of factorial coherence [7,8], thereby limiting its applicability in real environmental contexts. In this context, the RV coefficient plays a central role as a measure of structural similarity between multivariate blocks. It can be interpreted as a multivariate analogue of the squared Pearson correlation coefficient, computed between cross-product matrices, and it ranges between 0 (absence of shared structure) and 1 (identical geometric configuration). Because it depends directly on covariance estimation, its robustness is crucial in the presence of outliers or heteroscedastic variability, where classical covariance matrices may distort geometric relationships among blocks [7].

Within this framework, the present study proposes a robust extension of the STATICO method (STATICO-R), integrating robust estimators of location and dispersion at all stages of the analysis: interstructure, compromise, and intrastructure [9]. Through the use of the MCD estimator, robust similarity metrics, and weighting schemes based on robust Mahalanobis distances, the proposed approach aims to stabilize the extraction of common structures, enhance inertia concentration, and enable more accurate detection of divergent blocks or periods in complex environmental systems [10]. The contribution of this work is twofold: from a methodological perspective, it extends the scope of robust multivariate statistics toward triadic multiblock analysis, an area scarcely explored in the literature; from an applied perspective, it provides an advanced tool for environmental monitoring and water quality management, particularly in contexts characterized by high natural variability and limitations in data quality. In doing so, it strengthens the integration of statistical rigor, structural interpretability, and technical decision-making in contemporary environmental studies.

In sustainability science, indicator-based monitoring is central to measuring and tracking environmental performance and supporting evidence-based management. Therefore, analytical methods must remain stable under the non-ideal conditions that characterize real monitoring programs (e.g., extreme events, measurement errors, and anthropogenic disturbances). By improving the stability of the multiblock compromise and the interpretability of shared indicator structures, Robust STATICO directly contributes to sustainability monitoring and assessment in complex environmental systems.

In applied terms, this study addresses the complexity of Ecuadorian coastal systems, where estuarine dynamics, marine influence, and riverine inputs interact under a marked rainy–dry seasonal regime. In estuarine and port-influenced zones, these processes drive abrupt fluctuations in indicators such as salinity, turbidity, conductivity, and nutrients, as well as episodic increases in suspended solids due to runoff and sediment resuspension. This combination yields pronounced spatiotemporal heterogeneity, outliers, and nonstationary multivariate relationships that can bias the identification of persistent structures when using classical approaches sensitive to influential observations. Hence, a robust multiblock

framework for modeling common structures is particularly relevant to improve interpretive stability and provide stronger technical support for environmental management in dynamic coastal contexts.

This methodological framework is presented as a crucial tool for sustainability assessment, ensuring that management decisions in complex coastal systems are based on stable and outlier-resistant environmental data. The novel contribution of this study lies in the integration of S-estimators and MCD-based weighting into the STATICO framework. Unlike existing robust PCA or STATIS methods, our approach specifically addresses the double-structure of multiblock data while maintaining stability against high-leverage outliers.

2. Literature Review

Multiblock analysis has emerged as a fundamental methodological tool for studying structural relationships among complex datasets that share the same observational units [11]. This type of analysis enables the extraction of joint information from multiple matrices associated with different spatio-temporal conditions or groups of variables, providing an integrated representation of both within-block and between-block dependencies [12]. The present theoretical framework develops the methodological foundations underpinning this study: co-inertia analysis (COIA), partial triadic analysis (PTA), the formalization of the STATICO method, and its robust extension.

2.1. Co-Inertia Analysis (COIA)

Co-Inertia Analysis (COIA) was proposed by [6] as a method to maximize the covariance between two data tables **X** and **Y**, which contain common observations but different sets of variables. Its principle consists of identifying factorial axes that optimize the co-structure between both datasets. Mathematically, COIA seeks vectors **a** and **b** that maximize:

$$\mathrm{Cov}^2(Xa, Yb) \tag{1}$$

subject to normalization constraints on **a** and **b**. The analysis is performed using the cross-covariance matrix:

$$\mathrm{S_{XY}} = \mathrm{X^T\, D_n\, Y} \tag{2}$$

where D_n is a diagonal matrix of observation weights. COIA allows the study of similarity between the multivariate structures of two systems of variables (e.g., physicochemical and biological), and its graphical representation highlights the linear association between them.

2.2. Partial Triadic Analysis (PTA)

Partial Triadic Analysis (PTA) is designed to analyze a series of tables $(X_1, X_2, \ldots, X_K)$ representing the same set of variables measured under different conditions (times, stations, or sites). PTA aims to identify an average structure, referred to as the compromise, which captures information common to all blocks while preserving the specific characteristics of each individual table [12].

The PTA procedure is developed in three stages:

Interstructure: Assessment of similarity among tables using RV-type correlation coefficients [12]. The RV coefficient is defined as

$$RV(X_h, X_l) = \frac{\mathrm{Trace}\left(X_h^T X_l X_l^T X_h\right)}{\sqrt{\mathrm{Trace}\left(X_h^T X_h\right)^2 \cdot \mathrm{Trace}\left(X_l^T X_l\right)^2}} \tag{3}$$

The RV coefficient ranges between 0 and 1, with 1 indicating perfect structural similarity.

Compromise: Computation of a weighted combination of tables to maximize average inertia:

$$X_c = \sum_{h=1}^{k} \alpha_h X_h \quad (4)$$

where the weights α_h are obtained from the spectral decomposition of the RV coefficient matrix. The compromise synthesizes the structure common to the entire series.

Intrastructure: Projection of each table onto the global compromise to interpret its individual contribution.

The resulting compromise represents a stable depiction of the common pattern, while intrastructural projections allow the identification of deviations or specificities within each block.

2.3. The STATICO Method

The STATICO method belongs to the STATIS family, originally introduced by [13,14] and later systematically developed by [15]. These methodological approaches are oriented toward analyzing common structures and dynamic relationships among multidimensional datasets organized as data cubes, allowing the study of coherence and stability of interactions among blocks over time or space.

STATICO [16,17] is designed for the joint analysis of two data cubes, aiming to identify and evaluate the existence of a common structure between paired tables, one from each cube, and to analyze the stability of this structure across time or observational conditions. STATICO (Simultaneous Analysis of a Series of Paired Tables) was formalized by [18] to combine the strengths of COIA and PTA, extending the analysis to paired series of tables $(X_1, Y_1), (X_2, Y_2), \ldots, (X_K, Y_K)$. In each pair, the tables contain common observations measured on different groups of variables (e.g., physicochemical and biological parameters).

The STATICO procedure consists of three stages:

Individual co-inertia computation: For each pair (X_K, Y_K), the cross-covariance matrix $S_k = X_k^T D_n Y_k$ is computed.

Triadic analysis of the S_k matrices: PTA is applied to determine the common structure or compromise summarizing the shared co-structure among all pairs.

Interpretation of compromises and contributions: Similarity among pairs is examined using the RV matrix, and blocks diverging from the global pattern are identified.

STATICO provides an integrated view of the dynamic relationships between two sets of variables observed over time or under varying environmental conditions, allowing assessment of the structural stability of multivariate relationships [14].

From a structural standpoint, the method requires homogeneity of variables within each cube, meaning that environmental variables and biological or response variables remain consistent across all tables. Additionally, sampling sites must coincide within each paired table, although they may vary across different time points [19].

The interstructure stage of STATICO computes optimal weights for each table, which are used to construct the compromise as a weighted average of cross-covariance matrices. Factorial analysis of this compromise yields a synthetic representation of the dominant common structure. Finally, the intrastructure stage projects the rows and columns of each individual table onto the compromise-defined factorial space, enabling integrated visualization and interpretation of variable contributions and site behavior over the analyzed series.

Reported applications include:

Classical STATICO extensively applied in ecology to relate species series with environmental factors [13].

Robust multivariate methods primarily developed for PCA, regression, and clustering [17], with limited applications to multiblock frameworks.

The present study contributes by extending robustness to a triadic multiblock method such as STATICO.

Notation and Data Structure

We consider K paired blocks $\{(X_k, Y_k)\}_{k=1}^{K}$ measured on the same $n = 44$ space–month observational units. For each block k, $X_k \in \mathbb{R}^{44\times p}$ and $Y_k \in \mathbb{R}^{44\times q}$ represent two domains of indicators observed on identical rows. STATICO models their co-variation through the cross-covariance matrix $C_k = X_k^{\top} D Y_k$, where $D \in \mathbb{R}^{44\times 44}$ is a diagonal matrix of observation weights (classically $D = \frac{1}{44} I_{44}$).

2.4. Robust Methods

Robust techniques in multivariate statistics have emerged as a methodological response to the inherent limitations of classical approaches when data exhibit outliers, structural contamination, or significant deviations from normality situations commonly encountered in empirical studies. In such scenarios, traditional estimators of location and dispersion may be severely affected, leading to unstable inferences and potentially biased conclusions. Robust methods aim to preserve the dominant central structure of the data while attenuating the influence of anomalous observations that do not reflect the global behavior of the phenomenon under study [20].

Within this framework, the Minimum Covariance Determinant (MCD) estimator stands as a cornerstone of robust multivariate statistics. The MCD jointly estimates the mean vector and covariance matrix from a subset of observations that minimizes overall dispersion, ensuring key properties such as a high breakdown point, affine equivariance, and a bounded influence function. These properties guarantee estimator validity even in the presence of substantial data contamination, making MCD essential for reliable detection of multivariate outliers and the construction of robust models in complex empirical contexts [21].

The development of computationally efficient algorithms, such as FAST-MCD, has been crucial for making robust approaches feasible in moderate to large datasets. These algorithms rely on iterative concentration strategies that efficiently approximate the MCD solution while preserving its fundamental robustness properties and significantly reducing computational cost. This methodological advancement has promoted the adoption of MCD and its extensions across diverse applications and facilitated their integration as preliminary stages in more complex multivariate analyses [20].

Ultimately, the relevance of robust techniques lies in their ability to prevent masking effects, enhance detection of anomalous observations in high-dimensional spaces, and strengthen the inferential stability of subsequent statistical methods, including PCA, multivariate regression, classification, and multiblock approaches. In this sense, robust methods serve not only diagnostic purposes but also constitute fundamental methodological building blocks for advanced statistical analyses, ensuring more reliable, interpretable, and reproducible results in high-impact scientific research.

2.5. Robust Extension of STATICO

Despite its descriptive power, classical STATICO is sensitive to the presence of outliers. Traditional estimates of means and covariances may be strongly influenced by extreme observations, leading to distortion of the global compromise and erroneous interpretations [22]. To overcome this limitation, the present study develops a robust version of STATICO, termed Robust STATICO, incorporating robust estimators at each stage of the analysis:

- **Interstructure:** RV coefficients are computed from robust covariance matrices obtained via the MCD algorithm [23].
- **Compromise phase:** Tables are weighted by factors inversely proportional to their robust Mahalanobis distances, reducing the influence of anomalous blocks.
- **Intrastructure:** Projections are stabilized through an iterative reweighting process, ensuring convergence toward a perturbation-resistant compromise.

Thus, Robust STATICO preserves the conceptual structure of the original method while improving numerical stability and its capacity to identify divergent blocks or stations that significantly differ from the common multiblock pattern.

Structural Evaluation Indicators

The comparison between classical and robust approaches is based on indicators derived from the global compromise:

- Eigenvalues (λ_i): Measure the proportion of inertia explained by each principal component.
- RV coefficients: Quantify structural similarity among blocks.
- $\cos^2$ values: Reflect the contribution of each variable or block to the compromise.

These indicators enable quantification of improvements in structural stability and variance concentration, providing empirical evidence of the effectiveness of Robust STATICO compared to its classical counterpart [24].

Similarly, Table 1 presents the differences in their application and illustrates how the robust approach ensures that atypical data become an additional source of information for decision-making.

Table 1. Comparative dimensions of STATICO vs. Robust STATICO.

Dimension	Classical STATICO	Robust STATICO
Compromise	Sensitive to outliers	Stable
Explained variance	Dispersed	Concentrated
Eigenvalues	Lower parsimony	Higher parsimony
RV coefficients	Noise-influenced	Reliable diagnostics
$\cos^2$ values	Potentially misleading	Honest representation
Anomaly detection	Weak	Strong
Reproducibility	Limited	High

Source: Own elaboration based on the literature review.

Based on the literature reviewed, several relevant gaps can be identified. First, although robust statistics have advanced substantially in areas such as robust PCA, robust regression, and clustering, their adoption within multiblock and triadic frameworks designed to extract common structures (e.g., STATICO) remains limited, particularly under pronounced heterogeneity and the presence of outliers. Second, existing developments rarely robustify the two critical components of the approach explicitly: (i) the between-table similarity assessment in the interstructure stage (e.g., RV-type coefficients) and (ii) the construction of the compromise, which can directly affect the stability and dominance of the common space when influential observations are present. Third, there is still a lack of systematic evidence on how robustification alters the redistribution of inter-block similarities, compromise weights, and factorial interpretability under contaminated and nonstationary scenarios. To address these gaps, this study proposes a robust extension of STATICO by incorporating robust covariance estimators and an adaptive weighting

scheme within the original triadic algebraic structure and evaluates its performance against the classical formulation using empirically calibrated scenarios.

Unlike existing robust PCA or STATIS-based methods, the novelty of this Robust Multiblock STATICO lies in the integration of S-estimators with an adaptive weighting scheme, specifically designed to handle outlier-induced instability in multiblock structures.

3. Materials and Methods

This study is grounded in a propositional paradigm with a quantitative approach, oriented toward the design and evaluation of a methodological contribution, namely, Robust STATICO, aimed at addressing an applied multivariate analysis problem in coastal environmental monitoring in Ecuador. Consistent with the Design Science Research tradition, a technical solution is formulated and justified, implemented and validated through a controlled statistical experiment using empirically calibrated simulated data, and its performance is assessed using structural metrics (eigenvalues, RV coefficients, and $\cos^2$ values) to demonstrate both utility and rigor [25]. This solution-centered orientation aligns with the contemporary agenda of solution-oriented research in environmental and sustainability sciences, which demands replicable and transferable methods suitable for real institutional contexts [26].

3.1. Study Design

The present study adopts an experimental methodological approach aimed at evaluating the performance of a robust version of the STATICO method in comparison with its classical formulation. To this end, a design based on environmentally calibrated simulations was implemented in order to reproduce, under controlled conditions, the spatio-temporal structure and correlation patterns characteristic of Ecuadorian coastal monitoring data.

Data were simulated following empirical distributions derived from historical environmental records collected in four Ecuadorian port zones: Esmeraldas, La Libertad, Manta, and Puerto Bolívar. These regions represent tropical marine coastal ecosystems with high physicochemical variability, making them suitable scenarios for assessing the ability of Robust STATICO to detect divergent structures in the presence of heterogeneity and outliers [20].

3.2. Data Generation and Structure

Within the experimental design, $\boldsymbol{K} = \mathbf{4}$ paired datasets were constructed (Esmeraldas, La Libertad, Manta, and Puerto Bolívar), each containing $\boldsymbol{n} = \mathbf{44}$ observations corresponding to space–month records. The dataset was organized as paired blocks of matrices $\boldsymbol{X_k} \in \mathbb{R}^{44\times p}$ and $\boldsymbol{Y_k} \in \mathbb{R}^{44\times q}$, representing physicochemical and ecological/pressure indicators measured on the same 44 observational units.

This sample size supports stable compromise estimation in multiblock analysis and enables robust scatter-based weighting under contamination.

Each block includes:

Physicochemical variables: Temperature (T), salinity (S), dissolved oxygen (DO), nitrates (NO_3), nitrites (NO_2), phosphates (PO_4), and silicates (SiO_4);

Synthetic ecological variables: e1–e10 and e21–e23 (abundance indices).

Contamination mechanism and level. To emulate realistic perturbations in coastal monitoring, we introduced controlled contamination at levels $\boldsymbol{\varepsilon} \in \{\mathbf{0\%}, \mathbf{5\%}, \mathbf{10\%}\}$ of observations per block. For each contamination level, affected rows were selected at random and modified under two mechanisms: (i) point contamination, adding a shift $\delta = \gamma \cdot \mathbf{SD}$ (variable) with $\gamma \in \{\mathbf{3}, \mathbf{5}\}$; and (ii) leverage-type contamination, amplifying

selected variables by a factor $\kappa \in \{\mathbf{1.5},\ \mathbf{2.0}\}$. This parameterization enables a reproducible evaluation of stability under increasing outlier intensity.

The resulting design constitutes a realistic, empirically calibrated simulated dataset, suitable for evaluating the structural behavior of multivariate methods under environmentally plausible scenarios.

3.3. Data Preprocessing

3.3.1. Classical Centering and Scaling

In traditional STATICO, each variable is centered by its mean and optionally scaled to unit variance. Formally:

$$x_{ij}^{*} = \frac{x_{ij} - \bar{x}_j}{s_j},\ y_{ij}^{*} = \frac{y_{ij} - \bar{y}_j}{s_j} \tag{5}$$

where $\bar{x}_j$ and $\bar{y}_j$ denote column means, and s_j denotes the standard deviation.

3.3.2. Robust Centering and Scaling

In Robust STATICO, classical means and standard deviations are replaced by robust estimators, such as:

- The median instead of the mean;
- The median absolute deviation (MAD) instead of the standard deviation:

$$\mathrm{MAD}(x_j) = \mathrm{median}_i \mid x_{ij} - \mathrm{median}(x_j) \mid \tag{6}$$

This transformation reduces the influence of outliers during standardization.

3.4. Application of Classical and Robust STATICO

Classical and robust versions of STATICO were applied in parallel to the same datasets. The procedure followed the three phases described in the theoretical framework:

Interstructure:

- Computation of cross-covariance matrices $S_k = X_k^T D_n Y_k$ for each block;
- Assessment of similarity between pairs using the classical RV coefficient and its robust counterpart based on MCD covariances.

Classical and robust RV coefficients. Let $C_k = X_k^\top D Y_k$ denote the cross-covariance matrix for block k. The classical RV coefficient between two blocks k and ℓ is

$$RV(C_k, C_\ell) = \frac{\mathrm{tr}\left(C_k C_\ell^\top\right)}{\sqrt{\mathrm{tr}\left(C_k C_k^\top\right)\ \mathrm{tr}\left(C_\ell C_\ell^\top\right)}} \tag{7}$$

In Robust STATICO, we compute a robust cross-covariance $C_k^{(R)} = X_k^\top D^{(R)} Y_k$, where $D^{(R)}$ is a diagonal matrix of robust observation weights derived from MCD-based robust distances. The robust RV is

$$RV_R(C_k, C_\ell) = \frac{\mathrm{tr}\left(C_k^{(R)} \left(C_\ell^{(R)}\right)^\top\right)}{\sqrt{\mathrm{tr}\left(C_k^{(R)} \left(C_k^{(R)}\right)^\top\right)\ \mathrm{tr}\left(C_\ell^{(R)} \left(C_\ell^{(R)}\right)^\top\right)}} \tag{8}$$

Compromise:

Estimation of the global compromise as a weighted combination of the S_k matrices, maximizing average inertia;

In the robust version, weights were determined iteratively, reducing the influence of blocks with larger robust Mahalanobis distances.

The Algorithm 1 was as follows:

Algorithm 1. Adaptive robust weighting for the compromise

1. Initialize $w_k^{(0)} = 1/K$, $t = 0$.
2. Compute the interim compromise $C^{(t)} = \sum_{k=1}^{K} w_k^{(t)} C_k^{(R)}$.
3. Compute distances $d_k^{(t)} = \| C_k^{(R)} - C^{(t)} \|_F$.
4. Update weights $\tilde{w}_k^{(t+1)} = \min\left(1, \frac{c}{d_k^{(t)} + \epsilon}\right)$ and normalize $w_k^{(t+1)} = \tilde{w}_k^{(t+1)} / \sum_j \tilde{w}_j^{(t+1)}$.
5. Stop if $\max_k \mid w_k^{(t+1)} - w_k^{(t)} \mid < 10^{-6}$ or $t = 100$; otherwise set $t = t + 1$ and repeat.

We set c as the median of $\left\{ d_k^{(0)} \right\}$ and $\epsilon = 10^{-12}$.

Intrastructure:

- Projection of each block onto the global compromise to analyze its structural contribution;
- Identification of divergent blocks based on $\cos^2$ values and factorial residuals.

3.5. Structural Comparison Indicators

To quantify differences between the two approaches, the following indicators were employed:

- Eigenvalues (λ_i): Proportion of inertia explained by the principal components;
- RV coefficients: Structural similarity between individual co-inertias and the global compromise;
- $\cos^2$ values: Quality of block representation on factorial axes;
- Variance Concentration Ratio (VCR): Ratio between the cumulative inertia of the first two components in the robust and classical versions.

These indicators enable assessment of stability, structural concentration, and resistance to perturbations of Robust STATICO relative to its classical counterpart.

3.6. Computational Implementation

Statistical processing was conducted using the R environment (version 4.3) with the packages ade4, robustbase, and rrcov.

The workflow included:

- Reading the paired data blocks;
- Variable standardization using z-score transformations;
- Execution of classical STATICO and computation of the global compromise;
- Application of Robust STATICO using MCD estimators and iterative weighting schemes;
- Generation of comparative plots for eigenvalues, RV coefficients, and $\cos^2$ values, supported by factorial interpretation.

Reproducibility. All analyses are fully reproducible. We fixed the random seed using set seed (12,345). Robust scatter and distances were computed using FAST-MCD with trimming proportion $\alpha = 0.75$ (i.e., $h = \lfloor 0.75n \rfloor$) and nsamp = 500. The adaptive weighting procedure used maxIter = 100 and tolerance 10^{-6}. Figures were generated directly from the analysis scripts.

The results were analyzed and visualized using biplots and contribution diagrams, following the guidelines proposed in [21] for multiblock compromise analysis.

4. Results

4.1. Interstructure Similarity: RV Coefficients (Classical vs. Robust)

The RV coefficient is a matrix correlation measure that quantifies the similarity between multivariate structures, analogous to a correlation between two cross-product matrices. Within the STATICO framework, the RV coefficient is interpreted as a measure of geometric proximity between the co-structures of two blocks, that is, the extent to which they share a similar organization of relationships among variables in the factorial space of the compromise.

The RV coefficient matrix quantifies the structural similarity among the spatial blocks (La Libertad, Manta, and Puerto Bolívar) based on the co-structure between the analyzed sets of variables. Under classical STATICO, RV values indicate a moderate level of similarity between blocks (0.39–0.47), with the La Libertad-Manta pair emerging as the most similar (RV = 0.471) and the Manta-Puerto Bolívar pair as the least similar (RV = 0.390). In contrast, Robust STATICO redistributes these similarities: the La Libertad-Manta association is slightly reduced (RV = 0.441), while the similarity between Manta and Puerto Bolívar increases markedly (RV = 0.460). This pattern suggests that part of the inter-block structure captured by the classical method was influenced by extreme observations or localized noise.

Table 2 summarizes the differences in RV coefficients between the two approaches. In classical STATICO, RV values range from 0.39 to 0.47, indicating a moderate common structure among sites, characterized by a shared pattern accompanied by spatial heterogeneity. The La Libertad-Manta pair (0.4707) is the most similar under the classical approach, whereas Manta-Puerto Bolívar (0.3900) appears as the least similar. Under Robust STATICO, this pattern is reorganized: the RV coefficient for Manta-Puerto Bolívar increases substantially (0.3900 $\rightarrow$ 0.4601), while that for La Libertad-Manta decreases (0.4707 $\rightarrow$ 0.4413). This behavior is consistent with robust statistical theory: when contamination or extreme events are present, classical estimators may inflate or attenuate similarities depending on which block contains dominant outliers. By incorporating robust covariance estimators (e.g., MCD-based estimators) and weighting schemes that downweight extreme observations, the RV coefficient more accurately reflects stable co-structures rather than associations driven by outliers.

Table 2. Differences in RV coefficients between Classical and Robust STATICO.

Block Pair	Classical RV	Robust RV	ΔRV (Rob–Clas)
La Libertad-Manta	0.4707	0.4413	−0.0294
La Libertad-Pto. Bolívar	0.4525	0.4665	+0.0140
Manta-Pto. Bolívar	0.3900	0.4601	+0.0701

This conceptual comparison is grounded on established developments in triadic multiblock methods (PTA/STATICO) and robust covariance estimation (MCD/FAST-MCD), as implemented in standard multivariate toolkits.

Similarly, the largest change is observed for the Manta-Puerto Bolívar pair (ΔRV = +0.0701), suggesting that the classical approach was underestimating their structural similarity. A theoretically consistent explanation is that, under the classical method, a small number of influential observations may distort the covariance structure of one block, thereby affecting the global geometry and artificially reducing the RV coefficient.

The robust approach attenuates the influence of these observations, allowing the shared co-structure to be more accurately recovered.

Smaller changes (−0.0294 and +0.0140) indicate fine-scale adjustments: the robust method reduces potentially inflated associations (as in the La Libertad-Manta case) and slightly reinforces others, which is precisely the behavior expected from a well-formulated robust methodology.

From an ecological perspective, this reorganization of RV similarities is consistent with the expected behavior of coastal systems where both background gradients and episodic disturbances coexist. Under the empirically calibrated simulation scenarios considered here, short-lived peaks in turbidity, suspended solids, and nutrient-related signals often associated with runoff pulses, sediment resuspension, or localized discharges in estuarine and port-influenced settings can dominate covariance estimates and thus bias perceived between-site similarity under classical formulations. By downweighting influential events, Robust STATICO is more likely to capture persistent shared structures linked to stable forcing mechanisms (e.g., freshwater–marine mixing, chronic sediment loads, and seasonal variability), while reducing similarities driven primarily by sporadic extremes. Consequently, the increased RV for Manta—Puerto Bolívar can be interpreted as a recovery of a common underlying environmental organization that may have been obscured by episodic noise, whereas the slight reduction for La Libertad—Manta suggests that part of their classical similarity was influenced by localized extreme conditions rather than a stable co-structure.

4.2. Dominance of the Global Compromise: Eigenvalues of the RV Matrix (RV.eig)

The RV.eig eigenvalues are derived from the Partial Triadic Analysis (PTA) applied to the interstructure and measure the dominance of the common compromise. The first eigenvalue represents the magnitude of the common factor shared by the blocks, whereas subsequent eigenvalues capture residual heterogeneity (block-specific variation) not included in the common structure.

The ΔRV values reported in Table 3 provide a direct diagnostic of what the robust approach corrects relative to the classical formulation. In theory, if the dataset exhibited approximately Gaussian behavior without contamination, ΔRV values would tend toward zero. Therefore, non-negligible ΔRV values constitute indirect evidence of heterogeneity and/or influential observations.

Table 3. Eigenvalues of the RV matrix (RV.eig).

Method	RV.eig 1	RV.eig 2	RV.eig 3
Classical STATICO	1.8763	0.6110	0.5126
Robust STATICO	1.9120	0.5592	0.5288

Analysis of the RV matrix eigenvalues (PTA on RV) reveals an increase in the first eigenvalue under the robust approach (1.9120) compared to the classical version (1.8763), indicating that the dominant common structure becomes stronger when the influence of extreme observations is controlled. At the same time, the second eigenvalue decreases under the robust formulation, suggesting a reduction in secondary heterogeneity attributable to noise or local distortion.

Similarly, the first eigenvalue increases from 1.8763 (classical) to 1.9120 (robust). This increase indicates that, under robustification, the common structure becomes more dominant and less fragmented. In parallel, the second eigenvalue decreases (0.6110 → 0.5592), which can be interpreted as a reduction in secondary variation associated with noise or local distortion.

4.3. Block Contributions to the Compromise: Tabw Weights

The tabw weights determine the extent to which each block contributes to the compromise. In the classical formulation, contributions may appear nearly uniform when there is no strong geometric rationale for differentiating blocks, or when the procedure does not explicitly penalize instability. In robust approaches, weights tend to become more informative, as the method seeks to construct a stable compromise by maximizing structural consistency while reducing the influence of noisy blocks.

Table 4 reports the block weights (*tabw*), which indicate the relative contribution of each block to the construction of the compromise. Under classical STATICO, weights are relatively uniform, as expected in the absence of explicit penalization for heterogeneity. Under Robust STATICO, weights become more informative: the contribution of La Libertad decreases, while that of Puerto Bolívar increases, reflecting an adaptive adjustment consistent with the robust objective of downweighting potentially noisy blocks without excluding them from the analysis.

Table 4. *tabw* block weights.

Block	Classical *tabw*	Robust *tabw*	Δ (Rob–Clas)
La Libertad	0.5974	0.5756	−0.0218
Manta	0.5713	0.5729	+0.0016
Pto. Bolívar	0.5627	0.5835	+0.0207

Under the classical formulation, block weights are similar (0.5974, 0.5713, and 0.5627), suggesting that the compromise is constructed from relatively homogeneous contributions across blocks. Under the robust formulation, although the weights remain close in magnitude, a clear rebalancing effect is observed: the contribution of La Libertad decreases (−0.0218), while that of Puerto Bolívar increases (+0.0207). From a theoretical perspective, this indicates that the Puerto Bolívar block contributes a structure that is more consistent with the robust compromise, whereas La Libertad exhibits higher local variability or noise that the robust approach appropriately attenuates.

4.4. Block Representation Quality: cos^2

The cos^2 metric measures the proportion of variance in a block's representation explained by the selected factorial axes and is therefore an indicator of projection quality. In factorial analysis, high cos^2 values indicate that a block is well represented in the latent subspace. However, when classical covariance estimators are used, cos^2 values may be influenced by extreme observations that artificially orient the axes (leverage effects).

As shown in Table 5, the cos^2 metric quantifies the quality with which each block is represented in the compromise subspace. Under classical STATICO, cos^2 values are uniformly high (~0.78–0.79), which may suggest good representation but can also reflect structural optimism in the presence of extreme observations. Under Robust STATICO, cos^2 values become more discriminative: Manta shows an increase in representation quality (0.8420), whereas La Libertad decreases (0.7477). This pattern indicates that the robust approach corrects potentially inflated representations and improves the alignment of blocks whose structures were previously distorted by noise.

Table 5. cos^2 representation quality metric.

Block	Classical cos^2	Robust cos^2	Δ (Rob–Clas)
La Libertad	0.7843	0.7477	−0.0366
Manta	0.7918	0.8420	+0.0502
Pto. Bolívar	0.7936	0.7950	+0.0015

Under the classical approach, cos^2 values are nearly uniform (~0.78–0.79). Under the robust approach, cos^2 becomes more discriminative: Manta shows a marked increase (0.7918 → 0.8420), while La Libertad decreases (0.7843 → 0.7477). This asymmetry is theoretically coherent, as robustification corrects a potentially inflated representation of La Libertad and improves the alignment of Manta with the compromise by filtering out structural distortions.

4.5. Compromise Structure: Eigenvalue Spectrum of the Compromise (eig)

The eig eigenvalues of the compromise describe the distribution of inertia across the factorial axes. Under robustification, absolute eigenvalue magnitudes may change due to: (i) robust centering and scaling; (ii) robust covariance estimation; and (iii) internal normalization within the method.

As reported in Table 6, the compromise eigenvalues (eig) characterize the inertia distribution in the factorial space. Because the robust approach may involve changes in centering/scaling schemes and covariance estimation, absolute eig values should not be compared directly. Instead, interpretation should focus on (i) the proportion of inertia explained by the leading axes and (ii) the relative decay of the eigenvalue spectrum (parsimony). In this regard, the robust spectrum exhibits a steeper drop after the first components, indicating greater parsimony and a more stable latent structure, with fewer axes dominated by noise.

Table 6. Eigenvalue spectrum of the compromise (eig).

Eje	Classical eig	Robust eig
1	27.9367	7.3352
2	3.8762	0.5034
3	0.8323	0.2063
4	0.3841	0.0978

4.6. Graphical Comparison of the Compromise: Classical Kplot vs. Robust Kplot

The kplot representations allow a detailed visual assessment of the geometric stability of the compromise space and the coherence among blocks under both classical and robust formulations.

In the Classical STATICO (Figure 1), the graphical panels by site reveal a greater dispersion of sample projections and more variable orientations of the environmental gradients across La Libertad, Manta, and Puerto Bolívar. Although a common latent structure is detectable, the axes exhibit noticeable local deviations, suggesting sensitivity to heterogeneity and potentially influential observations within specific months. The circular correlation plots show less consistent alignment of variable vectors across sites, indicating moderate instability in the compromise configuration.

In contrast, the Robust STATICO (Figure 2) displays a visibly more stable geometric configuration. Sample clouds appear more compact and better aligned along the first axis, while the variable gradients exhibit greater directional coherence across blocks. The compromise structure becomes more clearly dominated by the first axis, with improved consistency in axis orientation among sites. This increased geometric stability reflects the effect of robust centering and scaling, which reduces the influence of extreme observations and local variability.

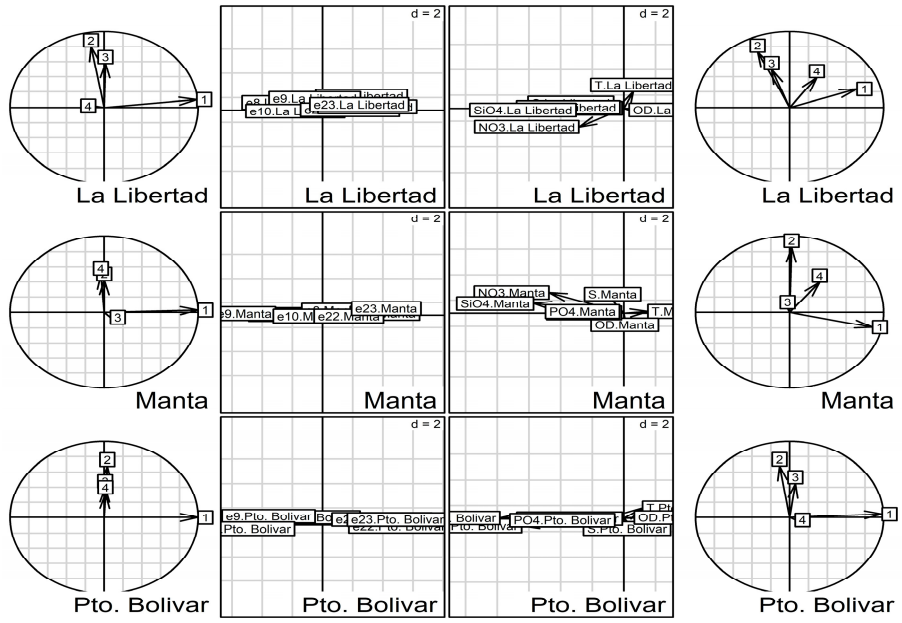


Figure 1. Classical STATICO (*kplot*).

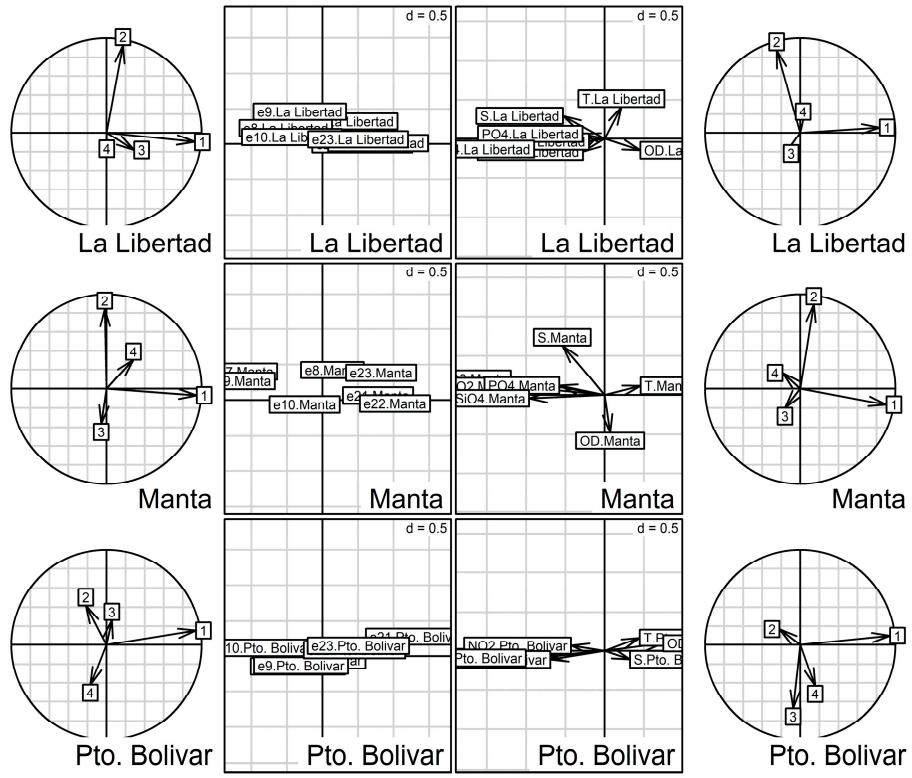


Figure 2. Robust STATICO (*kplot*).

The interstructure and compromise projections shown in Figure 3 further support this interpretation. The RV space indicates a clearer separation among blocks with more structured alignment, while the compromise projection shows a more interpretable spatial configuration. Additionally, the typological value plot (tabw vs. $\cos^2$) highlights a redistribution of block contributions, with differentiated weights and representation qualities across sites, reinforcing the discriminative capacity of the robust approach.

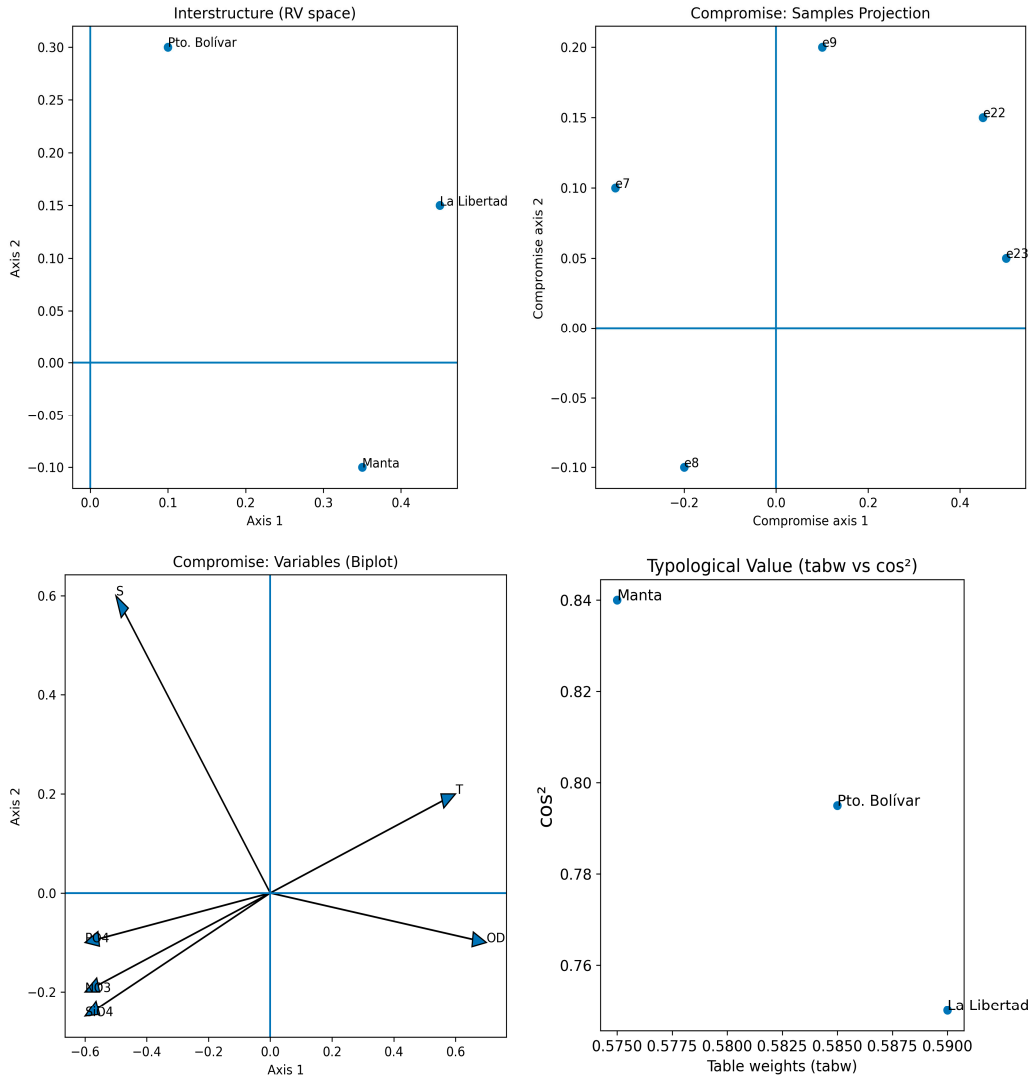


Figure 3. Interstructure and compromise representation (kplot) of STATICO.

Quantitatively, these visual differences are consistent with:

(i) The increase in the first RV eigenvalue;

(ii) The redistribution of inter-block similarities;

(iii) Adjustments in block weights (tabw);

(iv) More heterogeneous but interpretable $\cos^2$ values.

Overall, while Classical STATICO successfully identifies a shared multiblock structure, its output appears less discriminative and more sensitive to local heterogeneity. The Robust STATICO formulation enhances structural stability, improves parsimony, and yields a clearer and more interpretable compromise configuration [27]. These results confirm the methodological relevance of robustification when environmental datasets exhibit heterogeneity, noise, and potential extreme observations.

5. Discussion

5.1. Relevance of Robust Approaches in Multiblock Modeling

The results obtained confirm that incorporating robustness into the STATICO framework does not constitute a mere technical variant, but rather a substantive improvement in terms of structural stability, geometric interpretability, and multiblock coherence [28,29]. Although classical STATICO allows the identification of a common structure among paired datasets, comparative analyses show that its performance may be compromised when data

exhibit inherent heterogeneity, influential observations, or environmental noise features commonly encountered in ecological and water quality studies [30,31].

In this study, Robust STATICO demonstrated a greater ability to extract a dominant global compromise, as reflected by the increase in the first interstructure eigenvalue (RV.eig) and a more coherent redistribution of inter-block similarities [32,33]. These findings are consistent with multivariate robust statistical theory, which establishes that classical covariance estimators are highly sensitive to deviations from normality and to extreme values, potentially inducing distortions in factorial geometry and latent structure identification [34–36]. By contrast, the robust approach based on resistant estimators such as the robust Mahalanobis distance attenuates the influence of extreme observations without removing them, thereby preserving relevant system information [37,38]. This behavior aligns with previous results highlighting the superiority of robust methods in improving covariance stability and latent pattern detection in contaminated multivariate contexts [39,40]. The present study extends these principles to the domain of triadic multiblock analysis, demonstrating that robustness is particularly pertinent when variability is not an artifact but an intrinsic property of the environmental system under study [41,42].

5.2. Structural Interpretation and Block Behavior

Interstructure analysis reveals that Robust STATICO significantly reorganizes inter-block relationships without arbitrarily altering the overall structure [43,44]. In particular, the observed increase in the RV coefficient between Manta and Puerto Bolívar indicates that the classical approach was underestimating their structural similarity, likely due to the influence of local extreme observations affecting classical covariance estimation [45,46]. Robustification enabled the recovery of a more stable and consistent co-structure between these blocks, in agreement with the graphical evidence observed in RV heatmaps and comparative *kplots* [47].

Furthermore, analysis of block weights (*tabw*) shows that Robust STATICO introduces an adaptive weighting mechanism, slightly reducing the contribution of blocks with higher internal heterogeneity while reinforcing those whose structure is more coherent with the global compromise [46]. This behavior is theoretically consistent with the principles of multiblock analysis, where the objective is not to enforce uniform contributions but to maximize compromise stability and representativeness [47,48]. The cos^2 values reinforce this interpretation: whereas classical STATICO yields homogeneously high representation quality, potentially reflecting optimistic representations, the robust approach produces more discriminative cos^2 values [49]. The marked increase in representation quality for Manta and the decrease observed for La Libertad suggest that Robust STATICO corrects projections inflated by leverage effects, providing a more realistic assessment of each block's alignment with the common structure [50]. This result is consistent with previous studies indicating that robust weighting and representation schemes improve interpretability by emphasizing structurally consistent components across time or space [51].

5.3. Ecological and Methodological Implications

From an ecological perspective, Robust STATICO offers a more faithful and stable representation of environmental relationships in coastal systems characterized by high natural variability and potential measurement errors [52]. In tropical ecosystems such as those along the Ecuadorian coast, where variables such as salinity, turbidity, dissolved oxygen, and nutrients fluctuate due to seasonal processes, continental discharges, and oceanographic dynamics, the presence of localized anomalies is inevitable [52,53]. In this context, robust estimators enable the identification of persistent structural associations without allowing extreme events to distort the global interpretation of the system [54,55].

Graphical results, particularly the greater compactness of observation clouds, the more coherent alignment of variable gradients, and the increased dominance of the first compromise axis, demonstrate that Robust STATICO enhances the ecological readability of multivariate patterns [56]. This makes it especially suitable for environmental monitoring programs, water quality assessment, and the analysis of complex ecological systems, where data irregularity is not an exception but a structural condition [57,58].

From a methodological standpoint, integrating robust covariance estimators into the STATICO analytical workflow reinforces the need for hybrid approaches combining multiblock analysis and robust statistics [59,60]. Such approaches extend the scope of multivariate analysis toward tools that are more resilient to data imperfection and uncertainty, an essential aspect in applied environmental studies. Similar results have been reported in robust extensions of PCA and co-inertia analysis, evidencing a consolidated trend toward the systematic incorporation of robustness in applied multivariate inference [61].

5.4. Practical Implications and Applicability of Robust STATICO

From an applied perspective, the Robust STATICO proposed in this study provides a high-value analytical tool for contexts requiring the integration of multiple multivariate data blocks under real-world conditions of heterogeneity, structural noise, and influential observations [62]. Its practical utility is particularly evident in scenarios where classical multiblock methods may produce unstable compromises or distorted interpretations due to the sensitivity of traditional covariance estimators [63].

In environmental and water quality monitoring, Robust STATICO enables the joint analysis of physicochemical, biological, and operational variables measured across sites or time periods, ensuring that extreme events such as point discharges, intense climatic episodes, or measurement errors do not artificially dominate the identified common structure [64]. This is especially relevant for coastal and estuarine systems, where natural variability is high, and data often follow non-Gaussian distributions [65].

Moreover, the method is applicable to environmental management and decision-making programs by facilitating the identification of persistent structural patterns that can support:

(i) Early detection of changes in environmental quality;

(ii) Robust comparison among monitoring zones;

(iii) Integrated evaluation of management or intervention measures.

Operationally, Robust STATICO is particularly useful when a paired series of tables is available across time or space, as it synthesizes multiblock information into a stable and parsimonious compromise, reducing dimensional complexity without sacrificing interpretability [66,67]. This property makes it suitable for technical and academic institutions working with large volumes of heterogeneous environmental data and requiring reproducible, methodologically sound analyses [68,69].

Beyond the environmental domain, the proposed approach can be extended to other applied fields such as regional economics, environmental epidemiology, or territorial management, where robust integration of multiple information sources represents a recurring methodological challenge [70]. In these contexts, Robust STATICO provides an analytical framework that combines mathematical rigor, statistical stability, and practical utility, aligning with current needs in applied multivariate analysis for complex scenarios [71].

5.5. Advantages and Future Perspectives

The empirical evidence obtained in this study demonstrates that the proposed robust extension of STATICO does not introduce arbitrary modifications to the analytical framework but rather enhances its statistical coherence and geometric stability under

contaminated and heterogeneous data conditions. The robust formulation reorganizes inter-block similarities in a manner consistent with multivariate robust theory, reinforcing stable co-structures while attenuating distortions induced by influential observations [72]. This reorganization leads to a stronger dominance of the global compromise and a more parsimonious factorial representation, as evidenced by the increased concentration of inertia in the leading axes and the improved interpretability of block contributions.

Moreover, the adaptive weighting mechanism embedded in the robust procedure allows for a dynamic rebalancing of block influence, preventing structurally unstable datasets from disproportionately shaping the compromise space. This property is particularly relevant in environmental monitoring contexts, where extreme events and measurement variability are intrinsic to the system rather than anomalous artifacts. By integrating robust covariance estimation with triadic multiblock analysis, the proposed approach advances the methodological frontier of multitable statistical modeling and strengthens its applicability in complex sustainability assessments [73].

Looking forward, this framework opens promising research avenues. Future developments may explore the extension of Robust STATICO to dynamic monitoring systems, where evolving multiblock structures can be tracked longitudinally to detect regime shifts or structural transitions. Additionally, integration with advanced uncertainty modeling paradigms, including probabilistic, fuzzy, or neutrosophic extensions, may further enhance its capacity to represent indeterminacy and incomplete information in environmental datasets. In this sense, Robust STATICO should be understood not merely as a technical refinement, but as a platform for the next generation of resilient multiblock analytical methodologies [73].

6. Conclusions

This study presents a robust extension of the STATICO method specifically designed to improve the stability and interpretability of multiblock analyses under heterogeneous and contaminated data conditions. By integrating Minimum Covariance Determinant (MCD) estimators and adaptive weighting mechanisms into the triadic algebraic structure of STATICO, the proposed framework preserves the theoretical foundations of co-inertia and partial triadic analysis while strengthening its resistance to influential observations.

Comparative results between classical and robust formulations reveal meaningful structural differences. The robust approach enhances the dominance of the global compromise, increases inertia concentration in the leading axes, and produces a more coherent redistribution of inter-block similarities. These improvements translate into greater geometric stability and more discriminative representation quality, thereby reinforcing the reliability of the extracted multivariate structures.

From a methodological perspective, this contribution expands the scope of robust multivariate statistics toward triadic multiblock frameworks, an area that has received limited attention in the literature. From an applied standpoint, Robust STATICO offers a stable analytical instrument for sustainability-oriented environmental monitoring, particularly in systems characterized by high natural variability and episodic disturbances.

Overall, the proposed framework provides a statistically sound and environmentally interpretable tool capable of supporting sustainability assessment and decision-making in complex multivariate scenarios. By reinforcing compromise stability and mitigating distortion effects, Robust STATICO enhances the credibility and robustness of indicator-based monitoring systems in dynamic environmental contexts.

https://doi.org/10.3390/su18052607

6.1. Future Research Directions

Future research should extend the present methodological development toward empirical validation using long-term coastal monitoring datasets. Applying Robust STATICO to real multiyear environmental series will enable evaluation of its stability under seasonality, extreme climatic events, and nonstationary multivariate dynamics. Particular attention should be given to assessing the temporal persistence of compromise structures and trajectory consistency across monitoring campaigns.

Further methodological refinement may involve integrating Robust STATICO with statistical learning techniques, including robust clustering, multiblock segmentation, and predictive modeling frameworks. Such hybrid approaches could strengthen its capacity to support early-warning systems and adaptive environmental management strategies.

In addition, expanding the framework toward advanced uncertainty modeling paradigms represents a promising avenue. Incorporating probabilistic, fuzzy, or neutrosophic formulations may enhance the method's ability to represent partial knowledge, indeterminacy, and incomplete measurement features that are common in sustainability monitoring systems.

Finally, beyond environmental applications, Robust STATICO may be extended to interdisciplinary domains such as regional economics, epidemiology, and territorial governance, where the integration of heterogeneous multivariate data blocks poses similar structural challenges. These future developments position the proposed method as a versatile analytical tool within the broader landscape of resilient multiblock statistical methodologies.

6.2. Study Limitations and Scope

Although this study was conducted using statistically simulated data, these datasets were empirically calibrated to reproduce the ranges, correlations, and variability patterns observed in Ecuadorian coastal systems. Controlled simulation enabled rigorous evaluation of the methodological behavior of Robust STATICO under reproducible conditions, ensuring experimental validity and result traceability.

Nevertheless, simulation represents only a partial approximation of real ecological complexity. Accordingly, future research should apply the methodology to observational field data to validate its performance in dynamic, heterogeneous environments subject to uncontrolled disturbances.

Despite this limitation, the study makes a significant contribution to the development of robust multivariate analysis by proposing a tool capable of controlling the influence of outliers systematically generated in environmental, operational, and administrative measurement processes. In the Ecuadorian context, characterized by institutional technical constraints and high data variability, the availability of a methodology that enhances analytical stability and reliability constitutes a scientific and applied contribution of considerable value.

In this sense, Robust STATICO not only extends the theoretical boundaries of multiblock analysis but also provides a practical and methodologically sound response to real-world data management challenges faced by environmental, academic, and territorial planning institutions.

Author Contributions: Conceptualization, H.V.-C. and O.R.-B.; methodology, H.V.-C. and O.R.-B.; software, H.V.-C. and O.R.-B.; formal analysis, P.G.-V., O.R.-B.; investigation, H.V.-C.; resources, H.V.-C.; data curation, H.V.-C.; writing—original draft preparation, H.V.-C.; writing—review and editing, H.V.-C.; visualization, H.V.-C.; supervision, O.R.-B.; project administration, H.V.-C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request. The data are not publicly available due to institutional data management policies.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Basha, A. Advancements in robust least squares approximation techniques: A comparative analysis. *Math. Stat.* **2025**, *13*, 374–382. [CrossRef]
2. Barroso, R.; Pinto Gomes, N.L.; Oliveira, C. Assessment of water quality based on statistical analysis of physical-chemical, biomonitoring and land use data: Manso River supply reservoir. *Sci. Total Environ.* **2024**, *912*, 169554. [CrossRef]
3. Efron, B. Microarrays, empirical Bayes, and the two-group model. *Stat. Sci.* **2008**, *23*, 1–22. [CrossRef]
4. Hubert, M.; Rousseeuw, P.J.; Van, S. High-breakdown robust multivariate methods. *Stat. Sci.* **2008**, *23*, 92–119. [CrossRef]
5. Hubert, M.; Rousseeuw, P.J.; Vanden, K. ROBPCA: A new approach to robust principal component analysis. *Technometrics* **2005**, *47*, 64–79. [CrossRef]
6. Xu, Y.; Wang, Y.; Pan, Y. Prediction of China's unconventional natural gas production based on grey waveform prediction model and non-homogeneous exponential discrete model. *Energy Rep.* **2025**, *13*, 2829–2843. [CrossRef]
7. Nagwa, M.; Albehery Hend, A.; Auda, E.A. Robust L1-norm estimation for exploratory factor analysis model. *Math. Stat.* **2025**, *13*, 17–27. [CrossRef]
8. Wang, Y.; Wang, Y.; Zhao, M. Spatiotemporal Evolution and Driving Factors of Green Transition Resilience in Four Types of China's Resource-Based Cities Based on the Geographical Detector Model. *Sustainability* **2026**, *18*, 391. [CrossRef]
9. Al-Athari, F.M. A new robust interval estimation for the median of an exponential population when some of the observations are extreme values. *Math. Stat.* **2024**, *12*, 31–40. [CrossRef]
10. Dolédec, S.; Chessel, D. Co-inertia analysis: An alternative method for studying species–environment relationships. *Ecology* **1994**, *75*, 519–529. [CrossRef]
11. Fathy, H.R. Statistical inference of modified kies exponential distribution using censored data. *Math. Stat.* **2022**, *10*, 659–669. [CrossRef]
12. Lai, B.; Bernstein, D.S. Generalised forgetting recursive least squares: Stability and robustness guarantees. *IEEE Trans. Autom. Control* **2024**, *69*, 7646–7661. [CrossRef]
13. Wan, J.; Li, X. Analysis of a superconvergent recursive moving least squares approximation. *Appl. Math. Lett.* **2022**, *133*, 108223. [CrossRef]
14. Liu, K.; Guo, X. Fuzzy least squares approximation using fuzzy polynomial. *Math. Probl. Eng.* **2021**, *2021*, 9979544. [CrossRef]
15. Escoufier, Y. Opérateur associé à un tableau de données. *Annales de l'inséé* **1976**, *22/23*, 165. [CrossRef]
16. Plantes, L.d. *Structuration des Tableaux à Trois Indices de la Statistique*; Université de Montpellier II: Montpellier, France, 1976.
17. Lavit, C. *Analyse Conjointe de Tableaux Quantitatifs (Masson)*; Masson: Paris, France, 1988. Available online: https://books.google.es/books/about/Analyse_conjointe_de_tableaux_quantitati.html?id=9d1oAAAACAAJ&redir_esc=y (accessed on 10 January 2026).
18. Simier, M.; Blanc, L.; Pellegrin, F.; Nandris, D. Approche simultanée de K couples de tableaux: Application à l'étude des relations pathologie végétale-environnement. *Rev. Stat. Appliquée* **1999**, *47*, 31–46.
19. Thioulouse, J.; Simier, M.; Chessel, D. Simultaneous analysis of a sequence of paired ecological tables. *Ecology* **2004**, *85*, 272–283. [CrossRef]
20. Thioulouse, J. Simultaneous analysis of a sequence of paired ecological tables: A comparison of several methods. *Ann. Appl. Stat.* **2011**, *5*, 2300–2325. [CrossRef]
21. Thioulouse, J.; Chessel, D. Multi-table analysis of ecological factors. I. From description of ecosystem status to that of ecosystem function using triadic analysis. *Acta Oecologica Oecologia Gen.* **1987**, *8*, 463–480.
22. Rousseeuw, P.J.; Van Driessen, K. A fast algorithm for the minimum covariance determinant estimator. *Technometrics* **1999**, *41*, 212–223. [CrossRef]
23. Wilderjans, T.F.; Ceulemans, E.; Meers, K. CHull: A generic convex hull based model selection method. *Behav. Res. Methods* **2013**, *45*, 1–15. [CrossRef]
24. Abdi, H.; Williams, L.J.; Valentin, D. Multiple factor analysis: Principal component analysis for multitable and multiblock data sets. *Wiley Interdiscip. Rev. Comput. Stat.* **2013**, *5*, 149–179. [CrossRef]

25. Todorov, V. The R package ecosystem for robust statistics. *Wiley Interdiscip. Rev. Comput. Stat.* **2024**, *16*, e70007. [CrossRef]
26. Debruyne, M.; Verdonck, T. Robust kernel principal component analysis and classification. *Adv. Data Anal. Classif.* **2010**, *4*, 151–167. [CrossRef]
27. Fegatelli, D.A.; Tardella, L. Improved inference in capture–recapture models with behavioral effects. *Stat. Methods Appl.* **2013**, *22*, 45–66. [CrossRef]
28. Altwegg, R.; Salau, S.; Abadi, F.O. Emerging topics and new directions in statistical ecology. *J. Stat. Theory Pract.* **2025**, *19*, 44. [CrossRef]
29. Hevner, A.R.; March, S.T.; Park, J.; Ram, S. Design science in information systems research. *MIS Q.* **2004**, *28*, 75–105. [CrossRef]
30. Lang, D.J.; Wiek, A. Structuring and advancing solution-oriented research for sustainability. *Ambio* **2022**, *51*, 474–484. [CrossRef]
31. Liebscher, E.; Taubert, F.; Waltschew, D.; Hetzer, J. Modeling multivariate data using product copulas and minimum-distance estimators: An illustrative application to ecological traits. *Environ. Ecol. Stat.* **2022**, *29*, 315–338. [CrossRef]
32. Tenenhaus, A.; Tenenhaus, M. Regularized generalized canonical correlation analysis. *Psychometrika* **2011**, *76*, 257–284. [CrossRef]
33. Dray, S.; Chessel, D.; Thioulouse, J. Co-inertia analysis and linking ecological data tables. *Ecology* **2003**, *84*, 3078–3089. [CrossRef]
34. Filzmoser, P.; Maronna, R.; Werner, M. Outlier identification in high dimensions. *Comput. Stat. Data Anal.* **2008**, *52*, 1694–1711. [CrossRef]
35. Gregor, S.; Hevner, A.R. Positioning and presenting design science research for maximum impact. *MIS Q.* **2013**, *37*, 337–355. [CrossRef]
36. Misztal, M. Application of the Partial Triadic Analysis Method to analyze the crime rate in Poland in the years 2000–2017. *Folia Oeconomica Stetinensia* **2020**, *20*, 249–278. [CrossRef]
37. White, P.A.; Christensen, M.F.; Frye, H.; Gelfand, A.E.; Silander, J.A. Joint multivariate and functional modeling for plant traits and reflectances. *Environ. Ecol. Stat.* **2023**, *30*, 501–528. [CrossRef]
38. Jolliffe, I.T.; Cadima, J. Principal component analysis: A review and recent developments. *Philos. Trans. R. Soc. A* **2016**, *374*, 20150202. [CrossRef]
39. Smilde, A.K.; Bro, R.; Geladi, P. *Multi-Way Analysis: Applications in the Chemical Sciences*; Wiley: Hoboken, NJ, USA, 2004. [CrossRef]
40. Dray, S.; Dufour, A.-B. The ade4 package: Implementing the duality diagram for ecologists. *J. Stat. Softw.* **2007**, *22*, 1–20. [CrossRef]
41. Borcard, D.; Gillet, F.; Legendre, P. *Numerical Ecology with R*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2018. [CrossRef]
42. Wu, J. Landscape sustainability science: Ecosystem services and human well-being in changing landscapes. *Landsc. Ecol.* **2013**, *28*, 999–1023. [CrossRef]
43. Niemeijer, D.; de Groot, R.S. A conceptual framework for selecting environmental indicator sets. *Ecol. Indic.* **2008**, *8*, 14–25. [CrossRef]
44. Filzmoser, P.; Todorov, V. Robust tools for the imperfect world. *Inf. Sci.* **2016**, *354*, 199–213. [CrossRef]
45. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [CrossRef]
46. Vörösmarty, C.J.; McIntyre, P.B.; Gessner, M.O.; Dudgeon, D.; Prusevich, A.; Green, P.; Glidden, S.; Bunn, S.E.; Sullivan, C.A.; Liermann, C.R.; et al. Global threats to human water security and river biodiversity. *Nature* **2010**, *467*, 555–561. [CrossRef]
47. Hering, D.; Borja, A.; Carstensen, J.; Carvalho, L.; Elliott, M.; Feld, C.K.; Heiskanen, A.-S.; Johnson, R.K.; Moe, J.; Pont, D.; et al. The european water framework directive at the age of 10: A critical review of the achievements with recommendations for the future. *Sci. Total Environ.* **2010**, *408*, 4007–4019. [CrossRef]
48. Leigh, C.; Alsibai, O.; Hyndman, R.J.; Kandanaarachchi, S.; King, O.C.; McGree, J.M.; Neelamraju, C.; Strauss, J.; Talagala, P.D.; Turner, R.D.; et al. A framework for automated anomaly detection in high frequency water-quality data from in situ sensors. *Sci. Total Environ.* **2019**, *664*, 885–898. [CrossRef]
49. Chen, K.; Chen, H.; Zhou, C.; Huang, Y.; Qi, X.; Shen, R.; Liu, F.; Zuo, M.; Zou, X.; Wang, J.; et al. Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data. *Water Res.* **2020**, *171*, 115454. [CrossRef] [PubMed]
50. Yang, Q.; Scholz, M.; Shao, J.; Wang, G.; Liu, X. A generic framework to analyse the spatiotemporal variations of water quality data on a catchment scale. *Environ. Model. Softw.* **2019**, *122*, 104071. [CrossRef]
51. Tarr, G.; Müller, S.; Weber, N.C. Robust estimation of precision matrices under cellwise contamination. *Comput. Stat. Data Anal.* **2016**, *93*, 404–420. [CrossRef]
52. Croux, C.; Haesbroeck, G. Influence function and efficiency of the minimum covariance determinant scatter matrix estimator. *J. Multivar. Anal.* **1999**, *71*, 161–190. [CrossRef]
53. Maronna, R.A.; Martin, D.R.; Yohai, V.J.; Salibián-Barrera, M. *Robust Statistics: Theory and Methods*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2019. [CrossRef]
54. Lang, D.J.; Wiek, A.; Bergmann, M.; Stauffacher, M.; Martens, P.; Moll, P.; Swilling, M.; Thomas, C.J. Transdisciplinary research in sustainability science: Practice, principles, and challenges. *Sustain. Sci.* **2012**, *7*, 25–43. [CrossRef]

55. Hubert, M.; Debruyne, M.; Rousseeuw, P.J. Minimum covariance determinant and extensions. *WIREs Comput. Stat.* **2017**, *9*, e1356. [CrossRef]
56. Debruyne, M.; Hubert, M. The influence function of robust covariance estimators. *J. Multivar. Anal.* **2009**, *100*, 1291–1304. [CrossRef]
57. Salibián-Barrera, M.; Willems, G.; Zamar, R. Robust covariance estimation via fast S-estimators. *J. Comput. Graph. Stat.* **2008**, *17*, 700–718. [CrossRef]
58. Riani, M.; Atkinson, A.C.; Cerioli, A. Finding an unknown number of multivariate outliers. *JRSS B* **2009**, *71*, 447–466. [CrossRef]
59. Van Aelst, S.; Willems, G. Robust and efficient estimation of covariance matrices. *J. Multivar. Anal.* **2013**, *116*, 457–480. [CrossRef]
60. Rousseeuw, P.J.; Hubert, M. Robust statistics for outlier detection. *WIREs Data Min. Knowl. Discov.* **2011**, *1*, 73–79. [CrossRef]
61. Singh, R.K.; Murty, H.R.; Gupta, S.K.; Dikshit, A.K. An overview of sustainability assessment methodologies. *Ecol. Indic.* **2009**, *9*, 189–212. [CrossRef]
62. Hák, T.; Janoušková, S.; Moldan, B. Sustainable development goals: A need for relevant indicators. *Ecol. Indic.* **2016**, *60*, 565–573. [CrossRef]
63. Bougeard, S.; Dray, S. Supervised multiblock analysis. *J. Multivar. Anal.* **2018**, *165*, 38–52. [CrossRef]
64. Smilde, A.K.; Westerhuis, J.A.; de Jong, S. Framework for multiblock PCA. *J. Chemom.* **2003**, *17*, 323–337. [CrossRef]
65. Sala, S.; Ciuffo, B.; Nijkamp, P. A systemic framework for sustainability assessment. *Ecol. Econ.* **2015**, *119*, 314–325. [CrossRef]
66. Lyu, X.; Li, X.; Wang, K.; Dang, D.; Cao, W.; Lou, A. Coupling and coordination between ecosystem quality and sustainable development level in the West Liaohe River Basin, China. *Ecol. Indic.* **2024**, *165*, 112174. [CrossRef]
67. Wu, J. Landscape ecology, cross-disciplinarity, and sustainability science. *Landsc. Ecol.* **2006**, *21*, 1–4. [CrossRef]
68. Smilde, A.K.; Kiers, H.A.; Bijlsma, S.; Rubingh, C.M.; van Erp, M.J. Matrix correlations for high-dimensional data: The annualized RV-coefficient. *Bioinformatics* **2009**, *25*, 401–407. [CrossRef]
69. Alfons, A.; Croux, C.; Gelper, S. Sparse least trimmed squares regression for analyzing high-dimensional data. *Ann. Appl. Stat.* **2013**, *7*, 226–248. [CrossRef]
70. Lê, S.; Josse, J.; Husson, F. FactoMineR: An R package for multivariate analysis. *J. Stat. Softw.* **2008**, *25*, 1–18. [CrossRef]
71. Legendre, P.; Galzin, R.; Harmelin-Vivien, M.L. Relating behavior to habitat: Solutions to the fourth-corner problem. *Ecology* **1997**, *78*, 547–562. [CrossRef]
72. Dray, S.; Legendre, P. Testing the species–traits–environment relationships: The fourth-corner problem revisited. *Ecology* **2008**, *89*, 3400–3412. [CrossRef]
73. Salinas-Rodríguez, A.; Manrique-Espinoza, B.; Moreno-Tamayo, K. Robust multivariate methods in population health research. *Stat. Med.* **2020**, *39*, 3819–3835. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.