Science of the Total Environment 908 (2024) 168120

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Resilience and high compositional variability reflect the complex response of river waters to global drivers: The Eastern Siberian River Chemistry database

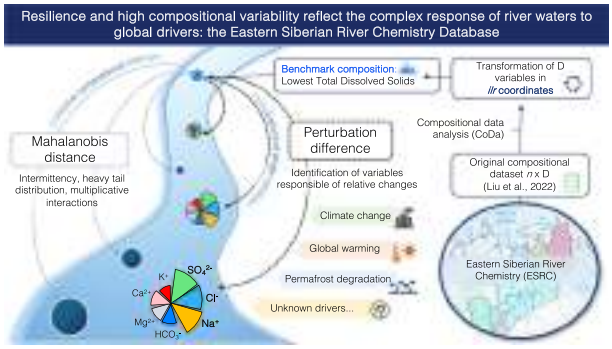
Caterina Gozzi [a,b,*], Antonella Buccianti [a,b,c]

[a] University of Florence, Dept. of Earth Sciences, Via G. La Pira 4, 50121 Firenze, Italy
[b] NBFC, National Biodiversity Future Center, Palermo 90133, Italy
[c] National Centre for HPC, Big Data and Quantum Computing, PNRR, Italy

HIGHLIGHTS

- Metrics for measuring changes from compositional benchmarks developed.
- Robust Mahalanobis distance intercepts environmental changes.
- Perturbation difference reveals compositional resilience.
- Heavy tails distributions are associated with instability and emergence of alternative states.
- Compositional resilience and variability mechanisms revealed in Eastern Siberian rivers.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Fernando A.L. Pacheco

Keywords:
Computation methods
Environmental geochemistry
Permafrost degradation
Variance
Climate change
Simplex geometry

ABSTRACT

The chemical composition of river waters represents an important matter of investigation to understand environment modifications in response to climate changes and global warming. Prolonged dry periods, heavy flood events, degradation of the lands and ice thawing, modify the chemical composition of river waters influencing the drivers governing the complex dynamics of river catchments where *everything comes together*. In this framework, Compositional Data Analysis (CoDA) offers methods in which the complex structure of the river water composition and the interrelationships among the various components are put into the proper context for their statistical analysis. In this research, we propose a new CoDA approach combining the robust Mahalanobis distance (*D*) calculus of ilr-transformed chemical variables and the perturbation difference, both with respect to a pristine compositional benchmark. The aim was to trace the change in the chemical composition of the Eastern Siberian River Chemistry Database where degradation of the permafrost for global warming produces important effects on natural waters. The findings indicate complex multiplicative laws and feedback mechanisms governing solutes in Eastern Siberian rivers, with high values of *D* found where permafrost is more discontinuous. Perturbations clearly discriminate chemical components more resilient to stresses induced by global changes (Ca^{2+}, Mg^{2+} and HCO_3^-) from those whose variability is not maintained under control (Cl^-, Na^+, SO_4^{2-}). These outcomes

* Corresponding author at: University of Florence, Dept. of Earth Sciences, Via G. La Pira 4, 50121 Firenze, Italy.
E-mail address: caterina.gozzi@unifi.it (C. Gozzi).

https://doi.org/10.1016/j.scitotenv.2023.168120
Received 30 August 2023; Received in revised form 20 October 2023; Accepted 23 October 2023
Available online 31 October 2023

open up a new scenario in searching for spatiotemporal resilience metrics to reveal rivers response to environmental changes.

1. Introduction

Complex dynamical systems can be characterized by tipping points at which it is possible to register significant, and sometimes catastrophic, changes (Scheffer et al., 2009; Grziwotz et al., 2023). Several examples can be found in literature for different situations, from ecosystems (Ardichvili et al., 2023) to financial markets and society (May et al., 2008; Lade et al., 2013). The way by which the systems reach such a threshold is an important field of investigation. Predicting critical transitions appears to be very difficult but fundamental to enhancing the resilience of natural and social systems. This difficulty is due to the absence of clear signals in the behavior of a variable used to monitor changes. Signals capable of anticipating the emergence of an abrupt shift and, consequently, the transition of the system from one state to another are difficult to find. Sometimes the study of the shape of the frequency distribution of a variable seems to be informative and the investigation of some statistics such as variance, skewness, kurtosis, as well as other more complex indices, are often used to detect in time or space the presence of critical transitions (Scheffer et al., 2012; Belle et al., 2017; Dakos and Kéfi, 2022).

In this perspective, the investigation of rivers water chemistry appears to be a challenge. Rivers are open thermodynamical systems able to exchange continuously matter and energy with the surrounding environment. They represent the response of the Earth's Critical Zone (ECZ) to external perturbations such as climate forcing and human pressure (Kirchner, 2009; Whitehead et al., 2009). By considering their position in the hydrosphere, rivers link geology and land features and represent the sentinels of any change due to global warming, land use or pollution. The Water Framework Directive (WFD, 2000/60/EC) requires EU Member States to achieve good status in all surface water bodies and groundwater by 2027. Environmental Quality Standards Directive are set to protect the most sensitive aquatic species, but also humans who, in turn, can be affected by secondary poisoning. Therefore, the WFD considers surface waters to be an integral part of the "One Health" project, which includes both natural and social systems. Chemical disequilibrium governs the water/rocks interaction processes thus distributing elements and chemical species among different reservoirs (Gaillardet et al., 1999; Shvartsev, 2013). Soil and rock weathering release cations such as Na^+, K^+, Si and Al^{3+} and anions (Cl^-, SO_4^{2-}, HCO_3^-). Solutes such as Cl^-, NO_3^-, and SO_4^{2-} can also arrive from other sources including rainfall, dust, and pollution (e.g., acid rain), although atmospheric deposition is considered low compared to weathering (Rogora et al., 2003, 2008). Existing literature has extensively investigated and documented river chemistry behavior under different weather conditions also considering seasonal changes and concentration–discharge relationships (Godsey et al., 2009, 2020; Zhi et al., 2020). Relations with climate control have also been investigated recently by Li et al. (2022) and Wang et al. (2024). Results indicate that, on a continental-scale, discharge predominantly controls mean concentrations of 16 solutes and that river chemistry is governed by the relative rates of solute addition (by reactions and input) and solute export. In a warmer climate, however, higher concentrations tend to deteriorate water quality even without human perturbations. Arctic and subarctic regions are particularly vulnerable to global warming, since freshwater-terrestrial systems will warm more rapidly than the global average (Prowse et al., 2006). Temperature and precipitation changes can stimulate modifications in the water balances and in several other aspects of the landscapes, e.g., river-ice breakup, increased infiltration capacity resulting from permafrost thawing, thermohaline circulation (Rouse et al., 1997; Prowse et al., 2015; Lique et al., 2016). Cold region rivers are often poorly studied, but a deeper understanding of their complex response to global changes is of pivotal importance. This is also the case of eastern Siberian rivers, where data and analysis of their environmental state are relatively sparse (Georgiadi et al., 2019; Kuzmin et al., 2014; Huh et al., 1998a, 1998b; Huh and Edmond, 1999). For these reasons, we decided to focus on the spatially extensive Eastern Siberian River Chemistry (ESRC) database to explore climate warming-induced changes in freshwater systems due to permafrost degradation. In order to accomplish this, we applied the principles of Compositional Data Analysis (CoDA) (Aitchison, 1982; Pawlowsky-Glahn et al., 2015; Pawlowsky-Glahn and Egozcue, 2020). CoDA methods have been applied to several research fields such as geology, biometrics, cultural heritage, medicine, with an exponential growth over the years. A recent bibliometric overview (Navarro-Lopez et al., 2022) highlighted that the main journals where CoDA is managed are indexed under the category of Geosciences and Multidisciplinary (23 papers), followed by Ecology (19), and Geochemistry and Geophysics (16). However, >40 research fields count papers that have used the CoDA approach.

The main objectives of this study are: i) to test new statistical methods, based on CoDA principles, for monitoring early changes in river water chemical composition, and ii) to assess the response of Siberian river chemistry to global changes. The proposed approach recognizes the strength of CoDA in taking into account the relationships among the components of the composition as required for investigating complex systems.

2. Materials and methods

2.1. The Eastern Siberian River Chemistry database

The Eastern Siberian River Chemistry (ESRC) database was used to test the possibility of identifying early warning signals able to check for modifications in the behavior of the analyzed variables as well as their resilience to changes (Dakos and Kéfi, 2022). The database was created by Liu et al. (2022) by merging 1434 multi-source data obtained from both published datasets and unpublished field studies. These include: (i) 159 datasets from the ArcticGRO water quality data (Holmes et al., 2021) and the GLObal RIver CHemistry (GLORICH) databases (Hartmann et al., 2014); (ii) 928 water chemistry data from 10 published studies in both English (Georgiadi et al., 2019; Kuzmin et al., 2014; Huh et al., 1998a, 1998b; Huh and Edmond, 1999), and Russian (Bochkarev, 1959; Shpakova, 1999; Berkin et al., 2009; Grebenshchikova et al., 2011; Sidorov, 1992); and (iii) 347 unpublished datasets provided by Gabysheva O.I. and Wang P. The dataset is accurately described in Liu et al. (2022) and publicly available at Springer figshare (Wang et al., 2022). The ESRC was constructed to improve the understanding of climate warming-induced changes in freshwater systems in permafrost-dominated eastern Siberia, where data are extremely scarce compared to western Siberia. It is important to note that the data do not cover all Eastern Siberia, since some basins are outside its boundaries (e.g., the Kolyma basin), while others (e.g., the Yenisey basin) are not included. However, for simplicity's sake and in accordance with Liu et al. (2022) and Wang et al. (2024), in the manuscript, we refer to the study area as Eastern Siberia. The database includes 9487 major ion concentrations (Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, SO_4^{2-}, HCO_3^-) in mg/L from 1434 water samples collected between May and August (84%) in eastern Siberia

spanning 1940–2019. Not all the parts of the compositions are known; missing values for K^+ are about 30.5 %, for other variables minor than 2.4 %. Consequently, in our analysis, only the complete data base was used, consisting of 846 cases. Climate variables (air temperature and precipitation data), discharges, trace element concentrations, and isotopes data at the basin and sub-basin scales are also provided in the ESRC database.

2.2. Geographical, climatic and geological setting

Eastern Siberia, as defined in this study, is located in the easternmost part of Eurasia and most of its area is located within the Arctic Circle and the sub-arctic region (Fig. 1). The average air temperature of eastern Siberia during 1940–2019 was - 10.6 ± 5.0 °C, and the average annual precipitation was approximately 312 ± to 92 mm (Wang et al., 2024). In the cold season, precipitation is dominated by snowfall with accumulation that begins in October and reaches its maximum in March after which the melts processes begin (Sazonova et al., 2004). The considered water samples mainly belongs to six large river basins of Eastern Siberia whose boundaries are illustrated in Fig. 1 and to a few more basins of smaller size (e.g., Khorbosunka, Olenyok). The Yana, Indigirka, and Kolyma river basins are located near the Artic Ocean in the northern part of the region, while the Angara and Selenga-Baikal basins are located near the Khangai Mountains in the southern part. The largest basin, the Lena, is one of the 15 largest rivers in the world in terms of length (4294 km), basin area, and average annual discharge. The studied rivers flow in the zone of northern taiga and tundra and the territory is underlain by an almost ubiquitous distribution of permafrost (about 90 % of the surface area) (Obu et al., 2019). This last is continuous to the north of the Vilyuy River, a major tributary to the Lena river, and discontinuous south of it. The uppermost permafrost horizon (~0.1–3.0 m) is subject to seasonal thawing in the warm season and freezing at subzero air temperatures (Gabysheva et al., 2022). An exceptional warmth was registered during the spring of 2020 over all Siberia (DeAngelis et al., 2023).

The study area covers the south-eastern part of the Siberian platform which is comprised by the Yenisey River on the west and the Lena River on the south and east (Meyer and Freeman, 2006). The Siberian Platform consists of a thick suite of siliciclastic (sandstones and shales) and carbonate (limestone and dolomite) rocks from the Mesoproterozoic to Mesozoic (Cocks and Torsvik, 2007; Barnet and Steiner, 2021). The central and southern portions of the Siberian Platform is also characterized by thick marine halite and gypsum sequences, which make this region one of the world's most promising areas for evaporitic industrial minerals (Steiner and Barnet, 2021). These deposits are associated with platformal sedimentation occurred from 550 to 500 Ma, extends from the west of Lake Baikal along the whole length of the Lena to Kempendyai in the Vilyui basin (Ronov, 1982), with also some exposed salts deposits (Huh et al., 1998a). On the northern edge of the Siberian Platform, diapiric salt domes of Devonian age can also be found around Port Nordvik on the Laptev Sea coast (Steiner and Barnet, 2021).

The easternmost part of the study area (Yana-Indigirka and Kolyma lowlands) is characterized by the Verkhoyansk–Kolyma folded area. This region is marked by the presence of tectonic structures, magmatism, and ore deposits. These are closely related to the Late Jurassic–earliest Early Cretaceous subduction–accretion and to the collision between the Kolyma–Omolon superterrane and the Siberian Craton (Parfenov and Kuzmin, 2001). The Yana–Kolyma gold belt is the largest in the Verkhoyansk–Kolyma metallogenic province with dikes varying from mafic to felsic compositions (Protopopov et al., 2019; Fridovsky et al., 2020).

Overall, according to the Global Lithological Map (GLiM) database (Hartmann and Moosdorf, 2012a, 2012b), nearly half of the study area consists of: i) siliciclastic sedimentary rocks (~23%) ii) mixed sedimentary iv) volcanic rocks (13 %) and, v) unconsolidated sediments (11 %) (Wang et al., 2024).

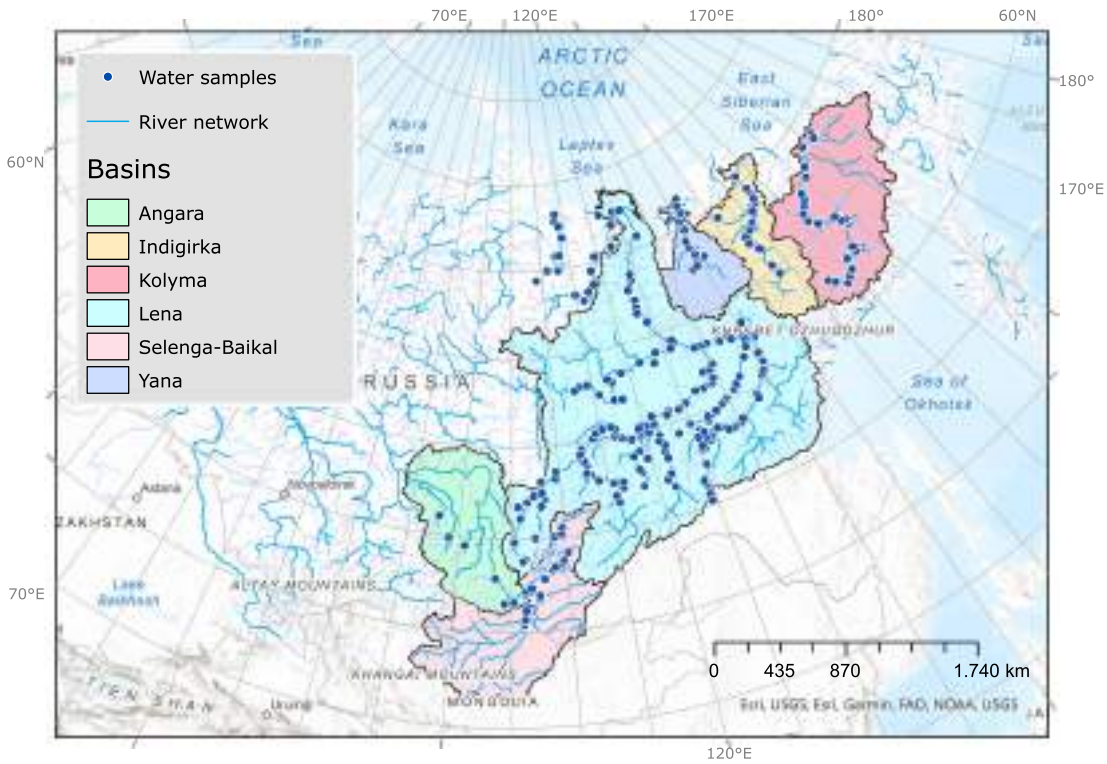


Fig. 1. Map showing the considered water samples locations of the Eastern Siberian River Chemistry (ESRC) database together with its six major river basins. The river network layer was derived from HydroBASINS and particularly from World Wildlife Fund's HydroSHEDS data (Lehner and Grill, 2013), while river basin boundaries were based on Liu and Huang (2022). Projected Coordinate System: WGS 1984 EPSG Russia Polar Stereographic.

2.3. Compositional data: what are they?

Compositional data describe parts of some whole. Most commonly, they are presented as vectors of proportions, percentages, frequencies, or concentrations. This is also the case of dissolved ion concentrations in water samples, which are expressed in mg/L, with liter as a references (Buccianti and Pawlowsky-Glahn, 2005; Sauro Graziano et al., 2020). The characteristics of compositional data prevent standard statistical techniques from being applied to such data in *raw* form (Pawlowsky-Glahn et al., 2015). Standard techniques are designed to be used with data that are free to range from -∞ to +∞, while geochemical data are always positive. As proportions it follows, then, that if one component increases, the other must decrease, whether or not there is a genetic link between them. It is tempting to think of them as real multivariate data, but this can lead to misinterpretations, some of which have been known since a century ago, but have been largely forgotten (Pawlowsky-Glahn and Egozcue, 2006). It is the relative variation of the components rather than their absolute variation that is important. This can be analyzed by considering log-ratios between the parts which are dimensionless or, alternatively considering the geometric structure of sample space.

2.4. The compositional data analysis approach

A n-part composition is a vector of n strictly positive real components, $[x_1 + x_2 + \cdots + x_n]$, such that $x_1 + x_2 + \cdots + x_n = \kappa > 0$ with κ generally equal to 100 (percentages) or 1 (proportions) or 10^6 for *ppm* and similar. The first insight about the nature of these type of data starts with Aitchison (1982) highlighting that log-ratios provide a one-to-one mapping from the constrained sample space of compositional data (called the simplex S^D) and the real space $\Re$. The algebraic-geometric structure of the simplex was recognized around the 2000s (Billheimer et al., 2001; Pawlowsky-Glahn and Egozcue, 2001a) leading to the proposal of the principle of working in coordinates (Mateu-Figueras et al., 2011). Compositions are thus represented by orthonormal coordinates (*ilr*, isometric log ratio coordinates) living in a real Euclidean space corresponding to the simplex sample space of compositions (Egozcue and Pawlowsky-Glahn, 2005).

One of the most important relapse of CoDA is the necessity to take into account the behavior of the whole composition since the analysis of isolate components leads to important misleading in the application of standard statistical methods, compromising the interpretation of natural phenomena. Therefore, in the logic of the dynamics of complex systems, CoDA appears to have its own role since the adequate management of the joint relationships among the parts of a composition becomes fundamental to track compositional changes.

Under CoDA, the tracing of compositional changes can be performed through the robust Mahalanobis distance from a given benchmark and then visualizing data related to river chemistry on a map. Before calculating the distance, data need to be transformed in *ilr* coordinates. In general, standard multivariate procedures assume that the majority of **Z** observations are generated by a multivariate normal distribution with mean vector μ and covariance matrix Σ. For a location estimator **t** and a covariance estimator **C**, the squared Mahalanobis distances between the observations expressed in coordinates and the respective location estimator **t**,

$$\mathrm{MD}(\mathbf{z}_i)^2 = (\mathbf{z}_i - \mathbf{t})' \mathbf{C}^{-1} (\mathbf{z}_i - \mathbf{t}), \qquad \text{for } i = 1, \ldots, n, \tag{1}$$

approximately follow a χ^2 distribution with $D-1$ degrees of freedom, χ^2_{D-1} (Filzmoser et al., 2018; Ronchetti, 2021) and a certain quantiles of it, like the 0.975-th quantile, $\chi^2_{D-1;0.975}$, can be used as a cut-off value to identify multivariate outliers (Filzmoser and Hron, 2008). The classical arithmetic mean (vector) and the (Pearson) sample covariance matrix are unsuitable as estimators **t** and **C** in Eq. (1) in the presence of outliers, since they are then themselves spoiled by the outliers. Thus, **t** and **C** should be robust estimators, like the MCD estimators $\mathbf{t}_{\mathrm{MCD}}$ and $\mathbf{C}_{\mathrm{MCD}}$ or the MM-estimator $\mathbf{t}_{\mathrm{MM}}$ and $\mathbf{C}_{\mathrm{MM}}$ (Maronna et al., 2006).

Another way to trace compositional changes is to use the perturbation difference, the group operation of the simplex S^D able to counterpart translation in real space (classical Euclidean statistics) and rotation on the sphere (circular statistics).

Given two D-part compositions $\mathbf{x}, \mathbf{y} \in S^D$, and $\oplus$ the symbol related to the application of the perturbation operator, $\mathbf{x} \oplus \mathbf{y}$ is defined by:

$$\mathbf{x} \oplus \mathbf{y} = \frac{[x_1 y_1, \ldots, x_D y_D]}{(x_1 y_1 + \ldots + x_D y_D)} = C[x_1 y_1, \ldots, x_D y_D] \tag{2}$$

where the "closure" operator C standardizes the perturbation vector by the sum of its components. Following Aitchison and Ng (2005) the operation $\oplus$ defines an Abelian group on the simplex with identity $e = (1/D)[1, \ldots, 1]$ and inverse $x^{-1} = C[x_1, \ldots, x_D]$. This mathematical property is fundamental to characterize the operation that changes a D-part composition **x** into a D-part composition **y**, also called perturbation difference (Pawlowsky-Glahn et al., 2015):

$$\mathbf{p} = \mathbf{y} \ominus \mathbf{x} = C\left[\frac{y_1}{x_1}, \ldots, \frac{y_D}{x_D}\right]. \tag{3}$$

The perturbation, together with the powering operation, analogous to scalar multiplication in real space, both defined in S^D, satisfy the requirements for operations of a vector space with the definition of an inner product, a norm and a distance (Billheimer et al., 2001; Pawlowsky-Glahn and Egozcue, 2001b). Consequently, any measure of *difference* between compositions **x** and **y** must be expressible in terms of perturbation difference (hereafter called only perturbation). In geochemistry, the perturbation approach finds easy application when a pristine composition $\mathbf{x_0}$ is subjected to a sequence of perturbations $\mathbf{p_1}, \ldots, \mathbf{p_n}$ to reaching its current state $\mathbf{x_n}$.

The two approaches are fundamentally different since the Mahalanobis distance is represented for each case by a single positive number while data related to perturbation are parts of a vector with D components, if D is the number of the analyzed variables. Thus in the first case the statistics to be monitored is the distance between compositions, while in the second case it is possible to verify which variables are more responsible of the relative change. The comparison of the results on a spatial scale appears to be very useful to identify variables and areas under stress.

In both cases (Mahalanobis distance and perturbation), the water composition characterized by the lowest Total Dissolved Solid (TDS) value in the ESRC data base was chosen as a benchmark to test the proposed method. The selected composition is drawn from the Indigirka basin and has a TDS of 12.59 mg/L ($Na^+ = 0.12$, $K^+ = 0.13$, $Ca^{2+} = 1.86$, $Mg^{2+} = 0.72$, $Cl^- = 0.19$, $SO_4^{2-} = 0.31$, $HCO_3^- = 9.25$). The maximum value of TDS is equal to 2142 mg/L from Lena basin. The difference from this compositional benchmark is then spatially monitored to discover areas and variables experiencing a major rate of change or more resilient behavior (Gozzi et al., 2020; Sauro Graziano et al., 2020). Alternatively, the reference composition can be given by other benchmarks as for example the robust mean of the database or other interesting compositions for defined purposes.

3. Results

3.1. Robust Mahalanobis distance from the lowest TDS-benchmark

For an overview of the robust Mahalanobis distance from benchmark composition over time, box plots are presented illustrating the D values aggregated by year from 1990 to 2018 (Fig. 2). A detailed statistical analysis of the time series was here prevented due to the sparse and uneven distribution of the data. Box-plots show small oscillations over time, with years characterized by higher (e.g., 1992–1995) and lower variability (e.g., 2006–2008). However, an increase in the distance can

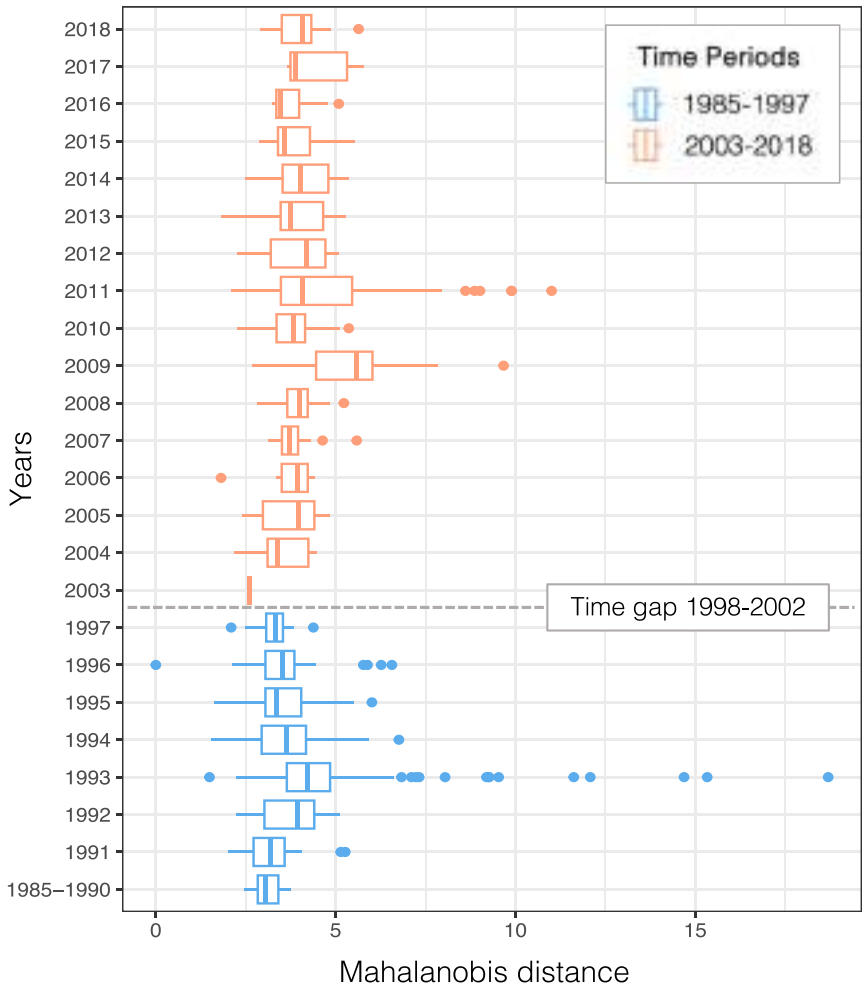


Fig. 2. Boxplot of Mahalanobis distances over the sampling years. The dashed line marked a time gap of four years in the dataset.

be observed over time, with a larger interquartile range around the median especially after 2009. Spearman correlation coefficient (R_s) between yearly medians of D and yearly medians of temperature, precipitation and potential evaporation records was also computed, but no significant associations were found (R_s=0.059, R_s=0.14 and R_s=0.33, respectively).

The robust Mahalanobis distance D is reported in Fig. 3 by considering the permafrost zonation modeled by Obu et al. (2018) to evaluate potential association of compositional changes in rivers with permafrost degradation. Each grid cell in the map is classified as continuous, discontinuous, sporadic, or isolated permafrost based on its permafrost probability. High values of D seem to be more common within the Lena basin where permafrost is more discontinuous or sporadic, even though few exceptions are found in the northern sector located in continuous permafrost and, in particular, for the Indigirka basin.

The Cullen and Frey graph reported in Fig. 4 suggests for D the presence of an asymmetric density distribution function, with data located in the field between the lognormal and gamma distributions.

The plot compares distributions in the space of the squared skewness and kurtosis synthesizing with these two statistics the properties of the probability distributions in a first attempt (Cullen and Frey, 1999). However, since skewness and kurtosis are not robust, to take into account the uncertainty of the estimated values a non parametric bootstrap procedure is performed by using the R (R Development Core Team, 2023) package *fitdistrplus* (Delignette-Muller and Dutang, 2015). The histogram in the upper-right corner of the plot (Fig. 4) also indicates a right-asymmetry of the distribution with a long tail. Values >5 mainly pertain to the Lena basin (81 cases), Yana (27), Indigirka (21) followed by Kolyma (8), other basins (7) and Selenga-Baikal (6). Accordingly, the first three basins (Lena, Yana, and Indigirka) are the ones which have experienced the greatest compositional changes relative to the benchmark during the considered time-span.

3.2. Perturbation difference from the lowest TDS-benchmark

Perturbation is another way to measure the compositional changes from a given compositional benchmark. While the robust Mahalanobis distance is represented by a single value for each case, perturbation is given by a vector where the relative changes undergone by each variable are reported. If the data are ranked for TDS values the content of the perturbation vector can be visualized by using comparative sequential plots as reported in Fig. 5, splitting the percentage contribution of the perturbation for each variable.

It is clear that the compositional changes from the chosen compositional benchmark tend to have a relative big effect for Na^+, Cl^- and SO_4^{2-} (reaching 90 % of the responsibility of the compositional variation) while K^+, Ca^{2+}, Mg^{2+}, and HCO_3^- undergo moderate and similar changes, all mainly under 25 % with respect to the other parts of the composition. It is also noteworthy the presence of a light decreasing trend in the perturbation intensity for K^+ with the increase of the *TDS* values, while an opposite behavior is recognized for Cl^- and, potentially, for Na^+. However, the results of the Runs Test, a statistical procedure for detecting non-randomness (Ronchetti, 2021), are significant only for Cl^- ($p < 0.05$, test performed on non-closed factors (Egozcue

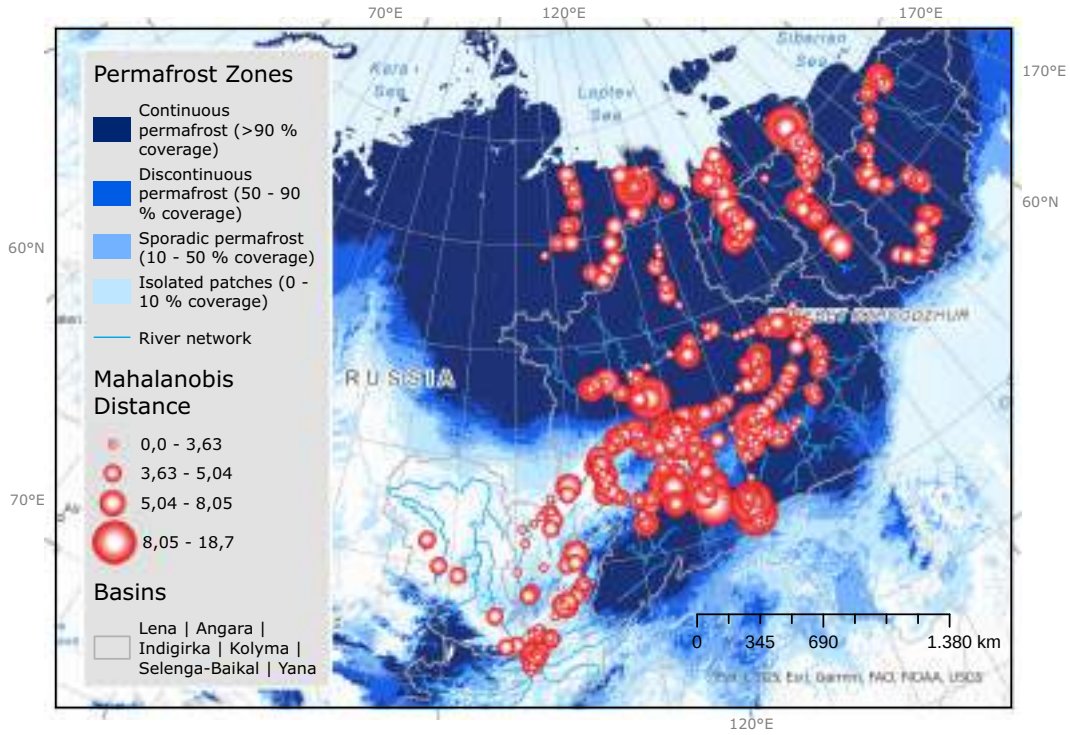


Fig. 3. Map of the Mahalanobis distance values along with the permafrost zonation proposed by Obu et al. (2018). Values of Mahalanobis distance are represented using graduated symbols and class breaks are determined using the natural breaks (Jenks) classification in ArcGIS Pro.

and Pawlowsky-Glahn, 2011)), confirming its role as conservative ion. Perturbation values for Na^+, Cl^- and SO_4^{2-} exhibiting higher relative variability and intermittency behavior with respect to Ca^{2+}, Mg^{2+}, K^+ and HCO_3^-, reveal conditions more prone and sensitive to environmental drivers and conditions, including seasonal variations (Scheffer et al., 2012).

In Figs. 6 and 7 the results of the percentage perturbations are mapped along with the spatial distribution of main lithotypes (Hartmann and Moosdorf, 2012a, 2012b), grouping the variables based on their similar behavior as deduced from Fig. 5.

For Ca^{2+}, Mg^{2+}, K^+ and HCO_3^- the relative percentage perturbation it is never higher than 40 % with the lower values pertaining to parts of the Lena river basin. Compositional changes are very similar for Ca^{2+}, Mg^{2+} and HCO_3^-, most likely representing what occurs to the carbonate cycle in terms of a higher contribution due to the evolution of the water/rock interaction processes. Particularly high Mg^{2+} perturbation values can be noticed especially along the Alder and Amga rivers. In this framework, K^+ shows its higher relative perturbation values in the southern parts of the Angara and Lena basins and in the Selenga-Baikal where plutonic rocks outcrops. Several hot spots are also found in the northern river basins located near to the Arctic Ocean.

By considering Na^+, Cl^- and SO_4^{2-} (Fig. 7), we can observe that the relative perturbation values sometimes exceed 90 % for Na^+ and SO_4^{2-}, while reaching a maximum of 59 % for Cl^-. However, while Cl^- and Na^+ show an uneven spatial distribution, SO_4^{2-} appears to have almost a similar behavior everywhere throughout Eastern Siberia, thus representing the chemical component most affected by compositional changes. In fact, about 244 cases on 846 show perturbation values higher than 50 % and 687 than 20 %. On the contrary, high Cl^- perturbation levels are monitored along the Lena and Vilyuy river's course and on the coast of the Laptev Sea, whereas for Na^+ high values are found mostly on exposed plutonic rocks south of Lake Baikal.

4. Discussion

Chemical weathering involves a complex system of reactions that move elements from different reservoirs such as atmosphere, hydrosphere, lithosphere, thus determining the hydrogeochemistry of river waters and the biogeochemical cycles of elements (Kleidon, 2010). These processes, governed by non-equilibrium thermodynamics, act as a "geological thermostat" affecting the global carbon cycle and log term climate stability (Zheng et al., 2023). In fact, continental weathering consumes atmospheric CO_2 and, when chemical processes are enhanced (e.g., uplift of mountain ranges), global cooling can be registered in the geological record, as happened in the Tibetan Plateau during the Cenozoic (Hren et al., 2007). Dissolved load in the river water is mainly generated by rock weathering, atmospheric inputs and anthropogenic activities, all these factors affecting the ECZ evolution. With global warming and the continuous increase of the human pressure on hydrological resources, it is fundamental to examine river chemistry to understand where processes exhibit more rapid changes or potential regime shifts (Gozzi et al., 2021; Gozzi and Buccianti, 2022).

The proposed method, developed under the CoDA framework, combines the calculus of the robust Mahalanobis distance with the perturbation difference both computed from a selected compositional benchmark. The method was applied to the ESRC database (Liu et al., 2022) including 846 data belonging to a severely stressed area due to permafrost degradation caused by global warming. For each sample major ion concentrations (i.e., Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, SO_4^{2-}, HCO_3^- in mg/L) were measured. In order to monitor the change from a pristine composition, that with the lower *TDS* value was used as a reference term. Alternatively, depending on the research objectives, any other composition of interest may be selected, including compositions based on law values related to health.

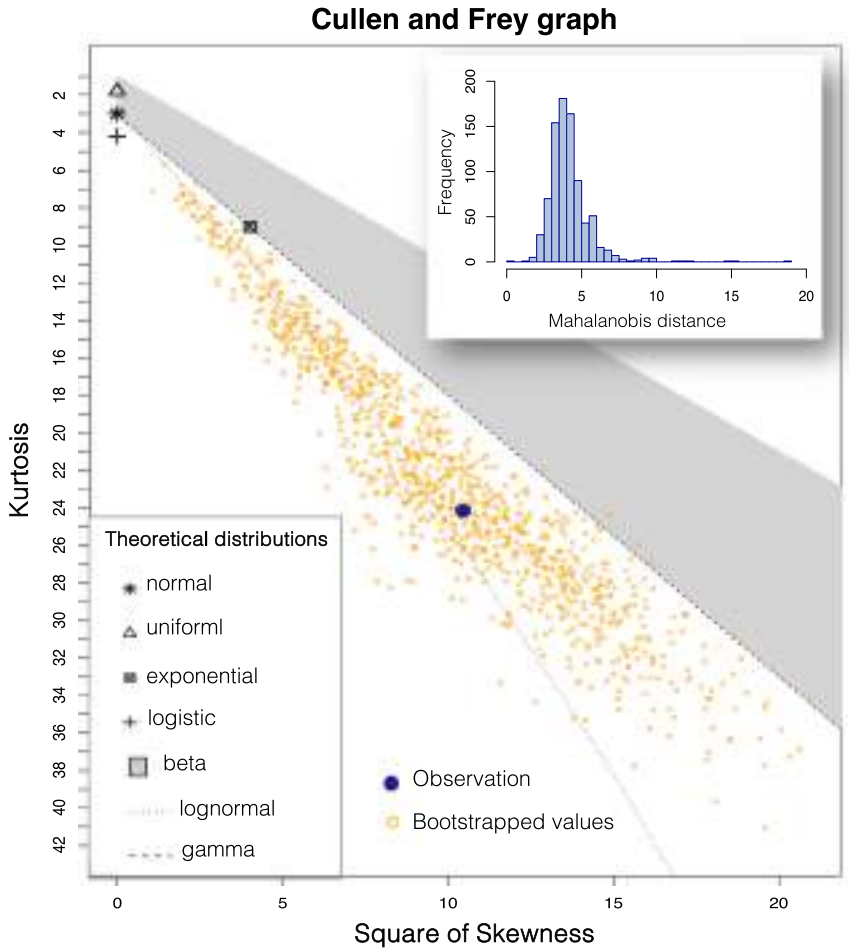


Fig. 4. Cullen and Frey graph (R package *fitdistrplus*) and histogram of robust Mahalanobis distance (upper-right corner).

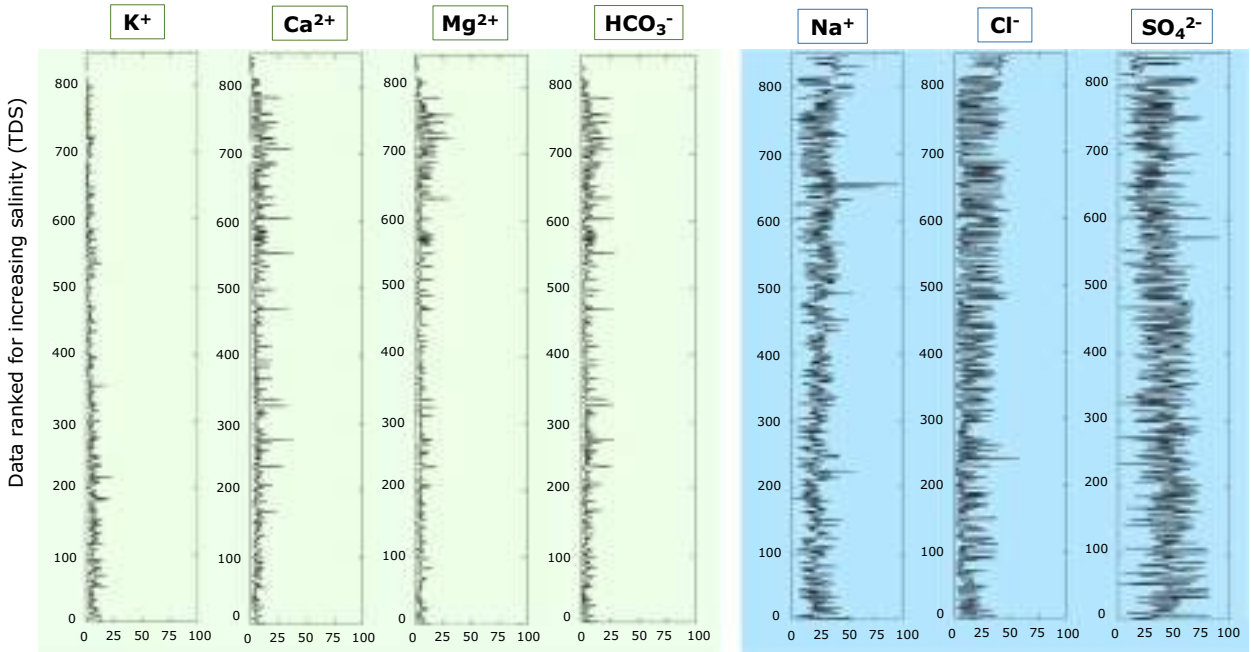


Fig. 5. Perturbation values ranked for increasing TDS values. Variables with similar variability are represented in different colors: light green - low variability (K^+, Ca^{2+}, Mg^{2+}, and HCO_3^-); light blue - high variability (Na^+, Cl^- and SO_4^{2-}).

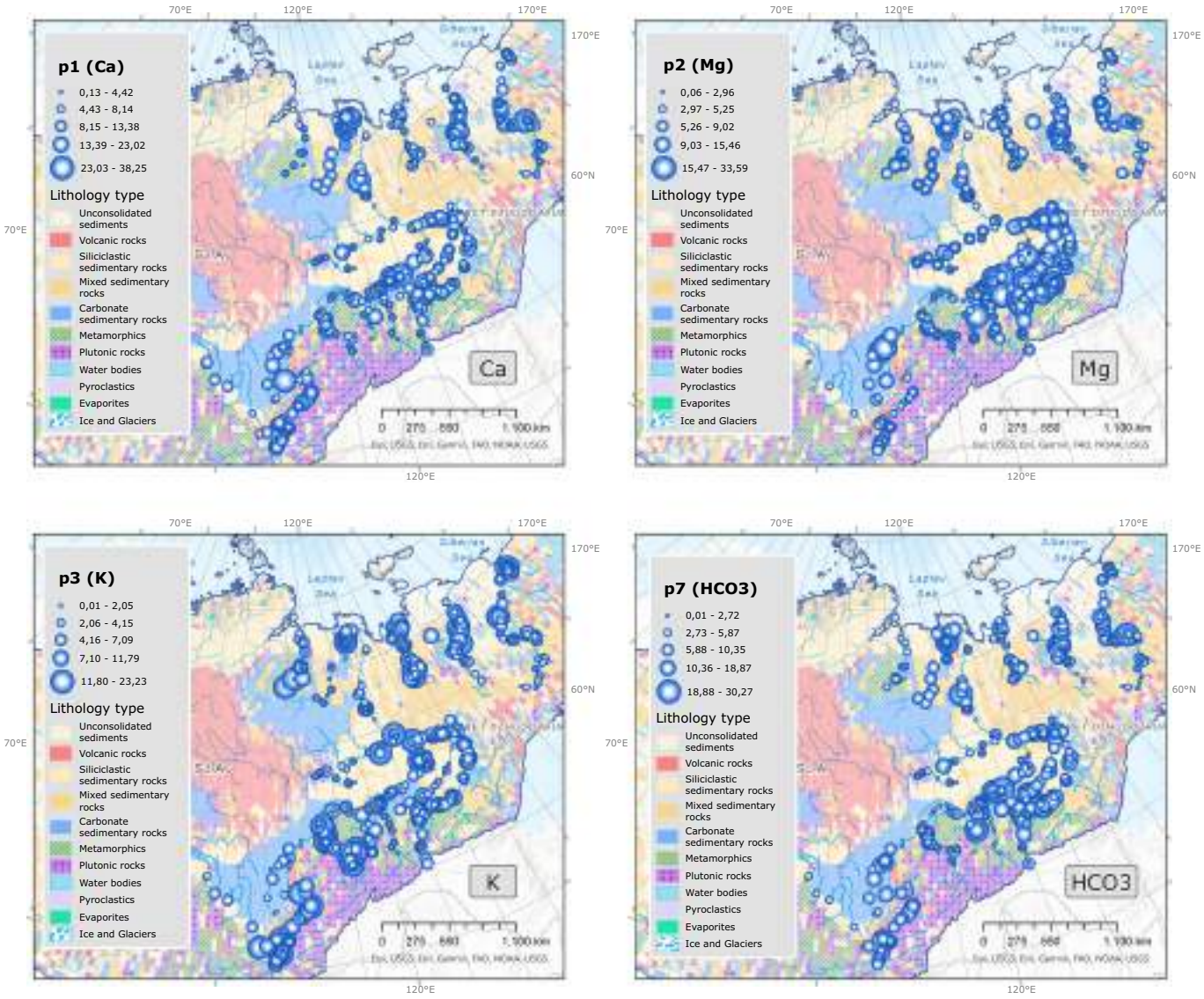


Fig. 6. Spatial behavior of perturbation values p_i for Ca^{2+}, Mg^{2+}, K^+ and HCO_3^-. Lithology types are derived and modified from the Global Lithological Map Database GLiM v1.0 (gridded to 0.5° spatial resolution) (Hartmann and Moosdorf, 2012a, 2012b). Projected Coordinate System: WGS 1984 EPSG Russia Polar Stereographic.

4.1. Distribution shape of Mahalanobis distance as resilience index

The map of the robust Mahalanobis distance values D allowed to obtain a first exploratory analysis about the whole intensity of the compositional changes compared with the permafrost type distribution (Obu et al., 2019). High values of D seem more common where permafrost is discontinuous or sporadic, such as in the Lena river basin (Fig. 3), indicating a potential role of permafrost degradation in increasing the *TDS* of river waters. In fact, when permafrost degrades, infiltration and groundwater storage increase, and flow paths become deeper, allowing highly mineralized groundwater to contribute more to streamflows (Walvoord and Striegl, 2007; Frey et al., 2007). In support of these findings, Wang et al. (2024) also found 2.3 times higher *TDS* in eastern Siberia river basins predominantly free of permafrost than in permafrost-dominated river basins. The few exceptions found in the northern sector of continuous permafrost could be instead associated to local process of salinization due to marine aerosols or presence of diapiric salt domes (e.g., Port Nordvik; Steiner and Barnet (2021)). The analysis of the shape of the frequency distribution of the D values have permitted to understand the mechanism generating the compositional changes. The Cullen and Frey plot of Fig. 4 confirms the presence of an asymmetric shape, with data located in the field between the log normal and gamma distributions (Aitchison and Brown, 1957). The presence of an asymmetrical distribution allows to exclude physical-chemical mechanisms governing the mixing of multiple sources, generally leading to a normal distribution (Andersson, 2021). This probability model describes fluctuations around a barycenter revealing a system that, notwithstanding the original heterogeneity, has achieved a homogeneity in river water composition (e.g., efficient mixing, long time of action; Allègre and Lewin (1995)). Conversely, the D probability model is more complex suggesting that the original heterogeneity continues to have a significant role and that homogeneity is not achieved. At the base of the emergency of asymmetry are multiplicative processes where non-linearity, feed-backs and heterogeneous interactions are the rule, often leading to a mixing between log-normality for the body of the distribution and power law for its right tail (van Rooij et al., 2013). This condition well describes the distribution of D values. Indeed, the Kolmogorv-Sminorv test for the log normality of all the data cannot be accepted ($p < 0.05$), whereas the fitting of a power law distribution to D indicates that, for $D_{min} > 3.87$ and $\alpha = 5.42$ (the scaling exponent), an heavy tail appears plausible (Fig. 8). The estimates were performed according to the goodness-of-fit-based method described in Clauset et al.

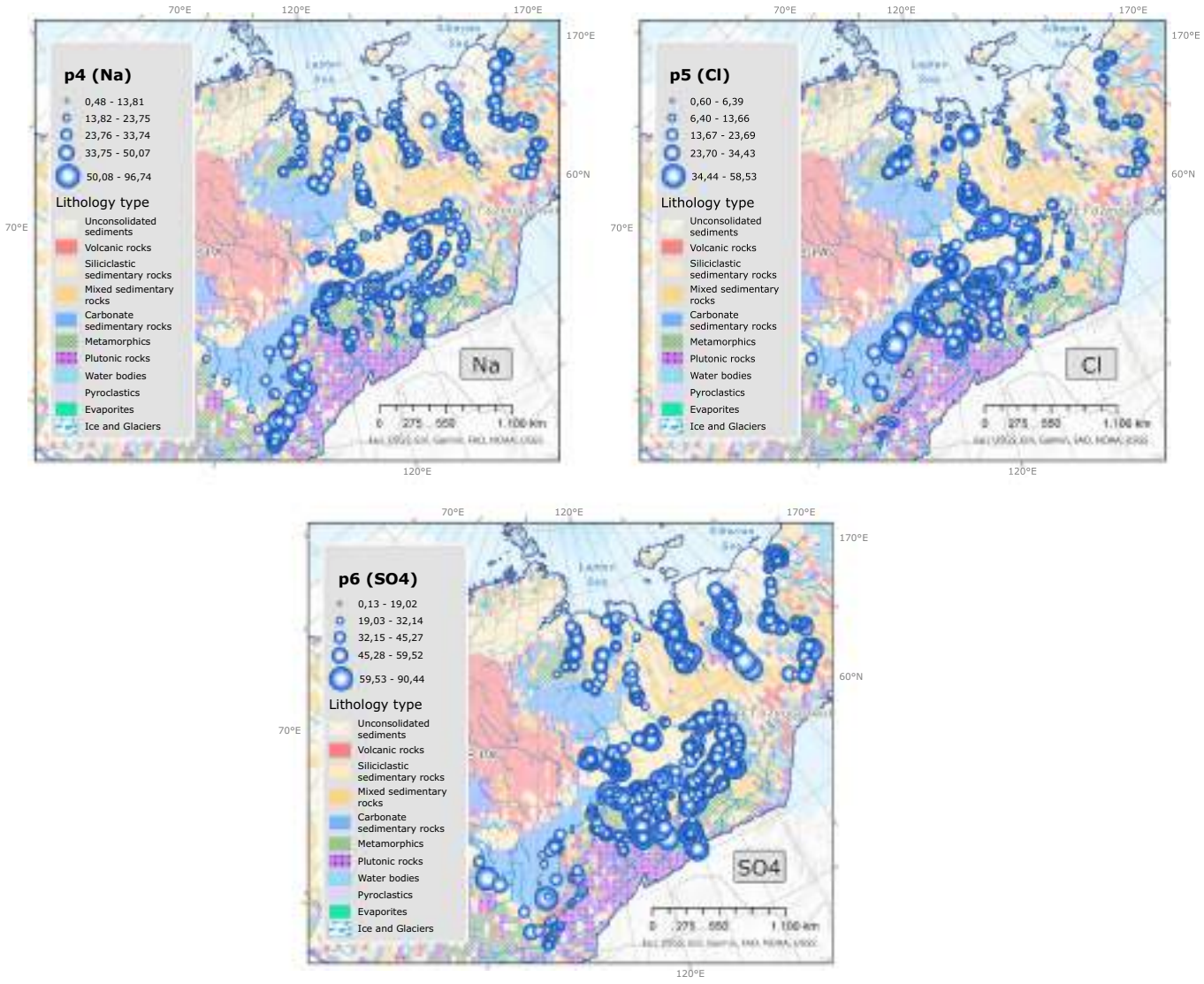


Fig. 7. Spatial behavior of perturbation values p_i for Na^+, Cl^- and SO_4^{2-}. Lithology types are derived and modified from the Global Lithological Map Database GLiM v1.0 (gridded to 0.5° spatial resolution) (Hartmann and Moosdorf, 2012a, 2012b). Projected Coordinate System: WGS 1984 EPSG Russia Polar Stereographic.

(2009). In this regard, it is worth remembering that a non negative random variable X is said to have a power law distribution if $Pr[X \geq x] \sim cx^{-\alpha}$ for $c > 0$ and $\alpha > 0$. Hence, in a log-log plot of $Pr[X \geq x]$, also known as the *cumulative distribution function*, asymptotically the behavior will be a straight line (Mitzenmacher, 2004).

Most of the data for which $D_{min} > 3.87$ pertaining to the Lena basin (254), Kolyma (78), other basins (45), Yana (33), Indigirka (23), Selenga-Baikal (18) and Angara (6). Compositional changes for Eastern Siberian rivers are thus characterized by multiplicative interactions. This means that the step of the compositional process modifying the benchmark is proportional to the benchmark itself, independently from the compositional differences. However, when D exceeds a certain threshold, feed-back mechanisms could increase concentration values hindering dispersion, generating heavy tail distributions. The occurrence of log normality mixed to power law reduces the ability of a system to resist or bounce back to the conditions that precede a possible shift since the depth of the *attractor basin* of the data tends to be lower, depending on the heaviness of the tails (Scheffer et al., 2012). Considering the density distributions of D separated according to the two available time periods (Fig. 9a), it is apparent that the tail has become heaviest in more recent years as a result of the appearance of a new basin of attraction (mode) in the data. The likelihood of bi-(multi)-modality can be considered a proxy of bistability and potential flipping between alternative states (Dakos and Kéfi, 2022).

In more detail, when an histogram with a normal distribution shape is upside down (Fig. 9b), the hole, if deep, represents the sink or attractor state from which is difficult to exit without external forcing (stable state). On the other hand, when an asymmetrical shape (log-normal or power law or their mixing) governs the frequency, slip planes or low thresholds can badly divide alternative stable states thus favoring a transition from one condition to another. This might generate flickering between the available states and increasing variance, triggering potential regime shifts (Wilhelm and Hanggi, 2003; Scheffer et al., 2012). In summary, the results highlight that processes able to influence solutes in Eastern Siberian rivers are acting in a fragmented way, generating intermittency and local variability, as well as a discontinuous spatial behavior (Colombo et al., 2018). Additionally, the heaviest tail of the distribution and the presence of bimodality suggests a lower resilience of the system to regime shifts in more recent periods, potentially linked to enhanced permafrost degradation caused by global warming (Wang et al., 2024). Global climate change and warming are the main anthropogenic impacts in this very remote region, indicating human impacts have no borders and far-reaching consequences worldwide. Although direct human activities in the Arctic and sub-arctic are

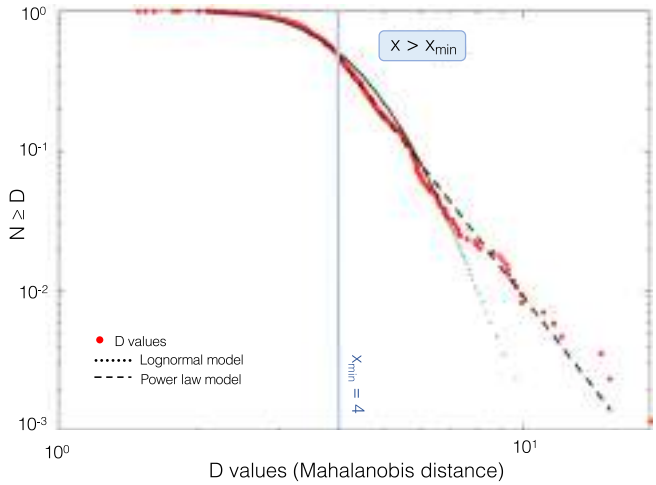


Fig. 8. Complementary cumulative distribution function for the $Pr[X \geq x]$ in a log-log plane (Mitzenmacher, 2004). The dotted line is the log normal model with mean and variance of the *D* calculated data; the dashed line represents the power law model for $x \geq 3.87$; the vertical light blue line is posed for $x = 4$.

generally assumed to be modest, they can interact with climate or environmental changes to either mitigate or exacerbate the human impacts (Huntington et al., 2007).

4.2. Perturbation differences

The calculation of the perturbation difference vector for each sample from the chosen benchmark provides information about the behavior of each variable within the composition. The sequential plots of the Fig. 5 combined with the maps of the perturbation difference values (Figs. 6 and 7) reveal that Ca^{2+}, Mg^{2+} and HCO_3^- are mostly stable with a moderately increasing variability when the *TDS* is higher. An opposite behavior is instead observed for K^+ with a slightly higher variability for lower *TDS* values. On the other hand, Na^+, Cl^- and SO_4^{2-} exhibit high variability and strongly intermittent behavior. Large fluctuations in an intermittent process are not a rare event such as in the Gaussian process and contribute to define the statistical moments of a probability distribution (Abramenko, 2014). In this case, the density distribution function is no longer determined by the first and second statistical moments (mean and variance), and high-order moments play a crucial role in shaping the heavy tails, generating multifractality (Zuo and Wang, 2016). The obtained results appear to indicate that the responsibility of the complex density function governing the *D* behavior (log normal plus power law) and, consequently, the compositional changes affecting the Eastern Siberia river waters, are primarily attributable to the behavior of the second group of variables (i.e. Na^+, Cl^- and SO_4^{2-}).

A wide literature suggests that the increase in precipitation induced by climate warming is the primary driver for the increase of Siberian river discharge and that concentrations of major ions tend to increase following permafrost degradation (Colombo et al., 2018; Feng et al., 2021; Liu et al., 2022). Based on our results, we can say that these effects generate high variability in Na^+, Cl^- and SO_4^{2-} with respect to Ca^{2+}, Mg^{2+}, K^+ and HCO_3^-. By comparing compositional changes to a defined benchmark (e.g., lowest TDS-benchmark), the different response of the variables is clearly evident.

The chemical stability of the components of the carbonatic cycle (Ca^{2+}, Mg^{2+} and HCO_3^-) likely indicate adaptive capacity (higher resilience) to change since all variations appear to be smoothed out. This behavior is also monitored in other river basins (e.g., Gozzi et al., 2018; Gozzi and Buccianti, 2022), and could be due to the rapid kinetics of dissolution/precipitation processes and mineral solubility that contrast the effect of the increasing discharge. The behavior of K^+ is generally complex, governed by silicate weathering, adsorption and decomposition processes. The vegetation uptake, together with the solubility of clay minerals, in general pose some limits to its concentration in river water but sensibility to flood event is well documented (Lucas et al., 2011). In western Siberia, Pokrovsky et al. (2015) reports evidence of the importance of plant litter and ground vegetation leaching, comparing K^+ concentrations as a function of latitude and finding an important relationship with the seasons. Hence, the continuous exchange between abiotic and biotic processes could maintain the variability of the elements under control. Differences in the behavior of the ions have been also discussed in Knapp et al. (2020) by the analysis of 30 events extracted from a 2-year time series of sub-hourly stream water solute measurements in Switzerland. It was demonstrated that the variability in concentration–discharge relationships among different hydrologic events and solutes could be linked to a range of environmental controls. On the whole, variables apparently better involved in bio-geochemical cycles seem more efficient in maintaining their variability inside defined ranges while for the others sound as if no successful tools are available to opposite to the concentration increase in waters.

On the other hand, the intermittency of Na^+, Cl^- and SO_4^{2-} opens a complete different scenario where the increase in variability and rate of change could anticipate regime shifts or critical transitions (Beck et al., 2018). The behavior of Na^+ is generally affected by silicate weathering, salts dissolution and cation exchange processes. The high variability registered in the relative compositional changes could be attributable to the multiples sources of the cation in the investigated area, spanning from silicate rocks, salts, mineralized groundwaters, sedimentary sulfide rocks and soil salinization phenomena, this last particularly in the cryoarid territories of the south of the Eastern Siberia (Chernousenko, 2021). Once it reaches the solution Na^+ tends to stay, increasing its concentration. However, permafrost degradation, involving an higher

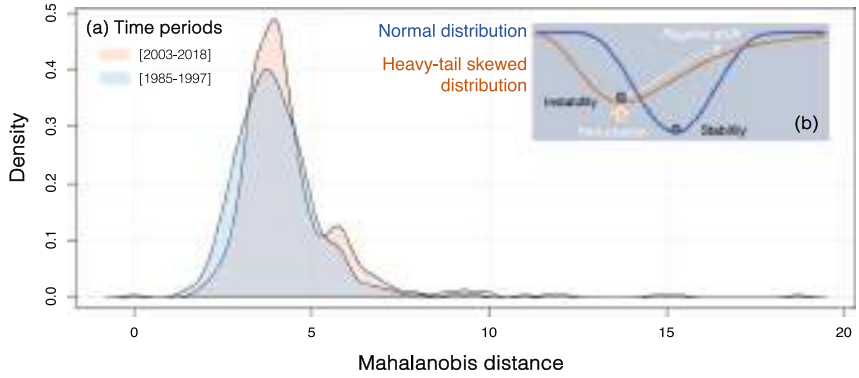


Fig. 9. a) Comparison of density plots for the two available time periods and b) simulated reversed densities for a normal and a heavy-tailed (skewed) distribution.

exchange with groundwater and underground can enhance its mobility and variability.

The concentration of SO_4^{2-} is strongly related to dissolution and oxidation processes. Permafrost degradation, by modifying the oxidation/reduction conditions, could be responsible of its high variability and intermittency. In this perspective, the complex feed-backs affecting human-induced climate change, hydrology and landscape modifications on the redox conditions were investigated in Herndon et al. (2020). Permafrost degradation is supposed to accelerate sulfide oxidation by increasing mineral exposure to weathering, depending of the lithological context. However, sharp increases of sulfate concentrations in water are hypothesized to be better explained by increased erosion with respect to the increase of the thaw depth (Kokelj et al., 2013). Thus, permafrost degradation can result in extreme erosion events that drive rapid mineral weathering affecting the biogeochemical cycling of redox-sensitive elements. In addition, concentration of SO_4^{2-} could also be influenced by human-induced pressures and pollution. Previous research, in fact, has demonstrated the impact of air-born particles (aerosols) of sulphuric and nitric acid from industrial sources carried through the atmosphere and rivers can contribute to the acidification of Siberian lake waters (Havens and Jeppesen, 2018; Biskaborn et al., 2021). Our result on SO_4^{2-} behavior could represent a further wake-up call.

Concerning Cl^- geochemistry, processes affecting this ion are related to evapo-transpiration and sea salt aerosols but there are not way to pose limits to its concentration in unsaturated aqueous solutions, due to its highly conservative behavior. The significant Cl^- perturbations in the Lena basin are likely caused by the interaction between groundwater and marine halite/gypsum sequences widespread on the Siberian Platform (Ronov, 1982; Huh et al., 1998a), a process which may be further enhanced by permafrost degradation. Cl^- is the only element in the composition that associates high variability and intermittency with a significant increasing trend with TDS, perhaps a second wake-up call.

5. Conclusions

Over 40 years after John Aitchison's first idea (Aitchison, 1982), Compositional Data Analysis (CoDA) has revolutionized the way constrained data are managed, representing the core for the analysis of geochemical data. CoDA provides methods where the complex compositional structure and the interrelationships among different components are properly analyzed within a statistical context.

In a rapidly changing world, impacted by climate modifications and global warming, it is urgent to define the correct way to monitor compositional changes. These last enable the discovery of resilient behaviors and traces of high variability and intermittency which could be harbingers of upcoming transitions between alternative states. The complex relationships between global warming and rainfall increase are, for example, at the base of the major ions content in waters of permafrost areas, as reported by the Intergovernmental Panel on Climate Change (IPCC). Concentration increases are limited by dilution effects caused by increasing runoff (Colombo et al., 2018). However a "chemostatic" behavior where solute concentrations in stream water are not determined exclusively by dilution is reported by Godsey et al. (2009).

In this work we proposed a new CoDA approach which combines the calculus of the robust Mahalanobis distance from a pristine compositional benchmark with that of perturbation difference from the same benchmark. The first one represents a fundamental tool to have a first exhaustive attempt for identifying collective behavior and the nature of the changes, while the second enable to explore the relative effects on the individual variables. We used the lowest TDS-benchmark as reference to trace how evolutionary geochemical paths evolve and then check their relapses over time and space. By analyzing the Eastern Siberian River Chemistry (ESRC) database (Liu et al., 2022) using the proposed approach the following results were obtained:

1. The Mahalanobis distance (D) from the lowest TDS-benchmark is described by a log-normal distribution. A heavy tail distribution is found exceeding a certain threshold of D. The tail has become heaviest in more recent years as a result of the appearance of a new basin of attraction (mode) in the data.
2. High values of Mahalanobis distances seem to be more common where permafrost is more discontinuous or sporadic.
3. The calculus of perturbations from the lowest TDS-benchmark revealed that not all the chemical components of river waters respond to global changes in the same way:

- Ca^{2+}, Mg^{2+}, HCO_3^- and K^+ are poorly perturbed and have low variability
- Na^+, Cl^- and SO_4^{2-} show instead high perturbations and variability.

From the obtained results the following findings were obtained:

1. Complex multiplicative laws and feedback mechanisms governing the dissolved solids in Eastern Siberian river. The system is experiencing a lower resilience to regime shifts in more recent periods, potentially linked to enhanced permafrost degradation caused by global warming.
2. Variables participating carbonatic cycle are able to respond to environmental drives maintaining the variability of the variation rate under control thus showing a higher resilience.
3. Conversely, Na^+, Cl^- and SO_4^{2-} are more prone to flickering since their high variability might generate multiple stable states (Scheffer et al., 2009). Sometimes this can represent an early warning signal since, for instance, the known environmental conditions of the hydrological cycle can no longer be maintained (Caldecott, 2022; Xu et al., 2022).

Our approach indirectly intercepted the most important features of the Eastern Siberian environment, discriminating between chemical components where variability can be or not maintained under control, under the pressure global drivers. It is notwithstanding that, singularly, temperature, precipitation or evapo-transpiration do not show significant relationships with D values. Nevertheless, as reported in Dakos and Kéfi (2022) resilience cannot be captured by a single number and research on multiple metrics should be developed to provide a composite evaluation. This is especially true for geochemical systems, where the experience is not as high as in ecological networks. This work wishes to represent a first step along this path of research under the unusual CoDA perspective in an area greatly impacted by global warming.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Caterina Gozzi: Conceptualization, Methodology, Visualization, Investigation, Writing – original draft, Writing – review & editing. **Antonella Buccianti:** Conceptualization, Resources, Validation, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The University of Florence is acknowledged for the financial assistance to the present research through University funds (A.B, C.G.). The authors acknowledge the support of the National Biodiversity Future Center (NBFC) and National Centre for HPC, Big Data and Quantum Computing to the University of Florence, Department of Earth Sciences, funded by the Italian Ministry of University and Research, PNRR, Missione 4 Componente 2, "Dalla ricerca all'impresa", Investimento 1.4, Projects CN00000033 and CN00000013. Special thanks goes to Ping Wang for providing the ESRC Database and for the scientific support. We would like to thank three anonymous reviewers for their valuable comments and suggestions, which helped us to improve the quality of the manuscript.

References

Abramenko, V., 2014. The multifractal nature of solar magnetism and the solar dynamo problem. Geomagn. Aeron. 7, 892–898.
Aitchison, J., 1982. The statistical analysis of compositional data (with discussion). Journal of the Royal Statistical Society Series B 44 (2), 139–177.
Aitchison, J., Brown, J., 1957. The Lognormal Distribution with Special References to its Use in Economics. Cambridge University Press, Cambridge.
Aitchison, J., Ng, K., 2005. The role of perturbation in compositional data analysis. Statistical Modeling 5, 173–185.
Allègre, C.J., Lewin, E., 1995. Scaling laws and geochemical distributions. Earth Planet. Sci. Lett. 132, 1–13. https://doi.org/10.1016/0012-821X(95)00049-I.
Andersson, A., 2021. Mechanisms for log normal concentration distributions in the environment. Sci. Rep. 11 (16418), 1–7.
Ardichvili, A., Loeuille, N., Dakos, V., 2023. Evolutionary emergence of alterntive stable states in shallow lakes. Ecol. Lett. 26 (5), 692–705. https://doi.org/10.1111/ele.14180.
Barnet, J.S., Steiner, B.M., 2021. Unravelling the complex geological evolution of one of earth's final remaining frontiers: east siberia. Geology Today 37, 12–17. https://doi.org/10.1111/gto.12336.
Beck, K., Fletcher, M., Gadd, P., Heijnis, H., Saubders, K., Simpson, G., Zawadzki, A., 2018. Variance and rate-of-change as early warning signals for a critical transition in an aquatic ecosystem state: a test case from Tasmania, Australia. J. Geophys. Res. Biogeo. 123, 495–508. https://doi.org/10.1002/2017JG004135.
Belle, S., Baudrot, V., Lami, A., Misazzi, S., Dakos, V., 2017. Rising variance and abrupt shifts of subfossil chironomids due to eutrophication in a deep sub-alpine lake. Acquatic Ecology 51, 307–319. https://doi.org/10.1007/s10452-017-9618-3.
Berkin, N., Makarov, A., Rusinek, O., 2009. Baikal studies: textbook [in Russian]. 291,. Irkutsk State University Press, Irkutsk.
Billheimer, D., Guttorp, P., Fagan, W., 2001. Statistical interpretation of species composition. J. Am. Stat. Assoc. 96, 1205–1214. https://doi.org/10.1198/016214501753381850.
Biskaborn, B.K., Narancic, B., Stoof-Leichsenring, K.R., Pestryakova, L.A., Appleby, P.G., Piliposian, G.T., Diekmann, B., 2021. Effects of climate change and industrialization on Lake Bolshoe Toko, eastern Siberia. J. Paleolimnol. 65, 335–352. https://doi.org/10.1007/s10933-021-00175-z.
Bochkarev, P.F., 1959. Hydrochemistry of the rivers of eastern siberia [in russian]. 155. Irkutsk, Vostochno-Sibirskoe knizhnoe izdatel'stvo.
Buccianti, A., Pawlowsky-Glahn, V., 2005. New perspectives on water chemistry and compositional data analysis. Math. Geol. 37, 703–727. https://doi.org/10.1007/s11004-005-7376-6.
Caldecott, J., 2022. Implications of earth system tipping pathways for climate change mitigation investment. Discover Sustainability 3 (37), 1–20. https://doi.org/10.1007/s43621-022-00105-7.
Chernousenko, G., 2021. Similarities and distinctions in the genesis of salt-affected soils in different regions of the south of eastern Siberia. IOP Conference Series: Earth and Environmental Science 862, 1–10. https://doi.org/10.1088/1755-1315/862/1/012001.
Clauset, A., Shalizi, C., Newman, M., 2009. Power law distribution in empirical data. SIAM Rev. 51 (4), 661–703.
Cocks, L.R.M., Torsvik, T.H., 2007. Siberia, the wandering northern terrane, and its changing geography through the palaeozoic. Earth Sci. Rev. 82, 29–74. https://doi.org/10.1016/j.earscirev.2007.02.001.
Colombo, N., Salerno, F., Gruber, S., Freppaz, M., Williams, M., Fratianni, S., Giardino, M., 2018. Review: impacts of permafrost degradation on inorganic chemistry of surface fresh water. Global Planet. Change 162, 69–83. https://doi.org/10.1016/j.gloplacha.2017.11.017.
Cullen, A., Frey, H., 1999. Probabilistic techniques in exposure assessment. Springer New Yoir, NY, USA.
Dakos, V., Kéfi, S., 2022. Ecological resilience: what to measure and how. Environ. Res. Lett. 17, 043003 https://doi.org/10.1088/1748-9326/ac5767.
DeAngelis, A., Schubert, S., Chang, Y., Lim, Y., Koster, R., Wand, H., Marquadat Collow, A., 2023. Drivers of the exceptional warmth over siberia during the spring of 2020. J. Climate 36 (15), 4837–4861. https://doi.org/10.1175/JCLI-D-22-0387.1.
Delignette-Muller, M., Dutang, C., 2015. Fitdistrplus: an r package for fitting distributions. J. Stat. Softw. 64, 1–34. https://doi.org/10.18637/jss.v064.i04.
Egozcue, J.J., Pawlowsky-Glahn, V., 2005. Groups of parts and their balances in compositional data analysis. Mathematical Geosciences 37, 795–828. https://doi.org/10.1007/s11004-005-7381-9.
Egozcue, J.J., Pawlowsky-Glahn, V., 2011. Basic concepts and procedures. In: Pawlowsky-Glahn, V., Buccianti, A. (Eds.), Compositional Data Analysis. Theory and Applications. John Wiley & Sons, ltd., pp. 12–28. https://doi.org/10.1002/9781119976462.ch2. Chapter 2.
Feng, D., Gelason, C., Lin, P., Yang, X., Pan, M., Ishitsuka, Y., 2021. Recent changes to arctic river discharge. Nat. Commun. 12 (6917), 1–9. https://doi.org/10.1038/s41467-021-27228-1.
Filzmoser, P., Hron, K., 2008. Outlier detection for compositional data using robust methods. Math. Geosci. 40, 233–248. https://doi.org/10.1007/s11004-007-9141-5.
Filzmoser, P., Hron, K., Templ, M., 2018. Applied Compositional Data Analysis: With Worked Examples in R. Springer International Publishing 280pp, Cham. doi: https://doi.org/10.1007/978-3-319-96422-5.
Frey, K.E., Siegel, D.I., Smith, L.C., 2007. Geochemistry of west Siberian streams and their potential response to permafrost degradation. Water Resour. Res. 43 https://doi.org/10.1029/2006WR004902.
Fridovsky, V., Yakovleva, K., Vernikovskaya, A., Vernikovsky, V., Matushkin, N., Kadilnikov, P., Rodionov, N., Karpinsky, A., 2020. Geodynamic emplacement setting of late Jurassic dikes of the Yana-Kolyma Gold Belt, NE folded framing of the Siberian craton: geochemical, petrologic, and U-Pb zircon data. Minerals 10, 1000. https://doi.org/10.3390/min10111000.
Gabysheva, O.I., Gabyshev, V.A., Barinova, S., 2022. Influence of the active layer thickness of permafrost in eastern Siberia on the river discharge of nutrients into the Arctic Ocean. Water 14. https://doi.org/10.3390/w14010084.
Gaillardet, J., Dupré, B., Louvat, P., Allègre, C., 1999. Global silicate weathering and CO_2 consumption rates deduced from the chemistry of large rivers. Chem. Geol. 159, 3–30.
Georgiadi, A.G., Tananaev, N.L., Dukhova, L.A., 2019. Hydrochemical regime of the Lena river in august 2018. Okeanologiya [Oceanology] 59, 881–884.
Godsey, S.E., Kirchner, J.W., Clow, D.W., 2009. Concentration–discharge relationships reflect chemostatic characteristics of US catchments. Hydrol. Process. 23, 1844–1864. https://doi.org/10.1002/hyp.7315.
Godsey, S.E., Hartmann, J., Kirchner, J.W., 2020. Catchment chemostasis revisited: water quality responds differently to variations in weather and climate. Hydrol. Process. 33 (24), 3056–3069.
Gozzi, C., Buccianti, A., 2022. Assessing indices tracking changes in river geochemistry and implications for monitoring. Natural Resource Research 31 (2). https://doi.org/10.1007/s11053-022-10014-1.
Gozzi, C., Sauro Graziano, R., Frondini, F., Buccianti, A., 2018. Innovative monitoring tools for the complex spatial dynamics of river chemistry: case study for the alpine region. Environ. Earth Sci. 77 (579), 1–11.
Gozzi, C., Sauro Graziano, R., Buccianti, A., 2020. Part–Whole Relations: New Insights about the Dynamics of Complex Geochemical Riverine Systems. Minerals 10. doi: https://doi.org/10.3390/min10060501.
Gozzi, C., Dakos, V., Buccianti, A., Vaselli, O., 2021. Are geochemical regime shifts identifiable in river waters? Exploring the compositional dynamics of the Tiber River (Italy). Sci. Total Environ. 785, 147268 https://doi.org/10.1016/j.scitotenv.2021.147268.
Grebenshchikova, V.I., Zagorulko, N.A., Pastukhov, M.V., 2011. Monitoring of the ionic composition of water at the Angara river from lake Baikal [in Russian]. Water Chem. Ecol. 4, 2–8.
Grziwotz, F., Chang, C., Dakos, V., van Nes, E., Schwarzländer, M., Kamps, O., Heßlera, M., Tokuda, I., Telschow, A., Hsieh, C., 2023. Anticipating the occurrence and type of critical transitions. Sci. Adv. 9, 1–12.
Hartmann, J., Moosdorf, N., 2012a. Global Lithological Map Database v1.0 (gridded to 0.5° spatial resolution). doi: 10.1594/PANGAEA.788537. Supplement to: Hartmann, Jens; Moosdorf, Nils (2012): The new global lithological map database GLiM: a representation of rock properties at the earth surface. Geochem. Geophys. Geosyst. 13, Q12004.
Hartmann, J., Moosdorf, N., 2012b. The new global lithological map database GLiM: a representation of rock properties at the earth surface. Geochem. Geophys. Geosyst. 13 https://doi.org/10.1029/2012GC004370.
Hartmann, J., Lauerwald, R., Moosdorf, N., 2014. A brief overview of the global river chemistry database, glorich. Procedia Earth and Planetary Science 10, 23–27. https://doi.org/10.1016/j.proeps.2014.08.005.
Havens, K., Jeppesen, E., 2018. Ecological responses of lakes to climate change. Water 10. https://doi.org/10.3390/w10070917. URL: https://www.mdpi.com/2073-4441/10/7/917.
Herndon, E., Kinsman-Costello, L., Godsey, S., 2020. Biogeochemical cycling of redox-sensitive elements in permafrost-affected ecosystems. Geophys. Monogr. 251, Wiley. volume 251. chapter 12, 245–265.
Holmes, R.M., McClelland, J.W., Tank, S.E., Spencer, R.G.M., Shiklomanov, A.I., 2021. Arctic great rivers observatory. Water quality dataset, version 20211118. URL. https://www.arcticgreatrivers.org/data.
Hren, M., Chamberpain, C., Hilley, G., Blisniuk, M., B., B., 2007. Major ion chemistry of the Yarlung Tsangpo–Brahmaputra river: chemical weathering, erosion, and CO_2 consumption in the southern tibetan plateau and eastern syntaxis of the Himalaya. Geochim. Cosmochim. Acta 71, 2907–2935.

Huh, Y., Edmond, J.M., 1999. The fluvial geochemistry of the rivers of eastern Siberia: III. Tributaries of the Lena and Anabar draining the basement terrain of the Siberian craton and the trans-Baikal highlands. Geochim. Cosmochim. Acta 63, 967–987. https://doi.org/10.1016/S0016-7037(99)00045-9.

Huh, Y., Panteleyev, G., Babich, D., Zaitsev, A., Edmond, J.M., 1998a. a. the fluvial geochemistry of the rivers of eastern Siberia: II. Tributaries of the Lena, Omoloy, Yana, Indigirka, Kolyma, and Anadyr draining the collisional/accretionary zone of the Verkhoyansk and Cherskiy ranges. Geochim. Cosmochim. Acta 62, 2053–2075. https://doi.org/10.1016/S0016-7037(98)00127-6.

Huh, Y., Tsoi, M.Y., Zaitsev, A., Edmond, J.M., 1998b. b. the fluvial geochemistry of the rivers of eastern Siberia: I. Tributaries of the Lena River draining the sedimentary platform of the Siberian craton. Geochim. Cosmochim. Acta 62, 1657–1676. https://doi.org/10.1016/S0016-7037(98)00107-0.

Huntington, H., Boyle, M., Flowers, G., Weatherly, J., Hamilton, L., Hinzman, L., Gerlach, C., Zulueta, R., Nicolson, C., Overpeck, J., 2007. The influence of human activity in the Arctic on climate and climate impacts. Clim. Change 82. https://doi.org/10.1007/s10584-006-9162-y.

Kirchner, J., 2009. Catchments as simple dynamical systems: catchment characterization, rainfall-runoff modeling, and doing hydrology backward. Water Resource Research 45 (2), 1–34.

Kleidon, A., 2010. Life, hierarchy, and the thermodynamic machinery of planet earth. Phys. Life Rev. 7, 424–460.

Knapp, J., von Freyberg, J., Studer, B., Kiewiet, L., Kirchner, J., 2020. Concentration–discharge relationships vary among hydrological events, reflecting differences in event characteristics. Hydrol. Earth Syst. Sci. 24, 2561–2576. https://doi.org/10.5194/hess-24-2561-2020.

Kokelj, S., Lacelle, D., Lantz, T., Tunnicliffe, J., Malone, L., Clark, I., Chin, K.S., 2013. Thawing of massive ground ice in mega slumps drives increases in stream sediment and solute flux across a range of watershed scales. J. Geophys. Res. Earth 118 (2), 681–692. https://doi.org/10.1002/jgrf.20063.

Kuzmin, M.I., Tarasova, E.N., Mamontova, E.A., Mamontov, A.A., Kerber, E.V., 2014. Seasonal and interannual variations of water chemistry in the headwater streams of the Angara River (Baikal) from 1950 to 2010. Geochem. Int. 52, 523–532. https://doi.org/10.1134/S0016702914070040.

Lade, S., Tavoni, A., Levi, S., 2013. Regime shifts in a social-ecological system. Theor. Ecol. 6, 359–372. https://doi.org/10.1007/s12080-013-0187-3.

Lehner, B., Grill, G., 2013. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. Hydrol. Process. 27, 2171–2186. https://doi.org/10.1002/hyp.9740.

Li, L., Stewart, B., Zhi, W., Sadayappan, K., Ramesh, S., Kerins, D., Sterle, G., Harpold, A., Perdrial, J., 2022. Climate controls on river chemistry. Earth's. Future 10, e2021EF002603, 1–17. https://doi.org/10.1029/2021EF002603.

Lique, C., Holland, M.M., Dibike, Y.B., Lawrence, D.M., Screen, J.A., 2016. Modeling the arctic freshwater system and its integration in the global system: lessons learned and future challenges. J. Geophys. Res. Biogeo. 121, 540–566 doi:10.1002/2015JG003120.

Liu, S., Huang, Q., 2022. River Basin Boundary Dataset. https://doi.org/10.6084/m9.figshare.19142621.v1.

Liu, S., Wang, P., Huang, Q., Gabysheva, O., Li, Z., Zhang, J., Kazak, E., Liu, Y., Bazarzhapov, T., Shpakova, R., Gabyshev, V., Pozdniakov, S., Frolova, N., 2022. A database of water chemistry in eastern Siberian rivers. Scientific Data 9 (737), 1–11. https://doi.org/10.1038/s41597-022-01844-y.

Lucas, R., Klaminder, J., Futter, M., Bishop, K., Egnell, G., Laudon, H., Högberg, P., 2011. A meta-analysis of the effects of nitrogen additions on base cations: implications for plants, soils, and streams. For. Ecol. Manage. 262, 95–104. https://doi.org/10.1016/j.foreco.2011.03.018 (environmental Stress and Forest Ecosystems: Case studies from Estonia).

Maronna, R., Martin, R., Yohai, V., 2006. Robust Statistics: Theory and Methods. John Wiley & Sons, New York. ISBN 978–0–470-01092-1.

Mateu-Figueras, G., Pawlowsky-Glahn, V., Egozcue, J., 2011. The principle fo working in coordinates, in: Pawlowsky-Glahn, V., Buccianti, A. (Eds.), Compositional Data Analysis: Theory and Applications, John Wiley & Sons, Ltd, Chichester. pp. 34–42. doi: https://doi.org/10.1002/9781119976462.ch3.

May, R., Levin, S., Sugihara, G., 2008. Ecology for bankers. Nature 451, 893–894.

Meyer, R., Freeman, P., 2006. Siberian Platform: Geology and Natural Bitumen Resources. Technical Report. U.S. Geological Survey Open-File Report 2006–1316. Available online at, URL. http://pubs.usgs.gov/of/2006/1316/.

Mitzenmacher, M., 2004. A brief history of generative models for power law and lognormal distributions data analysis. Internet Math. 1 (2), 226–251.

Navarro-Lopez, C., Linares-Mustaros, S., Mulet-Forteza, C., 2022. "The statistical analysis of compositional data" by John Aitchison (1986): a bibliometric overview. SAGE Open 12. https://doi.org/10.1177/21582440221093366.

Obu, J., Westermann, S., Kääb, A., Bartsch, A., 2018. Ground temperature map, 2000-2016. Northern Hemisphere Permafrost. https://doi.org/10.1594/PANGAEA.888600.

Obu, J., Westermann, S., Bartsch, A., Berdnikov, N., Christiansen, H.H., Dashtseren, A., Delaloye, R., Elberling, B., EtzelmÜller, B., Kholodov, A., Khomutov, A., Kääb, A., Leibman, M.O., Lewkowicz, A.G., Panda, S.K., Romanovsky, V., Way, R.G., Westergaard-Nielsen, A., Wu, T., Yamkhin, J., Zou, D., 2019. Northern hemisphere permafrost map based on TTOP modeling for 2000–2016 at 1km^2 scale. Earth–Science Reviews 193, 299–316. doi: https://doi.org/10.1016/j.earscirev.2019.04.023.

Parfenov, L., Kuzmin, M.E., 2001. Tectonics, Geodynamics and Metallogeny of the Territory of the Republic of Sakha (Yakutia). Moscow, Russia, Technical Report. Nauka/Interperiodika, p. 2001, 571p. (In Russian).

Pawlowsky-Glahn, V., Egozcue, J., 2001a. Geometric approach to statistical analysis on the simplex. Stochastic Environmental Research and Risk Assessment (SERRA) 15, 384–398. https://doi.org/10.1007/s004770100077.

Pawlowsky-Glahn, V., Egozcue, J., 2001b. Geometric approach to statistical analysis on the simplex. Stoch. Env. Res. Risk A. 15, 384–398.

Pawlowsky-Glahn, V., Egozcue, J., 2006. Compositional data and their analysis: an introduction. In: Compositional Data Analysis in the Geosciences: From Theory to Practice, volume 264. Geological Society, London, Special Publications, pp. 1–10.

Pawlowsky-Glahn, V., Egozcue, J., 2020. Compositional data in Geostatistics: a log-ratio based framework to analyze regionalized compositions. Math. Geosci. doi: https://doi.org/https://doi.org/10.1007/s11004-020-09873-2.

Pawlowsky-Glahn, V., Egozcue, J., Tolosana-Delgado, R., 2015. Modeling and Analysis of Compositional Data. Statistics in Practice. John Wiley & Sons Ltd., 272pp.

Pokrovsky, O., Manasypov, R., Loiko, S., Shirokova, L., Krickov, I., Pokrovsky, B., Kolesnichenko, L., Kopysov, S., Zemtzov, V., Kulizhsky, S., Vorobyev, S., Kirpotin, S., 2015. Permafrost coverage, watershed area and season control of dissolved carbon and major elements in western Siberian rivers. Biogeosciences 12 (21), 6301–6320. https://doi.org/10.5194/bg-12-6301-2015.

Protopopov, R., Trushelev, A., Kuznetsov, Y., 2019. 2019. State Geological Map of the Russian Federation. Scale 1:1,000,000 (Third Generation). Series Verkhoyansk-Kolyma, Sheet Q-54—Ust-Nera. Explanatory Note. Technical Report. VSEGEI Cartographic Factory, Saint Petersburg, Russia, 845p. (In Russian).

Prowse, T.D., Wrona, F.J., Reist, J.D., Gibson, J.J., Hobbie, J.E., Lévesque, L.M.J., Vincent, W.F., 2006. Climate change effects on hydroecology of arctic freshwater ecosystems. Ambio 35, 347–358. https://doi.org/10.1579/0044-7447(2006)35[347:cceoho]2.0.co;2.

Prowse, T., Bring, A., Mård, J., Carmack, E., 2015. Arctic freshwater synthesis: introduction. J. Geophys. Res. Biogeo. 120, 2121–2131. https://doi.org/10.1002/2015JG003127.

R Development Core Team, 2023. R: A Language and Environment for Statistical Computing. Version 4.2.0.. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/.

Rogora, M., Mosello, R., Arisci, S., 2003. The effect of climate warming on the hydrochemistry of Alpine Lakes. Water Air Soil Pollut. 148, 347–361. https://doi.org/10.1023/A:1025489215491.

Rogora, M., Arese, C., Balestrini, R., Marchetto, A., 2008. Climate control on sulphate and nitrate concentrations in alpine streams of northern Italy along a nitrogen saturation gradient. Hydrol. Earth Syst. Sci. 12, 371–381. https://doi.org/10.5194/hess-12-371-2008.

Ronchetti, E., 2021. The main contributions of robust statistics to statistical science and a new challenge. METRON 79, 127–135. https://doi.org/10.1007/s40300-020-00185-3.

Ronov, A.B., 1982. The Earth's sedimentary shell (quantitative pat- terns of its structure, composition, and evolution). Int. Geol. Rev. 24, 1313–1363.

Rouse, W., Douglas, M., Hecky, R., Hershey, A., Kling, G., Lesack, L., Marsh, P., Mcdonald, M., Nicholson, B., Roulet, N., Smol, J., 1997. Effects of climante change on the freshwaters of artic and subartic North America. Hydrol. Process. 873–902 https://doi.org/10.1002/(SICI)1099-1085(19970630)11:8<873::AID-HYP510>3.0.CO2-6.

Sauro Graziano, R., Gozzi, C., Buccianti, A., 2020. Is compositional data analysis (CoDA) a theory able to discover complex dynamics in aqueous geochemical systems? J. Geochem. Explor. 211, 106465 https://doi.org/10.1016/j.gexplo.2020.106465.

Sazonova, T., Romanovsky, V., Walsh, J., Sergueev, D., 2004. Permafrost dynamics in the 20th and 21st centuries along the east Siberian transect. J. Geophys. Res. Atmos. 109, 1–20.

Scheffer, M., Bascompte, J., Brock, W., Brovkin, V., Carpenter, S., Dakos, V., Held, H., van Nes, E., Rietkerk, M., Sugihara, G., 2009. Early warning signal for critical transitions. Nature 461 (3), 53–59. https://doi.org/10.1038/nature08227.

Scheffer, M., Carpenter, S., Lenton, T., Bascompte, J., Brock, W., Dakos, V., van de Koppel, J., van de Leemput, I., Levin, S., Van Nes, E., Pascual, M., Vandermeer, J., 2012. Anticipating critical transitions. Science 338 (6105), 344–348. https://doi.org/10.1126/science.1225244.

Shpakova, R., 1999. Formation of the Water Quality of the Lena River in the Modern Period [in russian]. Dissertation (Thesis). for the Degree of Candidate of Geographical Sciences.

Shvartsev, S., 2013. Water-rock interaction: implications for the origin and program of global evolution. International Jurnal of Sciences 2, 26–30.

Sidorov, I.S., 1992. The Special Features in the Formation of the Hydrohemicai Regime of the Lena Mouth Area and the Southeastern Part of the Laptev Sea (in Russian). Summary of Candidate of Geograph. Sci, Thesis, Rostov-on-Don, p. 26.

Steiner, B.M., Barnet, J.S., 2021. Resource potential of one of Earth's final remaining frontiers: the Siberian platform. Geol. Today 37, 18–22. https://doi.org/10.1111/gto.12337.

van Rooij, M., Nash, B., Rajaraman, S., Holden, J., 2013. A fractal approach to dynamic inference and distribution analysis. Front. Physiol. 4 (1), 1–11.

Walvoord, M.A., Striegl, R.G., 2007. Increased groundwater to stream discharge from permafrost thawing in the Yukon River basin: potential impacts on lateral export of carbon and nitrogen. Geophys. Res. Lett. 34 https://doi.org/10.1029/2007GL030216.

Wang, P., Liu, S., Huang, Q., Gabysheva, O., Li, Z., 2022. A database of water chemistry in eastern Siberian rivers. Figshare. Collection. https://doi.org/10.6084/m9.figshare.c.5831975.v1.

Wang, P., Huang, Q., Liu, S., Liu, Y., Li, Z., Pozdniakov, S.P., Wang, T., Kazak, E.S., Frolova, N.L., Gabysheva, O.I., Zhang, J., Bai, B., Yu, J., Min, L., Shpakova, R.N., Hao, L., Gabyshev, V.A., 2024. Climate warming enhances chemical weathering in

permafrost-dominated eastern Siberia. Sci. Total Environ. 906, 167367 https://doi.org/10.1016/j.scitotenv.2023.167367.
Whitehead, P., Wilby, R., Battarbee, R., Kernan, M., Wade, A., 2009. A review of the potential impacts of climate change on surface water quality. Hydrol. Sci. J. 54 (1), 101–123. https://doi.org/10.1623/hysj.54.1.101.
Wilhelm, T., Hanggi, P., 2003. Power-law distributions resulting from finite resources. Phys. A 329, 499–508.
Xu, X., Zhang, X., Riley, W., Xue, Y., Nobre, C., Lovejoy, T., Jia, G., 2022. Deforestation triggering irreversible transition in amazon hydrological cycle. Environ. Res. Lett. 17, 1–11. https://doi.org/10.1088/1748-9326/ac4c1d.
Zheng, X., Nel, W., Peng, J., Wu, W., 2023. Hydrochemistry, chemical weathering and their significance on carbon cycle in the heilong (Amur) river basin, Northeast China. Chemosphere 327 (138542), 1–13.
Zhi, W., Williams, K.H., Carroll, R.W.H., Brown, W., Dong, W., Kerins, D., Li, L., 2020. Significant stream chemistry response to temperature variations in a high-elevation mountain watershed. Communications Earth & Environment 1 (43), 1–10. https://doi.org/10.1038/s43247-020-00039-w.
Zuo, R., Wang, J., 2016. Fractal/multifractal modeling of geochemical data: a review. J. Geochem. Explor. 164, 33–41.