

Article

Implications of Water Quality Index and Multivariate Statistics for Improved Environmental Regulation in the Irtysh River Basin (Kazakhstan)

Ultuar Zhalmagambetova ¹, Daulet Assanov ², Alexandr Neftissov ³, Andrii Biloshchytskyi ^{4,5}
and Ivan Radelyuk ^{6,*}

¹ Department of Electrical Engineering and Automation, Toraighyrov University, Pavlodar 140000, Kazakhstan; ghalmagambetova.u@tou.edu.kz

² “Environmental Safety Center” Ltd., Ust-Kamenogorsk 070000, Kazakhstan

³ Research and Innovation Center “Industry 4.0”, Astana IT University, Astana 010000, Kazakhstan; alexandr.neftissov@astanait.edu.kz

⁴ Department of Science and Innovation, Astana IT University, Astana 010000, Kazakhstan; a.b@astanait.edu.kz

⁵ Department of Information Technology, Kyiv National University of Construction and Architecture, 03680 Kyiv, Ukraine

⁶ Department of Chemistry and Chemical Technologies, Toraighyrov University, Pavlodar 140000, Kazakhstan

* Correspondence: radelyuk.i@tou.edu.kz

Abstract: The selection of sites for permanent environmental monitoring of natural water bodies should rely on corresponding source apportionment studies. Tools like the water quality index (WQI) assessment may support this objective. This study aims to analyze a decade-long dataset of measurements of 26 chemical components at 26 observation points within the Irtysh River Basin, aiming to identify priority zones for stricter environmental regulations. It was achieved through the WQI tool integrated with geoinformation systems (GISs) and multivariate statistical techniques. The findings highlighted that both upstream sections of tributaries (Oba and Bukhtarma rivers) and the mainstream of the basin are generally in good condition, with slight fluctuations observed during flooding periods. Areas in the basin experiencing significant impacts from mining and domestic wastewater treatment activities were identified. The rivers Glubochanka (GL) and Krasnoyarka (KR) consistently experienced marginal water quality throughout the observation period. Various contaminant sources were found to influence water quality. The impact of domestic wastewater treatment facilities was represented by twofold elevated concentrations of chemical oxygen demand, reaching 22.6 and 27.1 mg/L for the KR and GL rivers, respectively. Natural factors were indicated by consistent slight exceedings of recommended calcium levels at the KR and GL rivers. These exceedances were most pronounced during the cold seasons, with an average value equal to 96 mg/L. Mining operations introduced extremal concentrations of trace elements like copper, reaching 0.046–0.051 mg/L, which is higher than the threshold by 12–13 times; zinc, which peaked at 1.57–2.96 mg/L, exceeding the set limit by almost 50–100 times; and cadmium, peaking at levels surpassing 1000 times the safe limit, reaching 0.8 mg/L. The adverse impact of mining activities was evident in the Tikhaya, Ulba, and Breksa rivers, showing similar trends in trace element concentrations. Seasonal effects were also investigated. Ice cover formation during cold seasons led to oxygen depletion and the exclusion of pollutants into the stream when ice melted, worsening water quality. Conversely, flooding events led to contaminant dilution, partially improving the WQI during flood seasons. Principal component analysis and hierarchical cluster analysis indicated that local natural processes, mining activities, and domestic wastewater discharge were the predominant influences on water quality within the study area. These findings can serve as a basis for enhanced environmental regulation in light of updated ecological legislation in Kazakhstan, advocating for the establishment of a comprehensive monitoring network and the reinforcement of requirements governing contaminating activities.

Keywords: environmental monitoring; environmental regulation; Irtysh River; Kazakhstan; multivariate statistics; principal component analysis; water quality index



Citation: Zhalmagambetova, U.; Assanov, D.; Neftissov, A.; Biloshchytskyi, A.; Radelyuk, I. Implications of Water Quality Index and Multivariate Statistics for Improved Environmental Regulation in the Irtysh River Basin (Kazakhstan). *Water* **2024**, *16*, 2203. <https://doi.org/10.3390/w16152203>

Academic Editor: Christos S. Akrotas

Received: 3 June 2024

Revised: 21 July 2024

Accepted: 28 July 2024

Published: 2 August 2024



Copyright: © 2024 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The guidelines outlined by the United Nations Environment Program for Vulnerability Assessment of Freshwater Resources to Environmental Change provide a framework for evaluating the condition of surface water bodies [1]. A core assessment component involves evaluating ecological health data, focusing primarily on water quality. The assessment of water resource vulnerability is crucial for supporting sustainable ecosystem conditions in river basins and focuses on two primary objectives: ensuring adequate environmental flows to sustain the health of the river system, and supplying water resources for interconnected natural assets like wetlands, forests, and floodplains. Water quality evaluation is commonly perceived as the presence of specific chemicals exceeding recommended limits. This parameter serves as a critical indicator of river basin health and necessitates a thorough examination for each natural water body [2].

The Irtysh River Basin (hereafter referred to as the IRB, the River, or the Basin) is a transboundary river system spanning 4248 km. This river has a drainage area of approximately 165,000 square kilometers. Originating at the border of Mongolia (<1%) and China (2.9%), the Irtysh River flows through Kazakhstan (53.1%) before entering Russia (44%), where it joins the Ob River and eventually drains into the Arctic Ocean [3]. The Basin sustains a population of around 15 million people and serves as a crucial water source for the region, playing a significant role in supporting social, economic, and industrial infrastructure, energy generation, and irrigation activities [4]. Within Kazakhstan, the Irtysh River Basin accounts for 30% of the total water runoff, representing the largest contribution to the country's water availability [5].

River water quality issues, due to their profound implications for ecosystem restoration and public health, are of particular concern. For example, the utilization of potentially unsuitable water from the Irtysh River can directly impact rural inhabitants, as highlighted by Tussupova et al. [6]. Two primary sources contribute to water quality challenges in Kazakhstan. First, intensive industrial operations along the riverbanks contribute to both historical contamination and ongoing pollution. Second, runoff from adjacent riparian areas further exacerbates these issues [7]. Industries in Kazakhstan have exploited regulatory loopholes leading to significant pollution of the river and its tributaries and raising serious concerns about water quality [8]. Previous research on the Irtysh River mainstream has revealed varying water quality conditions across different regions of Kazakhstan. While water quality in the Pavlodar region is generally suitable for various uses, including irrigation, fishing, domestic use, and recreation, the East Kazakhstan section is considered unsuitable for use without extensive treatment [9]. However, the region lacks in-depth investigations into the sources contributing to poor water quality in the Basin, providing primarily descriptive data and findings. Recent research by Krupa et al. [10] involved independent sampling and analysis of chemical parameters, revealing elevated concentrations of N-NO₂, Cu, Fe, Cr, and Zn along the mainstream. Similarly, a study by Ryskeldieva et al. [11] compared these findings with the distribution of Hg, Cd, Pb, Cu, and Zn along the mainstream and their clarkes in the hydrosphere, their spatial distribution, and change over time. The sole study dedicated to investigation of water quality in tributaries was conducted by Burlibayeva et al. [12] during 2010–2011. Their comprehensive evaluation of extreme concentrations of chemical components and biotesting techniques underscored significant impacts from the mining industry on water quality near Ust-Kamenogorsk. Biotesting results revealed high acute toxicity levels in tributary water samples. Even though chemical parameters met regulatory standards, most samples exhibiting survival rates were below 46.7%.

While natural purification processes generally maintain satisfactory water quality along the main course of the River, it is essential to prioritize water quality assessments in tributaries, particularly in areas vulnerable to contamination. Current methods and monitoring procedures for the Basin exhibit inefficiencies, leading to an inadequate understanding of the present state and necessitating a reevaluation. Implementing standardized monitoring and evaluation protocols throughout the basin would enable a more equitable

and objective assessment of regional challenges, facilitating targeted remediation efforts. Furthermore, analyzing historical and contemporary water quality data is crucial for effective river management and contamination risk mitigation. This analysis provides a critical baseline for identifying trends, potential pollution sources, and the effectiveness of existing and suggested management practices [13]. A comprehensive understanding of past and present water quality supports proactive environmental management strategies that safeguard these vital ecosystems. Additionally, there is a significant research gap concerning persistent organic pollutants, pesticides, pharmaceuticals, microplastics, and analogous contaminants in the Basin's water bodies.

The proper selection of sites for such investigations and the execution of permanent monitoring, as stipulated by the revised Ecological Code, should heavily rely on advanced source apportionment studies [14]. The water quality indices concept addresses such tasks, encouraging development of standardized systems for evaluating and communicating water quality issues, particularly in the context of pollution abatement [15]. Numerous water quality indices (WQIs) have been developed over decades. Among these, the Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI) has gained recognition as a robust tool for assessing the suitability of water bodies to support aquatic life and for evaluating overall river health [16]. This tool considers water quality parameter standards set by regulatory bodies or relevant authorities, tailored to address environmental issues and concerns [17] and to assure proper access to clean water, sanitation, and hygiene as a basic human right [18] in each particular area. One of the sound examples was the alignment of WQIs with Brazilian legislation by linking parameter information, aiding in environmental compliance and monitoring [19]. Furthermore, the integration of geographic information systems (GISs) and multivariate statistics enhances the capacity to analyze spatial trends in water quality parameters [20], identify pollution sources [21], improve pollution control [22], and intervene [23] and prioritize effective management strategies [24], as evidenced by numerous studies [25].

This work is dedicated to a comprehensive assessment of the health status of the Irtysh River Basin in its East Kazakhstan part. The authors aimed to pinpoint hotspots and sources of pollution that exert the most substantial influence on water quality in both the mainstream and tributaries of the River. For this purpose, a ten-year dataset of measurements encompassing 26 chemical components at 26 observation points was analyzed utilizing the water quality index (WQI), geographic information systems (GISs), and multivariate statistical tools. The central idea of this study is to uncover specific areas with low water quality and to identify particular sources of contamination that will be targeted for improved environmental regulation in alignment with the updated ecological legislation of Kazakhstan.

2. Materials and Methods

2.1. Study Area, Data Management, and Research Flow

The Irtysh River flows into Kazakhstan from China and is also fed by tributaries such as the Ishim, Ulba, Tobol, and other smaller rivers [12]. Precipitation, including snowfall, and groundwater originating from the Altay Mountains also significantly contribute to the river's flow [26]. The climate within the basin in Kazakhstan varies from mountainous regions in the southeast to flat plains in the northwest, influencing temperature regimes and precipitation levels. The continental climate prevalent in desert and semi-desert areas moderates significantly in mountainous and foothill regions. January stands out as the coldest month, with average temperatures falling between -12°C and -19°C , reaching minimums as low as -45°C and -49°C . Conversely, July represents the warmest month, with average temperatures ranging from 20°C to 22°C . While there is an overall trend of increasing temperatures annually, winters tend to show a slight cooling effect [27]. Annual precipitation levels vary between 400 and 650 mm, exhibiting notable differences between mountainous and foothill areas. Precipitation during the warmer months (March–October) surpasses that of the colder period (December–March).

This research focuses on the East Kazakhstan region of the Republic of Kazakhstan (Figure 1). A total of 26 monitoring sites were selected based on data provided by the National Hydrometeorological Service of Kazakhstan, “Kazhydromet”. Eight of these points are located along the mainstem (IR 1–IR 7), with one positioned at the initial entry point near the Kazakhstan–China border (KI). Five sampling points were established along the primary tributary, the Ulba River (UL 1–UL 5). The selection of these sampling sites was made considering upstream, within, and downstream positions relative to settlements, and the potential impact of both urban and rural areas on water quality. The same approach was employed when establishing observation points along the Krasnoyarka (KR), Clubochanka (GL), Breksa (BR), Tikhaya (TK), Oba (OB), and Bukhtarma (BU) rivers. Observation points located on the Tikhaya, Breksa, and Ulba rivers (UL 1, UL 2) are situated within the city of Ridder, a major metallurgical center of East Kazakhstan. Points UL 3–UL 5 and IR 1–IR 3 are situated within the administrative boundaries of Ust-Kamenogorsk, the central city of the East Kazakhstan region. Observation points IR 6–IR 7 are located near the city of Semey, the administrative center of the Abay region. Observation points on the Krasnoyarka and Glubochanka rivers are situated within the boundaries of large villages with operational domestic wastewater treatment facilities. All sampling points are located at a reasonable distance from the monitored objects, at the regulated boundary of sanitary zones defined as being at least 1000 m downstream from contamination sources.

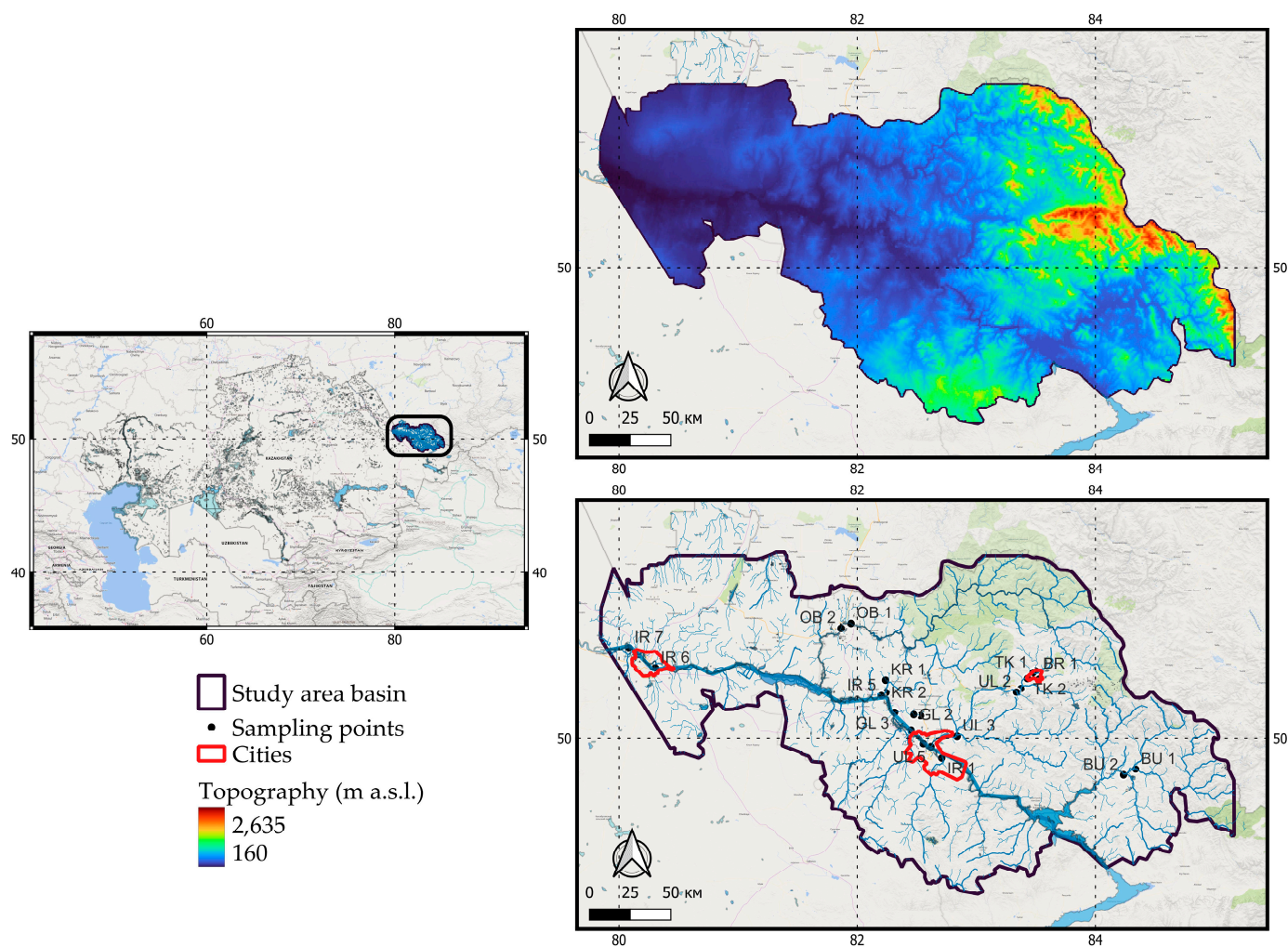


Figure 1. Study area.

Flow patterns indicate a consistent spatio-seasonal effect (Figure 2). Peak flow levels at the Ulba and Semiyarka stations typically peak in April, influenced by the practice of

storing water in upstream reservoirs during this month. In May, the release of water from these reservoirs leads to peak flows downstream from the study area. The peak flow at Kara Ertis station in June corresponds to the snowmelt from the Chinese part of the basin. The variation in average annual flow between the subsequent monitoring sites like the Kara Ertis, Ulba, and Semiyarka stations ($280 \text{ m}^3/\text{s}$, $94 \text{ m}^3/\text{s}$, and $967 \text{ m}^3/\text{s}$ respectively) can be explained by the relation of their locations to reservoir cascades and the inflow from the Ulba River. The average flow slightly decreases downstream from these points but remains relatively constant at the last observation points within Kazakhstan (Zhanabet and Priirtyshskoe, with average flows of $891 \text{ m}^3/\text{s}$ and $886 \text{ m}^3/\text{s}$, respectively).

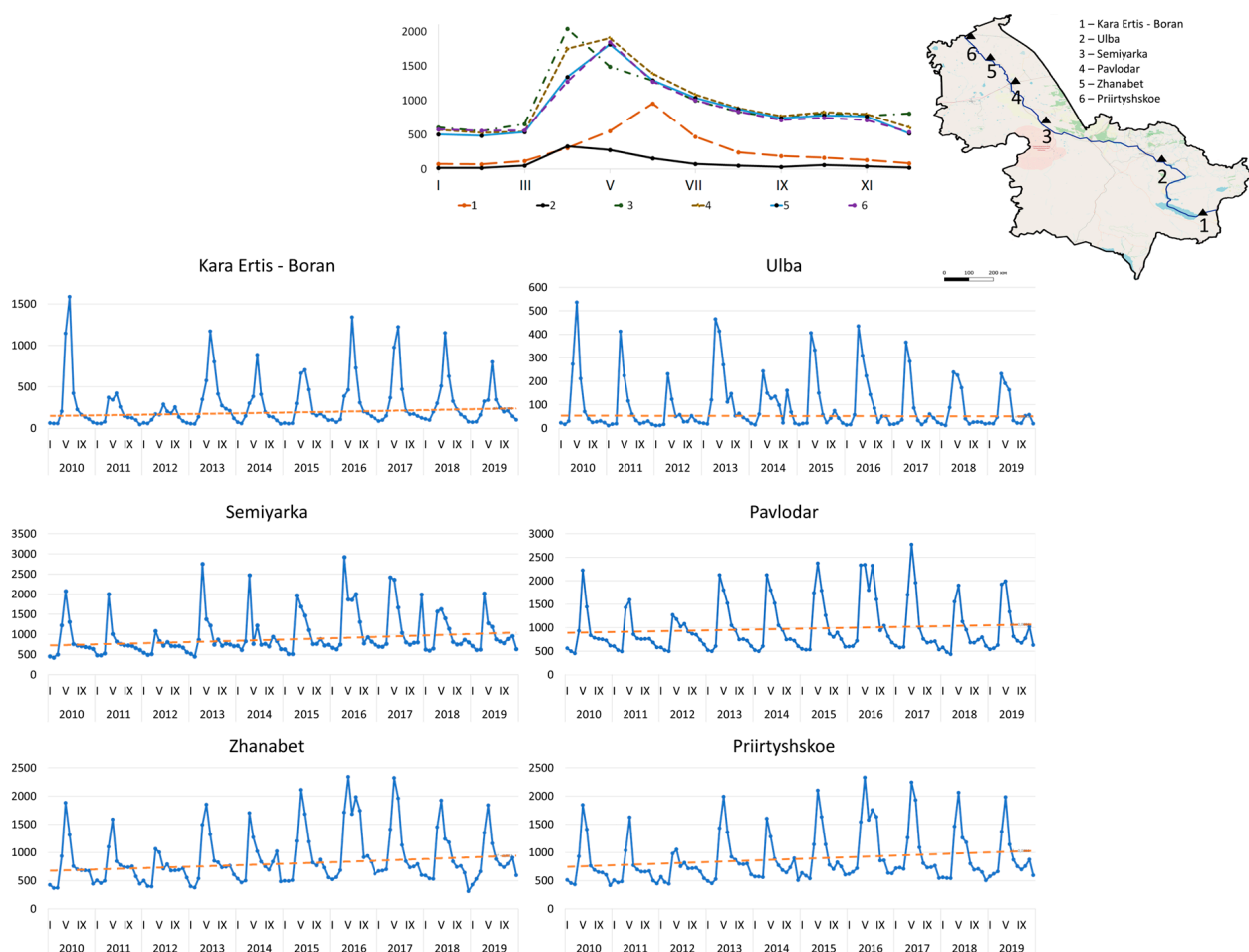


Figure 2. Flow rates (m^3/s) along the main course of the Irtysh River Basin in Kazakhstan. The top graph displays the monthly average flow rates at the stations. Source: [9] (Open Access).

The National Hydrometeorological Service of Kazakhstan, “Kazhydromet”, provided a comprehensive ten-year dataset spanning from 2011 to 2020 for this study. The equipment, laboratory, staff, and analytical methodologies adhered to national and international standards for quality control and quality assurance (QC/QA), accredited and controlled by ILAC (International Laboratory Accreditation Cooperation, Newton, Australia, <https://ilac.org/> accessed on 29 July 2024). The raw dataset consisted of measurements of 45 hydrological and chemical parameters throughout the entire observation period. Due to uneven measurements frequencies for different parameters, sampling events were selected based on the availability of regular chemical analyses for key contamination indicators. For seasonal fluctuations analysis this dataset was segmented into three distinct periods: cold season (including measurements taken in January and March, once per month), warm season (monthly measurements from June to October), and flooding season (monthly measurements in April and May). The final dataset included 17, 41, and 22 sampling events

for the cold, warm, and flooding seasons, respectively. The dataset contains the results of measurements of 26 chemical components, including total suspended solids (TSSs), pH, percentage of dissolved oxygen (% DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), NH_4^+ , NO_2^- , NO_3^- , PO_4^{3-} , Fe, Cu, Zn, Cr, Ni, Cd, Hg, As, Mn, total petroleum hydrocarbons (TPHs), phenols, F^- , Ca, Mg, Na, SO_4^{2-} , and Cl^- .

The research methodology employed the following steps. First, a water quality index (WQI) tool was utilized for a comprehensive assessment of water quality. Second, descriptive statistical analysis was conducted to interpret the WQI scores. Third, the potential sources impacting water quality were identified through the application of multivariate statistical techniques. These techniques included Pearson's correlation matrices and principal component analysis (PCA). These techniques evaluate the similarities in contaminants' patterns and group them based on their likely shared source of origin. Fourth, hierarchical clustering analysis was employed to summarize the findings of the source apportionment study and to identify spatio-seasonal patterns in water quality across the entire basin. This framework was applied to data from all three defined seasons (cold, warm, and flooding).

Mathematical and statistical computations were performed using Microsoft Office Excel 2016, IBM SPSS Statistics 26 software, and the "R" programming language (version 4.2.2., R Core Team). Spatial interpolation of the water quality index across the basin was analyzed using the inverse distance method in QGIS Desktop 3.20.3.

2.2. Water Quality Index

This study evaluates water quality in the basin using the Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI), a tool that simplifies complex water quality data into a more understandable format [28]. The CCME WQI comprises three main components: scope (F1), which considers the number of water quality parameters that do not meet objectives; frequency (F2), indicating how often objectives are not met; and amplitude (F3), reflecting the degree to which objectives are not attained. The index assigns a numerical value between 0 (poorest) and 100 (best) to depict water quality.

Parameter selection is typically determined by management goals and environmental characteristics specific to the research area [29]. Potential pollution sources in the IRB include natural factors like leaching from bedrock, the discharge of municipal wastewater, and mining activities [30]. The chosen chemical parameters cover all presumed contributors of natural and anthropogenic sources to environmental pollution in the region to ensure a comprehensive and unbiased assessment. While some parameters may have overlaps, multivariate statistical techniques facilitate the identification of source apportionment and related analyses. Establishing reference values is crucial for characterizing exceedances of permissible limits and assessing the health status of natural water bodies, such as rivers. These reference values can be sourced from various scientific, legislative, and regulatory frameworks, thus showcasing significant variability. The limits utilized in this study were derived from different sources through the comparison of water quality in the IRB with the conditionally ideal characteristics applicable to an excellent water status (see more in Supplementary Materials). As the primary aim of this research was to identify areas within the IRB with the poorest water quality and to evaluate seasonal trends, the established reference values may not represent universally applicable concentrations. These values might require adjustment when conducting holistic assessments of other river basins worldwide. The particular reference values were sourced from the research by Khan et al., which focused on evaluating water quality in rivers from a global perspective [31]. The reference value for dissolved oxygen was established based on research examining its impact on fish growth in aquaculture [32]. The guideline for the grey water footprint was used to set limits for man-made contaminants corresponding to a healthy water status [33]. Additionally, reference values for COD, NH_4^+ , and TPHs were obtained from the respective sources of the World Health Organization (WHO) [34] and the original source of CCME WQI [35].

2.3. Multivariate Statistics

To explore the multivariate relationships among variables and samples in the study area, correlation analysis, principal components analysis (factor analysis), and hierarchical cluster analysis were employed. Normalization of the dataset was conducted to mitigate the effects of unit variations.

Pearson's correlation coefficient was used to assess the relationships between each pair of variables, aiming to identify geochemical associations among different parameters. Correlation coefficients exceeding 0.5 were deemed statistically significant. Principal components analysis (PCA) was employed to extract the most influential parameters from a large dataset of inter-correlated variables and derive independent components. Factor analysis (FA) is a similar technique to PCA. However, PCA is presented as a linear combination of parameters. FA follows PCA and takes into account unobservable, hypothetical, and latent variables. They are included in the following equation with the special residual term (Equation (1)).

$$z_{ij} = a_{f1}f_{1j} + a_{f2}f_{2j} + \dots + a_{fm}f_{mj} + e_{fi}, \quad (1)$$

where z is the measured variable, a is the factor loading, f is the factor score, e is the residual term according to errors or other source of variation, i is the sample number, and m is the total number of factors.

A combination of sampling points was identified using cluster analysis, which grouped them based on similarities in chemical variable patterns. Hierarchical agglomerative cluster analysis utilized Ward's linkage distance, reported as $D_{\text{link}}/D_{\text{max}}$, Euclidean distance, and Q-mode to assess patterns in water quality. This approach facilitated the creation of a dendrogram that simplifies the analysis of similarities.

3. Results and Discussion

An overview of the findings is presented in Table 1. Each observation point can be categorized into one of five statuses: from 95 to 100 (excellent), from 80 to 94 (good) (indicating water quality is predominantly safeguarded with minimal threat with deviations from natural levels being infrequent), from 60 to 79 (fair) (indicating water quality is generally protected but intermittently threatened with occasional deviations from natural levels), from 45 to 59 (marginal) (indicating frequent threats to water quality with conditions frequently deviating from natural levels), or from 0 to 44 (poor).

An analysis of water quality data revealed that none of the sampling points achieved an "excellent" status according to the CCME WQI across all three seasons. The initial sampling point in Kazakhstan Kara Irtysh, located on the mainstream, exhibited the best water quality throughout the year. This finding supports the hypothesis that major pollution sources likely originate from tributaries. Similarly, the Bukhtarma River, originating in mountainous regions and unaffected by sources of contamination, demonstrated consistently good water quality. This trend was also noted upstream of the Breksa River.

The Krasnoyarka, Glubochanka, and Tikhaya rivers require particular attention. These rivers exhibited marginal and fair water quality statuses in 4, 8, and 6 out of 21 evaluations, respectively. The Ulba River also displayed concerning trends, with 9 out of 15 evaluations falling into the marginal and fair categories of the WQI (two and seven times, respectively). The lowest water quality was observed during the cold period, with 7 out of 26 observation points displaying marginal water quality statuses.

A ten-year retrospective analysis revealed persistent exceeding of established thresholds for trace metals (zinc, cadmium, and copper), calcium, and sulfates, which serve as indicators of natural processes within the river basin. Additionally, chemical oxygen demand levels surpassed acceptable limits, serving as a marker for the influence of domestic wastewater facilities (Figure 3). The highest zinc concentration was observed at the station KR 2 in 2010, exceeding the limit of 0.03 mg/L by nearly 100 times. Similarly, an alarming spike in zinc levels was detected at the GL 2 sampling point in April 2019, reaching a concentration of 2.025 mg/L, nearly 70 times the standard threshold. In September 2019 at

UL 2, zinc levels peaked at 1.57 mg/L, surpassing the permissible concentration by over 50 times. Figure 4 indicates a pattern of higher Zn concentrations during the cold seasons compared to the warm seasons. However, even during warm periods, Zn concentrations remained above the recommended limits.

Table 1. Water quality indices across different seasons in the studied area.

	Cold	Flooding	Warm
Krasnoyarka 1 (KR 1)	86	78	80
Krasnoyarka 2 (KR 2)	53	62	52
Glubochanka 1 (GL 1)	82	71	75
Glubochanka 2 (GL 2)	60	60	62
Glubochanka 3 (GL 3)	67	67	72
Ulba 1 (UL 1)	64	73	69
Ulba 2 (UL 2)	59	70	71
Ulba 3 (UL 3)	90	82	84
Ulba 4 (UL 4)	81	74	81
Ulba 5 (UL 5)	84	76	79
Breksa 1 (BR 1)	84	86	84
Breksa 2 (BR 2)	60	65	72
Tikhaya 1 (TK 1)	57	66	63
Tikhaya 2 (TK 2)	60	68	63
Oba 1 (OB 1)	84	79	87
Oba 2 (OB 2)	91	76	86
Bukhtarma 1 (BU 1)	91	82	87
Bukhtarma 2 (BU 2)	77	77	82
Irtys 1 (IR 1)	91	87	91
Irtys 2 (IR 2)	72	81	78
Irtys 3 (IR 3)	80	83	82
Irtys 4 (IR 4)	84	84	82
Irtys 5 (IR 5)	81	79	82
Irtys 6 (IR 6)	87	86	84
Irtys 7 (IR 7)	89	89	84
Kara Irtys (KI)	86	84	91

Similar to zinc, cadmium exhibited concerning exceedances of safe limits. The maximum Cd concentration (0.8 mg/L) was recorded in April 2018 at the station KR 2, exceeding the limit by a factor of 1000. Stations GL 2, TK 1, and UL 2 also displayed persistent Cd contamination with values consistently surpassing the allowable concentration of 0.0008 mg/L. For example, the Cd content at TK 1 in January 2016 reached 0.147 mg/L, exceeding the norm by more than 180 times. Copper concentrations also displayed extreme elevations with the highest value (0.051 mg/L) observed at the station KR 2 exceeding the limit of 0.004 mg/L by 13 times. At the GL 2 sampling point in 2014, copper concentrations reached 0.046 mg/L, surpassing the reference limit by 12 times. The data depict distinct seasonal variations with lower values observed during warmer periods and pronounced peaks during flooding seasons.

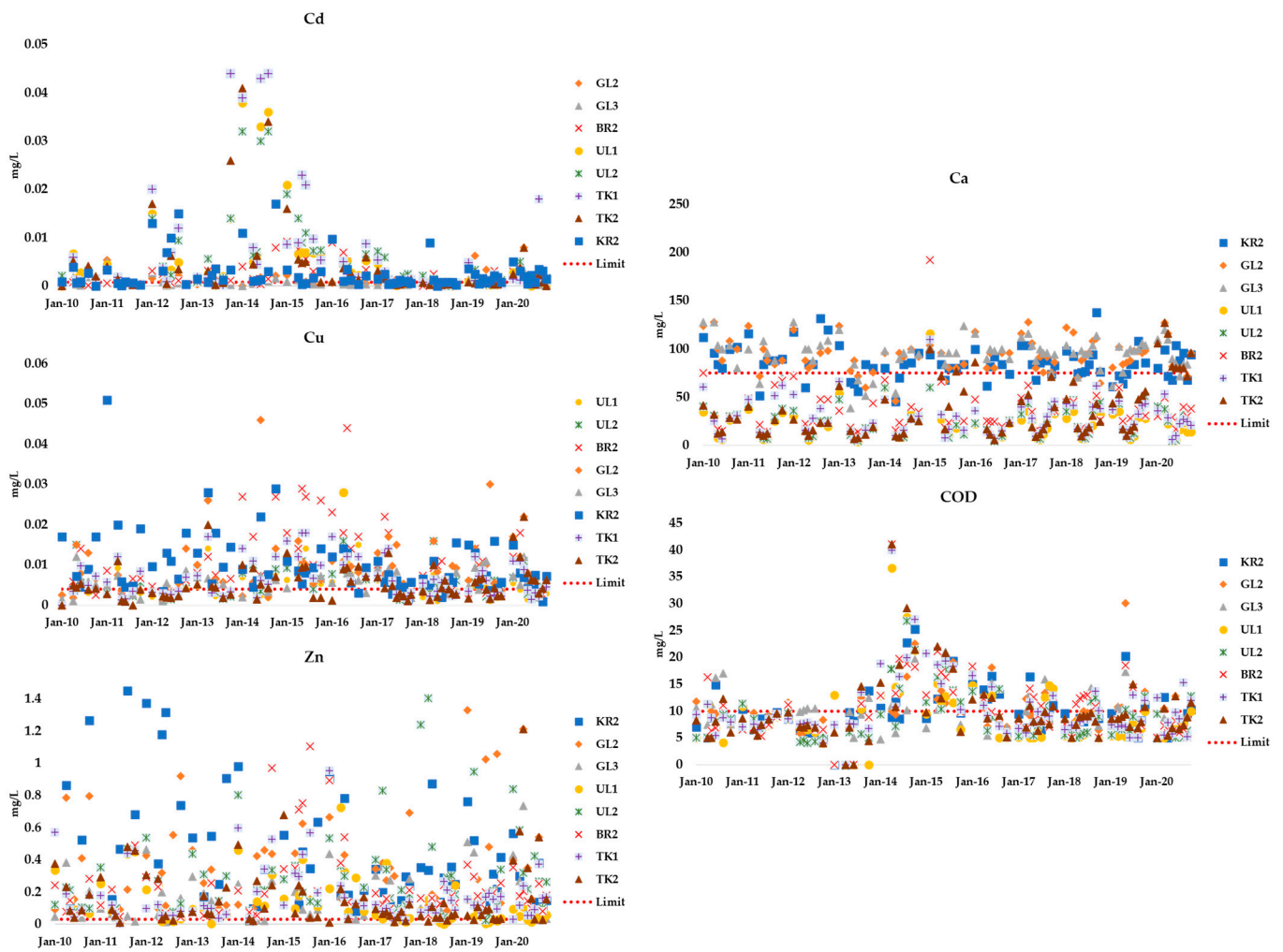


Figure 3. Temporal variation of selected chemicals within the studied period in the observation points with fair and marginal water quality status.

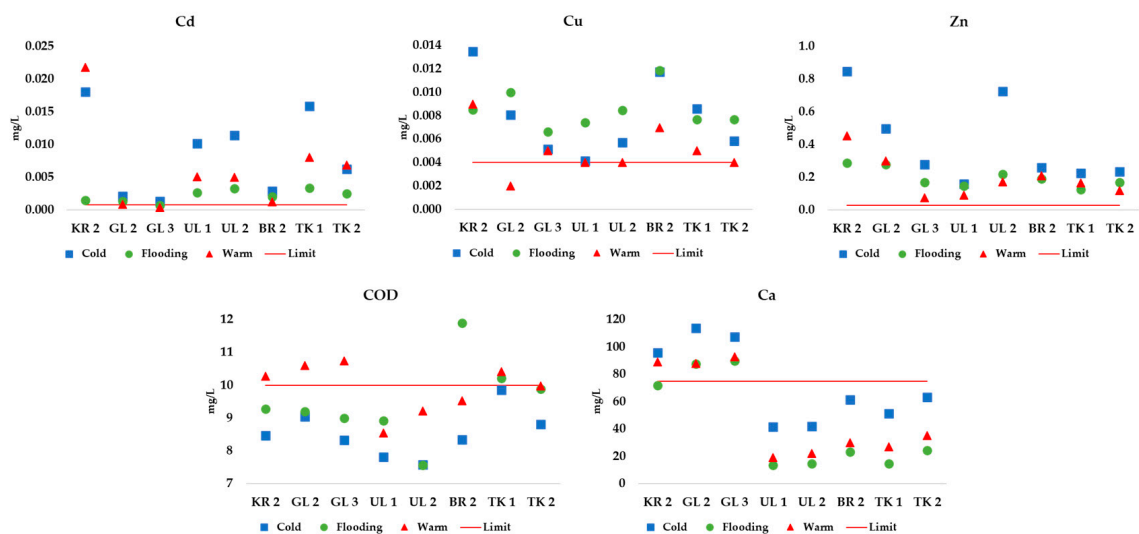


Figure 4. Average seasonal concentrations of indicative chemicals in the observation points with fair and marginal water quality status.

Consistent slight exceedances of recommended calcium levels were noted at the KR 2, GL 2, and GL 3 sampling stations. These exceedances were most pronounced during the

cold seasons for all sampling points with minimal values observed during flood periods likely due to dilution effects. In contrast, chemical oxygen demand (COD) demonstrated peak values during warmer seasons, while registering minimal levels during colder periods. KR 2, GL 2, and GL 3 emerged as the most impacted locations under the influence of this parameter.

3.1. Cold Seasons

Figure 5 illustrates the spatial distribution of the water quality index (WQI) within the study area during cold seasons. The key characteristic of the season is river ice covering, which represents a dynamic element in freshwater ecosystems, influencing water quality through various physical and biogeochemical processes [36]. While ice cover can offer temporary benefits, its impact on water quality is multifaceted. Reduced sunlight penetration resulting from the ice cover restricts the primary production by phytoplankton, leading to a decrease in dissolved oxygen levels in the water column [37]. This oxygen depletion can stress aquatic organisms, especially during prolonged periods of ice cover. Moreover, certain pollutants may become concentrated in the unfrozen water as they are excluded during ice crystal formation, posing a particular concern for contaminants with high water solubility [38].

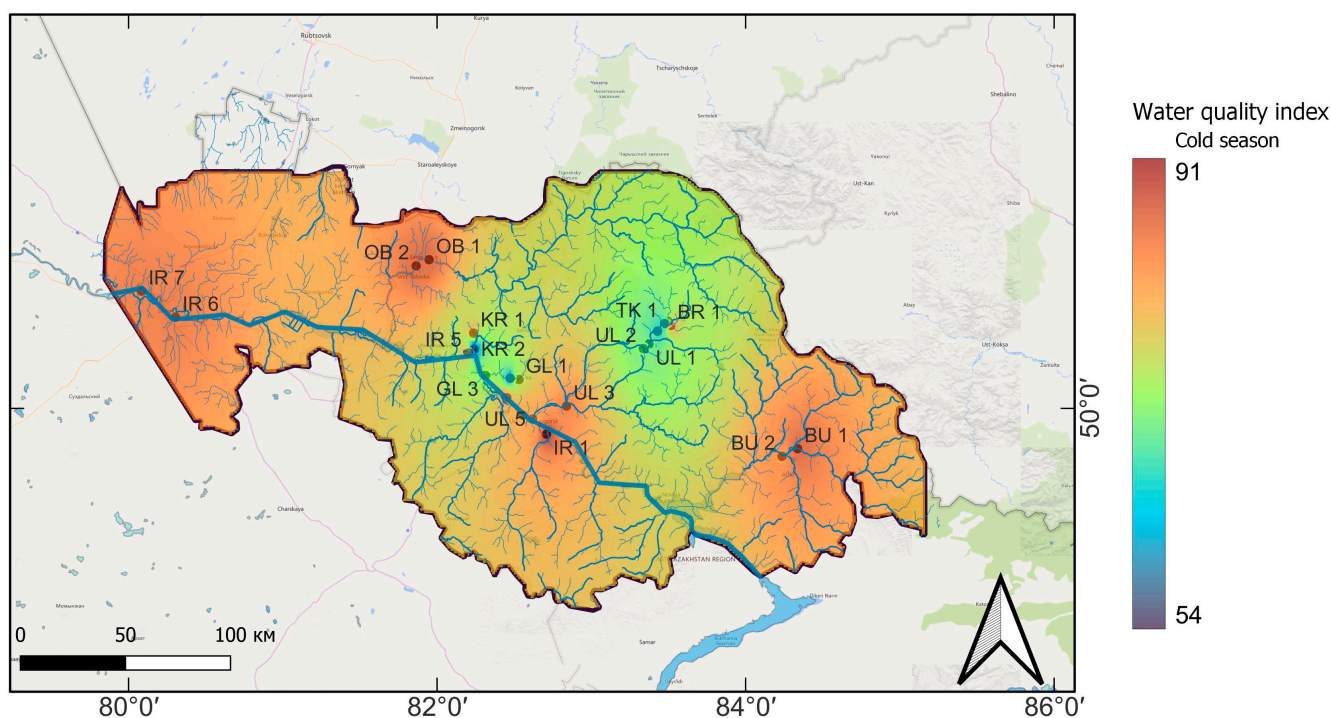


Figure 5. Distribution of WQI within the study area during cold seasons.

Table S1 summarizes the descriptive statistics of water quality parameters within the IRB during the cold periods. The lowest WQI value was recorded at sampling point KR 2, standing at 53, indicating substandard water quality primarily attributed to elevated levels of metals, specifically copper, zinc, and cadmium. The average concentrations observed over the monitoring period for these metals were 0.013 mg/L, 0.847 mg/L, and 0.018 mg/L, respectively, surpassing permissible limits by significant margins of 3, 28, and 23 times, respectively. Similarly, median values of 0.011 mg/L, 0.554 mg/L, and 0.034 mg/L also exceeded the permissible thresholds of 0.004 mg/L, 0.03 mg/L, and 0.0008 mg/L, respectively, for each parameter. Magnesium exhibited a peak concentration exceeding the maximum limit of 32 mg/L, while the average concentration equal to 24 mg/L remained within the permissible limit of 30 mg/L. Similarly, manganese levels averaged 0.13 mg/L, exceeding the permissible limit of 0.05 mg/L by nearly three times. Calcium concentrations

averaged 96 mg/L with a median of 99 mg/L, exceeding the limit of 75 mg/L. While the mean values of other parameters fell below the reference values, intermittent exceedances were observed, impacting the overall WQI. The maximum COD value recorded was 15 mg/L, while the average remained at 8.47 mg/L, staying within the permissible limit of 10 mg/L. Phosphate ion (PO_4^{3-}) levels surpassed the standard limit by approximately threefold, reaching a maximum of 0.29 mg/L, although the average concentration stayed below the specified limit at 0.06 mg/L. Biological oxygen demand averaged 1.79 mg/L, below the established norm with a single maximum exceedance of 3.61 mg/L.

Another sampling point, TK 1, exhibited marginal water quality with a WQI value of 57. Exceedances of permissible limits were observed for several parameters. Nitrite ion concentrations averaged 0.081 mg/L, exceeding the reference value by 1.35 times with a maximum concentration reaching 0.21 mg/L, surpassing the norm by 3.5 times. Phosphate ion concentrations also exceeded the established limit. The average concentration (0.29 mg/L) was nearly threefold higher than the norm, while the peak concentration (1.03 mg/L) exceeded it by tenfold. Ammonia ion concentrations displayed similar trends. The average value was marginally above the permissible level, slightly exceeding 1.5 mg/L. The maximum observed value equal to 3.38 mg/L exceeded the norm by more than three times. Zinc exhibited an average concentration of 0.226 mg/L, which surpasses the norm by almost eightfold, with a maximum observed value of 0.956 mg/L that is higher than the limit by 30 times. Cadmium levels exceeded the limit by 20 times, with an average concentration of 0.0159 mg/L and a maximum of 0.147 mg/L, surpassing the reference value by 180 times. Mean manganese concentrations stood at 0.08 mg/L, exceeding the norm by 1.5 times with a maximum concentration of 0.15 mg/L, surpassing it threefold. Furthermore, indicators of both natural and anthropogenic factors surpassed the limits multiple times during the observation period. COD concentrations exceeded the limit of 10 mg/L with the maximum value reaching 20.7 mg/L. Additionally, concentrations of calcium and sulfates surpassed the reference values by 1.5 times. The downstream point TK 2 also showed marginal water quality with exceedances observed in zinc, cadmium, and manganese (0.235 mg/L, 0.0063 mg/L, and 0.08 mg/L, respectively). Additionally, the average dissolved oxygen level remained below the recommended threshold, standing at 78% throughout the observation period.

The Tikhaya River converges with the Ulba River near observation point UL 1, which also exhibited poor water quality with a WQI value of 64. While the primary contaminants identified were zinc (mean value equal to 0.161 mg/L) and cadmium (mean value equal to 0.0102 mg/L), sporadic and varying elevations in other parameters were observed across all pollutant types. Similarly, poor water quality was observed at sampling point UL 2 with a WQI value of 59. Median values for copper, zinc, and cadmium exceeded the established norms by 1.5, 24, and 14 times, respectively. The median concentrations were 0.007 mg/L for copper, 0.586 mg/L for zinc, and 0.005 mg/L for cadmium. Maximum exceedances were observed at 0.009 mg/L for copper, 1.570 mg/L for zinc (surpassing norms by 52 times), and 0.066 mg/L for cadmium (exceeding norms by 82 times). Manganese concentrations also exhibited significant exceedances and surpassed the norm nearly fivefold on average with a maximum concentration of 0.52 mg/L, exceeding the limit by tenfold.

At the observation point GL 2, marginal water quality status was observed with a WQI value of 60. This status was primarily characterized by elevated average concentrations of copper (0.008 mg/L), zinc (0.497 mg/L), and cadmium (0.021 mg/L). While the exceedances were not persistent, peak concentrations further highlighted potential concerns. Copper reached a maximum of 0.017 mg/L, exceeding the norm by more than four times. Zinc exhibited the most significant exceedance with a peak concentration of 2.025 mg/L, surpassing the norm by over 67 times. Similarly, cadmium's peak concentration of 0.0063 mg/L exceeded the established limit by more than 7.5 times. Manganese concentrations averaged 0.13 mg/L with a maximum of 0.24 mg/L, exceeding the norm by fivefold. Magnesium levels, on average, stayed within the established norms. However, a peak concentration of 38 mg/L was observed, exceeding the maximum limit by 8 mg/L. Exceedances in the

maximum values of BOD and COD were noted at 3.1 mg/L and 14 mg/L, respectively, while the dissolved oxygen content was at 76%, below the recommended value of 80%.

Multivariate statistical analysis revealed distinct spatial patterns in water quality characteristics during the cold season. At KR 2, which showed the lowest WQI value, PC1 indicated positive correlations and loadings among Fe, Ni, and TPHs, while negatively associating with COD, BOD, pH, and TSS. These parameters, influenced by natural processes and local wastewater treatment plant effluents, predominantly govern the water quality within established limits. PC2 identified copper, zinc, and cadmium as significant contributors to water quality degradation, exhibiting negative correlations with fluoride and sulfate ions (Figure S1, Table S4).

Similarly, at the contaminated site GL 2, PCA highlighted high loadings of Zn, Cd, and Mn and their correlations with ammonia and nitrite ions (Figure S2, Table S5). The presence of metals could potentially inhibit nitrification and denitrification processes, exacerbated by oxygen depletion in winter, potentially leading to adverse environmental impacts, particularly from poorly treated effluent discharge. PC2 at this location grouped parameters such as phosphates, TPHs, pH, and BOD, suggesting the influence of domestic wastewater (even treated) under conditions of low dissolved oxygen and cold temperatures, potentially contributing to elevated phosphate levels.

At the upstream Ulba River location (UL 1), elevated concentrations of Zn and Cd were observed. These were influenced by both natural and anthropogenic factors (Table S6). The correlation matrix showed strong associations among parameters linked to domestic wastewater impacts: BOD, COD, NH_4^+ , NO_2^- , and TPHs (Figure S3). Downstream at UL 2, a significant deterioration in water quality was observed, with trace metals (Mn, Zn, Na, and Fe) dominating the first principal component (Table S7). The correlation matrix underscored strong interrelations among these metals, despite continued significant impacts from domestic wastewater, as reflected in PC2 with positive correlations and loadings of NO_2^- , NH_4^+ , and TPHs (Figure S4).

At the Tikhaya River near the Ulba confluence (TK 1), extensive contamination across various chemical parameters was observed with notable contributions from natural and mining sources. (Figure S5, Table S8). PC1 highlighted the loading of sulfates (SO_4^{2-}) and magnesium (Mg), along with metals like Mn, Cu, and Cd. PC2 revealed strong positive correlations among natural components: Na, Ca, and Cl^- . PC3 combines NO_2^- , NH_4^+ , and TPHs, indicating shared anthropogenic sources and elevated concentrations during sampling events. Slight improvement downstream at TK 2 was seen, but natural and mining factors continued to influence water quality, indicating its marginal status (Table S9, Figure S6). Finally, near the confluence of the Tikhaya and Ulba at BR 2, marginal WQI was mirrored by PCA and correlation matrix profiles resembling TK 1, suggesting common contamination sources. (Table S10, Figure S7).

Hierarchical cluster analysis (HCA) corroborates the spatial patterns identified by principal component analysis (PCA), as visualized in the dendrogram with three distinct clusters (Figure 6). Cluster 1 encompasses UL 1, UL 2, TK 1, TK 2, and BR 2, indicating a collective influence of localized natural processes, mining activities, and the discharge of domestic wastewater on water quality parameters. This cluster suggests a significant impact of both anthropogenic and environmental factors on shaping the water quality profiles at these monitoring sites. Cluster 2 includes observation points GL 2, GL 3, and KR 2. This cluster is characterized by distinctive features associated with notable contributions from domestic effluents along with influences from natural processes. Finally, Cluster 3 comprises observation points that exhibit favorable water quality conditions and a stronger tendency towards self-purification, potentially due to higher resilience to environmental stressors.

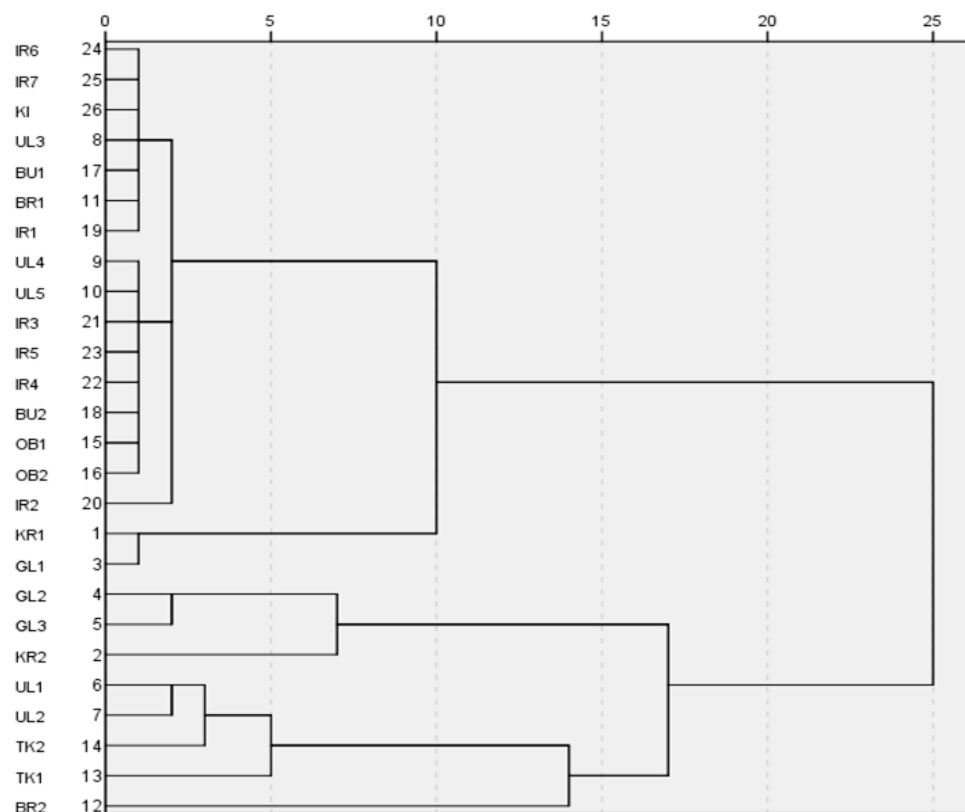


Figure 6. Dendrogram illustrating clustering of sampling sites based on water quality characteristics during cold seasons.

3.2. Flooding Seasons

Spatial analysis revealed distinct variations in water quality parameters between flooding seasons and cold periods (Figure 7). Observation points within Cluster 1 (UL 1, UL 2, TK 1, TK 2, and BR 2), previously identified through hierarchical cluster analysis, exhibited significant improvements in water quality, progressing from marginal to fair levels. This trend suggests a potential dilution effect caused by increased discharge during flooding events, which may temporarily decrease pollutant concentrations. Across the entire basin, a general deterioration in water quality from good to fair conditions was observed during flooding seasons compared to cold periods (Table S2). Upstream points within the main-stream maintained a more stable water quality status throughout the year, as they lacked significant contamination sources or possessed robust self-purification capabilities.

Among all observation points during flooding seasons, KR 2 and GL 2 exhibited the worst water quality, which was reflected by WQI values of 62 and 60, respectively. Total suspended solids displayed slightly but consistently elevated concentrations and were equal to 155 mg/L on average at KR 2. Potentially toxic elements also showed concerning elevations during the observed period. Copper levels exceeded the standard limit of 0.004 mg/L twofold, reaching 0.008 mg/L, with a peak of 0.028 mg/L, which is surpassing guidelines by sevenfold. Zinc levels also exhibited significant exceedances, with an average value 0.289 mg/L, which is ten times higher than the established limit of 0.03 mg/L. Similarly, cadmium concentrations exceeded the norm of 0.0008 mg/L twofold with an average of 0.0016 mg/L and a maximum concentration of 0.007 mg/L, exceeding the limit by nearly nine times. It is important to note that a single event of elevated mercury was observed, with a concentration equal to 0.0005 mg/L that is higher than the permissible limit by sixteenfold. Other parameters remained within acceptable ranges, while some displayed episodic elevations. Iron levels peaked at 0.93 mg/L, which is three times higher than the limit, while manganese peaked at 0.174 mg/L, exceeding the standard by fivefold, and magnesium reached 49 mg/L, surpassing the limit by 1.5 times. Although ammonia,

nitrite, and phosphate ions stayed within acceptable limits on average, their maximum observed values exceeded established norms. Maximum concentrations rose to 2.4 mg/L, 0.072 mg/L, and 0.17 mg/L, respectively, exceeding their corresponding limits by 1.6, 1.2, and 1.7 times, respectively.

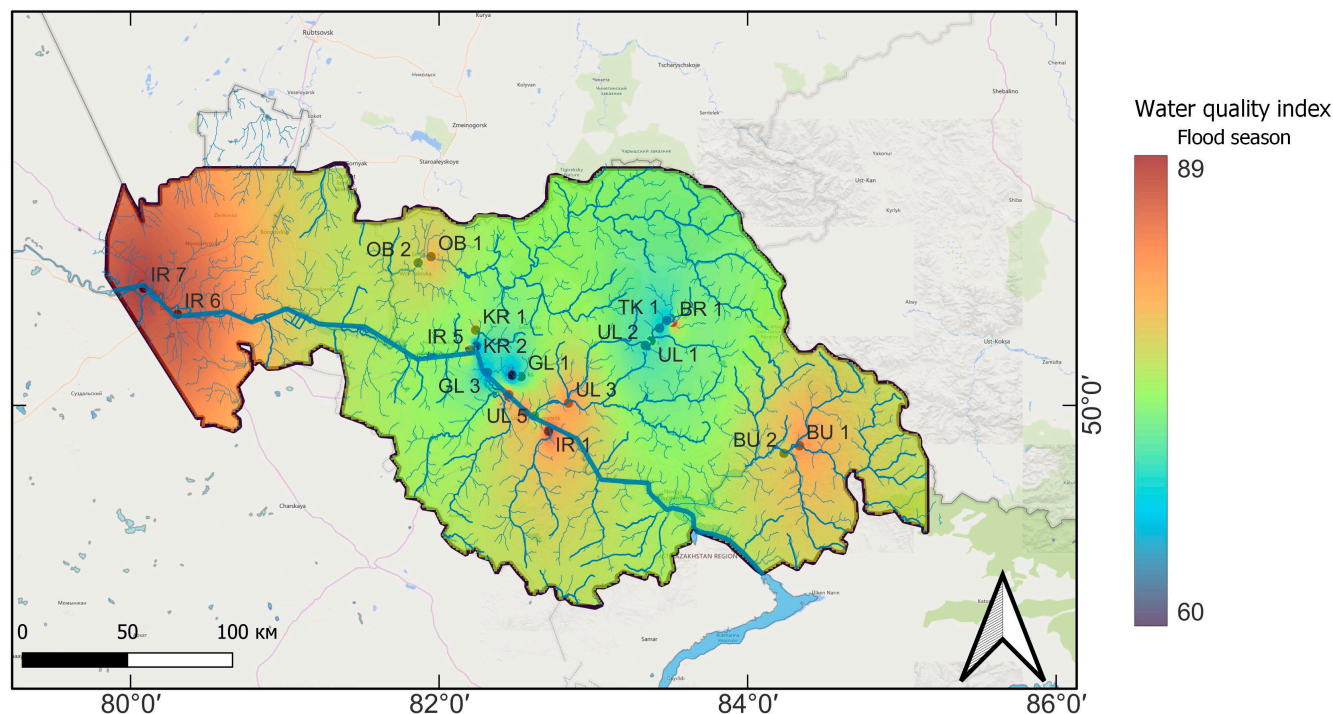


Figure 7. Distribution of WQI within the study area during flooding seasons.

At monitoring point GL 2, water quality was predominantly influenced by high concentrations of heavy metals. Average concentrations of copper (0.01 mg/L), zinc (0.297 mg/L), and cadmium (0.0014 mg/L) significantly surpassed the established reference values by 7, 40, and 10 times, respectively. The presence of mercury was also detected during one sampling event, with a concentration of 0.0004 mg/L, exceeding the permissible limit by more than 13 times. Furthermore, the average level of dissolved oxygen stood at 79%, consistently below the optimal threshold of 80%, indicating a slightly but consistently suboptimal oxygen content in the water.

Hierarchical cluster analysis (HCA) identified two distinct groups of observation points, each characterized by specific water quality attributes (Figure 8). Cluster 1 grouped sampling locations KR 2, GL 2, and GL 3, exhibiting analogous poor water quality conditions denoted by WQIs of 62, 60, and 67, respectively. This cluster also included KR 1 and GL 1, situated upstream, which consistently indicated high concentrations of Zn, Cd, and Mn, exceeding established thresholds. These metals contribute significantly to PC1 in the principal component analysis at the KR 2 point (Table S11). Strong interrelations among these contaminants and nitrate ions were observed, therefore hinting at the impact of flood-induced dilution effects on water quality deterioration (Figure S8). The correlation matrix for KR 2 further revealed an inverse relationship between fluoride (F^-) and several parameters, including pH, dissolved oxygen (DO), ammonia (NH_4^+), and nickel (Ni) (Figure S8). These parameters were collectively loaded onto PC2. Conversely, high positive loadings and correlations were evident for sulfate (SO_4^{2-}), calcium (Ca), magnesium (Mg), sodium (Na), and chloride (Cl^-), underscoring the substantial influence of natural factors on water quality at the observation point. A similar pattern was observed at GL 2 with a more consistent correlation matrix, indicating stronger inter-parameter relationships (Table S12). Strong positive correlations were observed among Cd, Zn, Ni, Mn, and Hg, which all contributed significantly to PC1 (Figure S9). Natural parameters

such as total suspended solids (TSSs), Mg, Ca, and low DO levels further exacerbated the adverse impact on water quality, rendering GL 2 particularly vulnerable. Similar trends in principal components and chemical parameters were observed downstream at GL 3, although slight improvements in water quality were noted (Figure S10, Table S13). Thus, the confluence of mining activities and prevailing natural conditions emerged as pivotal determinants shaping water quality dynamics within the IRB during flood seasons.

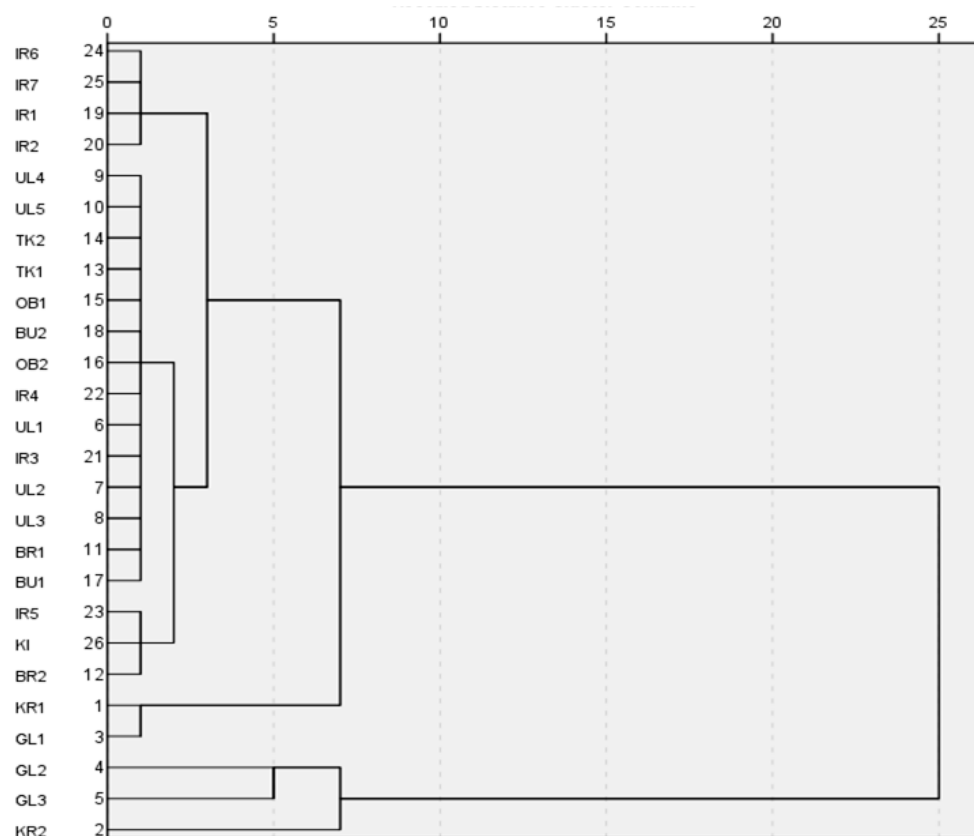


Figure 8. Dendrogram illustrating clustering of sampling sites based on water quality characteristics during flooding seasons.

3.3. Warm Seasons

Similar trends in water quality were observed during the warm season compared to cold periods (Figure 9, Table S3). Particularly, water quality improved after flooding seasons in the Oba and Bukhtarma rivers. This can be explained by dilution effects followed by a return to pre-flood conditions. Conversely, the Tikhaya, Krasnoyarka, and Glubochanka rivers returned to marginal water quality status based on WQI values after the dilution period.

The examination of KR 2 displayed the following indicative observations. Trace metal concentrations consistently surpassed established thresholds. These included Cu at 0.09 mg/L, Zn at 0.454 mg/L, Cd at 0.0218 mg/L, Mn at 0.101 mg/L, and Ca at 89 mg/L. The average COD concentration of 10.28 mg/L slightly exceeded the established limit of 10 mg/L. Episodic elevations were observed for several parameters. pH levels remained within the permissible limit, reaching a maximum value of 8.8. TSS reached a maximum of 170 mg/L, exceeding the norm by 20 mg/L. Peak COD values reached 22.6 mg/L, surpassing the norm by 2.5 times. Ammonium levels peaked at 1.75 mg/L, exceeding the reference value of 1.5 mg/L by 0.25 mg/L. Phosphate ions displayed a similar pattern, with peak concentrations reaching 0.34 mg/L, exceeding the norm by over threefold, while the average concentration remained at 0.05 mg/L.

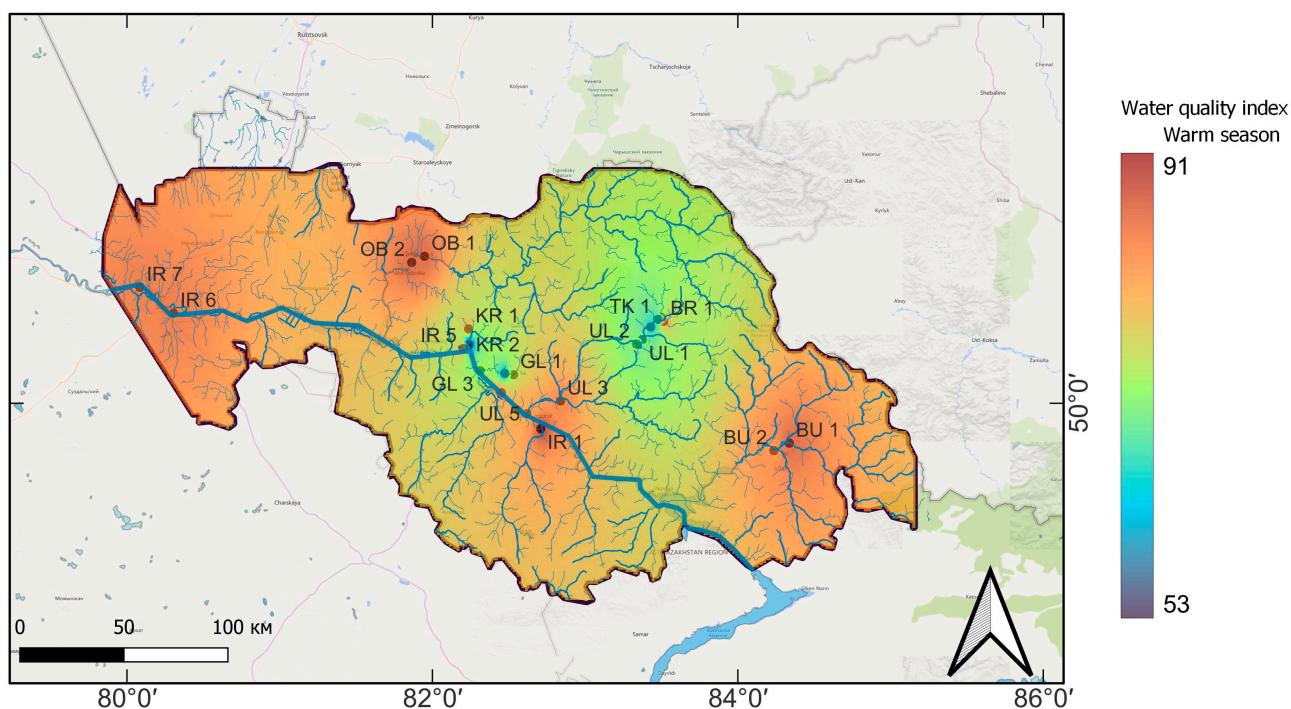


Figure 9. Distribution of WQI within the study area during warm seasons.

GL 2 exhibited trends similar to KR 2, with average COD (10.61 mg/L) slightly exceeding the established norm of 10 mg/L. pH levels remained within permissible limits, reaching a maximum value of 8.7. Metal concentrations surpassed established reference values, including Cu (0.046 mg/L), Zn (1.058 mg/L), Cd (0.0034 mg/L), Mn (0.288 mg/L), and Ca (111 mg/L). Similar to KR 2, episodic elevations were observed for several parameters. TSSs peaked at 450 mg/L, exceeding the norm by threefold. BOD reached 3.42 mg/L, surpassing the norm by 1.14 times. Maximum COD values reached 27.1 mg/L, exceeding the norm by threefold. NH_4^+ levels peaked at 3.03 mg/L, doubling the reference value of 1.5 mg/L. NO_2^- reached 0.16 mg/L, surpassing the recommended value by 2.7 times.

Shifting focus to TK 1, average indicators generally adhered to norms, with BOD averaging 3.56 mg/L and COD at 27.1 mg/L. Metal concentrations, however, exceeded established thresholds, with Cu and Cd surpassing reference values by 1.3 and 10 times, respectively. Additionally, peaks in ammonium levels were observed, reaching a high of 3.04 mg/L, while nitrite peaked at 0.14 mg/L, surpassing reference levels by 1.5 times, and phosphate concentrations recorded a maximal value of 0.8 mg/L, surpassing limits by eight times. Downstream at TK 2, elevated metal concentrations posed a significant threat, with Zn and Cd exceeding reference values by four and eight times, respectively.

The lowest WQI value recorded during the entire study period was observed at KR 2 during the warm season (WQI = 52). Principal component analysis (PCA) highlighted natural components like fluoride, calcium, chloride, sulfate, and sodium as significant factors across the first three principal components, thus indicating their impact on water quality. Strong correlations among these parameters underscore their substantial contribution to water quality (Figure S11). Additionally, positive correlations were also observed between these natural parameters and zinc and manganese, which exceeded established concentration limits. Another observation point exhibiting marginal water quality status was GL 2. PCA results for GL 2 revealed a contrasting pattern compared to KR 2 (Table S15). Anthropogenic components like ammonium, COD, BOD, and nitrite were prominent in PC1. Key correlations included a strong positive association between Cu and TSSs, and NH_4^+ and Mg, and a notable negative correlation between NH_4^+ and F^- . These findings suggest that elevated temperatures during the warm season may have influenced processes associated with all three pollution sources without a single predominant contributor. Simi-

lar water quality trends were observed downstream at GL 3, with reduced concentrations of potentially toxic metals. PCA results highlighted the significant contributions of various factors, including domestic wastewater and natural components like ammonia, nitrite, fluoride, calcium, sulfate, and manganese (Table S16). Additionally, moderate negative correlations were observed between ammonia and nitrite ions with natural components such as F^- and Ca (Figure S13). The presence of Zn and Cd in PC3 alongside Na suggests potential connections between natural geochemical processes and local mining activities. KR 1 and GL 1 also indicated improved water quality during warm seasons, although slight COD concentrations exceeding established standards indicated residual domestic wastewater influence. PCA results for these locations mirrored those of GL 3, highlighting the complexity of source apportionment (Table S17).

For TK 1, potentially toxic metals (Zn, Cd, and Cu) dominated PC1 (Table S18) and exhibited positive correlations with COD (Figure S14), suggesting both elevated metal concentrations and domestic wastewater influence (PC2). Phosphate, ammonia, and sulfate ions further supported the role of domestic sources impacting water quality at TK 1. PC2 also included sodium, bridging to PC3, which housed natural variables such as Ca, Mg, and Cl^- , all exhibiting strong positive correlations. Downstream sampling at TK 2 indicated a high loading of natural factors, including Ca, Na, Mg, SO_4^{2-} , and Cl^- in PC1 (Table S19), with all parameters displaying strong positive correlations (Figure S15). Notably, pH also correlated strongly with these parameters. PCA for BR 2 highlighted mining activities with Zn, Cd, Mn, and Cu contributing significantly to PC1 (Table S20), corroborated by strong positive correlations among these metals (Figure S16). Additionally, PC2 showed a high loading of natural factors, including Na, SO_4^{2-} , Mg, and Ca.

Cluster analysis identified two groups of observation points with different water quality statuses (Figure 10). Cluster 1 comprised observation points with improved water quality indicators influenced by natural and mining factors with minor contributions from domestic wastewater. Cluster 2 encompassed observation points with marginal to fair water quality statuses: BR 2, TK 1, and TK 2, demonstrating significant impacts from mining activities and natural components.

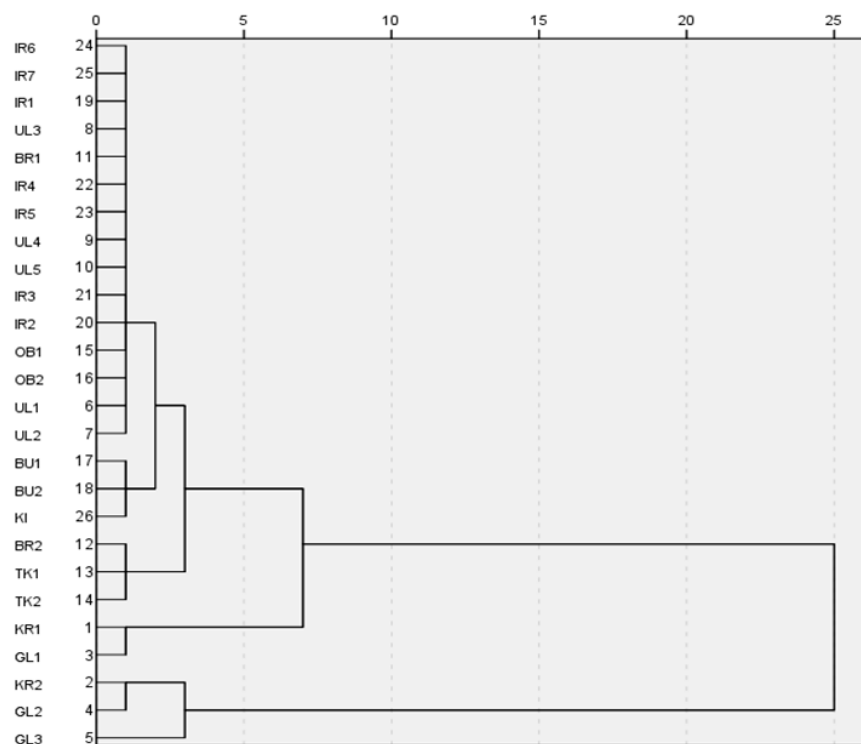


Figure 10. Dendrogram illustrating clustering of sampling sites based on water quality characteristics during warm seasons.

3.4. Discussion

This study highlights the significant impact of mining activities and inadequately treated domestic wastewater on the water quality of tributaries within the Irtysh River Basin. The study underscores the necessity of paying close attention to the potential transfer of toxic elements, as highlighted in a review by Babuji et al., which identifies hazardous substances such as Co, Cu, Mn, Pb, Se, Ni, V, and Zn that can migrate to soils and subsequently enter plants, particularly vegetables [39]. This suggests a critical consideration, especially as the river inundates the floodplain during the spring period, exacerbating water quality issues in the main water body and potentially leading to floodplain contamination in the basin [40]. Moreover, these elements remain at high concentrations even nowadays, according to the annual bulletin from Kazhydromet [41]. Regrettably, the bulletin fails to clarify the proposed sources of extreme concentrations, offering only a broad reference to the “main attribution to emissions from technological production and the impact of the unique soil composition in the area”. Therefore, Tileugabulov and Madani refer to the active and closed tailings storage facilities as the driving surface water contamination sources since the 1960s [42].

The findings demonstrated a significant impact of ore mining activities on the water quality within the basin. In 2021, mining companies in Ridder extracted over 17,000 tonnes of zinc and 1300 tonnes of lead. Production figures for raw materials in the region included over 110,000 tonnes of zinc and 155,000 tonnes of lead produced in Ust-Kamenogorsk, alongside over 107,000 tonnes of raw lead and 55,000 tonnes of copper [43]. Being large taxpayers, mining enterprises within the region might benefit from the prevalence of the Entrepreneurial Code over the Ecological Code. Industrial inspections were conducted every six months with companies receiving advance notice of at least thirty calendar days [44]. This prior notification allowed facilities to undertake corrective actions in the short term to achieve compliance during the inspection. There is only one known study focused on the water management of a mining company operating in the region [30]. Despite the company's increasing water intake, a significant rise in freshwater withdrawal was also identified. The company's withdrawals exceeded their government-issued water permit by up to 30%. This discrepancy underscores the outdated water allocation standards used by governmental bodies, failing to adapt to advancements in mining technology and the growing water demands. The authors recommended that government bodies adopt specialized mining discharge regulations that can progressively mitigate the environmental impact of mining on both water quality and quantity. Additionally, climate change poses a threat to mining industries in the region by affecting water supply and management through changes in precipitation patterns and temperature extremes [45]. This further exacerbates the existing challenges to water quality in the studied region.

Another concern comes from the impact of domestic wastewater treatment facilities. The study [46] claimed that most of the Kazakhstani small settlements rely solely on basic treatment facilities (mechanical and natural systems) that are often outdated and in poor condition, designed and built in the 1960s and 1980s. Moreover, according to a study by Karataev et al., only a small percentage (7%) of industrial wastewater and sewage underwent full treatment before being released back into waterways [5]. This might lead to excessive loads of organic pollutants and biogenic compounds. Thus, these facilities are not operating properly and are likely discharging poorly or even untreated wastewater with human health [47] and environmental threats [48] consequences.

Until recently, water management practices in Kazakhstan have suffered from a lack of science-based approaches at both governmental and industrial levels. There is an emergency for detailed characterization of wastewater, combined with the implementation of new environmental monitoring tools, such as contamination transport models, which can empower environmental agencies to effectively assess and control the impact of industrial activities and associated risks [49]. The implementation of permanent monitoring stations mandated by the Ecological Code requires enterprises to maintain automated “real-time” monitoring systems for emissions and pollutant discharges at contamination sources. This

data must be recorded and directly transmitted to a government regulatory body [50]. The findings of this study can serve as a guide for installing such monitoring equipment in crucial areas, particularly near the confluence of the Ulba, Tikhaya, and Breksa rivers (based on mining company data) and at locations with outdated wastewater treatment facilities (Krasnoyarka and Glubochanka rivers).

Another critical aspect requiring attention is the continuous update of the list of substances subjected to operational control. While environmental monitoring in developed countries ensures environmental safety and aligns with principles of sustainable development, leading to robust monitoring systems, Kazakhstani industries have often resisted such changes. For instance, there has been industry pushback against the inclusion of sensitive indicators in the list of monitored parameters [49]. These legislative loopholes have historically allowed polluting industries to further contaminate already compromised sites and water bodies. In line with recent legislative updates, it is imperative to prioritize identified hot zones as the primary areas for remediation and restoration efforts in the region.

4. Conclusions

This study is the first known attempt which successfully identified patterns in the water quality of the Irtysh River Basin in its East Kazakhstan part in the extensive spatial and seasonal levels. The water quality index indicated good water quality upstream of the Ob and Bukhtarma rivers and the Basin's main channel with slight declines during flood events. However, emerging areas like those influenced by ore mining (Tikhaya, Ulba, and Breksa rivers) and outdated wastewater treatment facilities (Glubochanka and Krasnoyarka rivers) exhibited concerning water quality issues. Multivariate statistical analyses supported these findings and highlighted the influence of natural factors, particularly during flooding seasons. These results provide valuable insights for governmental bodies. Prioritizing specific areas for enhanced environmental monitoring and stricter regulations for mining and manufacturing companies is crucial for safeguarding the Basin's water quality.

The main limitation of this work comes from the utilization of a slightly outdated dataset. Despite this limitation, its comprehensive ten-year timeframe mitigates potential risks and enables a holistic assessment. Another limitation, which can be identified as a prospective avenue for future research, considers the opportunity to expand the set of contaminants for monitoring purposes. This expansion may include such emerging pollutants as persistent organic compounds, pharmaceuticals, pesticides, etc., to enhance the comprehensiveness of environmental monitoring efforts. The authors emphasize the importance of prompt decision-making based on scientific evidence, as historical studies suggest potential delays in addressing environmental concerns in Kazakhstan. By proactively implementing these recommendations, policymakers can effectively address existing water quality challenges and ensure the long-term sustainability of the Irtysh River Basin.

Supplementary Materials: The supporting information for the descriptive statistics of the water quality within the Irtysh River Basin, and the results of multivariate statistical analysis, including correlation matrices and tables of factor analysis can be downloaded at <https://www.mdpi.com/article/10.3390/w16152203/s1>. Reference [51] is cited in the Supplementary Materials.

Author Contributions: U.Z.: formal analysis, investigation, writing—original draft; D.A.: data curation, formal analysis, validation, writing—original draft; A.N.: funding acquisition, project administration, resources, writing—review and editing; A.B.: funding acquisition, project administration, resources, writing—review and editing; I.R.: conceptualization, formal analysis, investigation, methodology, software, validation, visualization, writing—original draft. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan within the project BR21882258 “Development of Intelligent

Information and Communication Systems Complex for Environmental Emission Monitoring to Make Decisions on Carbon Neutrality”.

Data Availability Statement: Data have been provided by the National Hydrometeorological Service of Kazakhstan, “Kazhydromet”, upon a special request.

Acknowledgments: The authors are thankful to Ansar Kazambayev from the Department of Electric Power Industry, Toraighyrov University, Pavlodar for the support in data processing. The authors also express their appreciation to Ilyas Kazambayev from Astana IT University, Astana for his assistance in revising the manuscript.

Conflicts of Interest: Author Daulet Assanov was employed by the company “Environmental Safety Center” Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Huang, Y.; Cai, M. *Methodologies Guidelines. Vulnerability Assessment of Freshwater Resources to Environmental Change*; UNEP: Nairobi, Kenya, 2009.
- Fawell, J. Emerging contaminants, source water quality and the role of standards. *Water Pract. Technol.* **2015**, *10*, 432–437. [\[CrossRef\]](#)
- Huang, W.J.; Duan, W.L.; Nover, D.; Sahu, N.; Chen, Y.N. An integrated assessment of surface water dynamics in the Irtysh River Basin during 1990–2019 and exploratory factor analyses. *J. Hydrol.* **2021**, *593*, 125905. [\[CrossRef\]](#)
- Krasnoyarova, B.; Vinokurov, Y.; Antyufeeva, T. International water development problems in the transboundary Irtysh River basin: “new” solutions to old problems. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2019; p. 012049.
- Karatayev, M.; Kapsalyamova, Z.; Spankulova, L.; Skakova, A.; Movkebayeva, G.; Kongyrbay, A. Priorities and challenges for a sustainable management of water resources in Kazakhstan. *Sustain. Water Qual. Ecol.* **2017**, *9–10*, 115–135. [\[CrossRef\]](#)
- Tussupova, K.; Hjorth, P.; Berndtsson, R. Access to Drinking Water and Sanitation in Rural Kazakhstan. *Int. J. Environ. Res. Public Health* **2016**, *13*, 1115. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hrkal, Z.; Gadalia, A.; Rigaudiere, P. Will the river Irtysh survive the year 2030? Impact of long-term unsuitable land use and water management of the upper stretch of the river catchment (North Kazakhstan). *Environ. Geol.* **2006**, *50*, 717–723. [\[CrossRef\]](#)
- Radelyuk, I.; Tussupova, K.; Zhapargazinova, K.; Yelubay, M.; Persson, M. Pitfalls of Wastewater Treatment in Oil Refinery Enterprises in Kazakhstan—A System Approach. *Sustainability* **2019**, *11*, 1618. [\[CrossRef\]](#)
- Radelyuk, I.; Zhang, L.; Assanov, D.; Maratova, G.; Tussupova, K. A state-of-the-art and future perspectives of transboundary rivers in the cold climate—A systematic review of Irtysh River. *J. Hydrol. Reg. Stud.* **2022**, *42*, 101173. [\[CrossRef\]](#)
- Krupa, E.; Romanova, S.; Serikova, A.; Shakhvorostova, L. A Comprehensive Assessment of the Ecological State of the Transboundary Irtysh River (Kazakhstan, Central Asia). *Water* **2024**, *16*, 973. [\[CrossRef\]](#)
- Ryskeldieva, A.; Burlibaeva, D.; Yerdeshbay, A.; Kamelkhan, G.; Sarova, N. Clarke concentrations of heavy metals in surface waters of the transboundary river Yertis (Kazakhstan). *Water Sci.* **2023**, *37*, 18–27. [\[CrossRef\]](#)
- Burlibayeva, D.M.; Burlibayev, M.Z.; Opp, C.; Bao, A.M. Regime dynamics of hydrochemical and toxicological parameters of the Irtysh River in Kazakhstan. *J. Arid Land* **2016**, *8*, 521–532. [\[CrossRef\]](#)
- Bordalo, A.A.; Teixeira, R.; Wiebe, W.J. A Water Quality Index Applied to an International Shared River Basin: The Case of the Douro River. *Environ. Manag.* **2006**, *38*, 910–920. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cai, J.; Wang, X.; Cai, Y.; Wei, C.; Liao, Z.; Li, C.; Liu, Q. An integrated framework consisting of spatiotemporal evolution and driving force analyses for early warning management of water quality. *J. Clean. Prod.* **2024**, *462*, 142628. [\[CrossRef\]](#)
- Horton, R. An index number system for rating water quality. Water quality criteria-stream vs. effluent standards. *J. Water Pollut. Control Fed.* **1965**, *37*, 292–315.
- Paun, I.; Cruceru, L.; Chiriac, F.L.; Niculescu, M.; Vasile, G.; Marin, N.M. Water Quality Indices-methods for evaluating the quality of drinking water. In Proceedings of the 19th INCD ECOIND International Symposium—SIMI 2016, “The Environment and the Industry”, Bucharest, Romania, 13–14 October 2016; pp. 395–402. [\[CrossRef\]](#)
- Uddin, M.G.; Nash, S.; Olbert, A.I. A review of water quality index models and their use for assessing surface water quality. *Ecol. Indic.* **2021**, *122*, 107218. [\[CrossRef\]](#)
- Khan, I.; Zakwan, M.; Mohanty, B. Water Quality Assessment for Sustainable Environmental Management. *ECS Trans.* **2022**, *107*, 10133. [\[CrossRef\]](#)
- Cicilinski, A.D.; Virgens Filho, J.S. A new water quality index elaborated under the Brazilian legislation perspective. *Int. J. River Basin Manag.* **2022**, *20*, 323–334. [\[CrossRef\]](#)
- Benkov, I.; Varbanov, M.; Venelinov, T.; Tsakovski, S. Principal Component Analysis and the Water Quality Index—A Powerful Tool for Surface Water Quality Assessment: A Case Study on Struma River Catchment, Bulgaria. *Water* **2023**, *15*, 1961. [\[CrossRef\]](#)
- Yotova, G.; Varbanov, M.; Tcherkezova, E.; Tsakovski, S. Water quality assessment of a river catchment by the composite water quality index and self-organizing maps. *Ecol. Indic.* **2021**, *120*, 106872. [\[CrossRef\]](#)

22. de Andrade Costa, D.; Soares de Azevedo, J.P.; dos Santos, M.A.; dos Santos Facchetti Vinhaes Assumpção, R. Water quality assessment based on multivariate statistics and water quality index of a strategic river in the Brazilian Atlantic Forest. *Sci. Rep.* **2020**, *10*, 22038. [\[CrossRef\]](#)
23. Jo, C.; Kwon, H.; Kim, S. Temporal and Spatial Water Quality Assessment of the Geumho River, Korea, Using Multivariate Statistics and Water Quality Indices. *Water* **2022**, *14*, 1761. [\[CrossRef\]](#)
24. Panagopoulos, Y.; Alexakis, D.E.; Skoulikidis, N.T.; Laschou, S.; Papadopoulos, A.; Dimitriou, E. Implementing the CCME Water Quality Index for the Evaluation of the Physicochemical Quality of Greek Rivers. *Water* **2022**, *14*, 2738. [\[CrossRef\]](#)
25. Chidiac, S.; El Najjar, P.; Ouaini, N.; El Rayess, Y.; El Azzi, D. A comprehensive review of water quality indices (WQIs): History, models, attempts and perspectives. *Rev. Environ. Sci. Bio/Technol.* **2023**, *22*, 349–395. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Radelyuk, I.; Naseri-Rad, M.; Hashemi, H.; Persson, M.; Berndtsson, R.; Yelubay, M.; Tussupova, K. Assessing data-scarce contaminated groundwater sites surrounding petrochemical industries. *Environ. Earth Sci.* **2021**, *80*, 351. [\[CrossRef\]](#)
27. Kazhydromet. Kazakhstan Annual Bulletin of Climate Monitoring: 2019 (In Russian). Astana, Kazakhstan. 2020. Available online: <https://www.kazhydromet.kz/uploads/files/133/file/6005153f3e6f2ezhegodnyy-byulleten-monitoringa-sostoyaniya-i-izmeneniya-klimata-kazahstana-za-2019.pdf> (accessed on 26 May 2024).
28. Lumb, A.; Halliwell, D.; Sharma, T. Application of CCME Water Quality Index to Monitor Water Quality: A Case Study of the Mackenzie River Basin, Canada. *Environ. Monit. Assess.* **2006**, *113*, 411–429. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Yan, F.; Liu, L.; Li, Y.; Zhang, Y.; Chen, M.; Xing, X. A dynamic water quality index model based on functional data analysis. *Ecol. Indic.* **2015**, *57*, 249–258. [\[CrossRef\]](#)
30. Atakhanova, Z.; Meirambayeva, M.; Baigaliyeva, M. Mine Water Use in Kazakhstan: Data Issues, Risks, and Regulations. *Sustainability* **2024**, *16*, 2456. [\[CrossRef\]](#)
31. Khan, M.H.R.B.; Ahsan, A.; Imteaz, M.; Shafiquzzaman, M.; Al-Ansari, N. Evaluation of the surface water quality using global water quality index (WQI) models: Perspective of river water pollution. *Sci. Rep.* **2023**, *13*, 20454. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Mallya, Y.J. *The Effects of Dissolved Oxygen on Fish Growth in Aquaculture*; The United Nations University Fisheries Training Programme, Final Project: Reykjavik, Iceland, 2007.
33. Franke, N.; Hoekstra, A.Y.; Boyacioglu, H. *Grey Water Footprint Accounting: Tier 1 Supporting Guidelines*; UNESCO-IHE: Delft, The Netherlands, 2013.
34. WHO. *Petroleum Products in Drinking-Water: Background Document for Development of WHO Guidelines for Drinking-Water Quality*; WHO: Geneva, Switzerland, 2008.
35. Rickwood, C.J.; Carr, G.M. Development and sensitivity analysis of a global drinking water quality index. *Environ. Monit. Assess.* **2009**, *156*, 73–90. [\[CrossRef\]](#)
36. Lindenschmidt, K.-E.; Baulch, H.M.; Cavaliere, E. River and Lake Ice Processes—Impacts of Freshwater Ice on Aquatic Ecosystems in a Changing Globe. *Water* **2018**, *10*, 1586. [\[CrossRef\]](#)
37. Chambers, P.A.; Brown, S.; Culp, J.M.; Lowell, R.B.; Pietroniro, A. Dissolved oxygen decline in ice-covered rivers of northern Alberta and its effects on aquatic biota. *J. Aquat. Ecosyst. Stress Recovery* **2000**, *8*, 27–38. [\[CrossRef\]](#)
38. Pokrovsky, O.S.; Manasypov, R.M.; Loiko, S.V.; Krickov, I.A.; Kopysov, S.G.; Kolesnichenko, L.G.; Vorobyev, S.N.; Kirpotin, S.N. Trace element transport in western Siberian rivers across a permafrost gradient. *Biogeosciences* **2016**, *13*, 1877–1900. [\[CrossRef\]](#)
39. Babuji, P.; Thirumalaisamy, S.; Duraisamy, K.; Periyasamy, G. Human Health Risks due to Exposure to Water Pollution: A Review. *Water* **2023**, *15*, 2532. [\[CrossRef\]](#)
40. Rakhymberdina, M.Y.; Grokhotov, E.V.; Assylkhanova, Z.A.; Toguzova, M.M. Using Space Survey Materials for Modeling Hydrodynamic Accidents at Mining Enterprises in Kazakhstan. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *XLVI-5/W1-2022*, 193–198. [\[CrossRef\]](#)
41. Kazhydromet. Bulletin of Environmental Conditions of the East-Kazakhstani and Abay Regions (In Russian). Astana, Kazakhstan. 2023. Available online: https://www.kazhydromet.kz/uploads/calendar/146/year_file/65e84811bab16vko-byulleten-2023-rus-yaz--1.pdf (accessed on 26 May 2024).
42. Tileugabylov, A.; Madani, N. Evaluation of Water Contamination in the East Kazakhstan Mining Area Using Multivariate Geostatistics. In Proceedings of the EGU General Assembly Conference Abstracts, Vienna, Austria, 23–28 April 2023; p. EGU–1737.
43. Statistics. Industrial Production in the East Kazakhstan Region (1990–2022). Available online: <https://stat.gov.kz/api/iblock/element/region/22512/file/ru/> (accessed on 8 July 2024). (In Russian)
44. Kazakhstan. Entrepreneurial Code. Available online: <https://adilet.zan.kz/rus/docs/K1500000375> (accessed on 8 July 2024).
45. Golian, S.; El-Idrissy, H.; Stambuk, D. Using CMIP6 Models to Assess Future Climate Change Effects on Mine Sites in Kazakhstan. *Hydrology* **2023**, *10*, 150. [\[CrossRef\]](#)
46. Andraka, D.; Ospanov, K.; Myrzakhmetov, M. Current state of communal sewage treatment in the Republic of Kazakhstan. *J. Ecol. Eng.* **2015**, *16*, 101–109. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Kalmakhanova, M.S.; Diaz de Tuesta, J.L.; Malakar, A.; Gomes, H.T.; Snow, D.D. Wastewater Treatment in Central Asia: Treatment Alternatives for Safe Water Reuse. *Sustainability* **2023**, *15*, 4949. [\[CrossRef\]](#)
48. Ospanov, K.; Rakhimov, T.; Myrzakhmetov, M.; Andraka, D. Assessment of the Impact of Sewage Storage Ponds on the Water Environment in Surrounding Area. *Water* **2020**, *12*, 2483. [\[CrossRef\]](#)

49. Radelyuk, I.; Klemeš, J.J.; Tussupova, K. Sustainable Water Use in Industry—Reasons, Challenges, Response of Kazakhstan. *Circ. Econ. Sustain.* **2023**, *3*, 2267–2283. [[CrossRef](#)]
50. Kazakhstan. Ecological Code of the Republic of Kazakhstan. Available online: <https://adilet.zan.kz/rus/docs/K2100000400#z4470> (accessed on 26 May 2024). (In Russian).
51. George, J.; Thakur, S.K.; Tripathi, R.C.; Ram, L.C.; Gupta, A.; Prasad, S. Impact of coal industries on the quality of Damodar river water. *Toxicol. Environ. Chem.* **2010**, *92*, 1649–1664. [[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.