Source Apportionment of Dissolved Ions in a Watershed Using Neural Network and Non-Parametric Analysis

Kobe University: Okechi Selline

1.0 INTRODUCTION

Cold water environment issues are complex, involving snow accumulation and melting, river and lake freezing and thawing, unique climate and industrial pollution sources. Understanding the complex climatic ion dynamics within the catchment basins in these regions is crucial for effective water quality management, yet predicting these variations remains a significant challenge. These dynamics are governed by non-linear interactions, where seasonal weather patterns, directly influence hydrological pathways, altering the transport rates of dissolved ions into the stream network. As these relationship pathways dynamically switch between dilution and concentrated transportation across different seasons, traditional linear modelling often fails to capture the system's full complexity. Therefore, this study implements a sequential data-driven workflow for source apportionment of dissolved ions in Kushiro River Catchment basin.

2.0 RESEARCH OBJECTIVES

1. Develop a neural network model for predicting dissolved ion concentrations based on hydrological and land-use parameters.
2. Employ non-parametric analysis to confirm the statistical significance of potential source contributions.
3. Identify key environmental and anthropogenic factors influencing the distribution of dissolved ions within the Kushiro River Catchment basin.

3.0 METHODS

Neural Network Analysis (NNA) was used as the primary predictive tool to capture the complex, non-linear relationships driving dissolved ion transport. A separate, secondary analytical framework was executed to validate our predictions. Specifically, Principal Component Analysis (PCA) was employed to extract the underlying linear source factors and reduce data dimensionality, while the non-parametric Mann-Whitney test assessed the statistical differences across distinct seasons. By utilising these multivariate statistical methods independently, the study provided a rigorous, independent justification of the NNA outputs, transforming raw machine-learning predictions into a physically meaningful and statistically validated source apportionment framework. Data was collected from various locations within the Kushiro River Catchment basin. Sampling stations were classified into five categories, with the concentrations from each category serving as inputs and the concentration at the downstream end as the output. Our Neural Network Analysis (NNA) model assumed independent groups, with dissolved substances being transported directly to the downstream end. A modified sigmoid function was employed as the activation function to introduce non-linearity. Connection weights were adjusted based on the backpropagated error between observed and estimated outcomes. The supervised machine learning process involved immediate weight adjustment upon input data entry, minimising the error between the desired and predicted outputs. The NNA calculation was performed 10,000 times to reach a steady state for w1i and w2i, ensuring consistent

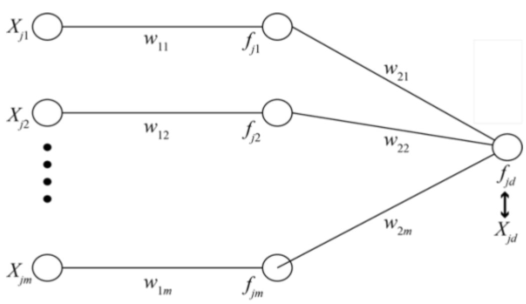

Fig.1: NNA model applied in this study

transportation rates of dissolved ions from each group to the downstream end. Additionally, PCA was utilised to elucidate the distribution of key ions in the Kushiro River basin. The dataset was normalised using the Z-score formula to standardise its mean to zero and standard deviation to one, thereby reducing the impact of varying measurement scales. PCA was performed on the dataset, with results displayed in both milligrams per litre (mg/L) and moles per litre (mol/L) for the winter and autumn seasons.

4.0 RESULTS

The results showed a consistent pattern in dissolved ion concentration with slight variations in group contributions during spring and autumn. Transportation rates were higher in spring. The rates were significantly higher when moles/l units were used compared to when mg/L units were used. The difference arises because concentration units influence how ion behaviour is represented and interpreted. Transport rates are affected by particle characteristics such as charge and size, which are more directly related to molar concentration. The neural network model showed a notable discrepancy for nitrate ions, with higher error values, prompting further investigation. To understand nitrate ions' impact on ion distribution in the Kushiro River catchment basin, samples were re-analysed excluding nitrate concentrations, and the findings were compared. In addition, when a two-scatter plot with two principal components was applied higher concentration of dissolved ions were observed in various stations. When nitrate ions were excluded from the analysis, notable shifts in station locations were observed, suggesting that nitrate ions significantly impacted these variations.

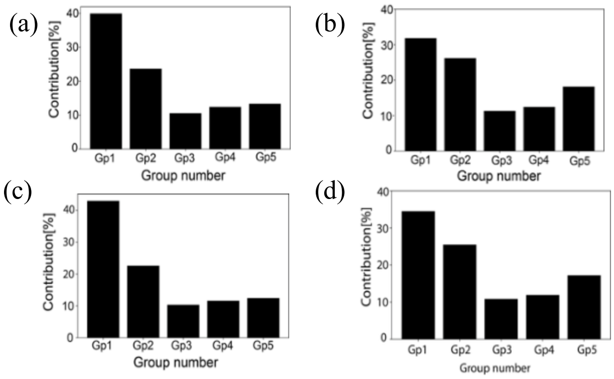

Fig. 2: Dissolved ion transportation rates (a)Winter(mg/l) (b)Autumn(mg/l) (c)Winter(mol/l) (d)Autumn(mol/l)

5.0 CONCLUSION

A combination of machine learning and non-parametric statistical tests effectively analysed the distribution of dissolved ions in the Kushiro River catchment basin. The Mann–Whitney U-test, suitable for non-normally distributed data, revealed that Group 1, located in a caldera lake region, had the highest ion concentrations. Seasonal fluctuations in ion concentrations were also noted, likely due to biological processes. Mol/L results provided a clear picture of ion distribution. These findings are significant in advancing data-driven approaches particularly in regions where conventional parametric techniques may be limited due to the non-linear nature of environmental systems.

6.0 REFERENCES

1. Komai, K Komai, K.; Nakayama, K.; Hiroki, S.; Sakaguchi, S.; Estimation of dissolved substances transport in Kushiro River basin using dissolved ion pattern, Journal of Japan Society of Civil Engineers Ser B1(Hydraulic Engineering), 2018, Vol. 74, Issue 4, pp.1505-1510.
2. Beitia, C.; Nakayama, K.; Maruya, Y.; Ohtsu, N.; Yamasaki, S.; Yamane, M.; Wyatt, A.; Komai, K. Suspended sediment transport estimation by X-ray fluorescence analysis. Journal of Hydrologic Engineering. 2016, 72, 1039-1044. doi: 10.2208/jscejhe.72. I_1039.
3. Jing Liu, Hu Ding, Min Xiao, Zhu-Yan Xu, Yuan Wei, Zhi-Hua Su, Lei Zhao, Jiao-Ting Peng, Heng Wang.; Xiao-Dan Wang. Climatic Variabilities Control the Solute Dynamics of Monsoon Karstic River: Approaches from C-Q Relationship, Isotopes, and Model Analysis in the Liujiang River. Water 2020, 12(3), 862. doi:10.3390/w12030862.
4. Bojarczuk, A.; Biernacka, A. Transformation of water chemistry in the Stara River under climate and land use changes in the Carpathians. *Sci Rep*.2025, **15**, 34463. doi.org/10.1038/s41598-025-17660-4
5. Roy, A. Two-stage principal component analysis on interval-valued data using patterned covariance structures. Advances in Data Analysis and Classification (2025). doi.org/10.1007/s11634-025-00650-9