The transport and accumulation processes of geochemical tracers in environmental compartments

by Jianguo Li

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of Daniel Ramp, Qiang Yu

University of Technology Sydney
Faculty of Science

February 2021

Certificate of Original Authorship

I, Jianguo Li declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Life Science/Faculty of Science at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.
This research is supported by an Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 3 February 2022

Acknowledgements

This four-year PhD journey is a special experience for me, everyday full of many possibilities and challenges. It was not very easy, sometimes even struggling, but I do have learned a lot during this period. I would like to thank many people and institutions involved to help me complete this project.

First of all, I want to express my sincere gratitude to my principal supervisor Professor Qiang Yu. He gave me many useful academic advices about my research, and also encouraged me to be confident when I am feeling difficult in my PhD journey, especially the last two years during the COVID pandemic. He is just like a lighthouse during my darkness, whenever I need help, he is always around. I would like to express my special appreciation to my co-supervisor, Associate Professor Daniel Ramp. I remember there was a slogan printed on Dan's office mug, 'think in a different way', and he did the same. His expertise from conservation always inspired me with my hydrology, and we worked a lot together with my project, especially the first 2 years. He taught me a lot of research skills, like R programming, and academic writing. He makes me one of the luckiest PhD in the world. My sincere appreciation also goes to Dr. Kate Brandis at UNSW. She went to UTS many times in person to help me design my project, and gave many helps at the BEEs lab, UNSW. I would also thank Professor Zongxing Li from CAS, Professor Ziyong Sun and Dr. Jianwei Bu from China University of Geosciences, and Professor Xiaocong Qiu from Ningxia University to share their data for my thesis.

I thank my colleagues from UTS, Dr. Hong Zhang, Dr. Song Leng, Dr. Jie He, Dr. Wenjie Zhang, Dr. Mingxi Zhang, Dr. Lijie Shi, Dr. Qinggaozi Zhu, Dr. Qinsi He, Dr. Qiaoyun Xie, Dr. Baoming Wang, Dr. Shihui Wen, Dr. Du Li, and Dr. Yuxia Liu, for their help and accompany. I also want to thank my colleagues from Centre For Compassionate Conservation, Dr. Eamonn Wooster, Dr. Chris Hasselerharm, Dr. Catlin Austin, Dr. Esty Yanco, Dr. Gavin Bonsen, Dr. Boyu Ji, Dr. Eric Lundgren, Dr. Arian Wallach, Dr. Finbarr, Dr. Loic Juillard, and Dr. Miriam Zemanova. I am also grateful to staffs at UTS, Shannon Hawkins, Maggie Chen, Simon Mitrovic, and Alfredo Huete. Thanks to Dr. Rendy Ruvindy and Dr. May Supasri, we always hiked, run and swam together, and this made my research life more colorful and released. Thanks also to Dr. Matt Balzer, Dr. Jordan Facey, Dr. Laurel Michie, we always had lunch together, and that time was always full of happy and comfortable.

I would like to express my gratitude to the colleagues at UNSW, Professor Richard Kingsford, Associate Professor Mark Ooi, Dr. Roxane Francis, Dr. Justin McCann, Sharon Ryall, and Chantelle, for their kind help when I was doing the feather-scanning work at

BEES, UNSW. And Justin helped fixed my bicycle many times, which guaranteed my normal commute to UNSW. Thanks also to my friends in the gym, in the fitness station of Wentworth Park, and to the beautiful Bondi beach, Coogee beach, and many other beautiful beaches where I often go outside of my research time, and brought me a lot of joy. All the memories of Sydney make me unforgettable.

I went to Australia for my PhD research when my baby daughter, Baiyutong Ruby Li was only 2 months old, and she is my motivation and always reminds me keep moving for the PhD no matter how difficult it is. I want to thank the understanding and support from my family. My wife, Liping Bai became a superhero during this time, because she not only needed to work, but also to take care of our baby infant. I also want to thank my dog Coco, for her understanding of my leaving and her company to the family when I was abroad. If I can complete my PhD, they have half the credits.

Finally, I would like to thank the UTS Top Up scholarship, and the CSC-UTS scholarship for financial support.

Publications arising from this thesis

Journal papers directly included in this thesis

Li J, Li Z, Brandis KJ, Bu J, Sun Z, Yu Q, Ramp D. 2020. Tracing geochemical pollutants in stream water and soil from mining activity in an alpine catchment. Chemosphere 242:125167. https://doi.org/10.1016/j.chemosphere.2019.125167 (Chapter 4)

Contents

List of Figures

List of Tables

Glossary

AIC	Akiake Information Criterion
ANOVA	Analysis of variances
APCS/MLR	absolute principal component score/multiple linear regression
BJD	Badain Jaran Desert
DEM	digital elevation model
EC	electric conductivity
EMMA	End member mixing analysis
GAM	generalised additive models
GBM	gradient boosted machine
GDAS	Global Data Assimilation System
GLM	generalised linear models
GMWL	global meteoric water line
GNIP	Global Network of Isotopes in Precipitation
GPCC	Global Precipitation Climatology Center
HM	heavy metal
HRB	Heihe River Basin
HYSPLIT	Hybrid Single-Particle Lagrangian Integrated Trajectory
LMM	linear mixed models
LMWLs	local meteoric water lines
MLR	multiple linear regression
MODIS	Moderate Resolution Imaging Spectroradiometer
NCEP/NCAR	National Centers for Environmental Prediction/National Center for Atmospheric Research
NDVI	normalized difference vegetation index
OLR	outgoing longwave radiation
PCA	principal component analysis
PMF	positive matrix factorization
QTP	Qinghai-Tibet Plateau
RDA	Redundancy analysis
RF	random forest
SDM	species distribution model
TDS	total dissolved solids
TEM	transmission electron microscopy
TPI	Topographic Position Index

TRI	Terrain Ruggedness Index
WCP	water chemical parameters
ZMSK	Zhamashike

Abstract

The ubiquitous geochemical tracers tend to persistently transport and accumulate in different environmental compartments, and a better understanding of geochemical tracer transport mechanism is essential for conservation and resource management. In this thesis, systematic data sets were collected from many typical environmental compartments both in China and Australia, including the catchment, desert, and typical mining sites at an alpine stream etc. Various geochemical tracers and methodologies, such as heavy metals and stable isotopes were applied to make a comprehensive exploration of the interactive impacts of human activity and natural processes on the transport and accumulation of geochemical tracers. The main findings of this thesis are:

(1) The species distribution model (SDM) can help improve the prediction accuracy of mapping geochemical tracers by considering the important explanatory variables. The prediction results of Chapter 2 and Chapter 3 confirmed the effectiveness and potential of leveraging SDMs from ecology to study heavy metal contamination in the field of hydrology, which offered new insights to understand the relationships between HMs and the human and physical environment.

(2) The combinations of interdisciplinary, multi-methods, various geochemical tracers, and indicators, facilitate the improvement of discovering the transport processes of tracers. In Chapter 3, the end member mixing analysis (EMMA) and SDM were combined to study the vertical hydraulic connections between surface water and groundwater. The water chemical tracers in river water were found to be able to serve as good explanatory variables to predict the HMs in soil in Chapter 4. And both elemental concentration and bioaccumulation ratio were used as the indicators to study the bioaccumulation processes in Chapter 6, which can provide more detailed information, otherwise the spatial attributes of the geochemical background information will be ignored. All these results supported this perspective.

(3) The model selection and model averaging methods constitute an advancement in quantitatively interpreting the relationships between isotopic signatures of precipitation and their local climatic variables. This method can supplement or precede more complex studies of hydrological cycles utilizing isotope tools.

Overall, the combination of various chemical tracers, as well as the introduction of powerful methodologies from related research fields, can effectively help improve the analysis accuracy. This thesis offers an improved understanding of the transport and accumulation of chemical tracers between different environmental compartments, and can provide useful

database at the data scarce areas. The results of this thesis can also provide scientific strategies for the regional development and management.

Keywords: environmental compartments, carriers, geochemical tracers, bioaccumulation, catchment, hydrological units, mining activity, spatial distribution models, model selection, arid and semi-arid regions, China, Australia

Chapter 1. Introduction

1.1 Research background

1.1.1 Transport of elements between environmental compartments

The different forms of chemical elements, also known as geochemical tracers, have been widely used as 'proxies' to study the earth system sciences. The interactions of geochemical tracers between different environmental compartments are driven by both the natural environment and anthropogenic activity. Potential sources of these tracers may be mobilized by a series of natural activities, including the volcanic activity (Process g in Figure 1-1), the weathered and eroded from the crustal material (Process e & h in Figure 1-1), and forest fires (Process f in Figure 1-1). Additionally, rapid economic growth over recent decades has necessitated greater levels of natural resources extraction and exploitation, and required to provide important chemical elements needed for developing populations. The development of technology has made it possible to extract more natural resources, but meanwhile the anthropogenic activity will extract unprecedented contaminants into the environment. These anthropogenic activity releases can give rise to higher concentrations of contamination, imposing great potential threats to human health and environmental sustainability.

Geochemical tracers have been reported to actively transport within multiple environmental compartments, such as the catchment (Hassellov and von der Kammer, 2008), urban system (Sarkar et al., 2021), agriculture (Meite et al., 2018), estuary (Fang et al., 2016), mining sites and public transport system etc., (Goth et al., 2019; Senduran et al., 2018; Stojic et al., 2017), spreading from local, regional, to global scales. The main transport processes of chemical elements in the cycling systems of different environmental compartments were presented in Figure 1-1. The atmospheric deposition from the atmosphere consistently brings chemical elements from the aerosols into the geosphere and hydrosphere (Nickel et al., 2014). The soil on the other hand plays as an elemental sink for the suspended particles, dust, and sand, and a place where the irrigation and seepage water pollutants accumulate. For instance, important sources from anthropogenic emission are mining (Process i in figure 1-1), disposal of sewage effluents (Process a in figure 1-1), fertilizers (Process d in figure 1-1), vehicle emissions, lead-acid batteries and aging water supply infrastructure (Process b & c in figure 1-1). The elements suspended particles from tailing soil in close proximity of industrial and mining areas can then travel to a long distance during the precipitation erosion and runoff generation processes (Li et al., 2020). In addition, modern fluvial sediments in hydrodynamic sorting process also influence the river water elemental contents via the water-rock interaction. The HMs converge into river water, which would be deposited on floodplain sediments (Hassellov

and von der Kammer, 2008) or transported farther downstream by the floods and channel aggregation (Hochella Jr et al., 2005). Eventually, elemental signatures will be biomagnified or bioaccumulated within biosphere, which involves aquatic organisms (Process l in Figure 1-1), insects, potentially mammals (Process k in Figure 1-1).

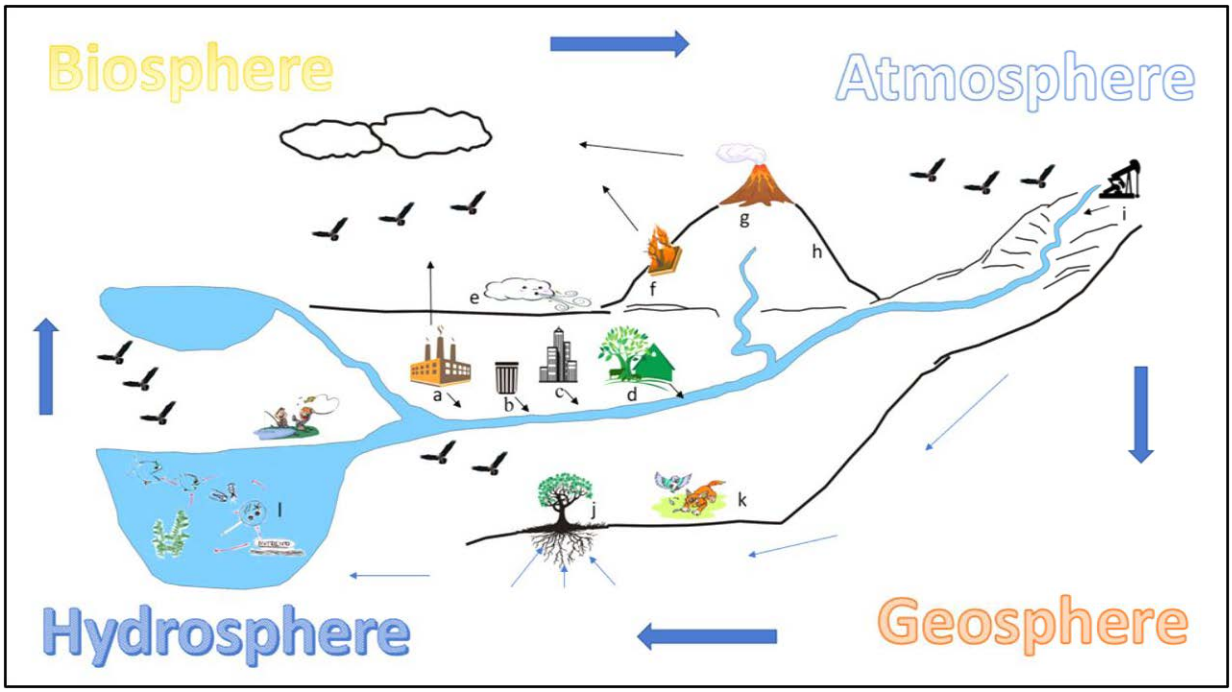


Figure 1-1. The sources and transport of chemical elements in the cycling systems of hydrosphere, biosphere, atmosphere, and geosphere. Transport processes include (a) stands for the disposal of sewage effluents, (b) waste, such as lead-acid batteries, (c) the combustion of energy fuel, (d) the pesticides and fertilizer from agriculture, (e) wind- blown dust, (f) forest and grass fire, (g) volcano activity, (h) natural crust, (i) mining activity, (j) absorption from the groundwater and soil, and (k) & (l) bioaccumulation in different trophic levels.

The chemical elements can exist in multiple forms in the natural environment. Heavy metals (HMs), hydrochemistry, and stable isotopes were focused on in this thesis, and literature review of their significance were presented in 1.1.2, and 1.1.3 part. And in 1.1.5 part, the application of novel non-lethal bioindicator method by bird feather was presented.

1.1.2 HMs and adverse effects

These chemical elements have been divided into many types according to the different purposes of research. For instance, from the perspective of environmental contaminants, the chemical elements are often divided into trace elements (Abdullah et al., 2015), heavy metals (Frantz et al., 2012; Furtado et al., 2019), metals and metalloids (Badry et al., 2019; López-Perea et al., 2019), etc. Some elements are essential and nutrients for the proper functioning

of the human body (Prashanth et al., 2015), including Ca, Co, Cu, Fe, Mg, and Zn, etc., while some elements are toxic to the human body, interfere with its functioning and undermine health, and a small amount could be fatal (Dudka and Miller, 1999), like As, Cd, Pb, Al, Li and Hg. These toxic elements have no known physiological functions, but are potentially fatal. However, some essential elements, if the intake exceeds the standard, will threaten health, such as Mn, Mo (0.05 μg/ day), Cr (50-200 μg/ day). In addition, the compounds of some essential elements can be very harmful, such as the hexavalent chromium compounds (Costa and Klein, 2006).

Contaminant elements can directly degrade air, water, and soil quality, and accumulate in animals through food consumption. Exposure to contaminant elements may cause endocrine and nervous disorders, certain physiological and behavioral abnormalities. On account of the widespread distribution and extensive use, Hg, Pb, Cr, As, and Cd are the most toxic elements.

The volcanoes, forest fires, ore extraction, and fossil fuels such as coal and petroleum, are the main sources of Hg (Pirrone et al., 2010). Mercury levels in the industrial environment are increasing due to mercury mining, discharge from hydroelectric, chlor-alkali plants, paper industries, refineries and dental clinics as well as thermometer factories (Brooks et al., 2008). The inorganic and organic mercury are accumulated in the bodies of organisms that prey on those that have bio-accumulated mercury, so the diets are also normally the main source in biosphere, such as fish and other seafood. The three main chemical forms of mercury are mercury vapour, inorganic mercury compounds and organic mercury compounds (Morel et al., 1998). Exposure to them has serious harmful effects on the central nervous system, developing nervous system and immune systems (Zahir et al., 2005). Mercury can also influence the reproductive system by threatening the development of the fetus's brain, or cause attention deficit and developmental delays in early life (Tas et al., 1996). The mercury may even directly harm the organs, such as kidneys, liver, lung, heart, eyes, skin and brain (Mahurpawar, 2015; Wolfe et al., 1998; Xu et al., 2018).

Lead is sourced from the wind-blown soil and road dust, and it is also from industrial sources, deteriorated paint, and the combustion of leaded gasoline and aviation fuel. The anthropogenic sources of chromium include cement-producing plants, manufacture of chromium-based products, emissions of chromium-based natural energy consumption, and cigarette smoke (Marschall and Schumacher, 2012). Lead can damage the central nervous system, and a statistically significant correlation between the lead exposure and violent crime in the United States also demonstrated this result (Nevin, 2000).

Arsenic is present in more than 200 different minerals, and mining activity is the major process that contributes to arsenic contamination of air, water and soil (Zhu et al., 2008). About one-third of the arsenic in the Earth's atmosphere is of natural origin from volcanic activity. Chromium and Arsenic have chronic carcinogenic effect at high concentration levels (Mandal and Suzuki, 2002). Cadmium is a naturally occurring toxic heavy metal with common exposure in industrial workplaces, plant soils, and from cigarette smoking. Cadmium and Cadmium compounds cause bone demineralization disease or impair lung function.

The mining activity is one of the most important sources for contaminant elements, which typically require large quantities of water to process ore (Johnson and Hallberg, 2005). Discharges of mine effluent and seepage from tailings may lead to serious contamination of freshwater resources (Candeias et al., 2015; Fetter et al., 2017). Mining pollutants can drain directly into local run-off or can be dispersed into a river via rainfall or other hydrological processes. Nonpoint source pollutants that infiltrate deep soil or enter streams can be transported into aquifers by the exchange process between groundwater and surface water.

1.1.3 Hydrochemistry and isotopic tracers

The hydrochemistry method is the most straightforward and cost effectively way to trace, dig and interpret the influence of hydrological cycles on geochemical tracers' transport information (Wang et al., 2018). The hydrochemistry tracers include major soluble ions, pH, EC, TDS, and stable isotopes etc. The hydrochemistry methods offer advantages over some physically based methods like the large water level gauging or monitoring networks, not only due to a much lower cost (Cook et al., 2013), but also providing more accurate information on the interaction and transportation over large spatial regions. These advantages enable water chemistry methods to be widely applied for ascertaining source and origin, or geochemical evolution processes (Liu et al., 2015), the interactions between different water bodies (Maurya et al., 2019; Wang et al., 2018), and monitoring the contamination transport and accumulation mechanisms in the hydrological cycles (Hatzinger et al., 2013; Khalil et al., 2015; Liu et al., 2019). The research scale has been spatially extended to Arctic and Antarctic areas (Hodson et al., 2002; Hoshina et al., 2016; Hoshina et al., 2014).

Water stable isotopes of $\delta^{18}O$, δD, and deuterium excess (d-excess=$\delta D-8\delta^{18}O$) offer a powerful proxy for evaluating the hydrological cycles, such as precipitation (Price et al., 2008; Risi et al., 2008), glacier (Kumar et al., 2018), permafrost water (Lacelle et al., 2014), surface water (Ogrinc et al., 2008), groundwater (Abbott et al., 2000; Wassenaar et al., 2009), and plant water (Bush et al., 2017), etc. Due to their stable properties, the water isotopic

signatures have been widely used to analyze various hydrological processes, like the hydraulic connections between surface water and groundwater (Yin et al., 2011), stream water recharge mechanism (Li et al., 2014), hydrograph separation (Lyon et al., 2009), and the influence of climate change on cryosphere (Kreutz et al., 2003; Li et al., 2016).

1.1.4 Bioaccumulation signals by bird feathers

Tracing the accumulation of chemical elements in organisms is an extremely complicated process, which will be affected by both the biotic and abiotic factors. To discover these processes, the prerequisite is to find a reliable entity that must be closely connected with both the environment and human activity. Birds are very good potential candidate to verify the above hypothesis. Birds are more sensitive to the surroundings than the human beings and other animals (Wretenberg et al., 2010). The foraging, breeding and other biological behaviors will be in close contact with the environment, which will be eventually recorded in the bird's body. Colonial birds' nests or roosts are in proximity at a particular location (Brandis et al., 2018), and this provides possibility for specific areas of study. In some areas where there is shotgun hunting, the birds that select hard grit for grinding food are extremely good sentinels to monitor the Pb contamination (Newth et al., 2013), as the grits are similar in weight and size to the shotgun pellets that deposit heavily in their habitats. Bird feathers are outstanding to trace the element bioaccumulation mechanism. Compared with some lethal sampling research methods like blood, liver, muscle, kidney and other internal organs (López-Perea et al., 2019), feathers are a nondestructive method. In addition, when compared with other non-invasive method such as eggshells, feathers can accumulate and record more information over a longer period (Abbasi et al., 2015). Previous bird feather bioaccumulation research covers wide trophic levels from herbivore to carnivore predators, as well as being over a large spatial range, even including Antarctica (Metcheva et al., 2011), and long-term temporal records from the museum feathers (Frederick et al., 2004). The 'bird feather' was inserted as the key word in Google Scholar and Web of Science, to search any online published literature that had applied the bird feather as bioindicator tools. The distribution map of sampling sites for the elemental bioaccumulation research using bird feathers according to these literatures were created as Figure 1-2. The environment signatures in bird feathers are presented as below.

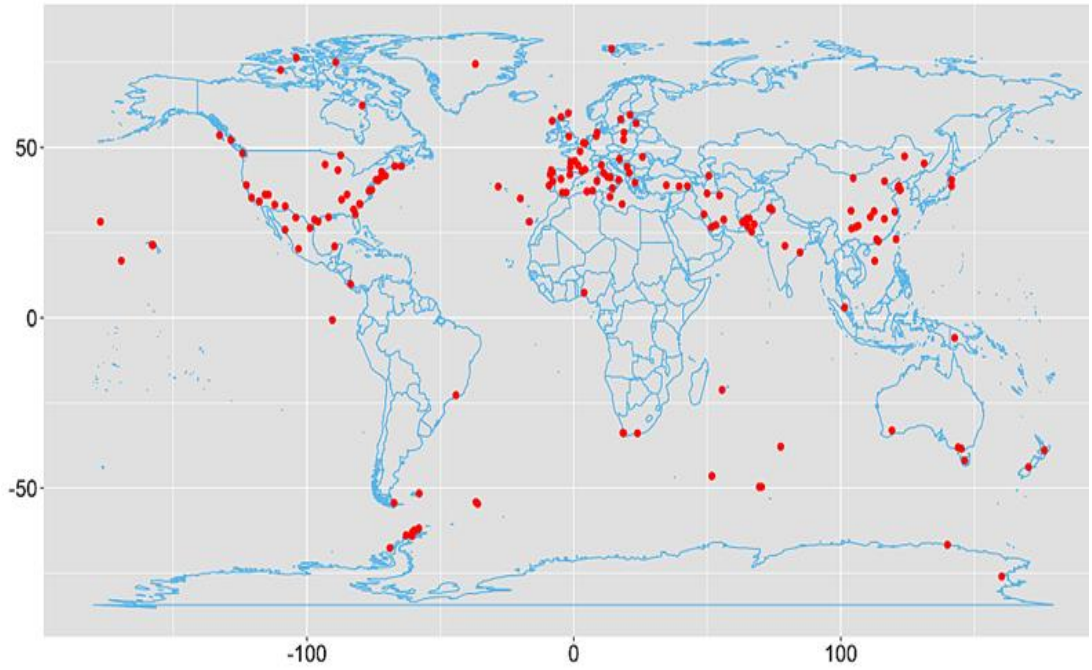


Figure 1-2. The distribution map of sampling sites for the elemental bioaccumulation research using bird feathers according to literature review.

1.1.4.1 Trophic signals

The stable isotopes in feather were able to trace the trophic signatures in food web (Thompson and Furness, 1995). The ratios of stable isotopes of N^{15}/N^{14} and C^{13}/C^{12}, always were prescribed as $\delta^{15}N$ and $\delta^{13}C$, can give detailed information of the structure and interaction among the food webs. The $\delta^{15}N$ in organisms showed strong trophic signals, for example, the declining $\delta^{15}N$ levels in Northern Fulmars Fulmarus indicated their declining trophic level (Thompson and Furness, 1995). The $\delta^{13}C$ gives weak trophic information, but it can better show the dietary sources from different "forage base". Previous research found that the waterbirds were more depleted in $\delta^{13}C$ for the terrestrial consumers, -28‰ (Jury et al., 1987), than marine consumers, -19~-24‰ (Peterson and Fry, 1987). Variations in $\delta^{13}C$ can also reflect the different foraging location with a specific atmosphere condition (Farmer and Leonard, 2011).

Trophic shifts over time can be analyzed using historical feather records, e.g., museum collections. Comparing the contemporary feather isotope ratios with historical records, the trophic level shifts can be identified (Farmer and Leonard, 2011), and the triggers can be inferred because many birds will change their feeding strategy in response to the climate change, as well as human activity. According to previous research, the flesh- footed shearwaters' $\delta^{13}C$ values in feathers decreased from -16.0±0.8‰ in 1936 to -17.4±1.6 ‰ in 2011 in response to the dietary items shifting, and the $\delta^{15}N$ also showed a decreased trend

from +16.9±1.5‰ in 1936 to +13.9±1.6 ‰ in 2011 (Ofukany et al., 2014). And Quillfeldt et al. (2010) found the Thin-billed prions become more dominant in poleward areas because the diet had been changed due to the climate change by comparing the feather δ ^{13}C isoscapes from different periods.

Except for the feather collected from the field, many controlled experiments had also been conducted under laboratory conditions (Mizutani et al., 1992; Pearson et al., 2003). The previous research had already provided many isotopic mixing models for the conservation research. The IsoSource mixing model was often applied to identify isotopic contributions of various sources of sediments (Phillips and Gregg, 2003). The SIAR (Stable Isotope Analysis in R) is able to analyze the smooth transition autoregressive and quantify diets (Parnell et al., 2010). Given the isotopic information of the consumers' tissues and food sources, the SIMMR (Stable Isotope Mixing Models in R, Parnell et al., 2010) and MixSIAR (Mixing Stable Isotope Analysis in R, Stock and Semmens, 2016), can help estimate the assimilated diet through the food web.

1.1.4.2 Movement information

Many birds migrate thousands of kilometers for dietary resources during breeding periods, or for favorable weather during the non-breeding periods, so better understanding of the migration strategy is essential for the conservation and resources management. The feather information will be matched to the isotopic gradient in the surrounding areas, so the signatures can help describe the movements of birds.

Unlike δ^{15}N which is indicative of trophic level, the δ^{13}C is potentially a good reference for tracing the geolocations. Based on the δ ^{13}C isoscape characters, Quillfeldt et al. (2010) found the pelagic seabird showed a moving pole-wards migration trend during the non-breeding period by using the stable isotope ratios of feathers grown at different periods. For the large spatial scale, many migrant birds are trans-equatorial to the opposing hemisphere, and current research has found that the ratio of δ^{13}C declined sharply with the latitude effect in the southern hemisphere (Phillips et al., 2009). There existed a 7‰ gradient and the δ^{13}C decreased from -23.2‰ in 53°S to -30.3‰ in 62°S (Quillfeldt et al., 2010).

The preference of stopover habitat selection has also been observed in the isotopic profile of δ^{13}C (Yohannes et al., 2014). Previous studies found that stable isotope signatures have good potential forensic tool for the movements, especially in mid to high-latitude continental regions where exist strong spatial isotopic gradients (Bowen et al., 2005), such as Northern American and Southern Ocean areas. However, Roscales et al. (2011) demonstrated that the novel feather stable isotopic methods were not suitable for the northeast Atlantic area due to

the weak gradients in $\delta^{13}C$ baseline values. This weak signature of stable isotopes is also a problem for the Australian continent (Miller and Trigoboff, 2001).

1.1.4.3 Contamination signatures

Incorporating the stable isotopes and heavy metals in the bird feathers provides a powerful method to trace the fate of contaminants. Although the $\delta^{13}C$ is mainly applied to trace the primary source of carbon diet, combining $\delta^{13}C$ together with HMs, can show the pathway of contaminations into food chain or food web.

Einoder et al. (2018) found that mercury in feather will increase with trophic $\delta^{15}N$ level, and arsenic and copper are higher than the level in food chain. Besides, the information of contaminants in feather can help monitor the pollutions at a large spatial scale. For example, Watanuki et al. (2015) used the Hg and $\delta^{15}N$ to trace the pollution a large spatial scale. Hampton et al. (2018) combined Pb and feather isotope information to trace the potential threaten from Pb- based bullets. Einoder et al. (2018) reported the higher lead loads in bird feather exceeding the toxic threshold. Watanuki et al. (2015) traced the mercury contaminants from the primary feather in Japan during the pollution events.

1.1.4.4 Climate signals

Climate change is also a critical driver of bird behaviors. The influence of meteorological factors could be analyzed based on the long historical feather information. For example, based on the comparison of feather stable isotopes collected from different periods, Lavers et al. (2014) found the predator populations of shearwater in Australia have been affected by the Leeuwin Current shifts and ENSO, and gave a prediction of the waterbirds based on the meteorological models prediction. The SOI (Southern Oscillation Index) was also applied to connect the relationships between feather contents and dietary items shifting. By comparing the feather $\delta^{13}C$ isoscape in different periods, Quillfeldt et al. (2012) reported the Thin-billed prions become more dominant in poleward areas because of the diet shifting caused by the sea temperatures increasing.

1.2 Research questions and knowledge gaps

The main research questions that this thesis focused on solving were explained in Figure 1-3, as well as the tracer methods of each chapter and the corresponding environmental compartments. Based on systematic data sets from China and Australia, various geochemical tracers were applied in this study to make a comprehensive exploration of the interactive impacts of human activity and natural processes on the transport and accumulation of

geochemical tracers. The specific research questions and current knowledge gaps that this thesis tried to fill are as follows:

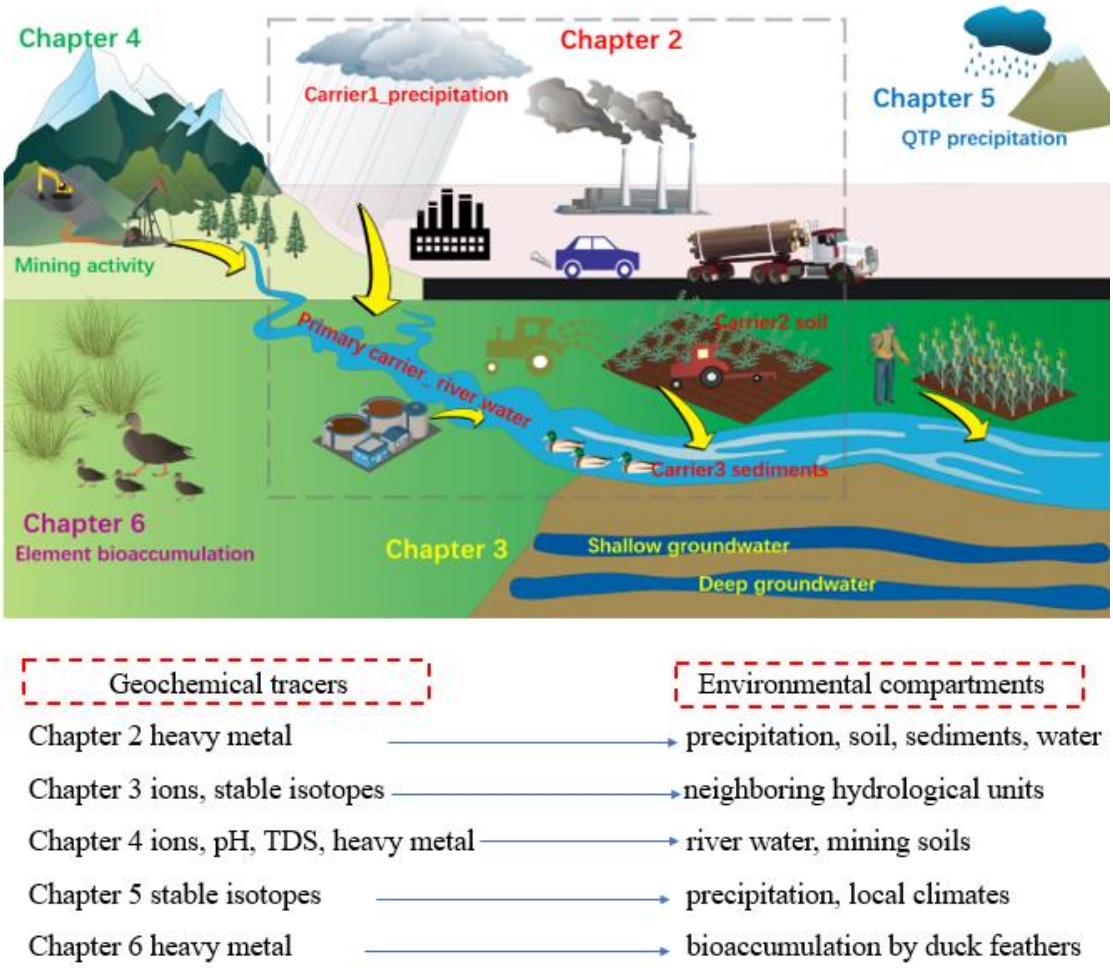


Figure 1-3. Framework showing proposed outline of each main chapters in this thesis, with chemical elements transport among various environmental compartments at different scales.

1.2.1 Transport of heavy metals between different carriers at catchment scale

HMs have been consistently entering the food chain, imposing great harm on environment and public health. However, previous studies on the transport mechanism and dynamics of HMs have been profoundly limited by the spatial inconsistence and scarcity of the field observations, especially over the complex environment such as large-scale catchments. There are three important 'gaps' in the current knowledge remain to be solved: 1) only single HM carrier was investigated in most of the previous studies, profoundly hindering our understanding of the interactions among different carriers, which play an important role in transporting the HMs; 2) insufficient explanatory variables of HMs were often selected to correspond with a specific carrier studied in previous research. Therefore, comprehensive

explanatory variables from separated environment and human factors are needed for a better understanding of HM contamination and transport mechanism; 3) current mapping methods of HMs are dependent on the spatial interpolation like Kriging, which are greatly affected by the intensity and spatial distribution of field data sampling. Therefore, comprehensive explanatory variables and efficient spatial modelling methods have been increasingly called for better identifying HMs and their carriers. These research questions were conducted in Chapter 2.

1.2.2 Water ions and stable isotopes to study the hydraulic connections between different hydrological units

The water quantity interaction and water contaminant transportation are two criteria for evaluating the hydraulic connections between different hydrological units. The connections could be even more complicated in the arid regions due to the extremely frequent interactions between surface water and groundwater, as well as the spatial inconsistence of river water across large scales. The active shallow groundwater is not only affected by the interactions from surface water and deep groundwater, but also a series of natural and human factors. An accurate understanding of the relationships between shallow groundwater and these explanatory variables has always been a knowledge gap that restricts hydraulic connection research in arid regions.

There are three main knowledge gaps currently. (i) The sub-basin scale exchange was mostly focused on by previous models (Freyberg et al., 2017; Khalil et al., 2015; Klaus and McDonnell, 2013; Liu et al., 2015), which failed to analyze the action-response relationship between different hydrological units. Although a few studies had carried out research for inter-basin exchanges, the research scope is limited to around the boundary (Wang et al., 2013; Zhao et al., 2012), where always cannot represent the characteristics of the entire unit. (ii) Only the water quantity interactions were focused on by the conceptual models of the EMMA (Guo et al., 2019; Wang et al., 2013). To the best of our knowledge, few research has applied this conceptual model for the water pollutant transportation between different hydrological units. (iii) Some pollution tracers, like nitrates, sulfates (Hu et al., 2019; Khalil et al., 2015), were often selected to correspond with a specific pollution source, such as the industry or agriculture sites (Guo et al., 2019). Explanatory variables with spatial attributes were always insufficient. Therefore, systematic sampling along the potential hydraulic connection routes, and introduction of novel spatial distribution model with considering the geological and hydrological factors are two essential requirements to fill these current knowledge gaps. These research questions were conducted in Chapter 3.

1.2.3 Water chemical parameters and soil heavy metals to trace mining pollutants

As the important anthropogenic sources of contaminant elements, the influence of mining activity had attracted many researches. Traditionally, heavy metals are used to trace pollutants in the soil, atmosphere, and aquatic systems, but they are often traced directly without considering conversion forces, and previous research has primarily focused on each of these transmission media separately. Spatially, although the effect of mining on water quality has been well studied on river areas proximal to mining activities in low altitude environments (Kondolf, 1997), the process of large-scale mining impacts within alpine catchments is more complicated and less studied. For example, mountainous geomorphology can significantly alter the export of weathering solutes and transport pathways (Rascher et al., 2018; Wellen et al., 2018). Furthermore, alpine topographic features make it more difficult to understand transport processes of contaminants in upstream environments. Thus, it is necessary to develop a novel method of tracing major water chemical parameters and soil heavy metals to fill the above knowledge gaps about the mining pollutants. These research questions were conducted in Chapter 4.

1.2.4 Variations of stable isotopes in precipitation influenced by local climatic factors

The study area of this topic was selected at the Qinghai-Tibet Plateau (QTP), where the hydrological cycles have experienced significant changes recent years. A thorough understanding of the relationships between isotopic compositions ($\delta^{18}O$ and δD) in precipitation and potential local climatic factors is a prerequisite requirement for discovering these hydrological processes in the Third Polar area. However, long-term continuous precipitation isotopic data in QTP, as well as the synchronous local climatic records were scarce due to the harsh environment, poor traffic and observation conditions. Thus, continuous sampling and synchronous local climatic variables over a long-term period at the transition zones of the QTP are two essential requirements to fill these current knowledge gaps. These research questions were conducted in Chapter 5.

1.2.5 Bioaccumulation mechanisms of heavy metals in waterfowl feathers

Understanding the mechanisms of element transportation and accumulation can help to monitor human health and environment conservation, but the accumulation of chemical elements in organisms is an extremely complicated process, which will be affected by both the biotic and abiotic factors. Their spread and accumulation will be accelerated, enhanced, and strengthened directly or indirectly by anthropogenic activity, and their relationship with the environmental background is more complicated. Bird feathers are a reliable tool to

discover these elemental bioaccumulation mysteries. There are four important 'gaps' in the current knowledge remaining to be solved. (i) The elemental geochemical background is ignored, and most of the current bioaccumulation research on feathers focused on independent elemental concentration, and tries to give the threshold of different elements in bird feather (López-Perea et al., 2019). In addition, the previous research about the bioaccumulation ratio focused on blood, hair, feces, urine and other organs, and only a few studies focused on the feather bioaccumulation ratio. (ii) The essential and toxic elements were not discussed separately, and compared with toxic elements, heavy metals, etc., the information that essential elements can indicate has rarely been analyzed. (iii) In terms of the explanatory variables, previous research focused on human activities and ignored the environment variables. Although some studies involved the environment, they were neither very systematic or comprehensive. (iv) Previous research stayed at the qualitative stage, with few quantitative analyses. These research questions were conducted in Chapter 6.

1.3 Significance of this thesis

The transport and accumulation of chemical elements are complicated processes, and the toxic elements can be consistently entering the food chain, imposing great harm on environment and public health. However, previous studies have been profoundly limited by the spatial inconsistence and scarcity of the field observations, as well as the mapping methods. Thus, a comprehensive framework of novel methods and datasets have been conducted in this thesis to fill these knowledge gaps for better mapping and assessing geochemical transport in different environment compartments, such as the catchment scale, different hydrological units, as well as the mining activity and bioaccumulation processes. The novel methodologies and experiment designs presented in this study can also be applied in other similar researches.

Additionally, this thesis was based on systematic field sampling over large spatial scales, as well as history record materials, remote sensing data, and geographic information system productions. The outcomes of this thesis will not only provide useful database for other research, especially at the data scarce areas, but also provide scientific strategies for the regional development and management.

1.4 Reference

Abbasi, N.A., Jaspers, V.L.B., Chaudhry, M.J.I., Ali, S., Malik, R.N., 2015. Influence of taxa, trophic level, and location on bioaccumulation of toxic metals in bird's feathers: a preliminary biomonitoring study using multiple bird species from Pakistan. Chemosphere, 120: 527-537.

Abbott, M., Lini, A., Bierman, P., 2000. δ18O, δD and 3H measurements constrain groundwater recharge patterns in an upland fractured bedrock aquifer, Vermont, USA. Journal of Hydrology, 228(1-2): 101-112.

Abdullah, M. et al., 2015. Avian feathers as a non-destructive bio-monitoring tool of trace metals signatures: a case study from severely contaminated areas. Chemosphere, 119: 553-561.

Badry, A. et al., 2019. Using an apex predator for large-scale monitoring of trace element contamination: Associations with environmental, anthropogenic and dietary proxies. Science of the Total Environment, 676: 746-755.

Bowen, G.J., Wassenaar, L.I., Hobson, K.A., 2005. Global application of stable hydrogen and oxygen isotopes to wildlife forensics. Oecologia, 143(3): 337-348.

Brandis, K., Bino, G., Spencer, J., Ramp, D., Kingsford, R., 2018. Decline in colonial waterbird breeding highlights loss of Ramsar wetland function. Biological Conservation, 225: 22-30.

Brooks, S., Arimoto, R., Lindberg, S., Southworth, G., 2008. Antarctic polar plateau snow surface conversion of deposited oxidized mercury to gaseous elemental mercury with fractional long-term burial. Atmospheric Environment, 42(12): 2877-2884.

Bush, R.T., Berke, M.A., Jacobson, A.D., 2017. Plant water δD and δ18O of tundra species from West Greenland. Arctic, Antarctic, and Alpine Research, 49(3): 341-358.

Candeias, C., Ávila, P.F., Da Silva, E.F., Teixeira, J.P., 2015. Integrated approach to assess the environmental impact of mining activities: estimation of the spatial distribution of soil contamination (Panasqueira mining area, Central Portugal). Environmental monitoring and assessment, 187(3): 1-23.

Cook, E.R. et al., 2013. Five centuries of Upper Indus River flow from tree rings. Journal of hydrology, 486: 365-375.

Costa, M., Klein, C.B., 2006. Toxicity and carcinogenicity of chromium compounds in humans. Critical reviews in toxicology, 36(2): 155-163.

Dudka, S., Miller, W., 1999. Accumulation of potentially toxic elements in plants and their transfer to human food chain. Journal of Environmental Science & Health Part B, 34(4): 681-708.

Einoder, L., MacLeod, C., Coughanowr, C., 2018. Metal and isotope analysis of bird feathers in a contaminated estuary reveals bioaccumulation, biomagnification, and potential toxic effects. Archives of environmental contamination and toxicology, 75(1): 96-110.

Fang, H., Huang, L., Wang, J., He, G., Reible, D., 2016. Environmental assessment of heavy metal transport and transformation in the Hangzhou Bay, China. Journal of hazardous materials, 302: 447-457.

Farmer, R., Leonard, M., 2011. Long-term feeding ecology of Great Black-backed Gulls (Larus marinus) in the northwest Atlantic: 110 years of feather isotope data. Canadian Journal of Zoology, 89(2): 123-133.

Fetter, C.W., Boving, T., Kreamer, D., 2017. Contaminant hydrogeology. Waveland Press.

Frantz, A. et al., 2012. Contrasting levels of heavy metals in the feathers of urban pigeons from close habitats suggest limited movements at a restricted scale. Environmental Pollution, 168: 23-28.

Frederick, P.C., Hylton, B., Heath, J.A., Spalding, M.G., 2004. A historical record of mercury contamination in southern florida (USA) as inferred from avian feather tissue: Contribution R-09888 of the Journal Series, Florida Agricultural Experiment Station. Environmental Toxicology and Chemistry: An International Journal, 23(6): 1474-1478.

Furtado, R., Pereira, M.E., Granadeiro, J.P., Catry, P., 2019. Body feather mercury and arsenic concentrations in five species of seabirds from the Falkland Islands. Marine pollution bulletin, 149: 110574.

Goth, A., Michelsen, A., Rousk, K., 2019. Railroad derived nitrogen and heavy metal pollution does not affect nitrogen fixation associated with mosses and lichens at a tundra site in Northern Sweden. Environmental Pollution, 247: 857-865.

Hampton, J.O., Laidlaw, M., Buenz, E., Arnemo, J.M., 2018. Heads in the sand: public health and ecological risks of lead-based bullets for wildlife shooting in Australia. Wildlife Research, 45(4): 287-306.

Hassellov, M., von der Kammer, F., 2008. Iron oxides as geochemical nanovectors for metal transport in soil-river systems. Elements, 4(6): 401-406.

Hatzinger, P.B., Böhlke, J.K., Sturchio, N.C., 2013. Application of stable isotope ratio analysis for biodegradation monitoring in groundwater. Current opinion in biotechnology, 24(3): 542-549.

Hochella Jr, M.F. et al., 2005. Direct observation of heavy metal-mineral association from the Clark Fork River Superfund Complex: Implications for metal transport and bioavailability. Geochimica et cosmochimica acta, 69(7): 1651-1663.

Hodson, A., Tranter, M., Gurnell, A., Clark, M., Hagen, J.O., 2002. The hydrochemistry of Bayelva, a high Arctic proglacial stream in Svalbard. Journal of Hydrology, 257(1-4): 91-114.

Hoshina, Y., Fujita, K., Iizuka, Y., Motoyama, H., 2016. Inconsistent relationships between major ions and water stable isotopes in Antarctic snow under different accumulation environments. Polar Science, 10(1): 1-10.

Hoshina, Y. et al., 2014. Effect of accumulation rate on water stable isotopes of near-surface snow in inland Antarctica. Journal of Geophysical Research: Atmospheres, 119(1): 274-283.

Johnson, D.B., Hallberg, K.B., 2005. Acid mine drainage remediation options: a review. Science of the Total Environment, 338(1-2): 3-14.

Jury, W.A., Focht, D.D., Farmer, W.J., 1987. Evaluation of pesticide groundwater pollution potential from standard indices of soil-chemical adsorption and biodegradation. Journal of Environmental Quality, 422-428.

Khalil, M.M., Tokunaga, T., Yousef, A.F.J.J.o.H., 2015. Insights from stable isotopes and hydrochemistry to the Quaternary groundwater system, south of the Ismailia canal, Egypt. 527: 555-564.

Kondolf, G.M., 1997. PROFILE: hungry water: effects of dams and gravel mining on river channels. Environmental management, 21(4): 533-551.

Kreutz, K.J., Wake, C.P., Aizen, V.B., Cecil, L.D., Synal, H.A., 2003. Seasonal deuterium excess in a Tien Shan ice core: influence of moisture transport and recycling in Central Asia. Geophysical Research Letters, 30(18).

Kumar, A., Tiwari, S.K., Verma, A., Gupta, A.K., 2018. Tracing isotopic signatures (δD and δ18O) in precipitation and glacier melt over Chorabari Glacier–Hydroclimatic inferences for the Upper Ganga Basin (UGB), Garhwal Himalaya. Journal of Hydrology: Regional Studies, 15: 68-89.

Lacelle, D., Fontaine, M., Forest, A.P., Kokelj, S., 2014. High-resolution stable water isotopes as tracers of thaw unconformities in permafrost: A case study from western Arctic Canada. Chemical Geology, 368: 85-96.

Lavers, J.L., Bond, A.L., Hutton, I., 2014. Plastic ingestion by Flesh-footed Shearwaters (Puffinus carneipes): Implications for fledgling body condition and the accumulation of plastic-derived chemicals. Environmental Pollution, 187: 124-129.

Li, J. et al., 2020. Tracing geochemical pollutants in stream water and soil from mining activity in an alpine catchment. Chemosphere, 242: 125167.

Li, Z. et al., 2016. Quantitative evaluation on the influence from cryosphere meltwater on runoff in an inland river basin of China. Global and Planetary Change, 143: 189-195.

Li, Z. et al., 2014. Study on the contribution of cryosphere to runoff in the cold alpine basin: A case study of Hulugou River Basin in the Qilian Mountains. Global and Planetary Change, 122: 345-361.

Liu, F. et al., 2015. Identifying the origin and geochemical evolution of groundwater using hydrochemistry and stable isotopes in the Subei Lake basin, Ordos energy base, Northwestern China. Hydrology and Earth System Sciences, 19(1): 551-565.

Liu, F. et al., 2019. Coupling hydrochemistry and stable isotopes to identify the major factors affecting groundwater geochemical evolution in the Heilongdong Spring Basin, North China. Journal of Geochemical Exploration, 205: 106352.

López-Perea, J.J. et al., 2019. Metals and metalloids in blood and feathers of common moorhens (Gallinula chloropus) from wetlands that receive treated wastewater. Science of the Total Environment, 646: 84-92.

Lyon, S.W., Desilets, S.L., Troch, P.A., 2009. A tale of two isotopes: differences in hydrograph separation for a runoff event when using δD versus δ18O. Hydrological Processes: An International Journal, 23(14): 2095-2101.

Mahurpawar, M., 2015. Effects of heavy metals on human health. International Journal of Research-Granthaalayah, 3(9SE): 1-7.

Mandal, B.K., Suzuki, K.T., 2002. Arsenic round the world: a review. Talanta, 58(1): 201-235.

Marschall, H.R., Schumacher, J.C., 2012. Arc magmas sourced from mélange diapirs in subduction zones. Nature Geoscience, 5(12): 862-867.

Maurya, P., Kumari, R., Mukherjee, S., 2019. Hydrochemistry in integration with stable isotopes (δ18O and δD) to assess seawater intrusion in coastal aquifers of Kachchh district, Gujarat, India. Journal of Geochemical Exploration, 196: 42-56.

Meite, F. et al., 2018. Impact of rainfall patterns and frequency on the export of pesticides and heavy-metals from agricultural soils. Science of the Total Environment, 616: 500-509.

Metcheva, R., Yurukova, L., Teodorova, S.E., 2011. Biogenic and toxic elements in feathers, eggs, and excreta of Gentoo penguin (Pygoscelis papua ellsworthii) in the Antarctic. Environmental monitoring and assessment, 182(1): 571-585.

Miller, N.G., Trigoboff, N., 2001. A European feather moss, Pseudoscleropodium purum, naturalized widely in New York State in cemeteries. The Bryologist, 104(1): 98-103.

Mizutani, H., Fukuda, M., Kabaya, Y., 1992. ^(13) C and^(15) N Enrichment Factors of Feathers of 11 Species of Adult Birds. Ecology, 73(4): 1391-1395.

Morel, F.M., Kraepiel, A.M., Amyot, M., 1998. The chemical cycle and bioaccumulation of mercury. Annual review of ecology and systematics, 29(1): 543-566.

Nevin, R., 2000. How lead exposure relates to temporal changes in IQ, violent crime, and unwed pregnancy. Environmental research, 83(1): 1-22.

Newth, J. et al., 2013. Poisoning from lead gunshot: still a threat to wild waterbirds in Britain. European Journal of Wildlife Research, 59(2): 195-204.

Nickel, S. et al., 2014. Modelling and mapping spatio-temporal trends of heavy metal accumulation in moss and natural surface soil monitored 1990–2010 throughout Norway by multivariate generalized linear models and geostatistics. Atmospheric Environment, 99: 85-93.

Ofukany, A.F., Wassenaar, L.I., Bond, A.L., Hobson, K.A., 2014. Defining fish community structure in Lake Winnipeg using stable isotopes (δ13C, δ15N, δ34S): Implications for monitoring ecological responses and trophodynamics of mercury & other trace elements. Science of the Total Environment, 497: 239-249.

Ogrinc, N., Kanduč, T., Stichler, W., Vreča, P., 2008. Spatial and seasonal variations in δ18O and δD values in the River Sava in Slovenia. Journal of Hydrology, 359(3-4): 303-312.

Parnell, A.C., Inger, R., Bearhop, S., Jackson, A.L., 2010. Source partitioning using stable isotopes: coping with too much variation. PloS one, 5(3): e9672.

Pearson, S.F., Levey, D.J., Greenberg, C.H., Del Rio, C.M., 2003. Effects of elemental composition on the incorporation of dietary nitrogen and carbon isotopic signatures in an omnivorous songbird. Oecologia, 135(4): 516-523.

Peterson, B.J., Fry, B., 1987. Stable isotopes in ecosystem studies. Annual review of ecology and systematics, 18(1): 293-320.

Phillips, D.L., Gregg, J.W., 2003. Source partitioning using stable isotopes: coping with too many sources. Oecologia, 136(2): 261-269.

Phillips, R.A., Bearhop, S., Mcgill, R.A., Dawson, D.A., 2009. Stable isotopes reveal individual variation in migration strategies and habitat preferences in a suite of seabirds during the nonbreeding period. Oecologia, 160(4): 795-806.

Pirrone, N. et al., 2010. Global mercury emissions to the atmosphere from anthropogenic and natural sources. Atmospheric Chemistry and Physics, 10(13): 5951-5964.

Prashanth, L., Kattapagari, K.K., Chitturi, R.T., Baddam, V.R.R., Prasad, L.K., 2015. A review on role of essential trace elements in health and disease. Journal of dr. ntr university of health sciences, 4(2): 75.

Price, R.M., Swart, P.K., Willoughby, H.E., 2008. Seasonal and spatial variation in the stable isotopic composition (δ18O and δD) of precipitation in south Florida. Journal of Hydrology, 358(3-4): 193-205.

Quillfeldt, P., Masello, J.F., McGill, R.A., Adams, M., Furness, R.W., 2010. Moving polewards in winter: a recent change in the migratory strategy of a pelagic seabird? Frontiers in Zoology, 7(1): 1-11.

Quillfeldt, P. et al., 2012. Impact of miniature geolocation loggers on a small petrel, the thin-billed prion Pachyptila belcheri. Marine Biology, 159(8): 1809-1816.

Rascher, E., Rindler, R., Habersack, H., Sass, O., 2018. Impacts of gravel mining and renaturation measures on the sediment flux and budget in an alpine catchment (Johnsbach Valley, Austria). Geomorphology, 318: 404-420.

Risi, C., Bony, S., Vimeux, F., 2008. Influence of convective processes on the isotopic composition (δ18O and δD) of precipitation and water vapor in the tropics: 2. Physical interpretation of the amount effect. Journal of Geophysical Research: Atmospheres, 113(D19).

Roscales, J.L., Gómez-Díaz, E., Neves, V., González-Solís, J., 2011. Trophic versus geographic structure in stable isotope signatures of pelagic seabirds breeding in the northeast Atlantic. Marine Ecology Progress Series, 434: 1-13.

Sarkar, D.J. et al., 2021. Occurrence, fate and removal of microplastics as heavy metal vector in natural wastewater treatment wetland system. Water Research, 192: 116853.

Senduran, C. et al., 2018. Mitigation and treatment of pollutants from railway and highway runoff by pocket wetland system; A case study. Chemosphere, 204: 335-343.

Stock, B.C., Semmens, B.X., 2016. Unifying error structures in commonly used biotracer mixing models. Ecology, 97(10): 2562-2569.

Stojic, N., Pucarevic, M., Stojic, G., 2017. Railway transportation as a source of soil pollution. Transportation Research Part D: Transport and Environment, 57: 124-129.

Tas, S., Lauwerys, R., Lison, D., 1996. Occupational hazards for the male reproductive system. Critical reviews in toxicology, 26(3): 261-307.

Thompson, D.R., Furness, R.W., 1995. Stable-isotope ratios of carbon and nitrogen in feathers indicate seasonal dietary shifts in northern fulmars. The Auk, 112(2): 493-498.

Wang, S. et al., 2018. Combination of CFCs and stable isotopes to characterize the mechanism of groundwater–surface water interactions in a headwater basin of the North China Plain. Hydrological Processes, 32(11): 1571-1587.

Wassenaar, L.I., Van Wilgenburg, S.L., Larson, K., Hobson, K.A., 2009. A groundwater isoscape (δD, δ18O) for Mexico. Journal of Geochemical Exploration, 102(3): 123-136.

Watanuki, Y. Yamamoto T, Yamashita A, Ishii C, Ikenaka Y, Nakayama SM, Ishizuka M, Suzuki Y, Niizuma Y, Meathrel CE, Phillips RA, 2015. Mercury concentrations in

primary feathers reflect pollutant exposure in discrete non-breeding grounds used by Short-tailed Shearwaters. Journal of Ornithology, 156(3): 847-850.

Wellen, C., Shatilla, N.J., Carey, S.K., 2018. The influence of mining on hydrology and solute transport in the Elk Valley, British Columbia, Canada. Environmental Research Letters, 13(7): 074012.

Wolfe, M.F., Schwarzbach, S., Sulaiman, R.A., 1998. Effects of mercury on wildlife: a comprehensive review. Environmental Toxicology and Chemistry: An International Journal, 17(2): 146-160.

Wretenberg, J., Pärt, T., Berg, Å., 2010. Changes in local species richness of farmland birds in relation to land-use changes and landscape structure. Biological Conservation, 143(2): 375-381.

Xu, X., Nie, S., Ding, H., Hou, F.F., 2018. Environmental pollution and kidney diseases. Nature Reviews Nephrology, 14(5): 313-324.

Yin, L. et al., 2011. Isotopes (δD and δ 18 O) in precipitation, groundwater and surface water in the Ordos Plateau, China: implications with respect to groundwater recharge and circulation. Hydrogeology Journal, 19(2): 429-443.

Yohannes, E., Arnaud, A., Béchet, A., 2014. Tracking variations in wetland use by breeding flamingos using stable isotope signatures of feather and blood. Estuarine, Coastal and Shelf Science, 136: 11-18.

Zahir, F., Rizwi, S.J., Haq, S.K., Khan, R.H., 2005. Low dose mercury toxicity and human health. Environmental Toxicology and Pharmacology, 20(2): 351-360.

Zhu, Y.-G. et al., 2008. High percentage inorganic arsenic content of mining impacted and nonimpacted Chinese rice. Environmental Science & Technology, 42(13): 5008-5013.

Chapter 2. Heavy metal habitat: a novel framework for mapping heavy metal contamination over large-scale catchment with a species distribution model

Abstract

Heavy metals (HMs) have been consistently entering the food chain, imposing great harm on environment and public health. However, previous studies on the transport mechanism and dynamics of HMs have been profoundly limited by the spatial inconsistence and scarcity of the field observations, especially over the complex environment such as large-scale catchments. As a result, comprehensive explanatory variables and efficient spatial modelling methods have been increasingly called for better identifying HMs and their carriers. In this study, a novel methodological framework for mapping HM contamination at catchment scale was proposed and applied, combing a species distribution model (SDM) with physical environmental and human factors. With the field samples of HMs from river water, wet deposition, soil and modern fluvial sediments across the Heihe river basin, our proposed framework was able to model the 'habitat' of each HM in the geographic space (termed as the HM 'habitat') and identify the HM 'hotspots' in the catchment. Results showed HM hotspots in the soil were identified in the middle reach of the Heihe river basin where intense human activities occur. The variables of the human footprint including distance to road, GDP, and nightlight were found to play the most important roles in soil HM transport among others. The output maps of HM habitats from soil, sediment, and wet deposition well agreed with the influence of industry contaminants, hydraulic sorting, and precipitation washout process respectively, indicating the potential of SDM in modelling the spatial distributions of the HM carriers. In addition, we found the distributions of HMs in the primary carrier of river water were attributed to both of the human and environmental variables. These variables, along with the maps of HM distributions derived from secondary carriers were thus selected as explanatory variables in SDM to predict the spatial patterns of the final HMs in river water, which was observed to have largely improved the prediction accuracy. These results confirm the effectiveness and potential of our framework to leverage SDMs from ecology to study heavy metal contamination at catchment scale, which offers new insights to understand the relationships between HMs and the human and physical environment.

Keywords: Heavy metal; species distribution models; Catchment; Carriers; Explanatory variables; Redundancy analysis; Model selection

Highlights:

* The output maps of HMs from soil, sediment, and wet deposition could respectively reflect the influence of industry contaminants, hydraulic sorting, and precipitation washout process

* HM transport from secondary carriers helped improve the prediction accuracy of HMs in river water through SDM

* SDM can be reasonably extended to the field of HM transport in water chemistry

2.1 Introduction

Heavy metals (HMs), ubiquitous and generally persistent in the environment, tend to accumulate in water cycle and bio-magnify in the food chain (Ali et al., 2019), imposing great potential threats to human health and environmental sustainability. HMs have been found to actively transport within multiple environmental compartments, such as the catchment (Hassellov and von der Kammer, 2008), urban system (Sarkar et al., 2021), agriculture (Meite et al., 2018), estuary (Fang et al., 2016), mining sites and public transport system etc., (Goth et al., 2019; Senduran et al., 2018; Stojic et al., 2017), spreading from local, regional, to global scales. Thus, understanding of HMs transport mechanism and dynamics is essential for conservation and resource management (Hassellov and von der Kammer, 2008).

Studies on heavy metal contamination within catchment have been increasingly applied to ensure a sustainable resource supply of catchments for people, stock, flora and fauna (Hochella Jr et al., 2005). However, such research has been challenged by the complex environment of catchment system, which consists of multi-dimensional carriers, such as the atmospheric deposition, soil, sediment, river water and biosphere etc., (Figure 2-1). For example, the atmospheric deposition in a catchment consistently brings HM contaminants from the aerosols into the river water, soil and eventually groundwater (Nickel et al., 2014). Soil on the other hand plays as a pollutant sink for the suspended particles, dust, and sand, and a place where the irrigation and seepage water pollutants accumulate. The HM suspended particles from tailing soil in close proximity of industrial and mining areas can then travel to a long distance during the precipitation erosion and runoff generation processes (Li et al., 2020). In addition, modern fluvial sediments in hydrodynamic sorting process also influence the river water HM contents via the water-rock interaction. Eventually, the HMs converged into river water, which would be deposited on floodplain sediments (Hassellov and von der Kammer, 2008) or transported farther downstream by the floods and channel aggregation (Hochella Jr et al., 2005). Therefore, the river water plays an active role in linking various carriers within a catchment, which we defined as the primary carrier in this study, while other carriers as secondary carriers. HMs transport within biosphere, which involves aquatic organisms, insects, potentially mammals, was not included here due to its intrinsic complex biotic factors (Hochella Jr et al., 2005) and lack of data.

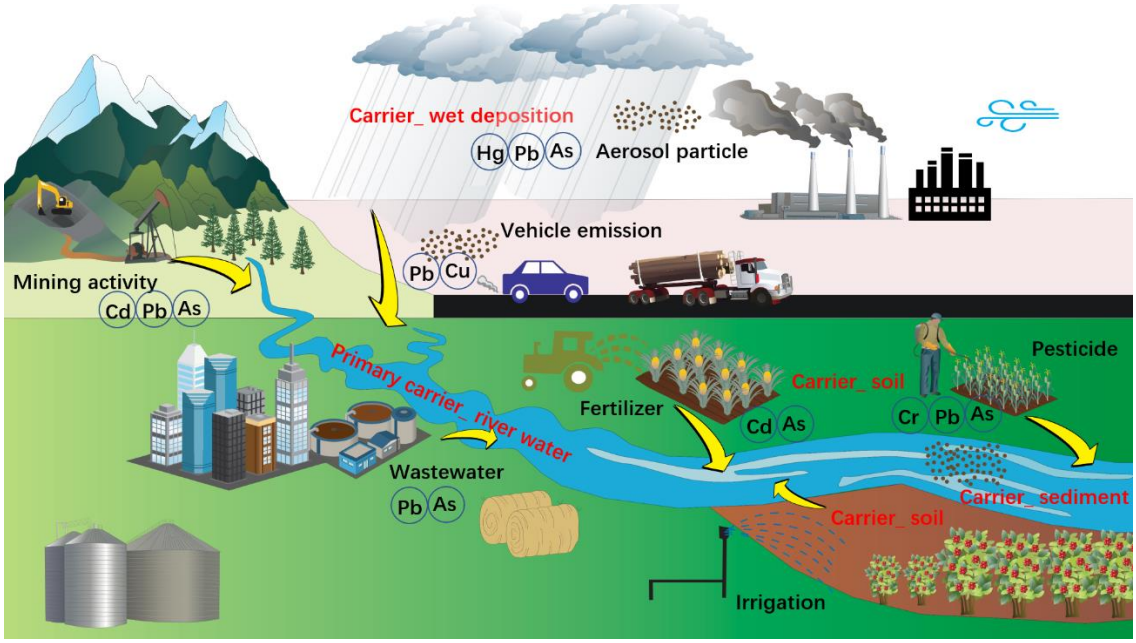


Figure 2- 1. The HMs transport among primary carrier of river water and secondary carriers of wet deposition, soil and sediment at catchment scale.

Heavy metal contamination and transport mechanism in catchment system have been investigated in the previous studies for addressing three main questions in the field: (i) What are the potential sources of the HMs and how the various carriers interact to spread them; (ii) What are the potential explanatory variables that drive HMs transport, at the perspectives of both the extrinsic factors and intrinsic mechanism; and (iii) What and where are the environmental impacts of the HMs. Key papers including their strength and limitations for these essential research questions are summarized in Table 2-1. For example, source apportionment methods using receptor models such as 'enrichment factors' and 'principal component analysis' models have been widely used to identify sources of HMs (Cable and Deng, 2018; Lv, 2019). These models however cannot quantify the relative contributions of different pollutant sources (Lv, 2019). Multivariate receptor models including 'principal component score/multiple linear regressions' and 'positive matrix factorization' thus were introduced to address this issue (Taiwo et al., 2014). However, these models require heavy prior knowledge of HM transport process, ignore spatial correlations among samples, and more importantly, often fail in identifying the connections between HMs and different explanatory variables for successful prediction of future HM contamination (Hu et al., 2020; Lv, 2019). Recent development of nanotechnology may offer potential solution for accurately identifying the relationships between HMs and explanatory variables, which yet is too expensive and not suitable for large-scale study areas (Hassellov and von der Kammer, 2008; Hochella Jr et al., 2005). Meanwhile, modern statistical methods including simple linear models and advanced machine learning have

been increasingly proposed to identify the relationships between HMs and explanatory variables (Hu et al., 2020).

As a result, although these previous studies have made significant contributions (details in the literature review of Table 2-1), there are three important 'gaps' in the current knowledge remain to be solved, which are largely due to the spatial inconsistence and scarcity of the field observations: 1) only single HM carrier was investigated in most of the previous studies, profoundly hindering our understanding of the interactions among different carriers, which play an important role in transporting the HMs; 2) insufficient explanatory variables of HMs were often selected to correspond with a specific carrier studied in previous research. Therefore, comprehensive explanatory variables from separated environment and human factors are needed for a better understanding of HM contamination and transport mechanism; 3) current mapping methods of HMs are dependent on the spatial interpolation like Kriging, which are greatly affected by the intensity and spatial distribution of field data sampling.

Therefore, a comprehensive framework of novel methods and datasets is needed to fill these knowledge gaps for better mapping and assessing heavy metal contamination in catchment especially at large-scale. Species distribution modelling (SDM) has been widely applied in research areas such as conservation biology, ecology and evolution. SDMs of various algorithms can map the past, current and future spatial distributions of species with species presence-only sample data and environmental data (e.g., climate, land use, soil etc), reflecting the suitability of their potential habitats (Guisan and Thuiller, 2005). SDMs can produce spatially continuous map of habitat suitability in the presence of methodological challenges: non-randomness of the occurrence data, inhomogeneity of the collection efforts, landscape heterogeneity in different scales etc (Ducci et al., 2015; Kearney and Porter, 2009). With the big data revolution, the applications of SDMs have been increasingly extended to other fields to create important new knowledge, such as geographical planning and design (Braunisch et al., 2011; Clemente et al., 2019; Walden-Schreiner et al., 2018), spatial distribution of human emotion (Li et al., 2021) etc. We thus believe SDMs can bring new perspectives for mapping the HM distributions, as important research similarities between species and HM do exist. (i) Multiple entities can coexist. For example, multiple species can share the same habitat while the multiple HMs discussed in this study own the same feature (e.g.,Pb, Hg, As, Cd, etc. often co-exist in the location of the catchment). (ii) Like species, HMs are being influenced by multi-dimensional explanatory variables, which can be classified into the natural and human factors. (iii) With corresponding explanatory variables, the patterns of both species and HMs can be predicted and presented as spatial maps (Elith and Franklin, 2013). (iv) Only the occurrence dataset of research entities is required. Thus, SDMs have great potential for mapping habitats of HMs and may bring new directions in this field.

In this study, we developed a novel framework for mapping HM contamination at large catchment scale combing a species distribution model with environmental and human variables. With different HMs (e.g., Pb, Hg, As, Cd, etc.) being considered as different 'species', we were able to model the 'habitat' of each HM in the geographic space (termed as the HM 'habitat') along with HM 'hotspot'. In particular, over the study area of Heihe catchment we aimed to: (i) investigate the spatiotemporal 'habitats' of HMs within in various carriers and their interactions; (ii) separate impacts from natural environment and human activities on secondary carriers (e.g., atmospheric deposition, soil, and sediment etc.); (iii) assess spatial HM transmission patterns of river water (the primary carrier) across the entire catchment, with modelled spatially consistent data. This study proposed to leverage SDMs models from ecology to map HM contamination, which can effectively build the environment-human-HM relationships to fill the current important knowledge gaps over large-scale catchment. This will offer new future perspectives and directions to the field of hydrological chemistry.

Table 2- 1. Summary of relevant studies conducted previously for investigating HM contamination and transport mechanism at catchment scale (the current paper is added for completeness). Three components in the 'Key results' column are identified by the code:(1) interactions among multiple carriers;(2) explanatory variables used;(3) spatial maps of HM distributions. NA represents 'not applicable' in the relevant research.

Study	Field sampling design/ requirements	Analytical methods	Key results
Lv (2019)	Prior designs with sampling density of 2km × 2km at a 1138 km^2 County	Receptor models, APCS/MLR and PMF, Kriging interpolation	(1) Only one carrier of soil was analysed (2) NA (3) Kriging maps with the factors resulting from receptor models
Nickel et al. (2014)	Sampling network throughout Norway is required	GLM, geo-statistics multivariate regression, Kriging interpolation	(1) NA (2) A set of potential predictors were classified based on literature review (3) Regression and residual maps with grid size of 5 km × 5 km
Hu et al. (2020)	Proximity	Three machine learning methods GBM, RF, GLM	(1) Soil and crop ecosystems (2) Soil properties and land use types. (3) NA
Hassellov and von der Kammer (2008)	Proximity	TEM	(1) Topsoil and river water (2) NA (3) NA

Sharma et al. (2008)	A prior reconnaissance survey was conducted before sampling	Risk index analysis, traditional statistical analyses	(1) Atmospheric deposition and vegetables (2) Only land use types were extracted (3) NA
Liang et al. (2017b)	Proximity	Curvilinear regression analysis	(1) Atmospheric deposition and soil (2) Only land use types were analysed (3) NA
Hochella Jr et al. (2005)	Proximity	TEM	(1) Floodplains and riverbed sediments (2) Only mining activity (3) NA
Quinton and Catt (2007)	Proximity	Traditional statistical analyses	(1) Soil and water (2) NA, only land use types (3) NA
Zang et al. (2021)	Sampling station and synchronous equipment is necessary	Receptor models, PCA, PMF	(1) Only one carrier of precipitation (2) Natural environmental factors and human influence (3) NA
Liang et al. (2017a)	High sampling density of seven sites/ km^2	PMF, Kriging interpolation	(1) Only one carrier of soil (2) Only land use types (3) Kriging interpolation maps
Li et al. (2017)	Grid distribution point method	Correlation analysis, PCA, Kriging interpolation	(1) NA (2) Only the land use types close to piles of mine tailings were analysed (3) Kriging interpolation maps
This study	Covering entire catchment; each dimension of carrier has their own sampling density	SDMs; RDA; GLM; GAM; model selection and model averaging	(1) Multiple carriers of river water, atmospheric deposition, soil, and sediment were analysed (2) Systematic explanatory variables to represent environmental and human influential factors; auxiliary variables of secondary carriers (3) Spatial maps of HM 'habitat', 'hotspot' and river transmission across the entire catchment modelled by a SDM

Proximity is a sampling strategy that chose a short distance apart between different dimensions of carriers at the same location. Abbreviations of the analytical methods used in Table 2-1 include absolute principal component score/multiple linear regression (APCS/MLR), positive matrix factorization (PMF), generalised linear models (GLM), gradient boosted machine (GBM), random forest (RF), transmission electron microscopy (TEM), principal component analysis (PCA), species distribution models (SDMs), redundancy analysis (RDA) and generalised additive models (GAM).

2.2 Study area

The Heihe River Basin (HRB) has always been considered as a natural laboratory in many research fields (Cheng et al., 2014). The HRB originates in the Qilian Mountains along the northeastern margin of the Qinghai-Tibet Plateau (QTP) that is known as the 'Third Pole of the World' (Fig. 2-2A). It flows 821 km, covers an area of approximately 14.3×10^4 km^2, and the mean annual runoff is 1.588×10^9 m^3 (Ge et al., 2013). It is separated into upper, middle, and lower reach areas by the Yingluoxia Hydrological Station (YLX) and the Zhengyixia Hydrological Station (ZYX) (green triangles in Fig. 2-2B).

The elevations decrease from above 5,000 m asl in the headwater area into nearly 900 m asl in the terminal lakes (Fig. 2-2C). Topographically, the basin ranges from the Qilian Mountains in the south, through the middle reaches in the Hexi Corridor Plain, to the foothills and Alxa plain, which partly border the Badain Jaran and Tengger deserts. The main geomorphological compartments include the glacier, alpine ice-snow and permafrost, water conservation forest, piedmont oasis, and desert oasis (Li et al., 2020). On the one hand, these numerous natural contexts provide a perfect research platform for the HMs transport among various carriers, but on the other hand they cause a lot of troubles for field sampling and observation. In other words, it becomes challenging to ensure the spatial consistent of carriers within a catchment by the conventional design and framework. Hydrologically, the upper reaches, containing 88% of the total annual water resources in the basin, is the main runoff generation area. Due to the extremely arid climatic and over exploited of water resource, the mainstream in the lower reaches becomes wide and sluggish, or even dried up in some sections, which hinders the consistent water sampling in the field.

In terms of the climate conditions, the mean annual precipitation in the upstream areas ranges from 200 to 700 mm, while mean annual temperature decreases with elevation from 4 to -3°C. The vegetation includes the mountain grassland, mountain forest and shrub, and alpine meadow (Li et al., 2020). The mean annual temperature in the middle stream district is about 3 to 7°C, and the mean annual precipitation ranges from 50 to 200 mm, and annual potential evaporation ranges from 2,000 to 3,000 mm. In the downstream, the average annual evaporation is 3,755 mm, the mean annual precipitation is 42 mm (Xi et al., 2010), and the ratio of annual potential evapotranspiration to the annual precipitation is up to 100 (Lu et al., 2021). Typical desert riverside forest, shrubbery, and pratoherbosa meadow vegetation grow along the river, mainly including Populus euphratica forest, P. euphratica-Tamarix sp. mixed forest and Tamarix sp. forest (Xi et al., 2010). The soil erosion in the middle and downstream areas is super serious and frequent, and these areas are the main source region of the Loess Plateau. Therefore, these

distinct natural backgrounds provide sufficient preconditions for extracting environmental explanatory variables.

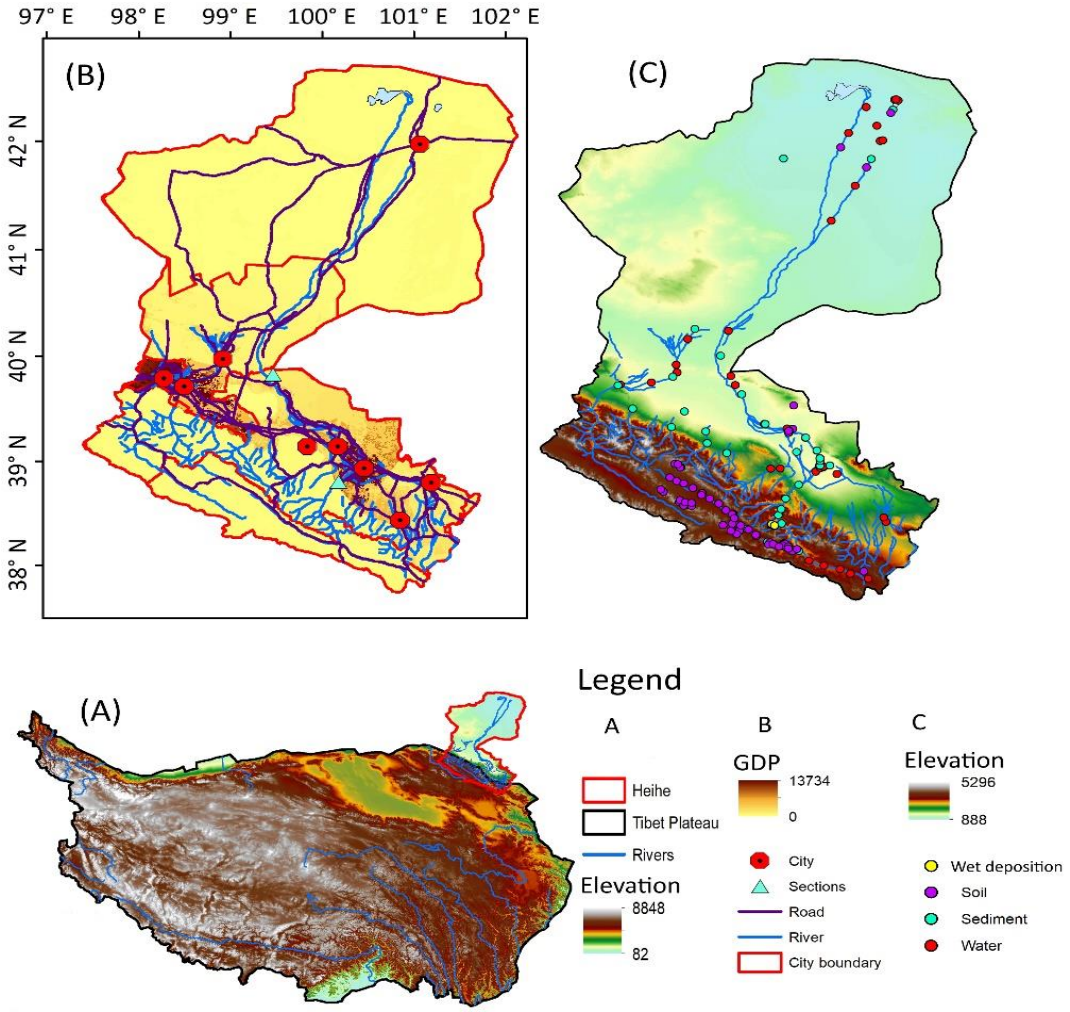


Figure 2- 2. (A) The Heihe River Basin (HRB) locates in the northeastern margin of the Tibet Plateau (TP) and the mainstream generates in the TP. (B) shows the main cities distribution, sections, road, and GDP (10^4 RMB km^{-2}) distribution. (C) The sampling sites in this study.

As for the social- economic conditions, the HRB covers three administrative provinces of Qinghai, Gansu, and Inner Mongolia autonomous region. The upstream area is dominated by animal husbandry, coupled with many small-scale mines. The mines had been placed into operation since the 1980s due to the special metallogenic conditions and resources (Li et al., 2020). Some mines were even located near glaciers where is the important water resource area. The conservation policy of the Qilian Mountain National Park had been adopted by the government in June 2017 to ban mining activity (Yan and Ding, 2020). The middle reaches, where accounts 95% of the population and 88.7% of the economy contributions (Fig. 2-2B), is renowned for its developed agriculture, nonferrous metals, steel, and petrochemical resources

(Zhang et al., 2020b). The agricultural development in the middle stream has a history of more than 2,000 years. It is a commodity grain base in the northwest region and has an important social and economic status. With the implementation of the 'Great West Exploitation Strategy' and 'Silk Road Economic Belt', agriculture and industry in Hexi Corridor have developed rapidly. The downstream district is a pastoral area mainly inhabited by ethnic minorities. The tertiary industries are developing rapidly in recent decades. The tourist population had exploded and increased from 0.03 million in 2000 to 1.1 million in 2015, and the tourism income accounted for two-thirds of the tertiary industry in 2015 (Lu et al., 2021). The agriculture chiefly includes cantaloupe and cotton farming, and animal husbandry with sheep and camels. These regional social- economic characteristics are obviously distinct within the catchment. Thus, this study area provides an ideal research site for this study.

2.3 Framework design

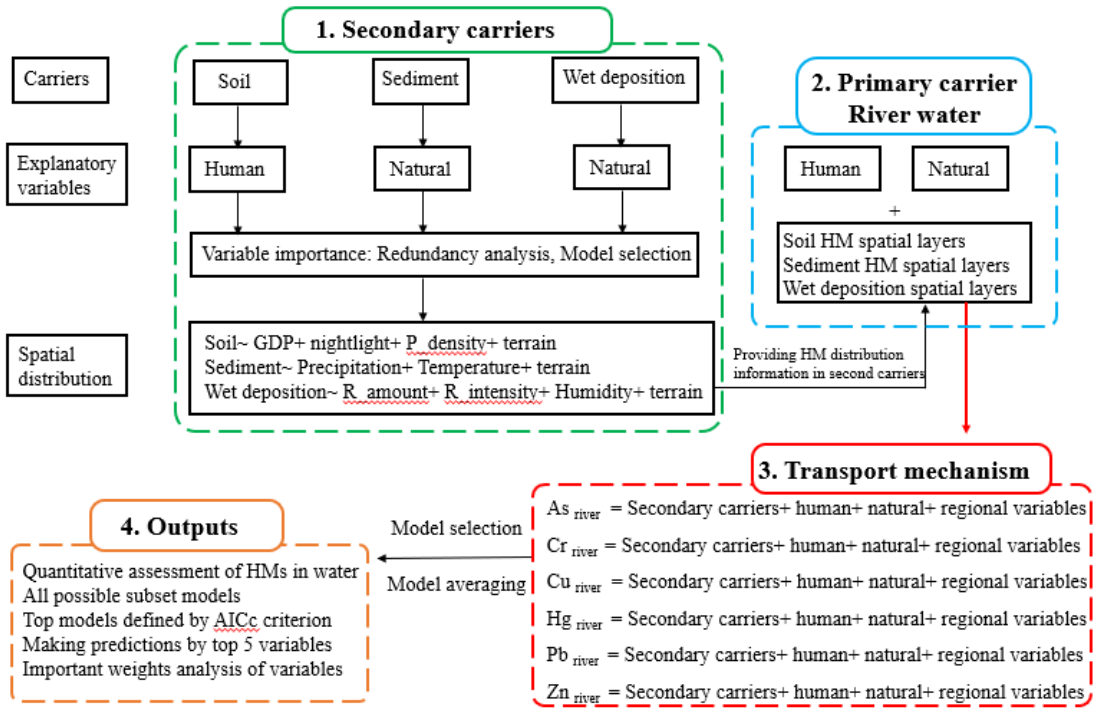


Figure 2- 3. Flowchart of this study with specific information for each step.

There were 4 main steps in this research (Figure 2-3). (i) Create spatial distribution layers of secondary carriers through SDM. In order to distinguish the effects of explanatory variables, soil sampling sites were close to the human activity locations, while sediment and wet deposition were far away from human activity. The importance analysis of explanatory based

on RDA and model selection will be then performed, and important variables will be used for prediction. (ii) Select systematic explanatory variables for primary carries, including human footprint variables and environmental variables, as well as the HM layers of secondary carriers. (iii) Relationships between HMs in river water and explanatory variables were analyzed by glm and gam. (iv) The spatial distributions of HMs in river were predicted by SDM, considering all the potential variables.

The methods and database used in this study were interpreted according to the main components of SDM. The main steps for SDMs include (i) conceptualization, (ii) data preparation; (iii) model fitting, (iv) performance assessment, and (v) outputs (Elith and Leathwick, 2009; Guisan and Zimmermann, 2000).

2.3.1 Model conceptualization

In ecology, the SDMs use species presence-only data and explanatory variables to assess species' niches and predict their potential habitats. In other research fields, the various entities are regarded as different 'species', and the areas showing the positive relations with research objectives are defined as 'habitat', such as the 'emotional habitat' (Li et al., 2021).

In this study, each HM, like As, Cd, Pb, Zn in the geographic space can be regarded as the 'species', and the spaces where show high concentrations are regarded as the 'Heavy metal habitat'. And we can depict the relationship between HMs and their explanatory variables, then predict HMs spatial distribution by the explanatory variables.

2.3.2 Data preparation

2.3.2.1 Field sampling of various carriers

For the purpose of this study, the primary carrier river water, secondary carriers soil, modern fluvial sediment, and atmospheric deposition were sampled within the HRB catchment (see Figure 2-2C). There was no sampling density or spatial consistent requirements for the primary carrier collection but tried to cover the entire catchment scale. A sets of river water samples were collected along the mainstream and tributaries of the HRB both in winter (December) and summer (July) periods. All river water samples were filtered with 0.22 μm polycarbonate membranes in the field into polythene bottles that had been thoroughly prewashed with deionized water. Samples were stored at 4 °C until being shipped to the laboratory for further HMs concentration analysis.

To further separate the influence from natural and anthropogenic variables, the secondary carriers were sampled with intentions. The sites of collecting the soils were around typical land

use types, for example the mining soils in the upstream, agricultural and industry soils in the middle stream, and the livestock field soils in the downstream areas (Figure 2-2C). To obtain representative samples, a series of standard soil sampling procedures were adopted (Li et al., 2020). Approximately 1 kg of fresh soil was collected using a clean plastic dustpan and brush and was stored in plastic bags at each site. On the contrary, the modern fluvial sediment samples were collected from riverbed and floodplain. All the sampling sites were far from urban areas and farmland to minimize the influence of human activity (Figure 2-2C). Each sample was taken from the subsurface below 10 cm depth by using a precleaned spade. Prior to analysis, both soil and sediment samples were air dried at ~20 °C; sorted through a 2 mm plastic sieve to remove gravel-sized stones, large plant roots and other debris; ground; homogenized with an agate mortar; and passed through a 200-mesh sieve (Zhang et al., 2020a). The atmospheric depositions were collected at a Qinghai spruce and Qilian juniper forest sites, with an elevation of 2850 and 3050 m asl respectively in the upstream of the catchment. The samples were collected after each rainfall event, and stored in clean 2 L polyethylene bottles, and refrigerated at 4 °C before analysis (Zang et al., 2021). The atmospheric deposition samples were filtered through a 0.45 μm polyester fiber membrane filter. The filtrates (rainwater) were transferred to clean polyethylene bottles and ultrapure HNO_3 was added to acidify the samples (pH < 2). The particulates collected on the filters (particulate matter) were dried at 40 °C, then weighed and refrigerated at 4 °C for further determination of elements.

The concentrations of HMs were measured by the inductively coupled plasma mass spectrometer (ICP-MS, 7500a, Agilent, Santa Clara, CA, USA). For quality control and assurance, blanks and standard reference materials (GBW07408 (GSS-8)) were processed using same digestion procedure. For most of the HMs, the corresponding relative standard deviation was below 5%. The grain size of the sediment was separated by bromoform with a specific gravity of 2.89 g/cm^3. A total of 37 sediments were used to analyze the relationship with ecology factors, and 25 out of them were obtained to analyze the relationships with lithology. All these samples were collected and analyzed around 2015, and finally a valid dataset of 94 water, 65 soils, 37 sediments and 8 wet depositions was collected for subsequent analysis.

2.3.2.2 Groups of explanatory variables

For the river water, we proposed two types of adequate and systematic explanatory variables based on the physical processes. One is the main explanatory variables, including the natural and human factors, the other is the auxiliary variables that we extracted from the corresponding HMs from the secondary carriers. The sets of explanatory variables were selected based on these principles: (a) variables can stand for typical natural and economic conditions in this catchment, such as the social- economy status; (b) variables are closely related to HMs transport

process, for example the frequent and strong erosions; and (c) variable has been proven by previous studies to have a profound impact on the HMs, such as the terrains around the mining sites. All these datasets were extracted from the social statistical, geographical information system products, and remote sensing dataset.

Main explanatory variables

The main explanatory variables were presented in Table 2-2. Land use data, terrain data, and meteorology data were all from the resources and environment data center of the Chinese Academy of Sciences (RESDC) (http://www.resdc.cn). Socio-economic data mainly included spatial distribution data of a 1 km grid GDP, nightlight, and population density. Transport data (railways, expressways, national highways and provincial roads) also came from RESDC. Population data (the ratio of urban to rural population) and economic development data (GDP and Engel coefficient) were obtained from The Statistical Bureau of Zhangye (2011). Finally, five natural variables and eleven human variables were extracted.

Table 2- 2. Potential natural and human explanatory variables for HMs. NA represents 'not applicable' in the relevant research.

Variables	Comments	Resolution	Unit
natural variables			
NDVI	Normalized difference vegetation index	1 km	NA
DEM	elevation	30 m	m asl
erosion	soil erosion classification	1 km	NA
temperature		500 m	℃
precipitation		500 m	10^{-1}mm
human variables			
GDP	gross domestic product	1 km	10^4 RMB km^{-2}
P_ density	population density	1 km	inhabitants km^{-2}
d_ road	distance to road	NA	m
night light	Range from 0 to 63	1 km	NA
land_1	forests	30m	m^2
land_2	grassland	30m	m^2
land_3	water	30m	m^2
land_4	unexploited land	30m	m^2
land_5.1	agriculture and plantations	30m	m^2
land_5.2	urban and rural residential	30m	m^2
land_5.3	industry, mining	30m	m^2

Auxiliary variables

The corresponding HMs concentration in secondary carriers was defined as the auxiliary variables to predict the primary carrier. Each carrier has their own sampling density during the filed collection process. We first calculated the respective determinant variables for each secondary carrier according to the SDMs model, and then predict their spatial distribution based on these determinant variables by SDMs. Once the spatial maps of each carrier are built within the catchment, the corresponding points can be extracted according to the location of the river water sampling point, and the geographic location can be synchronized. This is also our main methodology to solve gap1.

2.3.3 Model fitting

The choice of adequate modelling algorithms and desired model complexity should be guided by the research objective and by hypotheses regarding the specific entities-variables relationship (Elith and Franklin, 2013). According to the dataset characteristics, redundancy analysis (RDA), model selection and model averaging were applied in this model fitting process.

2.3.3.1 Redundancy analysis

RDA was introduced here to deal with multicollinearity and which variables should be included in the model. RDA is a powerful methodology to produce an ordination that regresses the impact of a matrix of multiple explanatory variables. The multiple linear regression is performed for the response variables in turn. The objects with similar variable values are ordinated closer together, while different values are projected apart.

Before analysis, the datasets are needed to be normalized to reduce the influence of different units. The qualitative variables are recorded as dummy variables (Legendre and Legendre, 2012), and they are different HMs species in this study. The explanatory variables with linear correlation will be removed. Then a permutation test with 1000 was conducted to examine the null hypothesis that no linear relationship exists between the response and explanatory variables. The constraining variables significant was examined by permutation test with 100000, and the determinant variables with higher significant value were remained when collinearity. The proportion of variance explained by each RDA axis is presented by an eigenvalue. Considering the extreme complexity of natural environment, the sum of the first two axes' eigenvalues above 10% is acceptable (Badry et al., 2019). To the best of our knowledge, the RDA was the first time to be applied in the HMs pollution area and will provide a more intuitive and direct concept for the HMs and variables.

2.3.3.2 Generalized linear models

Generalized linear models (glm) allowing the response variables to have error distribution models, have been widely used as an appropriate theoretic approach. Glm were applied in this study to firstly examine the ability of a reasonable combination of explanatory variables to explain variation of the HMs. Then the glm were applied to predict the spatial distribution of HMs based on the above determinant variables. For the wet deposition, the glm were only used for predicting spatial patterns, and the important variables were selected according to previous research (Stankwitz et al., 2012; Zang et al., 2021).

Table 2- 3. The generalized linear models used in this study.

determinant variables selection	
Soil	land use+ GDP+ Population density+ nightlight+ distance to road
Sediment_1	NDVI+ precipitation+ temperature+ erosion+ grain size
Sediment_2	Zircon+ Apatite+ Rutile+ Garnet+ Tourmaline+ Ilmenite+ Magnetite+ Amphibole+ Epidote+ Pyroxene+ Limonite+ Sphene
Water	human variables (GDP+ P_density+ nightlight+ d_road+ Land use) +Natural variables (NDVI+ precipitation+ temperature+ erosion) +Secondary carriers (wet deposition+ soil+ sediment)
Spatial distribution prediction	
Soil	[GDP+ nightlight+ P_density] + [terrain (dem+ slope+ tpi+ rough+ tri)]
Sediment	[precipitation+ temperature] + [terrain (flowdir+ dem+ slope+ tpi+ rough+ tri)]
Wet	[precipitation+ rainfall intensity+ humidity] + [terrain (dem+ slope+ tpi+ rough+ tri)]
water	[GDP+ d_road+ nightlight+P_density] + [precipitation+temperature] + [Secondary carriers] + [terrain (flowdir+ dem+ slope+ tpi+ rough+ tri)]

2.3.3.3 Generalized additive models

Generalized additive models (gam), where the response variable is not restricted to be linear in the explanatory variables, were then introduced to further estimate the smooth components of the glm models using smooth functions. The gam was introduced only for the primary carrier analysis in this study. According to the characteristics of our dataset, only five important variables were selected in the final models to depict the relationships between primary carrier HMs and their explanatory variables.

Prior to model fitting procedure, natural and anthropogenic variables were scaled by subtracting the mean and dividing by the standard deviation to enhance comparability of effect sizes across

variables measured in different scales. The collinearity between the explanatory variables was considered and only the most significant variables were retained for further analysis.

2.3.4 Model performance assessment

The models fitting should avoid overfitting or underfitting, and achieve a low generalization error that characterizes its prediction performance. In order to identify such a model, model selection and model averaging were used here to decide which model to select from candidate model families based on performance evaluations.

2.3.4.1 Model selection

When multiple model algorithms or candidate models are fitted, **model selection** was conducted to select the top model from a set of best models. The top model is the most parsimonious combination of explanatory variables using the cross- validation process. There are many selection criterions for model selection, and Akiake Information Criterion (AIC) is the most commonly used criteria, which is a measure of the goodness fit of an estimated statistical model (Burnham, K.P. and Anderson, D.R. 2002). The best models were then ordered by the criteria of AIC. The AIC is a measure of fit that penalizes for the number of variables. And **AICc** works better when the sample sizes are small:

$$\mathrm{AICc} = \mathrm{AIC} + \frac{2\mathrm{K}(\mathrm{K}+1)}{\mathrm{n}-\mathrm{K}-1} \qquad (1)$$

Smaller values indicate better fit and thus the AICc can be used to compare models. Thus, models with the lowest AICc are considered to be the top models, highlighting the included variables and model support relative to other models within 2 AICc.

2.3.4.2 Model averaging

The model averaging was then proposed for addressing the issues of uncertainty in the choice of probability distribution functions and the biased regression parameters during the model selection process (Burnham and Anderson, 2002). According to the characteristics of our dataset and the smooth functions, the selection of variables was based on: (a)whether corresponding to the research objectives well; (b) whether consistent with the related results in section 3.1 and 3.2; (c) sum of weights; and (d) whether collinear to each other. In this study, only the most important five explanatory variables were retained in the final model set to predict the relationships for each element in primary carrier.

2.3.5 Outputs and visualization

Once we have estimated this HMs-explanatory variables relationship, we can make predictions in space and in time by projecting the model onto available environmental layers.

The data extract, analysis, and maps visualization in this study were conducted in ArcGIS 10.5 (Environmental Systems Research Institute Inc.) and R v3.5.1 (R Core Team, 2018).

2.4 Results

2.4.1 Importance of explanatory variables for secondary carriers

2.4.1.1 Human driven factors via soil carrier

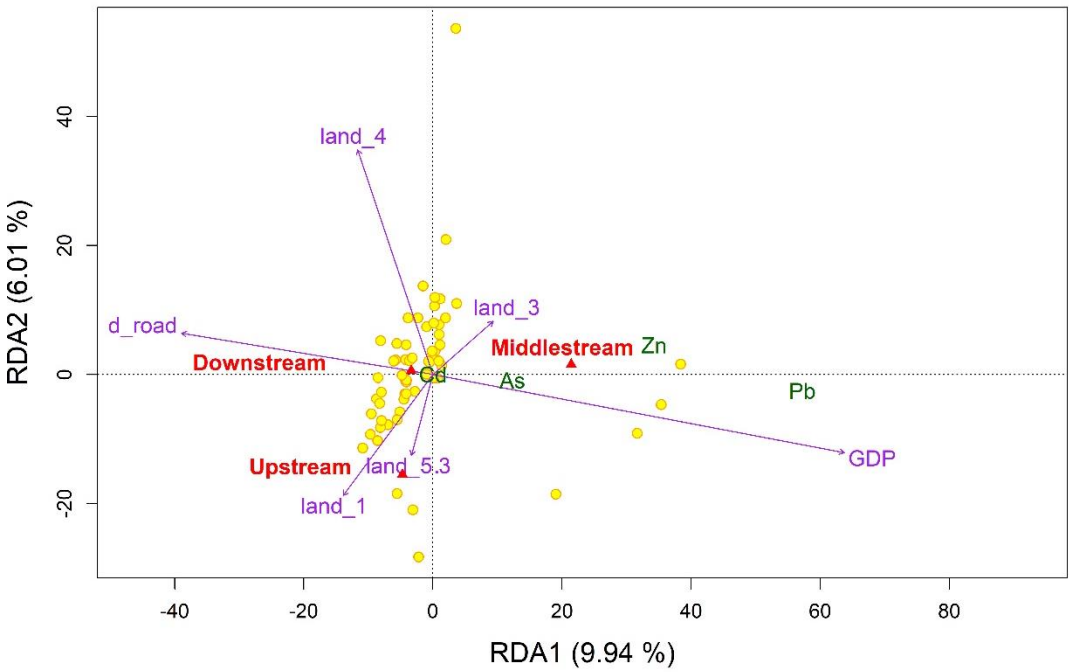


Figure 2- 4. Triplot of Redundancy analysis (RDA) showing the effects of human activity variables on soil heavy metals (R square=0.60). The yellow circles are the soil sampling sites, purple vectors represent the effect of explanatory variables.

For the HM concentrations in soil, all the peak values were recorded in the middle reach areas, where the average concentration of As in middle areas was even 20 times as that in the upstream areas (Table 2-4). According to the results of RDA, the GDP and 'distance to road' played an important role along the first axis, while other explanatory variables along the second axis (Figure2- 4). The elements of As, Zn and Pb correlated with the GDP, especially in the middle stream areas. On the opposite, the 'distance to road' was related to the downstream area. Cd

centered in the coordination system, showing a correlation with most of the explanatory variables. The land use types of forest and mining played an important role in the upstream area.

Table 2- 4. Average heavy metal concentrations of different carriers during summer period.

Carrier		As	Cd	Cr	Cu	Hg	Pb	Zn	unit
water	up	0.53	0.08	3.98	2.93	0.11	5.37	38.98	μg/L
	middle	0.68	0.31	4.67	3.54	0.33	6.98	32.42	
	down	0.56	0.25	3.62	3.57	0.25	5.74	49.17	
soil	up	15.36	1.64	35.44	34.87	0.04	24.72	107.62	mg/kg
	middle	320.55	2.70	91.68	46.03	0.06	1456.15	1030.23	
	down	1.45	0.21	29.12	14.88	0.02	4.72	55.57	
wet deposition		0.22	0.37	2.59	6.87	0.04	4.20	-	μg/L
sediments		-	-	255.17	-	-	28.79	-	mg/kg

Table 2- 5. Top models identified through model selection based on AICc. df= Degrees of freedom. Weight= Akaike weight.

HM	smoothed variables	df	loglik	AICc	weight
soil					
Arsenic	d_road+ GDP+ nightlight	5	-388.05	787.12	0.33
	d_road+ GDP+ land5.3+ nightlight	6	-387.58	788.6	0.16
Cadmium	d_road+ GDP+ land2+ P_density	6	-119.68	252.81	0.11
	d_road+ GDP+ land2	5	-121.03	253.07	0.1
Lead	d_road+ GDP+ nightlight	5	-495.92	1002.86	0.43
	d_road+ GDP+land_3+ nightlight	6	-495.46	1004.36	0.2
Zinc	d_road+ GDP+ nightlight	5	-471.43	953.87	0.11
	d_road+ GDP+ land_5.2+ nightlight	6	-470.28	954.01	0.1
sediment					
Chromium	2000μm+ mean size+ temperature	5	-185.93	384.16	0.14
	Garnet+ Zircon	4	-152.16	314.32	0.23
Nickel	63μm+ 500μm+ mean size	5	-140.64	293.58	0.24
	Garnet+ Zircon	4	-123.19	256.37	0.32
Lead	land_3+ precipitation	4	-108.89	227.27	0.09
	Magnetite+ Sphene	4	-79.28	168.55	0.66
water					
Arsenic	d_road+ precipitation	4	-9.12	27.19	0.09
Chromium	Cr.sedi+ Cr.wet	4	-77.48	163.92	0.22
Cooper	d_road+ land_2+ land_5+ nightlight	8	-57.93	135.65	0.16
Mercury	d_road+Hg.wet+land2+land3+land4+land5+nightlight	11	44.7	-59.86	0.04
Lead	d_road+GDP+land5.1+P_density+Pb.sedi+Pb.wet	8	-92.16	204.1	0.09
Zinc	d_road+GDP+land1+land2+land4+land5.1	8	-169.17	358.13	0.39

element.sedi: the related element concentration in sediment; element.wet: the related element concentration in wet deposition.

Top models with low AICc values were listed in Table 2-5. The top 2 models for each HM content in soil were listed in the Table 2-2. We hypothesized that the soil HM contents might be predominant by the different land use types, but only land use type of grassland showed correlation with Cd from the top 1 model. While 'distance to road', GDP and nightlight played a relatively more important role in correlated soil HMs than the different land use types.

2.4.1.2 Nature driven factor via sediment carrier

The tri-plot of Figure 2-5a showed that the grain sizes of 0-63, 250-500 μm explained together 54.8% of the total variance of the HM data. The 0-63 μm size showed significant influence in the downstream area. The upstream area was related to high precipitation. For the lithology RDA (Figure 2-5b), the Ilmenite, Garnet, and Magnetite played an important role in the first axis (36.54%). The upstream was related to Epidote rocks, while downstream was correlated to Sphene and Tourmaline.

The '500-2000μm grain size', the 'mean size', and 'temperature' were the best predicted parameters for Cr in the sediments (Table 2-5). And the 'Garnet' and 'Zircon' lithology influenced Cr heavily. For Ni, the '0-63 and 250-500 grain sizes' showed strongest correlation, as well as the same lithology factors from Cr. On the opposite contrary, the grain sizes showed no significant impacts on Pb, but the 'land use type of water' and 'precipitation' played an important role. And unlike Cr and Ni, the lithology type of Magnetite and Sphene predicted the Pb best (Table 2-5).

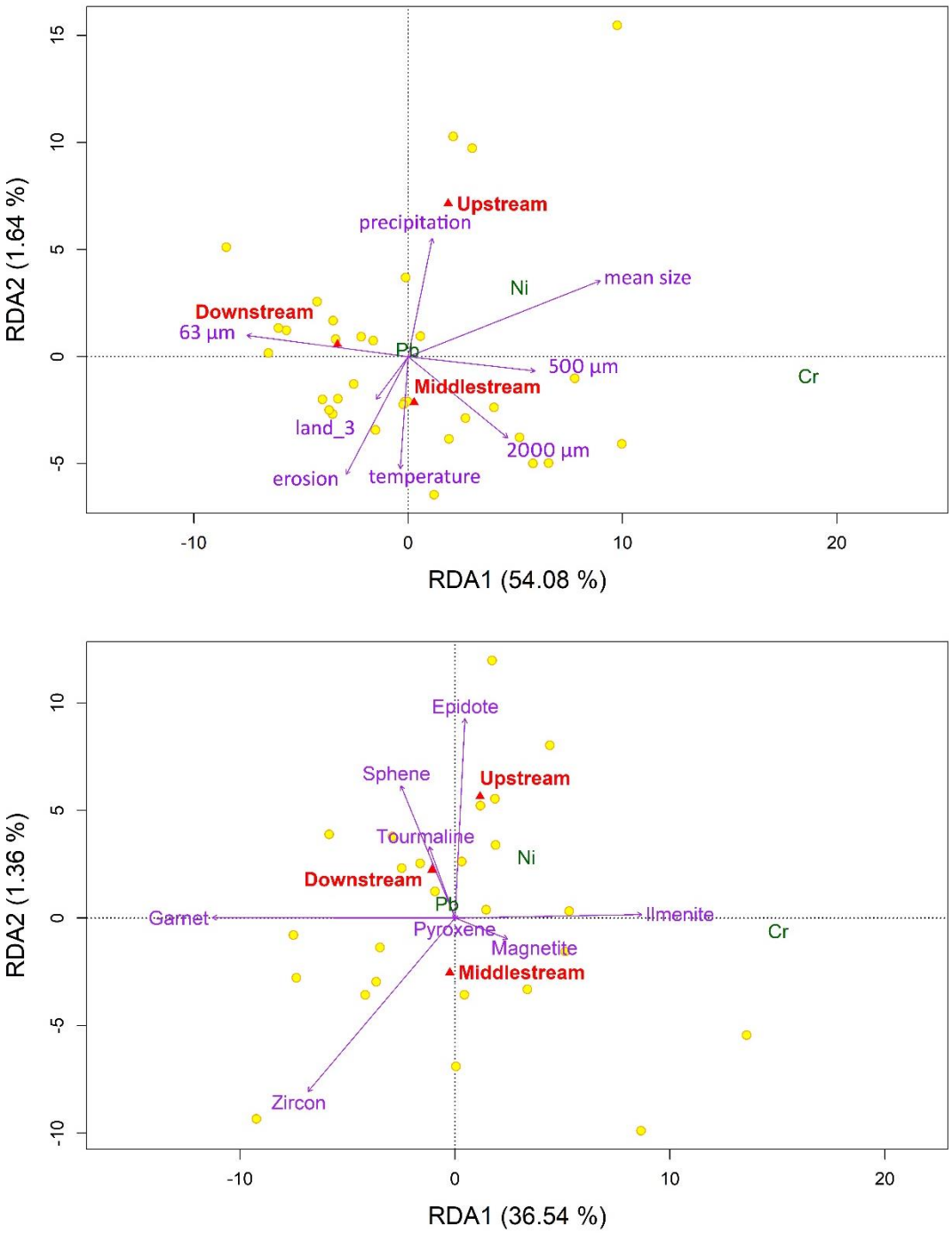


Figure 2- 5. Tri-plots of Redundancy analysis (RDA) showing the effects of grain sizes and lithology on sediment heavy metals (R square=0.56 and 0.38, respectively). The yellow circles are the sediment sampling sites. Purple vectors represent the effect of environmental variables, the red triangles are the catchment segments, and the green characters are plotted as the response variables. 63µm:0-63 µm; 500 µm:250-500 µm; 2000 µm:500-2000 µm; mean size: the mean size of all the sediments; erosion: soil erosion classification. The eigenvalue is the proportion of variance explained by each RDA axis.

2.4.2 Spatial distributions of secondary carriers modelled by SDM

2.4.2.1 HMs in soil

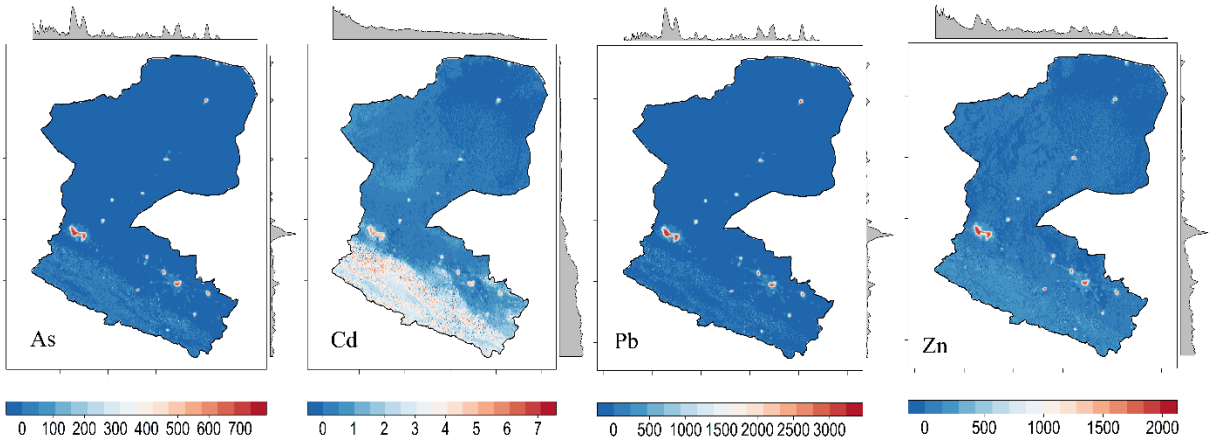


Figure 2- 6. Predicted distributions of soil heavy metals (Unit: mg/kg).

The explanatory variables we used to predict the HM distribution in soil are listed in Table 2-3, including GDP, nightlight, and population density, as well as the terrain characteristic factors. The concentrations of As, Pb, and Zn showed peak values around the big cities in middle areas, especially Jiuquan and Jiayuguan where are typical industrial cities in northwestern China. While the concentration of Cd had high values across the upstream areas (Figure 2-6).

2.4.2.2 HMs in sediment

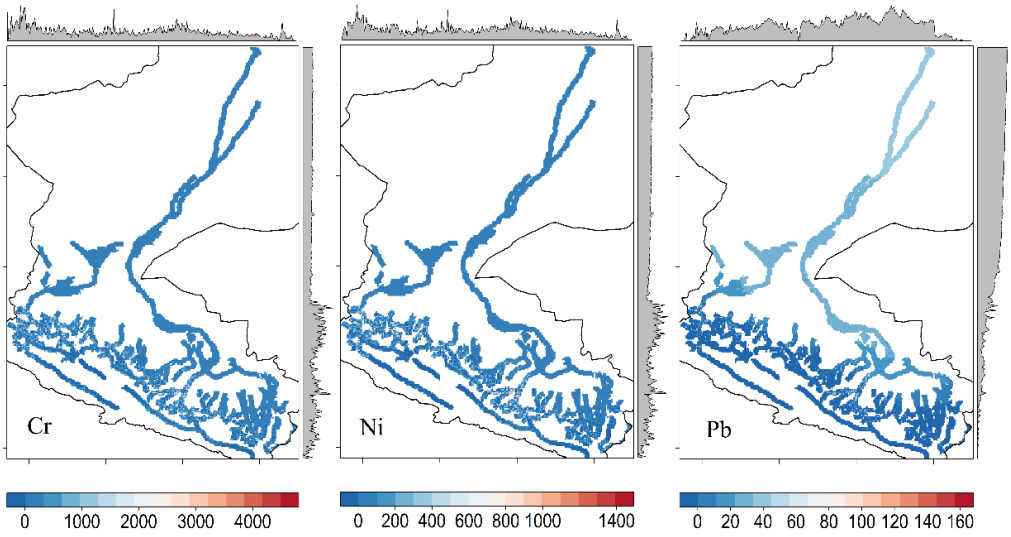


Figure 2- 7. Predicted distributions of sediment heavy metals (Unit: mg/kg).

The precipitation amount and temperature were selected as the explanatory variables according to the results of glm (Table 2-5), and the terrain factors were also considered in the SDM to

predict the spatial distribution of HMs in the sediment. The concentrations of all the HMs showed an increased trend from upper reaches to lower reaches, especially Pb showing the most obviously increasing characteristic in the downstream areas (Figure 2-7). However, high values of each element in the upstream were distributed in dots, reflecting the influence of the hydraulic sorting process.

2.4.2.3 HMs in wet deposition

Precipitation amount, rainfall intensity, humidity, and terrain factors were selected as the explanatory variables to predict the spatial distribution of HMs in wet deposition (Table 2-5). Output maps showed that the high values of HMs in wet deposition were distributed in the downstream areas, but relatively lower in the upstream and middle stream segments. These features might be explained by the scavenging effect and washout process of precipitation. In downstream areas where precipitation is scarce, HMs will continuously be accumulated in atmospheric particles, so high concentrations of HMs were observed during the limited precipitation events. On the contrary, the large amount and high frequency of precipitation in the upstream areas made the concentrations of HMs relatively lower.

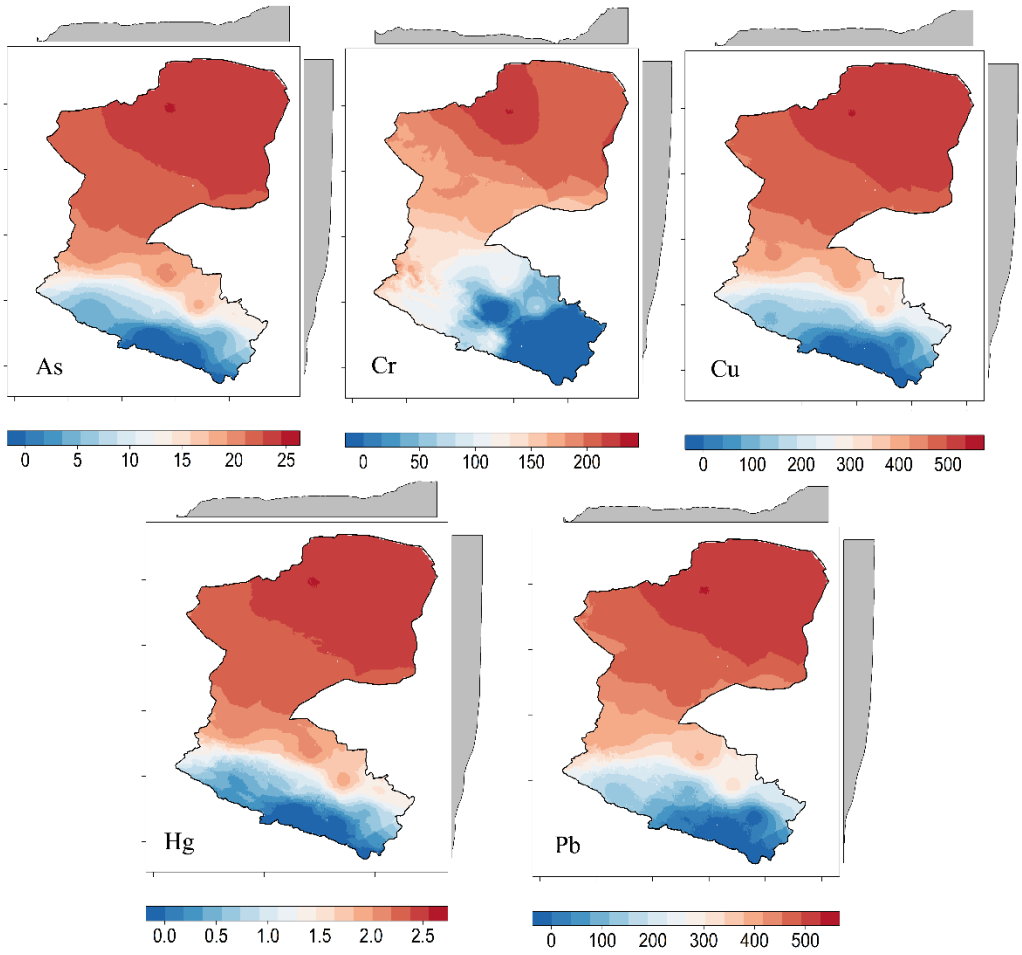


Figure 2- 8. Predicted distributions of wet deposition heavy metals (Unit: μg/L).

2.4.3 HM transmission at large scale catchment

In summer period, the highest values of As, Cd, Cr, Hg and Pb in river water were all from the middle stream segments, while the upstream areas had the lowest values for most of the HMs (Table 2-4).

2.4.3.1 Spatiotemporal distributions of heavy metals in river water

Except for Cu and Zn, the other four elements in the middle reaches were significantly higher than those in other sections. As and Cr in the middle reaches were significantly higher in winter than in summer, while Cu and Hg showed an opposite trend (Figure 2-9).

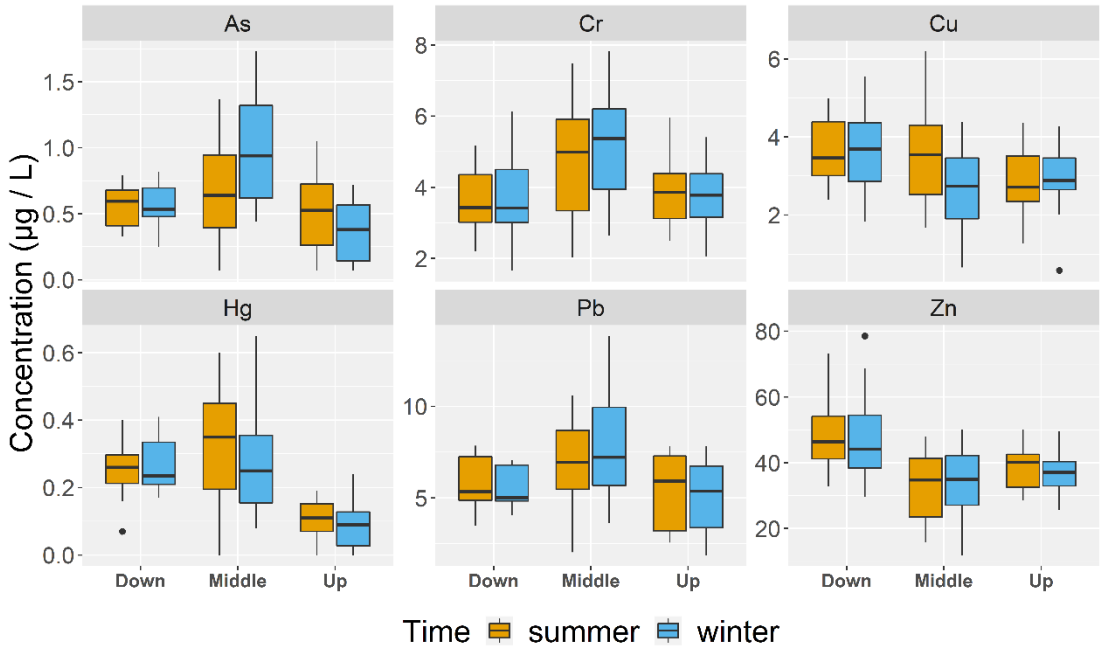


Figure 2- 9. The temporal and spatial distributional characteristics of water heavy metals. (Down: downstream, Middle: middle stream, Up: upstream).

2.4.3.2 Importance of variables for HMs in river water

According to the top models for water HMs (Table 2-5), As only showed correlation with 'distance to road'. Top model for Cr was explained by the Cr content from sediment and wet deposition. In terms of Hg, it was influenced by the 'distance to road', injection from the wet deposition, 'night light', as well as most the land use types except for forests.

The importation from sediments and wet deposition played an important role in the water Pb concentration, as well as the economic parameters, including 'distance to road', GDP, 'land use type of agriculture' and 'population density'. None of the ecology variables were presented in the top models for water HMs according to the AICc selection.

2.4.3.3 Relationships between HMs in river with all explanatory variables

Table 2- 6. Generalized Additive Models describing the response of water heavy metals to the major explanatory variables. Approximate significance of smooth terms.

HM	variables	Edf	ref.df	F	P	weights	N	DE
Arsenic	d_road	1.00	1.00	7.23	0.01	1	5	31.30%
	GDP	2.29	2.77	1.14	0.28	0.42	2	
	precipitation	1.34	1.57	0.10	0.12	0.13	1	
	As.wet	1.00	1.00	0.07	0.79	0.13	1	
	As.soil	1.96	2.35	1.59	0.18	0.12	1	
Chromium	d_road	1.57	1.94	0.65	0.07	1	5	48%
	GDP	1.00	1.00	4.62	0.04	0.54	2	
	precipitation	5.73	6.75	2.65	0.03	0.87	4	
	Cr.wet	1.00	1.00	3.05	0.09	0.63	3	
	Cr.sedi	1.00	1.00	0.29	0.59	0.13	1	
Cooper	d_road	1.00	1.00	2.02	0.16	0.63	8	41.50%
	GDP	1.49	1.78	0.49	0.61	0.5	7	
	precipitation	1.00	1.00	2.49	0.12	0.38	5	
	Cu.wet	1.00	1.00	3.02	0.09	0.24	3	
	nightlight	5.68	6.57	1.71	0.19	0.24	4	
Mercury	d_road	1.00	1.00	0.54	0.47	-	-	76.50%
	GDP	2.19	2.74	1.19	0.30	-	-	
	precipitation	2.27	2.80	5.54	0.01	1	2	
	Hg.wet	7.37	8.24	3.03	0.01	-	-	
	land_5.1	1.00	1.00	0.93	0.34	0.59	1	
Lead	d_road	1.00	1.00	1.46	0.23	0.51	4	46.10%
	GDP	1.00	1.00	1.93	0.17	0.39	3	
	Pb.wet	1.47	1.76	4.33	0.08	0.58	4	
	Pb.sedi	2.91	3.55	1.61	0.25	0.48	3	
	Pb.soil	2.35	2.80	3.35	0.02	1	7	
Zinc	d_road	1.00	1.00	3.99	0.05	0.33	2	41.30%
	GDP	1.00	1.00	2.26	0.14	0.67	4	
	Zn.soil	4.10	4.83	2.46	0.06	0.67	4	
	land_1	1.00	1.00	0.83	0.37	0.42	2	
	land_5.1	1.66	2.06	0.69	0.53	0.14	1	

Edf= Degrees of freedom estimated to model. Ref.df= Degrees of freedom estimated to waste. F= Significance of smoothed terms. P= P value. DE= Deviance explained. weights= sum of weights (relative importance values, for each explanatory variable). N= N

containing models (the number of times that each explanatory variable was included in the total models).

According to the research objectives and the importance analysis of the variables (Table 2-5), the top five significant variables were chosen as the final variables to explain the HMs in river water. The 'distance to road' showed negative relationships with As, Cu, Hg and Pb, but a positive influence for Zn, and a non-linear relationship with Cr (Figure 2-10). The GDP had positive relationship with Pb, but negative with Cr and Zn. While for As, Cu and Hg, the GDP showed a negative trend at first until the GDP increasing to 500, then turned into position relationships. The precipitation had positive relationships with As and Cu, but roughly negative relationships with Cr and Hg. The influence from wet deposition, showed positive relationships with As, Cu and Pb, but negative with Cr, and non-linear relationship with Hg. The influence of HMs from soil carrier played important roles in As, Pb, and Zn. It firstly showed a positive trend, but then turned into negative trend when the concentration of As and Pb were around 7.5 ppm.

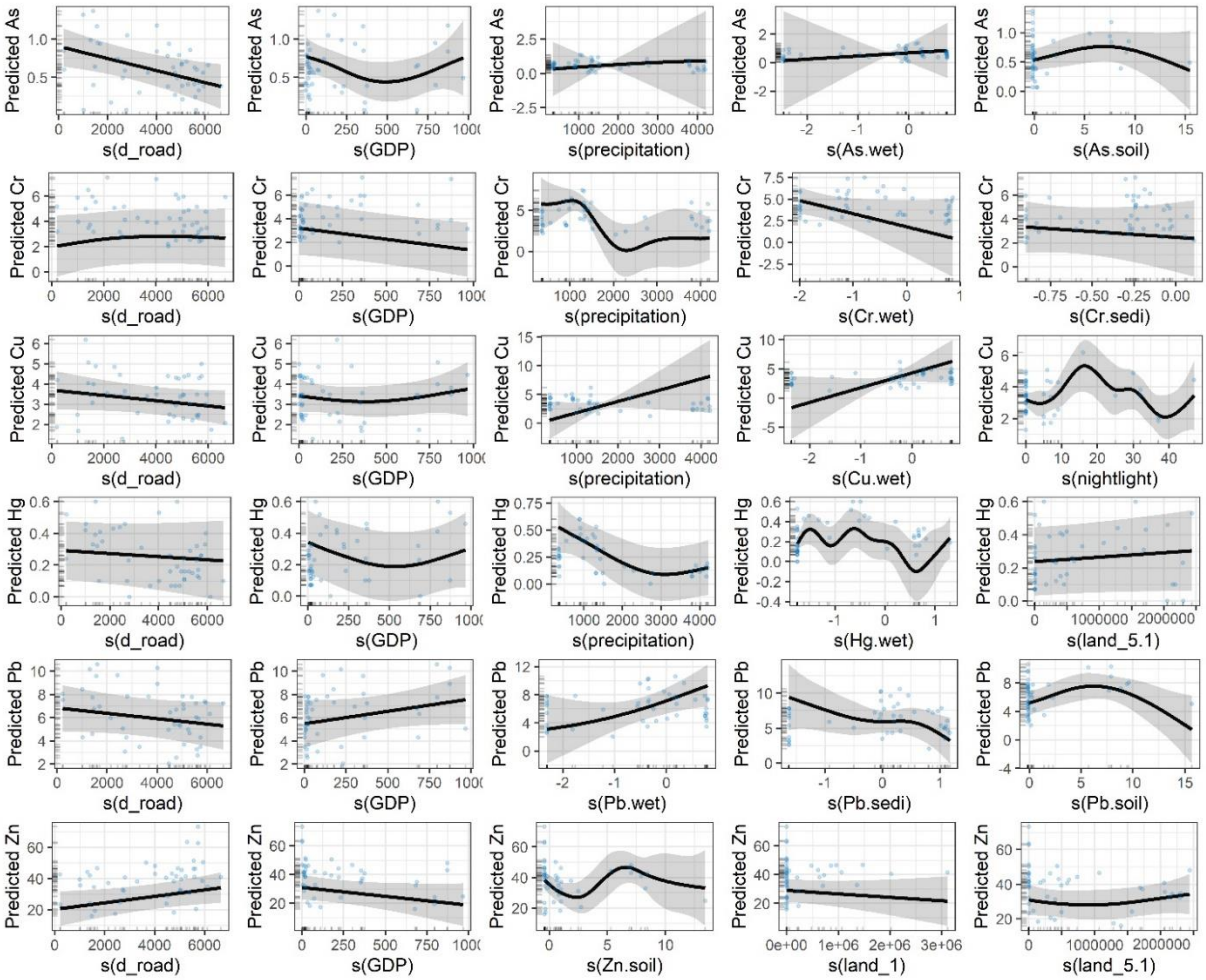


Figure 2- 10. Smoothed fits of relationships between variables and water heavy metals.

2.4.3.4 HM habitats and hotspots in river water

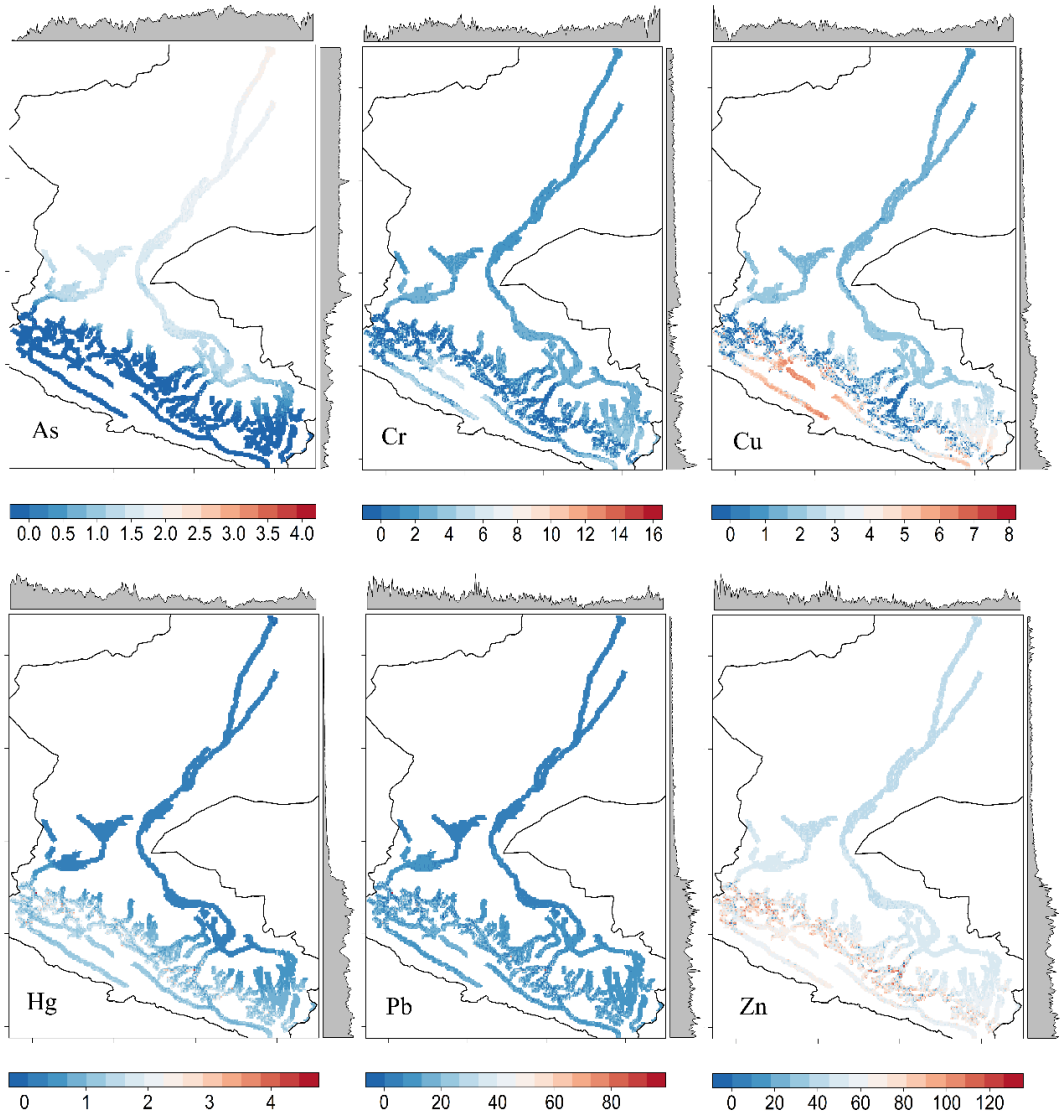


Figure 2- 11. Predicted distributions of heavy metals in river water, without considering the related heavy metals in secondary carriers (Unit: μg/L).

The human footprint variables, including GDP, population density, nightlight, distance to road, and land use types, together with the environmental variables, including NDVI, precipitation, temperature, and erosion classification were chosen as the explanatory variables in SDM to predict the HMs in river water. In order not to discount the influence of HM transport in catchment scale, the related HMs from secondary carriers were also considered in our study. And the output maps without and with considering the HMs in secondary carriers were presented in Figure 2-11 and Figure 2-12 respectively, and the latter method significantly improved the prediction accuracy. For instance, there were new high concentrations of As around the cities if we considered the contribution of arsenic transport from soil and wet deposition (Figure 2-12). Many high concentration points of Pb were

predicted in the upstream area when we considered the impact of Pb from the three secondary carriers, which further reflected the impact of human activities in the upstream. Additionally, the predicted concentrations of Hg and Pb showed an increased trend after we joined the HMs in secondary carriers.

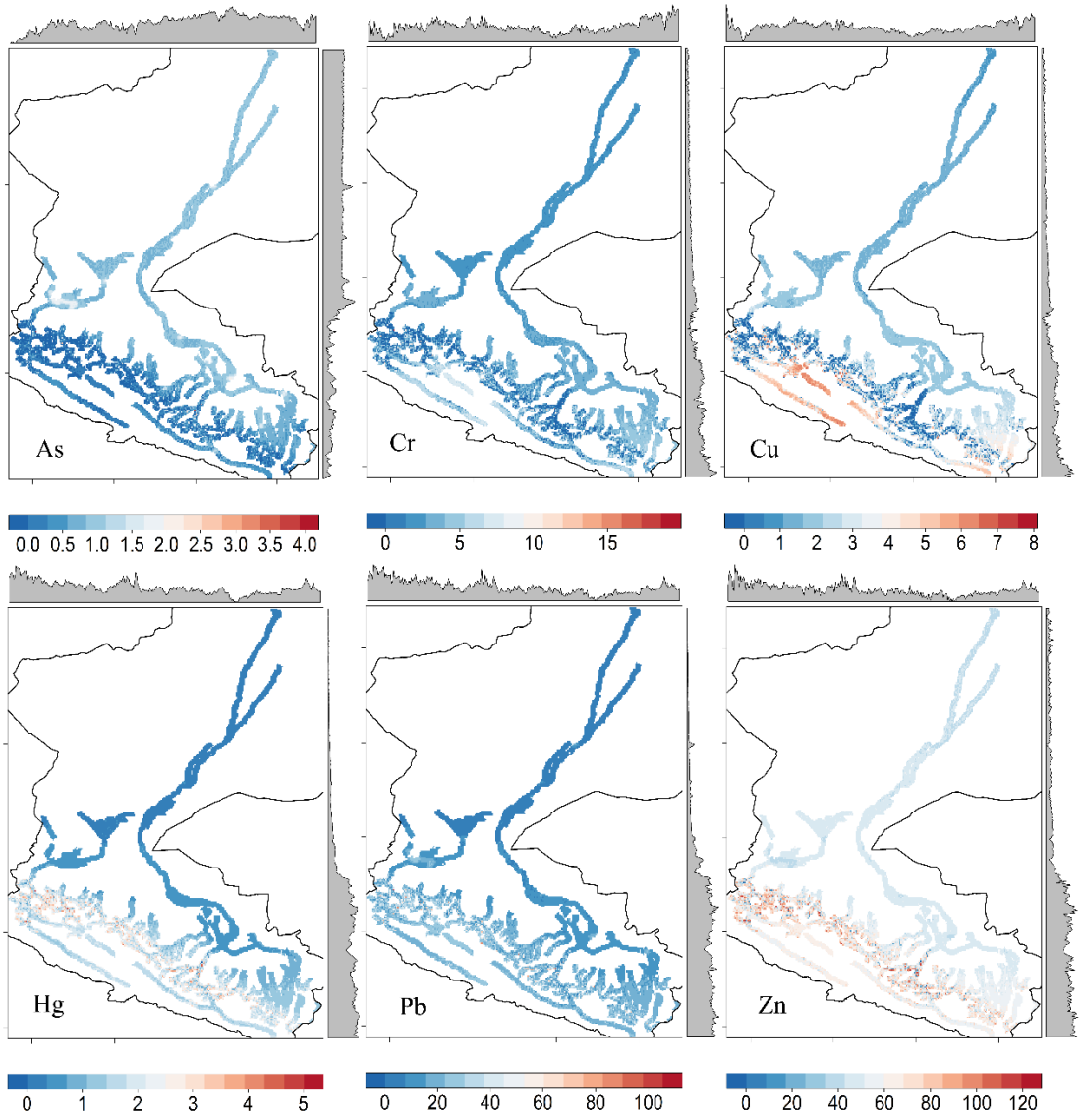


Figure 2- 12. Predicted distributions of heavy metals in river water, with considering the related heavy metals in secondary carriers (Unit: µg /L).

2.5 Discussion

2.5.1 Improvement of methodologies for HM mapping

Many methodologies have been developed to study the HM mapping and transport mechanism, and their pros and cons are discussed here. Source apportionment is one of the most important

methods to identify pollutant sources and quantify their contribution. Many published studies have employed the receptor models as source apportionment tools. Enrichment factors (EF) and Principal component analysis (PCA) models are two relatively simple receptor models, and have been applied widely in the HMs source apportionment research (Cable and Deng, 2018; Lv, 2019). PCA, as a linear technique for dimensionality reduction, could transform many HMs into few principal components (PC) corresponding to a certain source (Li et al., 2017). However, PCA cannot interpret the quantitative contributions of natural and human sources to HMs (Lv, 2019). The introduction of multivariate receptor models can help to quantify the source apportionment of HMs, such as the absolute principal component score/multiple linear regression (APCS/MLR) and positive matrix factorization (PMF). These receptor modellings use physical and chemical characteristics of pollutants to identify and apportion their contributing sources (Taiwo et al., 2014). However, there are four main limitations in these models. (i) Taiwo et al. (2014) underlined that the classification of potential sources and the understanding of different transport processes within each particular source are required before applying these models, considering the significant difference in potential origins. And the combination of two or more of these receptor models is strongly suggested to evaluate the uncertainties associated with different models. (ii) These methods ignore the spatial correlations between sampling points, which will increase the uncertainties and errors in subsequent calculations (Lv, 2019). (iii) Unlike the identification of comparable contributions from various sources, these methods are unable to assess the importance of the different variables in the modelling process (Hu et al., 2020). (iv) Further, few receptor models have succeeded in predicting the relationships between HMs and explanatory variables, and this detailed information is exactly necessary for governing the elevated levels of HMs contaminants. So, digging the relations between HMs and potential variables should be an important research direction to make up the limitations of receptor models.

Technically, some technologies have been developed to measure the HMs transport and transformation from nanoparticle scale. Studies have demonstrated that combining the transmission electron microscopy (TEM) and multielement detection is a more accurate and effective way to directly provide the morphological information of HM transport process on micro level (Hassellov and von der Kammer, 2008; Hochella Jr et al., 2005). However, these nanoparticle technologies require considerable cost, time, and effort, and may be very sensitive to experimental settings, which regrettably hinders their promotion to a wider catchment range. Statistically, the statistical analysis is fortunately a powerful method to find the relationships between HMs and explanatory variables. The traditional statistical algorithms, such as the respective linear models, have been applied in previous research (Nickel et al., 2014; Sharma et al., 2008). The advancement of machine learning and effective computing technology

provides a new perspective regarding HMs transport mechanism, including the generalized linear models (GLM), gradient boosted machine (GBM), and random forest (Hu et al., 2020), etc.

And in this study, the GLM, GAM, model selection, and model averaging which can quantitatively identify the relationships between different variables, were introduced. Compared to PCA and other receptor models, the RDA can not only qualitatively analyze the source, but also quantify the explanatory effect and analyze the relationship between samples in different clusters. The combination of interdisciplinary and multi-methods facilitates the improvement of the purposes of this research.

2.5.2 Overcoming the challenges of field sampling with SDM

The sampling strategy in field is a core scientific issue for studying the interactions among various carriers. Geographical continuity is an important prerequisite for establishing connections between different carriers. To ensure this, the main solution of previous research is the 'proximity' strategy during the sampling process (Table 2-1). The typical research includes the proximity between the soil and river water (Hassellov and von der Kammer, 2008; Quinton and Catt, 2007), soil and crops (Hu et al., 2020), atmospheric deposition and soil (Liang et al., 2017b), atmospheric deposition and vegetables (Sharma et al., 2008), flood plain and riverbed sediment (Hochella Jr et al., 2005).

However, current methods of studying carriers' interactions are challenging to ensure the spatial consistence of carriers, which will discount the accuracy of subsequent analysis. There are two limitations in this proximity design. Firstly, for each sampling point, the sampling conditions are site specific in situ, and this will definitely increase the difficult for collecting multi-dimensional carriers in one site. Complicated terrains, harsh weather, undeveloped public transport system, and other constraints at the field work make it very difficult or even impossible to collect multi-dimensional carriers at the same location (Li et al., 2020). Secondly, for the entire catchment, especially with large scales, different carriers' collecting sites cannot distribute at the same sampling density or rate under actual circumstances. A dearth of consistent data for different carriers is often credited as a major limitation for HMs interaction research.

The data requirement in SDM is the presence only data, that is, just recording where the presence of the species is observed. It can be seen that SDM has relatively low sampling requirements and is very friendly to scientific problems that cannot provide continuous sampling over large space. The sampling in this study was based on SDM requirement, and collected different carriers in typical areas across the entire catchment as much as possible.

2.5.3 Providing possibilities for systematic explanatory variables

Insufficient explanatory factors will affect the completeness and accuracy of the analysis results. The selection of appropriate indirect and direct potential variables has been identified as a key step of identifying the HMs sources and depicting the transport processes (Nickel et al., 2014). Estimating the explanatory variables has previously been undertaken using three main conventional approaches. (i) Collecting samples in typical land use areas, so land use is the prevailing explanatory factor (Li et al., 2017; Liang et al., 2017a; Liang et al., 2017b; Sharma et al., 2008). This means that a systematic reconnaissance survey was required to evaluate the land use types or other information before sampling (Sharma et al., 2008). (ii) This method is also applicable to point source pollution. Sampling needs to be taken next to typical site point with pollution activities (Hochella Jr et al., 2005; Liang et al., 2017b; Quinton and Catt, 2007). The mining operations are important point sources in previous research, like the waste rock and smelter waste (Liang et al., 2017b). (iii) For some specific experimental sites, a series of synchronous observation equipment is necessary. For example, at the location where the precipitation is collected, the synchronous equipment should be arranged to record the rainfall amount, intensity, duration, and wind speed, etc. (Zang et al., 2021).

Nevertheless, compliance sampling networks are not usually designed to obtain data for the above conventional methods. The observables, locations, and time periods to estimate HMs transport processes are also limited by the complex field situations and the lack of observation sites. Recently, Hu et al. (2020) proposed that the use of geographic information system, remote sensing products and other geographical information to extract explanatory variables is a very time-saving and efficient method. But this study only used land use maps as auxiliary variables and did not involve other human variables. More importantly, in addition to these two major categories of environment and human factors, few studies have used the secondary carriers as explanatory factors to explain and predict the primary carrier.

The Species distribution models (SDMs) were introduced and proposed here. SDMs can predict a species' probability of occurrence across a landscape by combining the species presence data with its explanatory variables (Kearney and Porter, 2009). Another important advantage of SDM is that it can superimpose and calculate many explanatory variables at the same time, and the promotion of GIS and other products provides convenience for the extraction of these variables information. In this study, we systematically collected almost all available GIS products in this area. Before choosing the final variable, we also evaluated the importance of each variable through RDA and model selection. Then the most influential explanatory variables were selected according to the importance results and research purposes. This research is also the first attempt to take the corresponding concentrations in the secondary

carriers into consideration, as a factor for the transport of HMs between different carriers to predict the distribution of HMs in river water. Comparing to the output maps without secondary HM concentrations, our final results can improve the prediction accuracy.

2.5.4 Potential of SDM in predicting the spatial distributions of HMs

The spatial distribution of HMs has been regarded as an important part of contaminant risk assessment (Li et al., 2017). The performance of current HM spatial prediction methods is not perfect enough. The HM spatial distribution is a powerful tool to estimate the environmental pollution. The Kriging interpolation is one of the most prevalent spatial interpolation methods in previous study (Li et al., 2017; Liang et al., 2017a). Strong evidence has shown that increasing the sampling density is the main approach to improve the accuracy of predictions (Lv, 2019). But these conventional interpolations ignore the influence of many determinants, and more attention should therefore be paid to investigate factors that could govern HMs transport and distribution (Li et al., 2017). Some studies had tried to tackle with this problem. Lv (2019) explored to apply the geostatistical techniques to estimate the HMs spatial distribution based on the factors resulting from the receptor models. Nickel et al. (2014) proposed to combine multivariate generalized linear models with Kriging interpolation to create HMs maps at a high level of spatial resolution. However, both these two studies tried to ensure the accuracy of prediction by increasing the sampling density (Table 2-1), which often encounters the field sampling problems. In addition, lots of preparatory work for these two studies was required and this undoubtedly increased the complexity of spatial prediction. For example, the explanatory variables must be calculated and selected through the acceptor models beforehand.

The prediction results of SDM in this study were not only considered the influence of the background such as topography, water flow direction, etc., but also taken into account the influence of human activities. This is a significant improvement over the predicted results of previous studies. For example, our soil sampling points are not distributed in Jiuquan and Jiayuguan areas, but the SDM model predicted that there were peak values in this place. The high values distribution of these HMs was consistent with t previous studies, which proved the feasibility and universal applicability of our model.

2.6 Conclusion

The spatial inconsistence and scarcity of the field observations are profoundly hindering our ability to map HMs over large areas and understanding of the interactions among their carriers. To address this issue, this study developed a novel methodological framework for mapping

heavy metal contamination over HRB with a species distribution model. We applied this framework with field samples of HMs to model the HM 'habitats' and identify contamination 'hotspots' across the catchment.

Results found that all the peak values of HM concentrations of soil samples were recorded in the middle reach areas. The variables of distance to road, GDP, and nightlight played a relatively important role in soil HMs. The output maps of HMs from soil, sediment, and wet deposition could respectively reflect the influence of industry contaminants, hydraulic sorting, and precipitation washout process, which proved the rationality of SDM in predicting the spatial distribution of secondary carriers. Concentrations of As and Cr of the river water in the middle reaches were significantly higher in winter than in summer, while Cu and Hg showed an opposite trend. Both the human footprint variables and environmental variables made contributions to the HMs in river water. All these variables, together with the HM transport from secondary carriers were selected explanatory variables in SDM to predict the HMs in river water, which helped improve the prediction accuracy. The methodological framework proposed in this study facilitates future rigorous exploration in Hydrological chemistry and pollution research.

2.7 References

Ali, H., Khan, E. and Ilahi, I. 2019. Environmental chemistry and ecotoxicology of hazardous heavy metals: environmental persistence, toxicity, and bioaccumulation. Journal of chemistry 2019.

Badry, A., Palma, L., Beja, P., Ciesielski, T.M., Dias, A., Lierhagen, S., Jenssen, B.M., Sturaro, N., Eulaers, I. and Jaspers, V.L. 2019. Using an apex predator for large-scale monitoring of trace element contamination: Associations with environmental, anthropogenic and dietary proxies. Science of the total environment 676, 746-755.

Braunisch, V., Patthey, P. and Arlettaz, R. 2011. Spatially explicit modeling of conflict zones between wildlife and snow sports: prioritizing areas for winter refuges. Ecological Applications 21(3), 955-967.

Burnham, K.P. and Anderson, D.R. 2002. A practical information-theoretic approach. Model selection and multimodel inference 2.

Cable, E. and Deng, Y. 2018. Trace elements in atmospheric wet precipitation in the detroit metropolitan area: levels and possible sources. Chemosphere 210, 1091-1098.

Cheng, G., Li, X., Zhao, W., Xu, Z., Feng, Q., Xiao, S. and Xiao, H. 2014. Integrated study of the water–ecosystem–economy in the Heihe River Basin. National science review 1(3), 413-428.

Clemente, P., Calvache, M., Antunes, P., Santos, R., Cerdeira, J.O. and Martins, M.J. 2019. Combining social media photographs and species distribution models to map cultural ecosystem services: The case of a Natural Park in Portugal. Ecological indicators 96, 59-68.

Ducci, L., Agnelli, P., Di Febbraro, M., Frate, L., Russo, D., Loy, A., Carranza, M.L., Santini, G. and Roscioni, F. 2015. Different bat guilds perceive their habitat in different ways: a multiscale landscape approach for variable selection in species distribution modelling. Landscape ecology 30(10), 2147-2159.

Elith, J. and Franklin, J. (2013) Encyclopedia of Biodiversity: Second Edition, pp. 692-705, Elsevier Inc.

Elith, J. and Leathwick, J.R. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual review of ecology, evolution, and systematics 40, 677-697.

Fang, H., Huang, L., Wang, J., He, G. and Reible, D. 2016. Environmental assessment of heavy metal transport and transformation in the Hangzhou Bay, China. Journal of hazardous materials 302, 447-457.

Ge, Y., Li, X., Huang, C. and Nan, Z. 2013. A Decision Support System for irrigation water allocation along the middle reaches of the Heihe River Basin, Northwest China. Environmental Modelling & Software 47, 182-192.

Goth, A., Michelsen, A. and Rousk, K. 2019. Railroad derived nitrogen and heavy metal pollution does not affect nitrogen fixation associated with mosses and lichens at a tundra site in Northern Sweden. Environmental Pollution 247, 857-865.

Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. Ecology letters 8(9), 993-1009.

Guisan, A. and Zimmermann, N.E. 2000. Predictive habitat distribution models in ecology. Ecological modelling 135(2-3), 147-186.

Hassellov, M. and von der Kammer, F. 2008. Iron oxides as geochemical nanovectors for metal transport in soil-river systems. Elements 4(6), 401-406.

Hochella Jr, M.F., Moore, J.N., Putnis, C.V., Putnis, A., Kasama, T. and Eberl, D.D. 2005. Direct observation of heavy metal-mineral association from the Clark Fork River Superfund Complex: Implications for metal transport and bioavailability. Geochimica et cosmochimica acta 69(7), 1651-1663.

Hu, B., Xue, J., Zhou, Y., Shao, S., Fu, Z., Li, Y., Chen, S., Qi, L. and Shi, Z. 2020. Modelling bioaccumulation of heavy metals in soil-crop ecosystems and identifying its controlling factors using machine learning. Environmental Pollution 262, 114308.

Kearney, M. and Porter, W. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecology letters 12(4), 334-350.

Legendre, P. and Legendre, L. (2012) Numerical ecology, Elsevier.

Li, J., Li, Z., Brandis, K.J., Bu, J., Sun, Z., Yu, Q. and Ramp, D. 2020. Tracing geochemical pollutants in stream water and soil from mining activity in an alpine catchment. Chemosphere 242, 125167.

Li, X., Yang, H., Zhang, C., Zeng, G., Liu, Y., Xu, W., Wu, Y. and Lan, S. 2017. Spatial distribution and transport characteristics of heavy metals around an antimony mine area in central China. Chemosphere 170, 17-24.

Li, Y., Fei, T., Huang, Y., Li, J., Li, X., Zhang, F., Kang, Y. and Wu, G. 2021. Emotional habitat: Mapping the global geographic distribution of human emotion with physical environmental factors using a species distribution model. International Journal of Geographical Information Science 35(2), 227-249.

Liang, J., Feng, C., Zeng, G., Gao, X., Zhong, M., Li, X., Li, X., He, X. and Fang, Y. 2017a. Spatial distribution and source identification of heavy metals in surface soils in a typical coal mine city, Lianyuan, China. Environmental Pollution 225, 681-690.

Liang, J., Feng, C., Zeng, G., Zhong, M., Gao, X., Li, X., He, X., Li, X., Fang, Y. and Mo, D. 2017b. Atmospheric deposition of mercury and cadmium impacts on topsoil in a typical coal mine city, Lianyuan, China. Chemosphere 189, 198-205.

Lu, Z., Feng, Q., Xiao, S., Xie, J., Zou, S., Yang, Q. and Si, J. 2021. The impacts of the ecological water diversion project on the ecology-hydrology-economy nexus in the

lower reaches in an inland river basin. Resources, Conservation and Recycling 164, 105154.

Lv, J. 2019. Multivariate receptor models and robust geostatistics to estimate source apportionment of heavy metals in soils. Environmental pollution 244, 72-83.

Meite, F., Alvarez-Zaldívar, P., Crochet, A., Wiegert, C., Payraudeau, S. and Imfeld, G. 2018. Impact of rainfall patterns and frequency on the export of pesticides and heavy-metals from agricultural soils. Science of the Total Environment 616, 500-509.

Nickel, S., Hertel, A., Pesch, R., Schröder, W., Steinnes, E. and Uggerud, H.T. 2014. Modelling and mapping spatio-temporal trends of heavy metal accumulation in moss and natural surface soil monitored 1990–2010 throughout Norway by multivariate generalized linear models and geostatistics. Atmospheric Environment 99, 85-93.

Quinton, J.N. and Catt, J.A. 2007. Enrichment of heavy metals in sediment resulting from soil erosion on agricultural fields. Environmental science & technology 41(10), 3495-3500.

Sarkar, D.J., Sarkar, S.D., Das, B.K., Sahoo, B.K., Das, A., Nag, S.K., Manna, R.K., Behera, B.K. and Samanta, S. 2021. Occurrence, fate and removal of microplastics as heavy metal vector in natural wastewater treatment wetland system. Water Research 192, 116853.

Senduran, C., Gunes, K., Topaloglu, D., Dede, O.H., Masi, F. and Kucukosmanoglu, O.A. 2018. Mitigation and treatment of pollutants from railway and highway runoff by pocket wetland system; A case study. Chemosphere 204, 335-343.

Sharma, R.K., Agrawal, M. and Marshall, F.M. 2008. Heavy metal (Cu, Zn, Cd and Pb) contamination of vegetables in urban India: A case study in Varanasi. Environmental pollution 154(2), 254-263.

Stankwitz, C., Kaste, J.M. and Friedland, A.J. 2012. Threshold increases in soil lead and mercury from tropospheric deposition across an elevational gradient. Environmental science & technology 46(15), 8061-8068.

Stojic, N., Pucarevic, M. and Stojic, G. 2017. Railway transportation as a source of soil pollution. Transportation Research Part D: Transport and Environment 57, 124-129.

Taiwo, A.M., Harrison, R.M. and Shi, Z. 2014. A review of receptor modelling of industrially emitted particulate matter. Atmospheric environment 97, 109-120.

Walden-Schreiner, C., Rossi, S.D., Barros, A., Pickering, C. and Leung, Y.-F. 2018. Using crowd-sourced photos to assess seasonal patterns of visitor use in mountain-protected areas. Ambio 47(7), 781-793.

Xi, H., Feng, Q., Si, J., Chang, Z. and Cao, S. 2010. Impacts of river recharge on groundwater level and hydrochemistry in the lower reaches of Heihe River Watershed, northwestern China. Hydrogeology Journal 18(3), 791-801.

Yan, K. and Ding, Y. 2020. The overview of the progress of Qilian Mountain National Park System Pilot Area. International Journal of Geoheritage and Parks 8(4), 210-214.

Zang, F., Wang, H., Zhao, C., Nan, Z., Wang, S., Yang, J. and Li, N. 2021. Atmospheric wet deposition of trace elements to forest ecosystem of the Qilian Mountains, Northwest China. Catena 197, 104966.

Zhang, J., Geng, H., Pan, B., Hu, X., Chen, L., Wang, W., Chen, D. and Zhao, Q. 2020a. Climatic zonation complicated the lithology controls on the mineralogy and geochemistry of fluvial sediments in the Heihe River basin, NE Tibetan Plateau. Quaternary International 537, 33-47.

Zhang, Y., Lu, Y., Zhou, Q. and Wu, F. 2020b. Optimal water allocation scheme based on trade-offs between economic and ecological water demands in the Heihe River Basin of Northwest China. Science of The Total Environment 703, 134958.

Chapter 3. Hydraulic connections between groundwater and surface water at different hydrological units in arid regions

Abstract

The hydraulic connections between different hydrological units play an important role in assessing the water resources, and these connections could be very complicated in the arid regions due to the extremely frequent interactions between surface water and groundwater, as well as the spatial inconsistency of river water across large scales. The active shallow groundwater is not only affected by interactions from surface water and deep groundwater, but also by a series of natural and human factors. An accurate understanding of the relationships between shallow groundwater and these explanatory variables has always been a knowledge gap that restricts hydraulic connection research in arid regions. The Badain Jaran Desert (BJD) and two adjacent catchments of the Heihe River Basin (HRB), Shiyanghe River Basin (SYH), from northwestern China were selected as the typical hydrological units. Surface water, shallow groundwater and deep groundwater were systematically sampled along with the potential groundwater flow directions. The end member mixing analysis (EMMA) was applied to analyze the potential water quantity connections between these different units, and then the species distribution model (SDM) was introduced to predict the spatial distributions of contaminant tracers, which considered both the geological and hydrological factors. Finally, the relationships between shallow groundwater and explanatory variables were explored by model selection and model averaging.

Results found that (a) the chemical evolution of most water types was mainly controlled by rock dominance, (b) the spatial distribution of highest values in shallow groundwater sulfates coincided with the locations of typical industrial cities according to our output maps, and (c) in addition to the sulfate replenishment from surface water and deep groundwater, the sulfates in shallow groundwater were also affected by regional variables such as population density, and the Mg2+ and EC contents in surface water. The introduction of the EMMA and SDM methodologies can help improve the accuracy of predicting the spatial distributions of groundwater. These results can also provide new ideas for revealing the hydraulic connections, and especially water pollution between different hydrological units in arid areas.

Key words: hydraulic connection, hydrological unit, arid regions, water quantity interaction, water quality, major soluble ions, stable isotopes, end member mixing analysis, species distribution model, Badain Jaran Desert, Heihe river basin, Shiyanghe river basin

Highlights:

* Systematically sampling typical water samples is the prerequisite for analyzing the hydraulic connections between adjacent hydrological units;

* The EMMA model can quantitatively analyze the water quantity interactions between surface water and groundwater in different hydrological units;

* The conceptual structure of EMMA provides a platform and theoretical basis for studying water contaminant transportation;

* Geological background and hydrological factors affect the spatial distribution of geochemical contents of groundwater;

* Sulfates in shallow groundwater were affected by replenishments from both surface water and deep groundwater, as well as some regional variables

3.1 Introduction

The hydrological cycle plays an important role in various environmental compartments of Earth (Mook and Rozanski, 2000). It not only provides water for life and production, but also helps maintain the ecological balance (D'Odorico et al., 2019). However, if the water body is polluted, the contaminants will be spread throughout the hydrological cycle, posing a great potential threat. Hydraulic connection is an important indicator that reflects the hydrological cycle (Lehr et al., 2015), and it includes two primary criteria, the water quantity interaction and water quality connection (Banks et al., 2011). Evaluating the mechanisms of hydraulic connections is essential for a clear understanding of the hydrological cycle, especially the water contaminant movement in and out of the aquatic ecosystems.

A hydrological unit refers to a specific drainage area with unique geographic characteristics that usually have obviously distinct hydrographic or topographic boundaries (Kwon and Kim, 2010; Seaber et al., 1987). These boundaries are just like administrative boundaries, so the hydrological unit is also defined as the 'hydrological province', such as the desert and adjacent catchment (Wang et al., 2018). There is little hydraulic connection between neighboring hydrological units in surface water due to the boundaries, but the conditions in groundwater could be more complicated. For example, the fracture networks (Chen et al., 2004), connected aquifers, or non-zero hydraulic heads difference between adjacent hydrological units (Gates et al., 2008), provide possibility for groundwater flow interactions across the unit boundary (Wang and Zhou, 2018). In addition, the infiltration recharge, various groundwater levels, unclear geological and aquifer conditions, further affect the connections between different hydrological units. This study defines surface water as Layer 1

(L1), and includes: river water, lake water, and reservoir water, while shallow phreatic groundwater is defined as Layer 2 (L2), and deep confined groundwater as Layer 3 (L3).

The hydrochemistry method is the most straightforward and cost- effective way to trace and interpret such hydraulic connections information (Wang et al., 2018). The hydrochemistry tracers include major soluble ions, pH, EC, TDS, and stable isotopes etc., but in order to distinguish isotopes in this study, the term “hydrochemistry method” excludes them. The hydrochemistry methods offer advantages over some physically based methods like large water level gauging or monitoring networks, not only due to a much lower cost (Cook et al., 2013), but also providing more accurate information on interaction and transportation over large spatial regions. These advantages enable water chemistry methods to be widely applied for ascertaining source and origin, or geochemical evolution processes (Liu et al., 2015), the interactions between different water bodies (Maurya et al., 2019; Wang et al., 2018), and monitoring the contamination transport and accumulation mechanisms in the hydrological cycles (Hatzinger et al., 2013; Khalil et al., 2015; Liu et al., 2019). The research scale has been spatially extended to Arctic and Antarctic areas (Hodson et al., 2002; Hoshina et al., 2016; Hoshina et al., 2014).

Water interaction and transportation have obvious regional characteristics because they are affected by a wide range of factors, such as different climate backgrounds, geology, land use conditions, etc. (Klaus and McDonnell, 2013). This research will focus on the hydraulic connections in arid and semi-arid regions. Compared with other climate zones, common features of arid land include low and irregular precipitation, high temperatures, high evaporation rates, and fragile ecological environments (Liu et al., 2015; Wang et al., 2015). These characteristics have caused multiple interactions and reuseof surface water and groundwater, which is the most unique hydrological phenomenon in arid regions. Arid land occupies more than one third of the earth's land surface, and can be found in the Midwestern United States, Australia, Northwest China, Israel, and Africa etc. A thorough understanding of the hydraulic connections between different hydrological units in arid regions is necessary for water resources regulation and water quality protection on a global scale, which is important for the decision-making processes regarding sustainable development of the region (SDG6) (Sadoff et al., 2020).

Previous studies on the interaction and transportation of water bodies between different hydrological units in arid and semi-arid areas mainly focused on solving the following questions. (i) Although there is little hydraulic connection between different hydrological units in L1, what are the conditions for L2 and L3? (ii) How to quantitatively interpret the vertical hydraulic relationships between L2, and L1, L3, including the transportation of water

contaminants? (iii) The L2, shallow groundwater is susceptible to human activities, but how to quantify these human influences? (iv) How to map the spatial distribution of hydrochemistry tracers in groundwater L2, L3, and improve the accuracy?

Previous studies' attempts to solve these general problems were listed in Table 3-1, including both pros and cons. For instance, the conventional hydrograph separation methods, including the End Member Mixing Analysis (EMMA) (Guo et al., 2019; Khalil et al., 2015; Klaus and McDonnell, 2013), Bayesian isotopic mixing model (Hu et al., 2019; Jin et al., 2019), and other tracer-based methods, can quantitatively discover the interactions between different water bodies. It is worth noting that the conceptualized models of flow contribution and runoff generation based on these mass balance methods, can help improve accuracy by minimizing the influence of infiltration rate, geology and other conditions that affect groundwater movement, especially for large spatial scales. The LMWL (local meteoric water line) method can intuitively and effectively reflect the contribution of local precipitation, as well as other water bodies (Wu et al., 2017; Zhao et al., 2012). The introduction of other novel technologies, such as membrane-vaporization tools (Freyberg et al., 2017), fault system models, as well as GPR (ground penetrating radar) and gravimetric analyses (Wang and Zhou, 2018), has helped improve the accuracy of research on hydraulic connections.

However, it will face the challenge that whether the sampling sites can represent the typical hydrogeological background over large scales due to the unique hydrological phenomenon in arid areas, such as the spatial inconsistence of river water and frequent interactions. In addition, most studies on the contaminant transportation focused on the point source pollutions (Hu et al., 2019; Minet et al., 2017), or along with the river body (Khalil et al., 2015), and the geographic continuity and spatial georeferenced attributes of the sampling sites have been ignored to evaluate groundwater quality on a larger spatial scale.

Table 3- 1. Summary of relevant studies conducted previously for investigating the hydraulic connections between surface water and groundwater of various hydrological units in arid regions (The current paper is added for completeness). Four components in the 'Key results' column are identified by the code:(1) Interactions between different boundary conditions;(2) Vertical interactions between surface water and groundwater;(3) Identify human influence;(4) Spatial maps of groundwater. NA represents ‘not applicable’ in the relevant research.

Study	Water bodies	Methods	Key results
(Klaus and McDonnell, 2013)	Precipitation, soil water, river, shallow groundwater	hydrograph separation, EMMA	(1) Watershed scale (2) NA (3) Mined and mining reclamation (4) NA

(Wu et al., 2017)	lake, spring, shallow groundwater	LMWL	(1) Desert, lakes (2) Only shallow groundwater (3) NA (4) NA
(Chen et al., 2004)	lake, spring, rainfall, groundwater	LMWL	(1) Desert, catchment (2) Only shallow groundwater (3) NA (4) NA
(Hoshina et al., 2016)	shallow ice cores, snow pits	correlation analysis	(1) Antarctica (2) Surface snow pits with deep ice cores (3) NA (4) NA
(Freyberg et al., 2017)	river, precipitation	membrane-vaporization tools	(1) Catchment scale (2) NA (3) NA (4) NA
(Wang and Zhou, 2018)	groundwater, lake, river water	Fault system models; LMWL	(1) Desert, catchment (2) NA (3) NA (4) NA
(Liu et al., 2015)	lake, groundwater	Piper diagram, LMWL, Gibbs	(1) Catchment scale (2) Quaternary groundwater, the shallow Cretaceous groundwater, and the deep Cretaceous groundwater (3) Groundwater exploitation (4) NA
(Khalil et al., 2015)	Groundwater, canal water, waste water	EMMA	(1) Canal scale (2) Quaternary aquifer, Miocene aquifer (3) Pumping for irrigation and reclamation activities (4) NA
(Guo et al., 2019)	groundwater	EMMA, LMWL	(1) Endorheic basins (2) L2, semi-confined groundwater, Confined groundwater (3) Irrigation-based agricultural practices (4) NA
(Wang et al., 2013)	rainfall, river, lake, spring, groundwater	EMMA, LMWL, Gibbs	(1) Catchment with neighboring desert (2) Layer1, Layer2, Layer3 (3) NA (4) Simple interpolation
(Zhao et al., 2012)	precipitation, spring, groundwater	LMWL	(1) Catchments with neighboring desert (2) NA (3) NA (4) Kriging interpolation

This study	precipitation, river, lake, spring, groundwater	Cluster, EMMA, spatial models, model selection, model averaging	(1) Mountains, a desert, with two adjacent catchments (2) Layer1, Layer2, Layer3 (3) Human activity to the shallow groundwater (4) Predicted spatial distribution, based on the hydrogeological information

Abbreviations of the analytical methods used in Table 3-1 include end member mixing analysis (EMMA), local meteoric water line (LMWL).

As a result, although previous studies have conducted systematic research, there are still many knowledge gaps due to the difficulty of sampling high spatiotemporal resolution data from these extremely arid environments. Furthermore, the variable landscape, complex geological structure, and a wide range of climate in these high mountains and remote deserts make it even harder to collect ideal water samples that can typically stand for the potential groundwater flow directions in these arid hydrological units. There are four main knowledge gaps currently. (i) The sub-basin scale exchange is mostly focused on in previous models (Freyberg et al., 2017; Khalil et al., 2015; Klaus and McDonnell, 2013; Liu et al., 2015), which fails to analyze the action-response relationship between different hydrological units. Although a few studies have carried out research for inter-basin exchanges, the research scope is limited to around the boundary (Wang et al., 2013; Zhao et al., 2012), which cannot always represent the characteristics of the entire unit. (ii) Only the water quantity interactions were focused on by the conceptual models of the EMMA (Guo et al., 2019; Wang et al., 2013). To the best of our knowledge, few research has applied this conceptual model for the water pollutant transportation between different hydrological units. (iii) Some pollution tracers, like nitrates, sulfates (Hu et al., 2019; Khalil et al., 2015), were often selected to correspond with a specific pollution source, such as industrial or agriculture sites (Guo et al., 2019). Explanatory variables with spatial attributes were always insufficient. (iv) Current mapping methods of geochemical tracers are dependent on the spatial interpolation, for example Kriging (Guo et al., 2015; Yang et al., 2011; Zhao et al., 2012), which are still problematic due to the ignorance of hydrological and geological backgrounds, such as topography, hydrogeology, land use, groundwater and surface water extraction, and climatic conditions etc.

Therefore, systematic sampling along the potential hydraulic connection routes, and introduction of novel spatial distribution model with considering the geological and hydrological factors are two essential requirements to fill these current knowledge gaps. In this paper, three typical hydrological units, the Badain Jaran Desert (BJD), and two adjacent catchments: the Heihe River Basin (HRB) and the Shiyanghe River Basin (SYH), from Northwestern China were selected as the representative cases for such an endeavor. Surface water, shallow groundwater and deep groundwater were systematically sampled along the

potential groundwater flow directions based on previous research (Chen et al., 2006; Ding and Wang, 2007; Wu et al., 2010). Hydrochemistry and stable isotope methos were applied to analyze the vertical interactions of water quantity between these three layers, then the water quality connections were explored with the theoretical basis of the EMMA framework structure. The species distribution model (SDM) was introduced to predict the spatial distributions of contaminant tracers, considering both the geological and hydrological factors. In particular, we aimed to (i) study the hydraulic connections between different water types by using hydrochemistry tracers. (ii) Predicting the spatial distribution of geochemical indicators in layer 2 by SDM based on systematic field data, including boreholes, monitoring wells, hydrogeological cross section maps, and other hydrogeology information. (iii) Evaluate the water quality interaction between different layers based on the conceptual model framework of EMMA. This study attempts to develop a regional groundwater flow model between different hydrological units that unify all the groundwater 'sub-basins' as well as mountainous and desert areas and highlights the supplementary and complementary information of geological and hydrological factors to help constrain the calibration of the flow model.

3.2 Methods and dataset

3.2.1 Study area

The Badain Jaran Desert (BJD), and two adjacent catchments of the Heihe River Basin (HRB), Shiyanghe River Basin (SYH) were selected as the research site, which covers an area of approximately $38\times10^4 km^2$. The hydraulic connections between BJD and neighboring units were still unclear (Wang and Zhou, 2018; Yao et al., 2015). According to the research objectives, we adopted three possible groundwater flow directions based on previous research (Figure 3-1A), to examine their potential connections.

3.2.1.1 Badain Jaran Desert

The BJD is home to the tallest sand dunes on earth, as well as nearly 100 hypersaline lakes and freshwater lakes in the extremely arid environment. The identification of recharge sources for these lakes has attracted a lot of research, but none of these theories have been widely accepted due to lack of evidence (Wang and Zhou, 2018). BJD is located in the Alxa Plateau, covers an area of approximately $4.92\times10^4 km^2$. In the southern part, the BJD is separated from HRB and SHY by the Heli Mountains, Beida Mountains, Longshou Mountains, and Heishantou Mountains and the southeastern part of BJD is adjacent to the Tengger Desert, partially separated by the Yabulai Mountains (Figure 3-1A).

The landscape of the BJD is characterized by continuous crescentic mega-dunes, with a relative height of 200-300 m. The BJD geological structure belongs to the Alxa block, with a mild landform. It is mainly composed of denuded low hills and inter-mountain depressions. Quaternary sediments generally cover the surface, forming widely distributed Gobi and deserts (Figure 3-2). The BJD is situated at the limit of the East Asian Monsoon boundary, and the annual precipitation is 50-60 mm, which is mostly concentrated in June-August. The annual evaporation is greater than 3500 mm, and the annual average temperature is 7-8℃. The population density of BJD is less than 1 person / 10 km^2, with basically no agriculture.

3.2.1.2 Heihe River Basin

HRB is the second largest inland catchment in China, with a total length of 821 kilometers and a drainage area of approximately 14. $29\times10^4 km^2$. The dry climate in HRB is mainly controlled by the circulation of the westerly and polar cold air masses. The precipitation in the middle reaches decreases from 250 mm in the eastern part to below 50 mm in the west, while evaporation increases from east to west, from below 2000 mm to more than 4000 mm. In the downstream area of HRB, the annual average precipitation is only 42mm, but the annual average evaporation intensity is 3755mm. The annual average temperature is 8.04℃, the highest temperature is 41.8℃, and the lowest temperature is -35.3℃.

The HRB straddles three different environmental compartments, the Qilian Mountains in the upper reaches, the corridor plains in the middle, and the Alxa Plateau in lower reaches which is adjacent to the BJD. The underlying surface of the middle and lower reaches is a deep Quaternary sedimentary layer which becomes a good underground water storage place. Neo-Tertiary or Cretaceous layers constitute the base of the basin, on which is deposited hundreds of meters or even more than a thousand meters of proluvial-alluvial facies Quaternary loose material, and abundant groundwater exists in it (Figure 3-2).

3.2.1.3 Shiyanghe River Basin

The SYH originated in the eastern section of the Qilian Mountains, has a total length of 250 km and a drainage area of approximately $4.16\times10^4 km^2$. In the middle reaches of the SYH, the annual precipitation is 150-300 mm and the annual evaporation is 1300-2000 mm. The annual precipitation in the downstream area is 50 mm, and the annual evaporation is 2000-2600 mm. In the SYH region, there are no perennial rivers, and groundwater storage in the alluvial aquifer depends on direct recharge from precipitation (Zhu et al., 2007). The total population of the SYH basin is 2.27 million, and the population density is 55 people / 10 km^2, which is about 3.4 times of the average population density in the Hexi Corridor.

3.2.2 dataset

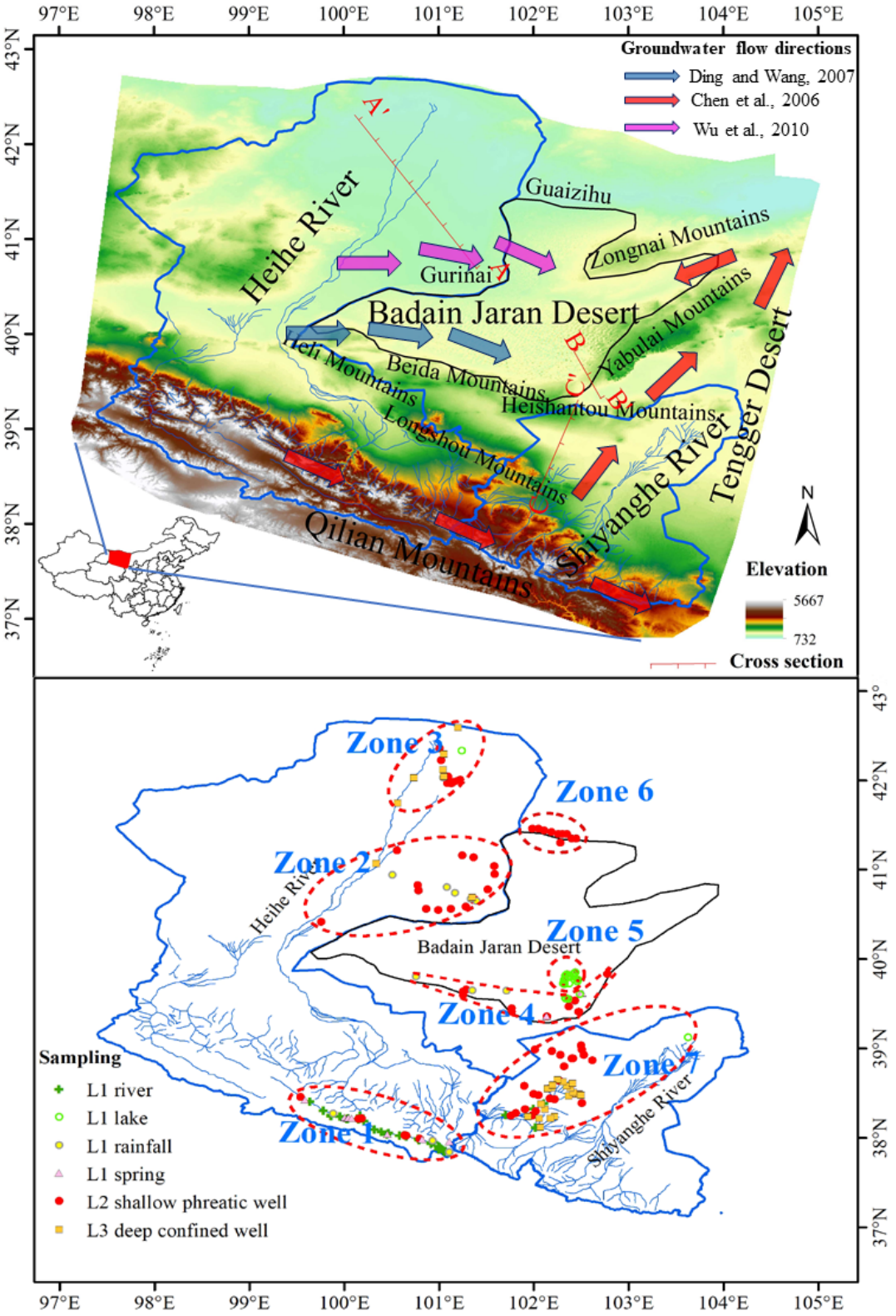


Figure 3- 1. (A) the cross sections, land forms, and groundwater flow directions in the study area ;(B) Sampling sites of different water bodies in each Zone.

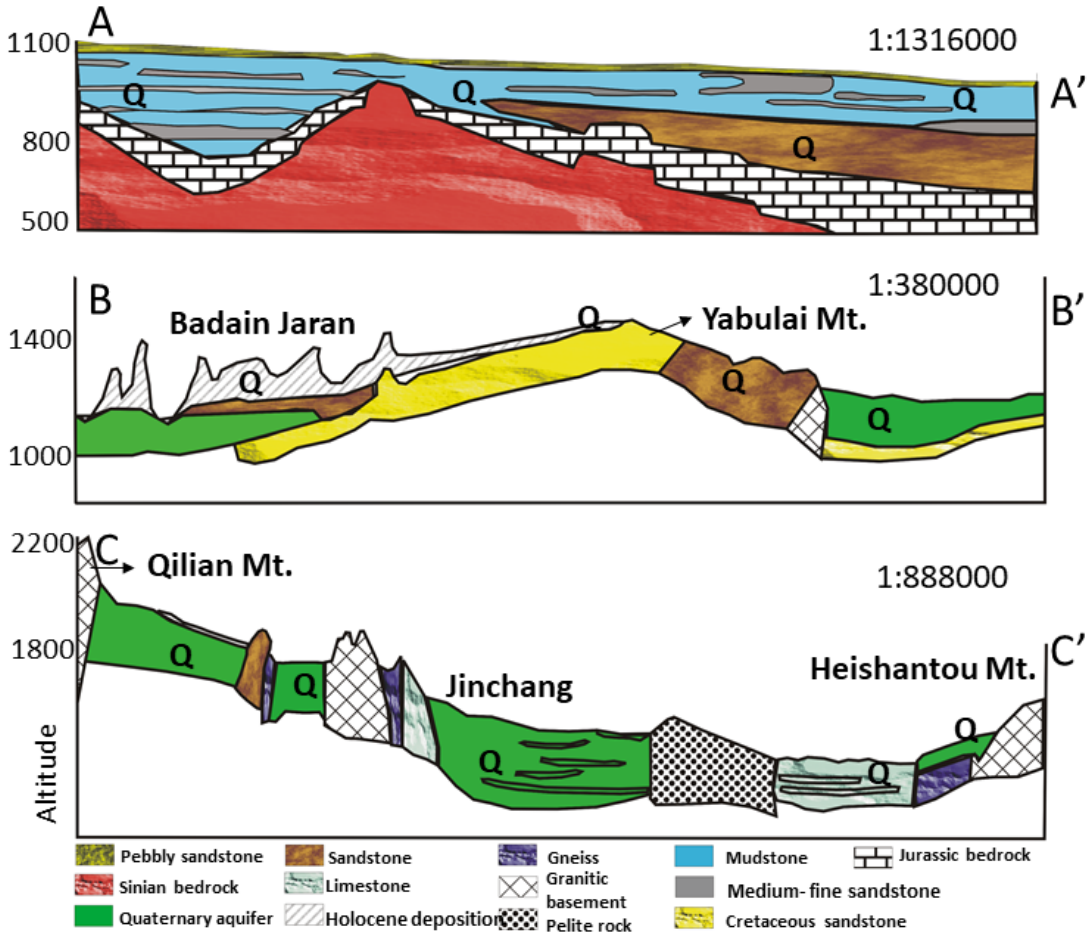


Figure 3- 2. The cross sections in Figure 3-1(A). AA' was adapted from Guo et al., 2019, BB' was adapted from Ma and Edmunds, 2006, CC' was adapted from Ma et al.,2010

3.2.2.1 Field sampling

According to the possible hydraulic connections between the three units (Figure 3-1A), the sampling areas were divided into seven sampling zones (Figure 3-1B). Zone 1(Z1) and Z7 were the upper reaches of the Qilian Mountains, and Z2 was the lower reaches of the HRB, including Gurinai next to the BJD. Z3 was the endorheic lake areas in the lower reaches of the HRB, Zone 4 was the bordering range of the BJD, while Z5 was the hinterland of BJD, and Z6 was Guaizihu. The sampling points were selected, where possible, to cover the typical area of potential hydraulic connections (Figure 3-1B). The final total number and distribution of samples were listed in Table 3-2.

The rainfall samples were obtained from long-term observation stations, or the events that occurred during the field work. In order to avoid collection of any locality-specific pollutants, the precipitation collectors were placed approximately 8 m high from the ground in the

stations. Surface water was collected from a depth of 5 cm below the water surface by submerging pre-cleaned polyethylene sample bottles. Groundwater samples were collected from long-term observation wells, industrial wells, and wells used for domestic purposes. Samples collected from depths of less than 20 m deep were grouped into shallow groundwater, while samples over 100 m were classed as deep groundwater. After collection, all the samples were immediately sealed and stored in a freezer at -18 ^{o}C until they were brought back to the laboratory for analysis. In this study, spring water was classified to L1, because it was difficult to confirm whether the due to spring water is shallow groundwater or deep groundwater.

3.2.2.2 Laboratory analysis

Stable isotopes, δ^{18}O, and δD of the sampled water were determined by isotope ratio mass spectrometry, and the results were reported relative to the VSMOW standard (Vienna Standard Mean Ocean Water) in per mille (d‰). Anions were analyzed using a Dionex-2500 ion chromatograph, while cations with a Dionex-600 ion chromatograph. The analytical precision reached 10-9 g/mL, and the ionic charge balance for all water chemistry was within 5%. In addition to the major ions and isotopes, the pH, total dissolved solids (TDS), and electrical conductivity (EC) were also used as the main indicators in our analysis and measured on the field.

3.2.3 Methods

Firstly, potential appropriate members for the 3.2.2 EMMA had been chosen through 3.1 and 3.2 in step 1, (Figure 3-3). The results of EMMA in 3.2.2 laid an important foundation for the follow-up studies on the water contaminant transportation. The relationships between contaminant tracers and potential explanatory variables were finally analyzed through 3.4, as well as the output maps of spatial distribution of groundwaters.

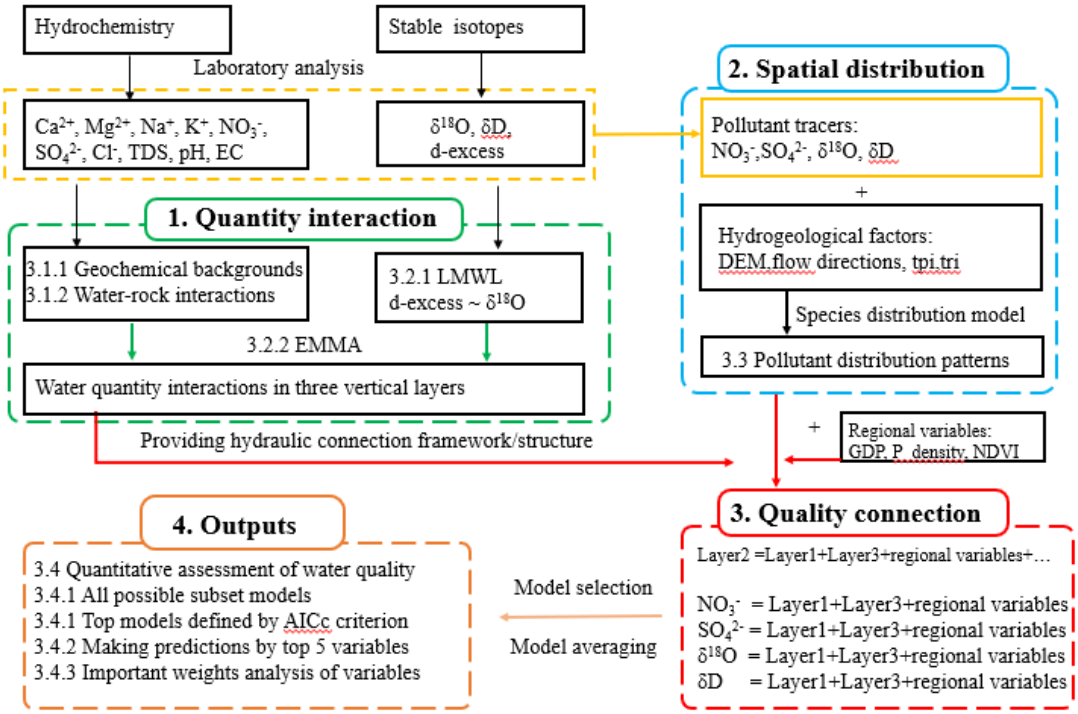


Figure 3- 3. Flowchart to show the specific steps of this study.

3.2.3.1 Hierarchical cluster analysis

The method of Euclidean hierarchical cluster was applied to classify all water bodies according to the chemical compositions, which allows us to have an overall understanding of the properties of the sample.

The hierarchical cluster analysis was performed using a set of dissimilarities for the objects being clustered. Initially, each object is assigned to its own cluster and then the algorithm proceeds iteratively, at each stage joining the two most similar clusters, continuing until there is just a single cluster. At each stage, distances between clusters are recomputed by the Lance–Williams dissimilarity update formula according to the particular clustering method being used.

3.2.3.2 End member mixing analysis

The EMMA method attempts to explain chemical ingredients of a given water as a mixture of a set of potential end members, according to the conservative tracers. The method is described as the following equations (Christophersen et al., 1990):

$$1 = a_1 + a_2 + a_3 \qquad (1)$$

$$C_t^1 = C_1^1 a_1 + C_2^1 a_2 + C_3^1 a_3 \qquad (2)$$

$$C_t^2 = C_1^2 a_1 + C_2^2 a_2 + C_3^2 a_3 \qquad (3)$$

where a_1, a_2, and a_3 are the discharge fractions of end members, C_t^1 and C_t^2 are the tracer concentrations of mixed groundwater. C_n^1 and C_n^2 are the tracer concentrations of the *n*th end member. In this study, the C_t^1 and C_t^2 are the tracer concentrations of shallow groundwater in the BJD, which was assumed to be recharged from different water types in the three hydrological units.

3.2.3.3 Species distribution models

SDM can map the past, current, and future spatial distributions of species, and is an innovative GIS-based method that combines species presence-only data with environmental variables to better map habitat. The research entity has been extended from species into other research fields, such as the visitors, recreation activity, or even the human emotions. The SDM was introduced in this study to aim to map the geochemical trace distribution in the groundwater. The geological and hydrological factors were considered into the SDM as the environmental variables.

3.3 Results

3.3.1 Hydrochemistry characteristics

3.3.1.1 Hierarchical clustering of waters

The lake waters of Layer 1 Zone 4 (L1Z4) and L1Z5 were separately classified into one type (Figure 3-4). These two types of lake water were located in the BJD area, which tended to be dominated by the anions Cl^-, SO_4^{2-} , and cations Na^+, K^+, with concentrations being 11680.35, 10192.73, 31696.65, and 1381.57 mg/L in L1Z4, and 36114.06, 8689.72, 46807.48 and 4163.42 mg/L in L1Z5, respectively (Figure 3-4; Table 3-2). The alkaline character of these two lakes was also shown by the pH values that were all above 9 (average values in Table 3-2). In addition, the abundances of $\delta^{18}O$, δD in these lakes were relatively heavier, indicating the strong evaporation effect. The chemical compositions of rainfall in L1Z2 were similar to those of lake in L1Z7, indicating the recharge from local precipitation to the lake in Z7 (Figure 3-4).

Table 3- 2. Average parameters for the ionic and isotopic compositions of different water samples in the seven zones of the study area.

Type	Zones	n	Ca	Mg	Na	K	NO3	SO4	Cl	18O	2H	EC	PH
Layer 1													
river	Zone 1	48	77.07	34.25	26.49	2.87	13.40	115.13	11.18	-8.37	-50.11	307.52	8.10
	Zone 7	4	20.29	74.93	23.69	2.13	4.67	138.94	14.46	-8.77	-54.53	NA	7.99
spring	Zone 1	5	85.75	46.37	41.85	3.64	19.04	126.23	27.55	-8.01	-47.49	333.27	7.78
	Zone 5	6	66.82	35.42	136.48	17.55	40.69	168.48	122.03	-2.86	-44.43	651.67	7.70
rainfall	Zone 1	4	26.58	1.08	3.65	1.85	5.46	21.02	4.21	-6.11	-39.43	111.23	8.35
	Zone 2	5	407.67	358.92	513.18	312.61	83.65	554.38	286.18	-1.78	-35.26	53.64	8.34
	Zone 3	1	69.77	6.66	39.00	1.86	15.73	74.85	33.45	0.70	-14.67	246.00	8.24
	Zone 4	3	27.74	34.38	1147.26	65.97	10.38	329.97	538.63	-5.15	-31.26	344.33	8.58
lake	Zone 3	1	67.97	652.89	1496.95	77.01	42.74	2154.61	927.89	-12.12	-79.54	7.93	8.50
	Zone 4	6	17.29	436.14	31696.65	1381.57	35.84	10192.73	11680.35	6.20	0.83	12.80	9.16
	Zone 5	22	26.61	83.76	46807.48	4163.42	85.91	8689.72	36114.06	1.70	-21.20	154.05	9.28
	Zone 7	1	769.95	629.38	1899.15	176.70	42.63	4089.51	832.54	-8.89	-59.70	10.49	7.62
Layer 2	shallow groundwater												
well	Zone 1	22	102.53	57.12	34.11	4.00	45.09	221.16	30.26	-8.00	-48.86	228.81	7.61
	Zone 2	19	46.71	39.84	333.48	19.99	46.14	326.23	264.37	-3.24	-50.06	1312.45	7.81
	Zone 3	11	108.76	116.91	289.21	8.41	28.63	393.90	128.67	-7.03	-54.09	3.73	7.56
	Zone 4	12	131.16	61.28	550.73	9.13	33.48	410.63	383.70	-7.49	-54.45	732.12	8.18
	Zone 5	10	51.07	27.80	223.77	10.53	55.03	226.09	163.97	-3.78	-47.39	710.40	8.10
	Zone 6	10	62.22	34.00	284.11	15.22	NA	296.82	332.60	-2.99	-55.18	1854.93	7.46
	Zone 7	23	37.01	103.00	240.21	9.56	39.67	362.78	219.60	-8.83	-67.30	NA	7.35
Layer 3	deep groundwater												
well	Zone 2	2	34.08	35.45	128.84	10.12	NA	93.82	169.32	-5.67	-51.82	888.35	7.67
	Zone 3	6	46.45	20.55	268.25	3.62	NA	327.42	220.69	-8.04	-68.57	1601.75	8.17
	Zone 7	20	32.98	92.48	68.46	4.47	NA	213.89	68.42	-8.51	-62.02	NA	7.43

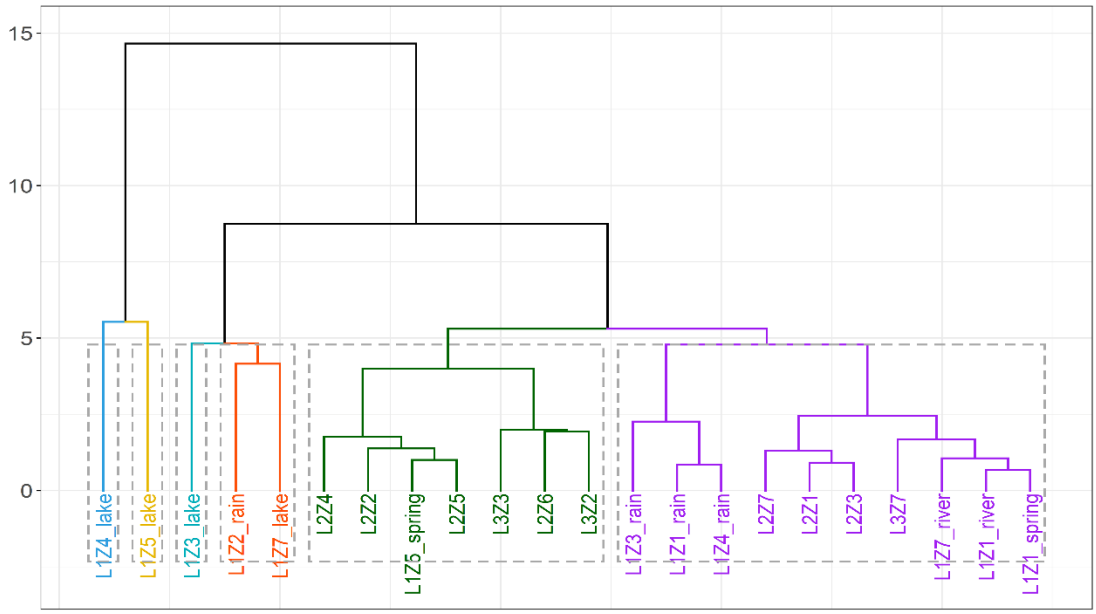


Figure 3- 4. Dendrogram of all water types in the study area. L1Z4_lake stands for the lake water of Zone 4 in Layer 1.

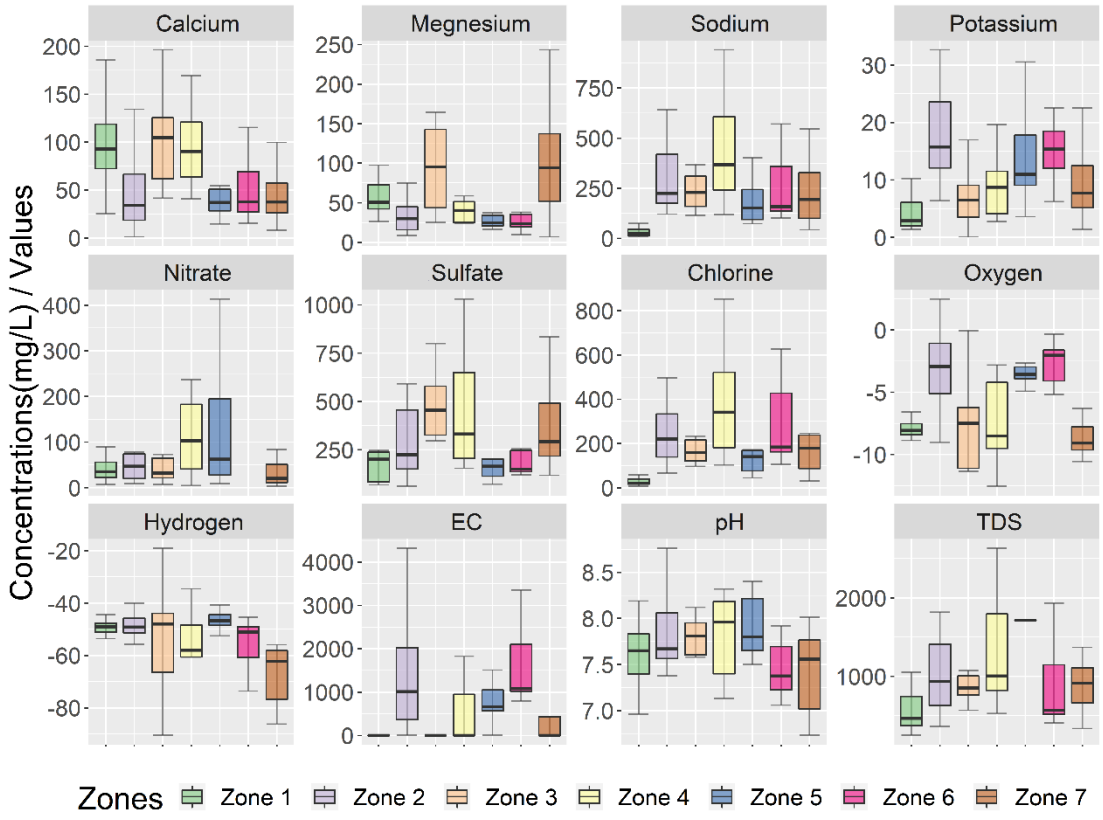


Figure 3- 5. Geochemical signature in shallow groundwater (Layer 2).

In terms of shallow groundwater in L2, the contents in L2Z2, L2Z4, and L2Z5 were in the same branch, indicating similar chemical characteristics. The pH, Na^{+}, SO_4^{2-} , and Oxygen of these three areas were significantly higher than other zones in L2 (Figure 3-4, Figure 3-5). The shallow groundwater in L2Z6, and the deep groundwater in L3Z2 and L3Z3 were located in a branch (Figure 3-4, Table 3-2), indicating the potential hydraulic connections between them. In addition, the groundwaters of L2Z1, L2Z3, and L2Z7 were in one branch (Figure 3-4), with higher concentrations of Ca^{2+} and Mg^{2+}, but lower in NO_3^{-}, Cl^{-}, TDS, and Oxygen (Figure 3-5).

The deep groundwater in L3Z7 was located in the same branch with river water and spring water in L1Z1(Figure 3-4), indicating the potential groundwater hydraulic connections (Figure 3-1A).

3.3.1.2 Water-rock interaction

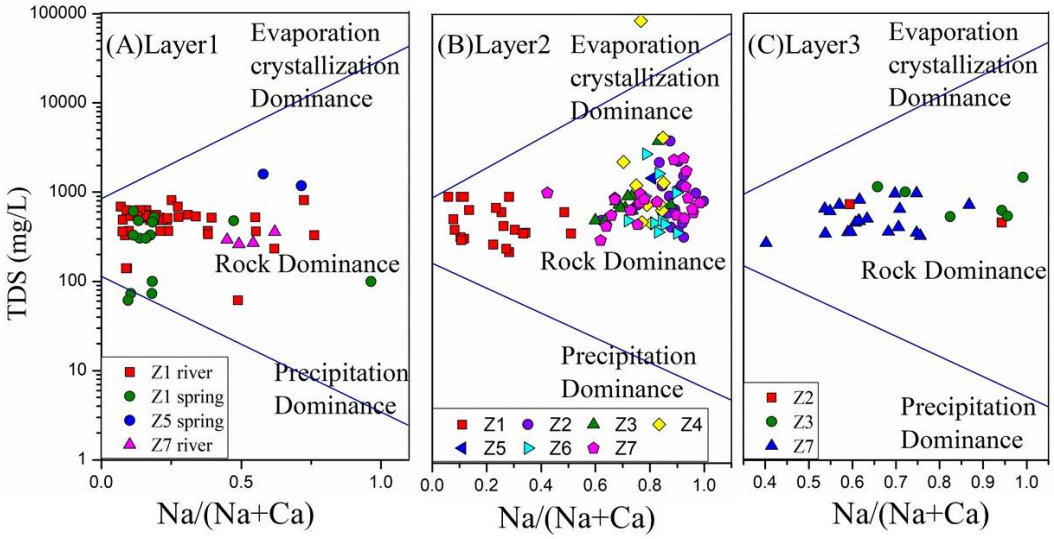


Figure 3- 6. Gibbs plots showing major processes controlling the water chemistry in the three layers.

Gibbs plot is a solute transport model that can calculate the groundwater salinity and changes in the concentration (Marandi and Shand, 2018). Except for some of the groundwater samples in L2Z4 that were dominated by evaporation crystallization (Figure 3-6B), the chemical evolution of the rest samples was mainly controlled by the rock dominance process (Figure 3-6), and this result was consistent with previous studies.

3.3.2 Isotopic signature

3.3.2.1 LMWL information

Local meteoric water lines (LMWLs) represent the long-term covariation of the average δD-δ^{18}O ratios at a specific site (Hughes and Crawford, 2012). In L1, the river water samples in L1Z1 and L1Z7, and spring water in L1Z2 were relatively concentrated, indicating the uniform recharge sources, while the distribution of other waters in L1 was scattered (Figure 3-7A), indicating complicated recharge sources. The rain water in L1Z1 and L1Z4 were located around the LMWL, while the rain water in L1Z2 and L1Z3 were not around the line. The abundance of isotopes of lake water from L1Z4 and L1Z5 in the summer period tended to be heavier than in other zones, in δD (-18 to +3.7‰) and δ^{18}O (4.6 to +7.5‰) due to the continual evaporation from the lake surface. The isotopic contents of lake water in L1Z5 were relatively depleted in winter. The lake water samples of L1Z3 were located in the lower left corner, extremely depleted, indicating an ancient water source, or with mixed supply sources with ancient water components.

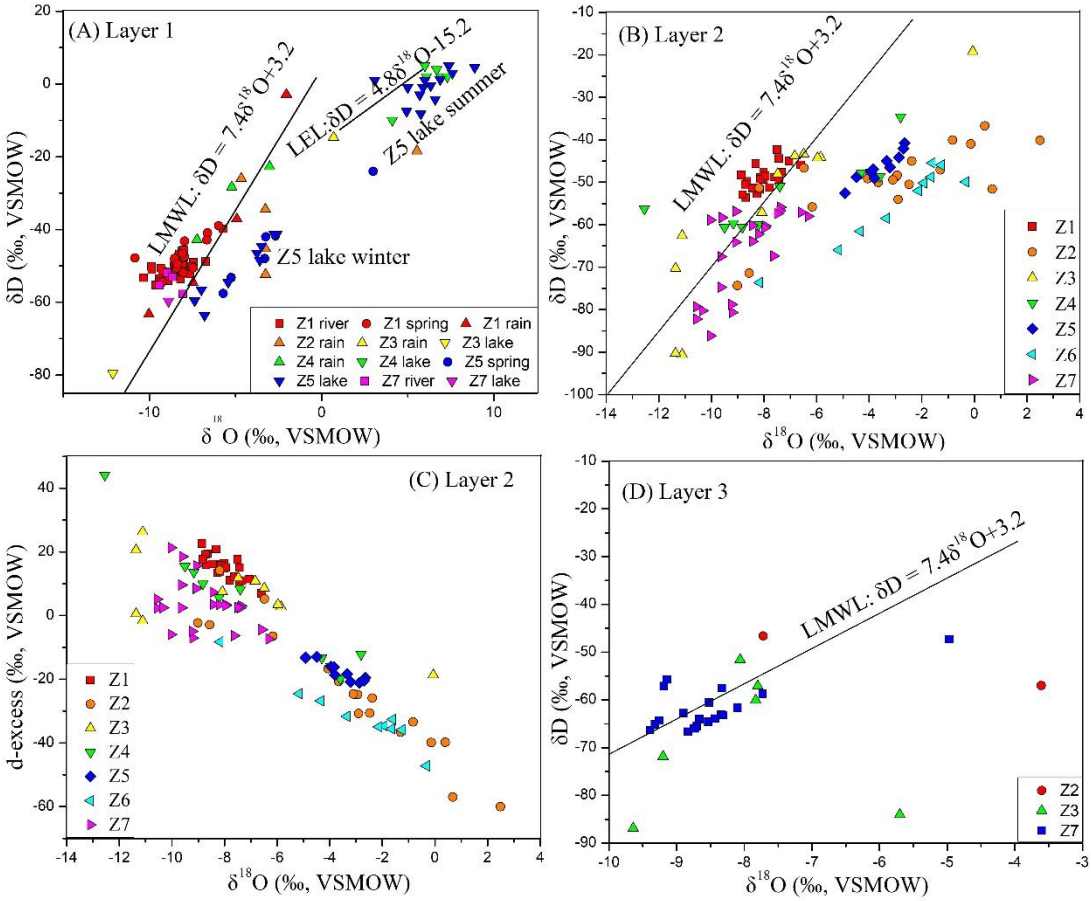


Figure 3-7. $\delta^{18}O$, δD, and d-excess relationship diagram of the surface and ground waters.

The distributions of shallow groundwater in L2Z1, L2Z5, and L2Z7 were relatively gathered (Figure 3-7B, 3-7C), showing the fixed supply sources, while other groundwater samples in L2 were scattered, with complicated recharge sources. The shallow groundwater samples in L2Z1 and L2Z7 were located around the LMWL, indicating the contributions of local meteoric recharge. The isotopic contents of samples in L2Z3 and L2Z7 were extremely depleted, indicating ancient water sources.

In L3, the aggregated distributions of the deep groundwater samples in L3Z7 (Figure 3-7D), indicated uniform recharge sources, while other samples in L3 were scattered. Similarly, the potential contributions of local precipitation to deep groundwater in L3Z7 were shown by the distribution patterns around the LMWL, and the ancient water recharged sources for groundwater samples of L3Z3 were indicated by their extremely depleted isotopic abundance.

3.3.2.2 Water quantity connections

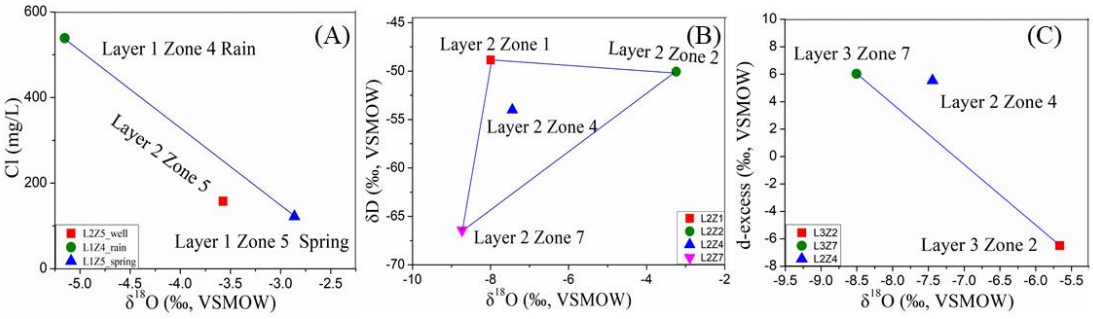


Figure 3- 8. Dual trace diagram of $\delta^{18}O$ and other environmental tracers for shallow groundwater in BJD with mean values of each end member.

Table 3- 3. Final tracers for end member mixing analysis (EMMA), with corresponding discharge fractions in the study area.

Shallow groundwater	Recharge components	EM quantity	EM	$\delta^{18}O$	Cl^-	δD	d-excess	Fraction (%)
zone 5	layer 1	2	rain_Z4	-5.15	538.63	-	-	8.57
			spring_Z5	-2.86	122.03	-	-	91.43
zone 4	layer 2	3	Z1	-7.99	-	-48.86	-	55.86
			Z2	-3.24	-	-50.06	-	16
			Z7	-8.73	-	-66.42	-	28.14
zone 4	layer 3	2	Z2	-5.67	-	-	-6.5	3.9
			Z7	-8.51	-	-	6.03	96.1

These EMMA can be used to identify sources of groundwater and estimate the percentage contributions of each member. We took shallow groundwater L2 as the objects to be supplied. In L1, the rain water in L1Z4 and spring water in L1Z5 were selected as the final two end members, with the concentration of Cl^- and abundance of $\delta^{18}O$ as the tracers (Figure 3-8A). The discharge fractions of rain and spring from surface water were 8.57% and 91.43%, respectively (Table 3-3).

In L2, the abundances of δD and $\delta^{18}O$ were used as the tracers, and the shallow groundwater in L2Z1, L2Z2, and L2Z7 were selected as the end members (Figure 3-8B), with the contributions being 55.86%, 16% and 28.14%, respectively (Table 3-3).

In L3, the deep groundwater in L3Z2 and L3Z7 were selected as the final end members, and the d-excess and $\delta^{18}O$ were used as the tracers (Figure 3-8C), with fractions being 3.9% and 96.1%, respectively. This result was consistent with Chen et al. (2004).

3.3.3 Spatial distribution of chemical tracers in shallow groundwater

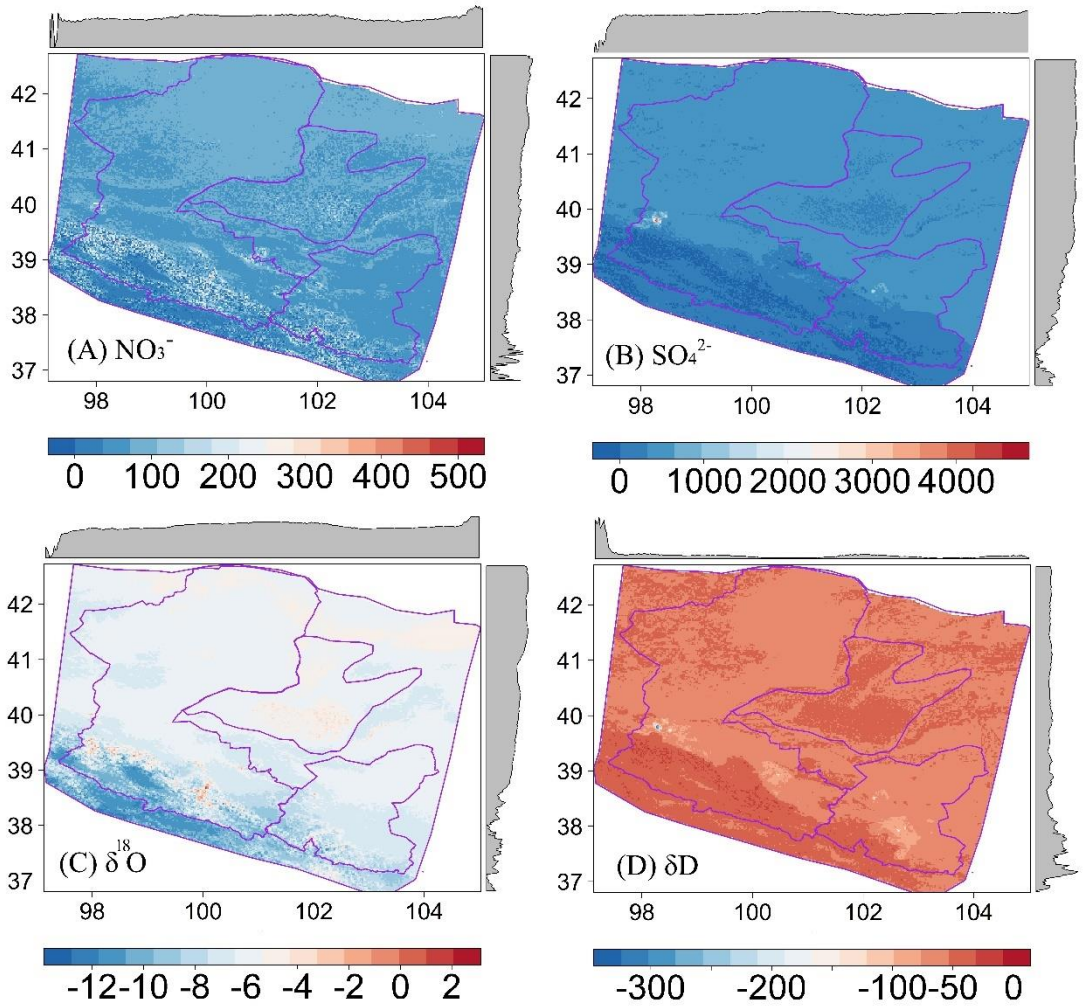


Figure 3- 9. Predicted spatial patterns of NO_3^-, SO_4^{2-}, $\delta^{18}O$, and δD in shallow groundwater (Layer 2) by species distribution model (SDM).

The spatial distribution of NO_3^-, $\delta^{18}O$, and δD of the shallow groundwater samples showed high values or enriched abundance in the upper reaches of the HRB and SYH, as well as the central part of BJD (Figure 3-9). The peak values of SO_4^{2-} were in the western part of the middle reaches of HRB, where located the typical industrial cities of the Jiuquan and Jiayuguan (Figure 3-9B). The abundance of δD tended to be lighter in the Jiuquan and Jiayuguan area, showing an extremely depleted trend (Figure 3-9D).

3.3.4 Shallow groundwater contaminants with explanatory variables

3.3.4.1 Top models for the contaminant tracers

Table 3- 4. Top models identified through model selection based on AICc. df= Degrees of freedom. Weight= Akaike weight.

indicators	smoothed variables	df	loglik	AICc	weight
Nitrates.L2	0.87 +(0.17)GDP+(-0.07)Mg.L1+(0.13)ndvi+(-0.21)pdensity	6	-135.43	283.63	0.06
Sulfates.L2	0.67+(0.12)Ca.L1+(-0.1)EC.L1+(0.18)GDP+(-0.15)Mg.L1+(-0.1)pH.L1+(0.24)Sulfates.L3	8	-69.26	155.86	0.15
Oxygen.L2	-0.54 +(0.19)Cl.L1+(-0.65)GDP+(-0.76)ndvi+(0.33)Oxygen.L1+(0.83)pdensity	7	-129.29	273.62	0.21
Hydrogen.L2	0.4+(-0.08)Ca.L1+(-0.24)Hydrogen.L3+(-0.43)pdensity+(0.14)TDS.L1	6	-60.34	133.44	0.47

The top models for the four contaminant tracers were identified by model selection from all possible combinations of explanatory variables. Top models with lowest AICc values are presented in Table 3-4. The variables in the top model of NO_3^- included the GDP, Mg^{2+} in surface water, NDVI, and population density, with the AICc value being 283.63, and the degrees of freedom (df) being 6 (Table 3-4).

When Ca^{2+}, EC, Mg^{2+}, pH in surface water, and SO_4^{2-} in deep groundwater, as well as when the regional variable GDP were selected as the smoothed variables, the AICc value of the top model for Sulfates in shallow groundwater was 155.86, and the df was 8 (Table 3-4). The $\delta^{18}O$ in top model could be explained by Cl^- and $\delta^{18}O$ in surface water, and the regional variables of GDP, NDVI, and the population density.

3.3.4.2 Relationships with final variables

Table 3- 5. Sum of model weights over all models including each explanatory variable.

Nitrates		s(ndvi)	s(Mg.L1)	s(Nitrates.L1)	s(pdensity)	s(Na.L1)	s(GDP)	s(EC.L1)	s(pH.L1)	s(Ca.L1)
	weights:	1	0.79	0.62	0.57	0.16	0.15	0.13	0.08	0.04
	N	19	14	11	11	4	3	2	2	1
Sulfates		s(EC.L1)	s(Mg.L1)	s(ndvi)	s(pdensity)	s(Sulfates.L3)	s(TDS.L1)	s(pH.L1)		
	weights:	1	1	1	1	1	0.48	0.32		
	N	3	3	3	3	3	1	1		
Oxygen		s(EC.L1)	s(ndvi)	s(Oxygen.L3)	s(pdensity)	s(Mg.L1)	s(GDP)	s(Oxygen.L1)		
	weights:	1	1	1	1	0.48	0.21	0.21		
	N	3	3	3	3	1	1	1		
Hydrogen		s(EC.L1)	s(GDP)	s(ndvi)	s(pdensity)	scale(Hydrogen.L1)	s(Na.L1)			
	weights:	1	1	1	1	1	0.32			
	N	2	2	2	2	2	1			

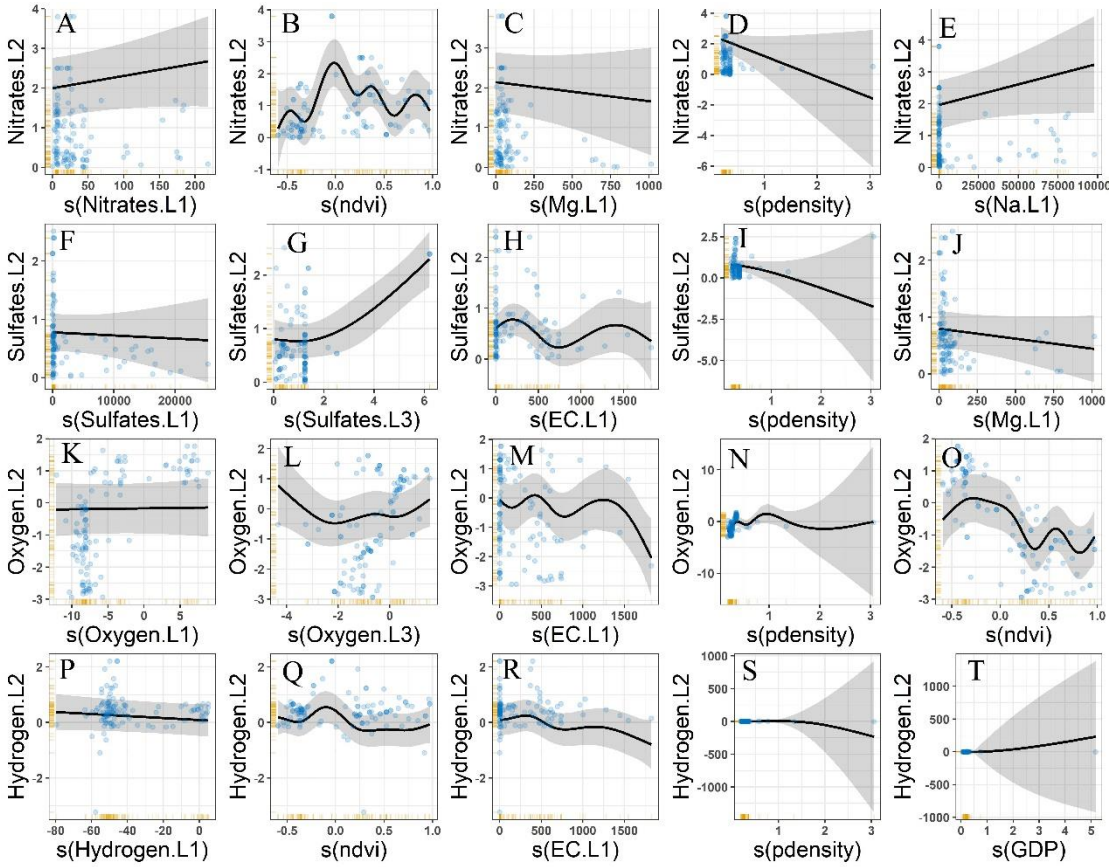


Figure 3- 10. Relationships between NO_3^-, SO_4^{2-}, $\delta^{18}O$, and δD and explanatory variables.

According to the research objectives and the importance analysis of the variables (Table 3-5), the final five explanatory factors were chosen as the final variables to explain the four contaminant indicators. Their relationships were analyzed by the gam and results were presented in Figure 3-10.

In terms of the influence of related tracers from different layers, the nitrates in surface water L1 showed a positive trend with nitrates in shallow groundwater L2. While the sulfates in L2 were slightly negative with sulfates in L1, they showed a roughly positive trend with sulfates in L3 when the concentrations of sulfates in L3 exceeding 1.5 mg/L. The abundance of Oxygen in L2 and L1 was basically the same, but the relationship with Oxygen in layer3 is complicated; while the Hydrogen in L2 showed a slightly negative relation with those in L1.

Because of the influence from human activity, all the four contaminant tracers showed a negative trend with the increasing of population density (Figure 3-10D, I, N, S); while Hydrogen showed a positive trend with GDP (Figure 3-10T). The relationship curves between tracers and NDVI were complex, and roughly reached the maximum value when NDVI around 0 (Figure 3-10B, O, Q).

In terms of the influence from soluble ions, the Mg^{2+} in surface water L1 was negatively related to both nitrates and sulfates in L2 (Figure 3-10C, J); while the Na^{+} showed a positive trend with nitrates (Figure 3-10E).

3.4 Discussion

3.4.1Combination of multiple environmental tracers

Combining the environmental tracers, like the major soluble ions, CFCs (chlorofluorocarbons) concentrations, $\delta^{18}O$, and δD, can help to gain a better insight of the hydraulic connections. For example, the $\delta^{18}O$, and δD are not significantly affected by geochemical reactions at normal shallow groundwater temperatures, and more conservative than the soluble ions (Huang et al., 2015). On the other hand, the soluble ions can yield valuable information about the invisible flow paths. In this study, conventional water-soluble ions and $\delta^{18}O$, and δD were combined to help elucidate the sources of shallow groundwater recharge and hydrogeochemical evolution processes.

SO_4^{2-} has been used as a proxy for marine biological activity in coastal regions (Curran et al., 2003; Sneed et al., 2011), but its sources might be from both natural environment and human activity. Only by analyzing the correlations between sulfates and other indicators, such as NO_3^-, can the sources of sulfates be confirmed (Li et al., 2020). NO_3^- is the main form of nitrogen in drinking water reservoirs. The sources of NO_3^-, the spatial and temporal varieties (Jin et al., 2019; Minet et al., 2017), as well as the transformation and transport pathways(Hu et al., 2019), could be quantified by some statistical methods. In this study, the SO_4^{2-} and NO_3^-were used as contaminant tracers to study the contaminant transportation between various hydrological units. Chloride is regarded as a conservative reference element in the shallow hydrological cycle because it is derived from atmospheric deposition (Cartwright et al., 2011). In this study, the concentration of Cl^- was used as the end member tracer to identify the contributions of surface water in the EMMA model.

3.4.2 SDM for groundwater prediction

The major difficulties to predicting the spatial distribution of groundwater are how to represent the various landscapes, complex geological structure, and multiple hydrological factors (Liu et al., 2019). For example, a geological structure usually contains geomorphological boundaries, elevations of aquifers, and geological frameworks (Kahle and Bartolino, 2007). Additionally, the knowledge about the hydrological system, and dynamic

information should be important index to evaluate groundwater models (Hill et al., 2013), such as the groundwater flow directions (Wang and Zhou, 2018)

The biggest advantage of SDM is that it can calculate many explanatory variables together, and it is good at organizing geographical datasest with various scales and resolutions into a unified framework. In this study, the SDM was introduced to predict the distribution of HMs in groundwater with multiple explanatory variables. The DEM dataset with a 1km resolution and hydrogeological cross section maps were used as the geological background, and indicators such as slope, aspect, terrain ruggedness index, topographic position index, etc., were all taken into consideration. The flow direction was also introduced as a hydrological factor for the model calculation. According to the output maps, the spatial distributions of all these four tracers can reflect the influence and restrictions of the geological background, which has improved the accuracy in the study area compared with previous studies (Guo et al., 2015; Yang et al., 2011; Zhao et al., 2012). In addition, the highest values of sulfates, and the depleted points of δD coincided with the location of the industrial cities, which also proves the rationality of our prediction results.

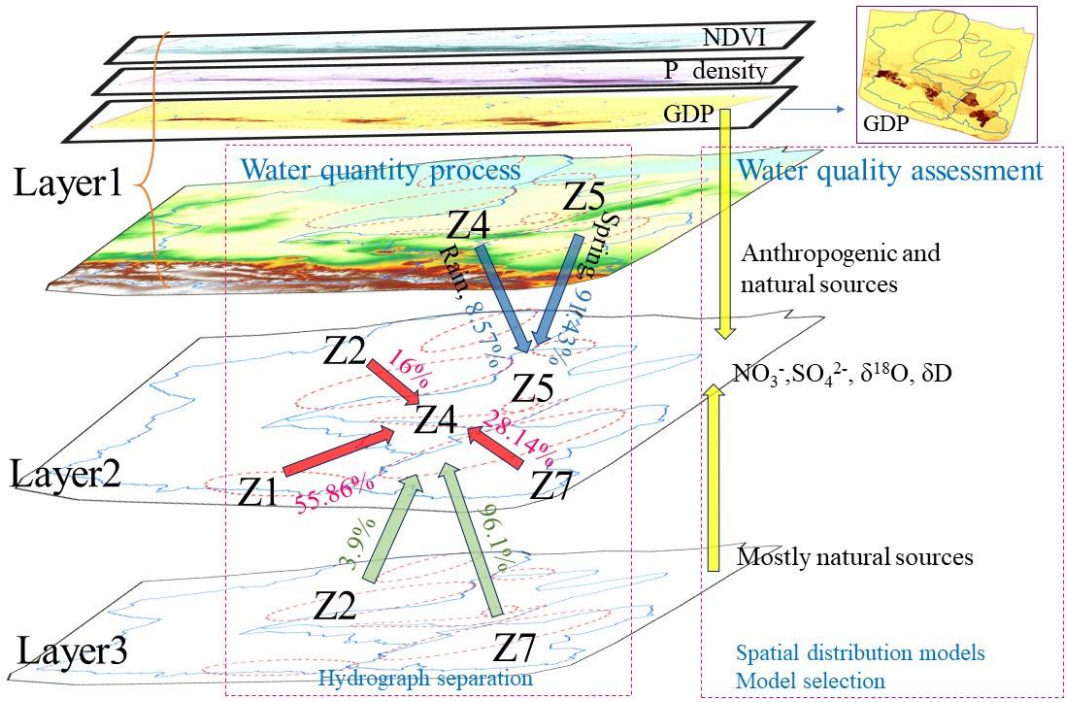


Figure 3- 11. Conceptual model of mechanisms controlling the water quantity and quality of Layer 2 and schema of the flow directions in the study area.

3.4.3 Possibility of EMMA for water quality analysis

The water quantity interaction between different hydrological units is a prerequisite for the water contaminant transportation. The conceptual model structure of interactions between surface water and groundwater by EMMA can be used for water quality analysis (Figure 3-11). Some previous studies had tried to explore the water quantity interactions in this region (Guo et al., 2019; Wang et al., 2013), but did not involve the water quality issues, and the scope of these studies only involved sub-basin scale.

A series of environmental degradation phenomena in arid regions, such as desertification, sand storms, vegetation degeneration, and soil salinization are all related to groundwater degradation (Lapworth et al., 2013). The main triggers for the over- exploitation and pollutions of groundwater are the anthropogenic influence. These human activities can usually be passed on by GDP and population density, especially in arid and semi-arid regions (Mukherjee et al., 2020). The promotion of GIS products makes it easy to introduce population density, GDP, distance to road, nightlight, soil erosion, NDVI, DEM, land cover, precipitation, temperature, and other information for modelling (Figure 3-11). In addition, this study created spatial distribution layers through SDM (Figure 3-9), which provided more information to further reveal the mechanism of water pollution.

3.4.4 Recharge sources of the BJD

The water sources of the BJD have always been a hot issue in historical research. This study did not focus on the water source, but instead used some hypotheses of potential groundwater flow directions as the basis for our sampling and research. Through research on hydraulic connections in this area, our results were consistent with some previous research. For example, we found that both the shallow groundwater and deep groundwater in the Qilian Mountains can potentially recharge the groundwater in the BJD area. This conclusion was similar with the conclusion of Chen et al. (2004) that proposed that there was rapid transfer of snow melt from the Qilian Mountains through fracture networks to the BJD. Although Chen's hypothesis has been rejected by many previous studies using different methods (Ma and Edmunds, 2006; Wu et al., 2017), we reached similar conclusions with Chen through water chemistry and isotopes.

We also found that Gurinai in the middle reaches of the HRB had a strong hydraulic connection with BJD groundwater, which was consistent with the conclusions of Gates et al. (2008) and Wang et al. (2013).

3.5 Conclusions

The unclear situation of aquifers and invisible groundwater flow directions are a challengeable obstacle restricting the study of hydraulic connections in arid and semi-arid regions. Surface water, shallow groundwater and deep groundwater were systematically sampled along the potential groundwater flow directions from BJD, HRB, and SYH. Then the EMMA and SDM were introduced to explore the hydraulic connections. Results found that the chemical evolution of most water types was mainly controlled by rock dominance. Taking the shallow groundwater in the BJD as the recharged object, the contribution of surface water was 8.57% from local precipitation, and 91.43% from regional spring water; while the contribution of deep groundwater was 3.9% from the middle reaches of the HRB, and 96.1% from the upper reaches of the SYH. The highest values of the spatial distribution of Sulfates coincided with the locations of typical industrial cities. In addition to sulfates replenishment from surface water and deep groundwater, the sulfates in shallow groundwater were also affected by regional variables such as population density and the Mg and EC contents in surface water. The EMMA and SDM models were introduced in this study to help improve the accuracy of predicting the spatial distributions of groundwater, which provides new ideas for revealing the hydraulic connections, especially water pollution between different hydrological units in arid areas.

3.6 References

Banks, E., Simmons, C., Love, A., Shand, P., 2011. Assessing spatial and temporal connectivity between surface water and groundwater in a regional catchment: Implications for regional scale water quantity and quality. Journal of Hydrology, 404(1-2): 30-49.

Cartwright, I., Hofmann, H., Sirianos, M.A., Weaver, T.R., Simmons, C.T., 2011. Geochemical and 222Rn constraints on baseflow to the Murray River, Australia, and timescales for the decay of low-salinity groundwater lenses. Journal of Hydrology, 405(3-4): 333-343.

Chen, J. et al., 2006. Formation mechanisms of megadunes and lakes in the Badain Jaran Desert, Inner Mongolia. Chinese Science Bulletin, 51(24): 3026-3034.

Chen, J.S. et al., 2004. Groundwater maintains dune landscape. Nature, 432(7016): 459-460.

Christophersen, N., Neal, C., Hooper, R.P., Vogt, R.D., Andersen, S., 1990. Modelling streamwater chemistry as a mixture of soilwater end-members—a step towards second-generation acidification models. Journal of Hydrology, 116(1-4): 307-320.

Cook, E.R. et al., 2013. Five centuries of Upper Indus River flow from tree rings. Journal of hydrology, 486: 365-375.

Curran, M.A., van Ommen, T.D., Morgan, V.I., Phillips, K.L., Palmer, A.S., 2003. Ice core evidence for Antarctic sea ice decline since the 1950s. Science, 302(5648): 1203-1206.

D'Odorico, P. et al., 2019. Global virtual water trade and the hydrological cycle: patterns, drivers, and socio-environmental impacts. Environmental Research Letters, 14(5): 053001.

Ding, H., Wang, G., 2007. Study on the formation mechanism of the lakes in the Badain Fijaran Desert. Arid Zone Res, 24(1): 1-7.

Freyberg, J.v., Studer, B., Kirchner, J.W., 2017. A lab in the field: high-frequency analysis of water quality and stable isotopes in stream water and precipitation. Hydrology and Earth System Sciences, 21(3): 1721-1739.

Gates, J.B. et al., 2008. Conceptual model of recharge to southeastern Badain Jaran Desert groundwater and lakes from environmental tracers. Applied Geochemistry, 23(12): 3519-3534.

Guo, X. et al., 2015. Stable isotopic and geochemical identification of groundwater evolution and recharge sources in the arid Shule River Basin of Northwestern China. Hydrological processes, 29(22): 4703-4718.

Guo, X. et al., 2019. Partitioning groundwater recharge sources in multiple aquifers system within a desert oasis environment: Implications for water resources management in endorheic basins. Journal of Hydrology, 579: 124212.

Hatzinger, P.B., Böhlke, J.K., Sturchio, N.C., 2013. Application of stable isotope ratio analysis for biodegradation monitoring in groundwater. Current opinion in biotechnology, 24(3): 542-549.

Hill, M.C. et al., 2013. Knowledge, transparency, and refutability in groundwater models, an example from the Death Valley regional groundwater flow system. Physics and Chemistry of the Earth, Parts A/B/C, 64: 105-116.

Hodson, A., Tranter, M., Gurnell, A., Clark, M., Hagen, J.O., 2002. The hydrochemistry of Bayelva, a high Arctic proglacial stream in Svalbard. Journal of Hydrology, 257(1-4): 91-114.

Hoshina, Y., Fujita, K., Iizuka, Y., Motoyama, H., 2016. Inconsistent relationships between major ions and water stable isotopes in Antarctic snow under different accumulation environments. Polar Science, 10(1): 1-10.

Hoshina, Y. et al., 2014. Effect of accumulation rate on water stable isotopes of near-surface snow in inland Antarctica. Journal of Geophysical Research: Atmospheres, 119(1): 274-283.

Hu, M., Liu, Y., Zhang, Y., Dahlgren, R.A., Chen, D., 2019. Coupling stable isotopes and water chemistry to assess the role of hydrological and biogeochemical processes on riverine nitrogen sources. Water research, 150: 418-430.

Huang, M., Hilderman, J.N., Barbour, L., 2015. Transport of stable isotopes of water and sulphate within reclaimed oil sands saline–sodic mine overburden. Journal of Hydrology, 529: 1550-1561.

Hughes, C.E., Crawford, J., 2012. A new precipitation weighted method for determining the meteoric water line for hydrological applications demonstrated using Australian and global GNIP data. Journal of Hydrology, 464: 344-351.

Jin, Z. et al., 2019. Quantifying nitrate sources in a large reservoir for drinking water by using stable isotopes and a Bayesian isotope mixing model. Environmental Science and Pollution Research, 26(20): 20364-20376.

Kahle, S.C., Bartolino, J.R., 2007. Hydrogeologic Framework and Ground-Water Budget of the Spokane Valley-Rathdrum Prairie Aquifer, Spokane County, Washington, and Bonner and Kootenai Counties, Idaho. US Geological Survey.

Khalil, M.M., Tokunaga, T., Yousef, A.F.J.J.o.H., 2015. Insights from stable isotopes and hydrochemistry to the Quaternary groundwater system, south of the Ismailia canal, Egypt. 527: 555-564.

Klaus, J., McDonnell, J., 2013. Hydrograph separation using stable isotopes: Review and evaluation. Journal of hydrology, 505: 47-64.

Kwon, H.-J., Kim, S.-J., 2010. Assessment of distributed hydrological drought based on hydrological unit map using SWSI drought index in South Korea. KSCE Journal of Civil Engineering, 14(6): 923-929.

Lapworth, D. et al., 2013. Residence times of shallow groundwater in West Africa: implications for hydrogeology and resilience to future changes in climate. Hydrogeology Journal, 21(3): 673-686.

Lehr, C., Pöschke, F., Lewandowski, J., Lischeid, G., 2015. A novel method to evaluate the effect of a stream restoration on the spatial pattern of hydraulic connection of stream and groundwater. Journal of Hydrology, 527: 394-401.

Li, J. et al., 2020. Tracing geochemical pollutants in stream water and soil from mining activity in an alpine catchment. Chemosphere, 242: 125167.

Liu, F. et al., 2015. Identifying the origin and geochemical evolution of groundwater using hydrochemistry and stable isotopes in the Subei Lake basin, Ordos energy base, Northwestern China. Hydrology and Earth System Sciences, 19(1): 551-565.

Liu, F. et al., 2019. Coupling hydrochemistry and stable isotopes to identify the major factors affecting groundwater geochemical evolution in the Heilongdong Spring Basin, North China. Journal of Geochemical Exploration, 205: 106352.

Ma, J., Edmunds, W.M., 2006. Groundwater and lake evolution in the Badain Jaran Desert ecosystem, Inner Mongolia. Hydrogeology Journal, 14(7): 1231-1243.

Marandi, A., Shand, P., 2018. Groundwater chemistry and the Gibbs Diagram. Applied Geochemistry, 97: 209-212.

Maurya, P., Kumari, R., Mukherjee, S., 2019. Hydrochemistry in integration with stable isotopes (δ18O and δD) to assess seawater intrusion in coastal aquifers of Kachchh district, Gujarat, India. Journal of Geochemical Exploration, 196: 42-56.

Minet, E.P. et al., 2017. Combining stable isotopes with contamination indicators: a method for improved investigation of nitrate sources and dynamics in aquifers with mixed nitrogen inputs. Water research, 124: 85-96.

Mook, W., Rozanski, K., 2000. Environmental isotopes in the hydrological cycle. IAEA Publish, 39.

Mukherjee, I. et al., 2020. Characterization of heavy metal pollution in an anthropogenically and geologically influenced semi-arid region of east India and assessment of ecological and human health risks. Science of the Total Environment, 705: 135801.

Sadoff, C.W., Borgomeo, E., Uhlenbrook, S., 2020. Rethinking water for SDG 6. Nature Sustainability, 3(5): 346-347.

Seaber, P.R., Kapinos, F.P., Knapp, G.L., 1987. Hydrologic unit maps.

Sneed, S.B., Mayewski, P.A., Dixon, D.A., 2011. An emerging technique: multi-ice-core multi-parameter correlations with Antarctic sea-ice extent. Annals of Glaciology, 52(57): 347-354.

Wang, L., Li, G., Dong, Y., Han, D., Zhang, J., 2015. Using hydrochemical and isotopic data to determine sources of recharge and groundwater evolution in an arid region: a case study in the upper–middle reaches of the Shule River basin, northwestern China. Environmental Earth Sciences, 73(4): 1901-1915.

Wang, P., Yu, J., Zhang, Y., Liu, C., 2013. Groundwater recharge and hydrogeochemical evolution in the Ejina Basin, northwest China. Journal of Hydrology, 476: 72-86.

Wang, S. et al., 2018. Combination of CFCs and stable isotopes to characterize the mechanism of groundwater–surface water interactions in a headwater basin of the North China Plain. Hydrological Processes, 32(11): 1571-1587.

Wang, X.-S., Zhou, Y., 2018. Investigating the mysteries of groundwater in the Badain Jaran Desert, China. Hydrogeology Journal, 26(5): 1639-1655.

Wu, X., Wang, X.-S., Wang, Y., Hu, B.X., 2017. Origin of water in the Badain Jaran Desert, China: new insight from isotopes. Hydrology and Earth System Sciences, 21(9): 4419-4431.

Wu, Y., Zhang, Y., Wen, X., Su, J., 2010. Hydrologic cycle and water resource modeling for the Heihe River Basin in Northwestern China. Science, Beijing.

Yang, Q. et al., 2011. Hydrological and isotopic characterization of river water, groundwater, and groundwater recharge in the Heihe River basin, northwestern China. Hydrological Processes, 25(8): 1271-1283.

Yao, Y. et al., 2015. Conceptual and numerical models for groundwater flow in an arid inland river basin. Hydrological Processes, 29(6): 1480-1492.

Zhao, L. et al., 2012. Origins of groundwater inferred from isotopic patterns of the Badain Jaran Desert, Northwestern China. Groundwater, 50(5): 715-725.

Zhu, G., Li, Z., Su, Y., Ma, J., Zhang, Y., 2007. Hydrogeochemical and isotope evidence of groundwater evolution and recharge in Minqin Basin, Northwest China. Journal of Hydrology, 333(2-4): 239-251.

Chapter 4. Tracing geochemical pollutants in stream water and soil from mining activity in an alpine catchment

This chapter is based on the following published paper:

Jianguo Li, Zongxing Li, Kate J. Brandis, Jianwei Bu, Ziyong Sun, Qiang Yu, Daniel Ramp. 2020. Tracing geochemical pollutants in stream water and soil from mining activity in an alpine catchment. Chemosphere 242:125167. https://doi.org/10.1016/j.chemosphere.2019.125167

Abstract

This research developed a method of tracing major water chemical parameters (WCP) and soil heavy metals (HM) to identify the processes of mining pollution in topographically complex landscapes. Ninety-nine spatially distributed water samples were collected to characterise the hydrochemical characteristics of an alpine river in north-west China. Sixty river WCP and fifty-six soil HM samples from areas near mining sites were then used to analyse the mining pollution process. Geographical and mining activity characteristics were derived from topographic and mine site information. The occurrence of sulphates (SO_4^{2-}) and nitrates (NO_3^-) in river water were highly correlated (up to 0.70), providing strong evidence of pollution from nearby mining activities. Levels of arsenic and cadmium were high in first and fifth order streams, where mining activities were most concentrated. The modelling results showed that geographical patterns and mining activity account for predicting HM distribution, and WCP can be reasonable predictors to trace soil mining pollution. This research can help improve the accuracy of predicting the mining pollution process.

Key words: alpine catchment, mining activity, soil heavy metals, water chemical parameters, geographical characteristics, stream orders, generalised additive models

Highlights

* Geographical characteristics of topography and mining activity strongly influence mining pollution in alpine catchments.

* Heavy metal distributions that account topographic patterning allow for spatial interpolation.

* Water chemical parameters can be reasonable predictors to trace soil mining pollution.

* Relationships between water chemical parameters and soil heavy metals can help understand the mining pollution process.

4.1 Introduction

Water resources are essential for supporting humans and ecosystems, makingwater availability and security a challenging global problem of great importance (Tilman et al. 2002; Vörösmarty et al. 2010). Water quantity and quality are two widely used criteria for evaluating water resource levels, both of which are highly sensitive to human activity (Novotny 2002; Scanlon et al. 2007). Deterioration in water quality can occur as a direct or indirect response to pollution, which conveys negative contaminants into aquatic systems and alters chemical composition, bacterial and nutrient accumulation, and sediment transport (Hounslow 2018). Rapid economic growth over recent decades has necessitated greater levels of mineral exploitation, required to provide important elements needed for developing populations (Duarte et al. 2019; Gredilla et al. 2019; Oliveira et al. 2019). However, mining activities typically require large quantities of water to process ore (Gunson et al. 2012), resulting in occurrences of mine effluent discharge and tailings seepage (Ramos et al. 2017; Nordin et al. 2018; Sánchez-Peña et al. 2018). This can lead to serious contamination of freshwater resources and polluted wastewater must often be managed for decades (Candeias et al. 2015; Fetter et al. 2017). Mining pollutants can drain directly into local runoff over land or can be dispersed into rivers via rainfall or other hydrological processes (Oliveira et al. 2018; Sánchez-Peña et al. 2018). Pollutants that infiltrate deep soil or enter streams can be transported into aquifers by exchange processes between ground and surface water. Groundwater can also be polluted directly if the aquifer system is destroyed through mining activity (Armienta & Segovia 2008). Although the effect of mining on water quality has been well studied on river areas proximal to mining activities in low altitude environments (Kondolf 1997), the process of large-scale mining impacts within alpine catchments is more complicated and less studied. For example, mountainous geomorphology can significantly alter the export of weathering solutes and transport pathways (Rascher et al. 2018; Wellen et al. 2018). Furthermore, alpine topographic features make it more difficult to understand transport processes of contaminants in upstream environments.

The upper reaches of alpine catchments are the powerhouse of freshwater generation and play an important role in enabling ecological balance and agricultural production in the middle and lower reaches (Gredilla et al. 2019). Runoff in these catchments is heavily influenced by changing climate conditions and extreme weather events (Arnell 2003; Painter et al. 2010; Li et al. 2016). However, because valuable mineral resources are often located in mountainous

locations, mining activities frequently occur in these locations, compounding other anthropogenic effects (Nordin et al. 2018; Sánchez-Peña et al. 2018). Traditionally, heavy metals (HM) are used to trace pollutants in the soil (Kabata-Pendias 2000), atmosphere (Francová et al. 2017), and aquatic systems (Tiwari et al. 2015; Tiwary et al. 2018). HM are often traced directly without considering conversion forces, and previous research has primarily focussed on each of these transmission media separately. However, combining these forces can help to understand the mechanisms that result in movement and exchange between them (Borrok et al. 2009). To provide a more holistic view of HM pollution, one solution is to examine a range of hydrochemical characteristics, as they are well suited to mapping hydrological processes (Adams et al. 2001) and pollutants from anthropogenic activities (Jalali 2007). This is because water chemical parameters (WCP) can trace sources of river water components (Helena et al. 2000), incorporating exchanges among soil, rock, and air, aswell as from additional sources (e.g. from human activity). Water ions (Ca^{2+}, Mg^{2+}, SO_4^{2-}, NO_3^-, etc.) and other physicochemical parameters, such as pH, electric conductivity (EC), and total dissolved solids (TDS), are important WCP that can be used to trace HM pollution, particularly for characterising water-rock interactions (Wellen et al. 2018). Therefore, improvements in understanding mining pollution may be gained by examining the interaction of WCP with trace elements from other carriers, such as soil HM within the environment.

In this study, samples of river water and soil from an alpine catchment in northwest China were analysed to consider two areas of interest. First, we characterised and traced hydro-chemical characteristics in the rivers, groundwater, melt, and precipitation within the catchment. Second, we used modelling techniques to predict the dispersal of soil HM from mining activities across the alpine catchment by quantifying the geographic variations and mining intensity information. This research highlights the correlation between the WCP and soil HM to interpret the process of mining pollution, and can help improve the accuracy of predicting the dispersal of soil heavy metals by combining the mining activity and geographical characters in a topographic mining site.

4.2 Methods and dataset

4.2.1 Study area

This study focused on the upper reaches of the Heihe River, China, as it is a topographically complex alpine landscape with a long mining history (Wei et al. 2018). The Heihe River has a length of 812 km, originating in the snow and glaciers of the Qilian Mountains in the north-eastern of the Qinghaie- Tibet Plateau, flowing into the Zhangye Basin through the Yingluoxia hydrological station, and finally entering the Ejinaqi oasis (Fig. 4-1a). As the river traverses

down from the plateau, beginning from glaciers above 4500m asl, perennial snow persists above 4000m asl while forests begin from elevations below 3000m asl. The mean annual temperature in the upper reaches is between -3 and 4 ^{o}C. The river region experiences a four-month wet season, occurring from June to September, during which more than seventy percent of annual precipitation occurs. Runoff from the upper reaches is the main source of flow generation, with an outbound runoff volume of $2.475 \times 10^{9} m^{3}$ annually, directly influencing water availability in the middle and lower reaches (Tian et al. 2018). The area is characterised by steep mountain valleys, with marked vertical zonations of vegetation, including glacier, alpine cold desert, marsh meadow, alpine shrub, and mountain grassland (Wang et al. 2009). The Zhamashike River (ZMSK) is the western branch of the upper stream, has a catchment area of approximately 5512 km^{2} and an average elevation of 3929m (Fig. 4-1b). Flow variation in the ZMSK is distinctly seasonal, with $44.43 \times 10^{4} m^{3}$/day flow in summer, but only 6.06 $\times 10^{4} m^{3}$/day flow in winter.

Geologically, the Heihe catchment is situated in the northern Qilian orogenic belt, between the Qilian and Alashan Blocks. The belt is an oceanic suture composed of subduction accretionary complexes. The ZMSK flows NW-SE through post-Devonian sedimentary cover, Quaternary cover, island arc volcanic rocks, and granites (Fig. 4-1c). Other covers include Neoproterozoic to early Paleozoic ophiolite sequences, and high-pressure metamorphic rocks. The primary metallogenic period of the area is early Paleozoic, while the lithology is composed of epi-metamorphic rock series, including clasolites, extrusive rocks, carbonates, and igneous rocks (Wei et al. 2018).

The region is rich in mineral resources, including coal, copper, lead, and zinc (Fig. 4-1b). Consequently, the study area has an extensive mining history over last four decades (Wei et al. 2018), with 23 mining sites possessing government licences during our survey in August 2013 (Table 4-1).

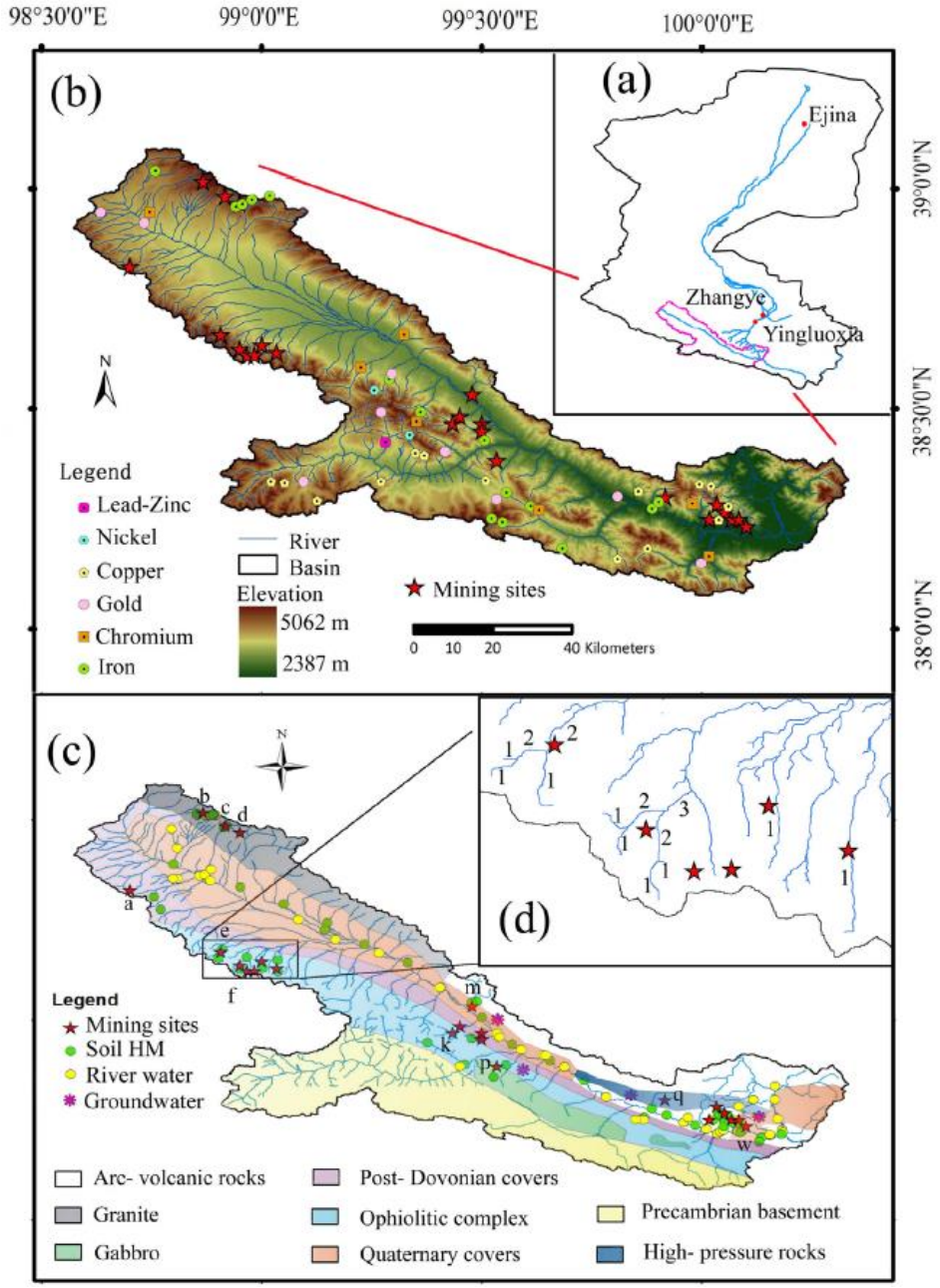


Figure 4- 1. Study site showing: (a) ZMSK in the Heihe River catchment (pink boundary line); and (b) location of known mineral deposits and mining sites; (c) soil and water sampling sites, lithology (adapted from Wei et al., 2018), and mining sites ID (following a longitude sequence from "a" to "w"); and (d) stream order systems.

4.2.2 Water and soil sampling

Ninety-nine water samples were collected between June 2012 and August 2013 across an elevation range of 2586- 4740m asl, comprising samples from rivers, wells, springs, precipitation, and meltwater from glaciers and snow (Fig. 4-1c). Sixty river samples (40 during summer, 20 during winter) were collected from the glacier terminus to the outlet along the main

river channel. Thirteen groundwater samples were collected, five from deep wells and eight from exposed springs. Eight glacieresnow meltwater samples were collected at the terminus of the Shiyi glacier during the ablation period. Eighteen precipitation samples were collected, one for each precipitation event that occurred at Shiyi glacier station. Due to the complicated topography characteristics and climate in this area, only five groundwater samples from deep wells, and eight from exposed springs were collected in summer. But all the springs were disappeared and the well water was run dry during winter fieldwork period. If we want to get more groundwater samples spatially in this area, the monitoring wells are needed to be drilled. But the drilling operation is difficult in this area. The current groundwater samples are not sufficient enough to give a good explanation for the interactions between surface water and groundwater. Therefore, we highlighted the interactions between the surface water and soil HM to explore the mining pollution. The interaction between surface water and groundwater with groundwater would be conducted in the future research.

All samples were stored at 4 ^{o}C prior to laboratory analysis. The geochemical composition of all samples was analysed using a Dionex-600 Ion Chromatograph (IC), while anions were analysed using a Dionex-2500 IC. Analytical precision was 95.6% for Ca^{2+}, 103.7% for Mg^{2+}, 94.3% for Na^{+}, 96.8% for K^{+}, and 98.6% for SO_4^{2-}, 97.2% for Cl^{-} and 96.4% for NO_3^{-}. We used fifty-six surface soil samples from a previous study (Bu et al., 2016), collected adjacent to mining sites in the catchment during August 2013 (Fig. 4-1c). A clean plastic dustpan and brush were used to collect samples from the top 20 cm of soil and each sample was stored in a plastic bag. All soil samples were air dried and kept at 20 ^{o}C until analysis. HM concentrations were measured using inductively coupled plasma mass spectrometry (ICP-MS). Sampling and analytical procedures followed previously published methodologies for tracing mining pollution and activity (Sanchís et al. 2015; Civeira et al. 2016a; Civeira et al. 2016b; Rodriguez-Iruretagoiena et al. 2016).

Table 4- 1. Detailed information for mining sites in the study catchment. The mine sites from "a" to "j" are considered in the upper part of the catchment, and almost of them are open pit mining. The "k" to "p" are in the middle part, and "l" to "w" are in the outlet part. OP: open pit mining; UG: underground mining (adapted from Wei et al., 2018).

ID	Extracted volume (10^4 t/year)	Status	Extraction method	Upper limit line (m asl)	Lower limit line (m asl)	Depth(m)	Area (km^2)	Wastes deposited volume (m^3)	Permission time range	Lithology
a	0.15	Abandoned	OP	4290	4190	100	0.57	84,530	07/2010–12/2014	Precambrian
b	0.33	Abandoned	OP	4582	4280	302	0.09	–	01/2013–01/2016	Granite
c	–	Completed	OP	4600	4300	300	0.50	22,001	12/2010–08/2013	Granite
d	28.33	Active	UG	4745	4240	505	1.12	–	11/2010–11/2013	Granite
e	1.30	Abandoned	OP	4308	4238	70	0.11	–	06/2011–07/2014	Mantle peridotite
f	3.06	Active	OP	4630	4362	268	1.30	83,260	07/2012–07/2017	Gabbro
g	0.35	Active	UG	4560	4280	280	0.31	22,003	07/2012–07/2013	Gabbro
h	1.25	Abandoned	UG	4539	4360	179	0.06	–	08/2011–08/2015	Gabbro
i	1.30	Active	OP	4330	4200	130	0.07	85,500	06/2011–07/2014	Gabbro
j	1.30	Abandoned	OP	4290	4220	70	0.03	–	06/2011–07/2014	Gabbro
k	5.25	Completed	Both	4460	4020	440	0.52	33,000	03/2010–10/2013	Post-Dovonian
l	–	Completed	UG	3959	3520	439	0.52	22,001	12/2010–10/2013	Post-Dovonian
m	3.25	Completed	UG	3650	3620	30	0.11	–	06/2011–06/2014	Quaternary covers
n	–	Completed	UG	3959	3520	439	0.52	–	12/2010–10/2013	Post-Dovonian
o	–	Completed	UG	3959	3520	439	0.52	–	12/2010–10/2013	Post-Dovonian
p	0.50	Abandoned	UG	4100	3920	180	0.12	32,007	12/2011–06/2014	Post-Dovonian
q	0.58	Active	UG	3130	3018	112	0.16	27,002	07/2012–07/2015	Arc-volcanic rocks
r	1.67	Completed	UG	3530	3000	530	0.13	37,007	10/2012–10/2013	Arc-volcanic rocks
s	10.67	Active	Both	3680	3150	530	0.54	–	12/2010–07/2013	Arc-volcanic rocks
t	6.50	Completed	UG	3460	3084	376	0.12	–	06/2011–06/2013	Arc-volcanic rocks
u	4.50	Completed	UG	3150	2790	360	0.19	–	03/2012–10/2015	Arc-volcanic rocks
v	8.00	Abandoned	UG	3159	2730	429	0.20	–	12/2010–01/2013	Arc-volcanic rocks
w	–	Active	UG	2694	2550	144	1.07	–	07/2011–09/2013	Arc-volcanic rocks

4.2.3 Stream order system

To understand variation in water pollutant concentration contributed to by mining activity across different stream scales, stream position and order was defined by the numerical order of branches in the river system (Horton, 1945; Strahler, 1957). Different methods for topologically ordering tributaries according to their distance from primary sources are available, however, we chose the Strahler method as it assigns the outermost tributaries as 'first order', and the stream orders for soil and mining sites are assigned by the proximity principle, making it valuable for analysing headwater quality (Horton 1945; Kang et al. 2008). Stream order increases when two streams of the same order merge. Where two links of different orders intersect, the higher of the two orders is given. Stream order was analysed using ArcGIS 10.5 (Environmental Systems Research Institute Inc.), using a 30m resolution digital elevation

model (DEM). Using the hydrology toolset in ArcMap 10.5, outlets were identified and then catchments and sinks extracted from the DEM, as well as flow direction and accumulation. An area of 10 ha was applied as the 'minimum accumulation area' to create streams and outlets for the stream order network.

During summer sampling, nine river samples were from first order streams, six samples from second, third and fourth-order streams, and thirteen from fifth-order streams (Fig. 4-1d, Table 4-2). During winter, four river samples were from first and second-order streams, two from third-order streams, three from fourth-order streams, and seven from fifth-order streams. There were twenty-two, eleven, five, four, and fourteen soil HM samples adjacent to first to fifth-order streams. Of the twenty-three authorised mining sites, nine were adjacent to first-order streams, one was proximal to a second-order stream, while the remaining thirteen were all adjacent to fifth-order streams.

Table 4- 2. The number of water samples, soil samples, and mining sites in or adjacent to different stream orders (first to fifth).

Orders	Stream summer	Stream winter	Soil samplings	Mining
First	9	4	22	9
Second	6	4	11	1
Third	6	2	5	0
Fourth	6	3	4	0
Fifth	13	7	14	13

4.2.4 Predicting heavy metal distributions

The regional distributions of HM were obtained by modelling the spatial patterning of HM from terrain and mining activity. The geographic complexity of the alpine catchment was captured by calculating topographic variation within the DEM using the 'terrain' function in the 'raster' package in R v3.5.1 (R Core Team, 2018): slope, aspect, flow direction, TRI (Terrain Ruggedness Index), TPI (Topographic Position Index), and roughness. TRI is the difference between the absolute elevation values of a cell and the values of its eight encompassing cells. TPI is calculated as the difference between the value of a cell and the mean value of its eight surrounding cells. Roughness is used to show the difference between the highest and the lowest cell values and its eight encompassing cells.

To account for variation attributed to the proximity and intensity of mining activities, we calculated two spatial functions from mine site locations, mine density and distance to the nearest mine (m). The function 'bkde2D' in R was used to perform a twodimensional kernel density estimation of mining activity, using a bandwidth of 0.05. Distance to the nearest minewas obtained using the function of 'distanceFromPoints', calculating the distance from a group of points to sampling points. Generalised linear models (glm) were run for each HM variable using topographic variables, elevation (m), slope (degrees), TPI, TRI, and roughness, and the two mine activity variables in R, using the 'mgcv' package. The 'predict' function was then used to create rasters of the predicted HM distribution across the entire catchment.

4.2.5 Tracing heavy metals through water chemical parameters

To interpolate values of HM to water sampling sites, we extracted raster values at all locations, using the 'extract' function in the 'raster' package in R. Generalised additive models (gam) were then used to model the relationship between HM and WCP using the 'mgcv' package to account for potential non-linear relationships. K^+, Na^+, TDS, and total ions were removed due to high correlation with other parameters (>0.7). Model selection was employed to identify the best candidate model from a combination of all possible models. Using model averaging approaches in the 'MuMIn' package in R, models were ranked in order of importance based on the Akaike information criterion (AIC) and those variables included in the best model set (AIC < 2) were used to find the best model for each HM.

4.3 Results

4.3.1 Hydrological background values

The total dissolved solids (TDS) in the river water ranged from 134.48 to 711.76 mg L^{-1}, with a mean value of 367.42 mg L^{-1} (Table 4-3); this is more than three times the global average TDS for river water (115mg L^{-1}, Zhu & Yang 2007). Well water had the highest mean TDS (489.8 mg L^{-1}). The highest mean electric conductivity (EC) was in meltwater (1074.75 mS cm^{-1}), followed by water from wells, springs, rivers, and precipitation (Table 4-3). The mean EC value for river water was 600.93 mS cm^{-1}, ranging from 244.5 to 1148 mS cm^{-1}. The rank order of cations in the river water, from highest to lowest concentration, was $Ca^{2+} > Mg^{2+} > Na^+ > K^+$. The rank order for anions, from highest to lowest concentration, was $SO_4^{2-} > Cl^- > NO_3^-$. SO_4^{2-}, Ca^{2+} and Mg^{2+} were the ions with highest concentration in river water, accounting for 45.8%, 27.2% and 10.4% of the total ionic concentration respectively. The water from the river, springs, wells, and meltwater was of the SO_4^{2-}-Mg^{2+} -Ca^{2+} type, with these ions accounting for more than 80% of the total ionic concentrations (Fig. 4-2). There were seasonal patterns evident,

with the concentrations of most river water samples from winter being higher than those from summer. The mean values of Mg^{2+}, Na^{+}, K^{+}, $NO3^{-}$, $SO4^{2-}$ in river water in winter were 35.81, 22.66, 2.28, 21.88 and 160.27mg L^{-1} respectively, whereas the mean values of those ions in summer were 23.31,11.41, 1.83, 7.40 and 102.14 mg L^{-1} respectively.

Table 4- 3. Average ionic composition (mg L^{-1}) of different water types from ZMSK River.

Samples	EC (μs/cm)	TDS (mg/L)	pH	Ca^{2+}	Mg^{2+}	Na^{+}	K^{+}	NO_3^{-}	SO_4^{2-}	Cl^{-}
River	600.93	367.42	8.08	71.91	27.47	15.16	1.98	12.22	121.52	14.85
max	1148.00	711.76	9.14	110.96	43.28	32.09	2.75	37.94	317.18	43.55
min	244.50	134.48	7.14	47.45	9.12	6.99	0.82	0.21	8.76	4.00
Spring	707.75	461.83	7.85	83.69	34.64	16.14	2.06	23.34	173.96	17.51
Well	788.00	489.80	7.74	99.81	56.79	55.15	5.32	65.28	193.92	44.57
Meltwater	1074.75	402.02	7.99	208.87	41.34	97.84	5.89	4.93	463.30	30.09
Precipitation	14.51	7.85	7.54	17.39	3.31	5.91	2.31	4.23	8.26	3.38
WHO limits	1500	NA	6.5-8.5	200	150	200	NA	NA	NA	250

The ionic concentrations of water samples in the ZMSK met all the requirements for drinking water according to World Health Organisation (WHO) standards (Lin et al. 2012). Water samples also met most of China's Environmental Quality Standards in GB383- 2002 for surface water and GB/T 14848-93 for groundwater (Han et al. 2016). All waters were alkaline, with pH values greater than 7.0 (Fig. 4-2). The pH value of river water in summer ranged from 7.14 to 8.33, with an average value of 7.87; this was lower than that of river water in winter, which ranged from 7.72 to 9.14, with an average of 8.5. Under the drinking water standards, pH values should be between 6.5 and 8.5. All summer samples were within this range and 10 of the 20 winter samples were above the range. The processes controlling stream solute chemistry are complicated (Kim et al. 2017), and the recharge sources may explain the higher pH in winter. This catchment is mainly recharged by wet precipitation in summer with a lower pH (Zhao et al. 2011), while in winter the groundwater with a high pH plays a more important role. Further research is needed for specific mechanism.

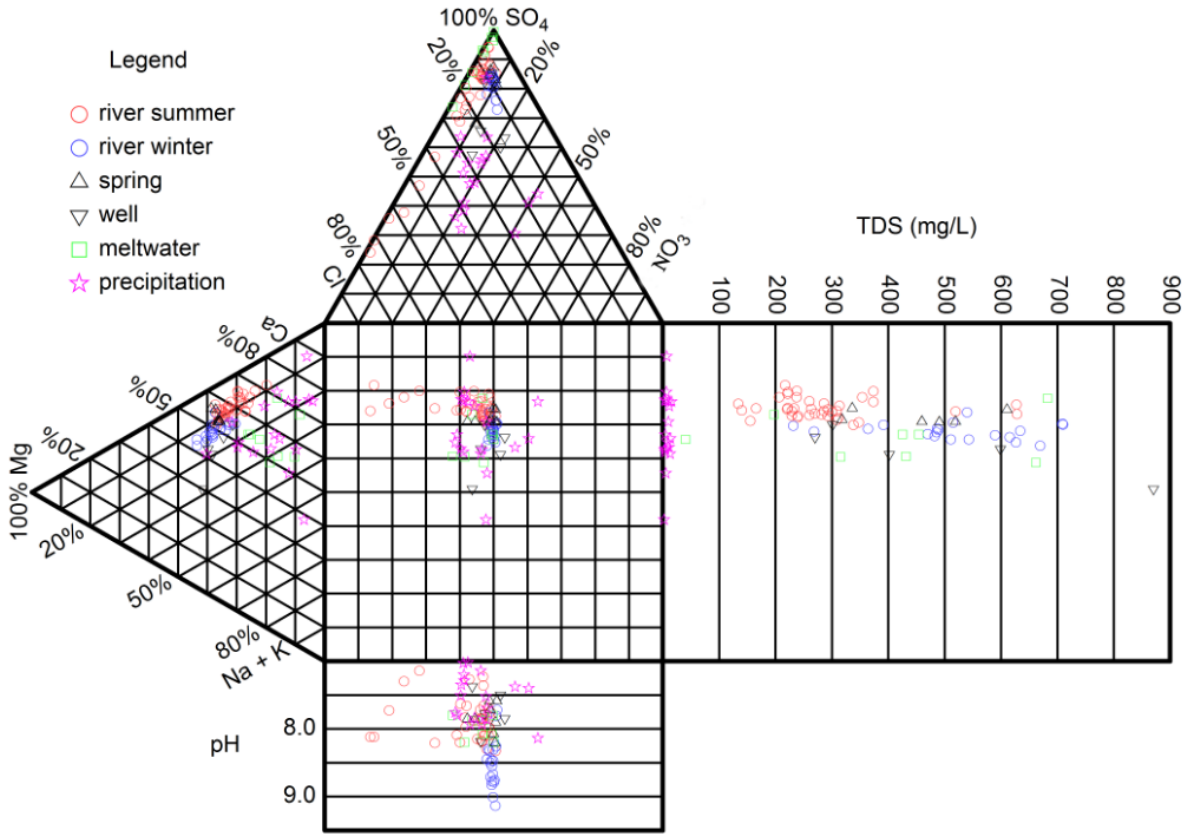


Figure 4- 2. Durov plot to graphically show cation and anion concentrations, pH, and TDS of the multiple water samples.

**All samples were distributed in the central rectangle, and each sample will have four corresponding points by dropping vertical lines from the central onto the four directions separately: the point in top triangle representing the anion percentage, the left triangle standing for the cation percentage, the bottom rectangle representing the pH value, and the right rectangle corresponding to the TDS value. The two ternary diagrams depicting the normalized percentage of anions and cations highlights layout characteristics, for example, that sulfates are dominant for most samples.*

4.3.2 Water chemical composition sources

Although the study area has been influenced by other human activities, the main anthropogenic pollutants were expected to be from mining, with the contribution from other activities likely to be negligible. To test this hypothesis, a Pearson correlation analysis was applied to study the relationships between ions, TDS, EC, and pH, and to identify possible sources. Considering the temporal variations in the water chemical composition, separate analyses were undertaken for river water in seasons. When the correlation coefficient between two elements is over 0.5, it is assumed that they have the communal possible sources. Shapiro-Wilk test was applied to

examine the normality. Due to the p values of Na^+, TDS and pH in river water during summer were 0.001, 0.000 and 0.02, they were removed for the Pearson correlation analysis. In summer, there were significant positive correlations between SO_4^{2-} and all other components except Cl^- (Table 4-4). SO_4^{2-} and Ca^{2+} were the ions with the highest concentration in the river water in summer, and the significant correlation between them indicated the contribution of gypsum to the composition of the river water. Nitrates (NO_3^-) usually indicate a human influence (Mayer et al., 2002), such as fuel burning by large machinery. $NO3^-$ was significantly correlated with most ions, especially Ca^{2+}, Mg^{2+}, Na^+, SO_4^{2-} and K^+ (Table 4-4).

Table 4- 4. Pearson correlation analysis (r values) of ions taken from river water during summer (left below) and winter (right top). * Indicates the correlation coefficient is significant at the 0.01 level (2-tailed).

S/W	Ca^{2+}	Mg^{2+}	Na^+	K^+	NO^{3-}	SO_4^{2-}	Cl^-	TDS	EC	pH
Ca^{2+}	1	0.242	0.521*	0.661*	0.823*	0.673*	0.497	0.293	0.301	-0.678
Mg^{2+}	0.815*	1	0.373	0.542*	0.512*	0.683*	0.803*	0.708*	0.698*	-0.007
Na^+	0.361	0.554*	1	0.883*	0.401	-0.197	0.293	0.471	0.503	-0.236
K^+	0.500*	0.789*	0.503*	1	0.433	0.09	0.577*	0.536	0.548*	-0.447
NO^{3-}	0.773*	0.713*	0.649*	0.675*	1	0.084	0.718*	0.494	0.503*	-0.678
SO_4^{2-}	0.753*	0.834*	0.293	0.688*	0.700*	1	0.568*	0.412	0.373	0.09
Cl^-	0.428	0.065	0.059	−0.256	0.205	0.095	1	0.562*	0.552*	-0.496
TDS	0.546*	0.498	0.410	0.367	0.576*	0.670*	0.185	1	0.997*	-0.104
EC	0.450	0.376	0.369	0.273	0.495	0.530*	0.122	0.924*	1	-0.096
pH	0.292	0.346	−0.036	0.285	0.139	0.235	−0.071	0.228	0.145	1

In the winter water samples (Table 4-4), SO_4^{2-} was significantly positively correlated with Ca^{2+} and Mg^{2+}, indicating the contributions of CaSO4 and MgSO4 dissolution to runoff. The correlations were less strong than in those from the summer samples, indicating that the weathering intensity of rock is weaker in winter. In addition, SO_4^{2-} was not significantly correlated with NO_3^- or K^+ in winter, which contrasts with strong correlations in summer. This was associated with less mining activity during winter, and the seasonal variation of river flow from summer to winter may also play an important role. NO_3^- was strongly positively correlated with Ca^{2+}, Mg^{2+} and Cl^-; however, NO_3^- was not correlated with SO_4^{2-}, Cl^- or K^+ in winter,

which again contrasts with strong correlations in summer. Finally, Mg^{2+} and Ca^{2+} were not significantly correlated in winter, but they were positively correlated in summer (Table 4-4).

4.3.3 Stream order system

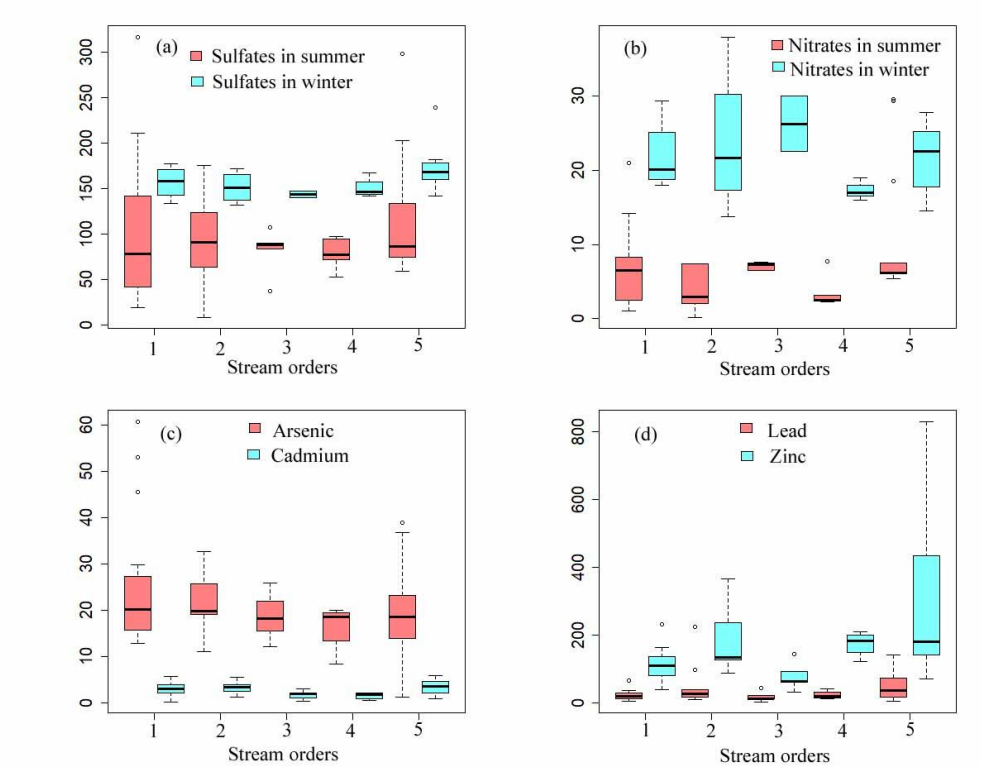


Figure 4- 3. Sulphates, nitrates, and soil HM concentrations in relation to stream orders in the ZMSK River, and the units for sulfates and nitrates are mg/L, for HMs being mg/kg.

First-order streams usually have wide riparian buffers that provide convenience for human activity. There are also no recharge sources from upstream, making them a good index to trace nonpoint source pollution. The median values of sulfates in summer did not show any variation trends, but rather showed higher individual values at the first and fifth-order streams (Fig. 4-3a). A similar distribution was observed for nitrates in summer (Fig. 4-3b). The high sulphate and nitrate concentrations in surface water and groundwater were likely caused by mining and agricultural activity (Rivett et al., 2008; Mativenga and Marnewick, 2018). Agriculture is limited in this location due to the cold weather and complicated landform (Gao et al., 2013). Thirteen mining sites in the alpine catchment were distributed near the main stream areas, where the elevation is lower and accessibility is greater for mining operations and transportation. This may explain the high concentrations of sulfates and nitrates in the last order streams. In contrast, most other mining points were distributed near the first-order streams. These areas were originally covered by bare rocks with low plant cover, which was more convenient for mining

exploitation. The high levels of pollutants in the original area also correlated with the level of mining activity.

The areas in which mining activity occurs have the highest concentrations of sulfates and nitrates, and other stream order systems have lower pollutant concentrations. Adverse natural conditions make mining activities more difficult in winter, so the distribution of sulfates and nitrates in winter did not show any obvious trend (Fig. 4-3a and 4-b). Regarding soil HM, the concentrations of As and Cd were highest in the first and fifth-order streams (Fig. 4-3c), while high concentrations of Pb and Zn were only found near the main stream areas (Fig. 4-3d).

4.3.4 Predicted dispersal of heavy metals

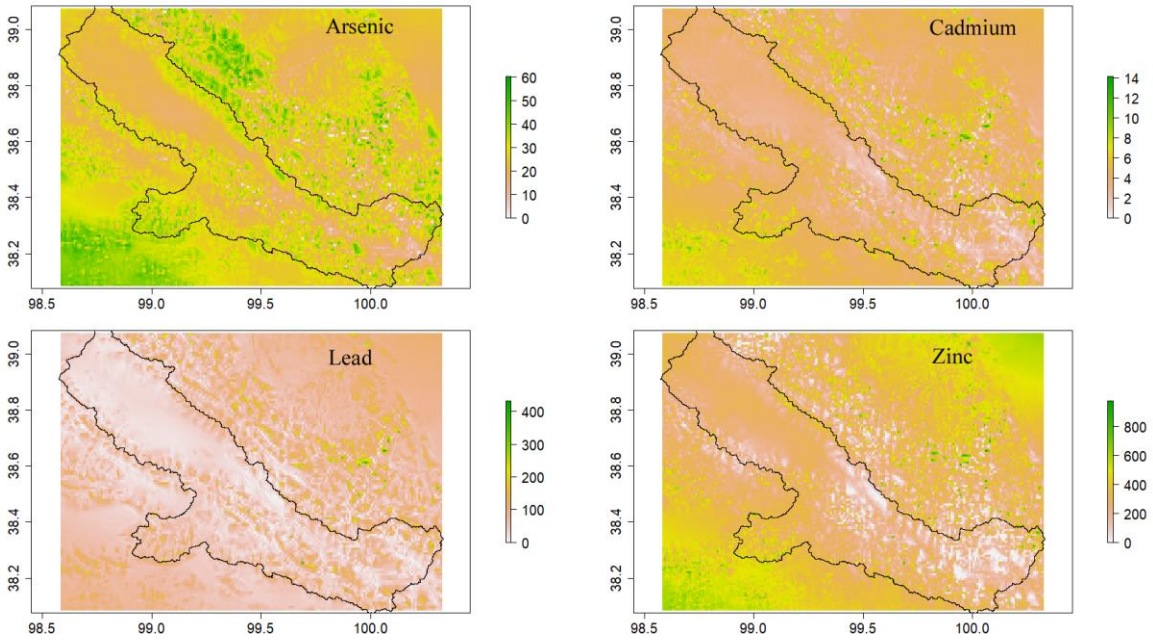


Figure 4- 4. The predicted spatial dispersal of HM in the ZMSK catchment, with the unit of mg/kg.

The dispersal of the HM contaminants from mine sites were predicted from geographical characteristics and mining activity characteristics (Fig. 4-4). HM values around mining sites were naturally highest, with most mining in the upper catchment centred around headwater areas. HM concentrations decreased from these headwater regions down to the main stream, reflecting attenuation processes over space. The prevalence of mining pollution in mountainous catchments is strongly driven by topography, stream dynamics, and land use. Compared to simple spatial interpolation in this area (Bu et al. 2016; Wei et al. 2018), predicted HM distributions can better reflect transport of mining pollution when using complex topographic information to address geospatial patterns.

4.3.5 Tracing heavy metals

Model selection identified the top models from all combinations of WCP for each HM. Models with the lowest AICc are listed in Table 4-5, highlighting the included variables and model support relative to other models within 2 AICc. When EC, NO_3^- and pH were selected as the smoothed variables, the AICc value of the top model for tracing arsenic was 322.34, and the degrees of freedom (df) was 13 (Table 4-5). The pH was also a smoothed variable in the top model for tracing cadmium and zinc. The top model for tracing lead was dominant by NO_3^- and SO_4^{2-}, and the AICc value was 526.19.

Table 4- 5. Top models identified through model selection based on AICc.

Heavy metal	Smoothed variables in best model	Degrees of freedom (df)	loglik	AICc	Akaike weight
Arsenic	EC+NO_3^-+Ph	13	-142.74	322.34	0.6
Cadmium	Ca^{2+}+pH	6	-76	166.16	0.6
Lead	NO_3^-+SO_4^{2-}	7	-254.45	526.19	0.45
Zinc	Ph+SO_4^{2-}	4	-352.01	712.75	0.17

**** The predictor variable values in the table are the number of times that variable was included in a subset of models and represents the relative variable importance of each variable. Symbols for predictor variables occurring in the best model set are total model number (TM), flow direction (FD), flow accumulation (FA), total ions concentration (TI) and degrees of freedom (df).***

Additional variables were included in models also within 2AICc of the best model. To not discount their influence on HM, these additional variables, where appropriate, were included in final models for each HM (Table 4-6). The final model for arsenic kept the same three variables as the top model, and the deviance explained (DE) value of these variables was 51.4%. The concentration of arsenic fluctuated in relation to NO_3^- concentration, reaching the minimum value when the concentration of NO_3^- was around 6mg L^{-1}, and showed a constant trend when the NO_3^- concentration was higher than 28 mg L^{-1} (Fig. 4-5a). Arsenic also showed a nonlinear relationship with pH, and increasing when pH was below 8.3, but decreasing after this point (Fig. 4-5b).

Both NO_3^- and pH were also smoothed variables for the other three HM, but showed the opposite trend. The NO_3^- displayed negative linear relationship with the concentrations of cadmium and zinc (Fig. 4-5f, l), while there was a positive linear relationship with lead (Fig. 4-5h). Conversely, pH was correlated with cadmium and zinc (Fig. 4-5d, n), but displayed negative linear relationship with lead (Fig. 4-5j).

The concentrations of lead decreased when the SO_4^{2-} concentration was less than 140mg L^{-1}, kept increasing after that point until the SO_4^{2-} concentration was around 240mg L^{-1}, and then decreased again (Fig. 4-5g). Compared to the top model for tracing zinc, the smoothed variables of NO_3^-, Ca^{2+}, Mg^{2+} and EC remained within 2AICc of the best model (Table 4-6). The concentrations of zinc fluctuated with SO_4^{2-} concentration, and two low peaks were reached when SO_4^{2-} concentration was 100 and 255 mg L^{-1} (Fig. 4-5k). The Mg^{2+} and EC showed positive linear relationship with zinc (Fig. 4-5o, p), while there was no significant change with Ca^{2+} (Fig. 4-5m). Zinc kept decreasing slightly, but remained constant when the Ca^{2+} concentration was over 70mg L^{-1}.

Table 4- 6. Generalised Additive Models (gam) describing the response of heavy metals (HM) to major water chemical parameters (WCP). Variable coefficients, Degrees of freedom estimated to model (Edf), Degrees of freedom estimated to waste (Ref.df), Significance of smoothed terms (F), P- value, Deviance explained (DE).

HM	Variable	Edf	Ref.df	F	P- value	DE
Arsenic	NO_3^-	8.738	8.977	3.394	0.001	51.4
	pH	2.124	2.654	1.991	0.153	
	EC	1	1	4.867	0.032	
Cadmium	pH	1	1	5.102	0.028	20.5
	Ca^{2+}	3.122	3.92	1.832	0.119	
	NO_3^-	1	1	0.303	0.584	
Lead	SO_4^{2-}	4.088	4.938	1.559	0.188	26.8
	NO_3^-	1	1	2.76	0.103	
	Ca^{2+}	1	1	0.127	0.723	
	pH	1	1	1.012	0.319	
Zinc	SO_4^{2-}	5.395	6.446	0.951	0.405	26.3
	NO_3^-	1	1	2.691	0.107	
	Ca^{2+}	1.741	2.224	0.438	0.776	
	pH	1	1	0.599	0.443	
	Mg^{2+}	1	1	1.364	0.249	
	EC	1	1	0.432	0.514	

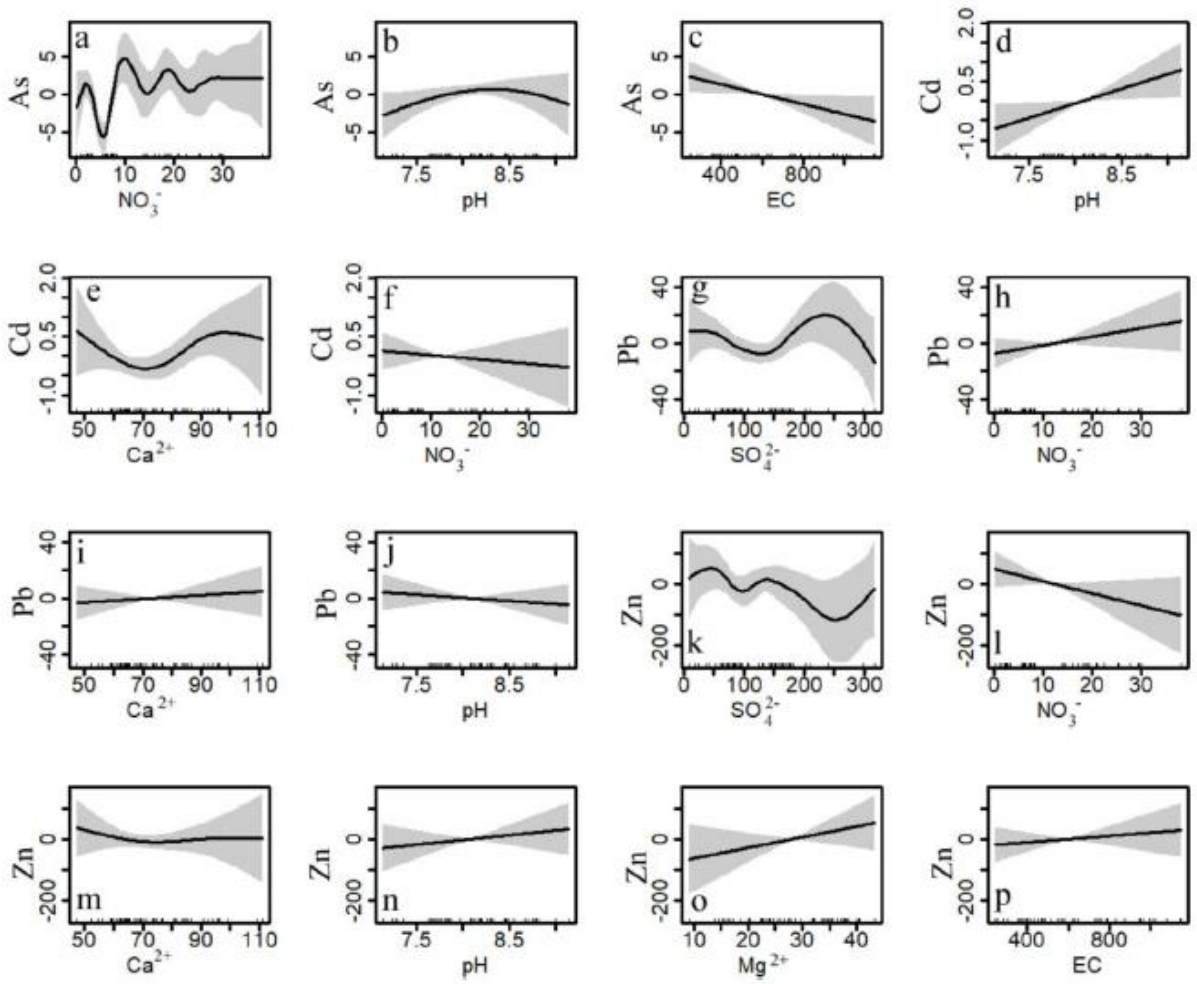


Figure 4- 5. Smoothed fits of relationships between WCP and HM. Tick marks on the x-axis are observed data points. The y-axis stands for the As (a to c), Cd (d to f), Pb (g to j), Zn (k to p).

4.4 Discussion

4.4.1 The possibility of water chemical parameters to trace mining activity

Chemical analysis of the main stream water met the standards for drinking and irrigation standards in the ZMSK, but previous research has found that tributaries were affected by local mining activity (Wei et al. 2018). Effective control of anthropogenic pollution here is crucial. The nitrate in river catchments is a common nitrogenous compound from the nitrogen cycle, but anthropogenic sources have greatly increased the concentration level, including discharges from fertilized agricultural lands (Matson et al. 1997), industrial waste water (Abdel-Halim et al. 2003), landfills, and energy consumption (Li et al. 2017). However, NO_3^- in the catchment was significantly positively correlated with most of other parameters during summer, reflecting the influence of human activities, such as energy consumption from mining operation equipment and vehicle emission from mining transportation. The NO_3^- here almost from the anthropogenic activity was further supported by the small content of NO_3^- in the river water.

Sulfates in riverwater are often formed as the result of thewater passing through rock or soil containing gypsum and other common minerals, or atmospheric deposition (Azabou et al. 2007; Shaffer et al. 2017). The high contents of SO_4^{2-} in river water showed the abundance of sulfide deposits in the catchment (Ma et al. 2009). Point sources, such as treatment and industrial discharges, and agricultural lands also contribute sulfates to river water (Tüfekci et al. 2007; Smolders et al. 2010). The significant positive correlation between SO_4^{2-} and most of other parameters showed the active activity for sulfates, but it is difficult to identify the natural or anthropogenic contributions. However, the coefficient of correlation between SO_4^{2-} and NO_3^- from the river water was up to 0.70 during summer, while only 0.08 during winter, indicating the human influence from mining activities during the warm periods. This inference was also supported by the stream order system analysis.

Most of the tracing elements, including the HM, have both natural and anthropogenic sources (Plank & Langmuir 1993; Pacyna & Pacyna 2001). Combing various tracing elements can give further insight on the contamination mechanism. This study explored to improve the accuracy of predicting the dispersal of soil HM by combining the mining activity and geographical characters. And then modelled the relationship between soil HM and WCP to analyse the mining pollution process. The analysis of WCP is much simpler and relatively inexpensive than analysing HM, so using the hydrological parameters is also an economic choice.

4.4.2 Mining pollution mechanism

Geographical characteristics, such as climate and topography parameters, play an important role in the mining pollution at an alpine catchment scale. The precipitation assumption directly determines the runoff volume in a precipitation-recharged catchment (Rodriguez-Iruretagoiena et al. 2016). Topography characteristics effect contamination transmission processes by influencing the stream directions and flow velocity. The soil erosion levels were determined by slopes, aspects, TPI, TRI, roughness and rainfall intensity. Large quantities of rocks are excavated to extract the desired mineral ore during operation at mining sites. The ore is then crushed into finely ground tailings and processed with various chemicals to extract the minerals. During rainfall events, a substantial quantity of sediment can be carried from the mining residuals on steep mountain valleys into the river, the other retained in the soil (Fig. 4-6).

Apart from natural factors, the variables ‘mining density’ and ‘distance to mining sites’ are necessary to be considered. The results also showed the importance of these two variables. Thus, mining pollution in alpine catchment area is a complex process, and more variables should be taken into account when analysing and modelling this process in order to provide more detailed and useful information. Very similar study of gold, coal and other mining results were reported

by previous research (Cerqueira et al. 2012; Oliveira et al. 2012; Arenas-Lago et al. 2013; Silva et al. 2013; Cutruneo et al. 2014).

The mining pollution mechanism in alpine catchments is a complicated process among the soil, groundwater, surface water, lithology, and sediments. From our study, we found that the WCP are satisfactory predictive variables to trace the HM pollution from the mining activity, and are able to provide a relationship between the soil HM and stream WCP. And more parameters are recommended to be considered in the future research to improve the quality of the model. The prevalence of mining pollution in mountainous catchment is strongly based on many factors, such as topography, stream dynamics and land use. The more components we are able to taken into account in the model, the more specific information we can interpret, and the closer to the real mining pollution process. In our model, the topographic complexity of the alpine catchment, including elevation, slope, aspect, flow direction, etc., was captured by calculating the topographic variation through variables derived from the 'terrain' function based on the digital elevation model (DEM). For two separate complex alpine catchments, if the DEM was available, then all the geological characters could be all extracted by this model. This model could work on different complex alpine catchments. The model behaviours will depend on the geographical characters and mining activity information.

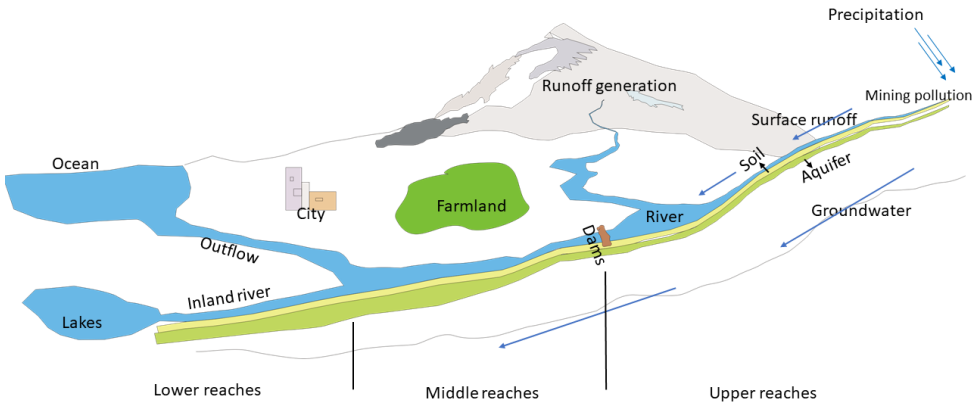


Figure 4- 6. Conceptual model of mining pollution in high mountain areas. The model shows the flow of water (blue arrows) from upper reaches to the middle and lower reaches. The headwater area is the main runoff- generation area, and the mining activity here has the potential to cause a series of water quality problems in middle and lower reaches.

4.5 Conclusions

The water from the river, springs, wells, and meltwater spatially distributed in the ZMSK were collected to characterise the hydro-chemical characteristics and assess the water quality. Then

the WCP and soil HM samples near mining sites were used to analyse the mining pollution process. This research presented a novel method of predicting the dispersal of soil HM by combining geographical factors and mining activity information, and then explored to trace soil HM from WCP around the mining alpine catchment.

The results showed that all water samples were SO_4^{2-}–Mg^{2+}–Ca^{2+} type throughout the year. The strong correlation between SO_4^{2-} and NO_3^- reflected the existing influence of mining activities. The levels of As and Cd were higher in the stream orders where the mining sites were concentrated. The prediction results showed that geographical and mining activity played an important role for the distribution of HM. The model selection results found that NO_3^- and pH were smoothed variables for all these four HM, and WCP are able to be reasonable predictors to trace soil mining pollution.

Increasing sampling density is a necessary method for improving the precision of predicting mining pollution in the future research. We strongly encourage similar research with the interactions between groundwater and surface water, more tracing elements, climate factors and lithology information to provide a more reliable method for analysing mining pollution mechanism at alpine catchments around the world.

4.6 References

Abdel-Halim S, Shehata A, El-Shahat M. 2003. Removal of lead ions from industrial waste water by different types of natural materials. Water Research **37**:1678-1683.

Adams S, Titus R, Pietersen K, Tredoux G, Harris C. 2001. Hydrochemical characteristics of aquifers near Sutherland in the Western Karoo, South Africa. Journal of hydrology **241**:91-103.

Arenas-Lago D, Vega F, Silva L, Andrade M. 2013. Soil interaction and fractionation of added cadmium in some Galician soils. Microchemical journal **110**:681-690.

Armienta MA, Segovia N. 2008. Arsenic and fluoride in the groundwater of Mexico. Environmental Geochemistry and Health **30**:345-353.

Arnell NW. 2003. Effects of IPCC SRES* emissions scenarios on river runoff: a global perspective. Hydrology and Earth System Sciences **7**:619-641.

Azabou S, Mechichi T, Sayadi S. 2007. Zinc precipitation by heavy-metal tolerant sulfate-reducing bacteria enriched on phosphogypsum as a sulfate source. Minerals Engineering **20**:173-178.

Borrok DM, Wanty RB, Ridley WI, Lamothe PJ, Kimball BA, Verplanck PL, Runkel RL. 2009. Application of iron and zinc isotopes to track the sources and mechanisms of metal loading in a mountain watershed. Applied Geochemistry **24**:1270-1277.

Bu J, Sun Z, Zhou A, Xu Y, Ma R, Wei W, Liu M. 2016. Heavy metals in surface soils in the upper reaches of the Heihe River, northeastern Tibetan Plateau, China. International journal of environmental research and public health **13**:247.

Candeias C, Ávila PF, Da Silva EF, Ferreira A, Durães N, Teixeira JP. 2015. Water–rock interaction and geochemical processes in surface waters influenced by tailings impoundments: impact and threats to the ecosystems and human health in rural communities (panasqueira mine, central Portugal). Water, Air, & Soil Pollution **226**:1-30.

Cerqueira B, Vega FA, Silva LF, Andrade L. 2012. Effects of vegetation on chemical and mineralogical characteristics of soils developed on a decantation bank from a copper mine. Science of the Total Environment **421**:220-229.

Civeira M, Oliveira ML, Hower JC, Agudelo-Castañeda DM, Taffarel SR, Ramos CG, Kautzmann RM, Silva LF. 2016a. Modification, adsorption, and geochemistry processes on altered minerals and amorphous phases on the nanometer scale:

examples from copper mining refuse, Touro, Spain. Environmental Science and Pollution Research **23**:6535-6545.

Civeira MS, Ramos CG, Oliveira ML, Kautzmann RM, Taffarel SR, Teixeira EC, Silva LF. 2016b. Nano-mineralogy of suspended sediment during the beginning of coal rejects spill. Chemosphere **145**:142-147.

Cutruneo CM, Oliveira ML, Ward CR, Hower JC, de Brum IA, Sampaio CH, Kautzmann RM, Taffarel SR, Teixeira EC, Silva LF. 2014. A mineralogical and geochemical study of three Brazilian coal cleaning rejects: demonstration of electron beam applications. International Journal of Coal Geology **130**:33-52.

Duarte AL, DaBoit K, Oliveira ML, Teixeira EC, Schneider IL, Silva LF. 2019. Hazardous elements and amorphous nanoparticles in historical estuary coal mining area. Geoscience Frontiers **10**:927-939.

Fetter CW, Boving T, Kreamer D 2017. Contaminant hydrogeology. Waveland Press.

Francová A, Chrastný V, Šillerová H, Vítková M, Kocourková J, Komárek M. 2017. Evaluating the suitability of different environmental samples for tracing atmospheric pollution in industrial areas. Environmental Pollution **220**:286-297.

Gredilla A, de Vallejuelo SF-O, Rodriguez-Iruretagoiena A, Gomez L, Oliveira ML, Arana G, De Diego A, Madariaga JM, Silva LF. 2019. Evidence of mercury sequestration by carbon nanotubes and nanominerals present in agricultural soils from a coal fired power plant exhaust. Journal of hazardous materials **378**:120747.

Gunson A, Klein B, Veiga M, Dunbar S. 2012. Reducing mine water requirements. Journal of Cleaner Production **21**:71-82.

Han D, Currell MJ, Cao G. 2016. Deep challenges for China's war on water pollution. Environmental Pollution **218**:1222-1233.

Helena B, Pardo R, Vega M, Barrado E, Fernandez JM, Fernandez L. 2000. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water research **34**:807-816.

Horton RE. 1945. Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. Geological society of America bulletin **56**:275-370.

Hounslow AW 2018. Water quality data: analysis and interpretation. CRC press.

Jalali M. 2007. Hydrochemical identification of groundwater resources and their changes under the impacts of human activity in the Chah basin in western Iran. Environmental Monitoring and Assessment **130**:347-364.

Kabata-Pendias A 2000. Trace elements in soils and plants. CRC press.

Kang S, Lin H, Gburek WJ, Folmar GJ, Lowery B. 2008. Baseflow nitrate in relation to stream order and agricultural land use. Journal of environmental quality **37**:808-816.

Kim H, Dietrich WE, Thurnhoffer BM, Bishop JK, Fung IY. 2017. Controls on solute concentration-discharge relationships revealed by simultaneous hydrochemistry observations of hillslope runoff and stream flow: The importance of critical zone structure. Water Resources Research **53**:1424-1443.

Kondolf GM. 1997. PROFILE: hungry water: effects of dams and gravel mining on river channels. Environmental management **21**:533-551.

Li J, Li Z, Feng Q. 2017. Impact of anthropogenic and natural processes on the chemical compositions of precipitation at a rapidly urbanized city in Northwest China. Environmental Earth Sciences **76**:1-14.

Li Z, Qi F, Wang Q, Song Y, Li J, Yongge L, Yamin W. 2016. Quantitative evaluation on the influence from cryosphere meltwater on runoff in an inland river basin of China. Global and Planetary Change **143**:189-195.

Lin CY, Abdullah MH, Praveena SM, Yahaya AHB, Musta B. 2012. Delineation of temporal variability and governing factors influencing the spatial variability of shallow groundwater chemistry in a tropical sedimentary island. Journal of Hydrology **432**:26-42.

Ma J, Ding Z, Edmunds WM, Gates JB, Huang T. 2009. Limits to recharge of groundwater from Tibetan plateau to the Gobi desert, implications for water management in the mountain front. Journal of Hydrology **364**:128-141.

Matson PA, Parton WJ, Power AG, Swift MJ. 1997. Agricultural intensification and ecosystem properties. Science **277**:504-509.

Nordin AP, Da Silva J, de Souza CT, Niekraszewicz LA, Dias JF, da Boit K, Oliveira ML, Grivicich I, Garcia ALH, Oliveira LFS. 2018. In vitro genotoxic effect of secondary minerals crystallized in rocks from coal mine drainage. Journal of hazardous materials **346**:263-272.

Novotny V 2002. Water quality: diffuse pollution and watershed management. John Wiley & Sons.

Oliveira ML, Da Boit K, Schneider IL, Teixeira EC, Borrero TJC, Silva LF. 2018. Study of coal cleaning rejects by FIB and sample preparation for HR-TEM: mineral surface chemistry and nanoparticle-aggregation control for health studies. Journal of Cleaner Production **188**:662-669.

Oliveira ML, Saikia BK, da Boit K, Pinto D, Tutikian BF, Silva LF. 2019. River dynamics and nanopaticles formation: a comprehensive study on the nanoparticle geochemistry of suspended sediments in the Magdalena River, Caribbean Industrial Area. Journal of Cleaner Production **213**:819-824.

Oliveira ML, Ward CR, Izquierdo M, Sampaio CH, de Brum IA, Kautzmann RM, Sabedot S, Querol X, Silva LF. 2012. Chemical composition and minerals in pyrite ash of an abandoned sulphuric acid production plant. Science of the total environment **430**:34-47.

Pacyna JM, Pacyna EG. 2001. An assessment of global and regional emissions of trace metals to the atmosphere from anthropogenic sources worldwide. Environmental reviews **9**:269-298.

Painter TH, Deems JS, Belnap J, Hamlet AF, Landry CC, Udall B. 2010. Response of Colorado River runoff to dust radiative forcing in snow. Proceedings of the National Academy of Sciences **107**:17125-17130.

Plank T, Langmuir CH. 1993. Tracing trace elements from sediment input to volcanic output at subduction zones. Nature **362**:739-743.

Ramos CG, Querol X, Dalmora AC, de Jesus Pires KC, Schneider IAH, Oliveira LFS, Kautzmann RM. 2017. Evaluation of the potential of volcanic rock waste from southern Brazil as a natural soil fertilizer. Journal of Cleaner Production **142**:2700-2706.

Rascher E, Rindler R, Habersack H, Sass O. 2018. Impacts of gravel mining and renaturation measures on the sediment flux and budget in an alpine catchment (Johnsbach Valley, Austria). Geomorphology **318**:404-420.

Rodriguez-Iruretagoiena A, de Vallejuelo SF-O, de Diego A, de Leão FB, de Medeiros D, Oliveira ML, Tafarel SR, Arana G, Madariaga JM, Silva LF. 2016. The mobilization of hazardous elements after a tropical storm event in a polluted estuary. Science of The Total Environment **565**:721-729.

Sánchez-Peña NE, Narváez-Semanate JL, Pabón-Patiño D, Fernández-Mera JE, Oliveira ML, da Boit K, Tutikian BF, Crissien TJ, Pinto DC, Serrano ID. 2018. Chemical and nano-mineralogical study for determining potential uses of legal Colombian gold mine sludge: experimental evidence. Chemosphere **191**:1048-1055.

Sanchís J, Oliveira LFS, De Leão FB, Farré M, Barceló D. 2015. Liquid chromatography–atmospheric pressure photoionization–Orbitrap analysis of fullerene aggregates on surface soils and river sediments from Santa Catarina (Brazil). Science of the Total Environment **505**:172-179.

Scanlon BR, Jolly I, Sophocleous M, Zhang L. 2007. Global impacts of conversions from natural to agricultural ecosystems on water resources: Quantity versus quality. Water resources research **43**.

Shaffer DL, Tousley ME, Elimelech M. 2017. Influence of polyamide membrane surface chemistry on gypsum scaling behavior. Journal of membrane science **525**:249-256.

Silva LF, de Vallejuelo SF-O, Martinez-Arkarazo I, Castro K, Oliveira ML, Sampaio CH, de Brum IA, de Leão FB, Taffarel SR, Madariaga JM. 2013. Study of environmental pollution and mineralogical characterization of sediment rivers from Brazilian coal mining acid drainage. Science of the total environment **447**:169-178.

Smolders AJ, Lucassen EC, Bobbink R, Roelofs JG, Lamers LP. 2010. How nitrate leaching from agricultural lands provokes phosphate eutrophication in groundwater fed wetlands: the sulphur bridge. Biogeochemistry **98**:1-7.

Tian Y, Xiong J, He X, Pi X, Jiang S, Han F, Zheng Y. 2018. Joint operation of surface water and groundwater reservoirs to address water conflicts in arid regions: an integrated modeling study. Water **10**:1105.

Tilman D, Cassman KG, Matson PA, Naylor R, Polasky S. 2002. Agricultural sustainability and intensive production practices. Nature **418**:671-677.

Tiwari AK, De Maio M, Singh PK, Mahato MK. 2015. Evaluation of surface water quality by using GIS and a heavy metal pollution index (HPI) model in a coal mining area, India. Bulletin of environmental contamination and toxicology **95**:304-310.

Tiwary R, Kumari B, Singh D. 2018. Water quality assessment and correlation study of physico-chemical parameters of Sukinda Chromite Mining Area, Odisha, India. Pages 357-370. Environmental Pollution. Springer.

Tüfekci N, Sivri N, Toroz İ. 2007. Pollutants of textile industry wastewater and assessment of its discharge limits by water quality standards. Turkish Journal of Fisheries and Aquatic Sciences **7**.

Vörösmarty CJ, McIntyre PB, Gessner MO, Dudgeon D, Prusevich A, Green P, Glidden S, Bunn SE, Sullivan CA, Liermann CR. 2010. Global threats to human water security and river biodiversity. nature **467**:555-561.

Wang N, Zhang S, He J, Pu J, Wu X, Jiang X. 2009. Tracing the major source area of the mountainous runoff generation of the Heihe River in northwest China using stable isotope technique. Chinese science bulletin **54**:2751-2757.

Wei W, Ma R, Sun Z, Zhou A, Bu J, Long X, Liu Y. 2018. Effects of mining activities on the release of heavy metals (HMs) in a typical mountain headwater region, the Qinghai-Tibet Plateau in China. International journal of environmental research and public health **15**:1987.

Wellen C, Shatilla NJ, Carey SK. 2018. The influence of mining on hydrology and solute transport in the Elk Valley, British Columbia, Canada. Environmental Research Letters **13**:074012.

Zhao L, Yin L, Xiao H, Cheng G, Zhou M, Yang Y, Li C, Zhou J. 2011. Isotopic evidence for the moisture origin and composition of surface runoff in the headwaters of the Heihe River basin. Chinese Science Bulletin **56**:406-415.

Zhu B, Yang X. 2007. The ion chemistry of surface and ground waters in the Taklimakan Desert of Tarim Basin, western China. Chinese Science Bulletin **52**:2123-2129.

Chapter 5. Local climatic factors controlling the isotopic composition of precipitation in the Qinghai-Tibet Plateau based on an eight-year observation

Abstract:

The hydrological cycles in the Qinghai-Tibet Plateau (QTP) have experienced significant changes over recent years. A thorough understanding of the relationships between isotopic compositions ($\delta^{18}O$ and δD) in precipitation and potential local climatic factors is a prerequisite requirement for discovering these hydrological processes in the Third Polar area. However, long-term continuous precipitation isotopic data in the QTP, as well as the synchronous local climatic records were scarce due to the harsh environment, poor traffic and observation conditions. A total of 430 precipitation samples together with nine local climatic variables from the Nam Co Lake over an 8-year were presented in this study for the first time to help interpret the influence of local climate patterns to the isotopic signal variations in precipitation. Combined with the classic local meteoric water lines (LMWLs) and environmental effects analysis, model selection and model averaging methods were introduced to explore the quantitative relationships between them.

Results found that on the annual scale, both the slope and intercept of the LMWL at Nam Co basin were higher than those of the global meteoric water line (GMWL), and these characteristics suggested that there were significant moisture sources from the continental landscape surface throughout the entire year at the monsoon-Westerly jet stream transition areas. The important explanatory variables of humidity and three temperature factors in the top model of d-excess also supported this conclusion. Seasonally, the lowest slope and intercept values of LMWL during the monsoon periods indicated the contributions of marine moisture. The lighter isotopic values in the monsoon season also reflected the strong rain out process by the high precipitation amounts in this period. We also found that isotopic signatures in different seasons were dominated by different local climatic factors according to the results of top models. For instance, the wind variables were important variables to explain the isotopic tracers during the Westerly season. The model selection and model averaging methods constitute an advancement in quantitatively interpreting the relationships between isotopic signatures of precipitation and their local climatic variables, which can supplement or precede more complex studies of hydrological cycles utilizing isotope tools. The long-term monitoring data sets in this study will enrich a multi-dimensional system of the hydrological cycles in the transition domain areas of QTP, as well as benefit the related model development and research.

Keywords: Nam Co Lake, Qinghai-Tibet Plateau, precipitation, stable isotopes, local climatic factors, altitude effects, amount effects, temporal variations, model selection

5.1 Introduction

Water stable isotopes of $\delta^{18}O$, δD, and deuterium excess (d-excess=$\delta D-8\delta^{18}O$) offer a powerful proxy for evaluating the hydrological cycles, such as precipitation (Price et al., 2008; Risi et al., 2008a), glacier (Kumar et al., 2018), permafrost water (Lacelle et al., 2014), surface water (Ogrinc et al., 2008), groundwater (Abbott et al., 2000; Wassenaar et al., 2009), and plant water (Bush et al., 2017), etc. Due to their stable properties, the water isotopic signatures have been widely used to analyze various hydrological processes, like the hydraulic connections between surface water and groundwater (Yin et al., 2011), stream water recharge mechanism (Li et al., 2014), hydrograph separation (Lyon et al., 2009), and the influence of climate change on cryosphere (Kreutz et al., 2003; Li et al., 2016b).

Precipitation is an important driving force to link the interactions of various water bodies, and a better understanding of what controls the isotopic values in precipitation is the prerequisite for studying hydrological cycles on local, regional, and global scales (Bowen et al., 2019; Duy et al., 2018; Kurita et al., 2009). The abundance of stable isotopes in precipitation reflects the atmospheric environment and the original water vapor sources during the process of precipitation formation. The stable isotopic compositions of precipitation are the results of the joint action of climatic variables at different spatial scales (Duy et al., 2018). On one hand, isotopic signatures in precipitation correlate to the long-distance transport of atmospheric moistures, which will experience dynamic changes due to various environmental backgrounds on the regional scales that mainly include the condensation and evaporation processes along the trajectory (Jouzel and Merlivat, 1984; Soderberg et al., 2013), air mass transport history (Kurita et al., 2009; Sjostrom and Welker, 2009), and temporal changes of the moisture sources (Krklec et al., 2018; Tian et al., 2007). On the other hand, the local climatic parameters that relate to the different local meteorological conditions or processes when precipitation events occur, will make isotopic signal variations more complicated (Duy et al., 2018). These local factors mainly include the precipitation amount, air temperature, the secondary evaporation below the cloud base (Crawford et al., 2017), and the moisture exchange between atmosphere and land (Ansari et al., 2020; Balagizi et al., 2018).

The Asian monsoon region has the most obvious monsoon climate patterns, and precipitation mainly originates from the Indian and Pacific Oceans (Shen and Lau, 1995). The Qinghai-Tibet Plateau (QTP) is an important geomorphic unit in the Asian monsoon region, and is

affected by two dominant weather systems, the Westerlies and monsoons. According to this climate characteristic, the QTP has always been divided into three zones, the Westerly domain area, the monsoon domain area, as well as the transition domain (Yao et al., 2013).

The hydrological cycles in QTP are very mysterious, notably by the rapid lake expansion and dramatic glacier retreat in the recent years (Liu et al., 2021; Song et al., 2014; Wu and Zhu, 2008). Most famously, the Siling Co lake, the current largest lake in Tibet, had experienced an area of 730 km^2 rapid expansion during the past four decades (Zhang et al., 2011). But what triggers drove the lake expansion are still unclear in this observation data-scare area (Kang et al., 2017). However, the QTP is the source of many important rivers in Asia (Liu et al., 2020), and this will impact the production and normal life in the middle and lower reaches in related countries and regions. Additionally, these unstable hydrological cycle events pose potential hazards, flooding pastures, villages, and forests within the QTP (Wang et al., 2021; Yuan et al., 2017), thus, it is very necessary and important to study the relationships between the stable isotopic compositions in precipitation and climatic factors in this area.

Previous studies on the relationship between climatic factors and isotopic values in Asian monsoons, especially in QTP areas, mainly focused on solving the following questions. (i) How various environmental effects affect the stable isotopic compositions in precipitation? (ii) What are the spatial and temporal variations of the isotopic signatures in precipitation? and (iii) what climatic factors control these isotopic values at a local or regional scale? Previous studies' attempts to solve these general problems are listed in Table 1, and include both limitations and benefits. For example, the local meteoric water lines (LMWLs) can reflect the long-term distribution of $\delta^{18}O$ and δD at the site (Putman et al., 2019), which can provide information of the distinctive local precipitation process and isotopic characteristics in a certain area. The introduction and development of many statistical and physical models not only help better discover the hydrological processes, but also can provide information about many important regional climatic parameters. For instance, information concerning the wind and humidity were available from the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR), air mass movement from Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) (Lone et al., 2019), and large-scale convective activities from outgoing longwave radiation (OLR) (Zhang et al., 2019). Although previous studies have conducted systematic research, there are still many knowledge gaps due to the difficult of sampling high spatiotemporal resolution data from the QTP area. Furthermore, the variable landscape, and complex geological structure in these high mountains and rural areas make it even harder to collect continuous precipitation samples.

Table 5- 1. Summary of relevant studies conducted previously for investigating relationships between the precipitation isotopic compositions and climatic variables in Asian monsoon areas, especially the QTP area (The current paper is added for completeness). Three components in the 'Key results' column are identified by the code: (1) Temporal scales (year), (2) Total numbers of precipitation samples, (3) Whether involve the influences of local climatic factors to precipitation isotopic composition, and their numbers, and (4) Whether the climatic factors are synchronously observed, and if not, what is the method used to collect these climatic data sets. NA means not available, or not mentioned.

Study	Area	methods	Key results
(Duy et al., 2018)	Mekong Delta, Vietnam	LMWLs, linear correlation, MLR, importance analysis; back trajectory modeling	(1)1.5 years, (2)74 samples from six sites, (3)Yes, but only temperature, precipitation amount, relative humidity (4)No, extracted from local climate records
(Laonamsai et al., 2021)	Thailand	LMWLs, environmental effects analysis; back trajectory modeling, regression analysis	(1)3 years, (2)702 samples from twenty-six sites, (3)No, only regional climatic factors, (4) No, extracted from GDAS
(Zhang et al., 2019)	Siling Co lake, QTP	LMWLs, back trajectory modeling, correlation analysis	(1)3 years, (2)159, (3) Yes, but only three local factors, including temperature, precipitation amount, and relative humidity, (4)No, extracted from GPCC, OLR; GDAS
(Li et al., 2020)	Zhimenda, QTP	LMWLs, environmental effects analysis, regression analysis	(1)2 years, (2)305 samples from three sites, (3) Yes, but only two local factors, temperature and precipitation amount, (4) NA
(Kang et al., 2017)	Nam Co lake, QTP	LMWLs, Isotope mass balance model	(1)3 years, (2)284, (3) Yes, but only precipitation amount, (4) Yes, local climatic data were recorded at Nam Co station
(Li et al., 2016a)	Wushaoling, QTP	LMWLs, environmental effects analysis, regression analysis, back trajectory modeling	(1)1 year, (2)468 from 5 sites, (3) Yes, but only two local factors, including temperature and precipitation amount, (4) NA
(Lone et al., 2019)	Indus river basin, Himalayas	NCEP/NCAR reanalysis, HYSPLIT back trajectory modeling, LMWLs, Lagrangian moisture diagnostic	(1)1 year, (2)110, (3)Yes, but only two local factors, temperature and precipitation amount,

			(4)No, extracted from NCEP/NCAR Reanalysis 1 dataset, and GDAS
This study	Nam Co lake, QTP	LMWLs, environmental effects analysis, model selection, GLM, GAM	(1)8, (2)430, (3)Nine local factors, (4)Yes, local climatic data were recorded at Nam Co station

Abbreviations in this table, local meteoric water lines (LMWLs), multiple linear regression (MLR), Global Data Assimilation System (GDAS), Global Precipitation Climatology Center (GPCC), outgoing longwave radiation (OLR), Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT), National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR), generalized linear models (GLM), and generalized additive models (GAM).

As a result, there are three main knowledge gaps currently. (i) There are rarely long-term continuous precipitation isotopic data in the QTP (details in Table 5-1), due to the harsh environment, poor traffic, and observation conditions. Although a few local meteorological conditions have examined the precipitation isotopic composition in the QTP, the previous studies mostly concentrated on the eastern part of the QTP (Yu et al., 2006), northeastern QTP (Li et al., 2016a), southern QTP (Tian et al., 2001b) etc., only a few attempts had been made to concentrate on the transition domain zones of the monsoon and Westerly transition zones (Zhang et al., 2019), where the large Siling Co Lake and Nam Co Lake are located. Additionally, some of the research only focused on some special period, like the summer monsoon periods (Gao et al., 2018). To the best of our knowledge, the current longest temporal scale for the precipitation isotopic compositions in the QTP was only three years (Kang et al., 2017). (ii) In terms of the climatic factors that potentially influence the isotopic signature in precipitation, most of the research only focused on the regional climatic scales, such as ENSO (Cai et al., 2017; Ishizaki et al., 2012); convective activities (Ansari et al., 2020; Risi et al., 2008b), the origins and properties of the air masses (Kurita et al., 2015; Uemura et al., 2012); as well as the distillation and secondary evaporation during moisture long-distance transport process (Tan, 2014; Tian et al., 2001a). However, relatively little attention has been paid to the local climatic factors (Duy et al., 2018). Furthermore, most of the local climatic data sets were extracted from the local climate records (Duy et al., 2018; Laonamsai et al., 2021; Lone et al., 2019; Zhang et al., 2019) (Table 5-1), rather than synchronously observed data for every single precipitation event, which will inevitably discount the accuracy of analysis. (iii) The relationships between isotopic values in precipitation and their potential explanatory variables were rarely quantitatively analyzed by

previous studies. Although some previous studies have reported some results qualitatively, few specific quantitative explanations of each factor to stable isotopes have been provided.

Given the above-mentioned knowledge gaps, continuous sampling and synchronous local climatic variables over a long-term period at the transition zones of the QTP are two essential requirements to fill these current knowledge gaps. In this study, the isotopic ratios in precipitation from valid 430 continuous events were collected from the Nam Co Lake over 8 years, and nine synchronous local climatic variables were recorded at the Nam Co station by the Chinese Academy of Sciences since 2005, including the precipitation amount, precipitation duration, relative humidity, temperature, maximum temperature during each event, minimum temperature during each event, pressure, wind direction and wind speed. Then the LMWLs, and environmental effects were applied to discover the background information of precipitation isotopic compositions and the model selection, generalized linear models (GLM), and generalized additive models (GAM) were introduced into the field of isotope hydrology, and tried to fill these current knowledge gaps. Specifically, the objectives of this study are (i) the typical patterns of precipitation stable isotopic signatures in Nam Co, (ii) the environmental effects in this area, (iii) temporal variations of isotopic signatures in Nam Co, and (iv) quantitatively relationships between isotopic values in precipitation and their potential explanatory variables. The model selection method is novel in the field of isotope hydrology (Allen et al., 2019), and can enhance their power to quantitatively interpret the relationships between isotopic signatures in precipitation and local climatic records. The long-term monitoring data sets in this study will enrich a multi-dimensional system of the hydrological cycles in the transition domain areas of the QTP, and benefit the related model development and research in these data-scarce locations. Results can also benefit regional development and conservation. The results of this study are also important to establish a scientific-informed strategy for regional development and management.

5.2 Materials and methods

5.2.1 Study area

The Nam Co Lake is an ideal site to study the impacts of local climatic factors on isotopic compositions in precipitation, because it is located within the transition zones where is jointly controlled by the monsoon and Westerly weather systems, as well as strong local recycling processes, such as the cloud-related processes (Yao et al., 2013). The Nam Co Lake is an endorheic basin, covers an area of approximately 2018.25 km^2 measured in 2009 (Zhang et al., 2013), and the maximum depth was over 90 m (Wang et al., 2009), with an elevation of 4718 m above sea level.

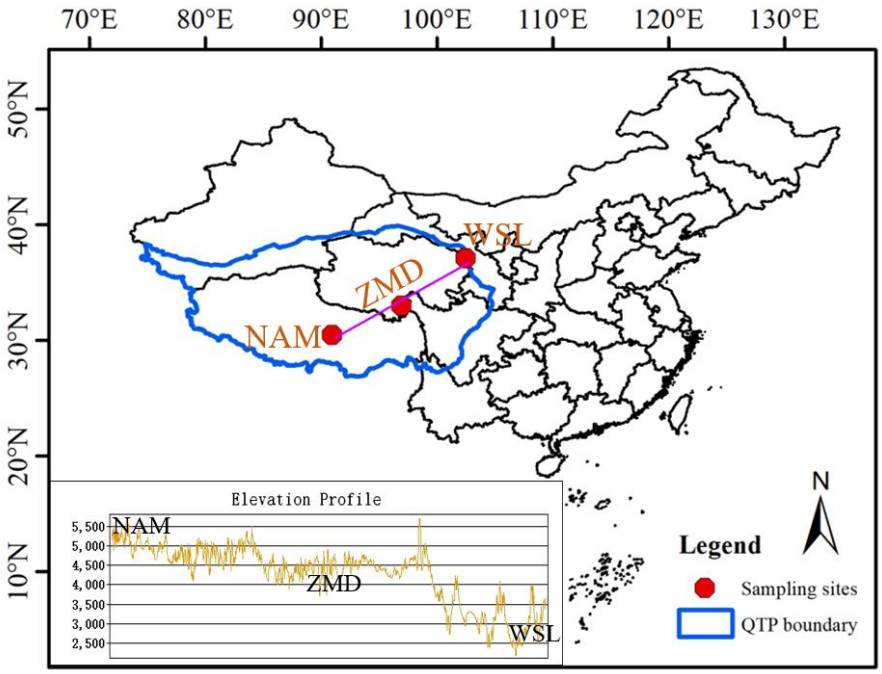


Figure 5- 1. Boundary of QTP (blue polygon) and locations of sampling sites. The NE-SW cross section denotes the three precipitation sampling sites, Wushaoling, Zhimenda, and Nam Co (detailed information in Table 5-2), used for studying the altitude effects in the QTP, with its elevation profile under left corner.

The climate conditions of this area feature a semi-arid subarctic plateau climate with strong seasonality, which has a mean annual air temperature of −0.6 °C and average annual relative humidity of ~53.6%. The multi-year mean annual precipitation is nearly 414.6 mm, with most of the precipitation events at this drainage occurring during the monsoon periods, which accounts for more than 80% of the annual precipitation (Kang et al., 2011). According to the interactions of Westerlies and monsoon predominate characteristics, four seasonal monsoon periods were divided in this study, including April–May as the pre-monsoon, June–August (JJA) as the monsoon periods, September as the post-monsoon, and October to next March as the Westerly domain season with a dry and cold air mass (Keil et al., 2010; Li et al., 2007; Zhang et al., 2019). On the set of monsoon seasons, the marine air masses move upward to the Himalaya mountains and are propagated into the QTP interior. The humidity and precipitation are significant during the monsoon period compared with other seasons. There are almost seven geomorphological units across the entire Nam Co drainage basin, among which the southwestern mountains, Nyainqentanglha mountains, and western highlands are the three significant areas, accounting for21%, 30%, and 32% of the total basin area, respectively (Keil et al., 2010). Wetland is also an important landscape feature in the watershed, with a surface area of approximately 600 km^2, including almost 200 km^2 along the lakeshore (Kang et al., 2017).

5.2.2 Field sampling and laboratory analysis

The Nam Co Multi-Sphere Observation and Research Station (Nam Co station, 30°46.44′ N, 90°59.31′ E) was established by the Chinese Academy of Sciences in 2005, and makes it possible to observe and investigate the long-term impact of climate change on water cycles on the QTP. Precipitation samples and synchronous local climatic data sets observations were started in 2005, and the record data sets between 2005 and 2012 were the first time used here for interpreting the relationships between precipitation isotopic compositions and potential local climatic patterns.

Precipitation samples were collected using an automated precipitation collector with a long narrow tube attached to the bucket container at Nam Co station. The event wise samples were collected immediately after the rain stopped (Kang et al., 2017). After collection, all precipitation samples of each rain event were sealed in a 15-mL polyethylene bottle and frozen at −18 °C to avoid any evaporation. Before analysis, they were stored in a refrigerator with the temperature set at −4 °C, so they could melt gradually without evaporation. The isotopic compositions of oxygen and hydrogen of precipitation are defined by,

$$\delta^{18}O = \frac{(^{18}O\ /\ ^{16}O)_{Sample} - (^{18}O\ /\ ^{16}O)_{V-SMOW}}{(^{18}O\ /\ ^{16}O)_{V-SMOW}} \times 1000\ ‰ \qquad (1)$$

$$\delta D = \frac{(^{2}H\ /\ ^{1}H)_{Sample} - (^{2}H\ /\ ^{1}H)_{V-SMOW}}{(^{2}H\ /\ ^{1}H)_{V-SMOW}} \times 1000\ ‰ \qquad (2)$$

where V-SMOW refers to the Vienna Standard Mean Ocean Water standard, and all precipitation samples were analyzed with respect to V-SMOW2 and Standard Light Antarctic Precipitation (SLAP), on a Liquid Water Isotope Analyzer (Picarro L2130-i isotopic H^2O Analyzer), with precisions for δ18O of±0.1‰ and δD of±0.4‰, respectively.

5.2.3 Model selection and model averaging

Model selection is the task of selecting a statistical model from a set of candidate models with combinations (subsets) of fixed effect terms in the global model based on optional rules for model inclusion. Model averaging is a robust method of addressing the model uncertainty. When multiple model algorithms or candidate models are fitted, model selection was conducted to select the top model from a set of best models. The top model is the most parsimonious combination of explanatory variables using the cross- validation process. There are many selection criterions for model selection, and Akiake Information Criterion (AIC) is the most commonly used criteria, which is a measure of the goodness fit of an estimated statistical model. The best models are then ordered by the criteria of AIC. The AIC is a

measure of fit that penalizes for the number of variables. And AICc works better when the sample sizes are small:

$$\mathrm{AICc} = \mathrm{AIC} + \frac{2\mathrm{K}(\mathrm{K}+1)}{\mathrm{n}-\mathrm{K}-1} \qquad (3)$$

Smaller values indicate better fit and thus the AICc can be used to compare models. Thus, models with the lowest AICc are considered to be the top models, highlighting the included variables and model support relative to other models within 2 AICc.

The model averaging was then proposed for addressing the issues of uncertainty in the choice of probability distribution functions and the biased regression parameters during the model selection process (Burnham and Anderson, 2002). According to the characteristics of our dataset and the smooth functions, the selection of variables was based on: (a)whether they corresponding to the research objectives well; (b) whether they were consistent with the related results in section 3.1 and 3.2; (c) sum of weights; and (d) whether they were collinear to each other. In this study, only the most important five explanatory variables were retained in the final model set to predict the relationships for each element in primary carrier.

5.3 Results

5.3.1 Local meteoric water line characteristics

The LMWLs always deviate from the global meteoric water line (GMWL, defined by $\delta D = 8\ \delta^{18}O+10$) (Craig, 1961), due to various climatic variables. The R square values of both seasonal and annual LMWLs in Nam Co station ranged from 0.95 to 0.99 (Figure 5-2), showing a very good linear relationship between $\delta^{18}O$ and δD on an 8-year temporal scale. The annual LMWL in Nam Co deviated slightly from GMWL, with higher values for both slope (8.34>8) and intercept (17.4>10) (Figure. 5-2A). The slope and intercept values in Nam Co were also higher than the neighboring sites, the Siling Co, Zhimenda, and Wushaoling stations (Table 5-2), as well as the stations in Vietnam and Thailand. These characteristics on an annual scale might correlate to the significant moisture sources from the continental landscape surface, which was always featured by higher isotopic values (Zhang et al., 2019).

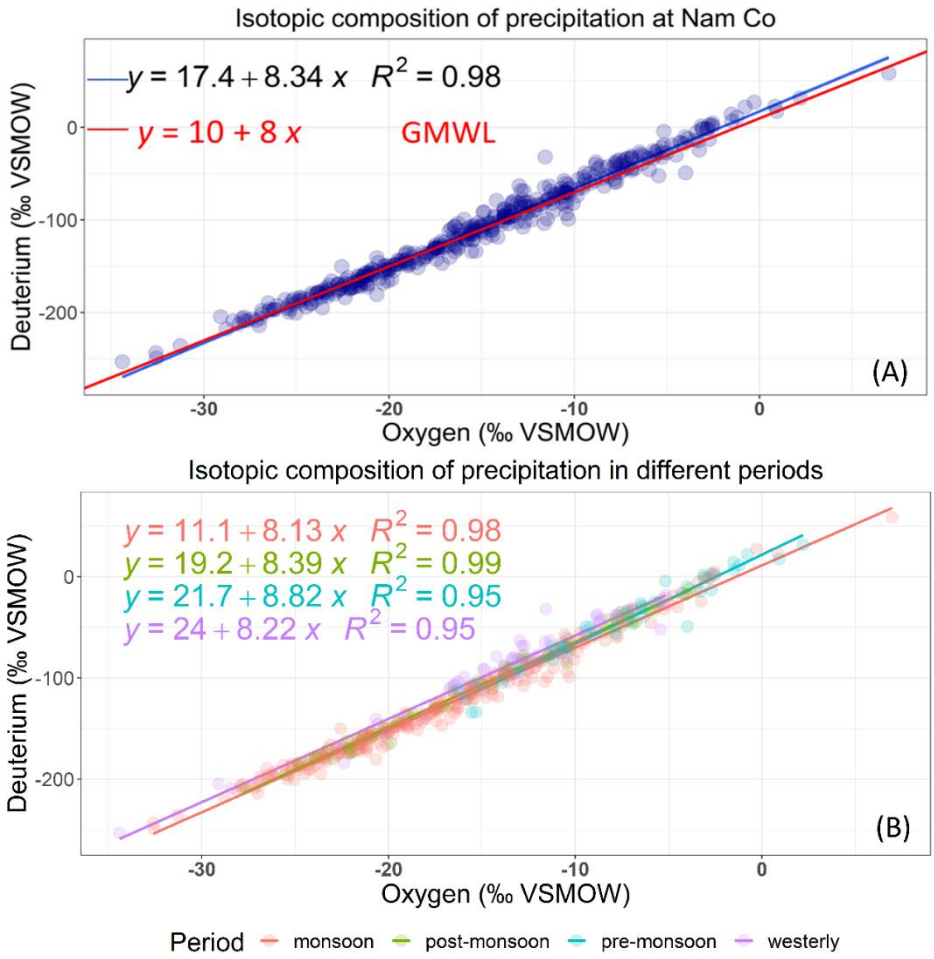


Figure 5- 2. LMWLs in Nam Co station full year (A) and during four seasonal periods (B), including the pre-monsoon (blue color), monsoon (red color), post-monsoon (green color), and Westerly (purple color).

Seasonally, the slope value achieved the highest values in the Pre-monsoon season (8.82), the lowest in the monsoon season (8.13), then it rebounded to 8.39in the post-monsoon, and was restored to 8.22 in the Westerly season (Figure 5-2A, Table 5-2). The slope value is always correlated to the isotopic values of the air masses that change all the time. If the vapor source is relatively constant over a long time, the slope values of LMWLs in this site should be close to 8. The slope values are always above 8 when high d-excess values are associated with lower isotope values. The lowest slope values of LMWL at Nam Co basin in monsoon periods reflects the strong marine contributions to the moisture sources, which brings high rainfall amounts with low d-excess values. The slope of the LMWLs is also regarded as an indicator of the re-evaporative processes below the cloud base (Peng et al., 2010). Notably, this result is different from the neighboring Siling Co Lake where the lowest slope appeared in the post-monsoon period, which might be correlated with the different secondary evaporation process.

Table 5- 2. Typical stations in QTP with LMWL information.

Site	Latitude	Longitude	Elevation	LMWL	Data period	Source
Nam Co	30.75	90.93	4730	$\delta D=8.34\delta^{18}O+17.4$	2005.8-2012.9	This study
				$\delta D=8.82\delta^{18}O+21.7$	pre-monsoon	
				$\delta D=8.13\delta^{18}O+11.1$	monsoon	
				$\delta D=8.39\delta^{18}O+19.2$	post-monsoon	
				$\delta D=8.22\delta^{18}O+24$	Westerly	
Siling Co	30.9	88.7	4681	$\delta D=8.19\delta^{18}O+10.72$	2014.1-2016.12	(Zhang et al., 2019)
Zhimenda	33.01	96.9	3580	$\delta D=8.03\delta^{18}O+11.64$	2016.6-2018.7	(Li et al., 2020)
Wushaoling	37.13	102.57	2800	$\delta D=7.72\delta^{18}O+11.6$	2013.5-2014.8	(Li et al., 2020)
Viernam	-	-	-	$\delta D= 7.56\ \delta^{18}O+7.26$		(Duy et al., 2018)
Thailand	-	-	-	$\delta D= 7.62\ \delta^{18}O+6.42$		(Laonamsai et al., 2021)

The intercept value of LMWL in Nam Co was 11.1 in the monsoon season, close to 10 of the GMWL, but the intercept values were all greater than 19 in other seasons, and the intercept value achieved the highest values in the Westerly season (reaching 24). The changing patterns of intercept are mainly denoted by the moisture sources, evaporative process below the cloud base, as well as the d-excess values (Dansgaard, 1964). The low values of intercept in the monsoon season can be explained by the marine vapor sources and the lowest values of d-excess in this period (Figure 5-3(D) in the next part). The strong sub-cloud re-evaporation and insignificant evaporative isotopic enrichment during the precipitation falling down process may also contribute to the low intercept values.

5.3.2 Environmental effects

5.3.2.1 Altitude effects

In order to study the influence of altitude effects on the stable isotopic compositions of precipitation in the QTP, we selected a profile with three sites that runs from northeast to southwest of the QTP, with altitudes ranging from 2800, 3580, to 4681 m above the sea level (Figure 5-1, Table 5-2). The altitude effects for precipitation $\delta^{18}O$ between Nam Co and Zhimenda was nearly -0.3‰ per 100 m of altitude, and -1.28‰ per 100 m of altitude for

Wushaoling and Zhimenda (Figure 5-2A). Previous studies showed that the altitude effects for precipitation $\delta^{18}O$ ranged from −0.1 to −0.6‰ per 100 m of altitude (Ichiyanagi, 2007; Laonamsai et al., 2021), so we can see the pronounced altitude effect in the QTP. The isotopic values of $\delta^{18}O$ and δD are usually observed to be gradually more negative in high-altitude areas than in the low-lying areas due to the altitude effect (Laonamsai et al., 2021), so the strong altitude effects in the QTP might indicate special landscapes and convective processes. There was little influence of altitude effects to the d-excess at different altitudes.

5.3.2.2 Influence of precipitation amount and temperature

Typically, the isotopic values in precipitation tend to be progressively depleted with an increasing precipitation amount, defined as an amount effect (Dansgaard, 1964). The $\delta^{18}O$ and δD compositions of precipitation at Nam Co area kept decreasing when the amount increased. When the precipitation was less than 2mm, the median value of $\delta^{18}O$ was around -15‰, and it decreased to -20‰ when the precipitation amounts exceeding 10mm (Figure 5-3 B). The median values of δ D showed a similar trend that decreased from -100‰ into -150‰ with the increase of precipitation amount from 2mm to 10mm. There was no significant influence of amount effects to d-excess. When the precipitation amount was greater than 7mm, the decreasing trends of isotopic compositions in precipitation were not obvious (Figure 5-3 B). This indicates that there are some other environmental effects in the Nam Co basin which can reduce the impacts of the precipitation amount effect after reaching this threshold, such as the secondary evaporation (Meng and Liu, 2010; Xiao et al., 2021), advection of evaporated vapor (Cole et al., 1999; Lee and Fung, 2008), or the continental moisture recycling (Krklec et al., 2018).

The temperature effects refer to the relationship between temperature and the $\delta^{18}O$ and δD compositions in precipitation. In the Nam Co area, the $\delta^{18}O$ and δD compositions reached the maximum values of -10‰ and -80‰, respectively, when the temperature ranged from 0 to 5°C (Figure 5-2C). When the temperature was greater than 5°C or less than 0°C, the isotopic contents in precipitation started to become more depleted.

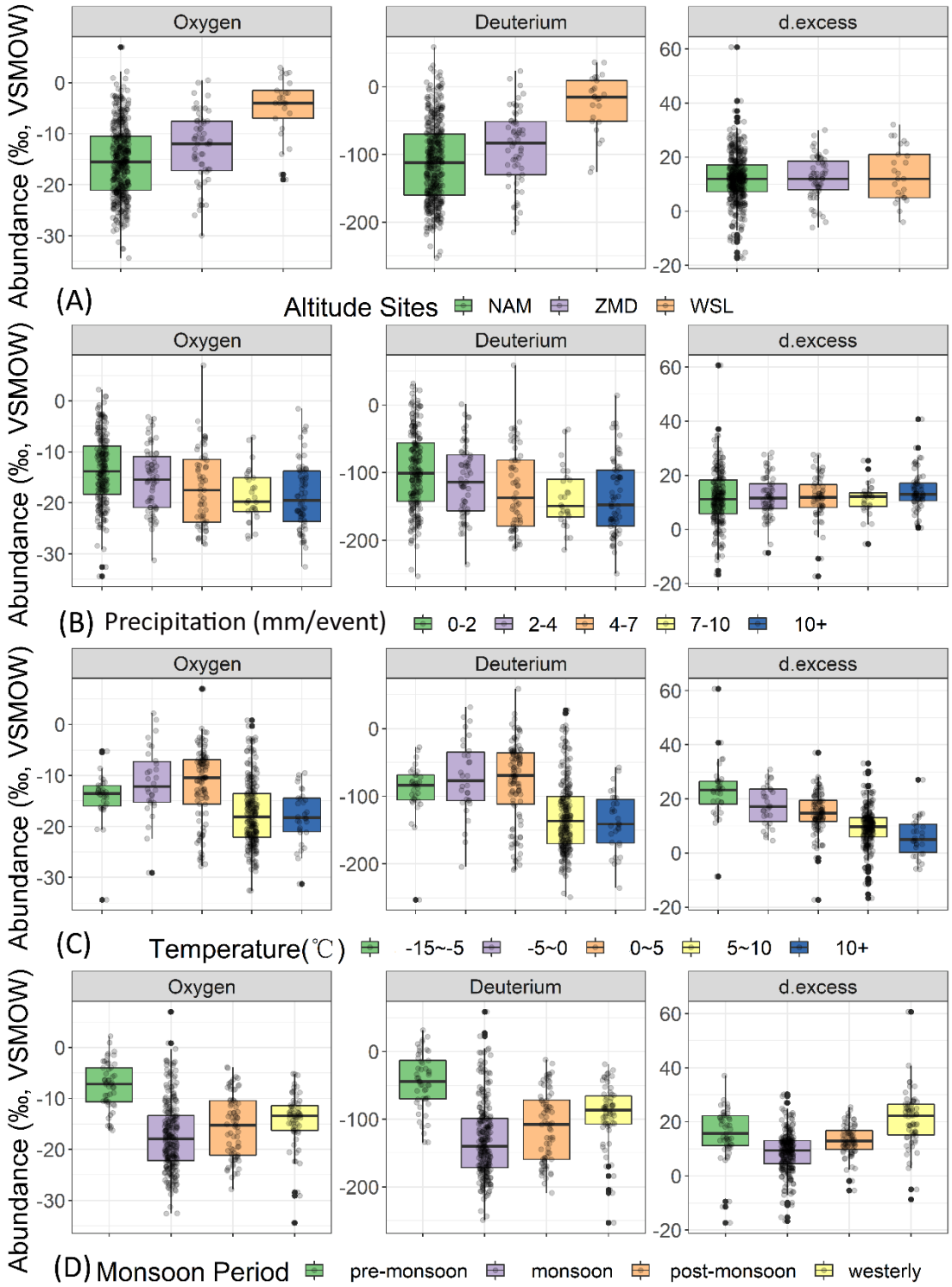


Figure 5- 3. Boxplots showing the influence of environment effects to isotopic variations.

5.3.3 Seasonal patterns of the isotopic variation

5.3.3.1 Isotopic compositions in different seasons

There were distinct seasonal variations in the $\delta^{18}O$, δ D, and d-excess composition of precipitation in Nam Co. The $\delta^{18}O$ and δ D had similar trends in different monsoon periods,

and their abundances were both very enriched in pre-monsoon, then most depleted during the entire monsoon season, and rebounded slightly in the post-monsoon and Westerly periods (Figure. 5-3D). The median value of d-excess also reached the lowest value in the monsoon season, but the highest value appeared in the Westerly season, rather than in the post-monsoon period.

The seasonal variations of isotopic compositions in precipitation correlated to the sub-cloud process and local climatic systems. The heavier isotopes in the pre-monsoon at the Nam Co basin are the consequence of a relatively low precipitation amount that will scavenge heavy isotopes from the ambient atmosphere. With the increasing precipitation amount in the monsoon season, the continuous rainout process under the cloud base made the isotopic compositions more depleted, thereby making the precipitation isotopic compositions even lighter (Hughes and Crawford, 2013). For the high content moisture, the height of reaching the dew point for the rising local air mass is relatively low, with low $\delta^{18}O$ and high d-excess. On the contrary, under relatively dry conditions, low precipitation and re-evaporation of raindrops will result in high $\delta^{18}O$ and low d-excess (Stewart, 1975). The slightly increasing isotopic values in the post-monsoon season can be explained by the weakening of convections along the trajectory with the retreat of monsoon. The weakening rain out effects driven by the decreasing precipitation amounts and precipitation events in post-monsoon and Westerly seasons, will keep the isotopic compositions below the cloud base at a heavier level than that in monsoon season. These results are consistent with the seasonal variations in the neighboring Siling Co basin (Zhang et al., 2019).

5.3.3.2 Top models for isotopes

Both the annual and seasonal top models for isotopes, in terms of the lowest AICc values by model selection are presented in Table 5-3, as well as their important explanatory variables. Annually, the top models for both $\delta^{18}O$ and δD at Nam Co Lake can be explained by the combinations of precipitation amount, minimum temperature, and wind speed, except that pressure is only an explanatory variable for $\delta^{18}O$ (Table 5-3). $\delta^{18}O$ and δD showed negative correlations with the precipitation amount, minimum temperature, but positive with wind speed. The top model of d-excess was explained by humidity, and three temperature variables, and d-excess had a positive relationship with the maximum temperature, but a negative relationship with temperature and minimum temperature. The ever-present significant influence of continental moisture throughout the year at the QTP might explain the explanatory variables of humidity and temperature in the d-excess top model (Zhang et al., 2019).

Table 5- 3. Top models identified through model selection based on AICc for isotopic compositions in precipitation at Nam Co basin. df= Degrees of freedom. Weight= Akaike weight.

indicators	Top models	df	loglik	AICc	weight	Period
Oxygen	(-0.74) precipitation+(-0.35) pressure+(-2.35) T_min+(0.85) W_speed-15.46	6	-1419.32	2850.84	0.26	full year
	(-1.5131) P_duration-7.5	3	-141.98	290.49	0.29	pre-monsoon
	(-1.37) humidity+(-2.15) precipitation+(-1.06) pressure+(-0.62) T_max+(-2.13) T_min+(0.4) W_direction-17.31	8	-779.53	1575.67	0.21	monsoon
	(1.99) pressure+(-1.07) temperature -15.54	4	-229.68	467.95	0.41	post-monsoon
	(0.69) W_speed-14.31	3	-192.04	390.5	0.23	Westerly
Deuterium	(-6.02) precipitation+(-22.18) T_min+(7.25) W_speed-111.56	5	-2328.52	4667.19	0.21	full year
	(5.2) humidity+(-16.28) P_duration-44.45	4	-247.71	504.33	0.21	pre-monsoon
	(-12.22) humidity+(-16.32) precipitation+(-8.57) pressure+(-5.47) T_max+(-18.84) T_min+(3.18) W_direction-129.54	8	-1296.38	2609.38	0.14	monsoon
	(16.01) pressure+(-13.21) temperature-111.17	4	-384.49	777.57	0.44	post-monsoon
	(2.91) W_speed+(1.03) P_duration-93.72	2	-325.4	655	0.43	Westerly
d-excess	(-0.85) humidity+(1.27) T_max+(-0.54) T_min+(-4.88) temperature+ 12.1	6	-1480.07	2974.41	0.08	full year
	(3.66) humidity+(-3.07) P_duration+(4.35) T_max+(-2.08) T_min+(-2.46) temperature+ 15.51	7	-168.66	354.06	0.27	pre-monsoon
	(-1.18) humidity+(-2.27) temperature+ 8.92	4	-821.17	1650.51	0.12	monsoon
	(-2.69) temperature+13.15	3	-221.99	450.32	0.31	post-monsoon
	(-1.96) humidity+(-3.79) temperature+(-1.56) W_direction+20.78	5	-227.34	465.75	0.22	Westerly

From the top models of the seasonal scales, different local climatic factors worked in different seasons. In the pre-monsoon season, all $\delta^{18}O$, δD, and d-excess showed a negative relationship with the precipitation duration, which can be explained by the rain out effects. The precipitation is featured by fewer events and lower amounts in the pre-monsoon period, and the shorter the periods of the precipitation duration, the heavier isotopic the signatures in precipitation are. There were more explanatory variables in monsoon season, which reflected

the more complicated climatic conditions at the Nam Co area. For example, both the top models of $\delta^{18}O$ and δD were interpreted by the combinations of precipitation amount, humidity, pressure, maximum and minimum temperature, and wind direction. While during the post-monsoon period, $\delta^{18}O$ and δD achieved the top models when explained by pressure and temperature, however the d-excess was only related to the temperature in the top model. Notably, all three isotopic signatures were correlated to the wind variables in their top models, and the dominant Westerly system might offer a plausible explanation for this pattern. For examples, $\delta^{18}O$ and δD had a positive relationship with wind speed, with the correlation coefficients being 0.69 and 2.91, respectively, while the d-excess showed a negative relationship with wind directions (-1.56).

In addition, it is found that d-excess was affected by temperature throughout the year, and their relations are all negative. In chronological order of pre-monsoon to Westerly, the coefficients were -2.46, -2.27, -2.69, and -3.79, respectively. The significant negative correlations between precipitation d-excess values and temperature at Nam Co indicated the signals of moisture evaporated from the continent.

5.3.4 Relationships with annual local factors

5.3.4.1 Local factor analysis by RDA

According to the results of the Redundancy analysis, the precipitation duration, pressure, maximum temperature, and wind speed played an important role in the first axis, by a contribution of 25.66% (Figure 5-4). By contrast, the precipitation amount and humidity worked for the second axis, with the contribution value of 13.67%. The isotopic signature in the monsoon periods was consistent with most of the local climatic variables, indicating the active land-atmosphere interactions and complicated climate conditions in this season. The wind patterns dominated in the pre-monsoon season, while precipitation amount was the most significant variable during the Westerly system.

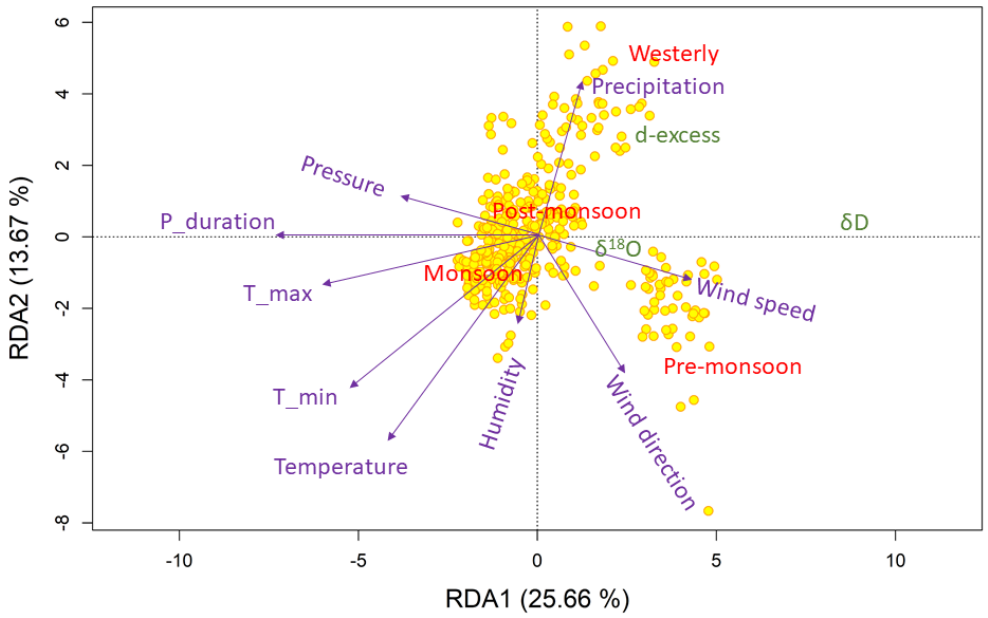


Figure 5- 4. Tri-plots of Redundancy analysis (RDA) of the relationships between isotopic signatures and explanatory variables.

Table 5- 4. Importance analysis of the local climatic predictors.

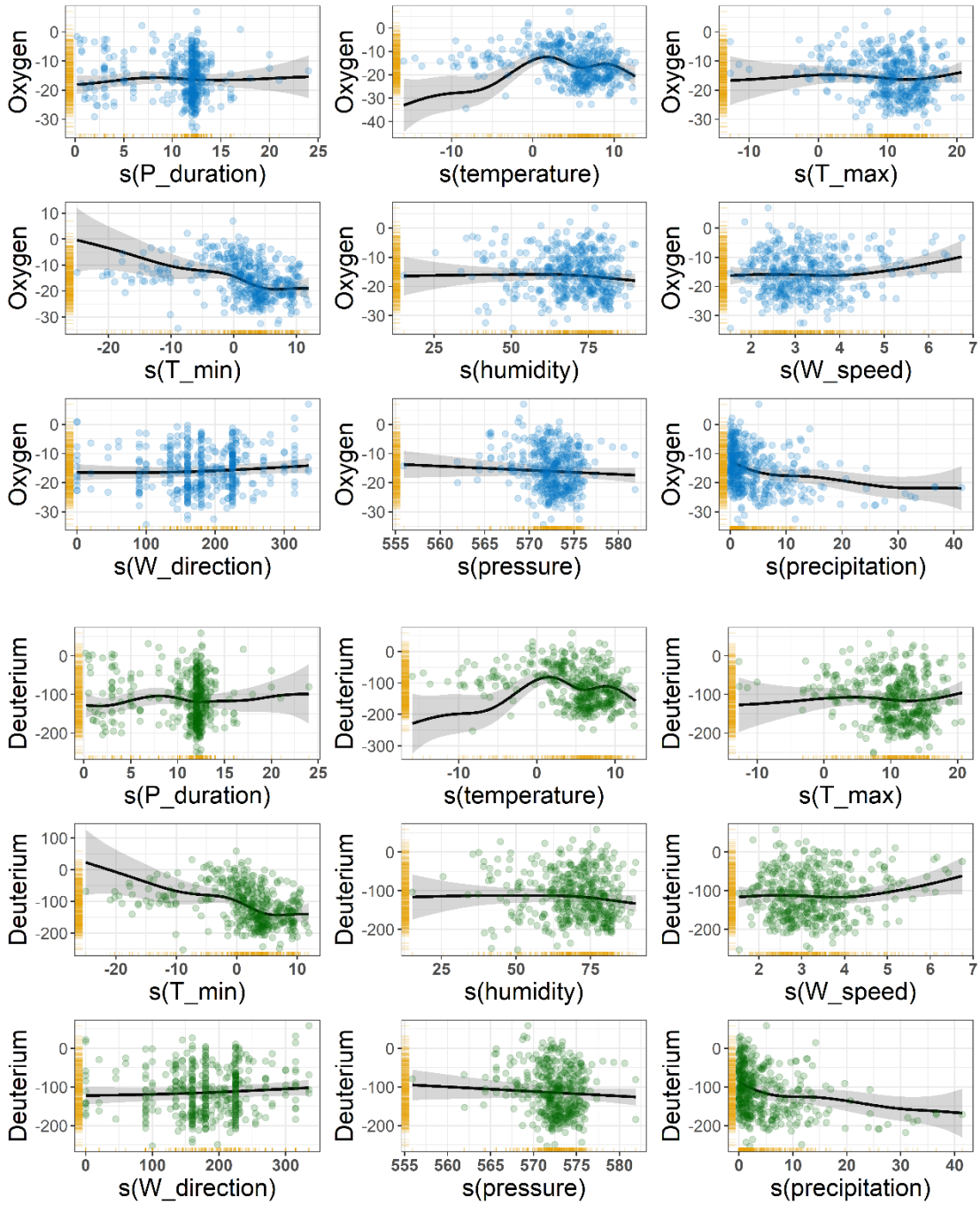
Tracers	weighs	s(precipitation)	s(T_min)	s(temperature)	s(W_speed)	s(T_max)	s(W_direction)	s(humidity)	s(pressure)	s(P_duration)
Oxygen	sum	1	1	1	1	0.65	0.63	0.59	0.37	0.05
	N	15	15	15	15	9	9	9	7	1
Deuterium	sum	1	1	1	1	0.43	0.41	0.84	0.27	0.07
	N	3	3	3	10	5	4	8	3	1
d-excess	sum	1	1	1	0.01	1	0.22	1	0.21	0.01
	N	3	3	3	1	3	1	3	1	1

5.3.4.2 Relationship analysis by GAM

According to the results of generalized additive models (GAM), the $\delta^{18}O$ and δD showed very similar response and variation patterns to all these nine local climatic factors (Figure 5-5). For instance, the $\delta^{18}O$ and δD signature showed a decreasing trend with the increasing of precipitation amounts, while d-excess showed a first slightly increasing and then decreasing trend, with the peak value of nearly -20‰ when the precipitation amount was 25cm.

There was no significant influence of precipitation duration on isotopic signatures on an annual temporal scale. When the precipitation duration was around 8 hours, $\delta^{18}O$ and δD reached a relatively enriched value of -15‰ and -100‰, respectively. When the duration was less or longer than 8 hours, the isotopic compositions started to show a negatively depleted trend. This was not consistent with the results on the seasonal variations of precipitation

isotopes, which might be related to that the monsoon precipitation, or other variables that weakened the effects of the amount on the annual scale. Generally, it can be found that the longer of precipitation duration, the lighter of the isotopic compositions.



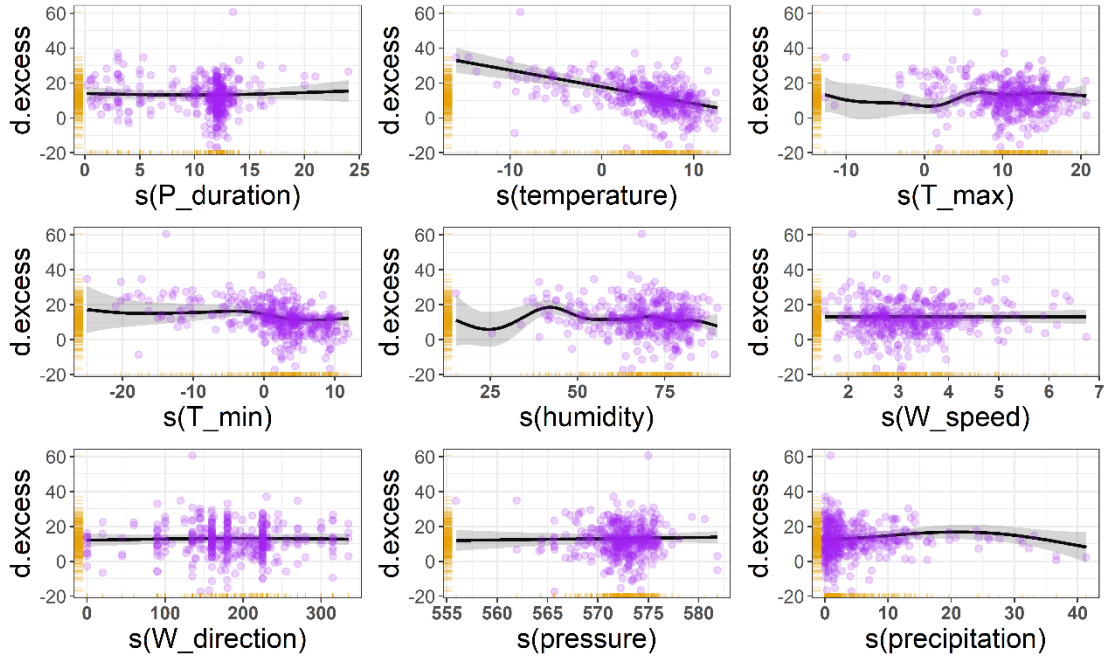


Figure 5- 5. Smoothed fits of relationships between $\delta^{18}O$, δD, d-excess, and explanatory variables.

Humidity was generally negatively correlated with $\delta^{18}O$ and δD, and their compositions showed an obvious downward trend when humidity was greater than 75%. Humidity presented a 'W' amplitude trend for d-excess. When humidity was less than 25%, d-excess dropped from 15‰ to 5‰, and then kept increasing until the maximum value of 20 ‰ when humidity was around 45%, then fell again, and remained flat after the humidity was greater than 50%. This complex relationship suggests some secondary fractionation processes at Nam Co, such as sub-cloud evaporation or secondary evaporation from open water bodies.

The average temperature showed an "M" trend effect on $\delta^{18}O$ and δD, but it has a negative linear correlation with d-excess. There was a totally negative relationship between the minimum temperature with the $\delta^{18}O$ and δD signatures, while the maximum temperature showed a similar trend with all these three tracers.

In terms of the wind speed, when the wind speed was less than 5m/s, it basically had very little effect on $\delta^{18}O$ and δD, and values remained at about -15‰, and -100‰, respectively. As the wind speed increased, the isotopic contents showed an increasing trend.

5.4 Conclusions

The difficulties of collecting continuous precipitation samples over a large time period are a challengeable obstacle restricting the study of relationships between isotopic compositions in

precipitation and potential local climatic variables in the monsoon-Westerly transition regions in the QTP. A total of 430 precipitation samples together with related local climatic variables over an 8-year were presented for the first time in this study to help discover the mystery of the hydrological process at Nam Co area. The model selection and model averaging methods were introduced to explore the quantitative relationships between them. Results found that there was a very good linear relationship between $\delta^{18}O$ and δ D, and these two isotopic tracers showed very similar response and variation patterns to all nine local climatic factors according to the results of GAM.

On the annual scale, both the slope and intercept of the LMWL at Nam Co basin were higher than those of the GMWL, and these characteristics suggested the significant moisture sources from continental landscape surface throughout the entire year at the transition areas of the QTP. The important explanatory variables of humidity and three temperature factors in the top model of d-excess also supported this conclusion. Seasonally, the lowest slope and intercept values of LMWL during the monsoon periods indicated the marine contributions of moisture. The lighter isotopic values in the monsoon season also reflected the strong rain out process by the high precipitation amounts originated from the ocean. We also found that different local climatic factors played important roles in different seasons according to the results of top models. The wind variables were important variables to explain the isotopic tracers during the Westerly season.

The EMMA and SDM models constitute an advancement in quantitatively interpreting the relationships between isotopic signatures in precipitation and their local climatic variables, which can supplement or precede more complex studies of hydrological cycles utilizing isotope tools. The long-term monitoring data sets in this study will enrich a multi-dimensional system of the hydrological cycles in the transition domain areas of QTP, and benefit the related model development and research.

5.5 References

Abbott, M., Lini, A., Bierman, P., 2000. δ18O, δD and 3H measurements constrain groundwater recharge patterns in an upland fractured bedrock aquifer, Vermont, USA. Journal of Hydrology, 228(1-2): 101-112.

Allen, S.T. et al., 2019. Global sinusoidal seasonality in precipitation isotopes. Hydrology and Earth System Sciences, 23(8): 3423-3436.

Ansari, M.A., Noble, J., Deodhar, A., Kumar, U.S., 2020. Atmospheric factors controlling the stable isotopes (δ18O and δ2H) of the Indian summer monsoon precipitation in a drying region of Eastern India. Journal of Hydrology, 584: 124636.

Balagizi, C.M., Kasereka, M.M., Cuoco, E., Liotta, M., 2018. Influence of moisture source dynamics and weather patterns on stable isotopes ratios of precipitation in Central-Eastern Africa. Science of the Total Environment, 628: 1058-1078.

Bowen, G.J., Cai, Z., Fiorella, R.P., Putman, A.L., 2019. Isotopes in the water cycle: regional- to global-scale patterns and applications. Annual Review of Earth and Planetary Sciences, 47: 453-479.

Burnham, K.P., Anderson, D.R., 2002. A practical information-theoretic approach. Model selection and multimodel inference, 2.

Bush, R.T., Berke, M.A., Jacobson, A.D., 2017. Plant water δD and δ18O of tundra species from West Greenland. Arctic, Antarctic, and Alpine Research, 49(3): 341-358.

Cai, Z., Tian, L., Bowen, G.J., 2017. ENSO variability reflected in precipitation oxygen isotopes across the Asian Summer Monsoon region. Earth and Planetary Science Letters, 475: 25-33.

Cole, J.E., Rind, D., Webb, R.S., Jouzel, J., Healy, R., 1999. Climatic controls on interannual variability of precipitation δ18O: Simulated influence of temperature, precipitation amount, and vapor source region. Journal of Geophysical Research: Atmospheres, 104(D12): 14223-14235.

Craig, H., 1961. Isotopic variations in meteoric waters. Science, 133(3465): 1702-1703.

Crawford, J., Hollins, S.E., Meredith, K.T., Hughes, C.E., 2017. Precipitation stable isotope variability and subcloud evaporation processes in a semi-arid region. Hydrological Processes, 31(1): 20-34.

Dansgaard, W., 1964. Stable isotopes in precipitation. tellus, 16(4): 436-468.

Duy, N.L., Heidbüchel, I., Meyer, H., Merz, B., Apel, H., 2018. What controls the stable isotope composition of precipitation in the Mekong Delta? A model-based statistical approach. Hydrology and Earth System Sciences, 22(2): 1239-1262.

Gao, J., He, Y., Masson-Delmotte, V., Yao, T., 2018. ENSO effects on annual variations of summer precipitation stable isotopes in Lhasa, southern Tibetan Plateau. Journal of Climate, 31(3): 1173-1182.

Hughes, C.E., Crawford, J., 2013. Spatial and temporal variation in precipitation isotopes in the Sydney Basin, Australia. Journal of Hydrology, 489: 42-55.

Ichiyanagi, K., 2007. Studies and applications of stable isotopes in precipitation. Journal of Japanese association of hydrological sciences, 37(4): 165-185.

Ishizaki, Y. et al., 2012. Interannual variability of H218O in precipitation over the Asian monsoon region. Journal of Geophysical Research: Atmospheres, 117(D16).

Jouzel, J., Merlivat, L., 1984. Deuterium and oxygen 18 in precipitation: Modeling of the isotopic effects during snow formation. Journal of Geophysical Research: Atmospheres, 89(D7): 11749-11757.

Kang, S., Yang, Y., Zhu, L., Ma, Y., Ma, W., 2011. Modern environmental processes and changes in the Nam Co basin, Tibetan Plateau. China: Beijing Meteorological Press.

Kang, S., Yi, Y., Xu, Y., Xu, B., Zhang, Y., 2017. Water Isotope framework for lake water balance monitoring and modelling in the Nam Co Basin, Tibetan Plateau. Journal of Hydrology: Regional Studies, 12: 289-302.

Keil, A. et al., 2010. Hydrological and geomorphological basin and catchment characteristics of Lake Nam Co, South-Central Tibet. Quaternary International, 218(1-2): 118-130.

Kreutz, K.J., Wake, C.P., Aizen, V.B., Cecil, L.D., Synal, H.A., 2003. Seasonal deuterium excess in a Tien Shan ice core: influence of moisture transport and recycling in Central Asia. Geophysical Research Letters, 30(18).

Krklec, K., Domínguez-Villar, D., Lojen, S., 2018. The impact of moisture sources on the oxygen isotope composition of precipitation at a continental site in central Europe. Journal of hydrology, 561: 810-821.

Kumar, A., Tiwari, S.K., Verma, A., Gupta, A.K., 2018. Tracing isotopic signatures (δD and δ18O) in precipitation and glacier melt over Chorabari Glacier–Hydroclimatic inferences for the Upper Ganga Basin (UGB), Garhwal Himalaya. Journal of Hydrology: Regional Studies, 15: 68-89.

Kurita, N., Fujiyoshi, Y., Nakayama, T., Matsumi, Y., Kitagawa, H., 2015. East Asian Monsoon controls on the inter-annual variability in precipitation isotope ratio in Japan. Climate of the Past, 11(2): 339-353.

Kurita, N., Ichiyanagi, K., Matsumoto, J., Yamanaka, M.D., Ohata, T., 2009. The relationship between the isotopic content of precipitation and the precipitation amount in tropical regions. Journal of Geochemical Exploration, 102(3): 113-122.

Lacelle, D., Fontaine, M., Forest, A.P., Kokelj, S., 2014. High-resolution stable water isotopes as tracers of thaw unconformities in permafrost: A case study from western Arctic Canada. Chemical Geology, 368: 85-96.

Laonamsai, J., Ichiyanagi, K., Kamdee, K., Putthividhya, A., Tanoue, M., 2021. Spatial and temporal distributions of stable isotopes in precipitation over Thailand. Hydrological Processes, 35(1): e13995.

Lee, J.E., Fung, I., 2008. "Amount effect" of water isotopes and quantitative analysis of post-condensation processes. Hydrological Processes: An International Journal, 22(1): 1-8.

Li, C., Kang, S., Zhang, Q., Kaspari, S., 2007. Major ionic composition of precipitation in the Nam Co region, Central Tibetan Plateau. Atmospheric Research, 85(3-4): 351-360.

Li, Z.-J., Li, Z.-X., Song, L.-L., Jin-Zhu, M., 2020. Characteristic and factors of stable isotope in precipitation in the source region of the Yangtze River. Agricultural and Forest Meteorology, 281: 107825.

Li, Z. et al., 2016a. Stable isotope composition of precipitation in the south and north slopes of Wushaoling Mountain, northwestern China. Atmospheric Research, 182: 87-101.

Li, Z. et al., 2016b. Quantitative evaluation on the influence from cryosphere meltwater on runoff in an inland river basin of China. Global and Planetary Change, 143: 189-195.

Li, Z. et al., 2014. Study on the contribution of cryosphere to runoff in the cold alpine basin: A case study of Hulugou River Basin in the Qilian Mountains. Global and Planetary Change, 122: 345-361.

Liu, W. et al., 2021. Rapid expansion of lakes in the endorheic basin on the Qinghai-Tibet Plateau since 2000 and its potential drivers. CATENA, 197: 104942.

Liu, Z., Yao, Z., Wang, R., Yu, G., 2020. Estimation of the Qinghai-Tibetan Plateau runoff and its contribution to large Asian rivers. Science of The Total Environment, 749: 141570.

Lone, S.A., Jeelani, G., Deshpande, R.D., Mukherjee, A., 2019. Stable isotope (δ18O and δD) dynamics of precipitation in a high altitude Himalayan cold desert and its surroundings in Indus river basin, Ladakh. Atmospheric Research, 221: 46-57.

Lyon, S.W., Desilets, S.L., Troch, P.A., 2009. A tale of two isotopes: differences in hydrograph separation for a runoff event when using δD versus δ18O. Hydrological Processes: An International Journal, 23(14): 2095-2101.

Meng, Y.-C., Liu, G.-D., 2010. Effect of below-cloud secondary evaporation on the stable isotopes in precipitation over the Yangtze River basin. Advances in water science, 21(3): 327-334.

Ogrinc, N., Kanduč, T., Stichler, W., Vreča, P., 2008. Spatial and seasonal variations in δ18O and δD values in the River Sava in Slovenia. Journal of Hydrology, 359(3-4): 303-312.

Peng, T.-R. et al., 2010. Stable isotopic characteristic of Taiwan's precipitation: A case study of western Pacific monsoon region. Earth and Planetary Science Letters, 289(3-4): 357-366.

Price, R.M., Swart, P.K., Willoughby, H.E., 2008. Seasonal and spatial variation in the stable isotopic composition (δ18O and δD) of precipitation in south Florida. Journal of Hydrology, 358(3-4): 193-205.

Putman, A.L., Fiorella, R.P., Bowen, G.J., Cai, Z., 2019. A global perspective on local meteoric water lines: Meta-analytic insight into fundamental controls and practical constraints. Water Resources Research, 55(8): 6896-6910.

Risi, C., Bony, S., Vimeux, F., 2008a. Influence of convective processes on the isotopic composition (δ18O and δD) of precipitation and water vapor in the tropics: 2. Physical interpretation of the amount effect. Journal of Geophysical Research: Atmospheres, 113(D19).

Risi, C. et al., 2008b. What controls the isotopic composition of the African monsoon precipitation? Insights from event-based precipitation collected during the 2006 AMMA field campaign. Geophysical Research Letters, 35(24).

Shen, S., Lau, K.-M., 1995. Biennial oscillation associated with the East Asian summer monsoon and tropical sea surface temperatures. Journal of the Meteorological Society of Japan. Ser. II, 73(1): 105-124.

Sjostrom, D.J., Welker, J.M., 2009. The influence of air mass source on the seasonal isotopic composition of precipitation, eastern USA. Journal of Geochemical Exploration, 102(3): 103-112.

Soderberg, K. et al., 2013. Using atmospheric trajectories to model the isotopic composition of rainfall in central Kenya. Ecosphere, 4(3): 1-18.

Song, C., Huang, B., Richards, K., Ke, L., Hien Phan, V., 2014. Accelerated lake expansion on the Tibetan Plateau in the 2000s: Induced by glacial melting or other processes? Water Resources Research, 50(4): 3170-3186.

Stewart, M.K., 1975. Stable isotope fractionation due to evaporation and isotopic exchange of falling waterdrops: Applications to atmospheric processes and evaporation of lakes. Journal of Geophysical Research, 80(9): 1133-1146.

Tan, M., 2014. Circulation effect: response of precipitation δ 18 O to the ENSO cycle in monsoon regions of China. Climate Dynamics, 42(3-4): 1067-1077.

Tian, L., Masson-Delmotte, V., Stievenard, M., Yao, T., Jouzel, J., 2001a. Tibetan Plateau summer monsoon northward extent revealed by measurements of water stable isotopes. Journal of Geophysical Research: Atmospheres, 106(D22): 28081-28088.

Tian, L. et al., 2007. Stable isotopic variations in west China: A consideration of moisture sources. Journal of Geophysical Research: Atmospheres, 112(D10).

Tian, L., Yao, T., Numaguti, A., Duan, K., 2001b. Relation between stable isotope in monsoon precipitation in southern Tibetan Plateau and moisture transport history. Science in China Series D: Earth Sciences, 44(1): 267-274.

Uemura, R. et al., 2012. Factors controlling isotopic composition of precipitation on Okinawa Island, Japan: Implications for paleoclimate reconstruction in the East Asian Monsoon region. Journal of Hydrology, 475: 314-322.

Wang, J. et al., 2009. Investigation of bathymetry and water quality of Lake Nam Co, the largest lake on the central Tibetan Plateau, China. Limnology, 10(2): 149-158.

Wang, S. et al., 2021. Reason Analysis of the Jiwenco Glacial Lake Outburst Flood (GLOF) and Potential Hazard on the Qinghai-Tibetan Plateau. Remote Sensing, 13(16): 3114.

Wassenaar, L.I., Van Wilgenburg, S.L., Larson, K., Hobson, K.A., 2009. A groundwater isoscape (δD, δ18O) for Mexico. Journal of Geochemical Exploration, 102(3): 123-136.

Wu, Y., Zhu, L., 2008. The response of lake-glacier variations to climate change in Nam Co Catchment, central Tibetan Plateau, during 1970–2000. Journal of Geographical Sciences, 18(2): 177-189.

Xiao, H. et al., 2021. Sub-Cloud Secondary Evaporation in Precipitation Stable Isotopes Based on the Stewart Model in Yangtze River Basin. Atmosphere, 12(5): 575.

Yao, T. et al., 2013. A review of climatic controls on δ18O in precipitation over the Tibetan Plateau: Observations and simulations. Reviews of Geophysics, 51(4): 525-548.

Yin, L. et al., 2011. Isotopes (δD and δ 18 O) in precipitation, groundwater and surface water in the Ordos Plateau, China: implications with respect to groundwater recharge and circulation. Hydrogeology Journal, 19(2): 429-443.

Yu, W. et al., 2006. Oxygen-18 isotopes in precipitation on the eastern Tibetan Plateau. Annals of Glaciology, 43: 263-268.

Yuan, Y. et al., 2017. The 2016 summer floods in China and associated physical mechanisms: A comparison with 1998. Journal of Meteorological Research, 31(2): 261-277.

Zhang, G., Xie, H., Yao, T., Kang, S., 2013. Water balance estimates of ten greatest lakes in China using ICESat and Landsat data. Chinese Science Bulletin, 58(31): 3815-3829.

Zhang, T. et al., 2019. Controls of stable isotopes in precipitation on the central Tibetan Plateau: A seasonal perspective. Quaternary international, 513: 66-79.

Chapter 6. Biotic, environmental and human footprint variables to explain the bioaccumulation mechanism with elemental signatures from waterfowl feathers

Abstract

Studying the bioaccumulation mechanism of the chemical elements in organisms is an important basis for biological conservation. We hypothesized that the bioaccumulation of toxic elements depends on the human activities, while essential elements are more closely related to the environmental conditions. To test this hypothesis, feathers from seven duck species were sampled as the research entities, and the concentrations of 16 toxic and essential elements were measured. Then the background geochemical sediments, a set of environmental and human footprint factors based on the biological behaviors were collected as the explanatory variables. The relationships between elemental concentrations in feather and explanatory variables were evaluated through redundancy analysis, correlation analysis and model selection. Studies found that environmental variables had a greater effect on essential elements, while human footprint variables had a great impact on both the toxic and essential ones. Through the top models, most of the elements were basically affected by both environmental and human footprint variables, but the essential element Cu was only affected by environmental factors, and the toxic elements Al, Cr, Li, and Pb, as well as essential Mn were only affected by human activities. The elements in the feathers were finally quantitatively explained by the biotic and abiotic variables, which proves that waterfowl feathers are a powerful tool to predict and monitor environmental pollution. This research provides new perspectives for a more accurate understanding of bioaccumulation mechanisms.

Key words: bioaccumulation mechanisms; duck feathers; bioaccumulation ratio; toxic elements; essential elements; environmental variables; human footprint; redundancy analysis; model selection

Highlights:

* Combination of elemental concentration and bioaccumulation ratio of duck feathers helps provide more valuable biological behavior information

* The elemental bioaccumulation is affected by both biotic and abiotic variables

* Bird feathers are a nonlethal tool with great promotional significance

* Quantitatively interpreting the relationship between element and explanatory variables can help clearly understand the bioaccumulation mechanisms

6.1 Introduction

Chemical elements are ubiquitous and persistent in different chemical forms throughout the environment (Ali et al. 2019). Their transport and accumulation processes will involve in many environmental compartments, like the atmosphere, lithosphere, hydrosphere, and biosphere (Innangi et al. 2019). A clear understanding of these mechanisms can help to effectively control the threat of elements to the environment.

These elements have been divided into many types according to the different purposes of research. For instance, from the perspective of environmental contaminants, the chemical elements are often divided into trace elements (Abdullah et al. 2015), heavy metals (Frantz et al. 2012; Furtado et al. 2019), metals and metalloids (Badry et al. 2019; López-Perea et al. 2019), etc. In this study, the elements were classified into two categories of toxic and essential elements, according to whether they are necessary to maintaining the human body. Some elements are essential and nutrients for the proper functioning of the human body (Prashanth et al. 2015), including Ca, Co, Cu, Fe, Mg, and Zn, etc., while some elements are toxic to the human body, interfere with its functioning and undermine health, and a small amount could be fatal (Dudka & Miller 1999), like As, Cd, Pb, Al, Li and Hg. These toxic elements have no known physiological functions, but are potentially fatal. However, some essential elements, if the intake exceeds the standard, will threaten health, such as Mn, Mo (0.05 μg/ day), Cr (50-200 μg/ day). In addition, the compounds of some essential elements can be very harmful, such as the hexavalent chromium compounds (Costa & Klein 2006).

Elements will eventually bioaccumulate or bio-magnify in the biosphere, including human beings (Innangi et al. 2019). Understanding the mechanisms of element transportation and accumulation can help to monitor human health and environment conservation but the accumulation of chemical elements in organisms is an extremely complicated process, which will be affected by both the biotic and abiotic factors. Their spread and accumulation will be accelerated, enhanced, and strengthened directly or indirectly by anthropogenic activity, and their relationship with the environmental background is more complicated. We hypothesized that the essential elements are more correlated with environmental factors, while the toxic elements are mainly sourced from human activity (Step 1 in Figure 6-1). To test the hypothesis, the prerequisite is to find a reliable entity that must be closely connected with both the environment and human activity.

Birds are very good potential candidate to verify the above hypothesis. Birds are more sensitive to the surroundings than the human beings and other animals (Wretenberg et al. 2010). The foraging, breeding and other biological behaviors will be in close contact with the environment, which will be eventually recorded in the bird's body. Colonial birds' nests or

roosts are in proximity at a particular location (Brandis et al. 2018), and this provides possibility for specific areas of study. In some areas where there is shotgun hunting, the birds that select hard grit for grinding food are extremely good sentinels to monitor the Pb contamination (Newth et al. 2013), as the grits are similar in weight and size to the shotgun pellets that deposit heavily in their habitats. Bird feathers are outstanding to trace the element bioaccumulation mechanism. Compared with some lethal sampling research methods like blood, liver, muscle, kidney and other internal organs (López-Perea et al. 2019), feathers are a nondestructive method. In addition, when compared with other non-invasive method such as eggshells, feathers can accumulate and record more information over a longer period (Abbasi et al. 2015). Previous bird feather bioaccumulation research covers wide trophic levels from herbivore to carnivore predators, as well as being over a large spatial range, even including Antarctica (Metcheva et al. 2011), and long-term temporal records from the museum feathers (Frederick et al. 2004). The distribution map of sampling sites for the elemental bioaccumulation research using bird feathers is displayed in Step2 ② in Figure 6-1, according to literature review. Some methods for feather monitoring have also received good attention and development recently (Peterson et al. 2019).

Elemental bioaccumulation mechanism by bird feathers have been investigated in the previous studies for addressing three main questions in the field: (i) What is the threshold value of the elemental concentration that can pose a health threat to birds; (ii) What are the potential explanatory variables that drive elemental bioaccumulation processes, at the perspectives of both biotic factors and abiotic environment; and (iii) How to quantitatively interpret the relationship between element and the explanatory variables. Key papers including their strengths and limitations for these essential research questions are summarized in Table 1. For example, previous studies found that the threshold of adverse effect with Se toxicity in birds was 1000 mg/kg (López-Perea et al. 2019); a safety threshold for Cd was set at 2 mg/kg (Burger & Gochfeld 1995); and there is no apparent reproduction impairment at feather Hg concentrations of 40 mg/kg (Weech et al. 2006). In addition, the introduction and development of statistical methodologies, like PCA, ANOVAs and correlation analysis, play a very good role in analyzing and explaining influencing factors (Frantz et al. 2012; Furtado et al. 2019). However, it is very difficult to uniformly delineate a threshold value that is harmful to birds, due to the fact that there are too many species of birds at various trophic levels, and the geochemical background values are always distinctly different in different regions.

As a result, although these previous studies have made significant contributions (details in the literature review of Table 1), there are four important ‘gaps’ in the current knowledge remaining to be solved. (i) The elemental geochemical background is ignored, and most of the

current bioaccumulation research on feathers focused on independent elemental concentration, and tries to give the threshold of different elements in bird feather (López-Perea et al. 2019). In addition, the previous research about the bioaccumulation ratio focused on blood, hair, feces, urine and other organs, and only a few studies focused on the feather bioaccumulation ratio. (ii) The essential and toxic elements were not discussed separately, and compared with toxic elements, heavy metals, etc., the information that essential elements can indicate has rarely been analyzed. (iii) In terms of the explanatory variables, previous research focused on human activities and ignored the environment variables. Although some studies involved the environment, they were neither very systematic or comprehensive. (iv) Previous research stayed at the qualitative stage, with few quantitative analyses.

Therefore, considering the above gaps, it is better to sample feathers from a kind of bird species that is best to be distributed in multiple trophic levels, and collect representative explanatory variables based on their avian biological behaviors. Ducks are outstanding from other waterfowl species due to: (i) The foraging behaviors of ducks include grazing, dabbling, or upending, etc., which provides more opportunities to be in close contact with soil, water, and other surrounding environment. (ii) Compared with the birds at the top of the trophic level, ducks can trace more essential elements. For example, previous studies have shown that the eagle is not an appropriate bio-monitor to detect local elements (i.e., Cu and Zn) contamination due to the wide foraging areas (Badry et al. 2019) but ducks always keep site fidelity, with individuals routinely occurring within closely overlapping areas. (iii) Ducks are the most popular legally shot species around the world. In military and legal hunting sites, the natural degradation and immobilization of Pb and Cu from bullets will potentially threaten resident and migrant wildlife. Ducks are exposed to Pb through ingestion of grit, soil intake from preening, or ingestion of contaminated food near ranges. As a result, bullets will potentially elevate the Pb and Cu contents in duck feathers (Hampton et al. 2018). In addition, (iv) there are many species in the duck family with many gradients in life habits and trophic levels, which can provide the possibility for digging more bioaccumulation information.

In this study, approximately 200 feathers from seven duck species were sampled from 14 distinct catchments in the state of Victoria, Australia, and the concentrations of 16 toxic and essential elements were measured (Step 2 in Figure 6-1).

Table 6- 1. Summary of relevant studies conducted previously for investigating elemental bioaccumulation mechanism in bird feathers (The current paper is added for completeness). Three components in the 'Key results' column are identified by the code: (1) Whether to ignore the bioaccumulation ratio but emphasize only the threshold value of concentrations; (2) Whether to discuss toxic and essential separately; and (3) Whether the human footprint and environmental variables are both considered.

Study	Species	Organism	Analytical methods	Key results
(López-Perea et al. 2019)	moorhen	blood, feathers	GLM	(1) No, only concentrations (2) No, metals and metalloids (3) Only human footprint from treated wastewater
(Abbasi et al. 2015)	48 bird species	feathers	PCA, ANOVAs	(1) No, only concentrations (2) No, only heavy metals (3) No, influence by taxa, trophic level, geographical location
(Badry et al. 2019)	Bonelli's eagle	feathers	GLM, PCA, RDA	(1) No, only concentrations (2) No, metals and trace elements (3) Yes, agriculture, urban mines, power plant
(Liu et al. 2019)	3 duck species	feathers	correlation analysis	(1) No, only concentrations (2) No, only trace elements (3) No, only human footprints
(Picone et al. 2019)	Kentish plover	feathers	ANOVA	(1) No, only concentrations (2) No, only trace elements (3) No, only human footprints
(Frantz et al. 2012)	feral pigeon	feathers	PCA	(1) Yes, feather/atmosphere ratio (2) No, only heavy metals (3) No, only human footprints
(Abdullah et al. 2015)	cattle egret	feathers, eggs	ANOVA	(1) No, only concentrations (2) No, only trace elements (3) No, only human footprints, like the industrial areas
(Furtado et al. 2019)	penguin, petrel, cormorant	body feathers	ANOVA	(1) No, only concentrations (2) No, only heavy metals (3) No, only human footprints
(Innangi et al. 2019)	sparrow	flight feathers	LMM	(1) No, only concentrations (2) No, only trace elements (3) No, explained by the ages, like juveniles, moulting, and adult
This study	7 duck species	flight feathers	RDA, GLM, GAM	(1) Combing concentration and ratio;

(2) 6 toxic, 10 essential;

(3) Biotic and abiotic

Analysis of variances (ANOVA), principal component analysis (PCA), linear mixed models (LMM), generalized linear models (GLM), generalized additive models (GAM), redundancy analysis (RDA).

Then the background geochemical sediments, a set of environmental and human footprint factors based on the biological behaviors were collected as the explanatory variables (Step 3 in Figure 6-1) to explore and fill the current knowledge gaps. In particular, the objectives of this study are: (1) combining the elemental concentration and bioaccumulation ratio, both the essential and toxic elements, uncover bioaccumulation information, and (2) quantitatively interpret the element bioaccumulation mechanisms using biotic and abiotic variables. This will offer new perspectives and directions to elemental bioaccumulation mechanisms.

6.2 Materials and methods

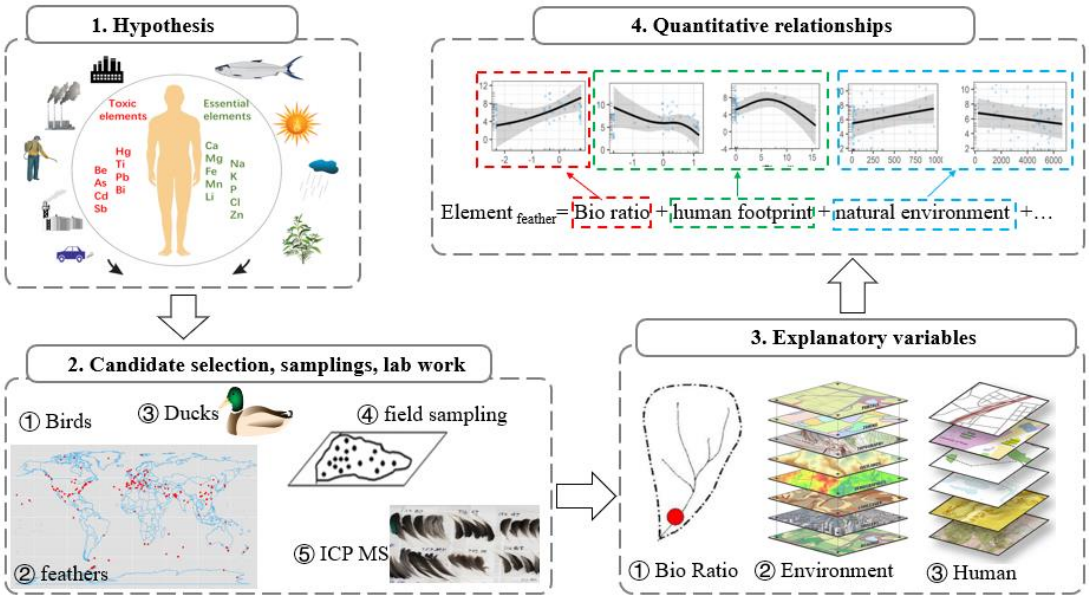


Figure 6- 1. Flowchart of this study with specific information for each step.

6.2.1 Study area

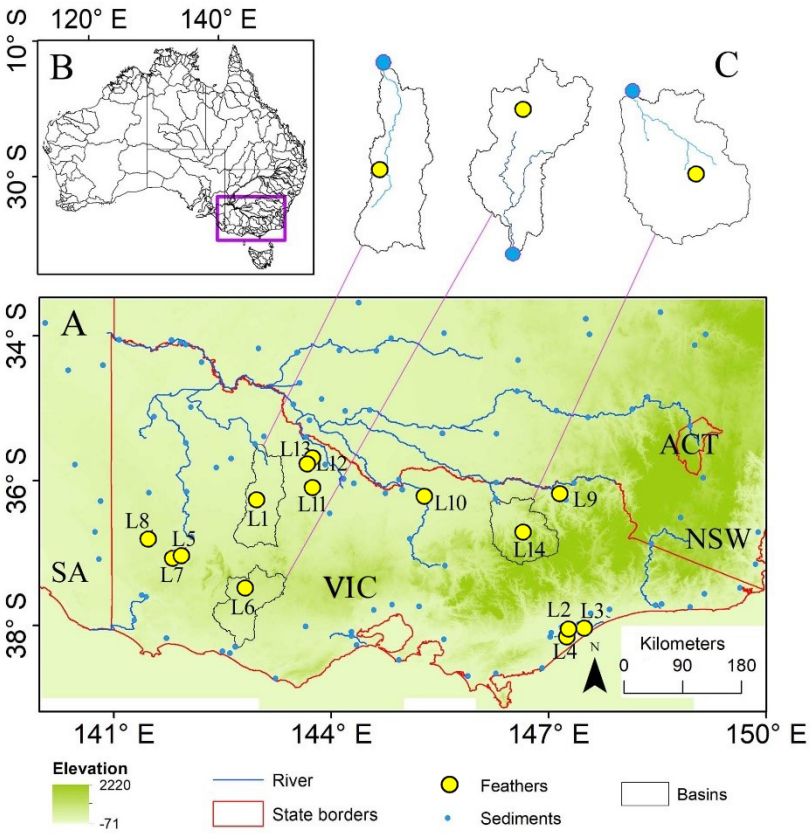


Figure 6- 2. (A) Study areas with feather and catchment outlet sediment sampling sites, and the site details are presented in Table 2. (B) The location of the study area (purple rectangle) within Australia with state borders and catchment boundaries. (C) Watershed delineation of three typical basins from all the 14 basins around the feather collection sites, to show spatial relationships between the feather collection sites and relative outlet sediment locations.

The study area is situated in the state of Victoria which lies in the southeastern part of the Australian continent, and is an ideal sampling point for the purposes of this study. The ocean lines here are assumed to be border limits for ducks as they live in areas of fresh water. The climate here is humid, and the catchment net is rich (Figure 6-2B), which means that many species of the ducks tend to stay in chosen areas and seldom venture outside them. The state of Victoria has a very active human population and rapidly increasing population density when compared to surrounding areas (Needham et al. 2020), and this provides a good platform for studying the impact of human footprints on ducks. In addition, shotgun hunting had not been restricted by the regulators during the sampling period.

The sites for feather and sediment collection were chosen to maximize the environmental and economic representation in southeastern Australia. For example, the population density in Lake 9 (L9) and L11 were above 800 people / km^2, while others were below 100 people /

km^2. The L2, L3, and L4 were in coastal areas, while the remaining samples were inland, with the distribution of rainfall amounts varying from 351 to 1110 mm.

Table 6- 2. Description of the sampling lakes.

Lable	Lake name	NDVI	rainfall	p_ frost	DEM	industry	mining	waste	p_density
L1	Lake Buloke	0.359	363.92	27	105	0	0	0	1
L2	Clyde Bank	0.550	577.12	23	0	0	0	0	45
L3	Blond Bay	0.608	610.95	18	0	0	0	0	87
L4	Dowd Morrass	0.500	560.64	22	-3	0	0	0	12
L5	Toolondo	0.488	477.49	33	158	0.028	0	0	58
L6	Lake Buninjon	0.513	526.95	24	232	0	0	0	39
L7	Kanagulk	0.508	505.06	33	173	0.018	0.868	0	14
L8	Wally Allens Swamp	0.497	466.66	35	158	0	0	0	12
L9	Lake Hume	0.566	870.30	65	182	0.005	0	0.183	1068
L10	Loch Garry	0.539	442.98	38	102	0.018	0	0	122
L11	Lake Lyndger	0.448	382.53	22	92	0.758	0.34	0.923	799
L12	Koorangie	0.410	351.67	19	71	0	2.602	0	9
L13	Bael Bael	0.411	351.67	19	71	0	2.605	0	8
L14	Lake Buffalo	0.705	1110.457	62	266	0	0	0	30

NDVI: normalized vegetable index; rainfall (mm/ year); P_frost: potential frost days annually; DEM: elevation; p_density: population density (people/ km^2); industry, mining, and waste are the areas of the related land use type in each buffer (km^2/ 78.5 km^2, the buffer refers to the method part).

6.2.2 Biotic dataset

The bioaccumulation could be driven by a set of physiochemical and biological variables, including the pH, temperature, habitat complexity, feeding behavior, as well as the exposure duration.

$$\mathrm{BAF}\% = \frac{\mathrm{C\ biota}}{\mathrm{C\ \ living\,environment}} \times 100\ \% \qquad (1)$$

The bioaccumulation ratio was calculated as a ratio of elemental concentration in tissues (mg/kg) relative to the living environment. Where the living environment is normally aquatic (mg/l) (Ahmed et al. 2019), soil (mg/kg), atmosphere (mg/kg) (Frantz et al. 2012), and sediment (mg/kg). The 'water at steady state' has always been used in the aquatic environment, and Ahmed et al. (2019) had given the threshold values of the bioaccumulation ratio when the tissue was fish muscles and living environment as the aquatic: BAF < 1000 indicating no probability of accumulation; 1000 < BAF < 5000 being the normal

bioaccumulation; while BAF > 5000 showing extremely bioaccumulation. But when the tissues are different, these ratio threshold values have no reference significance. In this study, to minimize the influence of duck movements, the sediments from each catchment outlet that is assumed to represent the chemical background value of the entire watershed, had been collected (Figure 6-2C).

6.**2.2.1 Feather collection and ICP laboratory analysis**

Feathers from seven duck species were collected in this study. We classified the duck trophic levels by the gizzard contents from literatures, where the food was divided into aquatic vegetation, terrestrial vegetation, micro- invertebrate, and macro-invertebrate (Table 3). The trophic levels of ducks are ranked as follows: PE (Pink eared, *Malacorhynchus membranaceus*) > MD (Mountain duck, *Tadorna tadornoides*) > BD (Black duck, *Anas superciliosa*) > GT (Grey teal, *Anas gibberifrons*) > CT (Chestnut, *Anas castanea*) > HH (Hardhead, *Aythya australis*) > WD (Wood duck, *Chenonetta jubata*). PE is an invertivorous species, and WD is a mostly herbivorous species, while other ducks are omnivorous species. Among these omnivores, the HH's food is basically from the aquatic environment.

Table 6- 3. Trophic levels of the duck based on gizzard contents from literature review, '0' means no related diet, while '1' stands for related contents in the gizzard.

Species	Aquatic vegetation	Terrestrial vegetation	Micro-invertebrates	Macro-invertebrates	trophic	Source
WD	0	1	0	0	Vegetarian	(Kingsford 1986)
BD	1	1	0	1	Omnivore	(Frith 1967)
CT	0	1	0	1	Omnivore	(Norman & Mumford 1982)
GT	0	1	0	1	Omnivore	(Norman & Mumford 1982)
HH	1	0	0	1	diving	(Frith 1967)
MD	0	1	0	1	Omnivore	(Davis & Wilcox 2013)
PED	0	0	1	1	Invertebrate feeder	(Kingsford & Norman 2002)

Duck feathers were collected in the gathering area of each species in March 2017. The duck feathers were then stored in the freezer laboratory of UNSW Sydney, and the temperature was kept below 0 ℃. Before the feathers were sent to the ICP MS laboratory, they were cleaned in water to remove the surface dirt by reverse water and dried in the fume hood. They were then analyzed using inductively coupled plasma mass spectrometry (ICP-MS, Model: ELAN9000, Perkin-Elmer, Germany). Through laboratory analysis, 6 toxic elements and 10 essential elements were obtained. The essential elements included Ca, Co, Cu, Fe, Mg, Mn, Mo, Ni, Se, and Zn; while the toxic elements included Al, As, Cd, Cr, Li, and Pb. Although Cr is

essential for humans, it had been classified as non-essential for animals(Malik & Zeb 2009), so we classified Cr as toxic.

6.2.2.2 Geochemical background samplings

The background geochemical data were extracted from the National Geochemical Survey of Australia (NGSA) project that had collected outlet overbank sediments from 1186 catchments across Australia (De Caritat & Cooper 2016). The overbank sediments were reasonably assumed to represent well-mixed, fine-grained composite samples of all major rock and soil types present in the catchment. At each outlet site, the 0-10 cm below the surface sediments were sampled. The sample collection process followed the protocols. Many criteria were taken to avoid any possible contamination or disturbance. In the laboratory, the samples were air-dried, homogenized, and the elemental compositions were analyzed in the Geoscience Australia's laboratories and other commercial laboratories. The detailed procedures for preparing the NGSA samples in the laboratory are described with an overview of Quality Assessment/Quality Control and Operational Health and Safety measures also provided.

6.2.3 Abiotic explanatory variables

To better understand the accumulation of elements in duck feathers, the sets of environmental and human footprint variables were selected based on these conditions:

(i) Variables are closely related to the biological behavior of ducks and other similar species. For example, the Green-wave hypothesis proves that some herbivorous birds will move with seasonal foliage dynamics that can be represented by the normalized difference vegetation index (NDVI). Heavy rainfall is an important resource for habitation and breeding. Except for the rainfall amount, moisture transport is another important trigger for water bird movement (McEvoy et al. 2017). The number and size of available farm dams as well as their proximity to agricultural land are important drivers of Australian Wood Duck abundance. Dams are sinks for many pesticides, fertilizers, etc., which can gather information on nearby agricultural activities.

(ii) Variables, such as the elevation and distance to mines have been proven by previous studies to have a profound impact on the element bioaccumulation.

(iii) The problem of collinearity between variables is temporarily ignored in the selection process of variables. If it is necessary to eliminate collinearity in further analysis, it should be screened according to the variable’s significance.

Finally, eleven environment variables and fifteen human footprint variables were extracted as showed in Table 4. A 5km radius around the sampling sites was then applied to extract the information according to the duck foraging fidelity ranges (McEvoy et al. 2019).

Table 6- 4. Environment and human footprint variables in this study.

Environment variables	Human footprints
NDVI	land use type1_conservation and natural areas
Rain of annual variations	Land use type2_agriculture and plantations
Rain of annual average	Land use type3_water bodies
rain of coldest season	Land use type4_human intensive uses
rain of warmest season	Land use type4.1_industry and manufacturing
temperature_annual average	Land use type4.2_farm infrastructure and residential
temp _ max warmest month	Land use type4.3_transport and communication
temp_ min coldest month	Land use type4.4_mining
potential frost days	Land use type4.5_waste treatment and disposal
Deuterium-excess	Land use type4.6_others
DEM	distance to mining
	distance to industry
	distance to dams
	distance to waste
	population density

6.**2.3.1 Environmental variables dataset**

(1) Normalized difference vegetation index

The normalized difference vegetation index (NDVI) has been widely considered to be an important indicator for evaluating the quality of "greenness" of vegetation coverage. The formulation of NDVI is:

$$\mathrm{NDVI} = \frac{\mathrm{NIR}-\mathrm{Red}}{\mathrm{NIR}+\mathrm{Red}} \qquad (2)$$

Where NIR is the near- infrared surface reflectance and Red is the red surface reflectance. The NDVI ranges from -1 to 1. Negative numbers are water bodies, 0 is bare rock or no vegetation coverage, and the vegetation coverage is improving when NDVI approaches to 1.

Moderate Resolution Imaging Spectroradiometer (MODIS) NDVI data (MOD13Q1, collection 6) in 2017 at 16-day, 250m resolution were downloaded from NASA online dataset (https://modis.ornl.gov/globalsubset). Quality assessment (QA) is further conducted for the NDVI data to filter the noise (https://lpdaac.usgs.gov/products/mod13q1v006/). We excluded the pixel reliability ranking of 3 which means the target is covered with cloud and not visible

(detailed information can reference “MODIS Vegetation Index User’s Guide (MOD13 Series)”).

(2) Rainfall and moisture

The rainfall information was extracted from the gridded annually rainfall fields at a resolution of 0.05° and ranging 1988 to 2017 from the Australian Government Bureau of Meteorology (BoM). The dataset was derived from several thousand ground meteorological station observations across Australia. And four indicators were used in this study:

(a) Annual rainfall,

(b) Rainfall annual variation, which shows how rainfall varies from year to year,

(c) Rainfall of the coldest quarter,

(d) Rainfall of the warmest quarter.

To further research the influence of moisture transportation, the stable isotope data set was introduced. Stable hydrogen (^{2}H and ^{1}H) and oxygen (^{18}O and ^{16}O) isotope signatures are powerful tools in the study of the hydrological cycle, as these two conservative tracers are essentially the water molecule itself. In this study, we used the Deuterium excess to show the moisture movements in the study area. Deuterium excess (d-excess) is defined by the relation:

$$\text{d-excess} = \delta^{2}H - 8\,\delta^{18}O \qquad (3)$$

The d- excess values of precipitation can provide insight to sources of moisture, atmospheric processes, and relative humidity. The low d-excess value (< 10) indicates evaporated rainfall whereas high d-excess value points towards recycled moisture. A decreasing trend was observed when moving from the seaside to the inland. The d-excess values were extracted from the gridded isoscape (‰) from the isoscape map of Australia, in which the precipitation isotopic composition data (1962-2014) was derived from 7 original rainfall collection sites of the Global Network of Isotopes in Precipitation (GNIP) data (accessible at http://www.iaea.org/water) and additional 8 sites of the Australian Nuclear Science and Technology Organization across the Australian continent (Hollins et al. 2018).

(3) Temperature

The temperature information was also extracted from the gridded database from the BoM, and the four indicators include:

(a) Annual mean temperature,

(b) Average daily maximum temperature of warmest month,

(c) Average daily minimum temperature of coldest month,

(d) Potential frost days annually that is calculated by adding the number of days with minimum temperature below a given threshold (potential frost days) from 1976 to 2005, then dividing this by the number of years. Potential frost days are mainly confined to a period from late autumn to early spring. Northern Australia is generally frost free, but the potential days increase sharply in the southeastern region.

6.2.3.2 Human footprint dataset

The **land use classification** was sourced from a product of the Australian Collaborative Land Use and Management Program (December 2018), with 50 m resolution (https://www.agriculture.gov.au/abares/aclump/land-use/catchment-scale-land-use-of-australia-update-december-2018). Four major classification was aggregated:

(a) conservation and natural environments,

(b) agriculture and plantations,

(c) water bodies with some of human projects, like dams and channels being included,

(d) human intensive use. As this study traced the human activity influences, the intensive uses were further categorized into 4.1 Manufacturing and industrial; 4.2 Residential and farm infrastructure; 4.3 Transport and communication; 4.4 Mining; 4.5 Waste treatment and disposal; 4.6 Other intensive uses, such as intensive horticulture, animal production and services.

Population density information was extracted from the Australian Bureau of Statistics,2017 data pool at 1 km^2 grid resolution. This dataset offers the distribution of population across the Australian landmass.

6.2.4 Methodology

6.2.4.1 Redundancy analysis

Redundancy analysis (RDA) is a powerful methodology to produce an ordination that regresses the impact of a matrix of multiple explanatory variables. The multiple linear regression is performed for the response variables in turn. The objects with similar variable values are ordinated closer together, while different values are projected further apart.

Before analysis, the datasets needed to be normalized to reduce the influence of different units. The qualitative variables are recorded as dummy variables (Legendre and Legendre 2012), and in this study they referred to different duck species. The explanatory variables

with linear correlation were removed and then a permutation test with 1000 shuffles was conducted to examine the null hypothesis that no linear relationship existed between the response and explanatory variables (Block et al. 2017). The constraining variables significant was examined by permutation test with 100000 shuffles. The proportion of variance explained by each RDA axis is presented by an eigenvalue. Considering the extreme complexity of the natural environment, the sum of the first two axes eigenvalues above 10% is acceptable in this study (Badry et al. 2019).

6.2.4.2 Generalized linear models and generalized additive models

Generalized linear models (glm) allowing the response variables to have error distribution models, have been widely used as an appropriate theoretic approach for biological research. Glm were therefore applied in this study to examine the ability of a set of explanatory variables to explain variation of the elemental concentrations in duck feathers. The explanatory variables included the biotic bioaccumulation ratio, environmental and human footprint variables.

Generalized additive models (gam), where the response variable is not restricted to be linear in the explanatory variables, were then introduced to further estimate the smooth components of the glm models using smooth functions. Then using the model averaging approach, the explanatory variables with the sum of weights≥ 0.5 were retained in the model set to predict the relationships for each element in duck feathers.

6.2.4.3 Model selection

Model selection was conducted to select the top model from the best model set. The top model is the most parsimonious combination of explanatory variables using the cross-validation process. The best models were then ordered by the criteria of the Akiake Information Criterion (AIC). The AIC is a measure of fit that penalizes for the number of variables. And AICc was introduced when the sample sizes are small:

$$\mathrm{AICc} = \mathrm{AIC} + \frac{2\mathrm{K}(\mathrm{K}+1)}{\mathrm{n}-\mathrm{K}-1} \qquad (4)$$

Smaller values indicate better fit and thus the AICc can be used to compare models. Thus, models with the lowest AICc are considered to be the top models, highlighting the included variables and model support relative to other models within 2 AICc.

Prior to statistical analysis, environmental and anthropogenic variables were scaled by subtracting the mean and dividing by the standard deviation to enhance comparability of effect sizes across variables measured in different scales. The collinearity between the explanatory variables was considered in this study and only the most significant variables

were retained for further analysis. The data extraction and analysis in this study were conducted in ArcGIS 10.5 (Environmental Systems Research Institute Inc.) and R v3.5.1 (R Core Team, 2018).

6.2.5 Research presumption of movement issue

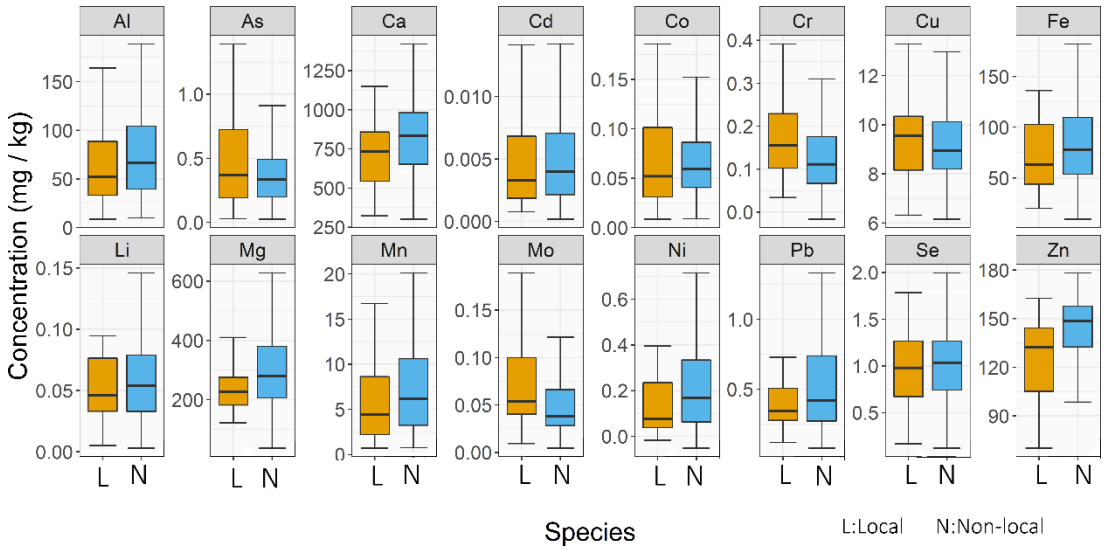


Figure 6- 3. Boxplot comparing the local and nonlocal species concentration in individual elements.

In terms of the movement, the WD and BD tend to live a more sedentary life comparing with other species (McEvoy et al. 2019). Data supports that the WD and BD, being more sedentary, could be regarded as local species, while all other species could be regarded as non-local.

Figure 6-3 shows that the concentration distributions of As, Cr, Cu and Mo in the local species were slightly higher than in the non- local ones, while other elements tended to show an opposite trend. Generally, there were no significant differences between the local and non-local species, indicating that movement may be not a significant factor influencing the element bioaccumulation mechanism. We focused on providing a methodology to test the hypothesis, and do not consider the movement issues. In addition, the sediments at the outlet were selected to represent the geochemical background of the entire watershed to limit the influence of movements.

6.3 Results

6.3.1 Biotic variables

6.3.1.1 Elemental concentrations

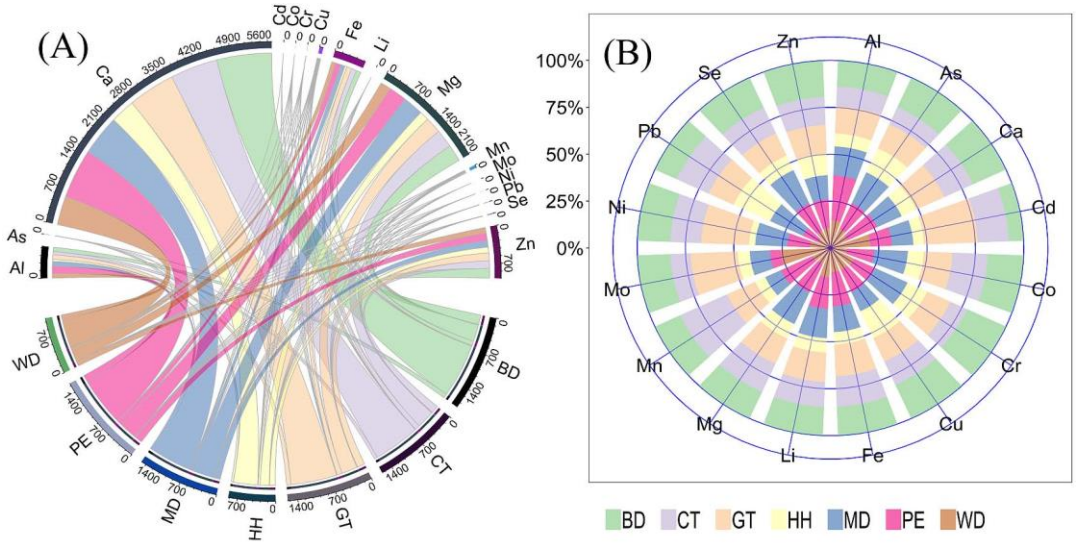


Figure 6- 4. Basic geochemical information for different duck species. (A) Chord diagram showing the element concentrations of the seven duck species, the width of the chord is the absolute value of element concentration; (B) polar coordination rose plot presenting the relative percent of the various elements in different species.

Table 6- 5. Element concentrations of different duck species, unit is mg/kg, and for Cd, Co, Li, and Mo is 10^{-2} mg/kg.

S	Al	As	Ca	Cd	Co	Cr	Cu	Fe	Li	Mg	Mn	Mo	Ni	Pb	Se	Zn
BD	82.6	0.57	1063.19	0.49	9.64	0.26	13.22	91.36	6.58	329	8.18	8.33	0.36	0.63	1.43	189.52
CT	61.53	0.33	920.49	1.03	9.45	0.11	9.14	95.56	5.69	352.31	15	5.07	0.34	0.93	1.23	145.56
GT	86.38	0.46	913.24	1.78	7.67	0.15	9.25	100.44	6.71	313.64	9.09	11.84	0.52	0.72	1.06	151.1
HH	41.11	0.46	477.66	0.35	4.35	0.13	10.61	39.54	3.31	181.88	3.26	3.04	0.07	1.06	0.7	110.97
MD	91.54	0.59	785.44	0.71	9.02	0.21	7.49	81.42	6.73	395.76	4.98	5.08	0.35	0.94	1.59	125.61
PE	142.43	0.49	957.51	0.70	7.32	0.14	8.77	107.44	7.19	349.55	10.19	2.68	0.24	0.44	1.31	148.23
WD	87.68	0.37	561.29	1.34	4.47	0.16	8.65	76.69	6.71	202.72	5.83	12.58	0.23	0.36	0.83	105.97

S, Species. BD, Black duck, *Anas superciliosa*; CT, Chestnut, *Anas castanea*; GT, Grey teal, *Anas gibberifrons*; HH, Hardhead, *Aythya australis*; MD, Mountain duck, *Tadorna tadornoides*; PE, Pink eared, *Malacorhynchus membranaceus*; WD, Wood duck, *Chenonetta jubata*.

The order of the total element concentrations in duck feathers did not follow the trophic level. It was sorted as BD> PE> CT> GT> MD> WD> HH, and the total concentrations followed by 1780.55> 1726.91> 1602.75> 1586.33> 1496.14> 1051.02> 867.57 mg/kg (Figure 6-4A).

For each single element, the hierarchy of concentrations was Ca> Mg> Zn> Al> Fe> Cu > Mn> Se> Pb> As> Ni> Cr> Co> Mo> Li> Cd, and the average level was 5678.81> 2124.85> 976.96> 593.26> 592.44> 67.14> 56.53> 8.15> 5.08> 3.27> 2.10> 1.16> 0.52> 0.49> 0.43> 0.06 mg/kg (Figure 6-4A, Table 6-5).

The invertebrate feeder PE, was the highest in toxic Al, Li, and essential element Fe, but lowest in Mo (Figure 6-4B), the concentrations being 142.43, 0.07, 107.44, and 0.03 mg/kg (Table 6-5). Of the four omnivorous species, MD had the highest values in toxic As, and essential Mg, Se, being 0.59, 395.76 and 1.59 mg/kg (Figure 6-4B, Table 6-5). GT achieved the peak values in toxic Cd and essential Ni, being 0.018 and 0.52 mg/kg (Figure 6- 4B, Table 5). However, the BD and CT only showed the maximum values in essential elements. In the total amount of elements, BD ranked first, which is related to its highest concentration in many essential elements, including Ca, Co, Cu and Zn, with the related concentrations being 1063.19, 0.10, 13.22 and 189.52 mg/kg (Figure 6-4B, Table 6-5). CT had the highest concentration in Mn.

The total element concentration of HH was the smallest due to its minimum values in most elements (Figure 6-4A), however, its Pb content was the highest and the Cu content was the second highest (Figure 6-4B), being 1.06 and 10.61 mg/kg (Table 6-5). HH mainly forages the aquatic plants and invertebrates but seldom eats terrestrial ones (Table 6-3), and the predation method is diving. The herbivorous WD only achieved the highest concentration in Mo (Figure 6-4B).

6.**3.1.2 Element bioaccumulation ratios**

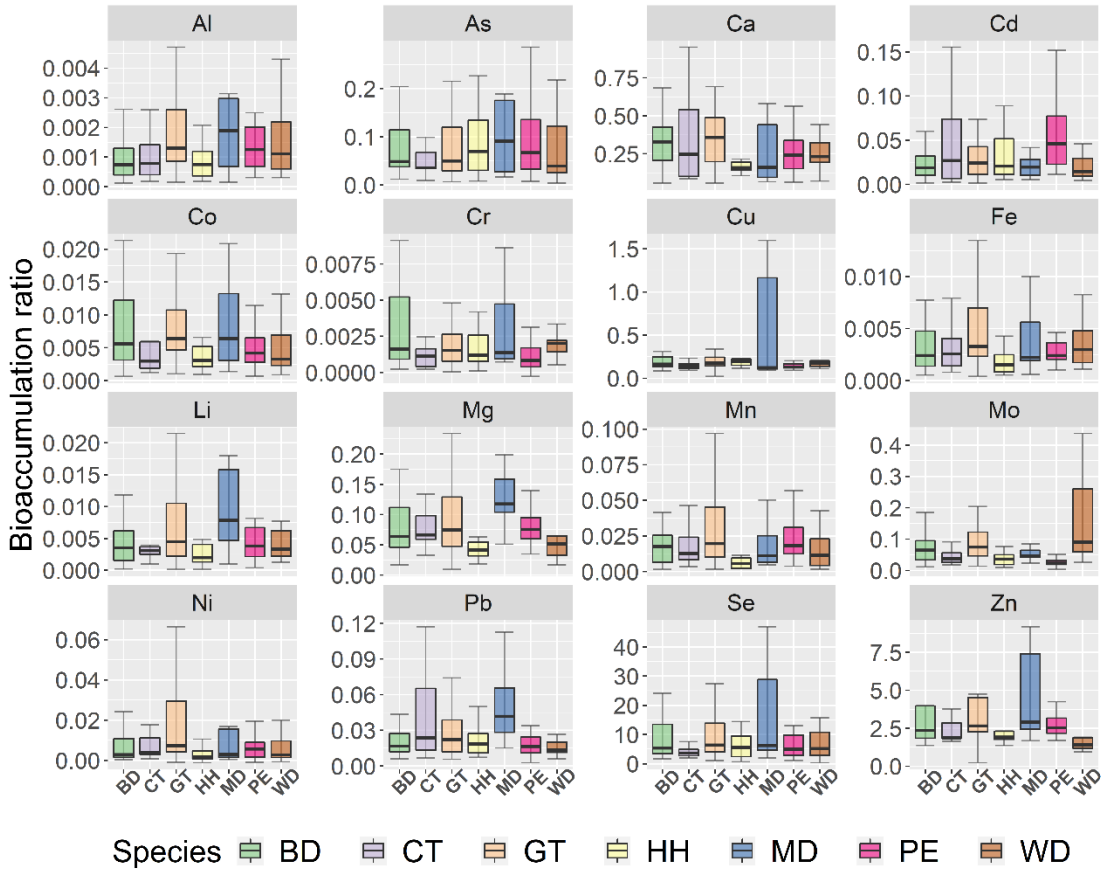


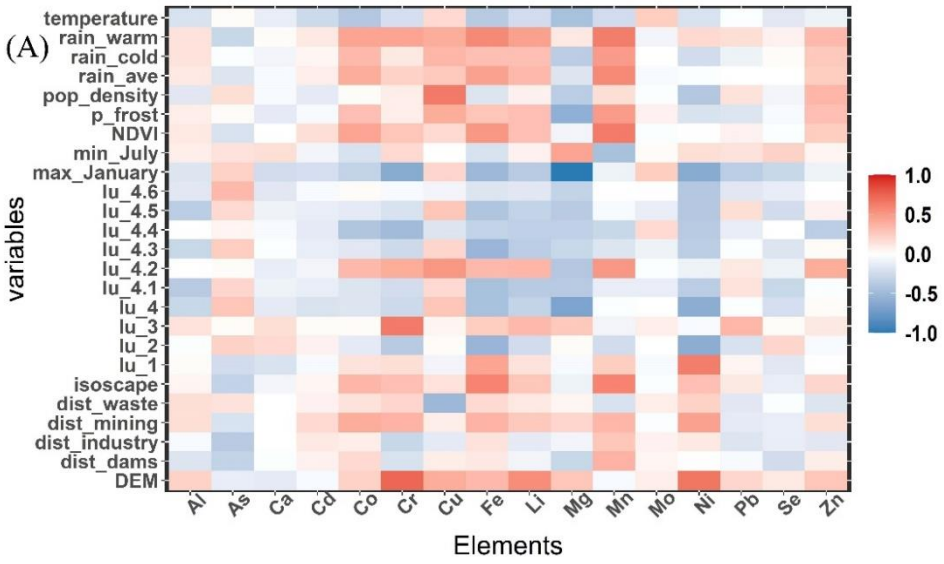
Figure 6- 5. Boxplots showing the bioaccumulation ratios of the seven duck species of different elements, and Cu was $\log_{10}$ (x + 1) transformed due to the over- squeeze.

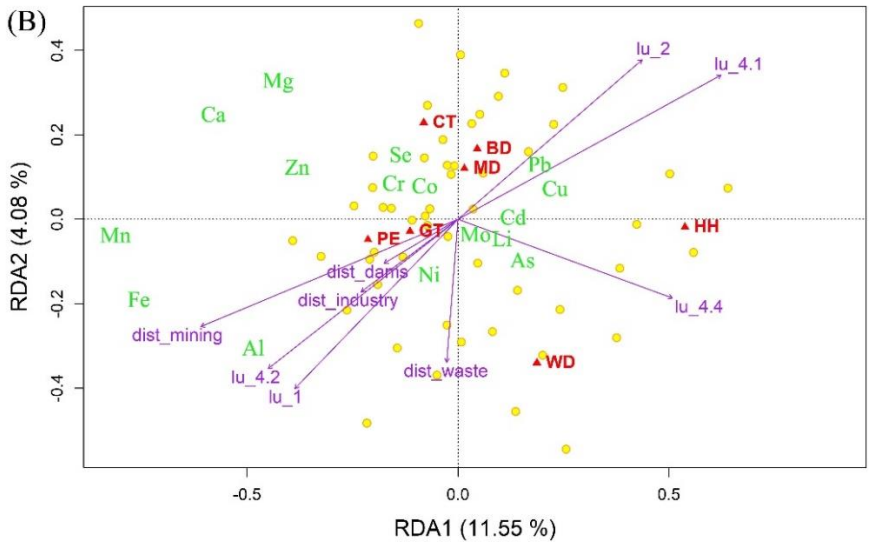
Although the concentrations of As and Cd in PE were not high (Figure 6- 4B), the ratios of these two elements were obviously more prominent than other species (Figure 6-5). On the contrary, the bioaccumulation ratios of Al, Li, Fe were not as obvious as their concentration values (Figure 6-5).

The bioaccumulation ratio distribution in MD showed a high trend in many elements, such as the toxic Al, As, Cr, Li and Pb, and essential Co, Cu, Mg, Se and Zn. With the exception of MD, the distribution of bioaccumulation ratio in other omnivorous species was mostly higher in essential elements. BD was higher in Co and Cr. The distribution of Fe, Ni and Mn in GT was generally higher (Figure 6-5).

WD is almost vegetarian and show relatively higher ratios in Fe and Mo, but having low ratios for toxic As, Cd, and Pb. Although the Pb and Cu concentrations of HH were higher, the related distributions of their bioaccumulation ratio were not so obvious (Figure 6-5).

6.3.2 Abiotic variables





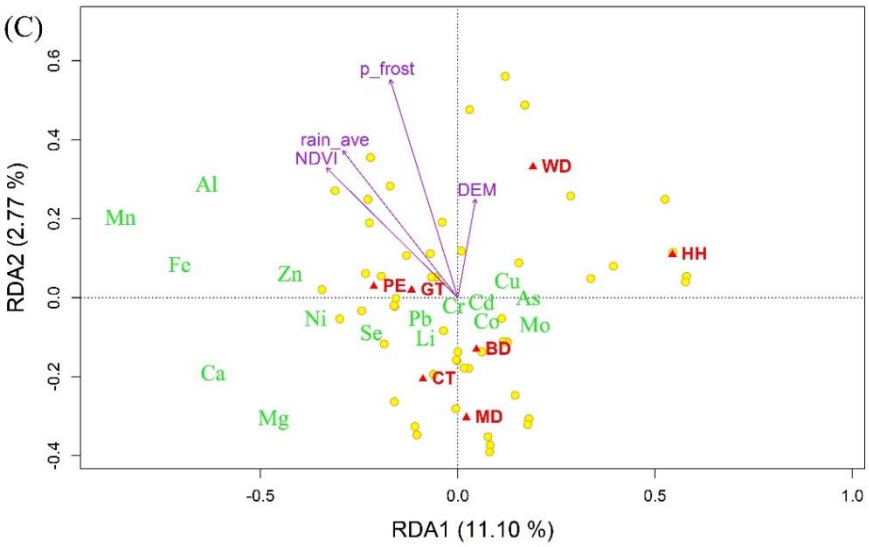


Figure 6- 6. (A) Correlative heat map of traits with their respective coefficients. (B) Triplots of Redundancy analysis (RDA) for human footprint and (C) environment variables.

The heatmap (Fig. 6-6A) showed the Pearson coefficients between the element concentrations and abiotic factors. The Land use types were positively correlated with As, Cu, and Pb, but roughly negatively correlated with other elements. Among them, 'land use type 4.3 farm' was obviously positively correlated with many elements, and 'land use type 3 water' showed a positive correlation with Pb. The concentration of Cu was positively related with 'population density', but negatively with 'distance to waste'. The correlation between environmental variables and essential elements was stronger than toxic ones. For example, the 'maximum temperature of January' was extremely negatively correlated with Mg, but "minimum temperature of coldest month July" showed a positive relationship with Mg. DEM was extremely positively correlated with Cr and Ni. Similarly, NDVI with Co, Fe and Mn was positively correlated.

According to the results of Redundancy analysis (RDA), the distribution of species was basically consistent with their biotic characteristics. The herbivorous WD and diving HH were dispersed separately by themselves, while other omnivore and invertebrate feeders were clustered together. (Figure. 6-6B & 6C).

In addition to "distance to waste" close to the second axis, other human footprint variables basically acted on the first axis. The "land use type 4.4 mining" was closely related to As, Cd, and Li. The "land use type 4.1 industry" and 'land use type 2agriculture' were closely related to Pb and Cu. The other variables acted in the same direction and were closely related to Al and Fe (Fig. 6-6B). The species were generally closer to the elements that they reached the highest concentration of, such as PE with Al and Fe, GT with Cd and Ni. HH was closer to the Pb and Cu, while WD, PE, GT with a lower Pb concentration were all far from Pb.

As for the environment variables, "DEM" and "potential frost day" were close to the second axis, but "NDVI" and "rain annual average" weakly acted on the first axis. The force direction of environment variables all deviated from the heavy metal area (Fig. 6-6C).

6.3.3 Bioaccumulation mechanism with both biotic and abiotic variables

6.3.3.1 Top models for elements

The top models were identified by model selection approach from the best explanatory variables for each element. Models with lowest AICc are listed in table 6, highlighting the included variables and model support relative to other models within 2 AICc. Essential element Cu was only affected by the environmental variables of 'potential frost days' and 'rainfall'. The toxic elements Al, Cr, Li, and Pb, as well as the essential Mn, were only

affected by the human footprint variables. Most other elements were basically influenced by the interaction of the two types of abiotic environmental and human footprint variables (Table 6).

Table 6- 6. Top models identified through model selection based on AICc.

Elements	Smoothed variables in best model	df	loglik	AICc	Aw
Al	Al_bio+dist_dams+dist_mining+dist_waste	16	-967.5	1970.65	0.08
As	As_bio+dist_dams+dist_waste+p_frost	19	158.49	-273.93	0.2
Ca	Ca_bio+dist_waste+lu_1+lu_2+lu_4+NDVI+p_frost	15	-1386.89	2808.77	0.2
Cd	Cd_bio+DEM+lu_2	12	655.32	-1282.66	0.05
Co	Co_bio+DEM+dist_dams+dist_industry+dist_waste	18	610.33	-1179.91	0.19
Cr	Cr_bio+dist_waste+lu_1+lu_2+lu_4	17	564.85	-1090.03	0.38
Cu	Cu_bio+p_frost+rain_ave	22	-278.36	609.33	0.01
Fe	dist_dams+Fe_bio+p_frost	16	-832.14	1701.43	0.4
Li	dist_dams+dist_waste+Li_bio	20	542.96	-1040.47	0.1
Mg	dist_dams+Mg_bio+NDVI+p_frost	16	-1111.8	2260.28	0.2
Mn	dist_waste+lu_2+Mn_bio	12	-595.54	1217.15	0.43
Mo	dist_dams+dist_waste+Mo_bio+NDVI+p_frost	19	536.05	-1027.94	0.13
Ni	dist_dams+dist_mining+dist_waste+lu_4+Ni_bio+rain_ave	19	234.26	-424.55	0.05
Pb	dist_dams+dist_waste+lu_4+Pb_bio	13	102.42	-175.41	0.13
Se	DEM+dist_dams+dist_mining+dist_waste+lu_1+Se_bio	18	-127.12	295.89	0.14
Zn	dist_dams+dist_industry+dist_waste+lu_1+NDVI+rain_ave+Zn_bio	18	-948.44	1939.22	0.16

df: Degrees of freedom, "Aw": "Akaike weight"

6.3.3.2 Interpreting element bioaccumulation by both biotic and abiotic variables

The gam was then introduced to further estimate the smooth components of the glm models using smooth functions. Then using model averaging approach, the explanatory variables with the sum of weights≥ 0.5 were retained in the model set to predict the relationships for each element in duck feathers.

The variables that entered the final best model were quantitatively evaluated for their impact on each element based on their "sum of weights" values (Fig 6-7). We then proceeded to simulate the final best model for each element through the gam model according to the

optimal options selected by AICc, and the results were presented in Figure 6-7 and supplementary Figure S1.

Biotically, the bioaccumulation ratio showed a positive association with most elements (Figure.S1 a-l), but with some inflection points for Li, Pb, Se and Zn (Figure. 6-7 a-d). There was a peak in some elements, such as with Li, Zn and Mo. Li reached a peak value of 0.5 mg/kg near ratio=0.13, and then decreased (Figure. 6-7a). Zn started to decrease after ratio=1.6 (Figure. 6-7d). Mo reached the first peak 0.7 mg/kg when ratio=0.8, then decreased until it reached the trough when ratio=1.2, and then it rose again (Figure.S1 k). There was a stable state for some elements, such as Se and Cu. Although Se also showed a constant rising state, it was stable at about 2.5 mg/kg when the ratio was between 10 and 30 (Figure. 6- 7c). After Cu reached its maximum value at ratio=7.5, a steady state began to appear (Figure.S1 g).

The environmental variables were mostly related to essential elements, except in the case of "potential frost day" with As and "DEM" with Cd (Figure. 6-7). The 'potential frost day' was positively related to Ca and Mo (Figure.S1 n, p), but negatively with As and Cu (Figure.S1 m, o). Its relationship with Mg fluctuated. When there were 15 potential frost days, Mg reached 600 mg/kg, then dropped, reached the lowest value in 25 days, and then rose again (Figure. 6-6i). The 'elevation' was basically positively related with Cd, Co and Mo (Figure.S1 q-s). The predicted concentration of Zn kept decreasing with the increase of rain (Figure.7j). The 'NDVI' had positive relationships with Cu and Zn (Figure. 6-7e, 7h), but negative with Mg (Figure. 6-7f). Mo first increased with NDVI, reaching the highest value when NDVI=-0.05, and then decreased (Figure. 6-7g).

The human footprint pressure affected all these elements after the weight was screened of the explanatory variables. The toxic Al, Li, and Pb, as well as the essential Cr, Fe, Mn, Ni, and Se only related to the human variables, with no significant environmental impact (Figure. 6-7). "Land use type 2agriculture" was generally positively correlated with all the correlated elements (Figure. 6- 7k, Figure.S1 z-ab). "Land use type 4 intensive use" was negatively correlated with Ca (Figure. 6-7 l), but positive with Cr, Cu, and the toxic Pb (Figure.S1 ac, ad, Figure. 6-7m). The "distance to mining" significantly influenced Se (Figure. 6-7n). The closer the distance to mining, the higher the content of Se, reaching the maximum value of 1.2 mg/kg at the central area of the mining site.

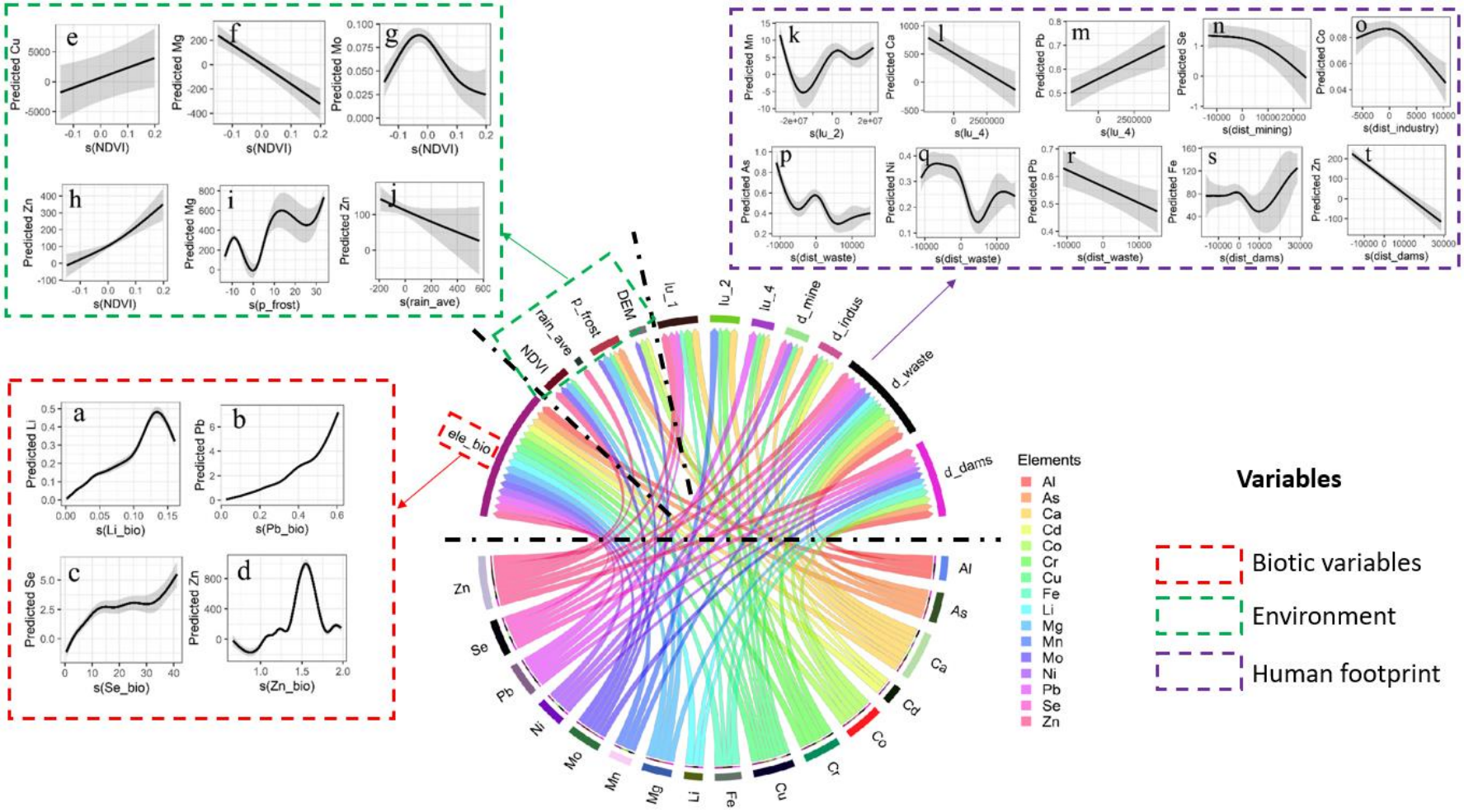

e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
a
b
c
d
s(NDVI)
s(p_frost)
s(rain_ave)
s(dist_waste)
s(dist_mining)
s(dist_industry)
s(dist_dams)
s(lu_2)
s(lu_4)
s(Li_bio)
s(Pb_bio)
s(Se_bio)
s(Zn_bio)
Predicted Cu
Predicted Mg
Predicted Mo
Predicted Zn
Predicted Mn
Predicted Ca
Predicted Pb
Predicted Se
Predicted Co
Predicted As
Predicted Ni
Predicted Fe
Predicted Li
rain_ave
p_frost
DEM
NDVI
ele_bio
lu_1
lu_2
lu_4
d_mine
d_indus
d_waste
d_dams
Zn
Se
Pb
Ni
Mo
Mn
Mg
Li
Fe
Cu
Cr
Co
Cd
Ca
As
Al
Elements
Al
As
Ca
Cd
Co
Cr
Cu
Fe
Li
Mg
Mn
Mo
Ni
Pb
Se
Zn
Variables
Biotic variables
Environment
Human footprint

Figure 6- 7. Chord diagram showing the quantitative relationships between the elements and explanatory variables according to the sum of weights from AICc, and the arrows showing the link between elemental concentration and relative attribution from the three different types of variables. The sum of each element is the sum of the weights of all explanatory variables ≥ 0.5. Same elements use the same color, the secondary bar near element shows the proportion effects from the explanatory variables. Smoothed fits of relationships between elements and explanatory variables. (Li_ bio stands for the bioaccumulation ratio of Li, same for Pb, Se and Zn_ bio).

The 'distance to industry' had a significant influence on Cd and Co, and the maximum concentration of Co was predicted around the industry sites (Figure. 6-7o). At a distance of 2000 m from the industry, the Cd was predicted to reach its lowest value, and then whether it was close to or far away from the industry, the value of Cd would rise (Figure.S1ah). When near the waste points, both the toxic and essential elements reached a peak (Figure. 6-7 p-r, Figure.S1 ak-at). The "distance to waste" affected most of the elements except Cd, Mg and Mo. On the contrary, when near the dams, most elements reached extremely low values, except for the predicted high values of Fe and Zn (Figure. 6-7s, 7t).

6.4 Discussion

6.4.1 Combing concentration and bioaccumulation ratio

Simply comparing the concentration values will ignore the influence of background geochemical information, which is not recommended. For example, if we analyze two types of bird feathers originating in the same geologic unit, like within the same hydrological basin, the elemental concentration can be directly compared but if the feathers come from different geologic units, the geochemical backgrounds or the bioaccumulation ratios should be considered. Thus, it is necessary to combine the elemental concentration and bioaccumulation ratio when studying the bioaccumulation processes, which can provide more detailed information, otherwise the spatial attributes of the geochemical background information will be ignored.

In PE, with a higher trophic level than other species in this study, the concentration of Cd was only ranked fifth among the seven species. The correlation between Cd and trophic level is easily to be ignored if we only consider the element concentration, however, if the bioaccumulation ratio of Cd is also considered, we will find that the distribution of Cd bioaccumulation ratio in PE feathers was the highest and significantly higher than other species (Figure 6-5). Similarly, the bioaccumulation ratios of Pb and Cu in HH were very low, but their element concentration were very high (Figure 6-4), which may indicate the influence of Pb - Cu bullets from hunting activity, or of the diving foraging manner.

GT reached peak values in toxic Cd and essential Ni concentrations, and the bioaccumulation ratios of Cd and Ni ratios were also significant. Cd and Ni showed a good correlation at p <0.001 level. In addition, the distribution of Li bioaccumulation ratio in GT also showed a high trend. The correlation between Cd and Ni may be associated with Ni-Cd batteries, and the important resource of Li is Li batteries. However, there is currently few literatures showing that GT tends to forage in landfills. Whether GT tends to approach the batteries needs more attention in the future research.

Some current research had proposed the bioaccumulation ratio threshold for organs other than feathers (Ahmed et al. 2019). Others had explored to provide unique thresholds for single element concentrations (Picone et al. 2019) but more accurate information can be expected only by considering both the element concentration and bioaccumulation ratios together. Therefore, we strongly suggest gradually establishing the ratio thresholds for feather in future research.

6.4.2 Influence of abiotic variables

Combining the environmental and human footprint variables can help to predict the elements more accurately. NDVI was one of the important environmental variables to predict Mo. The Mo content in duck feather reached highest when NDVI is around zero. Meanwhile Mo was also affected by "distance to dams", so we can infer that NDVI around zero is the bare land around dams.

Temperature variable, or specifically the potential frost days, had a significant influence on As. The positive correlation between temperature and As, may reflect the increase in natural sources of As, such as weathering, or in the increase in human activities that are more affected by temperature (Smith et al. 2003). However, the concentration of As was only related to distance to waste and dams, so we can confidently remove the influence of natural impact when predicting As.

Land use type 4, intensive use, was negatively correlated with Ca (Figure. 6-7l), but positive with Cr, Cu, and the toxic Pb. The intensity of human activity was negatively correlated with the essential elements, but positively correlated with toxic. It reflects those human activities are important sources of toxic elements.

The toxic elements Al, As, Li and Pb, were all affected by the distance of waste and dams. Pb was also affected by intensive use, and Cd was affected by distance industry. To strengthen the conservation of these waterfowls, it is necessary to control the waste discharge from anthropogenic activity.

6.5 Conclusions

This research explored the relationships between element signatures in feathers and biotic and abiotic variables. Results found that combining both the elemental concentration and bioaccumulation ratio can help improve the accuracy of analysis. Through the top models, most of the elements were basically affected by both environmental and human footprint variables, but the essential element Cu was only affected by environmental factors, and the toxic elements Al, Cr, Li, and Pb, as well as essential Mn were only affected by human activities. The environmental variables had a greater effect on essential elements, while human footprint variables had a great impact on both the toxic and essential ones. But there are also many other important factors in biotic variables, so we strongly recommend that future research considers the trophic level, gender, age and other biotic characteristics.

6.6 Supplementary materials

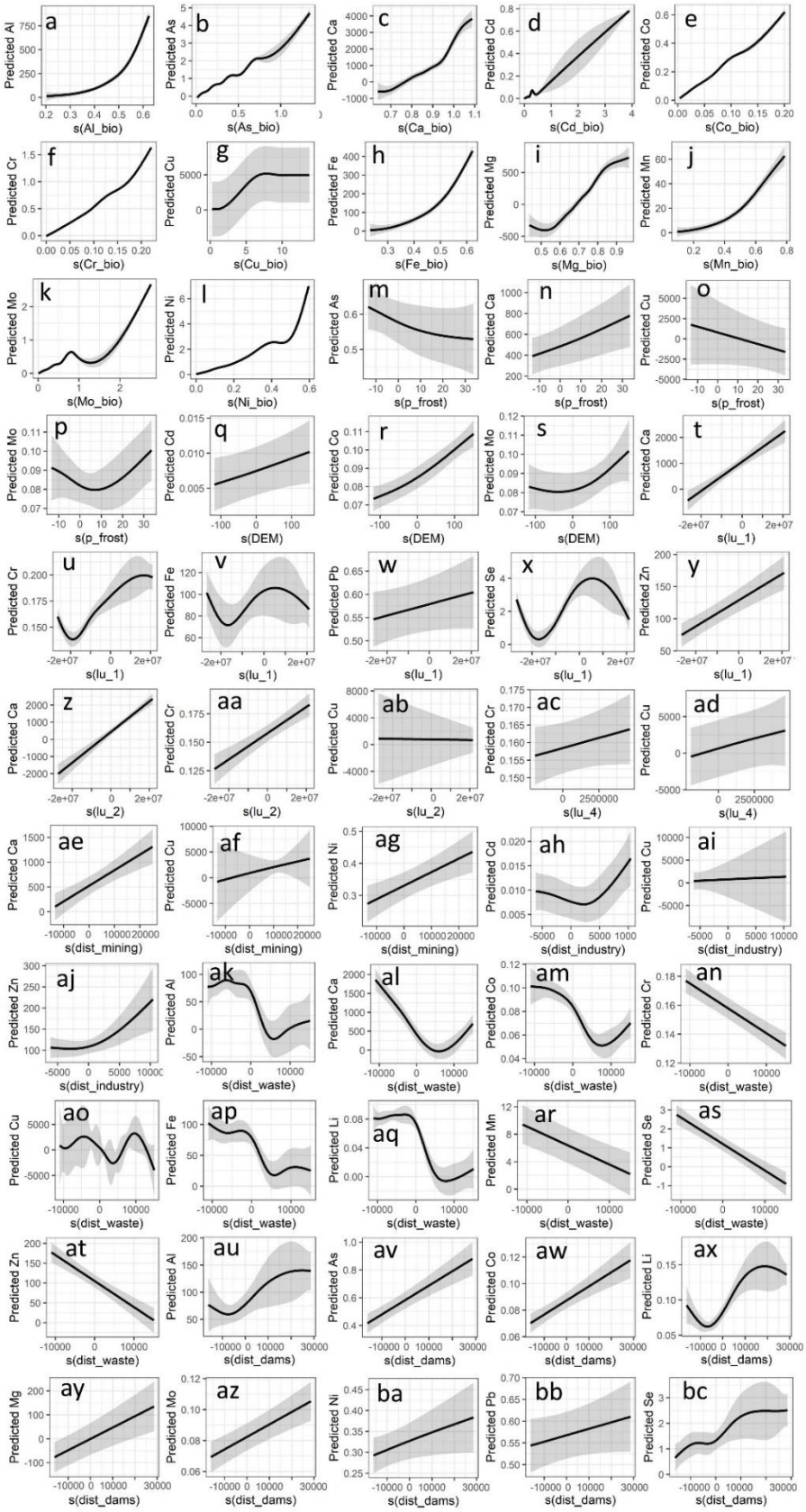


Figure. S1 Smoothed fits of relationships between elements and explanatory variables.

6.7 References

Abbasi NA, Jaspers VLB, Chaudhry MJI, Ali S, Malik RN. 2015. Influence of taxa, trophic level, and location on bioaccumulation of toxic metals in bird's feathers: a preliminary biomonitoring study using multiple bird species from Pakistan. Chemosphere **120**:527-537.

Abdullah M, Fasola M, Muhammad A, Malik SA, Bostan N, Bokhari H, Kamran MA, Shafqat MN, Alamdar A, Khan M. 2015. Avian feathers as a non-destructive bio-monitoring tool of trace metals signatures: a case study from severely contaminated areas. Chemosphere **119**:553-561.

Ahmed AS, Sultana S, Habib A, Ullah H, Musa N, Hossain MB, Rahman MM, Sarker MSI. 2019. Bioaccumulation of heavy metals in some commercially important fishes from a tropical river estuary suggests higher potential health risk in children than adults. Plos one **14**:e0219336.

Ali H, Khan E, Ilahi I. 2019. Environmental chemistry and ecotoxicology of hazardous heavy metals: environmental persistence, toxicity, and bioaccumulation. Journal of chemistry **2019**.

Badry A, Palma L, Beja P, Ciesielski TM, Dias A, Lierhagen S, Jenssen BM, Sturaro N, Eulaers I, Jaspers VL. 2019. Using an apex predator for large-scale monitoring of trace element contamination: Associations with environmental, anthropogenic and dietary proxies. Science of the total environment **676**:746-755.

Brandis K, Bino G, Spencer J, Ramp D, Kingsford R. 2018. Decline in colonial waterbird breeding highlights loss of Ramsar wetland function. Biological Conservation **225**:22-30.

Burger J, Gochfeld M. 1995. Biomonitoring of heavy metals in the Pacific Basin using avian feathers. Environmental Toxicology and Chemistry: An International Journal **14**:1233-1239.

Costa M, Klein CB. 2006. Toxicity and carcinogenicity of chromium compounds in humans. Critical reviews in toxicology **36**:155-163.

Davis RA, Wilcox J. 2013. Adapting to suburbia: bird ecology on an urban-bushland interface in Perth, Western Australia. Pacific Conservation Biology **19**:110-120.

De Caritat P, Cooper M. 2016. A continental-scale geochemical atlas for resource exploration and environmental management: the National Geochemical Survey of Australia. Geochemistry: Exploration, Environment, Analysis **16**:3-13.

Dudka S, Miller W. 1999. Accumulation of potentially toxic elements in plants and their transfer to human food chain. Journal of Environmental Science & Health Part B **34**:681-708.

Frantz A, Pottier M-A, Karimi B, Corbel H, Aubry E, Haussy C, Gasparini J, Castrec-Rouelle M. 2012. Contrasting levels of heavy metals in the feathers of urban pigeons from close habitats suggest limited movements at a restricted scale. Environmental Pollution **168**:23-28.

Frederick PC, Hylton B, Heath JA, Spalding MG. 2004. A historical record of mercury contamination in southern florida (USA) as inferred from avian feather tissue: Contribution R-09888 of the Journal Series, Florida Agricultural Experiment Station. Environmental Toxicology and Chemistry: An International Journal **23**:1474-1478.

Frith HJ. 1967. Waterfowl in Australia.

Furtado R, Pereira ME, Granadeiro JP, Catry P. 2019. Body feather mercury and arsenic concentrations in five species of seabirds from the Falkland Islands. Marine pollution bulletin **149**:110574.

Hampton JO, Laidlaw M, Buenz E, Arnemo JM. 2018. Heads in the sand: public health and ecological risks of lead-based bullets for wildlife shooting in Australia. Wildlife Research **45**:287-306.

Hollins SE, Hughes CE, Crawford J, Cendón DI, Meredith KT. 2018. Rainfall isotope variations over the Australian continent–Implications for hydrology and isoscape applications. Science of the Total Environment **645**:630-645.

Innangi M, De Rosa D, Danise T, Fozzi I, Giannotti M, Izzo M, Trifuoggi M, Fioretto A. 2019. Analysis of 11 trace elements in flight feathers of Italian Sparrows in southern Italy: A study of bioaccumulation through age classes, variability in three years of sampling, and relations with body condition. Science of the Total Environment **651**:2003-2012.

Kingsford R. 1986. Reproductive biology and habitat use of the Maned Duck Chenonetta jubata (Latham). University of Sydney.

Kingsford R, Norman F. 2002. Australian waterbirds—products of the continent's ecology. Emu **102**:47-69.

Liu L, Liu X-G, Sun Y, Pu Z-H, Xu H-Y, Li W-X, Wang Z-h. 2019. Trace elements in the feathers of waterfowl from Nanhaizi Wetland, Baotou, China. Bulletin of environmental contamination and toxicology **102**:778-783.

López-Perea JJ, Laguna C, Jiménez-Moreno M, Martín-Doimeadios RCR, Feliu J, Mateo R. 2019. Metals and metalloids in blood and feathers of common moorhens (Gallinula chloropus) from wetlands that receive treated wastewater. Science of the Total Environment **646**:84-92.

Malik RN, Zeb N. 2009. Assessment of environmental contamination using feathers of Bubulcus ibis L., as a biomonitor of heavy metal pollution, Pakistan. Ecotoxicology **18**:522-536.

McEvoy J, Ribot R, Wingfield J, Bennett A. 2017. Heavy rainfall triggers increased nocturnal flight in desert populations of the Pacific black duck (Anas superciliosa). Scientific reports **7**:1-9.

McEvoy JF, Hall GP, McDonald PG. 2019. Movements of Australian Wood Ducks (Chenonetta jubata) in an agricultural landscape. Emu-Austral Ornithology **119**:147-156.

Metcheva R, Yurukova L, Teodorova SE. 2011. Biogenic and toxic elements in feathers, eggs, and excreta of Gentoo penguin (Pygoscelis papua ellsworthii) in the Antarctic. Environmental monitoring and assessment **182**:571-585.

Needham C, Orellana L, Allender S, Sacks G, Blake MR, Strugnell C. 2020. Food retail environments in Greater Melbourne 2008–2016: Longitudinal analysis of intra-city variation in density and healthiness of food outlets. International journal of environmental research and public health **17**:1321.

Newth J, Cromie RL, Brown M, Delahay RJ, Meharg AA, Deacon C, Norton GJ, O'brien M, Pain DJ. 2013. Poisoning from lead gunshot: still a threat to wild waterbirds in Britain. European Journal of Wildlife Research **59**:195-204.

Norman F, Mumford L. 1982. Food of the Chestnut Teal, Anas Castanea, in the Gippsland Lakes Region of Victoria. Wildlife Research **9**:151-155.

Peterson SH, Ackerman JT, Toney M, Herzog MP. 2019. Mercury concentrations vary within and among individual bird feathers: A critical evaluation and guidelines for feather use in mercury monitoring programs. Environmental toxicology and chemistry **38**:1164-1187.

Picone M, Corami F, Gaetan C, Basso M, Battiston A, Panzarin L, Ghirardini AV. 2019. Accumulation of trace elements in feathers of the Kentish plover Charadrius alexandrinus. Ecotoxicology and environmental safety **179**:62-70.

Prashanth L, Kattapagari KK, Chitturi RT, Baddam VRR, Prasad LK. 2015. A review on role of essential trace elements in health and disease. Journal of dr. ntr university of health sciences **4**:75.

Smith E, Smith J, Smith L, Biswas T, Correll R, Naidu R. 2003. Arsenic in Australian environment: an overview. Journal of Environmental Science and Health, Part A **38**:223-239.

Weech SA, Scheuhammer AM, Elliott JE. 2006. Mercury exposure and reproduction in fish-eating birds breeding in the Pinchi Lake region, British Columbia, Canada. Environmental Toxicology and Chemistry: An International Journal **25**:1433-1440.

Wretenberg J, Pärt T, Berg Å. 2010. Changes in local species richness of farmland birds in relation to land-use changes and landscape structure. Biological conservation **143**:375-381.

7 Final conclusions and future research interests

7.1 Final conclusions

This study mainly investigated the chemical element transport processes and mechanism among various environmental compartments at different scales based on systematic data sets. The outcomes of this project can provide crucial information for the development of water resources management and conservation under climate change and human activity influence. The novel methodologies and experiment designs presented in this research can be further applied to other studies with similar explanatory variables for predicting and explaining the geochemical tracers transport and accumulate. The key findings of each chapter are as follows:

(1) Chapter 2 reported the introduction of SDMs into mapping the HM distributions at catchment scale. The results showed that the HM hotspots in the soil were identified in the middle reach of the Heihe river basin where intense human activities occur. The variables of the human footprint including distance to road, GDP, and nightlight were found to play the most important roles in soil HM transport among others. The output maps of HM habitats from soil, sediment, and wet deposition well agreed with the influence of industry contaminants, hydraulic sorting, and precipitation washout process respectively, indicating the potential of SDM in modelling the spatial distributions of the HM carriers. In addition, we found the distributions of HMs in the primary carrier of river water were attributed to both of the human and environmental variables.

(2) Chapter 3 found that the chemical evolution of most water types in the BJD and neighbouring catchments was mainly controlled by rock dominance. The spatial distribution of highest values in shallow groundwater sulfates, coincided with the locations of typical industrial cities according to our output maps. In addition to the sulfates replenishment from surface water and deep groundwater, the sulfates in shallow groundwater were also affected by regional variables such as population density and the Mg^{2+} and EC contents in surface water. The accuracy of predicting the spatial distributions of groundwater has been improved by introduction the EMMA and SDM methodologies.

(3) Chapter 4 reported that the occurrence of sulphates (SO_4^{2-}) and nitrates (NO_3^-) in river water were highly correlated (up to 0.70), providing strong evidence of pollution from nearby mining activities. Levels of arsenic and cadmium were high in first and fifth order streams, where mining activities were most concentrated. The modelling results showed that geographical patterns and mining activity account for predicting HM distribution, and WCP can be reasonable predictors to trace soil mining pollution.

(4) Chapter 5, results found that on the annual scale, both the slope and intercept of the LMWL at Nam Co basin were higher than those of the global meteoric water line (GMWL), and these characteristics suggested that there were significant moisture sources from the continental landscape surface throughout the entire year at the monsoon-Westerly jet stream transition areas. The important explanatory variables of humidity and three temperature factors in the top model of d-excess also supported this conclusion. Seasonally, the lowest slope and intercept values of LMWL during the monsoon periods indicated the contributions of marine moisture. The lighter isotopic values in the monsoon season also reflected the strong rain out process by the high precipitation amounts in this period. We also found that isotopic signatures in different seasons were dominated by different local climatic factors according to the results of top models. For instance, the wind variables were important variables to explain the isotopic tracers during the Westerly season.

(5) Chapter 6 reported that the environmental variables had a greater effect on essential elements, while human footprint variables had a great impact on both the toxic and essential ones. Through the top models, most of the elements were basically affected by both environmental and human footprint variables, but the essential element Cu was only affected by environmental factors, and the toxic elements Al, Cr, Li, and Pb, as well as essential Mn were only affected by human activities.

7.2 Future research interests

The future research interests are to: (1) study the responses of HMs transport process to natural environment and land use changes as long-term period heavy metal observation data sets were not mentioned in current study; (2) mapping the dynamics of pollutants in groundwater based on historical records by SDMs; (3) conduct future research with more tracing elements, climate factors and lithology information to provide a more reliable method for analysing mining pollution mechanism at alpine catchments around the world; (4) consider the influence of regional climatic factors to the isotopic signal variation of precipitation at Nam Co basin; and (5) provide further insights of bioaccumulation mechanism by more water bird species around Australia, as well as other explanatory variables, such as wet/dry spell lengths (i.e., consecutive rain/dry days) and numbers of rainfall days and total amount over a threshold.