Development of environment monitoring system and its application in the Baiyangdian Lake basin, North China Plain

February 2023

Wang Guangwei

Graduate School of Horticulture

CHIBA UNIVERSITY

（千葉大学学位論文）

環境モニタリングシステムの開発
および中国華北平原の白洋淀流域への応用

2023 年 2 月

千葉大学大学院園芸学研究科

環境園芸学専攻緑地環境学コース

王光偉

Contents

List of Figures

List of Tables

Abstract

Real-time monitoring of water environmental parameters for soil and water is important for ecological environmental protection. Automatic monitoring has gradually become the main way of environmental monitoring, but it is far from replacing manual monitoring at present. The focus of this study is to establish a water environment monitoring system that includes sensors, data transmission, store and analysis for environmental assessment, emergency management and resource development.

The main works of this thesis are as the following:

(1) Environmental monitoring system based on 5G (NB-IoT) technology was designed, developed and applied in the Baiyangdian Lake and one of its inflowing rivers (Fu River). The system consists of three major parts: sensing terminal, NB-IoT network communication unit and monitoring platform. The terminal device based on the main control chip of STM32 and communication module of BC95, can be compatible with sensors to monitor multiple parameters. The communication unit uses low-power design and the retransmission mechanism to increase the system working life and guarantee the data transmission with high quality. The monitoring platform was designed and developed in PHP programming language to realize the data collection, storage, and real-time display in cloud. It includes five modules, comprehensive information, basic data management, real-time data management, alarm management and system management, and realize the functions of integrated management, data export and statistics. The real-time flux of water and pollutants including total nitrogen (TN), and total phosphorus (TP) into the lake can be estimated by using the automatic monitoring data and manual monitoring data.

(2) The nonlinear filtering problem of Global Navigation Satellite System (GNSS)/ inertial navigation system (INS) under non-Gaussian noise has been studied. A modified maximum correlation entropy volume Kalman filter (MCCKF) based on the resampling-free sigma point update framework (SUF) is proposed to achieve robust multi-sensor information fusion. The effectiveness of the algorithm was verified through numerical simulations and field tests on GNSS/INS. The experiment results show that RMCCKF has a better performance than MCCKF, which reduces the root mean square error of the land robot heading from 1.77 degrees to 0.26

degrees, in addition to improving the localization results. This idea is applicable to unmanned vehicles, enhancing the reliability of autonomous navigation and meeting the needs of long-term autonomous work of unmanned vehicles.

(3) The load estimation and source analysis of major pollutants in Baiyangdian Lake basin and the flux of pollutants to the lake through the major rivers were conducted. And the spatial and temporal distribution of pollutants in the lake were also analyzed. The output coefficient method has been used to calculate the non-point source pollution load of Baiyangdian Lake basin in 2020. Furthermore, LOADEST model was used to estimate the fluxes of COD, TN, and TP. The result of the model in Fu River was compared with that estimated by monitoring data combing with the in-situ NB-IoT monitoring system. The results are similar and in acceptable range which suggested that the automatic system can provide continuous monitoring benefiting for assessment of pollutants. If the automatic monitoring technology in this study can be used to realize the synchronous daily monitoring of water quality and quantity, the time average dispersion error caused by the asynchronous data frequency can be avoided in the pollutant flux calculation process, and the deviation caused by the time domain discrete sampling of continuous flux can be reduced, thus significantly improving the accuracy of flux calculation.

Though the monitoring technology of autonomous navigation has not been applied in the study area yet at this point, it does provide a possible way for monitoring the regional environmental parameters by unmanned vehicles. Combing with the in-situ automatic monitoring system based on NB-Iot technology, the environment monitoring system was developed in this study. In the future, the platform embedded in the LOADEST model can analyze the monitoring data automatically.

Keyword: Autonomous navigation; Baiyangdian Lake; Pollutants source; Load estimation; NB-IoT; Environment monitoring; Export coefficient model; LOADEST model.

要旨

土壌と水は地域生態環境保全にとって非常に重要である。そのためにリアルタイム環境モニタリングの自動化が急務になっている。本研究では、環境要素の実測、データの転送・保存・分析・表示などの仕事をまとめて完成できる環境モニタリングシステム構築に焦点を当て、5 G（NB-IoT、狭帯域 IoT）技術はデータ伝送の安定性を高め、ロバストなマルチセンサ情報融合アルゴリズムを開発し、モニタリングデータに基づき中国河北省白洋淀流域の水環境を評価した。

具体的なことが次のようにまとめられた。

（1）5 G（NB-IoT）技術を用いて白洋淀流域環境モニタリングシステムを構築した。このシステムは、センシング端末、NB−IoT ネットワーク通信ユニット、観測プラットフォームの 3 つの主要な部分から構成されている。STM 32 マスターチップ及び BC 95 通信モジュールが組み込まれたセンシング端末は、複数のセンサーの信号を同時に対応できる。NB−IoT ネットワーク通信ユニットが低消費電力設計とデータ再伝送機能を採用しているため、ユニットの寿命が長く、データ伝送の品質も高い。観測プラットフォームは PHP プログラミング言語で設計開発され、データの収集、記憶、リアルタイム表示を実現した。このプラットフォームには、統合情報、基礎データ管理、リアルタイムデータ管理、アラーム管理、システム管理の 5 つのモジュールが含まれています。これら 5 つのモジュールは、統合管理、データの出力、統計解析などの機能を備えている。

（2）非ガウスノイズ条件での GNSS/INS の非線形フィルタリングを調べた。ロバストなマルチセンサ情報の融合を実現するため、再サンプリング無しシグマ点更新フレーム（SUF）に基づく補正最大相関エントロピー体積カルマンフィルタ（MCCKF）が提案された。GNSS/INS の数値シミュレーションおよびフィールドテストより、アルゴリズムの有効性が検証された。また、RMCCKF（再スケーリング MCCKF）は MCCKF より優れ、測位精度を改善したことが分かった。これより自律ロボットのナビゲーション精度を向上させ、巡回モニタリングが可能になった。

（3）開発した環境モニタリングシステムは白洋淀への流入河川である府河に設置され、河川水位、水温、電気伝導度のリアルタイムモニタリングが成功した。これより観測データの非同期性に起因する時間平均分散誤差を回避できた。また、白洋淀湖の主要

汚染物質の負荷・フラックス計算と発生源分析を行い、汚染物質の時空間分布も把握した。出力係数法を用いて 2020 年の白洋淀流域における非点源汚染負荷を計算し、白洋淀湖の汚染負荷を求める。今後 LOADEST モデルを観測プラットフォームに組み込むことによって、モニタリングデータの自動解析を実現する。

キーワード：インテリジェント監視、NB−IoT 技術、自律的なナビゲーション汚染物質源負荷、出力係数法、LOADEST モデル。

Acknowledgement

First of all, I would like to thank my doctoral supervisor, Prof. Changyuan Tang for his kindly guidance on this research work. As an interdisciplinary student, Prof. Tang paid much more attention and patience on my thesis. Prof. Tang has a broad multi-disciplinary vision, especially he has very abundant experience and unique insights in the design and application of equipment. He has always maintained the spirit of research and full enthusiasm in new research fields, which is inspiring me all the time.

I would like to thank Dr. Zhenglun Yang and Dr. Han Zhang of Chiba University for their great help in this study. Thanks are also given to Dr. Congke Gu, Dr. Cuicui Lv, Dr. Zhixiong Zhang, and Dr. Wenbo Zheng from Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Dr. Bingbo Cui of Jiangsu University for their guidance and help in my scientific research. I would also like to thank Dr.Yizhang Zhang of the Tianjin Branch of the Chinese Academy of Environmental Sciences and all my colleagues for their support.

The work of this paper was supported by the National Key Research and Development Program of China (2021YFD1700500, 2018YFD0406502, 2018YFD0800306) and the 100-Talen Project of Chinese Academy of Sciences. The research group of water and solute cycle from Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences provide the field observation data.

Thanks are given to my parents and sister for their selfless support. Thank my wife, Dr. Shiqin Wang, and my daughter for their encouragement, which gives me the confidence and motivation to continue my studies. Thank you for your support and selfless love. Finally, thank all the people who helped me.

Chapter 1 Introduction

1.1 Significance of this study

With the development of the water management, the monitoring of the water environment and its related environmental indicators is the basis, and it can provide data to support the reasonable environmental protection. Application of Internet of Things (IoT), Big Data, Cloud Computing, Machine Learning (ML) and other latest information technologies are leading various industries into a turning period. The application of latest information technology helps to achieve environmental monitoring to meet the needs of environmental assessment and management. Real-time monitoring can effectively ensure the safety of drinking water, irrigation water. It is of great significance for regional flood control and drought relief, water pollution management and control, water ecological restoration, as well as for ensuring the effectiveness of early project construction of water environment management.

1.2 Research Status

To monitor hydrological and water environment indicators, the combination of manual monitoring and automatic monitoring is adopted as required. In the 1970s, developed countries initiated the automatic monitoring of surface water (rivers, lakes, etc.) environment, mainly for the hydrological and water quality data of major basins. In addition, the automatic monitoring system aimed at monitoring pollution sources has also been used widely. More attention has been paid in the real-time and dynamic monitoring of water environment. Recently, innovative technologies such as the IoT and mobile Internet provide opportunities and means for building new water environment monitoring systems.

Nowadays, the comprehensive and automatic water environment monitoring systems have been operated in many basins in the world. For example, the Critical Zone Observatory initiated by the U.S. Foundation and the European Terrestrial Environmental Observations (TERENO) platform. In 2017, Iacono L et al. pointed out that an important reason for limiting the application of wireless sensing technology in water environment monitoring in large-scale is the high price of some wireless sensing devices. Therefore, they proposed to achieve the application of wireless sensor network in

water resource protection, but no specific design scheme is proposed (Chen et al., 2015, Wang et al., 2017).With the development of multi parameter water quality sensors, which are becoming more compact, cheap, reliable, and low power within a limited budget. This makes it possible to deploy more multi-parameter sensors in water environment. The monitoring network can continuously capture the spatial distribution characteristics and time evolution of key parameters of water environment, which is helpful to timely understand the change process of water environment and to monitor emergency water pollution events. Combined with mobile nodes such as water environment monitoring unmanned vessels, the coverage of regional water environment monitoring networks can be improved further. Although monitoring of the water environment based on the IoT has attracted extensive attentions in the world, the research in this field is still facing challenge. For example, how to ensure effective data transmission of monitoring equipment in the field; how to design high reliability and low-cost automatic navigation to meet the long-term automatic monitoring of unmanned vehicles for providing continuous dynamic detection means to monitor environmental elements in large water bodies. Finally, the selection of water environment assessment models and the comprehensive construction of analysis methods still need to be further studied to ensure the smooth progress of multi-monitoring.

1.3 Technology Overview

1.3.1 Narrow Band Internet of Things（NB-IoT）

NB-IoT (Narrow Band Internet of Things), a technology standard defined by the 3GPP standardization organization, is a narrowband RF technology designed for 5G IoT. NB-IoT technology was first proposed jointly by Huawei company and Vodafone, a British telecom operator, and proposed NB-M2M (Machine to Machine) to 3GPP, the NB-IoT standardization organization, in May 2014, and it fused with NB-OFDMA narrowband orthogonal frequency division multiple access technologies in May 2015 to form NB-IoT (In September 2015, the NB-IoT technology was fused with NB-LTE technology to form the NB-IoT technology scheme). After complex testing and evaluation, the physical layer standard of NB-IoT technology was fully defined in March 2016. On June 16, 2016, the 3GPP Radio Access Network Technical Specification Group meeting adopted the

NB-IoT technical protocol, all the NB-IoT physical layer specifications were frozen, and the standardization work was completed. After completing the development of NB-IoT performance standards in September 2016, Huawei Company, which participated in the standardization of NB-IoT technology, launched the first commercially available NB-IoT chip developed by the industry in a flash. To promote further development of the mobile IoT market, the new LPWAN network technology represented by NB-IoT is being extensively researched and applied in various fields. In April of the same year, China Telecom, Huawei, and others cooperated to jointly develop a new generation of IoT smart life solutions. With the joint efforts of China Mobile, Vodafone, Huawei, Ericsson, Qualcomm, ZTE, MediaTek, and other enterprises, the uplink transmission solutions of multiple vendors were organically integrated into a unified technical solution, which facilitated the successful completion of standardization. With the joint efforts of the whole industry, NB-IoT technology has matured rapidly, with remarkable improvement in terms of technical standards, network deployment, and application implementation. European operators cooperate with Huawei to focus on NB-IoT and actively conduct field tests and existing network upgrades. Deutsche Telekom conducted existing network NB-IoT upgrades and deployed NB-IoT networks in Germany and the Netherlands respectively in March 2017, with applications in smart heating, smart meter reading, smart garbage cans, etc. Vodafone conducted existing network NB-IoT upgrades in 2017 and deployed NB-IoT in EU 11 countries to deploy NB-IoT, has opened smart meter reading and smart parking business rights in Spain. Telecom Italia opened a dozen sites for outfield testing at the end of 2017 and commercialized in some cities in the second quarter of 2017. Spain and France Telecom (Orange) conducted outfield testing in Belgium and Chile, respectively, at the end of 2017.

Overall, major operators and chip vendors of the world are studying NB-IoT technology. The global mobile IoT main market is not only related to network-related hardware, software, services, and connectivity, but also the market related to various IoT applications across industries will grow significantly in the future, with broad prospects in key areas such as industrial control, transportation and logistics, and public utilities.

NB-IoT is a narrowband 5G IoT technology based on a cellular network, so called the next-generation IoT technology. The signal strength of NB-IoT has increased by 20dB compared with the existing IoT signal, which is equivalent to the ability to penetrate an additional wall. That means

the ability to send signals from underground facilities, which currently has the characteristics of low power consumption, low cost ,wide coverage, and large connection, is very suitable for the demand of field monitoring scenarios with small data volume. Therefore, the development of this technology provides a good application prospect in the field of the environment monitoring.

1.3.2 GNSS/INS Integrated Navigation System

1.3.2.1 Global Navigation Satellite System (GNSS)

GNSS (Global Navigation Satellite System) is the collective name for the global navigation satellite systems including GPS (USA), GLONASS (Russia), Compass (China) and Galileo (Europe).

Inertial Navigation System (INS) means that the velocity and position of an object in the navigation coordinate system are calculated from the Inertial Measurement Unit (IMU) module.

GNSS systems can achieve metre-level positioning accuracy and centimetre-level positioning accuracy using RTK/CORS ground-based augmentation stations. GNSS has the advantage of not accumulating positioning errors, but GNSS is a radio navigation system and is susceptible to interference. For moving objects, the signal can sometimes be blocked and the GNSS positioning will be interrupted. Inertial navigation, on the other hand, is relatively more accurate and less susceptible to interference, but errors can accumulate. Combining inertial navigation INS and GNSS navigation can achieve complementary advantages and good navigation results. combined INS/GNSS navigation is an important type of combined inertial navigation system.

1.3.2.2 Robust multi-sensors information fusion

With the rapid expansion of self-driving cars, the stability of GNSS in the face of signal fading is of great concern. Because of the complementally error properties of GNSS and INS, their integration of them is a preferred candidate for land vehicle navigation when the satellite signal is blocked from time to time. Within the design of composite navigation, data fusion methods play a moderating role, of which KF and its variants are widely used. Under the Gaussian assumption of noise and predictive distribution, KF is the best choice for linear state space filtering problems. However, there are no completely linear models in practice, and model instability and non-Gaussian noise will severely degrade the KF-type filter performance if not properly addressed.

A number of sub-optimal filtering techniques have been developed under the KF framework, such as the Extended Kalman Filter (EKF), which linearises the model function at a single pre-identified point. The nonlinearity error of EKF can be further eliminated by employing the

deterministic sampling algorithm, which is termed as sampling-based KF, such as unscented Kalman filter， cubature Kalman filter (CKF) (Arasaratnam and Haykin, 2009), high-degree CKF (Jia and Xin, 2019), to name a few. The sampling-based KF approximates the probability density function (PDF) of the underlying state directly, which however inherits the problem of KF, i.e., updates the state iteratively by propagating Gaussian moments only and depends strongly on the prior information of the filter model (Cui et al., 2018, Hu et al., 2020). In the process of non-Gaussian and nonlinear filtering problems, a particle filter (PF) is an ideal tool for approximating the whole shape of the posterior PDF utilizing a large number of weighted particles (Arulampalam et al., 2002). However, the computational complexity of PF is huge in case the dimension of the state model is high, which needs improvement for real-time application (Schon et al., 2005).

In order to improve the KTF robustness under non-Gaussian noise, Robust Huber-based filters were developed, which outperform their KF counterparts when the measurement noise is non-Gaussian (Chang et al., 2015, Karlgaard and Schaub, 2007). A filter update framework based on the maximum correlation criterion (MCC) is then proposed for state estimation under non-Gaussian processes and measurement noise (Chen et al., 2015, Wang et al., 2017). By using regression techniques, the MCC-based renewal is further extended to the non-linear measurement equation, and modelling results show that the MCC-based filter is heavily dependent on the choice of kernel bandwidth (Li et al., 2020b, Liu et al., 2019b). Both MCC- and Huber-technique generated robust filters exhibit degraded performance when the Gaussian assumption is valid, as they do not fully exploit the features of the non-Gaussian distribution. In order to model noise and state with greater accuracy, student t-based filters (STF) (Roth et al., 2013, Tronarp et al., 2019) and robust student t-based KFs (RSTKF) (Huang et al., 2017) have been proposed, in which STF reduces the sensitivity of the KF to individual observed outliers and RSTKF reduces the sensitivity of the KF to a single observation by using variational Bayesian (VB) transformation of the student t-based state space model (SSM) into a hierarchical Gaussian SSM to enhance the STF. in the design of the adaptive KF, unknown noise variance, state and model parameters can be estimated concurrently by VB (Sarkka and Nummenmaa, 2009). However, the degree of freedom parameters of the RSTKF can neither be approximated to absolute infinity nor estimated by VB based on a limited number of observation, so both adaptive and hybrid distribution approximations have been designed to improve the KTF under non-Gaussian noise (Huang et al. 2017, Li et al. 2020a, Zhu et al. 2021). Most of the robust filters mentioned above focus only on the update of the measurement of the KTF and ignore the uncertainty present in the state dissemination phase. In contrast to measurement noise, where both process noise and measurement noise are unknown and non-Gaussian in real models, process noise cannot be estimated based on a limited number of observations. State prediction covariance is

estimated instead of process noise (Huang et al., 2018), but this degrades the KTF performance as it is required to adjust for process noise (Chang et al., 2021). The integration of Global Navigation Satellite Systems (GNSS) and Inertial Navigation Systems (INS) will be affected by environmental interference and severe vehicle mobility, producing scattered particle noise and heavy tail noise. Robust MCC-based filters have been widely used for INS-based combined navigation, and simulation results show that they perform well with non-Gaussian noise (Deng et al., 2020, Liu et al., 2019a, Liu et al., 2018b, Xu et al., 2020). It has been noted that nonlinear functions with instantiated points can conserve information on non-Gaussian and odd-order terms of Gaussian moments (Yuanxin et al., 2005, Jia and Xin, 2019). Yet, in the generation of sigma points, only the even-order terms of Gaussian moments are spread in an iterative update framework based on the sampled KF. A robust CKF is developed for GNSS/INS suffers from signal outages, and a novel sigma-point update framework (NSUF) is employed to suppress the filtering covariance inflated by observation missing (Cheng et al., 2014, Cui et al., 2018, Cui et al., 2021, Tian and Cheng, 2013). The NSUF is further extended by using prediction sigma points error from the likelihood function, and the posterior sigma points error is updated by the linear transform of the model prediction residue (Huang et al., 2016). The direct modification of sigma points without Gaussian information-based resampling is named resampling-free SUF (Straka and Dunik, 2020), and the prediction residue-based methods are a special implementation of the resampling-free SUF.

Huang et al.'s (2016) resampling-free SUF was performed in published work recently, which showed that the sigma point remnants of the similarity function are more indeterminate than those predicted by the kinetic function (Cui et al., 2017, Cui et al., 2019). In this study, an estimated framework based on MCC is used to improve the robustness of the CKFs based on the non-resampled SUF, and this increases the robustness of the estimation updates as well as reduces the sensitivity of the non-resampled SUFs to the changing estimation noise. The main achievements of this work include: 1) MCC-based covariance rescaling improves the transformation of the system function instantiation point-based non-resampled SUF; 2) the measurement update of the MCC is augmented by a non-resampled SUF that channels more information into the filtering returns of the MCC; and 3) the validity of the approach is verified by field tests and numerical simulations of the filtering model.

This method is also applicable to the precise navigation of unmanned ships, short for unmanned surface vehicles, which are surface robots that can navigate fully automatically on the water surface according to a preset mission based on sensor information (Karlgaard and Schaub, 2007), and was first used mainly in the military field. As early as the 1970s the U.S. military has been loading unmanned ships on naval systems.

1.3.3 Non-point source environmental assessment

1.3.3.1 Classification of non-point source pollution sources

Water environmental quality was affected by either the point source and non-point source pollutants. The point source pollutants are easy to be monitored. Due to the randomness of non-point source pollution and the uncertainty of the scope and pathway of pollutant discharge, the most effective and direct method is to establish simulation models when monitoring the non-point source was difficult. Quantitative source analysis of non-point source pollution in basins is the fundamental basis for pollution control from sources. It was necessary to access the environmental quality when combing the model and monitoring the non-point source pollution.

Non-point source pollution includes soil sediment particles, nitrogen and phosphorus nutrients, pesticides, and atmospheric deposition. Pollutants can enter the receiving environment through surface runoff, soil erosion, and agricultural drainage, and are characterized by randomness, extensiveness, lag, ambiguity, and difficulty in monitoring (Liu et al., 2018a).

According to the regional and process characteristics of the occurrence of non-point sources, they can usually be classified into types of agricultural non-point source pollution, urban domestic pollution, surface runoff pollution from mining areas, surface runoff pollution from forest areas, other soil erosion and loss, and atmospheric rainfall and dust pollution. Among them, agricultural non-point source pollution has the greatest impact, from agricultural runoff loss, discharges of village household waste, rural livestock and poultry breeding manure, pesticides and organic matter, etc. The pollutants produced are large in quantity and complex in composition. Temporally, the pollutant discharge is intermittent, with pollutants accumulating on sunny days and discharging on rainy days; spatially, it is strongly spatially heterogeneous due to the topography, the changes in land use, and the associated agricultural activities.

Surface runoff caused by rainfall is the most important non-point source pollution source, including non-permeable and permeable surface runoff. Surface runoff is the main driver of soil erosion, which loses the surface organic matter layer of soil, while dissolved pollutants and suspended particulate matter enter water bodies with surface runoff. The intensity of soil erosion depends on rainfall intensity, topography, land use practices, and vegetation cover.

1.3.3.2 Calculation of non-point source pollution load

The discharge of non-point source pollutants is extensive, random, uncertain, and characterized by multiple sources and routes (Chen et al., 2018b), leading to difficulties in monitoring and load measurement, especially at the basin scale where non-point source pollution has prominent multiple compound pollution characteristics. However, obtaining accurate quantitative data on agricultural surface source pollution is in turn the basis and key to implementing pollution control management.

Research on agricultural surface pollution measurement started in the United States in the 1960s and 1970s, and a series of surface pollution measurement models, including output coefficient models and mechanistic models, were developed in the United States. Young et al. (1989) from Institute of Agricultural Research, U.S. Department of Agriculture, published an article on "Models for Agricultural Surface Source Pollution Assessment", which was the first attempt to use computer models to analyze agricultural surface source pollution problems. In the 21st century, agricultural surface source pollution measurement research has attracted extensive attention worldwide, especially in Asian countries such as Japan, Korea and China, which have been most active in recent years. Since agricultural surface pollution is closely related to hydrometeorology, soil structure, crop type and environmental management level, pollution measurement methods must be adapted to the actual situation of the study area in order to have rigorous and objective measurement results. Therefore, in addition to general research on model application, validation and comparison, scholars usually choose to improve the model based on the actual situation in the study area. For example, Boulange et al. (2014) introduced the Japanese pesticide concentration model (PCPF, Pesticide Concentration in Paddy Field) into the SWAT model and established a new pesticide degradation and transport transformation model (PCPF- 1 C SWAT) for paddy fields.

Quantification of non-point source pollution systems at the basin scale is the key to effectively promote non-point source pollution control. After decades of development, nowadays, studies on pollution loads include urban runoff pollution, pollution load prediction, and model coupled simulation directions. For example, Yazdi et al. (2021) studied the influence of land use characteristics in cities on pollution loads in basins and found that TSS was significantly greater in open space and industrial land than in other land, and TN concentration was positively correlated with impermeability and TP concentration with lawn cover. Li et al. (2021a) conducted a study

related to pollution model coupling and verified the applicability of LSTM-BP model in basin-scale pollution load simulation. Kang et al. (2021) predicted the urban large basin pollution load and reduced the treatment cost of total Zn pollution. In summary, the pollution load studies have been focused on point source pollution, and data-based studies such as model simulations, with non-point source pollution, especially urban non-point source pollution studies, are the key directions of pollution load studies.

Mechanistic models are constructed based on the process of developing non-point source pollution and are simulated using mathematical models, and this method considers both the input and output of pollutants. For example, SWAT model (Chen et al., 2018a, Uribe et al., 2018), AnnAGNPS model (Zema et al., 2018), and SPARROW model (He, 2014, Li et al., 2017) are among the mechanistic models. Mechanistic models consider the process of pollutant migration transformation and can better predict the pollutants (Ouyang et al., 2017, Xueman et al., 2020). Yet, such models require large amounts of data to confirm whether they are applicable in the local geographic area, which makes their application and scaling difficult (Adu and Kumarasamy, 2018).Aggregate models, also known as "black box models", are constructed based on empirical relationships between pollution loads and influencing factors (Wang et al., 2007, Wang et al., 2010b). The pooled model uses multiple pooling parameters to describe the nutrient dynamics of a basin and constructs empirical equations based on hydrological, water quality, rainfall, and other monitoring data to estimate the non-point source pollution load.

1.3.3.3 Application of the output coefficient model

Among the non-point source pollution aggregate models, the output coefficient model is commonly used for their explicit physical significance and simplicity of presentation, and is the classical empirical model for agricultural surface source pollution accounting (Rong et al., 2017). The output coefficient model is a simple method used to quantitatively identify non-point source pollution from different land use types (Shen et al., 2011, Zhang et al., 2014, Wang et al., 2020). In the 1970s, scholars constructed a model for the relationship between lake nutrient loads and land use based on data from 928 basins (Uttormark et al., 1974), which was the earliest statistical model established to simulate non-point source pollution loads. Johnes (1996) based on the research results of the earlier statistical model, and then further refined and deepened it to form the classical output

coefficient model. The basic idea is to use the relatively easy to obtain data of each pollution unit (such as human, livestock and poultry or unit land area, etc.), establish the relationship between pollution units and surface source pollution output through multivariate linear correlation analysis, calculate the pollutant output, and then accumulate the pollution load of different unit source types to estimate the potential generation of surface source pollution within the study area. This is a simple non-point source pollution load estimation method. Although the model ignores the complex migration and transformation process of surface source pollution, its advantages include simplicity of operation, low data requirements, avoidance of the problem of large amount of data required to support physical model simulations, and robustness, which make it widely used in China. When other scholars apply the model, they usually superimpose some environmental factors to modify the output coefficients in order to better adapt to the local environmental conditions to improve the simulation accuracy of the model (Rong et al., 2017, Li et al., 2017, Guo et al., 2008). The improved output coefficient model has been widely applied in the directions of spatial and temporal distribution of pollutant loads, identification of pollution critical source areas, and formulation of pollutant prevention and control measures.

The output coefficient model has been widely used in China because of its flexibility and high adaptability to different regions. The calculation area of the output coefficient method can be a well-defined boundary basin. For example, Ren and Guo (2015) applied this method for the study of surface source pollution load in the Baoxiang River basin in Yunnan, and Li et al. (2016) for the estimation of the catchment area of Xinlicheng Reservoir, a water source in Changchun, Yang et al. (2015) for the estimation of Chongqing section of the Three Gorges reservoir area and Zhang Likun et al. (2017) for the non-point source pollution in the Hulan River basin risk analysis, etc. The calculation area of the output coefficient model can also be different levels of administrative units, such as the estimation of agricultural surface source pollution in Beijing by Liu et al. (2011) and the analysis of agricultural surface source pollution in Guangdong Province by combining GIS and ECM models by Ye et al. (2013). ECM studies have good applicability in different spatial and temporal scales from small to large scales.

1.4 Purpose of the study

The purposes of the study are (1) to build a stable IoT network to monitor the major water

environmental parameters which has been used in the Baiyangdian lake and one of the inflowing rivers (Fu River); (2) to develop high precision navigation algorithm for unmanned vehicles which can be used in spatial monitoring of water quality; and (3) to establish a method for estimating pollution load of the Baiyangdian lake basin and pollutant fluxes flowing into the lake when comparing the results of the model with those by combing with the monitoring data by IoT monitoring system and manual observation data.

1.5 System composition and research content

In view of the specific water environment and characteristics of the Baiyangdian Lake basin, combined with the situation of each river in the basin, with the goal of building a monitoring system suitable for the environmental assessment and emergency management of the basin, this study mainly focuses on the construction of monitoring system and data analysis technology. The system structure is shown in Figure 1.1.

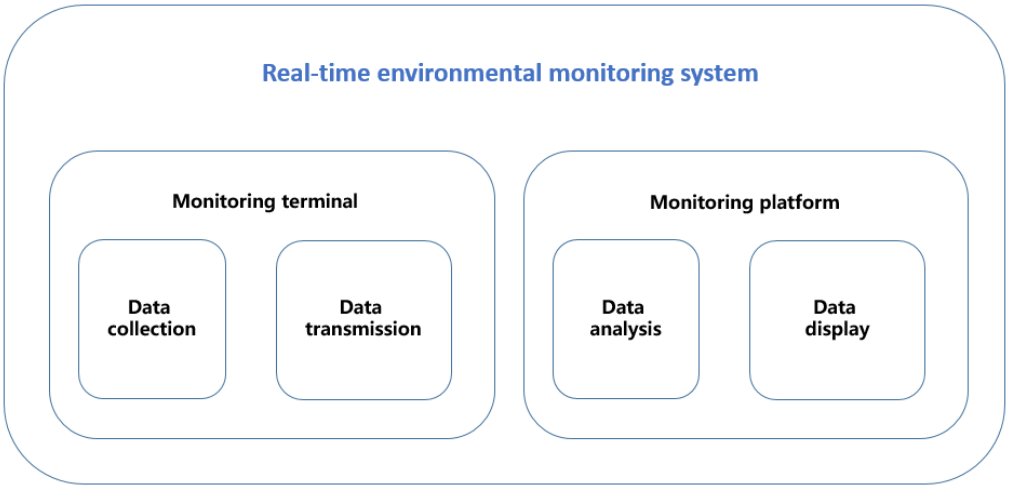

Figure 1.1 The structure of monitoring system

The system is composed of monitoring terminal and monitoring platform. The monitoring terminal includes communication module, data collection module, storage module, processing module and sensors. The monitoring platform includes data display and data analysis modules. This research focuses on the NB-IoT based water environment monitoring terminal, data transmission stability, high-precision integrated navigation and environmental assessment model in detail.

The dissertation contents are shown in Table 1.1 and Figure 1.2, respectively.

Table 1.1 Content structure and scientific subjects discussed in the study

Chapter	Contents	Scientific subjects
Chapter 1	1) Reviews on environment monitoring, IoT, autonomous navigation, and water environment assessment.	
Chapter 2	1) Establishment of a monitoring system by using NB-IoT technology. 2) Development of a retransmission mechanism and a data transmission platform.	1) Characteristics of NB-IoT for data collection. 2) Efficient data transmission through a retransmission mechanism
Chapter 3	1) Development of robust multi-sensors information fusion algorithm. 2) Numerical simulation and field test on GNSS/INS.	1) Nonlinear filtering for GNSS/INS under non-Gaussian noise 2)An improved method of maximum correntropy cubature Kalman filter (MCCKF).
Chapter 4	1) The introduction of the case study area 2) The flux calculation of major pollutants into the lake in 2020 by using continuous flow data and discontinuous water quality data	Calculation method of pollutant flux from hydrological and water quality monitoring data was demonstrated for Baiyangdian Lake.
Chapter 5	1. The pollutants flux entering the Baiyangdian Lake in 2020 was calculated using the point source data and non-point source load calculation model. 2. Comparing of the above two calculation methods.	The calculation method, process and effect of non-point source models were demonstrated for Baiyangdian Lake. Comparison of simulative and monitoring method was made.
Chapter 6	General conclusions and further works	

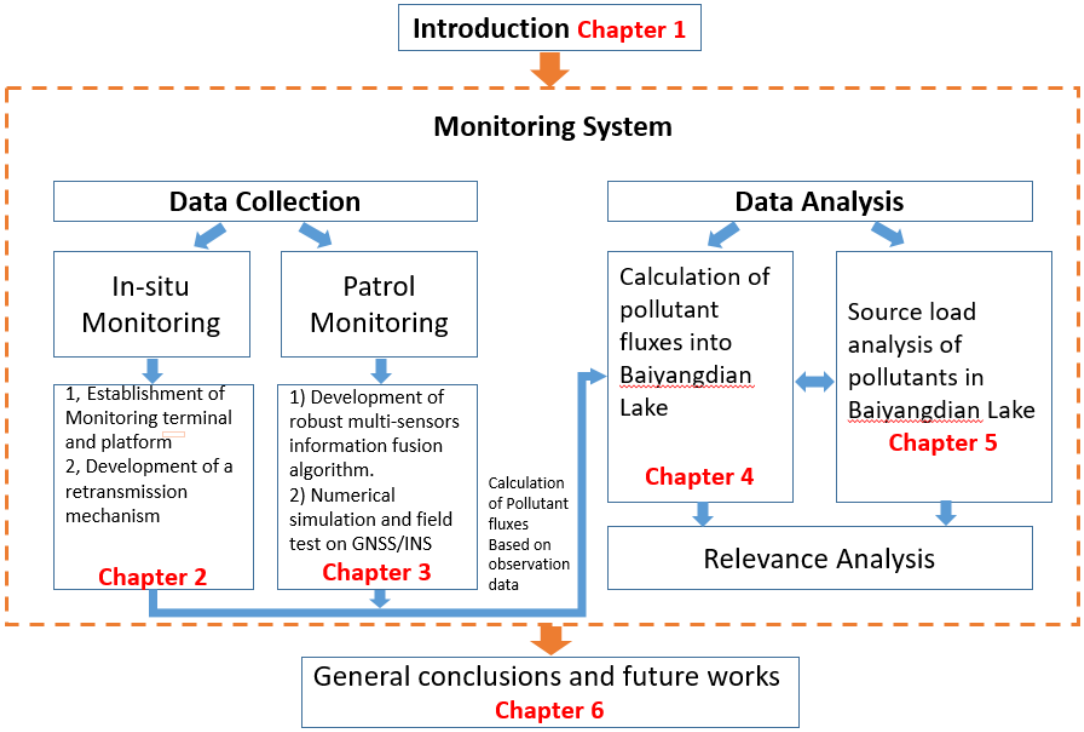

Figure 1.2 The overall research scheme of the dissertation

Chapter 2 Establishment of in-situ environmental monitoring system based on NB-IoT technology

High power consumption, high-cost and relatively limited number of connections exist in application of the 2G/3G/4G wireless communication technology. Short-range IoT transmission technologies such as WiFi, Zigbee and Bluetooth, have weak coverage and limited communication distance. In this study, several wireless communication technologies of Low Power Wide Area Network (LPWAN) were compared. A narrowband IoT was used to design and develop the surface-groundwater monitoring system. The system is based on NB-IoT (narrowband cellular) communication technology to realize an IoT system where a single gateway can receive data from multiple sensors to monitor multiple parameters of environment over long distances (up to 20 Km). The system are in the long time scale, with scientific and reasonable time interval (the highest time resolution of 30 minutes), which has been successfully applied in the long-term continuous monitoring of the key monitoring section in observing of water depth (being used to calculation of water level for surface water and groundwater), water temperature (T) and electrical conductivity (EC) and also it can be used in monitoring other water quality/quantity indicators. The automatic monitoring system of surface water and groundwater quantity/quality can accurately grasp the status and characteristics of water environment.

The key components of this chapter are as follows.

(1) The key technologies of NB-IoT communication technology are analyzed, and the overall scheme design of NB-IoT wireless sensing system for surface-groundwater monitoring is proposed. The system consists of four major parts: sensing terminal, NB-IoT network, transmission mechanism and software monitoring platform. Applying the sensing terminal based on the main control chip STM32, communication module BC95, the node embedded software is developed and the low-power processing mode is analyzed, which has been applied successfully in monitoring of surface water and groundwater level, water temperature (T) and electrical conductivity (EC), and also be used in monitoring the water content, T and EC of soil water. In addition, according to the monitoring needs, dissolved oxygen (DO), pH, nitrate and other monitoring sensors of water environment indicators can be induced into this system.

(2) The monitoring software and database management system was developed by editing the upper computer software to realize the data monitoring, storage, and management functions. The real-time monitoring and early warning function was conducted through the real-time data obtained from the system and the data managed by the database. Using the down-command mode and Iot platform with online upgrade, the firmware of the main control board and module of data acquisition can be changed to realize remote upgrade, change of parameters such as acquisition frequency and transmission frequency, which saved operation and maintenance costs and a lot of human and material resources. Data collection and reporting monitoring, storage and real-time display are realized on the local server system. The software system includes a total of comprehensive situation page, data management, alarm management and system management functions, realizing the functions of integrated management of each part, data export and statistics.

(3) Automatic monitoring instruments of surface water by using above technology were installed along the inflow river, Fu River and the Baiyangdian Lake from 2019 under NB-IoT network environment. Water data from the another 5 surface water monitoring points in Baiyangdian National Agricultural Science Zone in Xiongan New Area were also verified by comparing manual monitoring data in the field. The results show that the change trend of monitoring data is consistent with the change trend of manual sampling data, which indicates that the automatic monitoring network of Fu River and Baiyangdian Lake Water System can effectively capture the change pattern of surface water characteristics, the monitoring system is stable and the monitoring results are reliable, which improves the technical level and efficiency of environmental monitoring and has high feasibility and practical value.

2.1 Architecture and feature of NB-IoT technology

2.1.1 Network System Architecture

The NB-IoT network system architecture consists of a wireless network, a core network, a service platform (e.g., IoT platform), an application server (AS), a basic communication suite, and terminals and user cards (Qian and Wang, 2012). The functions of these components were described as follows.

(1) Core network

The core network consists of packet domain core network, circuit domain core network and

subscriber database (HSS), where the packet domain core network consists of MME, S-CW , P-GW and other functional units (Ge and Zhao, 2014), mainly providing packet domain data transmission and capacity opening functions; the circuit domain core network consists of MSC/VIR, SMS center and other functional units, mainly providing SMS transmission functions; the subscriber database consists of HSS and other functional units, mainly providing user contract data functions.

(2) Business Platform

The service platform includes the user-oriented IoT capability open enablement platform, connection management platform, air write card platform, service gateway and other platforms. In order to open network capability to users, IoT capability open enabling platform provides device access, data storage, data routing and forwarding functions for terminal devices, and provides data push, device management, data query, command issuance and other functions for upper layer applications.

(3) Wireless Network

The NB-IoT radio access network (E-UTRAN) consists of the eNodeB, which allows users to access the network downward through the air interface, and upward through the S 1-MME interface to the signaling plane to the MME and the user plane to the S-GW through the S 1-U interface (Chen and Gan, 2016).

(4) Terminals

The NB-IoT terminal needs to meet the technical requirements of basic function, communication function, service function, RF performance, card interface capability, electromagnetic compatibility, etc., and has the ability to access the NB-IoT network.

(5) Application server

The application server is used to complete the functions of storage, forwarding and management of data related to vertical industries.

2.1.2 NB-IoT technology features

(1) Massive connectivity

NB-IoT adopts a narrowband transmission design with a system bandwidth of 200 kHz. In the frequency domain, the minimum granularity of uplink resources is 3. 75 kHz and the minimum downlink is 180 kHz (Fang and Gang, 2014). NB-IoT supports the connection of up to 50,000 User Equipment (UE) to the core network per cell.

(2) Ultra-low power consumption

Low-power technology is a distinctive feature of NB-IoT network, which is especially important in low-frequency data transmission and battery power supply and meets the actual needs of Internet of Everything. The low power consumption of terminal transmission is achieved through optimization of both hardware and software.

In the module hardware design, the power consumption is reduced by further improving the integration of chips, RF front-end devices and other modules to reduce the insertion loss of the channel; at the same time, the loss of devices and circuits is reduced through the development of high-efficiency amplifiers and high-efficiency antenna devices by various manufacturers; the architecture is optimized mainly in the standby power supply mechanism, which can turn off the power supply circuit of chips and modules that do not need to work in the standby state. In the standby mode, the power supply circuit of chips and modules that do not need to work can be turned off. According to the business needs, consider the use of low-power processors, control the processor main frequency, computing speed and standby mode to reduce terminal power consumption.

The NB-IoT terminal introduces two new power saving features including Power Saving Mode (PSM) and Extended Discontinues Reception (eDRX) (Wang et al., 2016), both of which are negotiated after the user terminal initiates the request and connects to the NB-IoT core network. The choice can be made by selecting either mode for system design, or by combining the power saving features of both for integrated application.

(3) Coverage enhancement

Under the same frequency band, NB-IoT has 20 dB gain compared with existing mobile communication networks, which is equivalent to a 100 times improvement in signal reception

capacity and a great improvement in network coverage (Huang, 2017).

(4) Ultra-low cost

Ultra-low cost is an important feature of NB-IoT, as a mobile IoT technology designed for LPWAN, is expected to have the same or lower module cost than 2G module. The ultra-low cost of NB-IoT terminal devices relies on the protocol design and product implementation with full low-complexity consideration. The NB-IoT main chip architecture currently follows a highly integrated multi-unit approach, with a single antenna, simplified RF front-end, and low requirements for crystal oscillation, which can ensure a low-complexity terminal implementation. Besides including BP (NB-IoT baseband), AP (application processor), RAM (random memory), SD (storage device), PM (power management), RF (radio frequency), additionally PA (power amplifier), GNSS (global navigation satellite system) can be included depending on the demand (Zou et al., 2017). Through a highly integrated design approach, the peripheral devices in NB-IoT terminal development and the terminal design difficulty can be reduced effectively. On the other hand, NB-IoT chips will have strong cost competitiveness due to the simplicity of NB-IoT technology and the low requirements for each functional module in the chip architecture.

2.2 Monitoring scheme design of NB-IoT

2.2.1 IoT system design

The architecture of IoT can be divided into four layers: sensing layer, network layer, application layer, and public technical support (Figure 2.1).

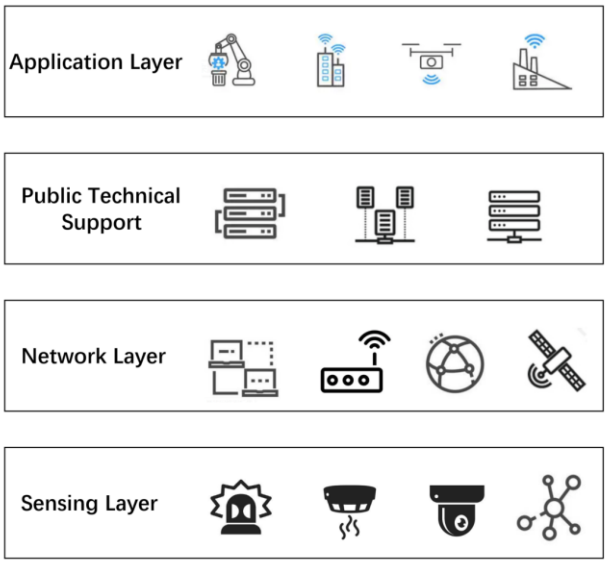

Figure 2.1 IoT Architecture Diagram

(1) Sensing layer

The sensing layer deploys the basic equipment of the IoT, which consists of a variety of devices including environmental quantity sensors, switch state sensors, electronic tags, video sensors and intelligent terminals to sense and measure environmental and equipment information, and transmits the collected sensing information through wired communication or wireless communication, converges to nodes such as gateways, and analyzes the data before finally reporting in a unified manner. It belongs to the bottom layer of the system logic architecture, which is used to deploy basic hardware facilities, including sensors, low-power devices, power supplies, signal processing circuits, etc. There are three key parts inside the sensing terminal: environmental sensor, MCU (Micro Controller Unit), and NB-IoT communication module. The environmental sensors will continuously collect the changing parameters of the environment, calculate, and process them through the embedded program of MCU, and finally report the results through the NB-IoT communication module. Most of the time, the NB-IoT communication module is in a deep body sleep state to save battery power, and only when sensing data is detected and reported, the NB-IoT communication module is awakened and sends data (Xiao and Song, 2013). Compared with other schemes, the NB-IoT-based wireless sensing system does not require an aggregation gateway, which saves development costs and facilitates subsequent sensing module expansion. In addition, as NB-IoT is based on authorized spectrum, communication reliability is greatly guaranteed, and security is enhanced.

(2) Network layer

The network layer can transmit the sensor information reported by the sensor layer upward to the application layer with high reliability and security to reach a broader interconnection. It can realize the most suitable communication method based on different network characteristics according to different application scenarios and business needs. Sensing layer will generate a certain amount of data streams according to demand after being triggered. The data streams are transmitted to the base station through the NB-IoT-based network and then connected to the Internet or other external networks.

(3) Application layer

The application layer parses, stores and processes the sensing information transmitted by the

sensing layer, and can use data mining, cloud computing and other technologies to realize intelligent query and visual display of the system and human-object interaction.

(4) Public technical support

Public technical support contains security technology, identification codes, standards and specifications and other public support technologies.

2.2.2 Hardware structure of the terminal

According to the overall system scheme structure and demand analysis, the following NB-IoT terminal is used in this system. The terminal hardware circuit consists of the following parts: low-power MCU, NB-IoT module, data acquisition circuit, power supply circuit, bus interface, memory module, 485 communication circuit, and clock circuit. Combined with the functional requirements of the online monitoring system, this study selects STM32L073RZT6 as the microcontroller and BC95-G module as the NB-IoT communication module for the monitoring system to achieve the lowest possible power consumption while ensuring stable data transmission. The physical diagram of the NB-IoT hardware circuit and the overall structure of the terminal are shown in Figure 2.2 and 2.3. The bus interface adopts RS-485, and the protocol is Modbus-RTU. Through RS-485 interface, multiple sensing data can be sent through one NB-IoT terminal.

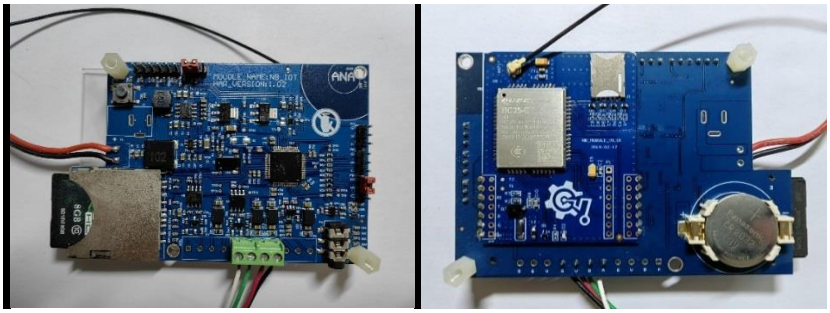

Figure 2.2 NB-IoT terminal hardware.

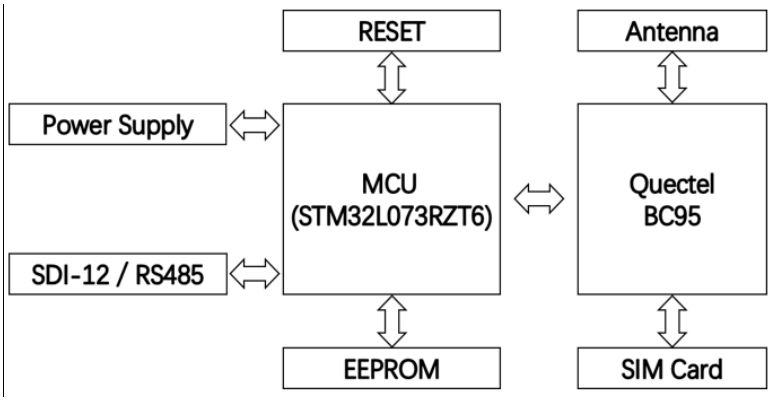

Figure 2.3 NB-IoT terminal structure.

In the IoT system, the selection of wireless communication module is crucial, which determines the communication performance and data transmission accuracy of the system. By analyzing and comparing the performance parameters of various modules, this study adopts the BC95 module (Figure 2.4) as the communication module of NB-IoT terminal. The BC 95 module has the characteristics of high performance and low power consumption, and its size is only 19.9 mm × 23.6 mm × 2.2 mm. Based on the design of reducing the product size and module cost, the BC 95 module can maximize the demand for small-sized communication modules for NB-IoT terminals, which is conducive to system integration design (Zhou, 2015).

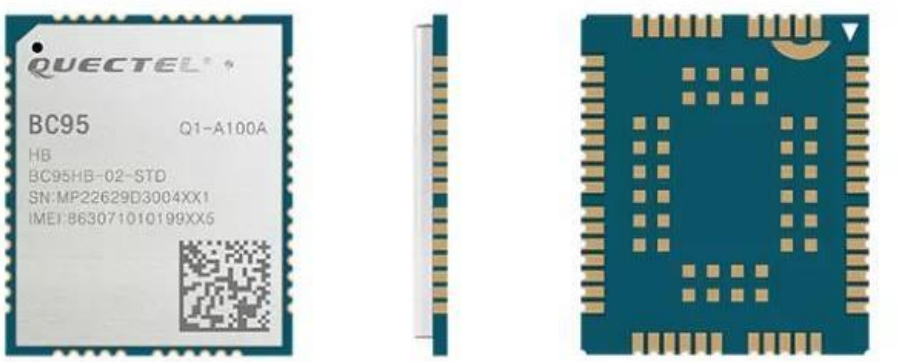

Figure 2.4 Quectel BC95.

The BC 95 module has a compact size, low power consumption, and a large operating temperature range. The BC 95 module hardware package circuit is convenient for soldering and easy to embed into product applications, which can well meet the requirements of IoT applications and facilitate customers to design and upgrade their products quickly and flexibly. The application of SMT chip technology makes the module highly reliable (Wang and Hu, 2011). Through the access of NB-IoT wireless network, the BC95 module can connect to the NB-IoT base station to realize wireless communication. The communication flow of the module in the NB-IoT terminal is shown in Figure 2.5.

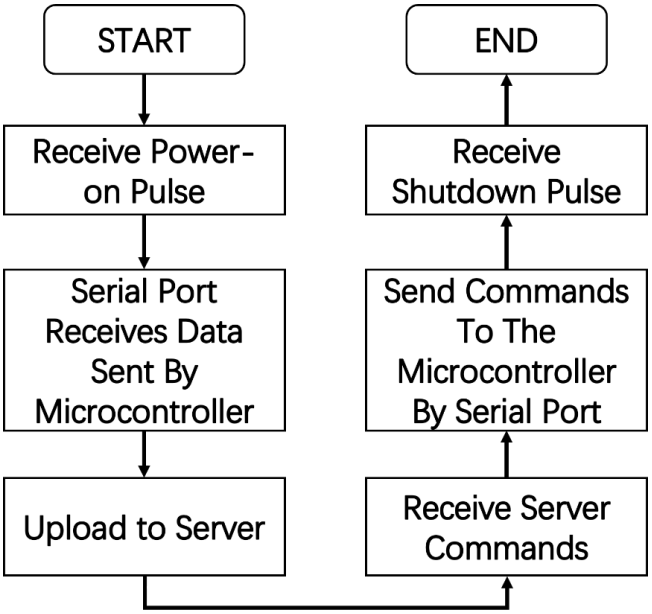

Figure 2.5 BC95 communication flow.

The BC95 module incorporates power saving technology with current consumption as low as 5uA in PSM (power saving mode). The following table shows the main performance of the BC95 module.

Table 2.1 BC95 module main performance.

Parameters	Descriptions
Electricity supply	VBAT voltage range: 3.1 V ~ 3.2 V. A typical power supply voltage: 3.6 V.
Transmission power	23dBm±2 dB
Temperature range	-30℃~ 75℃
USIM card port	Only support 3.0 V external USIM card
Main serial port	For the AT command transmission and data transmission baud rate to 9600 BPS: when used for software upgrade the baud rate to 115200 BPS.
Debugging serial ports	Used for software debugging, and set the baud rate

2.2.3 System software design

(1) Software development environment

The water quality monitoring platform in this study adopts B/S architecture, and the key technologies include front-end page development, back-end REST interface development, database design, container deployment, etc. The front-end is developed based on Element UI, the back-end development language is PHP, the data is stored in MySQL, and the database model uses Navicat Data Modeler to design the table structure. The components involved in the system, such as MySQL and Nginx, are built into images to facilitate the system's deployment, operation, and maintenance. Docker composer is used to startup and shut down all services in the system. Eclipse integrated development environment is employed for platform research and development to improve

development and debugging efficiency. The Eclipse development interface is shown in Figure 2.6.

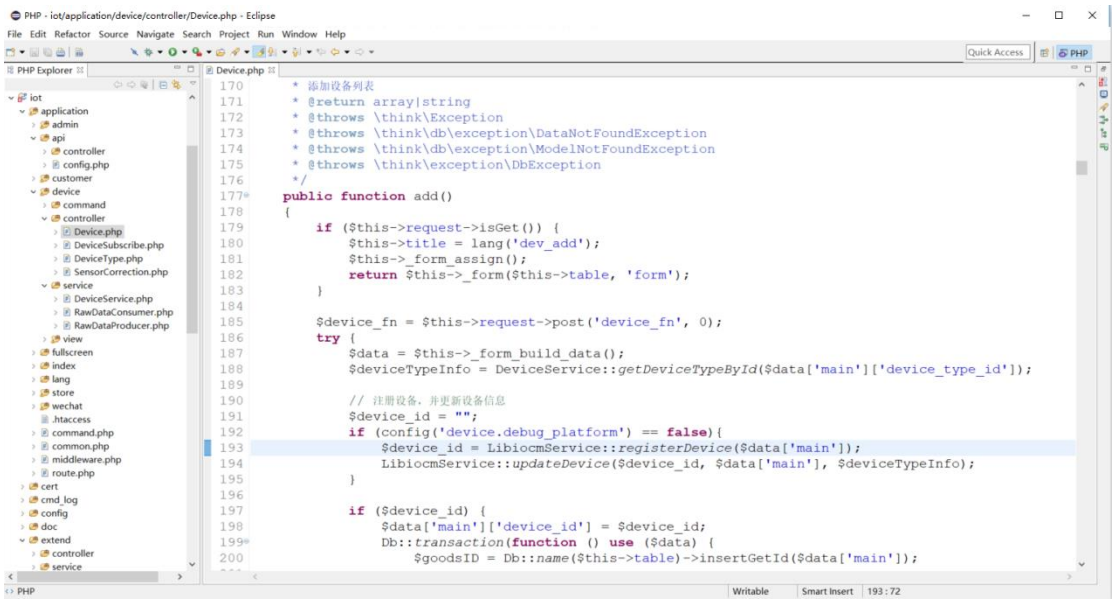

Figure 2.6 Development environment of embedded platform.

(2) Node design flow

The BC95 module uses AT commands to achieve communication. The AT instruction of the main control chip is sent to the BC95 communication module in the form of a string, which is sent to the server by the terminal device, and the server will also respond and give feedback. The workflow diagram of the main function of the system software design is shown in Figure 2.7.

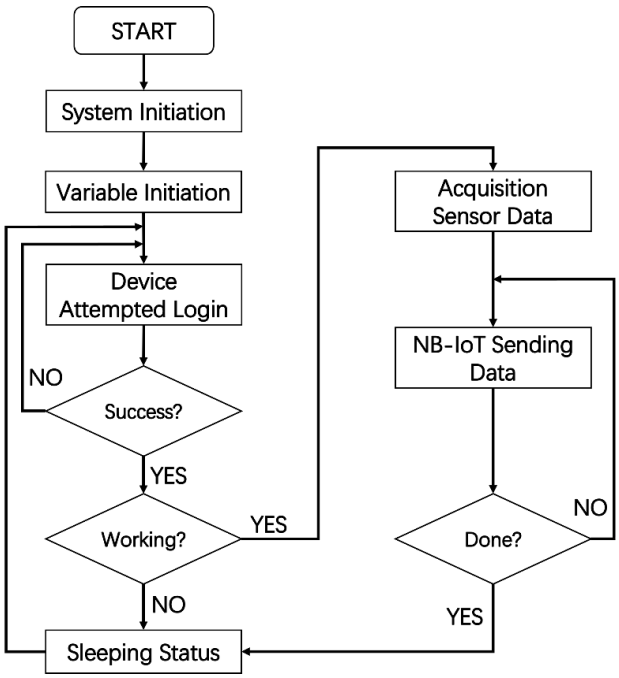

Figure 2.7 System main function workflow.

The system is initialized, which includes system initialization and variable initialization. System initialization refers to the system initialization of each module of the terminal, including clock initialization, I/O port initialization and serial port initialization (Luo, 2010). After initialization, a series of AT commands are sent for terminal testing to confirm normal MCU and NB-IoT module communication, normal serial port communication, and normal SIM card information reading, to confirm whether the NB-IoT module is in full working mode, whether it is successfully registered to the network, and to confirm whether the signal is normal. After the initialization of the program is completed, the microcontroller collects and packages the sensing data, then sends and uploads the data through the NB-IoT communication module, and then transmits it to the server for data reception via the NB-IoT network. After completing one data transmission, the terminal device enters the sleep state and waits to be woken up when the next data is reported.

(3) Network attachment

Before the NB-IoT module can communicate, the terminal device needs to be connected to the NB-IoT base station and registered to the NB-IoT network. The NB-IoT module uses protocols such as COAP/UDP to communicate (Zhang et al., 2013), and the MCU can send a series of AT commands to the NB-IoT module to achieve network attachment and registration. The flow of NB-IoT terminal enrollment is shown in Figure 2.8.

First, system send AT command to test the current network connection to confirm whether it has been successfully attached to the network and the serial port is communicating. The NB-IoT module will select the base station with the most matching frequency in the vicinity for connection after judgment and will keep trying until the connection is successful if it encounters problems such as weak signal during the connection process and the connection fails. The process of testing network connection is done by sending AT command: AT+CGSN=1 to query module IMEI code, AT+CFUN? to query module function status, and AT+CFUN=1 to turn on module function.

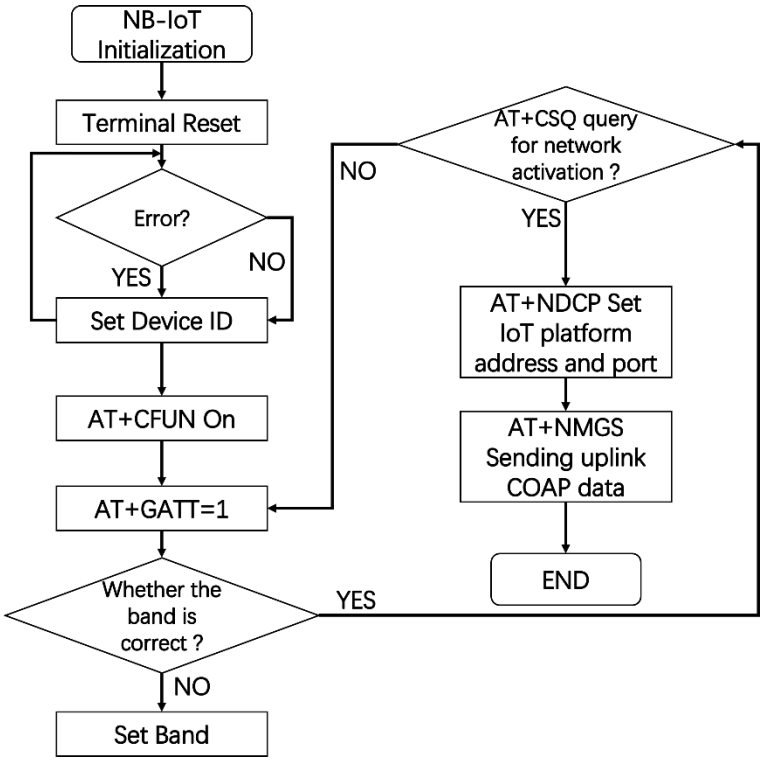

Figure 2.8 NB-IoT Admission Process.

After successful connection with the core network, the network is activated and registered before it can connect with the server for communication. The network registration test process is done by sending AT, command: AT+NBAND? It is to query the current frequency band, sending AT+CSQ>5 to confirm the current signal strength, sending AT+CGATT=1 to activate the network, and sending AT+CEREG? it is to query whether the network is registered. The system will enter sleep mode after completing one registration to reduce power consumption while waiting for the server to reply.

(4) Low-power processing mode

The terminal performs uplink packet upload at certain time intervals, the device state is changed, and the heartbeat packet upload of the terminal is used to maintain the terminal state at the platform side. The low-power processing service model that can be adopted by the system is as follows:

1) PSM is used for the standby state, and the length of PSM active time can be configured to 0, which can save power to the maximum.

2) The terminal implements the RAI function to reduce the power consumption during the operation of the RRC inactivity timer (Chen, 2014), and the monitoring period of the sensing nodes

is set reasonably to reduce the power consumption of the sensors.

3) The command information of the server to the terminal such as time calibration and restart is performed after the data transmission of the terminal data reporting status change is completed to avoid frequent waking up of the terminal from the dormant state.

4) Periodic heartbeat packet uploading of terminals can be replaced by TAU cycles to reduce power consumption of heartbeats (Xie, 2018).

2.3 Retransmission mechanism

2.3.1 Problem definition

The IoT platform connects hardware devices downward and application systems upward, playing a connecting role. The communication between the IoT platform and the device as well as with the application system is asynchronous. Firstly, the device sends data to the IoT platform through the southbound interface, and then the IoT platform sends the data to the application system through the northbound interface, thus realizing the data transmission (Figure 2.9).

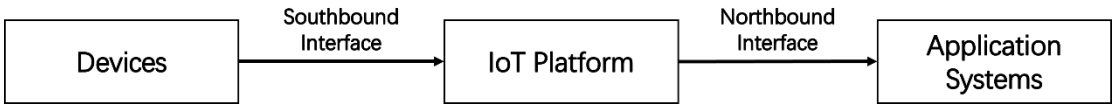

Figure 2.9 Communication flow of the IoT platform.

The above scheme will have serious packet loss when the network is unstable, and it is difficult for the application system to sense the packet loss, and it is also impossible to trigger the device to retransmit the packet loss data through the feedback mechanism. On this basis, this study designs a packet loss retransmission scheme based on the IoT platform through the asynchronous session mechanism, in which the application system automatically sends a command to the IoT platform when packet loss is detected, and then the IoT platform forwards the command to the designated device, which receives the command and retransmits the loss data.

2.3.2 Solutions

(1) Overall solution

A packet loss retransmission scheme is designed for the problem that the IoT platform loses serious packets when the network is unstable and cannot retransmit the loss data (Figure 2.10).

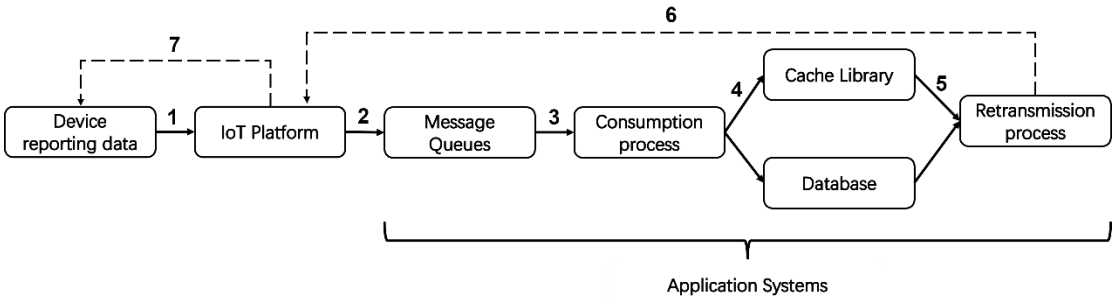

Figure 2.10 Packet loss retransmission scheme based on IoT platform.

① The device transmits data to the IoT platform through the southbound interface based on networks such as 4G/5G/NB-IoT. To achieve packet loss retransmission, the device needs to report the device attribution project ID, device site ID and this communication session ID in addition to the sensor data.

② After the data reaches the IoT platform, the IoT platform sends the data to the application system through the northbound interface. To increase the robustness of the application system, the original data reported by the device is written directly to the message queue, and the data parsing and exception handling are done by the data consumption process.

③ The consumption process takes the original data from the message queue, parses the data, handles exceptions, etc.

④ The consuming process directly writes the normal data to the database, and writes the information of the checked packet loss data to the cache library.

⑤ The retransmission process gets the packet loss information from the database and the cache library, then generates retransmission commands according to the specification of the IoT platform.

⑥ The retransmission process sends the retransmission information to the IoT platform by northbound interface. The IoT platform forwards the retransmission command to the specified device.

(2) Logic of consuming process

The consumption process is responsible for getting and parsing the original data from the message queue where the original data is stored (Figure 2.11). For normal data, the data is parsed and then directly entered into the database, and for packet loss data, the packet loss information is written to the cache library after parsing.

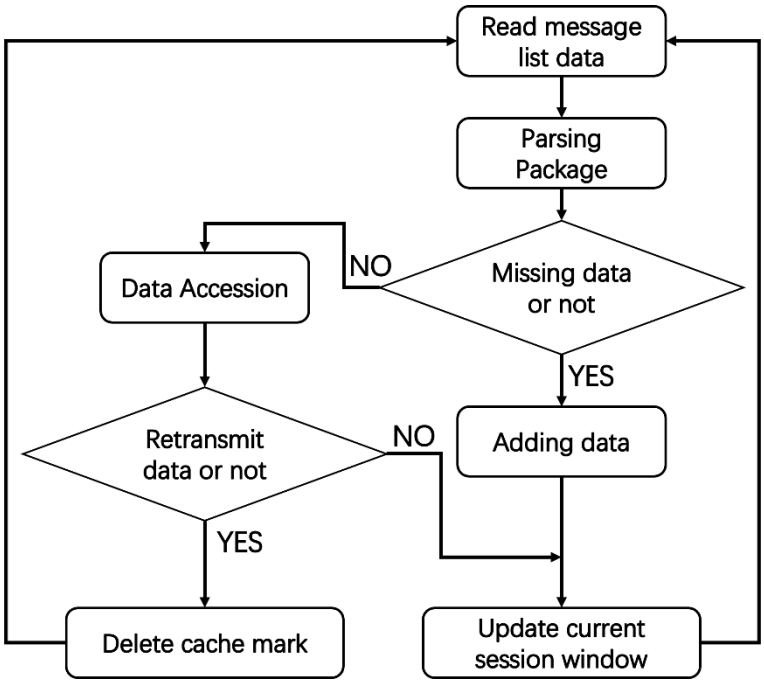

Figure 2.11 Data consumption process processing logic.

The processing steps of the data consumption process are as follows (Figure 2.12).

① Get a piece of data from the message queue.

② Parse out the data header and packet body from the data packet.

③ Determine whether there is data loss, if there is no data loss, the data will be entered into the library, if there is data loss, the item ID, site ID and session ID of the lost data will be written to the retransmission table.

④ Determine if the data is a retransmitted packet, if it is , delete the cache mark and return to

⑤ If it is not a retransmitted packet, update the current session ID of the current device.

(3) Retransmission process processing logic

The retransmission process is responsible for getting the required retransmission packets from the missing information table and generating retransmission commands to send to the IoT platform.

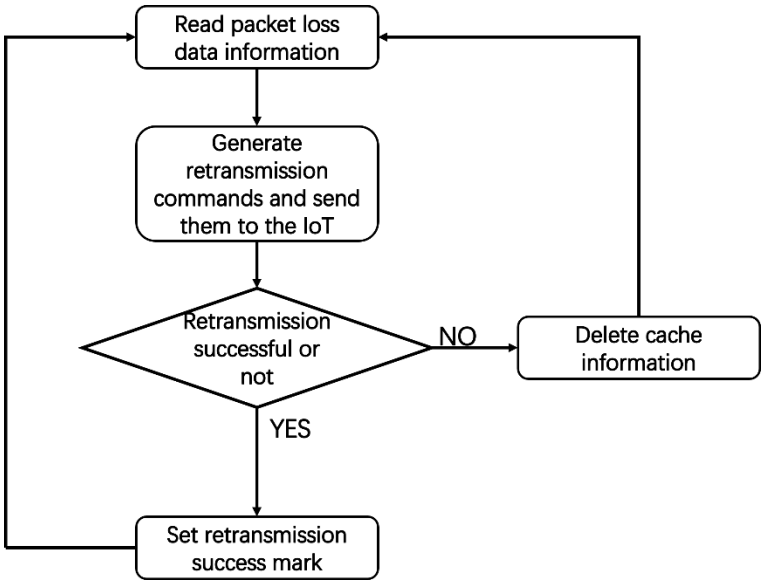

Figure 2.12 Retransmission process processing logic.

The specific processing steps of the retransmission process are as follows.

① Obtain the packet loss data information from the database.

② Generate the retransmission command based on the packet loss data information and send the command to the IoT platform.

③ Judge whether the command is successful or not. If the retransmission command is successful, set the retransmission success mark and delete the information in the cache.

④ If the retransmission command fails, delete the recorded cache information and continue to process the next lost data, and after one round of iterative cycle, issue the retransmission command again for the unsuccessful packets. If the retransmission still fails after N attempts, the retransmission is abandoned.

2.3.3 Contribution of Retransmission mechanism

Because the distance and location between the seven monitoring points and the IoT cellular base station are different, the effect of wireless communication will also be different. The retransmission mechanism can provide up to three retransmissions of packet lost data. If all three retransmissions fail, it will be considered as data loss. Taking Wangting site, located in the upstream of the Fu River, as an example, during the period from July 2019 to November 2021, a total of 878 sets of packet loss data were retransmitted, including 851 successful sets and 27 failed sets, with a success rate of about 97%. In other words, this method recovers 97% of packet loss data. Table 2.2 shows that the average retransmission success rate of seven monitoring points is about 96%.

Table 2.2 Success-rate of retransmission mechanism.

	A	B	C	D	E
1	station_id	success_cnt	failed_cnt	total_cnt	success_rate
2	Wangting	851	27	878	0.97
3	Nanliukou	76	2	78	0.97
4	Xixiangyang	51	4	55	0.93
5	Erqiao	201	11	212	0.95
6	Baiyangdian	181	1	182	0.99
7	Outlet of breeding base	290	11	301	0.96
8	Inlet of breeding base	82	5	87	0.94
9					0.96

2.4 NB-IoT monitoring platform application

2.4.1 Background

With the gradual development and maturity of the Internet of Things (IoT), the need for basic data collection, transmission and management in various fields is becoming more and more urgent. It is urgent to build a platform that integrates common hardware, sensors and data management to efficiently provide a practical data collection solution for users. Based on this, a set of environmental big data monitoring and analysis platform is developed and designed in conjunction with telecom IoT platform to provide data collection and management services for various users.

Figure 2.13 shows the architecture diagram of the environmental big data monitoring and analysis platform. The monitoring and analysis platform relies on the communication connection capability of telecom IoT, interconnects with IoT devices through the southbound interface of telecom IoT, and with the monitoring and analysis platform through the northbound interface of telecom IoT platform.

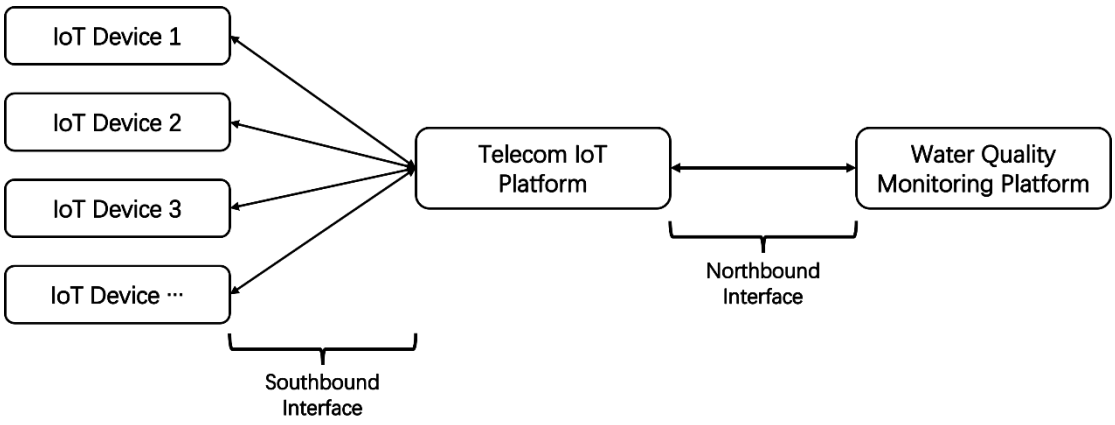

Figure 2.13 Architecture diagram of environmental big data monitoring and analysis platform.

2.4.2 Display elements

Figure 2.14 shows the large screen display page of the platform, which displays the overall information of the data by monitoring system. The top left corner of the large screen shows the distribution of daily data collection volume. The middle left is the list chart of data collected by current monitoring sensors. The middle shows the distribution of monitoring sites in the system by means of a map. The upper right corner shows the estimated pollution level calculated from the data collected by the current monitoring sensor. Middle right shows the data collected by the current monitoring sensors. The six figures below show the trend of the data collected by the current monitoring sensors.

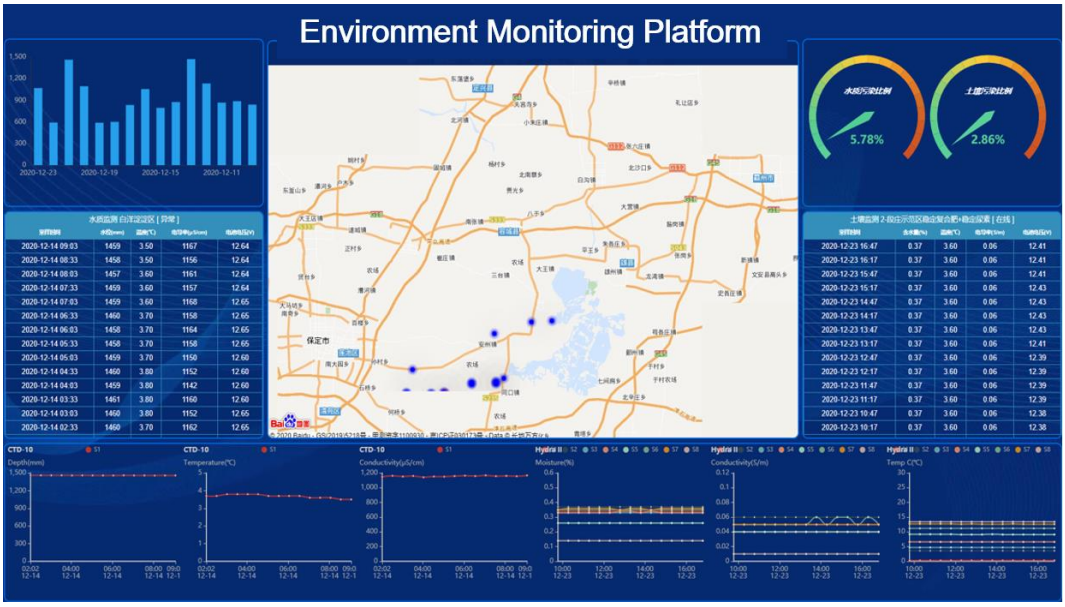

Figure 2.14 Screenshot of environmental big data monitoring and analysis system.

2.4.3 Device Management

2.4.3.1 IoT management list

The device management module contains monitoring devices, remote control devices, IoT online list and command issuance interface.

(1) Monitoring device module

Monitoring device is an entry page, which can be screened by device name, attribution item and device type.

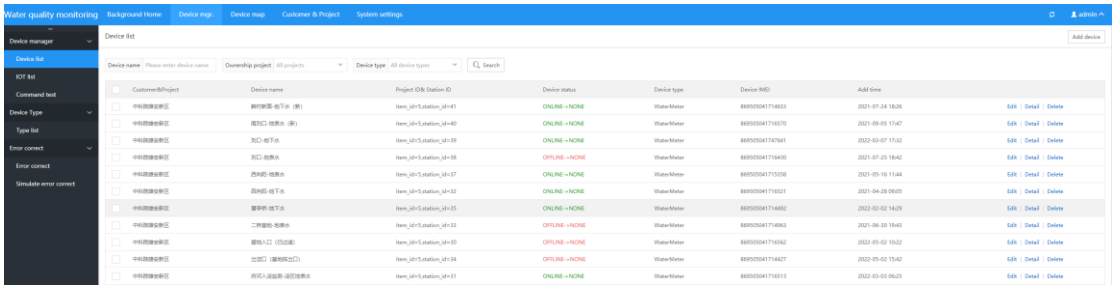

Figure 2.15 IoT management list.

(2) Device property modification

Editable device information include device type, attribution items, device IMEI, and device location information, etc.

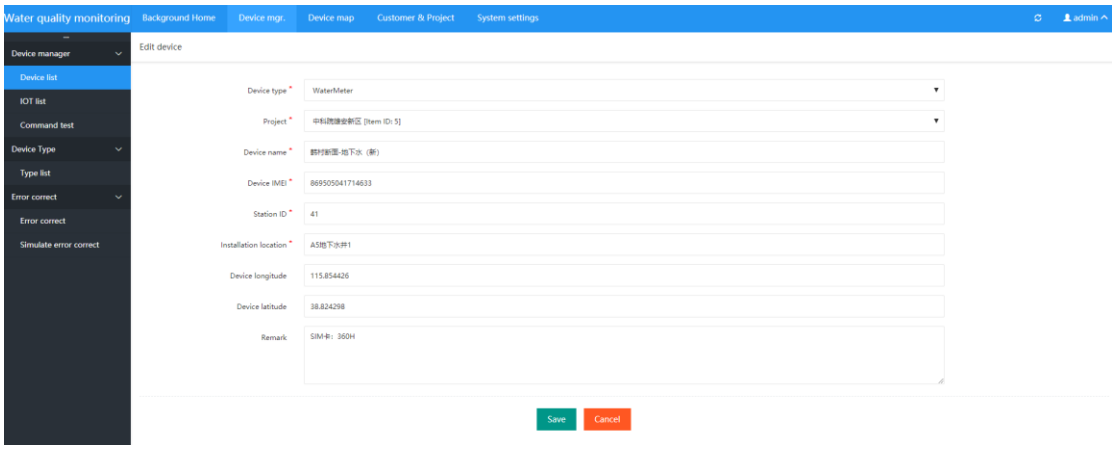

Figure 2.16 Editing page of device information.

(3) Device details

The device details page displays the basic attribute information of the device on the left side, and the data information related to the device on the right side, such as the monitoring data of the device, the original data uploaded to the platform, the data sent for retransmission, and the history commands for retransmission.

The data page displays the time series data collected by each sensor in a curve chart, bar chart, pie chart, etc. The data in the charts can be analyzed in perspective by zooming in the upper right corner, area selection, data viewing, refreshing, and chart downloading buttons. It can select data of different time periods for visual analysis and export data of corresponding time periods to Excel.

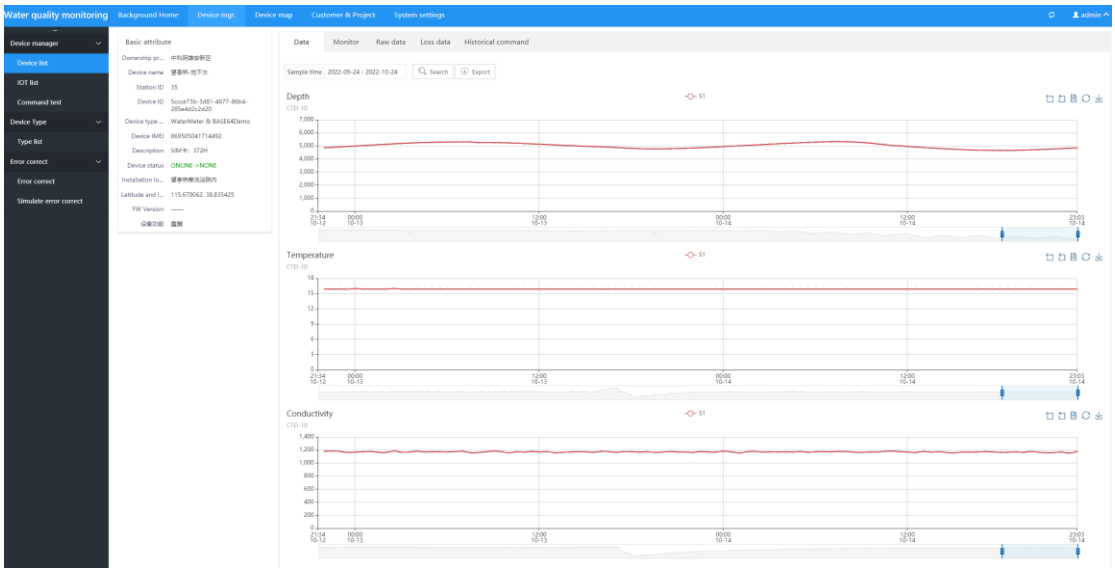

Figure 2.17 Device details page.

The data monitoring page displays the status timing data of setting battery and each sensor, such as main battery voltage, voltage data of coin cell, SD card status, sensor connection information, low voltage alarm, etc.

Figure 2.18 Device status monitoring page.

The original data page shows the original data uploaded by the device through the IoT platform and information related to signal strength. It is used to track the integrity and correctness of the data uploaded to the water quality monitoring platform. It can export the raw data of the specified time period to Excel.

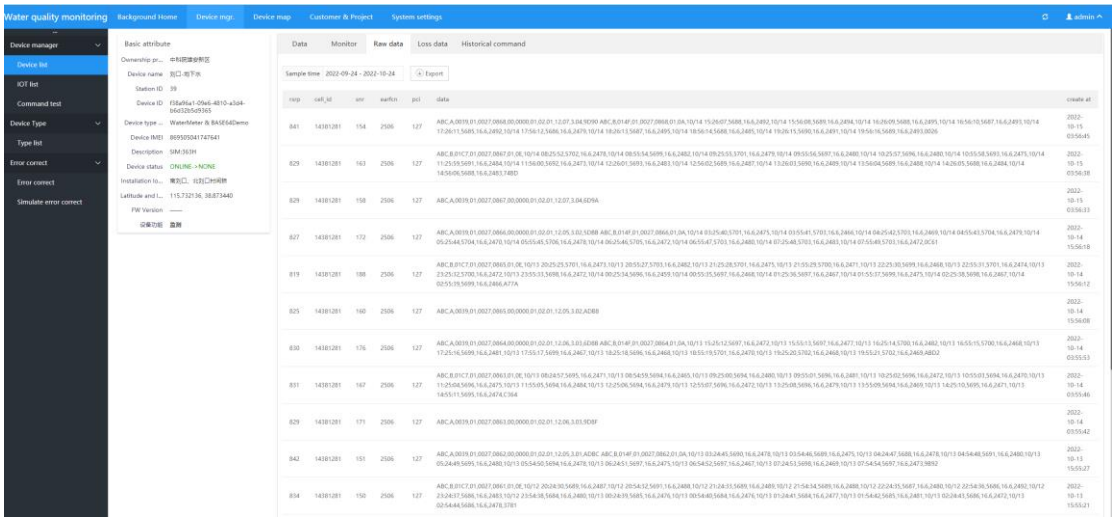

Figure 2.19 Device details of raw data page.

The History Data page displays retransmission history data. It includes the project ID, station ID, session ID, Number of retransmissions, retransmission time, and whether the retransmission was successful. At the same time, it can export the retransmission data of the specified time-period to Excel.

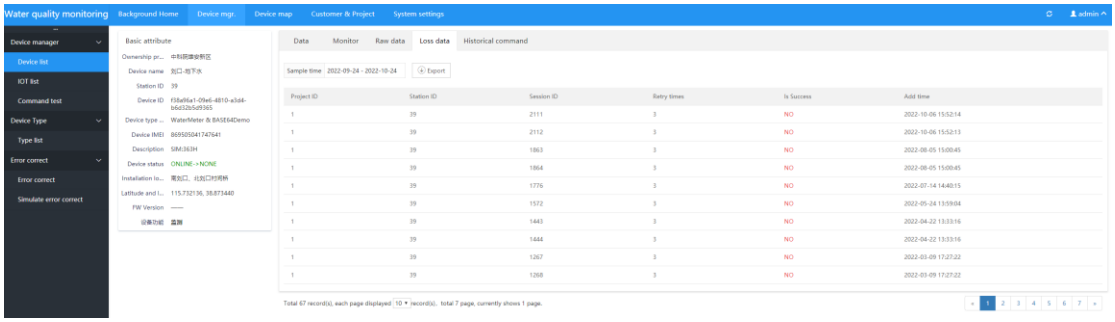

Figure 2.20 Device details of historical data page.

(4) IoT List Module

Get the information of the list of devices that have been registered to the telecom IoT platform, including device name, device status, device ID, device IMEI, the name of the manufacturer to which the device belongs, device type, device model, communication protocol between the device and the IoT platform, and installation location. At the same time, devices can be deleted directly from the Telecom IoT platform via the Delete button.

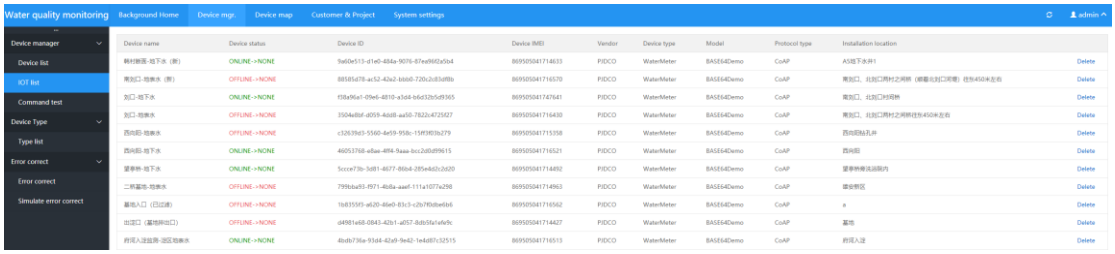

Figure 2.21 IoT platform registered devices list page.

(5) Command test module

The command test module is used to send control commands to the IoT platform to test whether the communication connection of the device is normal. Each time a command is sent, it is necessary to provide the device ID (unique identification ID of the telecom IoT platform device), the device communication protocol service ID, the invoked method and the parameters corresponding to the method.

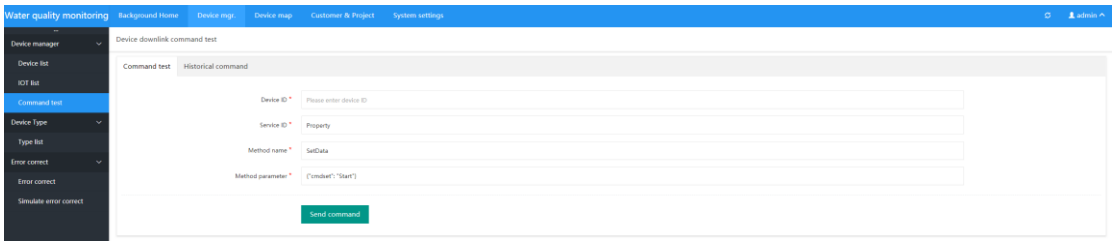

Figure 2.22 Command test page.

2.4.3.2 Device Types

The Device Type module, used to maintain and manage device type information. The device type list shows the type of name, device model, vendor ID, vendor name, protocol type, and time added for each device type. With the Delete button, you can delete the specified equipment type, or specify multiple equipment types to batch delete multiple equipment type information.

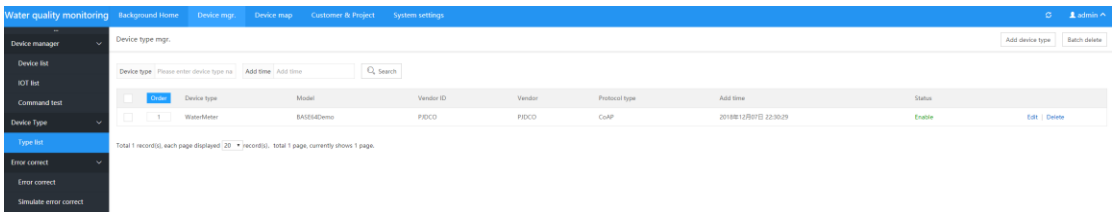

Figure 2.23 Device Type List.

The Device Type Edit page allows you to edit information such as device type name, model number, device LOGO, vendor ID, vendor name, protocol type, and device description.

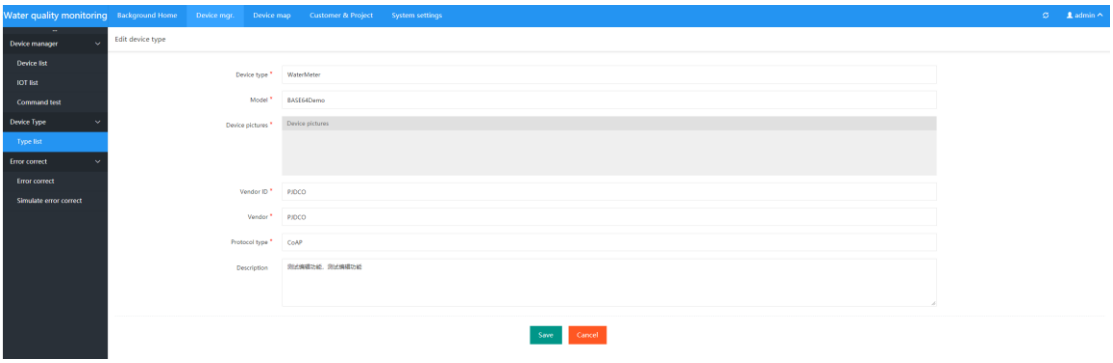

Figure 2.24 Device Type Edit page.

New Device Type page allows you to add a new device type.

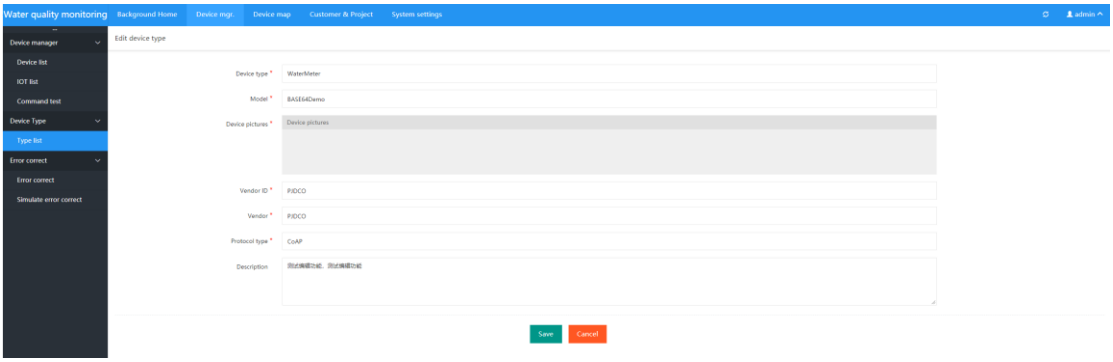

Figure 2.25 Page for adding device types.

2.4.3.3 Error Correction

The Error Correction page is used to make corrections for fixed error of the specified sensor.

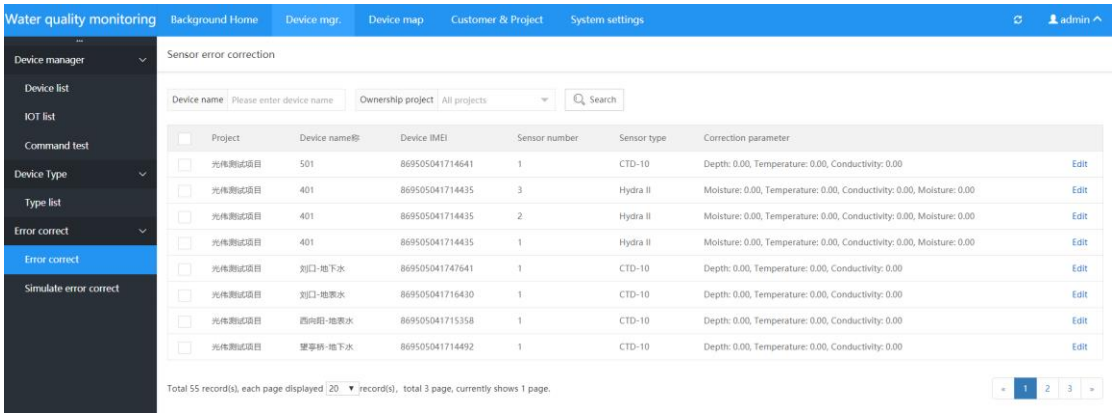

Figure 2.26 Error Correction Page.

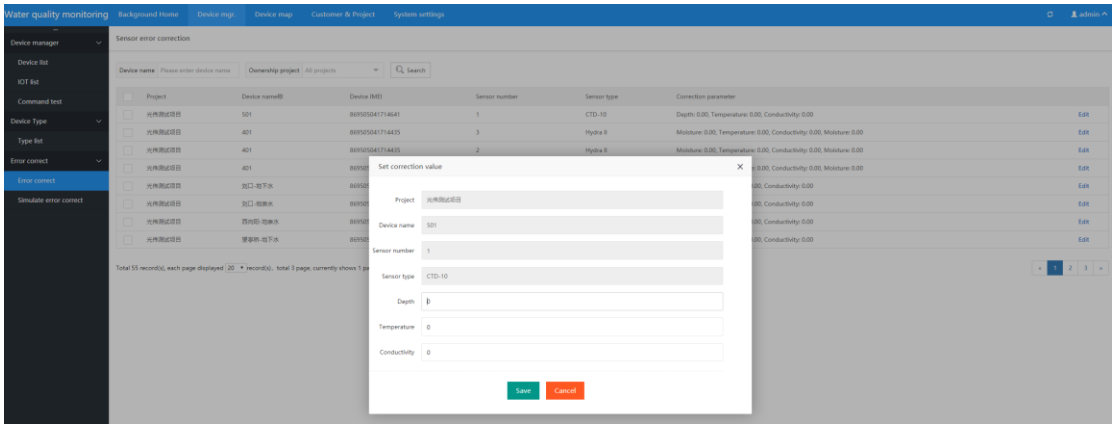

Figure 2.27 Error Correction Rules Edit Page.

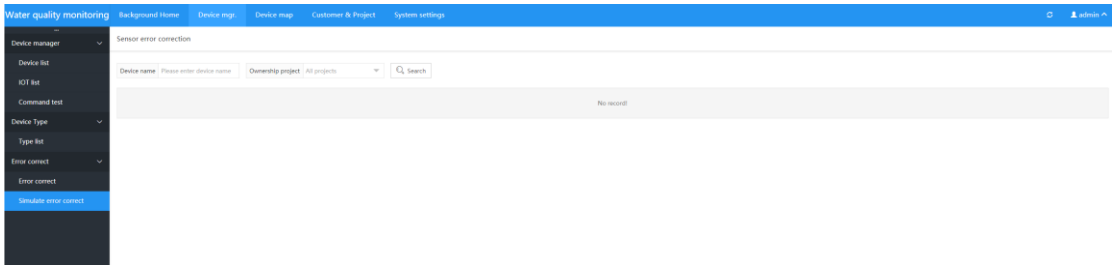

Figure 2.28 Analog sensor error correction page.

2.4.3.4 Device Map

The Device Map module displays the location of devices on the map with graphical information based on the latitude and longitude of the devices. This module provides an overall view of the device distribution in the platform.

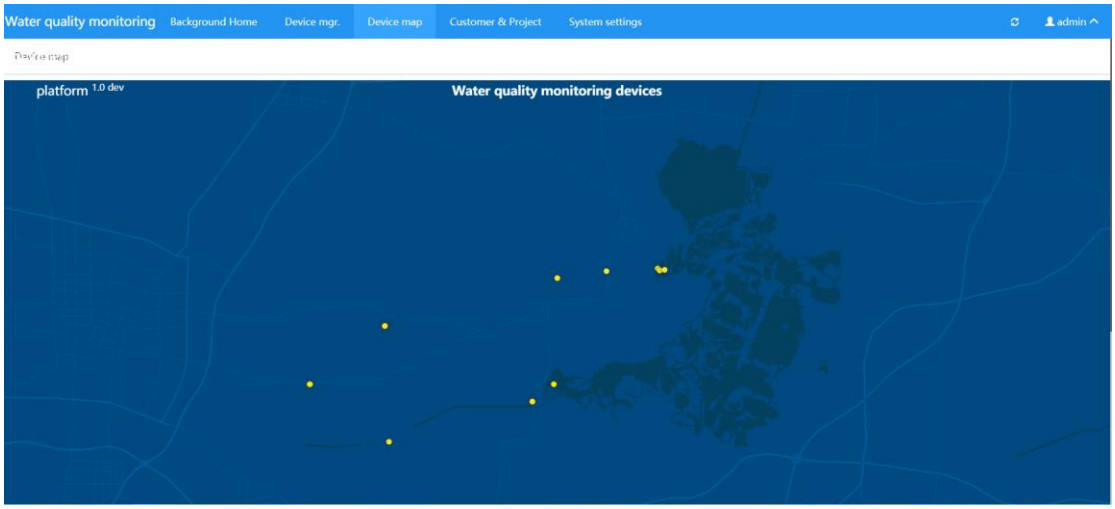

Figure 2.29 Device Map Page.

2.4.3.5 Customer Projects

The Customer Project module is used to manage customer information, project information, and association information between customers and projects. The customer list page allows you to add customers, delete and batch delete customers, edit customer information, and associate customers to corresponding projects.

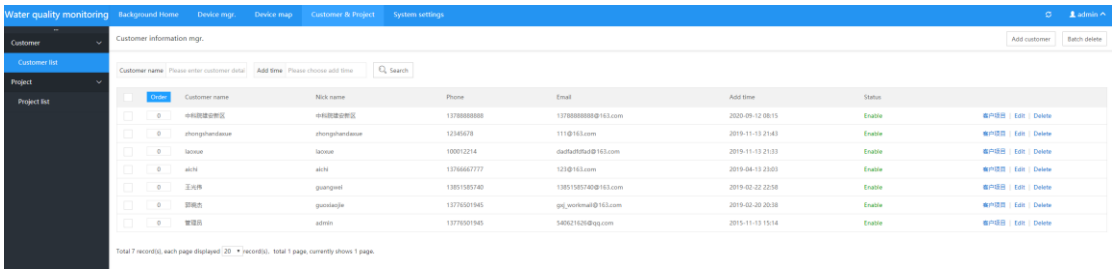

Figure 2.30 Client List.

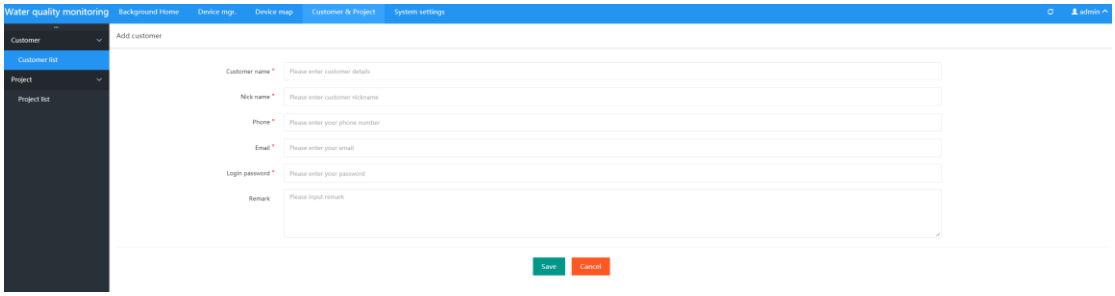

Figure 2.31 Adding Customers page.

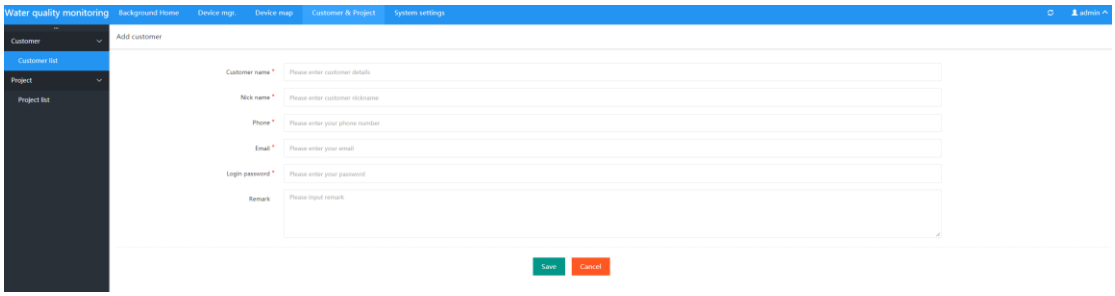

Figure 2.32 Edit Customers page.

The Project List page mainly displays the project information in the system. You can create projects, delete projects, and edit project information. You can also set the binding information between users and projects.

Figure 2.33 Project list page.

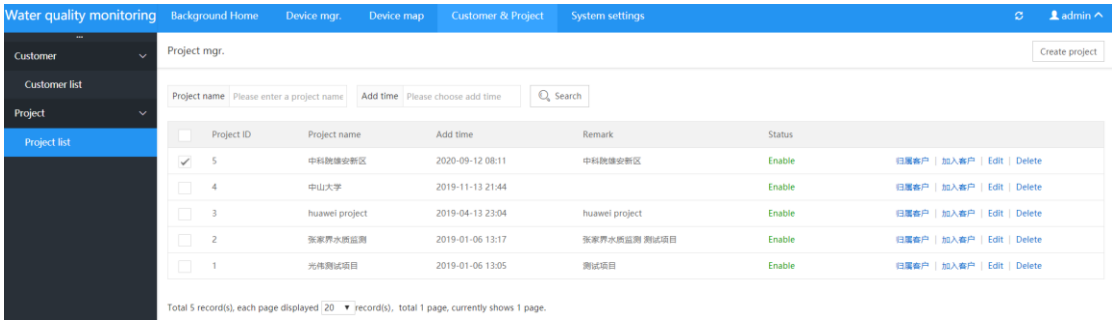

Figure 2.34 Add Customer to Project page.

2.5 Conclusions

A In Situ online monitoring system was developed, focusing on the application of NB-IoT system in environmental monitoring, proposing a data retransmission mechanism to ensure stable transmission of monitoring data in the field under the condition of unstable wireless signal, and according to the statistics of 5 monitoring points, the retransmission mechanism ensures about 96% of the retransmission success rate and greatly improves the reliability of data transmission. At the same time, further realize the low power consumption operation of the system and obtain stable monitoring data. This system has been successfully applied in the monitoring the water characteristics of level, temperature, and EC in Baiyangdian Lake and its upstream inflowing river of Fu River in Xiongan New Area.

Chapter 3 Key Technologies of Autonomous Navigation for Mobile Monitoring

To satisfy the long-time autonomous work of unmanned vehicle, the reliability of autonomous navigation is enhanced by using robust multi-sensors information fusion algorithm, which provides a continuous and dynamic patrolling means for mobile monitoring in large areas. The problem of nonlinear filtering for GNSS/INS under non-Gaussian noise is studied in this chapter, where an improved maximum correntropy cubature Kalman filter (MCCKF) is proposed based on resampling-free sigma-point update framework (SUF). By modifying the sigma points directly and building constrain on the covariance of MCCKF, the resampling-free SUF retains part of the non-Gaussian and high-order information of dynamic model into the posterior sigma points. The proposed algorithm is named as RMCCKF, where the re-scaled measurement and state prediction covariance by maximum correntropy criterion are imported into the construction of posterior sigma points. The proposed algorithm is verified by using numerical simulation and field test on GNSS/INS. Experiment results indicate RMCCKF achieves better performance than MCCKF, where the root mean square error of land vehicle heading is reduced from 1.77 degree to 0.26 degree besides an improved position result.

The pilot phase of this research method was tested on unmanned vehicles and obtained good results. However, this solution is applicable to any unmanned vehicles, including unmanned car, unmanned boats, and unmanned aircraft, etc. to provide them with high-precision navigation data. This method can be used for water quality monitoring unmanned ships to provide low-cost and high-precision navigation solutions and obtain dynamic monitoring data of large areas of water.

3.1 Maximum correntropy and cubature kalman filter

3.1.1 Overview of maximum correntropy criterion

The correspondence of two random variables X, $Y \in \mathcal{R}$ could be specified as

$$V(X,Y) = E[\kappa(X,Y)] = \int \kappa(x,y) dF_{XY}(x,y) \tag{3.1}$$

where $E[\cdot]$ is the anticipation operator, $\kappa(x,y)$ is the transposition-invariant Mercer Kernel and $F_{XY}(x,y)$ is the combined distribution of X, Y. In this study, the applied Gaussian kernel is given as

$$\kappa(x,y) = G_e(x-y) = exp\left(-\frac{e^2}{2\sigma^2}\right) \tag{3.2}$$

where $e = x - y$, and σ denotes the kernel bandwidth. Typically, correspondence can only be estimated

$$\hat{V}(X,Y) = \frac{1}{N}\sum_{i=1}^{N} G_\sigma\big(e(i)\big) \tag{3.3}$$

where $e(i) = x(i) - y(i)$, and $\{x(i), y(i)\}_{i=1}^{N}$ is the samples drawn from F_{XY}. By developing the Gaussian kernel in a Taylor series expansion, we obtain

$$V(X,Y) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n \sigma^{2n} n!} E[(X-Y)^{2n}] \tag{3.4}.$$

This indicates that the corresponding term contains a balanced total of all random variable X-Y moments of even order. With a reasonable choice of kernel bandwidth, the upper order moments of X-Y can be preserved and are insensitive to observed anomalies.

3.1.2 Problem formulation of CKF

The state estimation is to determine x_k at current time k from the noisy $y_{1:k}$, where $y_{1:k}$ represent the measurement up to and including the one at time k. Under Bayesian filtering framework the state estimation is separated into two stages, i.e., time propagation and measurement update, where the former is written as

$$p\big(x_{k|k-1}\big|y_{1:k-1}\big) = \int_{\mathbb{R}^n} p(x_k|x_{k-1}) p(x_{k-1}|y_{1:k-1}) d\, x_{k-1} \tag{3.5}$$

and the measurement update can be formulated as

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{\int_{\mathbb{R}^n} p(y_k|x_k)p(x_k|y_{1:k-1})dx_k} \tag{3.6}$$

where the numerator denotes the constant of normalization. Due to the Markov property of the state model and the conditional separateness of y_k from past observations, the measurement history can be removed from the numerator and denominator. Equations (3.5) and (3.6) cannot be used to estimate the states straight away and have no closed-form answers when the kinetic and measurement functions involved are non-linear and the noise is non-Gaussian.

By means of certainty sampling methods, the CKF can be used to estimate the state prior moments PDF and the probability function with third order accuracy in a Taylor series, and CKF shows better numerical stability than UKF in case the state-dimension is high. Assume that the discrete-time state space model with the aggressive noise under consideration can be described as follows:

$$x_k = f(x_{k-1}) + w_{k-1} \tag{3.7}$$

$$z_k = h(x_k) + v_k \tag{3.8}$$

where $x_k \in \mathbb{R}^n$, $z_k \in \mathbb{R}^p$ are the status and measurement vector at time k, and w_{k-1}, v_k are the process noise and measurement noise respectively, with $w_{k-1} \sim N(0, Q_{k-1})$ and $v_k \sim N(0, R_k)$, where $x \sim N(\mathrm{x}, \mathrm{P})$ represents that random variable x satisfy Gaussian distribution with mean x and variance P. The $f(\cdot)$ and $h(\cdot)$ function represents the dynamic function and measurement function, respectively, and satisfy $f(\cdot): \mathbb{R}^n \to \mathbb{R}^n$, $h(\cdot): \mathbb{R}^n \to \mathbb{R}^p$. Suppose the initial state $x_{0|0}$, w_{k-1} and v_k are mutually independent. Given the posterior PDF at time k-1$p(x_{k-1}) = N\big(\hat{x}_{k-1|k-1}, P_{k-1|k-1}\big)$, Q_{k-1}, R_k and observation z_k, the procedure for the CKF used for the a posterior PDF approval at time k is equated as Algorithm 1. The CKF sigma point ξi is given by

$$\xi_i = \begin{cases} \sqrt{\mathrm{n}} e_i, & i = 1, \cdots, n \\ -\sqrt{\mathrm{n}} e_{i-n}, & i = n+1, \cdots, 2n \end{cases}$$

Where $e_i \in \mathbb{R}^n$ is the i-th column of I_n, and $N = 2n$ delegates the number of samplings. $X_{i,k|k-1}$ and $X_{i,k|k}$ indicates the samplings for moments matching of random variable transformed by dynamic and measurement function.

Algorithm 1: Cubature Kalman filter

Require: $\hat{x}_{k-1|k-1}$, $P_{k-1|k-1}$, Q_{k-1}, R_k

Prediction:

for i = 1, ⋯, N do

$$X_{i,k|k-1}=\text{chol}\left(P_{k-1|k-1}\right)\xi_i + \hat{x}_{k-1|k-1}\text{ , }\omega_i=1/N \tag{9}$$

end for

$$\hat{x}_{k|k-1}=\sum_{i=1}^{N}\omega_i f\left(X_{i,k|k-1}\right) \tag{10}$$

$$P_{k|k-1}=\sum_{i=1}^{N}\omega_i f\left(X_{i,k|k-1}\right)\left(f\left(X_{i,k|k-1}\right)\right)^T - \hat{x}_{k|k-1}\hat{x}_{k|k-1}^T + Q_{k-1} \tag{11}$$

Update:

for i = 1, ⋯, N do

$$X_{i,k|k}=\text{chol}\left(P_{k|k-1}\right)\xi_i + \hat{x}_{k|k-1},\ \omega_i=1/N \tag{12}$$

end for

$$\hat{z}_{k|k-1}=\sum_{i=1}^{N}\omega_i h\left(X_{i,k|k}\right) \tag{13}$$

$$P_{k|k-1}^{zz}=\sum_{i=1}^{N}\omega_i h\left(X_{i,k|k}\right)\left(h\left(X_{i,k|k}\right)\right)^T - \hat{z}_{k|k-1}\,\hat{z}_{k|k-1}^T+R_k \tag{14}$$

$$P_{k|k-1}^{xz}=\sum_{i=1}^{N}\omega_i X_{i,k|k}\left(h\left(X_{i,k|k}\right)\right)^T - \hat{x}_{k|k-1}\,\hat{z}_{k|k-1}^T \tag{15}$$

$$K_k = P_{k|k-1}^{xz}\left(P_{k|k-1}^{zz}\right)^{-1} \tag{16}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1}+K_k\left(z_k - \hat{z}_{k|k-1}\right) \tag{17}$$

$$P_{k|k} = P_{k|k-1} - K_k P_{k|k-1}^{zz}\,K_k^T \tag{18}$$

Remark 1. In the implementation of CKF the samplings are resampled twice according to the approximated Gaussian moments, which takes into account the additive noise whereas drops the extra information of $f\left(X_{i,k|k-1}\right)$ and $h\left(X_{i,k|k}\right)$ in addition to the Gaussian moments in the framework of iterative updating.

Remark 2. The derivation of (16)-(18) is based on the linear least mean square error (LMMSE) criterion, which is insensitive to non-Gaussian noise (e.g. observed outliers and scattered grain noise). Although the third-order volume rule approximates the third-order nonlinear transformation in Taylor series, the CKF inherits the linear Bayesian update of the KF, which is less productive in the presence of less measurement noise, making the new measurement update the preferable choice for nonlinear updates.

3.2 Robust measurement updates based on MCC

The measurement update of the CKF is sensitized to non-Gaussian noise. a robust measurement update framework based on the MCC was designed in this study and the MCCKF was deployed on this basis. the filter model was specified according to (3.7) and (3.8)

$$\begin{bmatrix} \hat{x}_{k|k-1} \\ z_k \end{bmatrix} = \begin{bmatrix} x_k \\ h(x_k) \end{bmatrix} + \delta_k \tag{3.9}$$

where

$$\delta_k = \begin{bmatrix} -\left(x_k - \hat{x}_{k|k-1}\right) \\ v_k \end{bmatrix} \tag{3.10}$$

and then

$$E\{\delta_k \delta_k^T\} = \begin{bmatrix} P_{k|k-1} & 0 \\ 0 & R_k \end{bmatrix} = \begin{bmatrix} B_{p,k|k-1} & 0 \\ 0 & B_{R,k} \end{bmatrix} \begin{bmatrix} B_{p,k|k-1} & 0 \\ 0 & B_{R,k} \end{bmatrix}^T = B_k B_k^T \tag{3.11}$$

A multiplication of both sides of (3.9) by B_k^{-1} gives the following equation for the regression.

$$y_k = d(x_k) + e_k \tag{3.12}$$

where $y_k = B_k^{-1} \begin{bmatrix} \hat{x}_{k|k-1} \\ z_k \end{bmatrix}$, $d(x_k) = B_k^{-1} \begin{bmatrix} x_k \\ h(x_k) \end{bmatrix}$ and $e_k = B_k^{-1} \delta_k$. Then define the value function based on MCC as

$$J_{MCC}(x_k) = \frac{1}{L} \sum_{i=1}^{L} G_\sigma \left(y_{k,i} - d_i(x_k) \right) \tag{3.13}$$

where $L = n + p$, $y_{k,i}$ represents the i-th element of y_k, and $d_i(x_k)$ denotes the i-th row $d(x_k)$. Then, the optimal x_k can be solved by

$$\frac{\partial J_{MCC}(x_k)}{\partial x_k} = 0 \tag{3.14}$$

then we have

$$\left(\frac{\partial d(x_k)}{\partial x_k} \right)^T C_k (d(x_k) - y_k) = 0 \tag{3.15}$$

where $C_k = \begin{bmatrix} C_{k,x} & 0 \\ 0 & C_{k,z} \end{bmatrix}$ is used to re-weight the covariate array of residual error and the reconstruct measurement. Define $e_{k,i} = y_{k,i} - d_i(x_k)$, then $C_{k,x} = diag\left(G_\sigma(e_{k,i}), \cdots, G_\sigma(e_{k,n}) \right)$, and $C_{k,z} = diag\left(G_\sigma(e_{k,n+1}), \cdots, G_\sigma(e_{k,n+p}) \right)$. Due to the nonlinearity contained in h(x_k), (3.15) can be expanded further in the present estimation $\hat{x}_{k|k}^{t-1}$ as

$$\left(\frac{\partial d(\hat{x}_{k|k}^{t-1})}{\partial \hat{x}_{k|k}^{t-1}} \right)^T C_k^{t-1} \left(d(\hat{x}_{k|k}^{t-1}) + \frac{\partial d(\hat{x}_{k|k}^{t-1})}{\partial \hat{x}_{k|k}^{t-1}} (x_k - \hat{x}_{k|k}^{t-1}) - y_k \right) = 0 \tag{3.16}$$

where $\hat{x}_{k|k}^{t-1}$ and C_k^{t-1} are the fixed-point algorithm iterates over the variables corresponding to t-1. Define $D_k^{t-1} = \frac{\partial d(\hat{x}_{k|k}^{t-1})}{\partial \hat{x}_{k|k}^{t-1}}$, we have

$$\hat{x}_{k|k}^{t} = ((D_k^{t-1})^T C_k^{t-1} D_k^{t-1})^{-1} (D_k^{t-1})^T C_k^{t-1} (D_k^{t-1} \hat{x}_{k|k}^{t-1} - d(\hat{x}_{k|k}^{t-1}) + y_k) \tag{3.17}$$

Define $H_k^{t-1} = \frac{\partial h(\hat{x}_{k|k}^{t-1})}{\partial \hat{x}_{k|k}^{t-1}}$, and use the matrix inversion lemma, (3.17) can be further expressed as

$$\hat{x}_{k|k}^{t} = \hat{x}_{k|k-1} + K_k^{t-1} \left(z_k - h(\hat{x}_{k|k}^{t-1}) - H_k^{t-1} (\hat{x}_{k|k-1} - \hat{x}_{k|k}^{t-1}) \right) \tag{3.18}$$

where

$$K_k^{t-1} = P_{k|k}^{t-1}(H_k^{t-1})^T\left(H_k^{t-1}P_{k|k}^{t-1}(H_k^{t-1})^T + R_k^{t-1}\right)^{-1} \tag{3.19}$$

$$P_{k|k}^{t-1} = B_{p,k|k-1}\left(C_{k,x}^{t-1}\right)^{-1}B_{p,k|k-1}^T \tag{3.20}$$

$$R_k^{t-1} = B_{R,k}\left(C_{k,z}^{t-1}\right)^{-1}B_{R,k}^T \tag{3.21}$$

$$C_{k,x}^{t-1} = diag\left(G_\sigma\left(e_{k,1}^{t-1}\right),\cdots,G_\sigma\left(e_{k,n}^{t-1}\right)\right) \tag{3.22}$$

$$C_{k,z}^{t-1} = diag\left(G_\sigma\left(e_{k,n+1}^{t-1}\right),\cdots,G_\sigma\left(e_{k,n+p}^{t-1}\right)\right) \tag{3.23}$$

$$e_{k,i}^{t-1} = y_{k,i} - d_i\left(\hat{x}_{k|k}^{t-1}\right) \tag{3.24}$$

The posterior covariance matrix can be updated as

$$P_{k|k}^t = (I_n - K_k^{t-1}H_k^{t-1})P_{k|k-1}(I_n - K_k^{t-1}H_k^{t-1})^T + K_k^{t-1}R_k(K_k^{t-1})^T \tag{3.25}$$

Based on the derivation of MCC-based KF update, the MCCKF can thus be summarized as Algorithm 2, where K_k^{t-1} and H_k^{t-1} are replaced with the moments approximated by the statistic linear regression. It is noticed from Algorithm 2, the final output using the fixed-point iterative algorithm performs the posterior covariance $P_{k|k}^t$ only once, which reduces the calculation flexibility and does not significantly degrade the efficiency of the MCCKF in our simulations.

Algorithm 2: MCC-based Cubature Kalman filter

Require: $\hat{x}_{k-1|k-1}$, $P_{k-1|k-1}$, Q_{k-1}, R_k , σ , ε

Prediction:

Calculate $\hat{x}_{k|k-1}$, $P_{k|k-1}$using equations (9)-(11)

Update:

Calculate $\hat{z}_{k|k-1}$, $P_{k|k-1}^{xz}$ using equations (12), (13), and use Cholesky decomposition to get $B_{p,k|k-1}$, $B_{R,k}$

Let $t = 1$, $\hat{x}_{k|k}^{0}$=$\hat{x}_{k|k-1}$ and repeat

Calculate $C_{k,x}^{t-1}$, $C_{k,z}^{t-1}$ using equations (32)-(34)

R_k^{t-1}=$B_{R,k}C_{k,z}^{t-1}B_{R,k}^{T}$

$P_{k|k}^{t-1}$=$B_{p,k|k-1}C_{k,x}^{t-1}B_{p,k|k-1}^{T}$

$P_{k|k-1}^{zz}$ =$\sum_{i=1}^{N}\omega_i h(X_{i,k|k})\left(h(X_{i,k|k})\right)^{T}-\hat{z}_{k|k-1}\hat{z}_{k|k-1}^{T}$+$R_k^{t-1}$

H_k^{t-1}=$\left(P_{k|k-1}^{xz}\right)^{T}P_{k|k-1}^{-T}$

$K_k^{t-1}=P_{k|k-1}^{xz}\left(P_{k|k-1}^{zz}\right)^{-1}$

$\hat{x}_{k|k}^{t}=\hat{x}_{k|k-1}$+$K_k^{t-1}\left(z_k-h\left(\hat{x}_{k|k}^{t-1}\right)-H_k^{t-1}\left(\hat{x}_{k|k-1}-\hat{x}_{k|k}^{t-1}\right)\right)$

Update t = t+1 until $\dfrac{\left\|\hat{x}_{k|k}^{t}-\hat{x}_{k|k}^{t-1}\right\|}{\left\|\hat{x}_{k|k}^{t-1}\right\|}\leq\varepsilon$

$P_{k|k}^{t}$=$\left(I_n-K_k^{t-1}H_k^{t-1}\right)P_{k|k-1}\left(I_n-K_k^{t-1}H_k^{t-1}\right)^{T}$+$K_k^{t-1}R_k\left(K_k^{t-1}\right)^{T}$

3.3 Development of the design of the methodology

Since the MCCKF succeeds the SUF of the CKF, i.e., it performs the resampling algorithm twice in one filter cycle, which discards the non-Gaussian and odd-order term information of the instantiation point. In addition, the linear update framework of the CKF is less efficient when the prior has large uncertainty and the measurement noise is small. In this section, the resampling-free SUF is used to enhance the MCCKF, which not only increases the robustness of gain matrix updates in the resampling-free SUF, but also retains more information in the propagation of uncertainty in the estimated iterative nonlinear states.

In order to modify the sigma points directly, the method of left multiplication of sigma points presented in (Cheng et al., 2014), which is a concrete implementation of the resampling-free SUF, is used. Assume that we have the following equation.

$$\mathrm{y}=\mathrm{g}(x)+\epsilon \tag{3.26}$$

This shows that the random variable x is converted by mapping $\mathrm{g}(\cdot)$ with additive noise $\epsilon\sim N(0,\Sigma_\epsilon)$.

Let $\{\mathcal{X}_i, \mathcal{W}_i\}_{i=1}^{N}$ denotes the sigma points and weights corresponding to x. The sigma points of y are then directly modified using left multiplication with the following equation (Straka and Dunik, 2020)

$$\mathcal{Y}_i = \hat{\mathrm{g}} + \mathcal{M}(\mathcal{G}_i - \hat{\mathrm{g}}) \tag{3.27}$$

where $\mathcal{G}_i = g(\mathcal{X}_i)$, $\hat{\mathrm{g}} = \sum_{i=1}^{N} \mathcal{W}_i \mathcal{G}_i$, and the gain matrix $\mathcal{M} = \sqrt{P_{gg} + \Sigma_\epsilon}\left(\sqrt{P_{gg}}\right)^{-1}$, where $P_{gg} = \sum_{i=1}^{N} \mathcal{W}_i (\mathcal{G}_i - \hat{\mathrm{g}})(\mathcal{G}_i - \hat{\mathrm{g}})^T$.

Define $\left\{X_{i,k-1|k-1}^{+}, \omega_i\right\}_{i=1}^{N}$, as the temporal k-1 posterior sigma points and weights, and applying (3.27) to the instantiated dynamic and measurement functions, we have

$$X_{i,k|k} = \bar{x}_F + \sqrt{P_{FF} + Q_{k-1}}\left(\sqrt{P_{FF}}\right)^{-1}\left(f\left(X_{i,k-1|k-1}^{+}\right) - \bar{x}_F\right) \tag{3.28}$$

$$Z_{i,k|k} = \bar{z}_H + \sqrt{P_{HH} + R_k}\left(\sqrt{P_{HH}}\right)^{-1}\left(h\left(X_{i,k|k}\right) - \bar{z}_H\right) \tag{3.29}$$

where

$$\bar{x}_F = \sum_{i=1}^{N} \omega_i f\left(X_{i,k-1|k-1}^{+}\right) \tag{3.30}$$

$$P_{FF} = \sum_{i=1}^{N} \omega_i \left(f\left(X_{i,k-1|k-1}^{+}\right) - \bar{x}_F\right)\left(f\left(X_{i,k-1|k-1}^{+}\right) - \bar{x}_F\right)^T \tag{3.31}$$

$$\bar{z}_H = \sum_{i=1}^{N} \omega_i h\left(X_{i,k|k}\right) \tag{3.32}$$

$$P_{HH} = \sum_{i=1}^{N} \omega_i \left(h\left(X_{i,k|k}\right) - \bar{z}_H\right)\left(h\left(X_{i,k|k}\right) - \bar{z}_H\right)^T \tag{3.33}$$

The sigma points are then propagated through a linear Bayesian update framework

$$X_{i,k|k}^{-} = X_{i,k|k} + K_k\left(z_k - Z_{i,k|k}\right) \tag{3.34}$$

Define the sigma points error matrix at time k as $\tilde{X}_{i,k|k} = X_{i,k|k}^{-} - \bar{x}_{k|k}$, where $\bar{x}_{k|k}$ corresponding to $\left\{X_{i,k|k}^{-}, \omega_i\right\}_{i=1}^{N}$ and have

$$\bar{x}_{k|k} = \sum_{i=1}^{N} \omega_i X_{i,k|k}^{-} \tag{3.35}$$

$$\bar{P}_{k|k} = \sum_{i=1}^{N} \omega_i \left(X_{i,k|k}^{-} - \bar{x}_{k|k}\right)\left(X_{i,k|k}^{-} - \bar{x}_{k|k}\right)^T \tag{3.36}$$

By setting $\tilde{X}_{i,k|k} \sim N(0, K_k R_k (K_k)^T)$ to explain the increase in variance of the direct sigma integral revisions from $X_{i,k|k}^{-}$ to $X_{i,k|k}^{+}$, we get

$$X_{i,k|k}^{+} = \bar{x}_{k|k} + \sqrt{\bar{P}_{k|k} + K_k R_k (K_k)^T}\left(\sqrt{\bar{P}_{k|k}}\right)^{-1} \tilde{X}_{i,k|k} \tag{3.37}$$

where $X_{i,k|k}^{+}$ is then used as the new sigma points for MCCKF at next filter period. The justification of setting the variance increase as $K_k R_k (K_k)^T$ has been given in (Straka and Dunik, 2020). An enhanced MCCKF, referred to as RMCCKF in the context, is developed by importing MCC-based

measurement updates into the resampling-free SUF. algorithm 3 summarizes the RMCCKF.

Algorithm 3: Resampling-free SUF-based MCCKF

Require: $\hat{x}_{k-1|k-1}$, $P_{k-1|k-1}$, Q_{k-1}, R_k , σ , ε

Initialize prediction:

Calculate $\hat{x}_{k|k-1}$, $P_{k|k-1}$, $f\left(X^{+}_{i,k-1|k-1}\right)$ use equations (9)-(11)

Update $X_{i,k|k}$ use (38), (40) and (41)

Update:

Update $Z_{i,k|k}$ use (39), (42) and (43)

Calculate $\hat{z}_{k|k-1}$, $P^{xz}_{k|k-1}$ as

$\hat{z}_{k|k-1}=\sum_{i=1}^{N}\omega_i Z_{i,k|k}$

$P^{xz}_{k|k-1}=\sum_{i=1}^{N}\omega_i\left(X_{i,k|k}-\hat{x}_{k|k-1}\right)\left(Z_{i,k|k}-\hat{z}_{k|k-1}\right)^{T}$

Initialize t = 1, $\hat{x}^{0}_{k|k}=\hat{x}_{k|k-1}$ and repeat update $\hat{x}^{t}_{k|k}$, $P^{t}_{k|k}$ as in Algorithm 2

Calculate $X^{+}_{i,k|k}$ use (44)-(47) based on $h\left(X_{i,k|k}\right)$ and R^{t-1}_{k}

Prediction:

Update $X_{i,k+1|k+1}$ use (38), (40) and (41) based on $f\left(X^{+}_{i,k|k}\right)$

Calculate $\hat{x}_{k+1|k}$ and $P_{k+1|k}$ as

$\hat{x}_{k+1|k}=\sum_{i=1}^{N}\omega_i X_{i,k+1|k+1}$

$P_{k+1|k}=\sum_{i=1}^{N}\omega_i\left(X_{i,k+1|k+1}-\hat{x}_{k+1|k}\right)\left(X_{i,k+1|k+1}-\hat{x}_{k+1|k}\right)^{T}+Q_k$

Return to **Update** for next filtering period

Remark 4. In the no-resampling SUF implementation, the gain matrix relies on the state of the KTF and the covariance of the noise. The linear transformation of the distance matrix between the instantiated and critical points preserves the non-Gaussian information contained in the distance matrix. Note that the state dimension has no direct effect on $X_{i,k|k}$ and $X^{+}_{i,k|k}$, thus alleviating the non-local sampling problem to some degree. In the presence of non-Gaussian noise in models (3.7) and (3.8), they can be extended to $X_{i,k|k}$by (3.34).

Remark 5. The MCC-based measurement update rescales the imprecision spread by using the kernel functions as shown in (3.20) and (3.21), which allows the gain matrix used to resample the free SUF to be more robust, thus suppressing filter divergence in the presence of non-Gaussian noise. Put differently, by combining the MCC-based measurement update with the resampling-free SUF, not only does the sigma point error matrix include non-Gaussian information, but the gain matrix is not sensitive to non-Gaussian noise.

3.4 Performance evaluation and analysis

In this section, the proposed RMCCKF's performance is demonstrated by nonlinear numerical examples and on-board GNSS/INS field tests.

3.4.1 Numerical simulation

A univariate nonsmooth growth model (UNGM) is used to justify the implementation of RMCCKF performance. The UNGM state model can be written as

$$x_k = 0.5x_{k-1} + 25\frac{x_{k-1}}{1+x_{k-1}^2} + 8cos\big(1.2(k-1)\big) + w_{k-1} \tag{3.38}$$

$$y_k = \frac{x_k^2}{20} + v_k \quad k = 1, \cdots, M \tag{3.39}$$

where M is the total simulation step, w_{k-1}, v_k are the process noise and measurement noise. Suppose the data series is generated with initial state $x_0 \sim N(0,1)$, and set M = 500. During the simulation, ε = 1e-4 is chosen for the MCCKF and RMCCKF, and the kernel bandwidth of the MCCKF and the kernel bandwidth of the RMCCKF are chosen by trial and error. First, we take into account the case where both procedure noise and measurement noise are Gaussian dominated, i.e., $w_{k-1} \sim N(0, Q_{k-1})$ and $v_k \sim N(0, R_k)$, where $Q_{k-1} = 1$, $R_k = 1$. The kernel bandwidths of MCCKF and RMCCKF are $\sigma = 0.4$ and $\sigma = 0.8$, this is the same as the case of non-Gaussian noise for a fair comparability. The root mean square error (RMSE) and the average RMSE (ARMSE) are used as measures of performance with the following equations

$$RMSE(m) = \sqrt{\frac{1}{M}\sum_{k=1}^{M}\big(x_k - \hat{x}_{k|k}\big)^2} \tag{3.40}$$

$$ARMSE = \frac{1}{MC}\sum_{m=1}^{MC} RMSE(m) \tag{3.41}$$

where $m = 1, \cdots, MC$ denotes the Monte Carlo run with MC = 50 in the simulation. The estimation errors under Gaussian noise is shown in Figure1. Notice that, when the noise subject to Gaussian distribution, MCCKF does not perform better than CKF which results from that the LMMSE is an optimal criterion under Gaussian noise. However, because of the nonlinearity of UNGM state model, all of the algorithms output approximation state to the true state. RCKF does not improve the performance compared with CKF, which indicates that there is no additional information in instantiated points based on the small number cubature points. The resampling-free SUF improve the uncertainty propagation efficiency of MCCKF, and RMCCKF shows better performance than CKF and MCCKF.

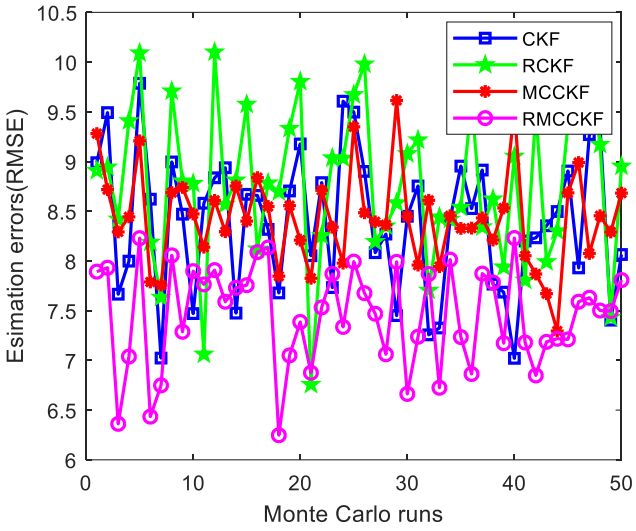

Figure 3.1 Estimation errors （Gaussian noise）.

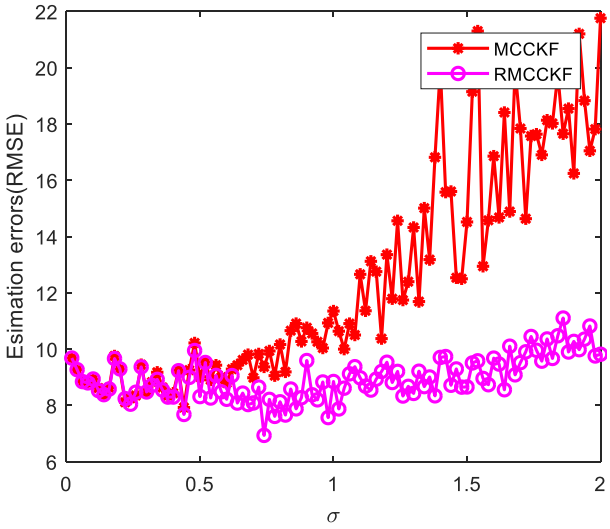

Figure 3.2 Kernel bandwidth selection （non-Gaussian noise）.

To verify the RMCCKF under non-Gaussian noise, the process noise and measurement noise are set to a mixture of Gaussian distributions to simulate the non-Gaussian distribution, that is

$$w_{k-1} = 0.9N(0, Q_{k-1}) + 0.1N(0,100 \times Q_{k-1}) \tag{3.42}$$

$$v_{k-1} = 0.8N(0, R_k) + 0.2(0,100 \times R_k) \tag{3.43}$$

where $Q_{k-1} = 1, R_k = 1$. It is noted from Figure4.2 that the resampling-free SUF is less sensitive to σ, $\sigma = 0.4$ and $\sigma = 0.8$ are selected for MCCKF and RMCCKF in the simulation. The estimation errors of different filters are shown in Figure3.3. The MCC-based measurement update shows a clear advantage and both the MCCKF and RMCCKF obtain better results than their LMMSE-based peers. In addition, the RMCCKF outperforms the MCCKF, suggesting that the no-

resampling SUF provides an improvement over the MCC-based measurement update when both process noise and measurements obey a non-Gaussian distribution. The ARMSEs presented in Table 3.1 further validate this conclusion. In the presence of non-Gaussian noise, as we have seen, all filters exhibit performance deterioration, while the MCC-based measurement update shows much less performance deterioration compared to the LMMSE-based filter. By introducing a non-resampling SUF into the MCCKF, the ARMSE of the RMCCKF improves by about 11.5% and 4% over the MCCKF in the presence of Gaussian and non-Gaussian noise, respectively.

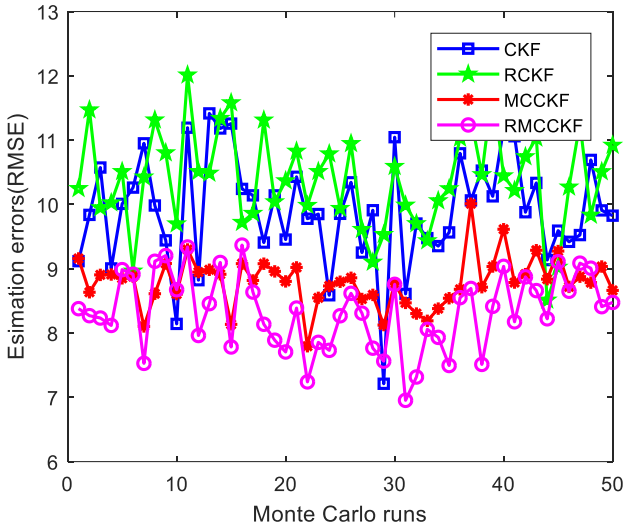

Figure 3.3 Estimation errors （non-Gaussian noise）.

Table 3.1 Armse of different filters.

Algorithm	CKF	RCKF	MCCKF	RMCCKF
Gaussian noise	8.38	8.78	8.43	7.45
Non-Gaussian noise	9.91	10.39	8.79	8.34

3.4.2 GNSS/INS field test

In this section, a vehicle mounted GNSS/INS field test is conducted to evaluate the performance of RMCCKF, in which the filter is operated in a loosely coupled GNSS/inertial guidance system architecture. The earth-centered earth-fixed frame (*e*-frame) is chosen as the navigation frame, and the IMU is fixed in the body frame (*b*-frame) that measures the angular motion and linear motion of vehicle with to inertial frame (*i*-frame) resolved in *b*-frame. The navigation error $\mathbf{x} = [\delta\boldsymbol{\psi} \quad \delta\mathbf{v} \quad \delta\mathbf{p} \quad \boldsymbol{b}_a \quad \boldsymbol{b}_g]$ is used as a filter state to represent the attitude error,

velocity and position, and deviation of the gyroscope and accelerometer, respectively. The inertial guidance system indirect error model resolved in the e frame is used as the dynamic model, which is defined as follows

$$\delta\dot{\boldsymbol{\psi}} = -\boldsymbol{\omega}_{ie}^{e}\otimes\delta\boldsymbol{\psi} + \mathrm{C}_{b}^{e}\boldsymbol{b}_{g} \tag{3.44}$$

$$\delta\dot{\mathbf{v}} = -\left(\mathrm{C}_{b}^{e}\mathbf{f}_{ib}^{b}\right)\otimes\delta\boldsymbol{\psi} - 2\boldsymbol{\omega}_{ie}^{e}\otimes\delta\mathbf{v} + \frac{2g_0}{r_{eS}^{e}}\frac{\mathbf{p}}{|\mathbf{p}|^2}\mathbf{p}^{T}\delta\mathbf{p} + \mathrm{C}_{b}^{e}\boldsymbol{b}_{a} \tag{3.45}$$

$$\delta\dot{\mathbf{p}} = \delta\mathbf{v} \tag{3.46}$$

where $\mathbf{a}\otimes$ represents the skew symmetric matrix of vector $\mathbf{a}$, and $|\mathbf{a}|$ is the magnitude value of $\mathbf{a}$; $\boldsymbol{\omega}_{ie}^{e}$ describes the rotation velocity of e-frame with respect to i-frame resolved in e-frame, $\mathbf{f}_{ib}^{b}$ is the specific force measured by accelerometers, C_{b}^{e} represents the rotation transform from b-frame to e-frame, and g_0 is the acceleration due to the local gravity, r_{eS}^{e} is the geocentric radius at the earth surface. The measurement model is formulated based on the position and velocity output of GNSS and INS, and the lever arm effect between the two systems is considered.

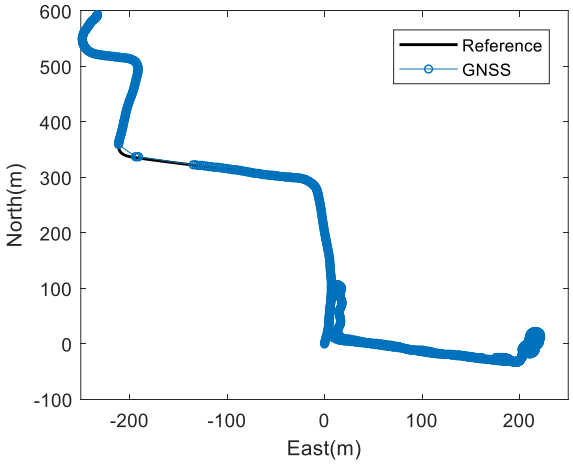

Figure 3.4 Car-mounted test trajectory.

In the field test experiment, a NovAtel SPAN system is applied to record vehicle sensor output and the proposed algorithm is evaluated in a post-processing way. The error characteristic of experiment setup are as follows: bias and scale factor of gyroscope are $1°/h$ and 100ppm, respectively, and its angular random walk is $0.071°/\sqrt{h}$; the bias and scale factor of accelerometer are 0.3mg and 300ppm, respectively. The output frequency of GNSS and IMU are 5Hz and 200Hz, and the data sets are logged with the output of GNSS epoch. The raw data are post-processed by the Waypoint Inertial Explorer software and our proposed algorithm respectively, where the former result is adopted as the reference trajectory in our simulation. The experiment trajectory is shown

in Figure 3.4, and as we can see the GNSS suffers from signal outages from time to time due to the tree blocking and multipath disturbances from environment. The raw IMU data is shown in Figure3.5, and it is notable that there is obvious observation outlier and time-varying noise.

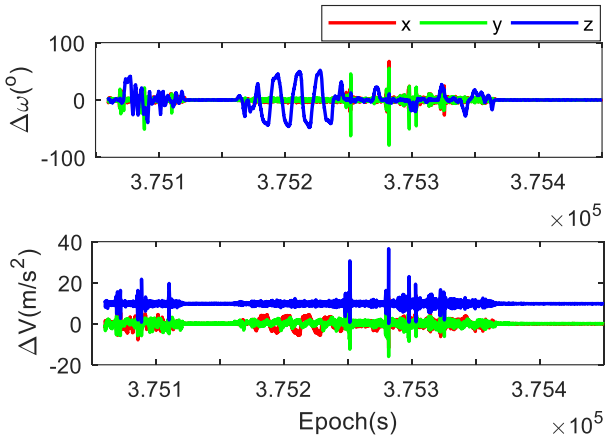

Figure 3.5 Raw IMU output data.

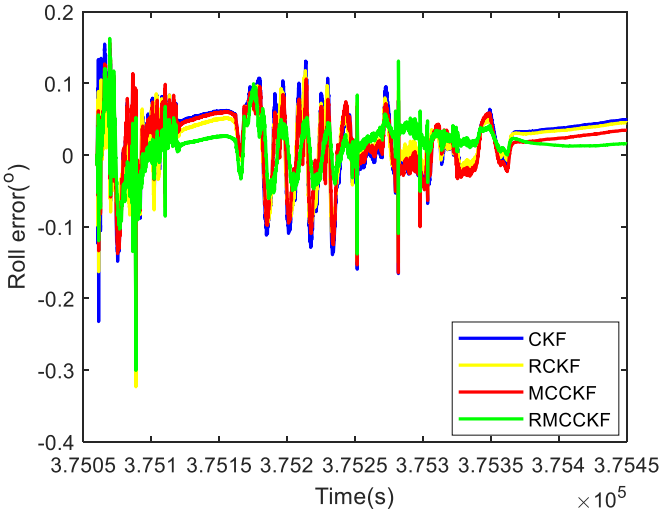

Figure 3.6 Roll error of different filters.

In the simulation, the kernel bandwidth σ and ε for MCCKF and RMCCKF are selected as 300 and 1e-3, respectively, and all the filters share the same parameter configuration to make a fair comparison. The attitude result of different filters is shown in Figure3.6-8, as we can see MCCKF achieves better result than CKF, which indicates that the re-scaled covariance of MCC-based update improves the performance of linear Bayesian update framework. Notice that RMCCKF achieves the best result in terms of attitude, which demonstrates that by combining the resampling-free SUF with MCC-based update, both the two have been improved, which coincides with the comments on *Remark 4*. The correction information of attitude come from the off-diagonal element of prediction

covariance, therefore its update performs in the prediction stage. In the simulation, the initial uncertainty of attitude is set as 1 degree. It is notable in Figure3.6 and Figure3.8, unlike MCCKF which may suffer from large initial uncertainty, RMCCKF improves the predication stage of CKF and it converges faster than CKF and MCCKF.

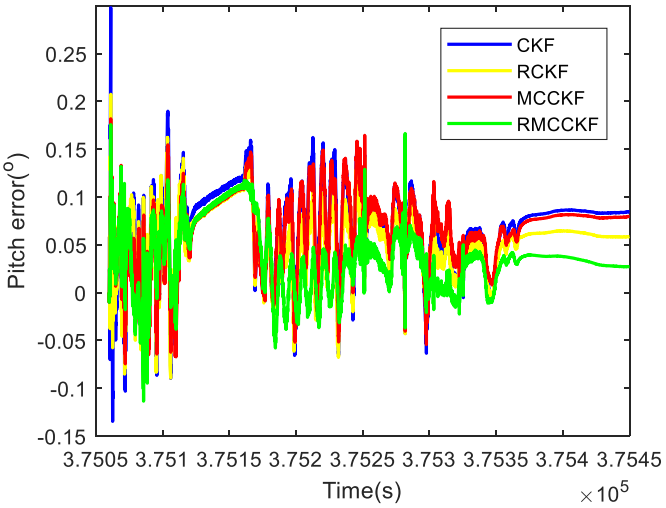

Figure 3.7 Pitch error of different filters.

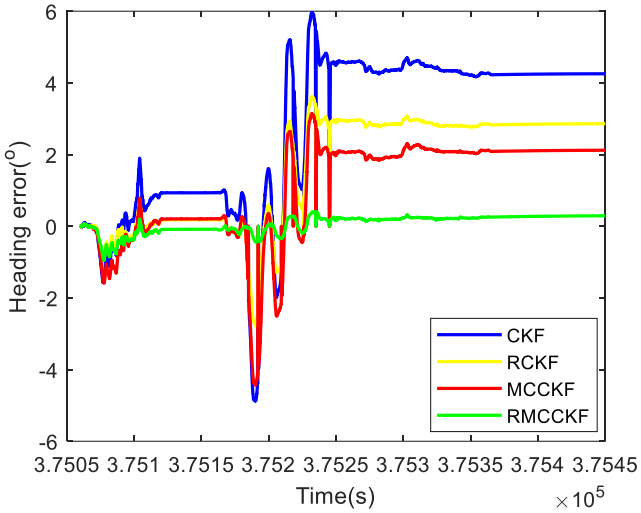

Figure 3.8 Heading error of different filters.

The position errors are shown in Figures 3.9 and 3.10, and the RCKF outperforms the RMCCKF in the presence of long signal interruptions. by directly revising the sigma points, the RCKF constructs two loops for forward filtering, the first for state and uncertainty propagation based on Gaussian moments, and the other for uncertainty propagation by using instantiated points . As can be seen in Figure 3.10, the RCKF does not effectively absorb the observations when the vehicle is running with frequent angular motions, resulting in a significant decrease in position.

More detailed results for the navigation parameters are presented in Table 3.2, where RMSE is used as a performance metric and for brevity the North and East positions are referred to as NP and EP, respectively.We can see that the MCCKF obtains improved attitude results compared to the CKF, where the heading error is reduced from 4.48° to 1.77°. By introducing the no resampling SUF, the RMCCKF improves the heading by 84.3% compared to the MCCKF, and the steady-state errors in pitch and heading are significantly reduced. Note that the position results are similar for the MCCKF and CKF, with the RMCCKF improving position in both position directions. And the no resampling SUF limits the covariance of the CKF, thus limiting the covariance expansion of the sigma point inconsistency.

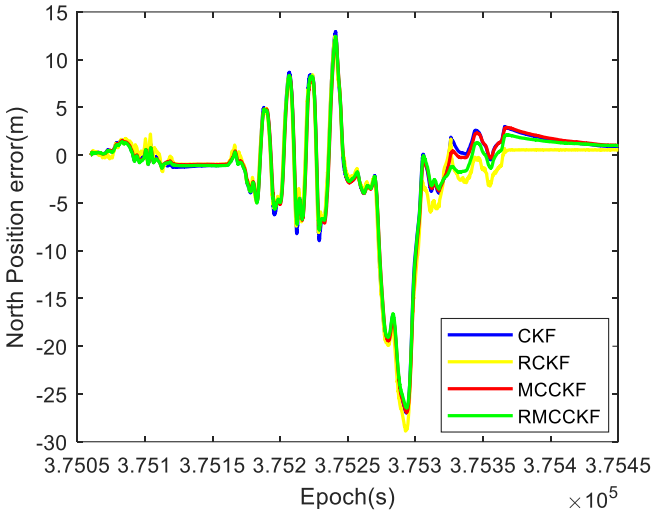

Figure 3.9 position error（North）.

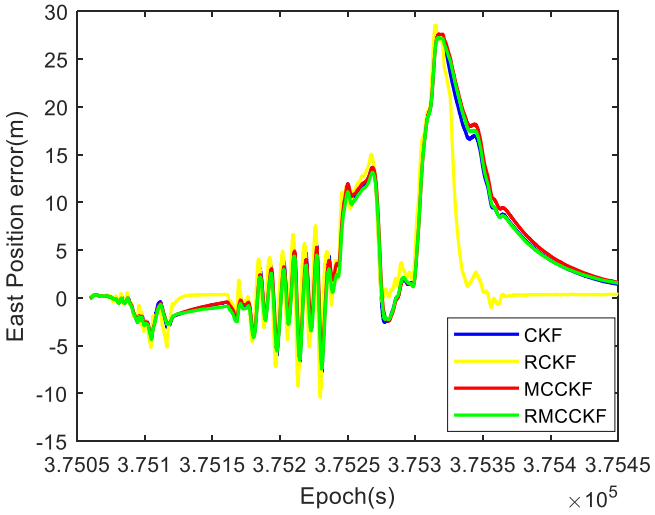

Figure 3.10 position error （East）.

The position improvement of the MCC-based CKF is caused by the attitude improvement

because the position error is coupled with the attitude error. In the simulation, the time cost computed by MATLAB of CKF, RCKF and RMCCKF are 2.85 s, 2.69 s and 4.79 s, respectively, corresponding to 390 s of trajectory data.

Table 3.2 RMSE of navigation parameters for different filters.

Algorithm	EP (m)	NP (m)	Roll (°)	Pitch (°)	Heading (°)
CKF	8.50	6.20	4.00e-2	7.94e-2	4.48
RCKF	6.55	6.46	4.24e-2	6.36e-2	2.24
MCCKF	8.80	6.14	4.28e-2	7.57e-2	1.77
RMCCKF	8.54	6.03	4.07e-2	4.96e-2	0.26

3.5 Conclusions

To improve the robustness of resampling-free sigma-point update framework (SUF), the MCC-based measurement update is imported into the new SUF, where the re-scaled noise and prediction covariance is employed in the formulation of SUF transform. Simulation based on numerical example and car-mounted filed test data are performed to verify the superiority of proposed algorithm. The results show that the RMCCKF not only accelerates the convergence of the MCC-based update with high initial uncertainty, it improves the observed assimilation of the RCKF. Compared with the MCCKF and CKF, the RMCCKF significantly improves the attitude results of land vehicles, especially the heading.

The accuracy of navigation in unmanned vehicle greatly relies on heading and position outcomes, and the results of this study provide ideas for the design of driverless car navigation systems. In this study, the RMCCKF is implemented in a post-processing manner for evaluation and is being considered for a real-time implementation of an embedded navigation system. This approach is generally useful in the use of unmanned vehicles.

This method can provide a high precision and low-cost navigation solution for unmanned ships. Its purpose is that the patrol monitoring by unmanned ships makes up for the limitations of space in-situ monitoring and can realize the real-time dynamic monitoring of the entire region through reasonable analysis methods, that is, it completes the expansion from point to area, which is of great significance for basin management.

Chapter 4 Calculation of pollutant fluxes into the Baiyangdian Lake

4.1 Introduction of the pollutant flux estimation

Lakes are important for providing surface water resources, which have various functions such as water conservation and climate regulation by increasing biodiversity. As the convergence region of rivers, lakes usually accept a large amount of pollutants imported by the inflowing rivers, and the amount of imported pollutants has an important impact on the quality of lake water. Therefore, it is important for regional water management to computer the fluxes of pollutant and to clarify the contribution of each incoming river to the lake. Instantaneous pollutant fluxes are usually subject to dramatic fluctuations due to external factors. Most studies of pollution in natural water bodies and their basins need to focus on pollutant fluxes (e.g., annual fluxes) over long time periods, and the relationship between flow and corresponding pollutant concentration monitoring data over long time periods needs to be studied, therefore relevant monitoring of hydrologic and water quality parameters is crucial in the calculation of fluxes over long time periods. Although there are long-term monitoring of major parameters for water quantity and water quality at hydrological stations of the rivers and lakes by the relevant departments of government, the frequency of monitoring of water quality is usually carried out not as high as the continuous flow and water level records by using automatic monitoring system.

In this chapter, due to the limited types of data available for online monitoring of other rivers, only the pollutant flux of Fu River has been calculated by the monitoring data of this study. Daily average flow data for key river crossings and the water quality data with one month frequency in 2020 in major rivers of the Baiyangdian Lake basin were used to estimate the influent fluxes of total nitrogen (TN), ammonia nitrogen (NH_4^+-N), total phosphorus (TP) and COD to the Baiyangdian Lake by using LOADEST model. Finally, the fluxes of water, TN and TP calculated in Chapter 3 will be compared with the results of this chapter.

4.2 Introduction of Baiyangdian Lake basin

The vector map of Baiyangdian Lake boundary used in this paper is derived from the Global Change Science Research Data Publishing System (*http://www.geodoi.ac.cn/WebCn/doi.aspx?Id=1556*), and this boundary is used in Arcgis software to obtain a map of the counties contained in the Baiyangdian Lake through coordinate conversion and mask extraction (Figure 4.1).

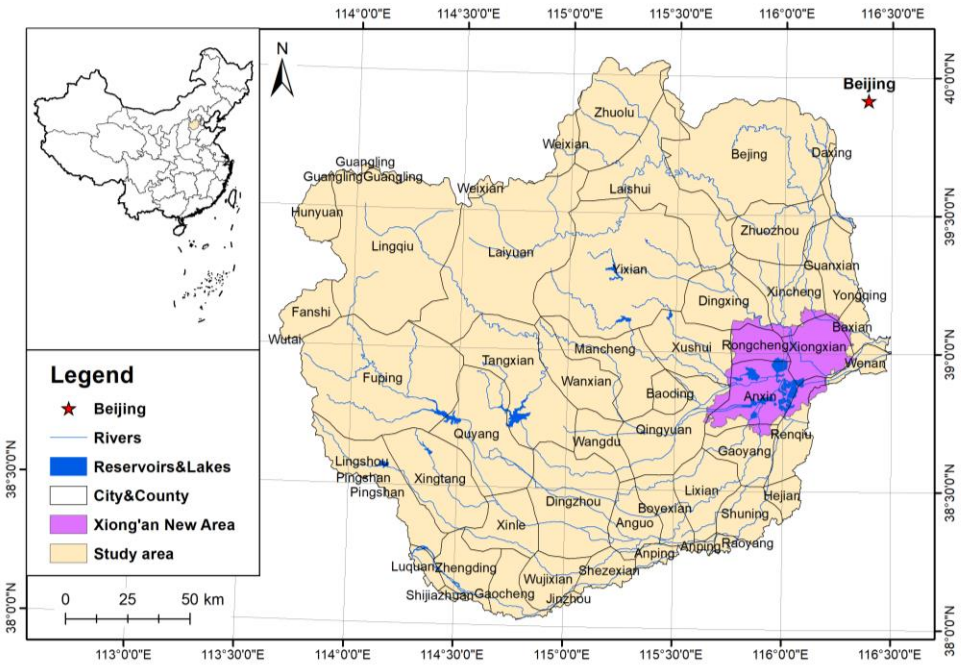

Figure 4.1 Baiyangdian basin administrative division and water system.

The Baiyangdian Lake basin has a total area of about 30,000km^2 and is located in Baoding City, where Xiong'an New Area located in the downstream of the basin. The annual precipitation is concentrated, usually in the summer and autumn months of July and September (Figure 4.2), and its climatic conditions are mainly influenced by warm and humid air masses from the Pacific and Indian Oceans and dry and cold air masses from Siberia.

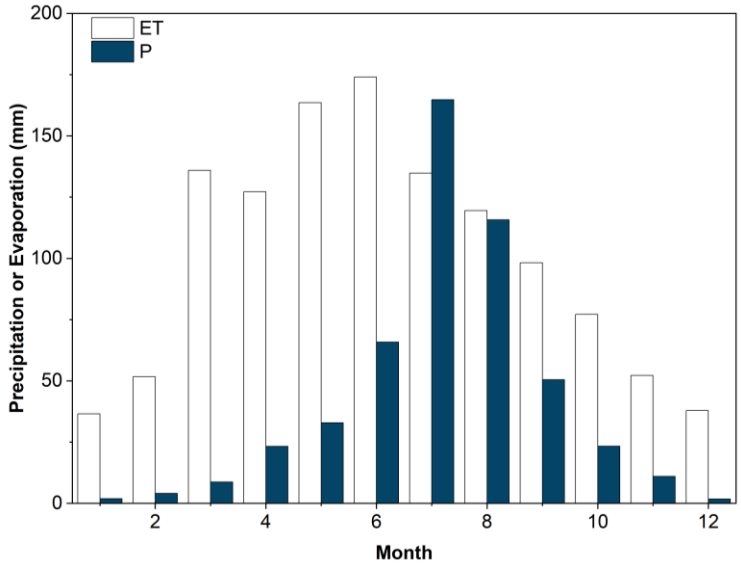

Figure 4.2 Year statistics of precipitation and evaporation in the study area (1977-2021)

The topography of the Baiyangdian Lake basin is sloping from northwest to southeast, showing a gradual transition from mountains to hills to plains (Figure 4.3). The Baiyangdian Lake, which is part of the Daqing River system of the Haihe River Basin, is the largest freshwater wetland in the North China Plain and is a national priority wetland for ecological protection. Baiyangdian Lake, the largest freshwater lake in North China located in the downstream of the basin, has a total area of 366 square kilometers and is dotted with 143 large and small deposits, with an average storage capacity of 1.32 $\times 10^8$ m^3. There are five large reservoirs in the upper reaches of the Bayangdian basin (Hengshanling Reservoir, Koutou Reservoir, Wangkuai Reservoir, Xidayang Reservoir, and Angezhuang Reservoir), of which Hengshanling Reservoir, Koutou Reservoir, and Wangkuai Reservoir are in the southern tributaries, which are injected into the Zhulong River via the Ci River and Sha River, and finally into the Baiyangdian Lake, while Xidayang Reservoir is injected into the Baiyangdian Lake via the Tang River. In the northern tributary, Angezhuang Reservoir is injected into Baiyangdian Lake via the water system in the north tributary (Zhongyishui River, Beijuma River, Nanjuma River and Baigouyin River). The rivers in the Baiyangdian Lake basin are generally characterized by distinct upper river channels and a staggered network of lower rivers, with the highest density of the network concentrated in the Baoding city and the Baiyangdian Lake.

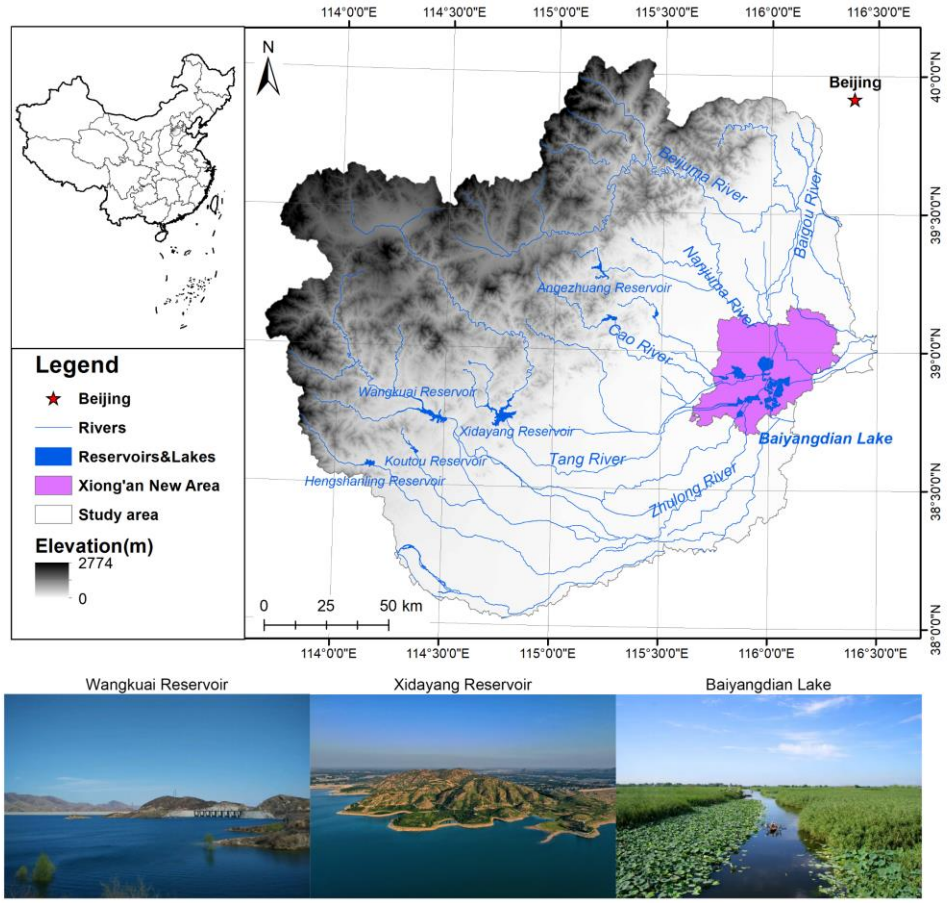

Figure 4.3 Topography and water system of Baiyangdian Lake

4.3 Introduction of Baiyangdian Lake

Baiyangdian Lake (38°43'~39°02'N, 115°38'~116°07'E) is located in the middle of the Daqing River system, east of Baoding city (45 km) and near Beijing (162 km) and Tianjin (155 km), and is an important lake reservoir in the South Branch of the Daqing River Basin of the Haihe River (Du et al., 2018). The Lake area is surrounded by a dike, and the overall terrain of the is inclined from northwest to southeast. The lake area consists of seven large lakes of more than 6.67km^2, such as Baiyangdian, Shaochedian, Chiyudian and Yangjiaodian, and a total of 143 small and medium-sized lakes distributed in a staggered manner (Zhu et al., 2018). Baiyangdian Lake is a shallow lake with an average depth of 2.5 m, a length of 39km from east to west and a width of 28.5 km from north to south (Ji et al., 2019), a total area of 366 km^2, and an average annual water storage of 1.32×10^9 m^3. The topography of Baiyangdian Lake is unique, with more than 3,700 ditches, 120,000 mu of reeds, and villages (Figure 4.4).

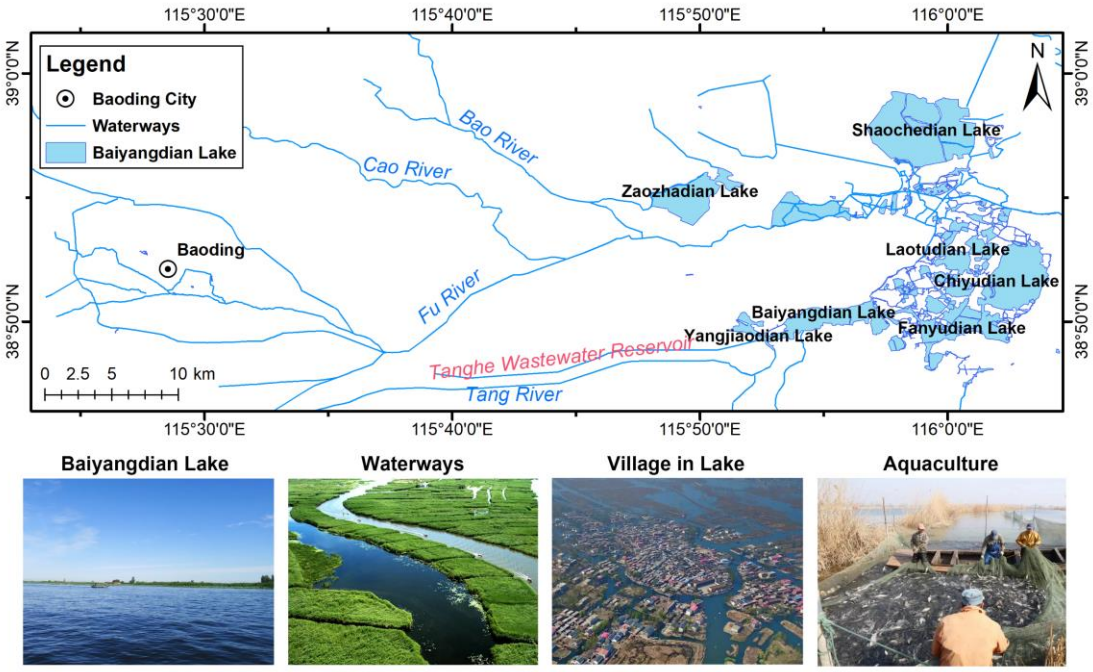

Figure 4.4 Distribution of Baiyangdian Waters

Baiyangdian Lake, as an important flood mitigation area in the middle reaches of the Daqing River system, has been responsible for flood storage of nine upstream rivers entering the lake with an average annual runoff of 3.57×10^9 m^3. However, in recent years, due to the construction of many flood control, flood removal, regulation and irrigation projects upstream, many rivers in flowing into the lake have been in a state of dryness and disconnection all year round, and the flow of rivers into the lake has been decreasing year by year (Wang et al., 2010a). At the same time, the Baiyangdian Lake has been facing the dilemma of no water recharge for a long time due to the year-long drought and low rainfall in North China (Jun and Yongyong, 2017). Due to the long years of insufficient water the downstream floodgate of the area is released for a short period of time and is in a state of prolonged closure. In addition, the complex and changing geomorphological conditions of the Baiyangdian Lake area and the sloping topography from northwest to southeast lead to the accumulation of pollutants in the water environment, which has a serious impact on the ecological environment of the lake.

The pollution of Baiyangdian Lake is widely distributed, with many and complex pathways. The main sources of pollutants in the lake can be divided into point source pollution and non-point source pollution. Point source pollution is caused by exogenous pollution from upstream into the lake and rivers as well as input pollution from industrial wastewater from factories to Baiyangdian

Lake. There is also a series of non-point source pollution such as domestic sewage discharge from improperly treated basin towns and rural areas, input of agricultural pollution from the area through surface runoff, development of tourism and pollution input caused by aquatic and livestock farming. According to the water quality monitoring data of Baiyangdian Lake from the Hebei Provincial Environmental Monitoring Center, the conventional water quality indicators such as COD, TN and TP at the monitoring points in the lake can no longer meet the functional requirements of the water environment, which makes the contradiction of the ecological construction of Baiyangdian Lake more acute.

Table 4.1 Pollution sources and pathways of Baiyangdian water bodies

Emission type	Pollution Sources	Pollution pathways
Point source pollution	Lake-fed rivers	Upstream rivers entering the lake sink pollutants into the precipitation. Among them, Fu River is the most polluted and is the discharge channel for Baoding's domestic and industrial sewage.
	Agriculture	Soil erosion and the sink of pollutants such as pesticides and fertilizers used by villagers around Baiyangdian Lake and the village farmland in the lake to the water body.
	Livestock breeding	Aquaculture, livestock and poultry farming feed and waste from the farming process pollute Baiyangdian Lake.
	Cities & Towns	Pollutants from surrounding towns sink into the lake through surface runoff.
	Tourism Industry	The rapid development of tourism, Baiyangdian Lake motor boats and the number of tourists increased year by year, pollution increased.
Non-point source pollution	Sediments	Baiyangdian Lake sediments are heavily silted up and there is a risk that endogenous contamination of sediments will be released into the overlying water column and enter the food chain through benthic fauna and crisis human health.
	Village within the lake	The domestic sewage of the village is mostly discharged directly into the lake without treatment, which has a serious impact on the water ecology of Baiyangdian Lake. Especially, the successful implementation of a series of measures such as exogenous pollution control and ecological water replenishment after the establishment of Xiong'an New Area has made the pollution problem of the village gradually become the main pollution problem of Baiyangdian water environment.

4.4 Water System and Sub-basin Division of Baiyangdian Lake

There are nine rivers entering Baiyangdian, which are Baigouyin River, Cao River, Lilong River, Tang River, Fu River, Bao River, Qingshui River, Ping River and Xiaoyi River, the historical name of "the end of the nine rivers". After several natural and man-made changes, Tang River no longer flows into Baiyangdian Lake. At present, Fu River and Xiaoyi River keep delivering water to Baiyangdian Lake, while Baigouyinyin River and Bao River seasonally deliver water to Baiyangdian Lake. The remainder of the river has been disconnected. In this study, the data collection covered the entire Baiyangdian (Figure 4.5). Since not all rivers in the watershed flow

into Baiyangdian, Fu River, Xiaoyi River, Baigouyin River and Bao River were selected as sub-basins by ArcSWAT, as well as pollutant loads and sources in Baiyangdian were also analyzed (Figure 4.6).

Figure 4.5 The study area

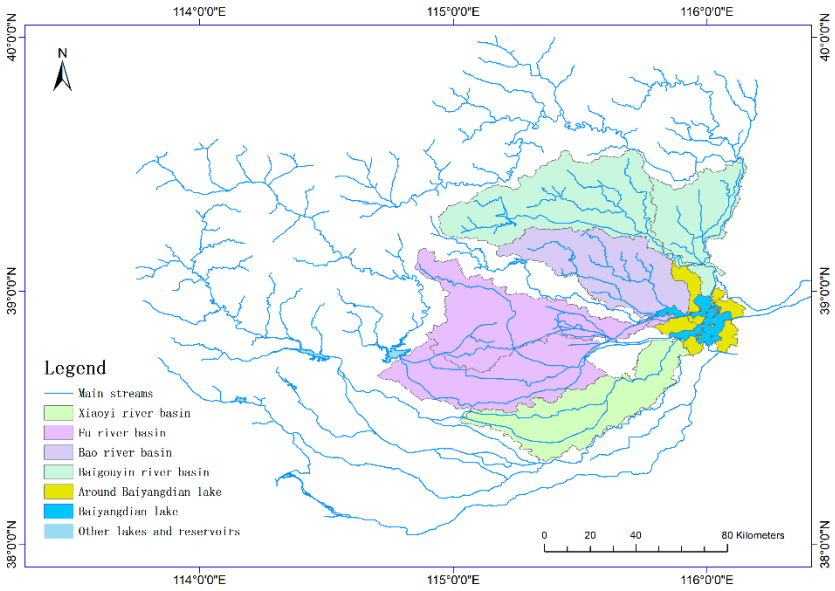

Figure 4.6 The division of sub-basins

4.5 Model for estimating pollutant fluxes into the lake

4.5.1 The model of LOADEST

The water quality data used in this chapter are derived from the monthly water quality monitoring data of the departments of Chinese government. Obviously, the monthly water quality monitoring data cannot represent the average concentration of pollutants in that month which will be used to calculate the pollutant load. To accurately estimate the pollutant load to the lake, this paper uses the LOADEST (Load Estimator) model to simulate the daily pollutant load. The model of LOADEST is a procedure used to estimate the load of pollutant in a river, and a multivariate nonlinear regression equation was established by the pollutant load and flow in the river (Cohn, 1995). This model provides a total of 11 regression equations to describe the variation of solute fluxes in river waters, and the choice of specific equations in the model is made by the software from the predefined equations. The optimal model is selected based on the results of the Akaike Information Criterion (AIC) and the SPPC (Schwarz Posterior Probability Criterion) criteria obtained from the model. The SPPC value is not used directly in the selection of the model, but only used for comparison. After the selection of the optimal model, the estimation of the model parameters can be determined by choosing the most appropriate one of the three methods for evaluating the parameters provided by the LODEST model. Three methods can be used for parameter estimation, consisting of: (1) Minimum likelihood estimation, MLE (Cohn et al., 1989); (2) Adjusted maximum likelihood estimation, AMLE (Cohn et al., 1992); (3) Least absolute deviation, LAD (Powell, 1984). The corresponding method was selected for parameter estimation based on the normality of the residuals and the completeness of the data. The normality testability of the model residuals can be discriminated based on the probability curve correlation coefficient (PPCC) value. The closer the PPCC value is to 1, the more the residuals tend to be normally distributed.

4.5.2 Pollutant Flux Estimation Formula

The LOADEST model is used for flux estimation, which relies on the 11 regression equations already available in the model. If there are incomplete, too little, valid data for the river, or anthropogenic disturbances such as sluice gates along the river, the failure might exist in fitting the simulated data and observed data due to the computational non-convergence. Then the pollutant

flux estimation equations need to be used for estimation in this case.

Three equations can be used to calculate the total flux for different cases as follows:

$$W_1 = K \sum_{i=1}^{n} \frac{C_i Q_i}{n} \tag{4.1}$$

$$W_2 = K \sum_{i=1}^{n} C_i \overline{Q_P} \tag{4.2}$$

$$W_3 = K \frac{\sum_{i=1}^{n} C_i Q_i}{\sum_{i=1}^{n} Q_i} \overline{Q_r} \tag{4.3}$$

K is the time transition coefficient computed from the frequency of monitoring time and flux period; C_i is the instantaneous monitored collection of pollutants; Q_i is the transient movement of water; $\overline{Q_r}$ indicates the average data in the expected period; $\overline{Q_P}$ denotes the average flux data of C_i; n indicates the amount of data.

Equation 4.1 weaker the time variation of runoff; Equation 4.2 highlights the contribution of runoff to every pollutant measurement of concentration; Equation 4.3 highlights the mean temporal change in total runoff, with a weaker response to runoff variation; each equation is used in different situations.

4.6 Analysis of monitoring data

4.6.1 Monitoring design of surface water and groundwater

The monitoring site of river water and groundwater along the river in the field was selected to set up the monitoring system. The rules of the water quantity and quality monitoring for surface water and groundwater are as follows.

(1) Monitoring section must be representative, its position and the number of points should reflect the environmental quality of water bodies, spatial and temporal distribution of pollutants and the law of change and try to obtain the best representation with fewer sections.

(2) Monitoring sections of river should avoid stagnant water areas, backwater areas and outfalls, and try to choose a stable river (sediment) bed, straight river sections, wide river surface and smooth flow of water.

(3) Monitoring site should consider the traffic conditions, economic conditions, implementation security and the hydrological information is easy to obtain, etc., to ensure the

feasibility and convenience of monitoring.

Groundwater monitoring points were chosen based on the following rules.

(1) Integrity and comparability. In the overall and macroscopic scale, it should be able to control different hydrogeological units, and it must be able to reflect the changes in groundwater in the region and changes in groundwater contaminants, and monitoring position focus on different aquifers in the study area.

(2) Representation and balance. The principle of the density of monitoring points is dense for the main study area, the general area is sparse. Monitoring points are dense in areas with serious pollution and sparse in areas with less pollution. As far as possible, the minimum monitoring points to obtain sufficient representative groundwater environmental information.

(3) Feasibility and continuity. Consider the representation of the monitoring results and the feasibility and convenience of the actual monitoring, give priority to the existing observation holes for monitoring, and maintain the continuity of the groundwater monitoring point networks to determine the flow field and its water quality changes.

The surface water monitoring points are located along the route from Fu River to Baiyangdian Lake (Figure 2. 35), which Fu River is the only river that has water all year round among the 9 rivers flowing into the Baiyangdian Lake. It accepts domestic and industrial sewage from Baoding City, with a low sewage treatment rate and a high rate of entering the river. Additionally, the water transfer from upstream or basin outside (such as South Water Transfer to North, water from Yellow River) also flowing from this river to the Baiyangdian Lake. 1105m^3 of domestic sewage and wastewater flows into the river every day, accounting for 45.24% of the average flow of the river. Sewage discharge from agricultural, rural, and scattered enterprises along the river, and agricultural surface pollution from excessive chemical fertilizers on the farmland around the Baiyangdian Lake directly threaten the safety of the water environment. Aquaculture is an important economic industry in the wetland area, and its accompanying discharge of breeding wastewater, etc. leads to an increase in the concentration of ammonia nitrogen, total nitrogen, and total phosphorus in the waters of the lake. The establishment of automatic surface water and groundwater monitoring points along Fu River in the upstream, midstream, downstream and lake area respectively can reflect the changes of surface water quantity and quality in real time from the river to the lake. It also provides an intuitive

reference for future water environment monitoring in the Xiong'an New Area.

Groundwater monitoring points were selected to set several key positions along the vertical direction of Fu River (Figure 2.35). The monitoring points are distributed in the surface water sampling points on the south and north bank of Fu River to monitor the changes of groundwater quantity and quality, which can provide groundwater data support for Baiyangdian Lake environmental protection.

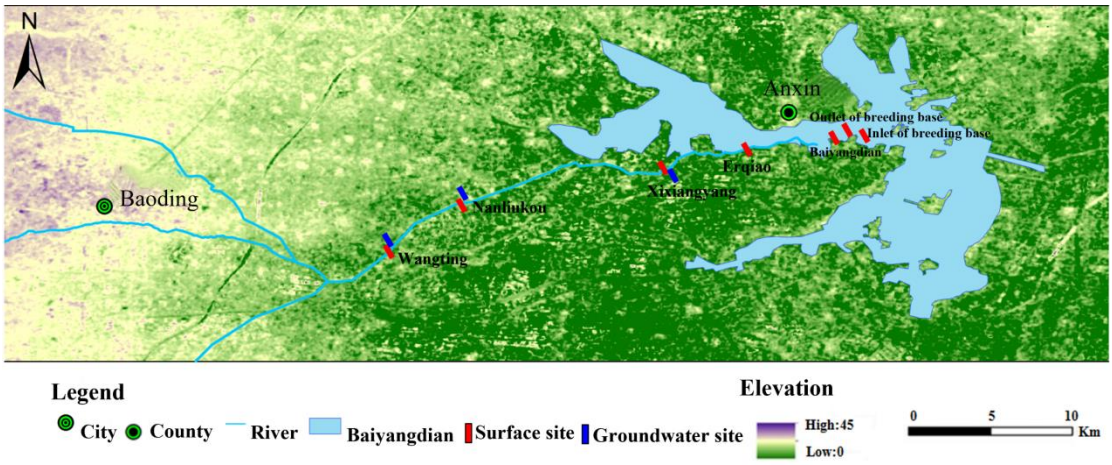

Figure 4.7 Automatic monitoring sections of the water quantity and quality for surface water and groundwater from Fu River to Baiyangdian Lake.

4.6.2 Monitoring Network for Surface Water and Groundwater

The automatic monitoring devices of water quantity and quality for surface water and groundwater were installed at 5 sections from Fu River to Baiyangdian Lake since 2018. The devices in field were shown in Figure 2.35. Beside the automatic monitoring sites, we collected water samples at 7 sites along Fu River and in Baiyangdian Lake to nvestigate the change of parameters of water environment in Baiyangdian Lake and its upstream inflowing river (Fu River). Currently, the surface water and groundwater monitoring equipment is operating normally (Figure 2.36). We evaluated the automatic monitoring data of the river when comparing with the analysis data with water sampling. It was found that the fitting relationship of automatic monitoring data of conductivity (EC) and the analysis data of nitrate or other data of water chemicals can be used to calculate the dynamic changes of nitrate of water in a long time series.

Figure 4.8 Environment around the surface water and groundwater automatic monitoring devices from Fu River to Baiyangdian Lake

4.6.3 Application of water monitoring system

The low-power and wide-area cellular narrowband 5G network (NB-IoT) technology is used to build the Internet of Things monitoring system for surface water and groundwater quality to achieve automatic data collection and transmission. This study set up 7 sites with monitoring of surface water quantity and quality along Fu River to Baiyangdian Lake, including 4 sites along Fu River, Wangting (F2), Nanliukou (F3), Xixiangyang (F5-1), Erqiao (F6), and 3 sites in Baiyangdian Lake, located near the inlet B1. At the same time, three automatic groundwater quantity and quality monitoring points are set up near the surface water monitoring points along the route from Fu river to Baiyangdian Lake, including Wangting (F2), Nanliukou (F3) and Xixiangyang (F5-1). The spatial distribution map of monitoring points and the surrounding environment of the points are shown in Figure 4.7 and 4.8. Monitoring indicators include water level, water temperature (T), electrical conductivity (EC). The frequency of online monitoring of water quality is 30 minutes. Figure 4.9 shows the data of 4 monitoring points of surface water.

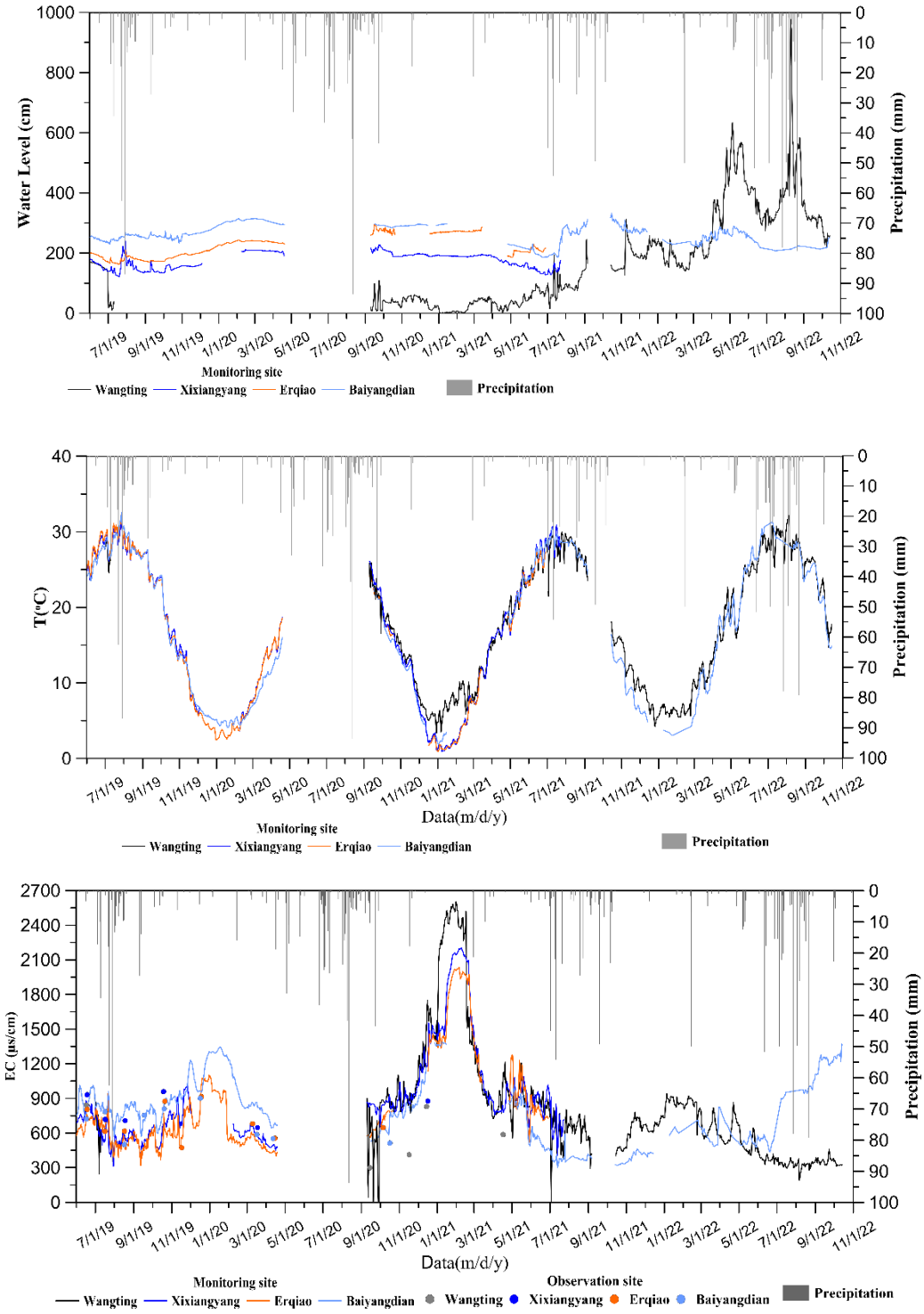

Figure 4.9 Data of automatic monitoring network.

Two soil water monitoring devices was installed near Fu River to collect soil water content, EC and temperature, respectively. The monitoring depths were 50cm, 75cm, 100cm, 150cm, 200cm, and 250cm, respectively (Figure 4.10， Figure 4.11). This was used to testify the reliability of the monitoring system which can not only be used in surface water, groundwater in situ, but also in soil

water by connecting different sensors. In monitoring of soil water, Hydro-probe sensor was connected into our monitoring system.

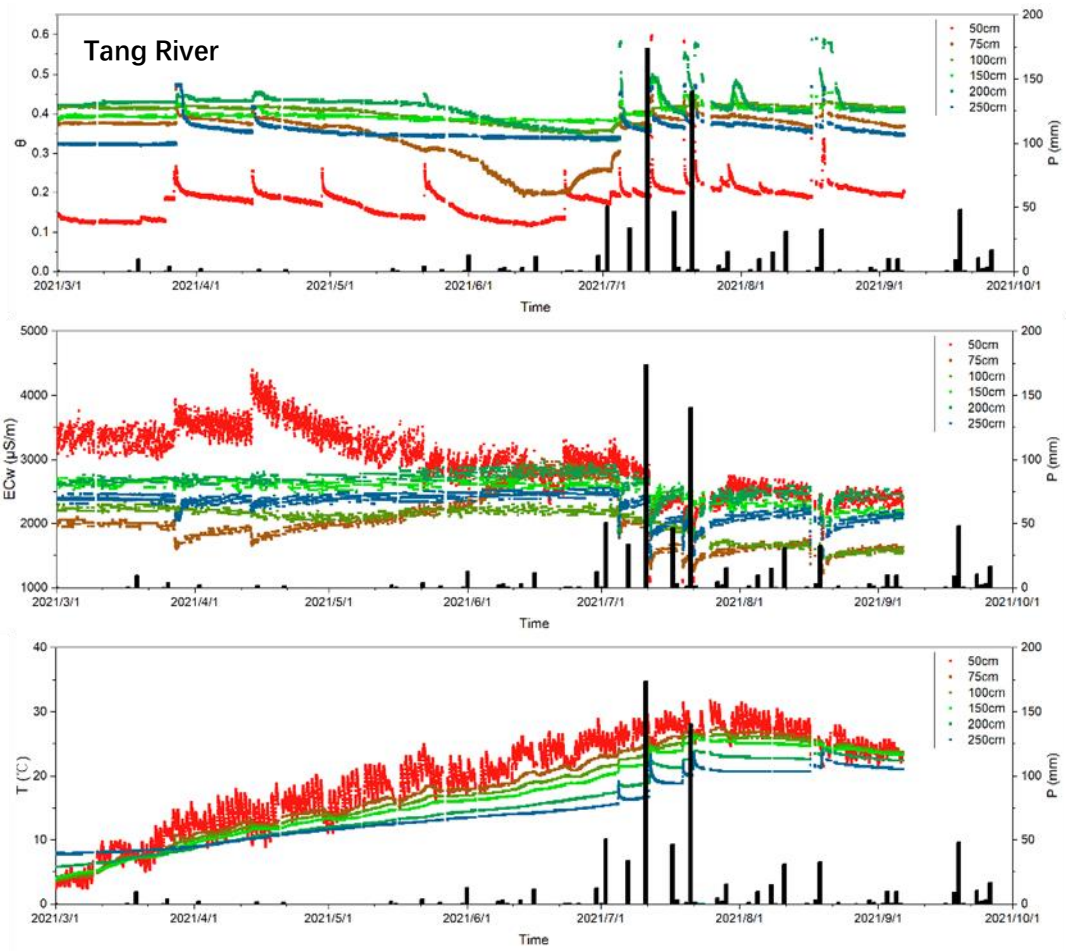

Figure 4.10 Data of soil water monitoring（Tang River）

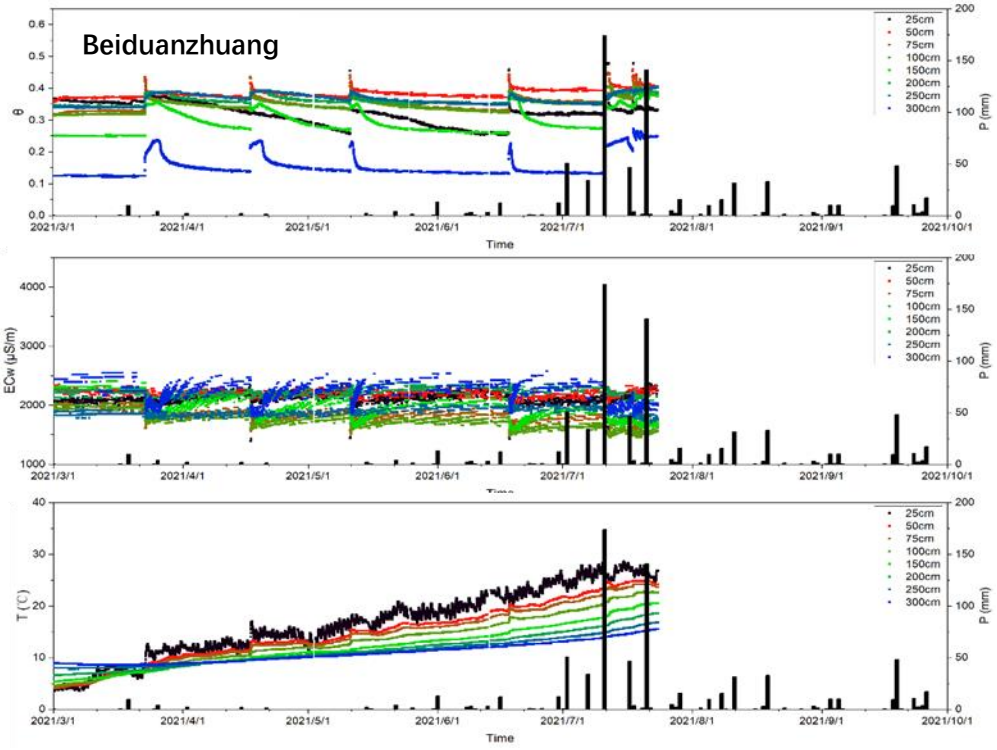

Figure 4.11 Data of soil water monitoring（Beiduanzhuang）

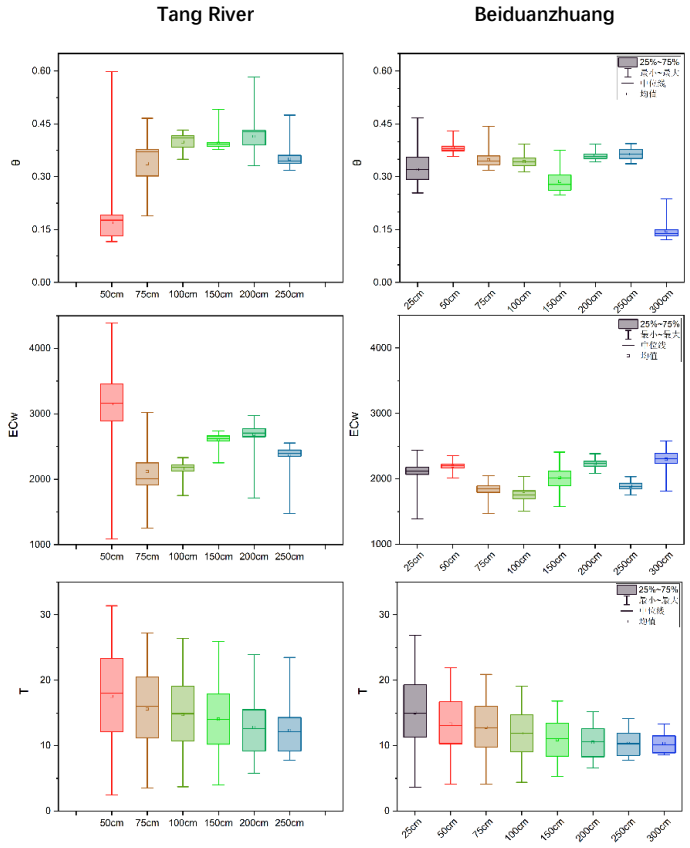

Figure 4.12 Analysis chart of monitoring soil data

The macropores in the shallow soil of the ancient Tang River channel lead to large water loss

and evaporation in the shallow soil, and the water content is greatly affected by the external environment; Affected by evaporation and historical sewage irrigation, the EC value of shallow soil pore water is high; In addition, the large variation of temperature reflects the rapid infiltration of precipitation and irrigation water.

In the homogeneous soil of northern Beiduanzhuang, the water movement rate of each layer of the soil is similar, resulting in the similar space-time distribution of water content above 250 cm, and the difference of soil water EC value is not large; The soil temperature of each layer presents normal distribution, and the temperature variability is low as the depth increases (Figure 4.12).

4.6.4 Calculation of pollutant fluxes from Fu River into the Baiyangdian Lake

The general flux of TP and TN can be calculated using the following two equation:

$$W_2 = K \sum_{i=1}^{n} C_i \overline{Q_P} \quad (4.2)$$

Where, W_2 is the flux of TP or TN, W_{TN} is the flux of TN; S is the cross-sectional area of the river;

C_i is the transient monitored pollutant concentration; Q_i is the transient flow of water; $\overline{Q_r}$ denotes the average data in the expected period; $\overline{Q_P}$ denotes the average flow data of C_i; n denotes the data volume.

Figure 4.13 shows the relationship of water depth and flux.

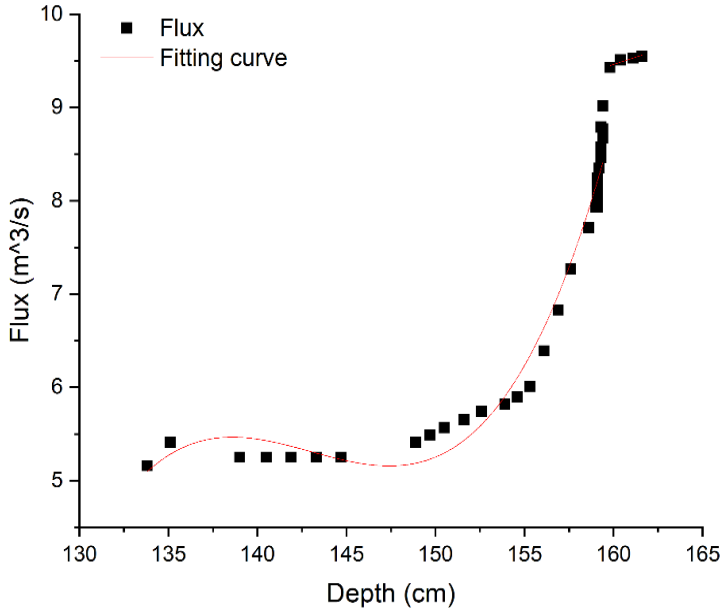

Figure 4.13 Relationship of water depth and flux

Figure 4.14 shows that the highest water flux occurred in winter (from November to January) when water transfer increased the water flux and the second highest water flux occurred in rainy season from July to August in 2020 and 2022. While the highest water flux in 2021 occurred in rainy

season with abundant rainfall during that year. The TN flux shows the similar trend with water flux with highest value in December and second highest value in rainy season. This suggests the water transfer resulted in the increasing TN flux into Baiyangdian Lake. The highest TP flux occurred in October of 2021, and the second highest TP flux occurred in Jun 2020, then in Jun 2021. During the observation period from July 2020 to Sep 2022, the total water flux, TN flux and TP flux were 308 million m^3, 1377 tons and 18.6 tons, respectively. The annual water flux, TN flux and TP flux in 2021 were calculated to be 162 million m^3, 780 tons and 10.6 tons, respectively.

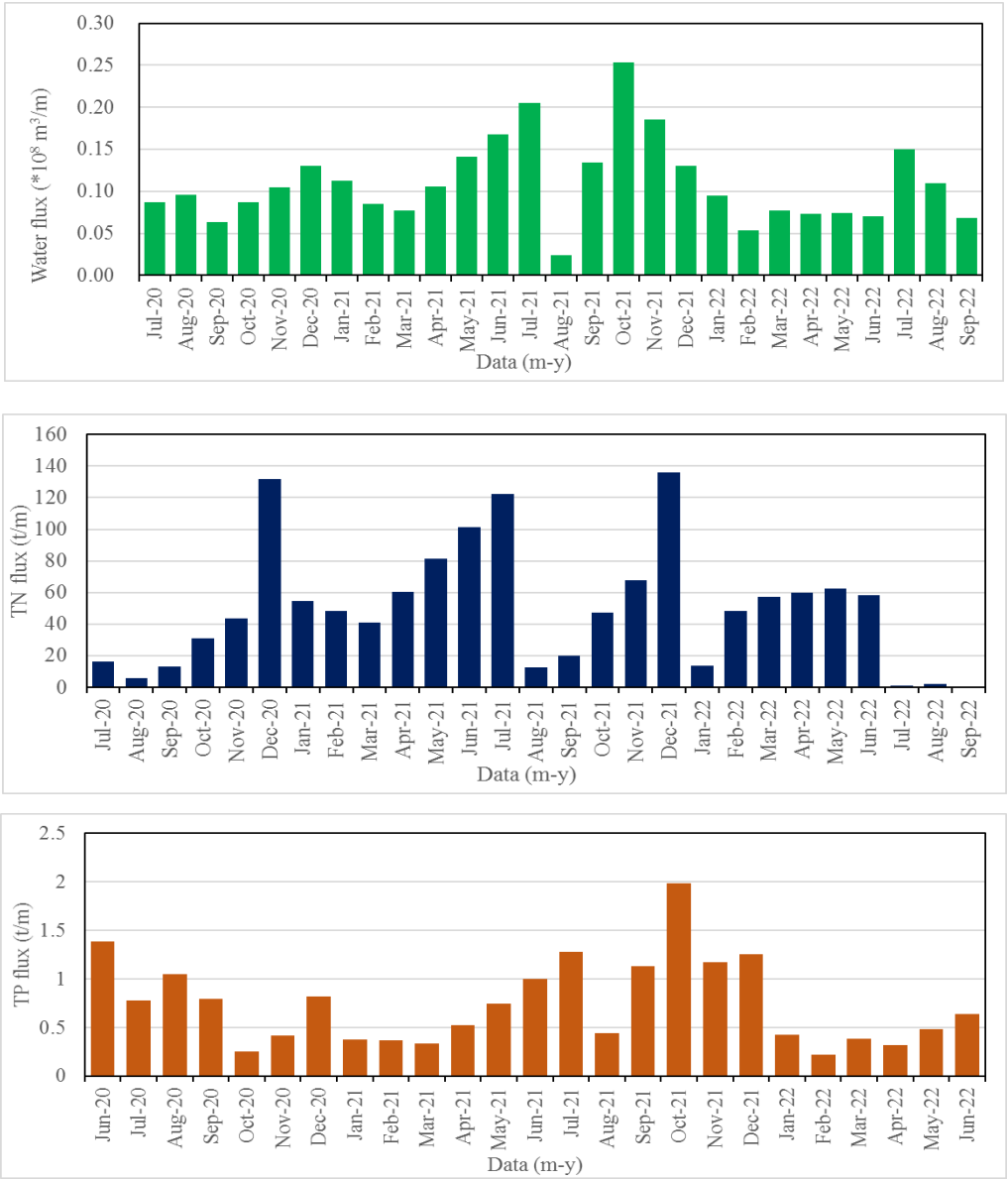

Figure 4. 14 The Water flux, TN flux and TP flux from Fu River.

4.7 Total pollutants transported into Baiyangdian Lake

The exogenous pollutants in Baiyangdian Lake are brought in by the river water through the river entering the lake. The pollutant flux into the Fu River can be calculated to determine the pollutant flux more accurately into the lake. The pollutants in Baiyangdian water area are mainly brought out by the river through the outlet river, so the pollutant outlet flux can be calculated from the pollutant flux of the outlet river.

With the intensification of human activities and the impact of climate change and other factors, many rivers entering Baiyangdian Lake have dried up or become seasonal rivers. In recent years, only Fu River and Xiaoyi River have water flowing into the lake all the year round, while Baigouyin River and Baohe River have water flowing into the lake during the replenishment period, and the downstream of other rivers and reservoirs are in a state of cutoff.

4.7.1 Flux of pollutants from Fu River

The source stream of Fu River originates from Yimuquan Village in Baoding City, flows southeastward through the western and southern urban areas of Baoding City. Then Fu River flows eastward, joins Huandi River and Xinjinxian River, turns northeast, passes Wangting Town, and flows into Baiyangdian at Nanliuzhuang Village, Anxin County. The total length of Fu River is 62.1km and the drainage area is 643.2km^2.

This study used the monthly water quality data from Anzhou Water Quality Station during 2018 to 2020, and the daily average flow data from Xiaowangting/Dong'an Hydrological Station during 2018 to 2020. The water quality measurement stations and hydrological measurement points did not coincide. And the flow change of Fu River was affected by human activities since there were many sluices along the river. These facts were leading to the non-convergence of the fitting using the LOADEST model, so the time-period flux is selected for calculation.

According to the flux results, the change trend of the total amount of COD_{Cr}, COD_{Mn} and BOD in the Fu River is relatively consistent, and the peaks appear in the rainy season from July to August. In addition, the water replenishment in April and the end of each year from 2018 to 2020 also caused an increase in the total amount of sedimentation, but only 1/3 to 1/2 of the input in the rainy season. Nutrient salt pollutants (ammonia nitrogen, total nitrogen, and total phosphorus) also had obvious peaks in the rainy season. Upstream water replenishment had no obvious effect on the flux changes

of ammonia nitrogen and total phosphorus, and there was only a slight increase during water replenishment. The change trend of total nitrogen was like that of COD_{Cr}, and it increases obviously during water replenishment. This may be because the reclaimed water discharged by the upstream sewage treatment plant is an important part of the replenishment water. At present, the current sewage treatment plant in Baoding only has rigid requirements for ammonia nitrogen and total phosphorus in the discharged water, but not for total nitrogen. Therefore, ammonia nitrogen and total phosphorus are less affected by the discharge of reclaimed water, and total nitrogen will increase as the proportion of reclaimed water in the river increases.

All pollutants flux was significantly increased during the rainy season, indicating that rainfall has a key role in promoting the flow of surrounding pollutants into the river, which means that upstream The main source of pollution entering the Fu River is non-point source pollution.

Table 4.2 shows the annual flow, annual water replenishment amount and annual pollutant flux of the Fu River from 2018 to 2020. Comparing with 2018, the total amount of all kinds of pollutant entering the sediment has decreased significantly in 2020, among which ammonia nitrogen and total phosphorus had both decreased by more than 50%. Combining the total flow in 2018 and 2020, it could be considered that the improvement of water quality was the key reason for the decline of pollutants amount of entering the lake from the Fu River.

Table 4.2 flux of water and pollutants from Fu River to Baiyangdian Lake during 2018-2020

Year	Annual flow ($10^8 \times m^3$)	Annual water replenishment ($10^8 \times m^3$)	Total amount of pollutant [a](t)					
			COD_{Cr}	COD_{Mn}	BOD	NH_3-N	TN	TP
2018	1.98	1.98	4167.8	1081.2	697.7	333.8	1398.9	59.4
2019	1.75	1.75	3034.5	999.9	445.7	168.3	949.2	27.8
2020[b]	1.63	1.42	2813.3	789.6	327.5	195.7	844.8	27.6
Sum	5.37	5.15	10015.6	2870.4	1470.9	697.8	3193.0	114.9

a, the water quality data was from Anzhou Water Quality Station, and the hydrological data was from Xiaowangting/Dong'an Hydrological Station.
b, the flux was 0 from October 16th to December 31st, 2020.

4.7.2 Flux of pollutants from drying river of Zhulong River, Cao River and Ping River

The Zhulong River is the section of the river after the confluence of the Shahe, Maghe and Mengliang rivers in the north of Junshen Village in Anguo City, which flows northeast through Anping County, Bono County and Li County to the north of Zhangboshizhuang Village in Gaoyang County and flows into the Lake. The total length of the river is 79.24 km, with a total basin area of 9430km^2. It is a seasonal river with a large sand content.

Cao River, which belongs to the Daqing River system in the Haihe basin, originates from Wuhuiling in the Taihang Mountains in Yixian County, Baoding City, and was originally a tributary of Xu River, called Cao River in Caohe Town, and then changed its flow to merge into Fu River and enter Zaozhadian Lake.

Ping River is located in the northeast of Baoding City, originating from Nanxing Village in the southwest of Dingxing County, flowing through Dingxing, Xushui, Rongcheng and Anxin counties, and injecting into Zhaoli Dian in Anxin County through the southeast of Heilongkou, with a total basin area of 443 km² and a total main-stream length of 30 km.

The relevant data of the above three rivers are currently available as the bi-monthly water quality data of Beiguo Village, Caohe Town, and Xiahexi sitefrom 2016 to 2020. According to the data of the water quality observation point at the observation site, these three riversare in a dry state from 2016 to 2020, and no river water enters Baiyangdian Lake, so the flux of the Zhulong River, Cao River and Ping Riverinto the lake is no longer calculated, which all were referred as zero.

4.7.3 Flux of pollutants from Bao River

Bao River originates from the east side of Langya Mountain, flows southeast from Shaling to Xishan North and then turns northeast, flows to Tang Lake and then turns southeast, flows into Xushui County after Qucheng, flows through Anxin County, Rongcheng County and other places into Zaozhadian Lake, the river length in the territory is 33 km.

There were water quality data of Xushui station once every two months from 2016 to 2020. The flow data show that the Bao River was a seasonal river, and the flow flux during 2018-2020 was greater than 0 (non-disconnected), which are shown in Table 4.2.

The data show that the middle line of South-North Water transfer Project for the Baiyangdian Lake ecological water replenishment was transferred into Baiyangdian Lake via Baohe reservoir (Xushui) and Bao River, and a total 0.19 $\times 10^8$ m^3 of water was transferred in 2018, 0.31 $\times 10^8$ m^3 in 2019 and 0.36 $\times 10^8$ m^3 in 2020. We could confirm that the period with water flow listed in Table 4.3 is in accordance with the time of water transfer, and the total amount of water passing through Xuzhou station in that year was basically equal to the total amount of water transferred in the corresponding year (Table 4.4). Additionally, it was presumed that the water data of Xushui observation station were data of water transfer and water produced from seasonal rainfall.

Table 4.3 Periods when the average daily flow of the Bao River is greater than 0 as observed at the Xushui water quantity station in 2018-2020

Year	Month	Time Period
2018	May	1-31
	June	1-30
2019	September	1-13，17-30
	October	1-31
	November	1-15
	December	11-31
2020	January	1-31
	February	1-27
	May	24-31
	June	1，7-30
	July	1-31

The government department has built the water quality monitoring station in Xushui, downstream of the Bao River. The frequency of water quality analysis in 2018 and 2019 was once every two months, and that in 2020 was once a month. However, the water quality data in December 2019 and April and December 2020 recorded as 0. The water quality data of the Bao River in August, October and November 2020 was greater than 0, but the daily average flow data from August to December of 2020 were 0. For such cases it was referred as that there was no flux into the lake, and the water quality data with zero flow data were not adopted.

Compared with the natural runoff, the period of water transfer was short, with the large flow rate and stable water quality. Therefore, the flow rate was the main factor affecting the flux. At the same time, due to the lack of water quality and quantity data in some months, it did not meet the requirements of the model fitting. The final choice of flux estimation method was using Then the water quality data with flow rate during that period was used to calculate the flux. Since there was no water quality data of the Bao River entering the Baiyangdian Lake, it was impossible to confirm the final proportion of runoff through the Xu Shui water quantity station to inlet of the lake. Therefore, the estimated flux here delegating the upper limit of the flux of pollutants from the Bao River entering the Baiyangdian Lake, and the actual value should be smaller than the estimated result. The calculated results are as follows.

Table 4.4 Annual fluxes of water and pollutants in the Bao River for 2018-2020.

Year	Length of river with water	River flow ($\times 10^9 m^3$)	Transferred water ($\times 10^9 m^3$)	Annual volume of pollutants (t)				
				COD_{Cr}	COD_{Mn}	NH_4-N	TN	TP
2018	56	0.19	0.19	150.2	46.9	1.0	-	0.94
2019	94	0.30	0.31	262.7	83.0	5.3	10.3	1.4
2020	123	0.36	0.36	286.7	89.9	9.3	53.6	0.99

NOTE: The actual total amount of pollutants entering Baiyangdian Lake from the Bao River should be less than these values at the Xushui monitoring station. There is no TN data in 2018 and 2019, and the amount of TN shown in the table was that brought by the transferred water from December 2019 to February 2020.

4.7.4 Flux of pollutants from Daqing River and Baigouyin River

The main-stream of the Daqing River is located in the northeast of Baiyangdian Lake and starts from the confluence of the Baigouyin River and the Nanjuma River, which is divided into the Nan- and Beijuma Rivers at Zhangfang, and the Beijuma River, which is called the Baigouyin River after the Xiaoqing River and the Liuli River (Dashi River). The main-stream of the Daqing River flows down to the hub of the Xingaifang.

Xingaifang hub was constructed in 1970, which is the main project of river management of the northern branch of the Daqing River. The flood diversion gate corresponds to the Xingaifang floodway, the Baigouyin River gate corresponds to the Baigouyin River, and the Daqing River irrigation gate corresponds to the original river channel of the Daqing River (Figure 4.15).

Figure 4.15 Xingaifang Hub.

The hydrological data used in this chapter are the water quality data of the main stream of Daqing River, including data from Dongcicun station of Baigouyin River and Behedian station of Nanjuma River once every two months from 2016 to 2019 and once a month in 2020.

Although no data on the water quality of Baigouyin River were available, the water quality

measurement point of the main-stream of the Daqing River was adjacent to the Baigouyin River gate, which was the starting point of the Baigouyin River. Therefore, it could be assumed that there is no significant variation for water quality from the main stream water to diversion water through different division gate. The flow volume entering the Baigouyin River during the period were then calculated according to the proportion of the water crossing the Baigouyin River gate. Table 4.3 showed that the total flow of the main stream of the Daqing River was small, and the dry of the river was mainly in the first half of the year. While the flow was usually generated at the beginning and the second half of the year at the Baigouyin river gate, and the flow varies greatly from year to year. According to the information about the ecological water transer in Xiongan New Area, the Baigouyin River is the final section of the "Nanjuma River-Baigouyin River" recharge route into the Baiyagndain lake, and the recharge water source of this recharge route mainly comes from Angezhuang Reservoir, Wanglong Reservoir, South Water Transfer North Yi Water Retreat and the upstream sewage treatment plant, with 130 million m^3 of water in 2018 and 0.39 million m^3 in 2019. The water transfer was 39 million m^3 and 28 million m^3 in 2019 and 2020, respectively. Table 4.5 show that the amount of water flowing through the river gate of Baigouyin River in 2018 and 2019 was basically equal to the amount of water transferred by the upstream ecology, however, the amount of water flowing through the river in 2020 was significantly larger than the amount of water transferred, due to the inflowing of other sources such as upstream rainfall.

Table 4.5 Overflow time and water flux of the main-stream of Daqing River and Baigouyin River

Year	Daqing River			Baigouyin River			Proportion
	Overflow time		Flux	Overflow time		Flux	
	Month	Days	(10^8m^3)	Month	Days	(10^8m^3)	
2018	1-2	44	0.031	-	-	-	-
	2-5	85	0.067	-	-	-	-
	7-12	180	1.88	8-12	143	1.31	69.7%
2019	1-5	134	0.47	1	26	0.16	34.0%
	7-9	76	0.15	-	-	-	-
	10	13	0.0072	-	-	-	-
	10-12	65	0.25	11-12	61	0.23	92.0%
2020	1-2	37	0.013	-	-	-	-
	7-12	158	0.84	7-12	148	0.80	95.2%

According to the existing water quality data and flow data, LOADEST model was used to fit the water quality indexes of the main-stream of Daqing River. After test and verification, the fitting of COD_{Cr}, COD_{Mn} and total phosphorus indicators were completed. The NH_4^+-N data fluctuates greatly due to the impact of flow break. Additionally, the TN was detected in January 2020, and the

data with too small volumes after the month of flow break was removed. The fluxes of NH_4^+-N and TN were calculated by using flux estimation formula.

Based on LOADEST predefined model, the model parameters and loadings were estimated using water body data from 2016 to 2020. The regression model with the smallest AIC and SPPC statistics was selected to obtain the average daily flow data for the main-stream of the Daqing River. The results of the model fitting are shown in Table 4.6.

Table 4.6 LOADEST model fitting results

Pollutants	LOADEST Loading Calculation Regression Equations	Statistics Indicators		
		R^2	SCR	NSE
COD_{Cr}	$\ln(L) = a_0 + a_1 \ln Q + a_2 dtime$	89.29%	-0.910	0.889
COD_{Mn}	$\ln(L) = a_0 + a_1 \ln Q + a_2 dtime$	90.39%	-0.2367	0.905
TP	$\ln(L) = a_0 + a_1 \ln Q + a_2 \ln Q^2 + a_3 dtime$	75.17%	-0.0447	0.666

where L is the pollutant flux estimated by the model (kg/d), a_0, a_1, a_2, and a_3 are equation coefficients, Q is the runoff volume, and *dtime* is the fractional date after centering. The test statistical indicators reflect a good model fit.

4.8 Total pollutant flowing out of Baiyangdian Lake

The water volume of the Zhaowangxin River is almost entirely controlled by the Zao Linzhuang water-gate at the outlet of Baiyangdian Lake, which is not applicable to the LOADEST model that simulates natural river fluxes, so the period flux equation was used to account for lake discharge fluxes.

Table 4.7 shows the water quantity and pollutant fluxes, which are based on the instantaneous flow and water level data from January 2020 to October 2021 at the Shigezhuang hydrological station downstream of the gate once a day (during the flat and abundant water periods) or once every 10 days (during the dry water period), and the month-by-month water quality data from February 2020 to September 2021.

Table 4.7 Flux and pollutant volumes of Zhaowangxin River in 2019-2020

Year	Month	Days	Flux (10^8m^3)	COD_{Cr}	COD_{Mn}	NH_4-N	TN	TP
2020	5-6	37	0.082	238.2	65.7	1.31	6.08	0.25
	7	7	0.031	94.9	28.5	1.35	3.79	0.17
	8	14	0.170	492.4	139.2	8.83	20.37	0.85
	9-10	24	0.408	898.2	273.5	2.86	35.44	1.63
	11-12	51	0.261	478.8	132.3	2.61	20.02	0.78
Total	-	133	0.952	2202.4	639.2	16.96	86.60	3.68
2021	2-3	38	0.235	410.4	131.3	2.11	18.76	0.70
	6-10	133	9.25	68983.3	21234.3	406.42	6673.94	129.05
Total	-	171	9.49	69393.7	21365.6	408.53	6692.71	129.75

The results show that the time of opening and releasing water in 2020 was mainly occurred after the rainy season, especially from August to October, and the flux of pollutants produced was also large. A total amount of 25.1 million m^3 water passed through Shigezhuang hydrological station during the water transfer period in November-December upstream, which was 1.5 times more than that in August.

4.9 Pollutant fluxes in Baiyangdian Lake

The total amount of five pollutants (TP, TN, NH4-N, CODCr, CODMn) entering and leaving Baiyangdian during the whole natural year of 2020 was selected based on the flux calculation results of the rivers entering and leaving the lake. The accounting results of the inflowing fluxes of water and pollutants from Bao River, Xiaoyi River, Baigouyin River and Fu River were summarized, and the results are shown in Table 4.8, Table 4.9 and Fig 4.16.

Table 4.8 Fluxes and pollutants of rivers entering the lake in 2020

Rivers	Flux (10^8m^3)	Replenishment (10^8m^3)	Pollutants (t)				
			COD_{Cr}	COD_{Mn}	NH_4-N	TN	TP
Fu River	1.63	1.42	2813.3	789.6	195.7	844.8	27.6
Xiaoyi River	0.65	0.65	919.1	255.7	25.4	223.6	4.2
Baigouyin River	0.80	0.28	1063.2	294.2	10.1	150.5	3.7
Bao River	0.36	0.36	286.7	89.9	9.3	53.6	0.99
Total	3.44	2.71	5082.3	1429.4	240.5	1272.5	36.5

Table 4.9 Proportion of flow and pollutants of rivers entering the lake in 2020

Rivers	Flux (%)	Replenishment (%)	Pollutants (%)				
			COD_{Cr}	COD_{Mn}	NH_4-N	TN	TP
Fu River	47.38	52.40	55.35	55.24	81.37	65.49	75.64
Xiaoyi River	18.90	23.99	18.08	17.89	10.56	17.57	11.51
Baigouyin River	23.26	10.33	20.92	20.58	4.20	11.83	10.14
Bao River	10.47	13.28	5.64	5.39	3.87	4.21	2.71

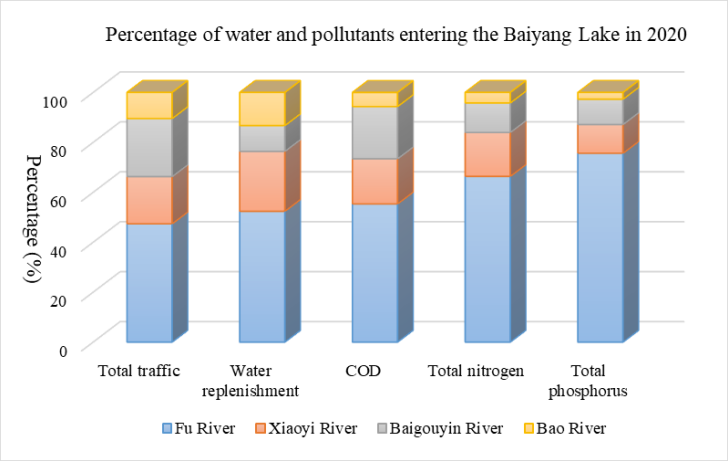

Figure 4.16 Water and pollutant percentages of different rivers reaching Baiyangdian Lake in 2020

The best water quality among the four rivers is the Bao River, where the percentage of each pollutant is less than 10% of the total pollutants in the four rivers. The volume of water in Fu River inflowing the lake in 2020 was largest, which is similar to the sum of the other three rivers, and about 87% of the water in Fu River was recharged by upstream. The volume of water from Baigouyin River entering the lake was 8×10^7 m³ in 2020, of which only 35% came from upstream of the water. And the input of NH_4-N from Fu River accounted for 81.37% of the total input of pollutants to the river.

From the proportional relationship of water volume, the input water of Fu River accounted for 47.38% of the river's water volume into the lake in 2020. Nevertheless, the percentage of input of each pollutant is more than the percentage of water, suggesting that the pollutants per unit of water volume transported into the lake was higher in Fu River than that in other rivers. The total amount of pollutants transported per 1 million m³ of water into the lake was calculated and results are listed in Table 4.10, and it was more obvious that Fu River has higher pollutants flux. For example, the COD_{Cr} transported per unit of water in Fu River was 1.22 times higher than that in the Xiaoyi River, 1.3 times higher than that in the Baigouyin River, and 2.17 times higher than that in the Bao River. Thus, overall, the pollutant content of the Fu River is higher on average than other rivers and should be taken seriously. The amounts of pollutants (COD_{Cr}, COD_{Mn}, NH_4-N, TN, TP) transported per unit of water in Xiaoyi River was less than that in Fu River. The difference on fluxes for COD_{Cr} and COD_{Mn} between Fu River and Xiaoyi River was small, while the difference on fluxes for NH_4-N, TN and TP was larger. The presence of more villages and farmland along the Fu River may be the

main reason, especially between Baoding City and Baiyangdian, producing more non-point source pollution, mainly nitrogen and phosphorus. In addition, the fact that the Fu River receives sewage discharges from Baoding City is another reason for the high pollutant flow. Among the four rivers that enter the lake, the Baoding River has the best water quality.

Table 4.10 the amounts of pollutants carried per unit of water in the lake in 2020.

Rivers	Pollutant fluxes（$t/10^6m^3$）				
	COD_{Cr}	COD_{Mn}	NH_4-N	TN	TP
Fu River	17.26	4.84	1.20	5.18	0.17
Xiaoyi River	14.14	3.93	0.39	3.44	0.06
Baigouyin River	13.29	3.68	0.13	1.88	0.05
Bao River	7.96	2.50	0.26	1.49	0.03

The water quantity and the total amount of pollutants transported for the outflowing river Zhaowangxin River calculated are as shown in Table 4.11. Zhaowangxin River exported about 95 million m^3 of water from Baiyangdian Lake, accounting for 27.67% of the total amount of water entering the lake from the river. The output of both COD_{Cr}, and COD_{Mn} accounted for about 44% of the volume into the lake, while the output of nitrogen and phosphorus nutrients only accounted for about 10% of the volume into the lake. In terms of the total amount of pollutants transported per unit of water, the river water entered the lake and flowed gradually to the outlet of lake, and the COD content in the water increased, while the nutrient salt content decreased significantly.

Table 4.11 Total and proportion of pollutants transported out of the lake by rivers in 2020.

Rivers	Flux	Pollutants				
		COD_{Cr}	COD_{Mn}	NH_4-N	TN	TP
Zhaowangxin River	$0.95\times10^8m^3$	2202.4t	639.2t	16.96t	86.6t	3.68t
Water balance	$2.49\times10^8m^3$	2879.9t	790.2t	223.54t	1185.9t	32.81t
Excretion ratio	27.67%	43.33%	44.72%	7.05%	6.81%	10.08%
Pollutant fluxes（$t/10^6m^3$）		23.13	6.71	0.18	0.91	0.04

4.10 Compare of flux by LOADEST model with that by automatic monitoring data

It has been found that Fu River is a major inflowing river which carries large amount of pollutants such as COD_{Cr}, $COD_{Mn,}$ NH_4-N, TN and TP. The Fu River has a great impact on water quality of Baiyangdian Lake. That is why the Fu River was selected to set up the automatic monitoring system in-situ by using the NB-IoT technology. In Chapter 2, the automatic monitoring system was set up along the Fu River from upstream to downstream and the dynamics of EC (delegating the status of total water quality) and water depth (being used for calculation of water

level), and T during the monitoring period were identified. The inflowing flux of water and pollutants (TN, TP) can be calculated by using the water quantity (estimated by variation of water depth) and concentration of pollutants.

Comparing the calculation results of water flux, TN flux and TP flux in Fu River in this chapter and results by monitoring data in Chapter 2. Using the model, the total annual water flux of Fu River is 163 million m^3, the TN flux is 844.8 tons, and the TP flux is 27.6 tons in 2020. By using the monitoring data, the total annual water flux of Fu River is 162 million m^3, the TN flux is 780 tons, and the TP flux is 10.6 tons in 2021. As the calculation period is in different years, the difference in different is acceptable. Additionally, the results of water flux, TP and TN flux are in range of believable. Even the difference, the online monitoring data collected in this study provide a way to evaluate the flux of water and pollutant. Futhermore, the online and real-time monitoring with high collection frequency and continuity give an effective way to evaluate the water quality. And the calculation method is simple and faster than the LOADEST-model-based calculation method in this chapter, which is more suitable for rapid environmental monitoring, early warning, and other occasions.

4.11 Conclusions

In this chapter, firstly, the flux (water, TN, TP) of Fu River is calculated using the monitoring equipment data of this study, and then, the water flow and water quality data of key sections of inflow and outflow rivers in 2020, which are collected by hydrological sites from government, were used to estimate the flux of major pollutants of COD_{Cr}, $COD_{Mn,}$ NH_4-N, TN, TP inflowing to or outflowing from Baiyangdian Lake in 2020. The LOADEST model and the pollutant flux calculation formula were combined to estimate the flux of input and out from the Lake.

Nonlinear regression equations of the LOADEST model was used to estimate the river flow and pollutant flux. And good fitting results were obtained for the observed and the calculated data that met the requirements. For data with missing values, the corresponding pollutant flux calculation formula was selected according to the characteristics of pollutants and rivers to calculate pollutant flux.

The exogenous pollutants in Baiyangdian Lake were brought in by river flowing into the lake. The pollutants in Baiyangdian Lake are brought out by the river through the outlet river, so the

pollutant outlet flux can be calculated from the pollutant flux of the outlet river. As most rivers that flows into Baiyangdian Lake have dried up or become seasonal rivers. In recent years, only Fu River and Xiaoyi River have water entering the lake all year round, and during the water transfer period both the Baigouyin River and the Bao River have water flowing into the lake. Zhangwangxin River is the only outlet of Baiyangdian Lake.

For above 4 inflow rivers, the fluxes of 5 major pollutants（TP, TN, NH_4-N, COD_{Cr}, COD_{Mn}）in the lower reaches of these inflow rivers and the upper reach of 1 outflow river were calculated. It was found that among the rivers entering the lake, the 4 inflow rivers carried a amount of $3.44\times 10^8 m^3$ of water to Baiyangdian Lake, with 5082.3 t of COD_{Cr}, 1429.4 t of COD_{Mn}, 240.5 t of NH_4-N , 272.5 t of TN and 36.5 t of TP in 2020. The water quality of Bao River was the best of all inflow river, with the smallest amounts of pollutants entering the lake in 2020, while the average pollutant concentration of Fu River was the highest, and the contributions to the total amount of pollutants entering Baiyangdian Lake in 2020 exceeded 50%. The flux of pollutants out of the lake in 2020 is significantly smaller than the flux into the lake, indicating that Baiyangdian Lake has retained a large amounts of pollutants. With the biogeochemical processes in Baiyangdian Lake, the proportion of nutrients in the outflow water decreased significantly, and the proportion of COD increased.

Fu River has the most important impact on the water quality of the Baiyangdian Lake. The water flux, TN flux and TP flux in Fu River was compared with those calculated based on monitoring data by NB-IoT technology. Though there is some differences between results from LOADEST model and by the monitoring data, the differences are acceptable, as the result delegates the water and pollutants flux in different year. The online monitoring data collected in this study provide a way to evaluate the continuous flux of water and pollutant. And the calculation method is simple and faster than the LOADEST-model-based calculation method in this chapter, which is more suitable for rapid environmental monitoring, early warning, and other occasions.

Chapter 5 Source analysis and load analysis of pollutants in Baiyangdian Lake basin

5.1 Introduction

The flux of pollutants flowing into rivers is related to the source and load of pollutants in the whole basin. The point source and non-point source in the Baiyangdian Lake basin composed of the sources of those pollutants. According to the characteristics of the Baiyangdian Lake basin with the impact of the intensive agriculture and dramatic human activities, the non-point source and the load of the Baiyangdian Lake basin was calculated by using the output coefficient method, RUSLE model and PLOAD model and the point source data. As there are only 4 rivers discussed above flowing into the Baiyangdian Lake, which affects the quality of water, the source and load was of the sub-basins of these rivers were also calculated. Additionally, the load of the pollutants in the sub-basins was compared with the flux of pollutants flowing into the rivers which was calculated in Chapter 4. The gaps between the load and the flux inflowing river were discussed in this chapter.

5.2 Pollution load estimation methods

5.2.1 Export Coefficient Modelling (ECM)

The ECM is widely suggested based on its accessible theoretical basis that total pollutants in a watershed contribute to nonpoint source pollution loads (Ongley et al., 2010, Johns, 1996). In this modeling method, different output coefficients were obtained considering animal husbandry and rural livelihood factors, such as livestock numbers and distributed conditions, land use types, and rural residents, to obtain an analysis of agricultural non-point source pollution loads and source. Non-point source pollution loads from rural livelihoods, livestock and agricultural farming in Baiyangdian watershed in 2020 were calculated based on the ECM method, and the final agricultural and rural surface water loads in Baiyangdian were corrected based on the correlation between water quantity data from upstream water stations and downstream water quantity data.

The ECM equation:

$$L_j = \sum_{i=1}^{m} E_{ij} \times A_i \qquad (5.1)$$

where, L_j is the pollutant's average export coefficient, kg/(km^2·a); i is the land use type, there are m types; E_{ij} is the pollutant's export coefficient j type of land use of category i (kg/km^2) or the discharge factor of livestock and poultry of category i type (kg/a) or the output coefficient of per person (kg/a); A_i is the land use type i in the river (km^2) or the number of livestock or population.

5.2.2 Revised Universal Soil Loss Equation (RUSLE)

The RUSLE is used for adsorbed pollutants load (e.g., TN, TP, and COD) which is released into water bodies through soil erosion (Chen et al., 2013, Wischmeier and Smith, 1978).RUSLE is an empirical soil erosion model designed based on the common soil erosion equation as shown in Equation 5.2.

$$W_{ads} = r_{ads} \times \left(\sum_{u=1}^{U} X_u \times A_u \times C_s \times t \right) \times 10^{-3} \qquad (5.2)$$

$$X_u = R \times K_u \times LS \times S_u \times M_u \qquad (5.3)$$

Where, W_{ads} is the adsorbed pollutant load, t/a; r_{ads} is the transportation ratio of sediment; u is the type of land use; X_u is the modulus per year of soil erosion, denotes mean lost soil due to erosion, it is computed by equation 5.3, t/hm^2; A_u is the type u of land use, km^2; C_s is the Contextual concentration of contaminants in the soil, g/kg; t is the concentration rate of pollutants; R is the drivers of erosion by precipitation; K_u is the factor of soil erosion; LS is the factor of topographic; S_u is the conservation factor of soil and water; M_u is the plant cover and management factors.

The rainfall erosion driver R in Equation 5.3 measures the effect of rainfall on the soil and captures the abortion probability associated with the moisture event. R can be evaluated by the following equation:

$$R = \sum_{i=1}^{12} 1.735 \times 10^{1.5 \log \frac{P_i^2}{P_u} - 0.8188} \qquad (5.4)$$

Where, P_u is the average annual precipitation; P_i is the average monthly precipitation.

The soil erosion factor K_u in equation 5.3 is the rate of soil erosion measured on standard plots per unit rainwater erosion index, usually calculated by the soil characteristic attribute properties:

$$K = \left\{0.2 + 0.3exp\left[-0.0256S_d\left(1 - \frac{S_i}{100}\right)\right]\right\} \times \left[\frac{S_i}{C_i + S_i}\right]^{0.3} \times \left\{1.0 - \frac{0.25C_0}{C_0 + exp\left(3.72 - 2.95C_0\right)}\right\} \times \left[1.0 - 0.7\left(1 - \frac{S_d}{100}\right)\right] \Big/ \left\{1 - \frac{S_d}{100} + exp\left[-5.51 + 22.9\left(1 - \frac{S_d}{100}\right)\right]\right\} \quad (5.5)$$

S_d is the amount of sand content, %; S_i is the amount of silt content, %; C_i is the amount of clay content, %; C_0 is the amount of organic carbon content, %.

The LS is topography factor in Equation 5.3 represents the effect of topography on soil erosion. Where the slope length factor L indicates the effect of slope length on erosion and the slope length factor S captures the effect of slope length on erosion.LS can be computed by the following equation.

$$LS = \left(\frac{Q_A \times L_c}{22.13}\right)^{0.4} \times \left(\frac{sin\,\theta}{0.0896}\right)^{1.3} \quad (5.6)$$

Q_A is the flow accumulation; L_c is the cell size; θ is the slope, M_u

The factor M_u for vegetation cover and management in Equation 5.3 is used to capture the effect of vegetation and human management practices on soil erosion rates, which includes crop cultivation and farm management in agricultural areas and the suppression of soil erosion by vegetation cover in forested areas. The factor's value varies with plant growth due to seasonal factors. M_u can be calculated by:

$$M = \begin{cases} 1 & m = 0 \\ 0.658 - 0.3436 \lg m & 0 < m < 78.3\% \\ 0 & m > 78.3\% \end{cases} \quad (5.7)$$

Where, m is the mean annual vegetation cover.

Using the RUSLE model, the parameters of the research area were computed and a series of raster data were acquired, as shown in Figure 5.1. Other used parameters sources were recorded in Table 5.1.

Table 5.1 Parameters sources used in RUSLE model.

No.	Data catagories	Data sources
1	Long-time Rainfall	China Meteorological Administration and U.S. National Climate Data Center
2	Soil property	Soil dataset of China from the Second National Soil Census
3	Soil silty	Harmonized World Soil Database version 1.1, HWSD
4	Terrain	ASTER GDEM 30m
5	Vegetation coverage	Spatial distribution dataset of China Annual Vegetation Index (NDVI).
6	Land use	Spatial distribution dataset of remote sensing monitoring of land use types in China in 2020

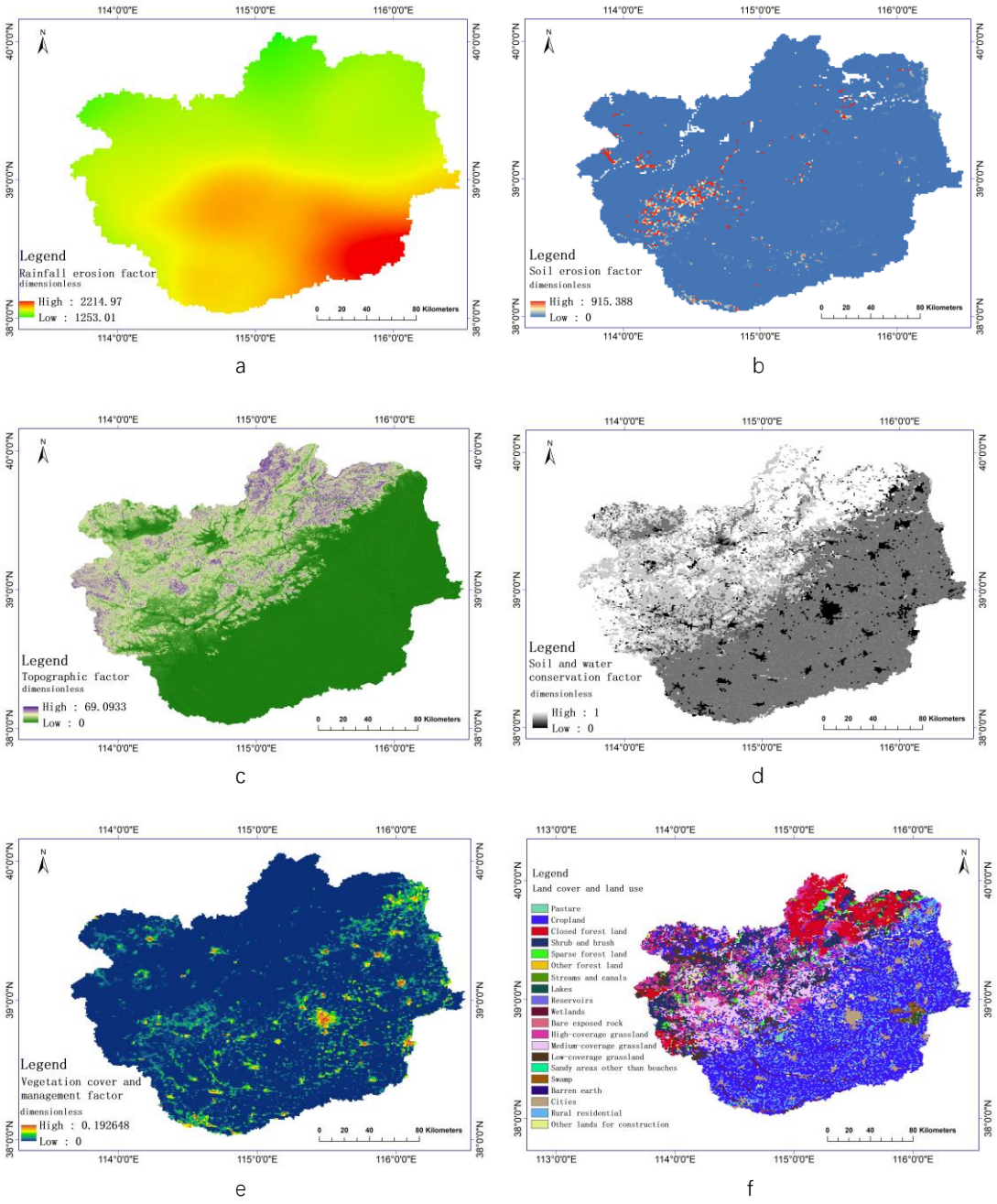

Figure 5.1 Computed results of parameters of RUSLE model in the research area.

5.2.3 PLOAD model

The PLOAD model was used to calculate nonpoint source pollution loads for different sub-watersheds and land uses based on annual rainfall, land use, and optimal watershed management practices (Hill, 2001, Shoemaker, 1997). In this study, the PLOAD model was used to calculate the pollution load of dissolved pollutants from precipitation. The equation is:

$$W_{RJ} = \lambda_{RJ} \times \sum_{u=1}^{U} \left(P \times P_J \times R_{Vu} \times A_u\right) \times C_u \times 10^{-5} \quad (5.8)$$

W_{RJ} is the dissolved pollutants load, t/a; λ_{RJ} is the river entry coefficient; P is the precipitation, mm/a; P_J is the runoff coefficient correction; R_{Vu} is the average runoff coefficient of the land use type u; A_u is the land use type u, hm^2; C_u is the average output pollutant concentrations in precipitation under different land use type u, mg/L.

Using the runoff curve number (SCS-CN) model to compute runoff coefficients (Kang and Yoo, 2020). The model can represent the effects of various soil types, various land use patterns and previous soil moisture content on precipitation runoff. The fundamental equations of the model are as follows:

$$\begin{cases} Q = \dfrac{(P - 0.2S)^2}{P + 0.8S}, & P \geq 0.2S \\ Q = 0, & P < 0.2S \end{cases} \quad (5.9)$$

$$S = \frac{25400}{CN} - 254 \quad (5.10)$$

Q is the surface runoff, mm; P is the precipitation, mm; S is the surface potential storage, mm; CN is the composite parameter of the basin features, which is related to soil properties, pre-humidity, and land use. The ultimate runoff coefficient is the ratio of Q to P. The parameters related to the rainfall runoff coefficient for the research area are shown in Figure 5.2. According to the references (Cui, 2011, Li et al., 2021b), pollution concentration, stream entry coefficient and runoff correction coefficient were used.

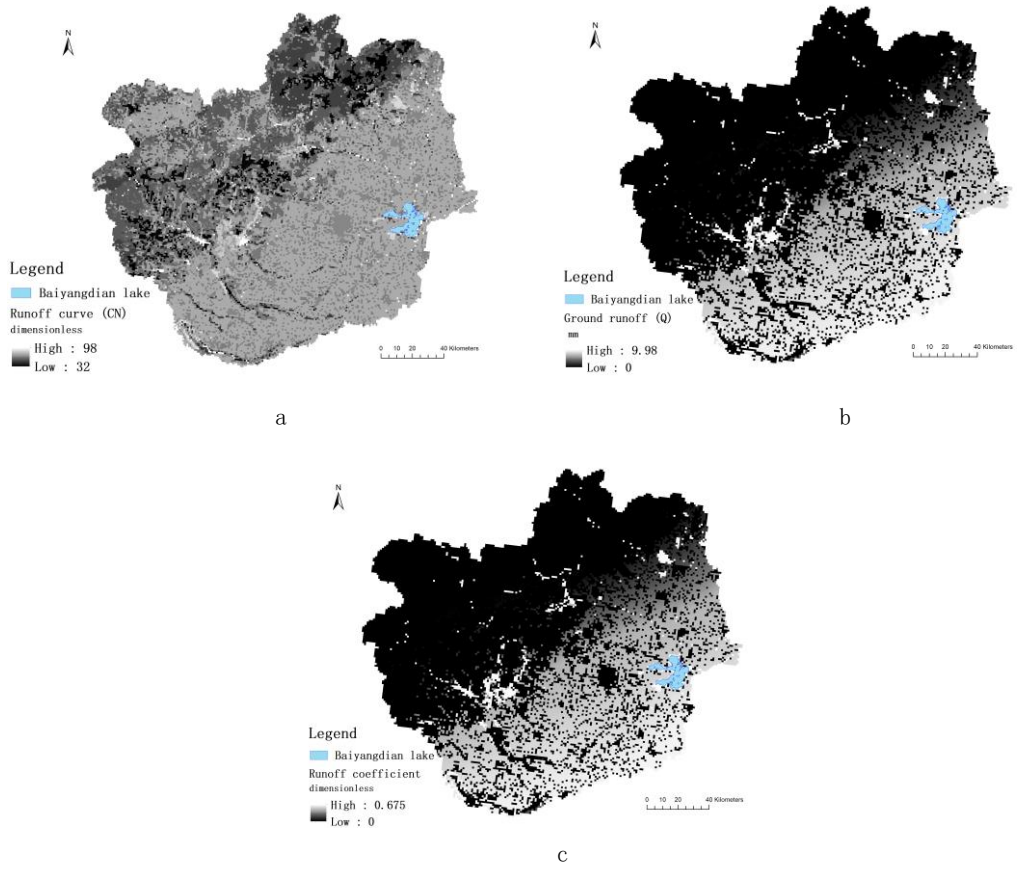

Figure 5.2 Rainfall runoff coefficient parameters in the research area.

5.2.4 Point source load

Point source pollution includes sewage from industrial enterprises and sewage treatment plants. Industrial wastewater pollution load is calculated by the following equation 5.11:

$$W_P = \lambda_P \times \left(\sum_{k=1}^{K}\sum_{n=1}^{N}(Q_{IW,k,n} \times C_{IW,k,n} \times 10^{-6}) + p_{LW} \times P_{LW} \times (1 - f_{LW}) \times 10^{-3}\right) \quad (5.11)$$

Wp is the load of point source pollution, t/a; λ_P is the pollution of point source; K and N are the respective quantities of industrial pollution source industry types and industrial enterprises; $Q_{IW,k,n}$ is the wastewater emission volume k of industry n enterprises, m^3; $C_{IW,k,n}$ is respective average pollutant discharge concentration, mg/L; p_{LW} is the production factor of per capita pollution, kg/(per person·a); P_{LW} is the population of the catchment area; f_{LW} is the wastewater treatment rate. The respective data sources for the collection are shown in Table S1 in the Supplementary Material.

5.3 Non-Point Source Pollution Load Accounting

Many of the inflowing rivers into the lake of the Baiyangdian Lake are dry or diverted. Therefore, there dry rivers cannot deliver pollutants to the lake, and the pollution load produced by

the entire basin is not equivalent to the load to the lake. To calculate the actual impact of rivers on Baiyangdian Lake, we need to further refine the sub-basin according to the actual situation, and the sub-basins corresponding to rivers with actual inflow to the lake are selected as the object of study to investigate the non-point source loads.

Rivers originating from west mountain areas are stopped and stored in reservoirs. This greatly reduced the amounts of particulate pollutants produced by erosion carried to downstream. Because the incoming water is stopped, the connection between the upstream and downstream of the river is interrupted, resulting in no pollutants from the upstream basin entering the water bodies downstream directly. Therefore, the sub-basin of the Baiyangdian Lake was segmented using the ArcSWAT model on basis of the digital elevation data and actual river conditions.

5.3.1 Agriculture and rural non-point source pollution load accounting

Domestic waste and domestic water are the main sources of non-point source pollution in rural areas (Zhang et al., 2022). The dominant factor in the intensity of nonpoint source pollution comes from the residential population. The pollution intensity is much higher in densely populated areas. The rural domestic nonpoint source pollution load based on relevant information is computed as shown in Figure 5.3.

As shown in Figure 5.4. An important source of agricultural nonpoint source pollution comes from livestock and poultry, which emit large amounts of nitrogen, phosphorus, and organic pollutants into the environment (Li et al., 2021b). Non-point source loads from livestock farming were derived based on the approximate number of livestock in each region and county of the Baiyangdian watershed, taking into account feeding methods, manure and urine disposal methods.

Excessive use of fertilizers and pesticides during agricultural cultivation leads to residues on the surface of cultivated land and non-point source pollution through surface runoff into natural water bodies (Ouyang et al., 2018). The main determinants of planting nonpoint source pollutant loads come from planting crops and management practices. The intensity of surface source pollution is computed from land use data and the area planted is acquired from the modified values to get the yield coefficient, as shown in Figure 5.5.

Figure 5.6 shows the total non-point source pollution load in rural areas obtained for the three types of rural living, livestock farming and agricultural cultivation.

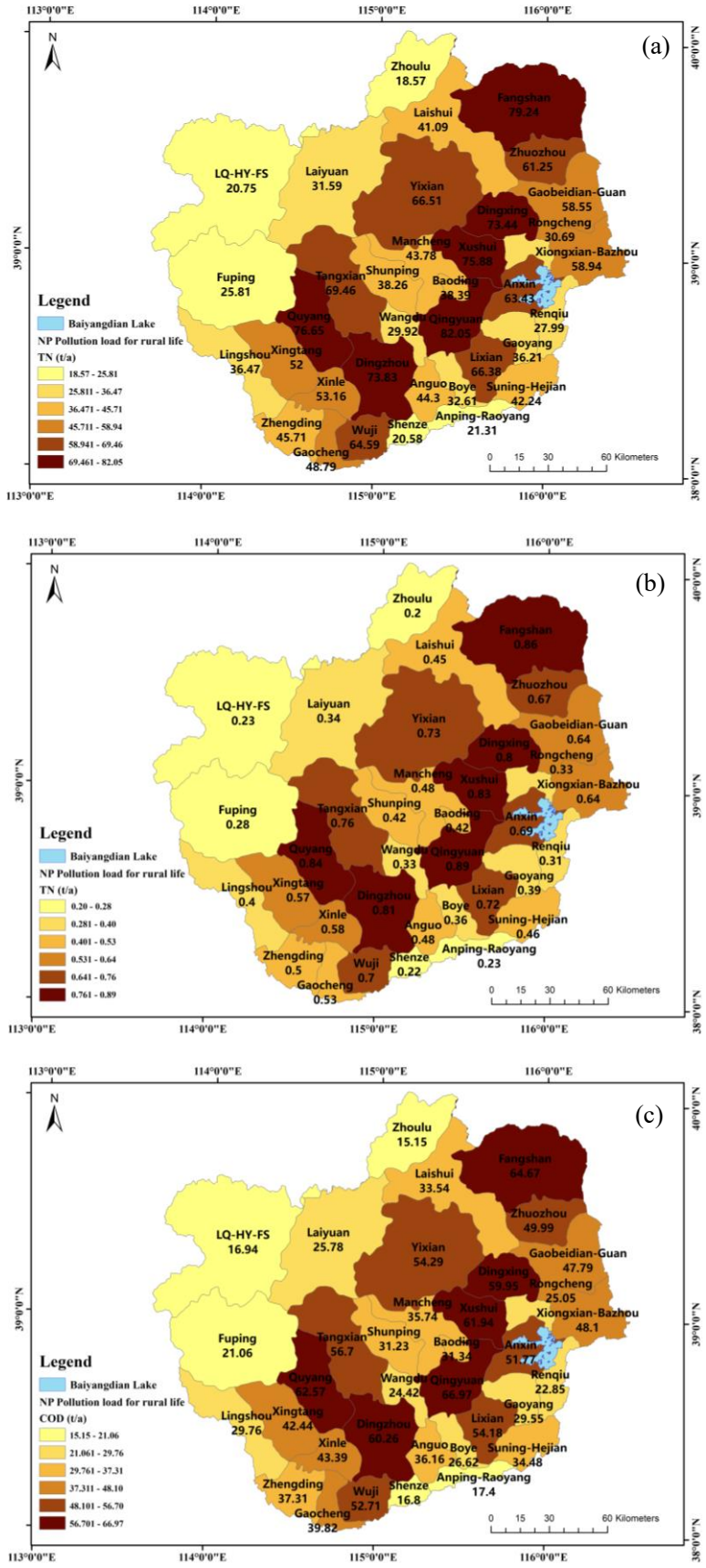

Figure 5.3 Baiyangdian watershed districts and counties rural living non-point source pollution load. (a: TN load; b: TP load; c: COD load).

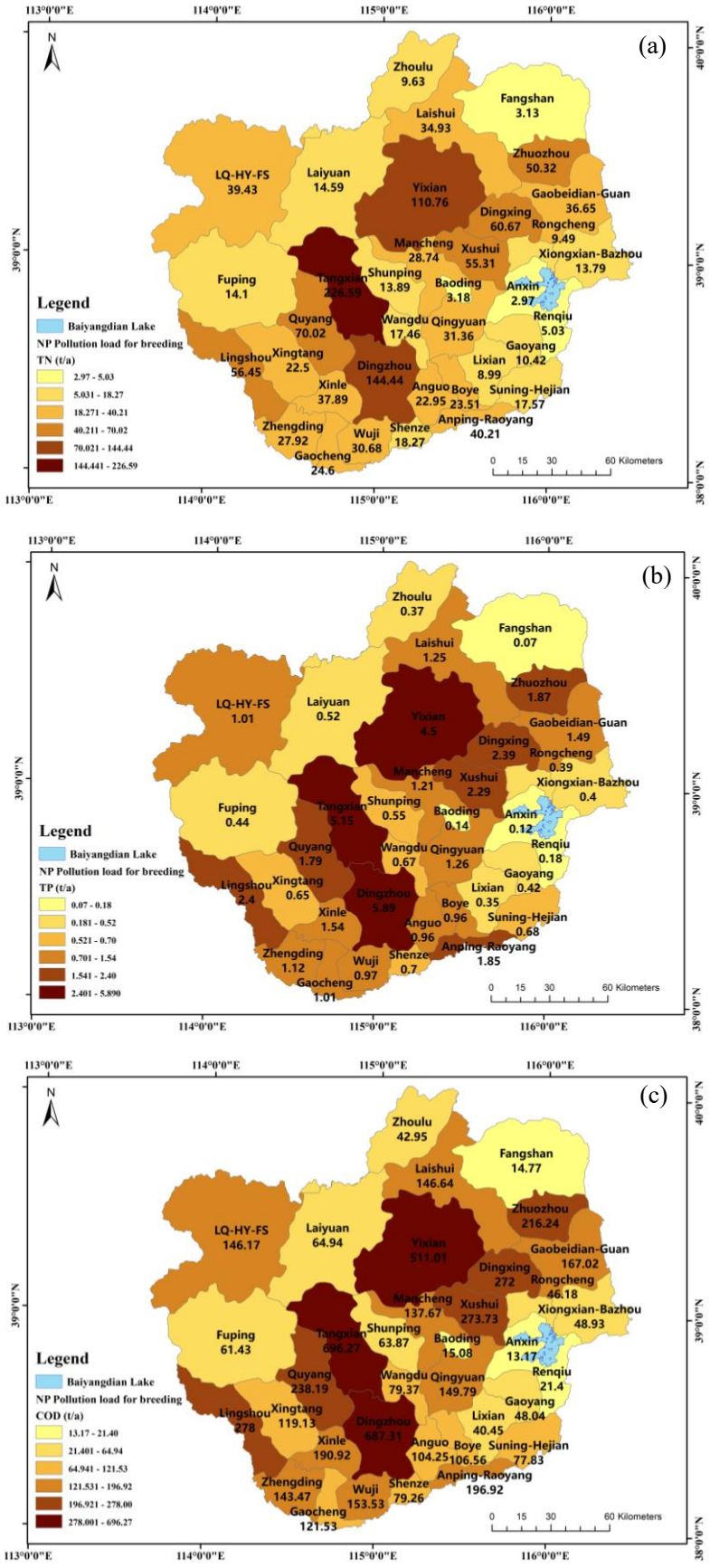

Figure 5.4 Baiyangdian watershed districts and counties livestock and poultry non-point source pollution load. (a: TN; b: TP; c: COD load)

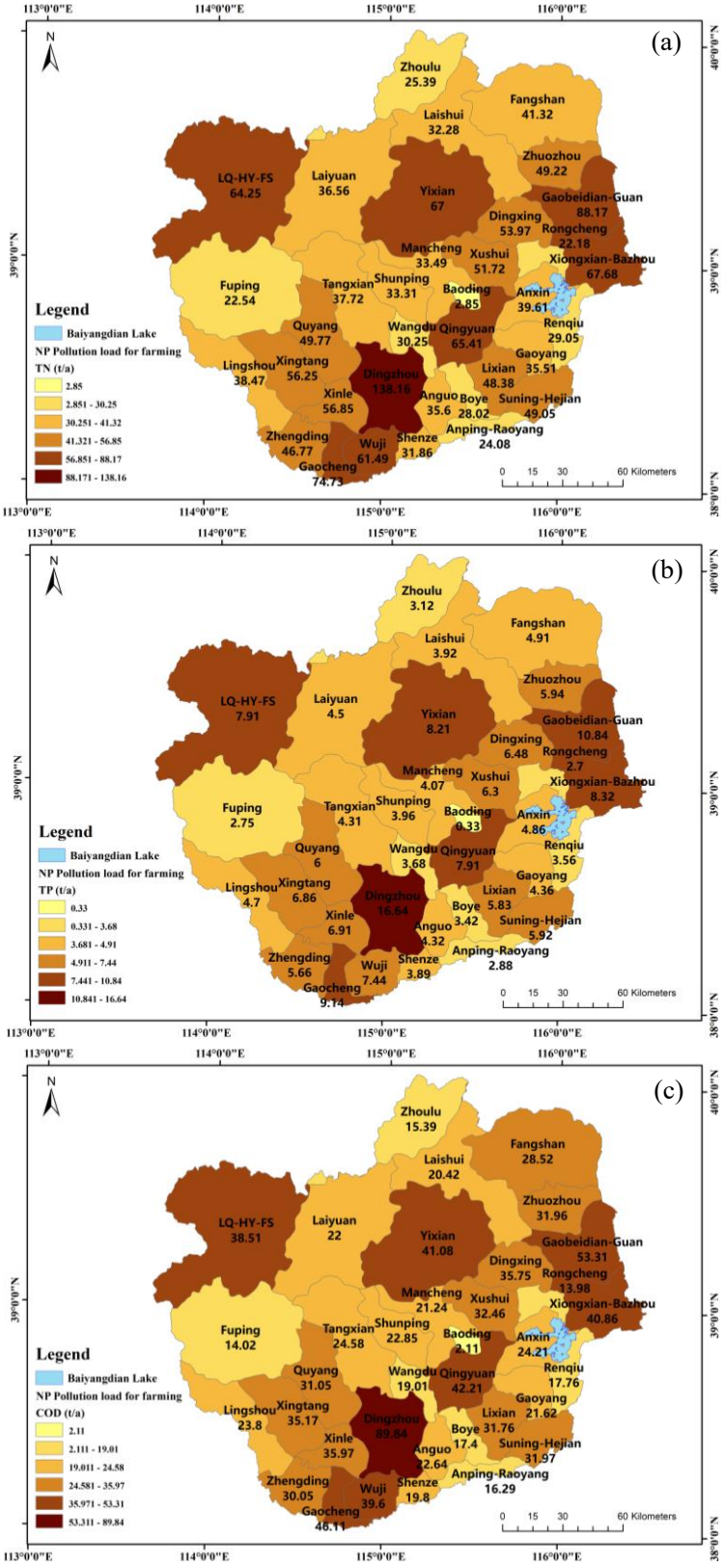

Figure 5.5 Baiyangdian watershed districts and counties agricultural cultivation non-point source pollution load. (a: Total nitrogen load; b: Total phosphorus load; c: COD load)

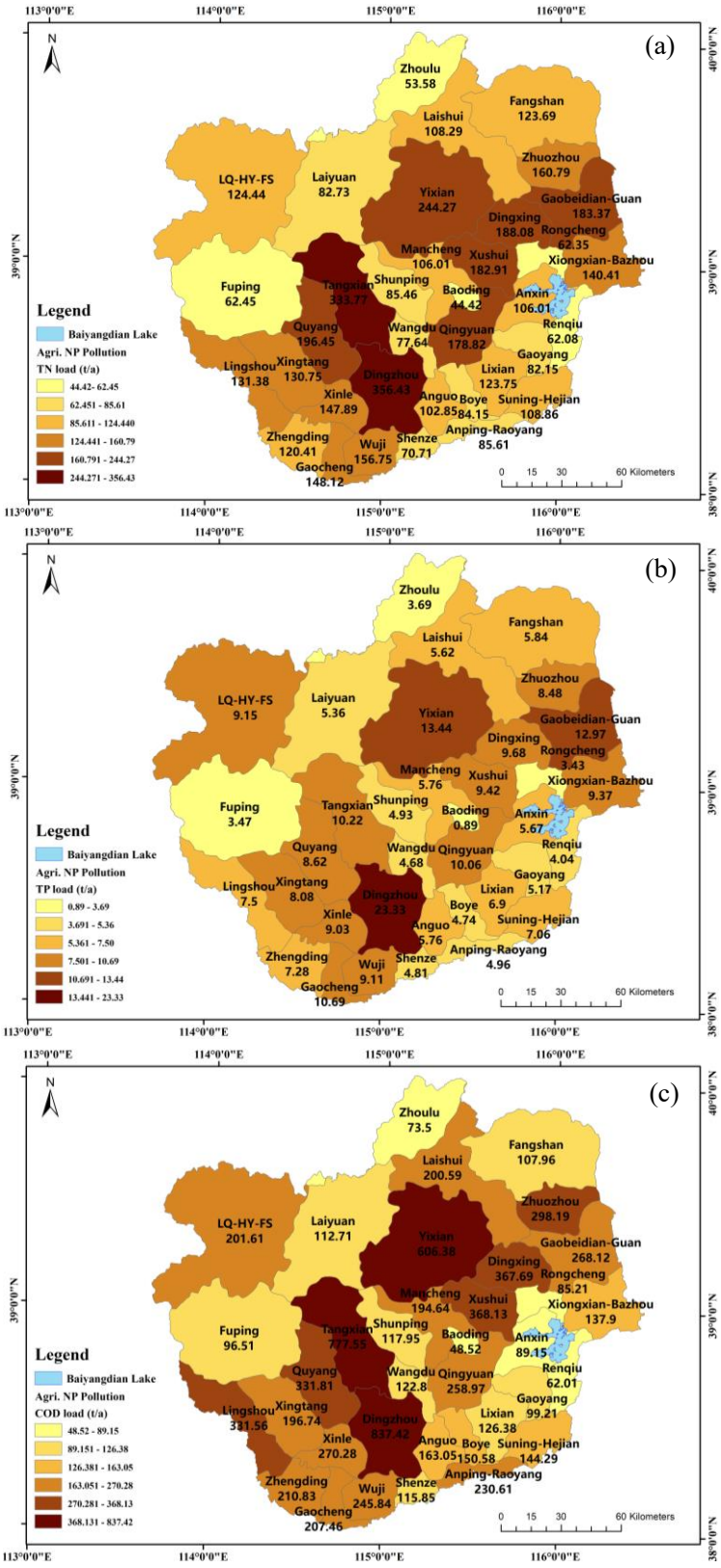

Figure 5.6 Total agricultural and rural surface source pollution in Baiyangdian watershed districts and counties. (a: TN load; b: TP load; c: COD load)

On the whole, Dingzhou City, Tang County and Yi County are the three areas with the highest agricultural and rural surface loads in the Baiyangdian Basin, accounting for 19.6%, 17.5% and 26.9% of the total agricultural and rural surface loads. Based on the Baoding Economic and Statistical Yearbook (2020), the total area under crop cultivation in Dingzhou City in 2019 was as high as 159,000 hectares, including 117,000 hectares of grain and 18,000 hectares of vegetables; poultry production was 6.53 × 106 heads, of which 870,000 were live pigs (Bureau, 2018). In addition, Dingzhou city has the fourth highest rural population and the highest pollution production among all cities and counties in the Baiyangdian basin. Likewise, Tang and Yi counties have large-scale agricultural and cultivation activities that lead to high non-point pollution sources.

The Baiyangdian watershed consists of all the rivers that naturally flow into Baiyangdian. Due to the breakage of many rivers that no longer flow into Baiyangdian, the pollution load generated by the whole watershed cannot be equivalent to the sediment load, so Baiyangdian was divided into four sub-basins (Fuxie River, Xiaoyi River, Baigou Yin River and Bao River). The corresponding potential sedimentation of agricultural and rural non-point source pollution was adjusted, with TN load, TP load and COD load of 1073.82 t, 102.72 t and 3349.33 t, respectively (Table 5.2). In particular, the contribution of livestock and poultry farming to COD was over 70%. The contribution of agricultural cultivation to TP was 73.12%; the contribution of the three non-point sources to TN was almost equal, with a slightly higher contribution of 37.48% from rural living.

Table 5.2 Non-point source load （agricultural and rural areas）

Pollutant	Amount (t)	Percentage (%)		
		Rural life	Livestock and poultry breeding	Agricultural cultivation
Total nitrogen	1073.82	37.48	29.59	31.93
Total phosphorus	102.72	6.93	17.95	73.12
COD	3349.33	17.15	70.16	11.69

Livestock farming has a remarkable influence on COD pollutants in lower water bodies, so effective treatment of aquaculture pollution should be considered, with special emphasis on the management of small-scale dispersed farming to control the discharge of organic pollutants. In addition, excessive amounts of agrochemicals applied in agricultural cultivation lead to large amounts of nutrients remaining on the surface of agricultural soils and flowing into natural water bodies through rainfall, leading to eutrophication. Consequently, the adoption of science-based methods of fertilizer application and intensified farm management practices have become desirable

measures to reduce nonpoint source pollution in agricultural cultivation.

5.3.2 Pollutant load accounting for adsorption state by surface runoff scouring

Rainfall leads not only to the migration of surface water-soluble pollutants into water bodies, it also leads to the entry of soil particles into water bodies through erosion of topsoil (Wischmeier and Smith, 1978). Pollutants like N, P and organic matter attached to the surface of soil particles have become an important source of non-point source pollution in water bodies. The surface soils of agricultural fields in the Baiyangdian watershed are usually contaminated with adsorbed organic matter and inorganic N and P due to long-term tillage. Uncovered outdoor farmland soils are easily washed away by precipitation, which leads to surface soil erosion. Thus, the pollution load due to soil erosion should be considered when calculating the non-point source pollution caused by rainfall in the Baiyangdian watershed.

To distinguish dissolved pollutants in rainfall runoff from particulate pollutants carried by soil erosion, the TN and TP contents of the surface soil in the soil database are used to deduct the water-soluble adsorbed N and adsorbed P contents, and the resulting difference is used as the particulate pollutant content; for the background values of soil COD pollutants, the soil organic matter content is used directly as the particulate pollutant content.

According to equation (5.2) for each parameter, the adsorbed pollutant load from soil erosion in Baiyangdian watershed is shown in Figure 5.7.

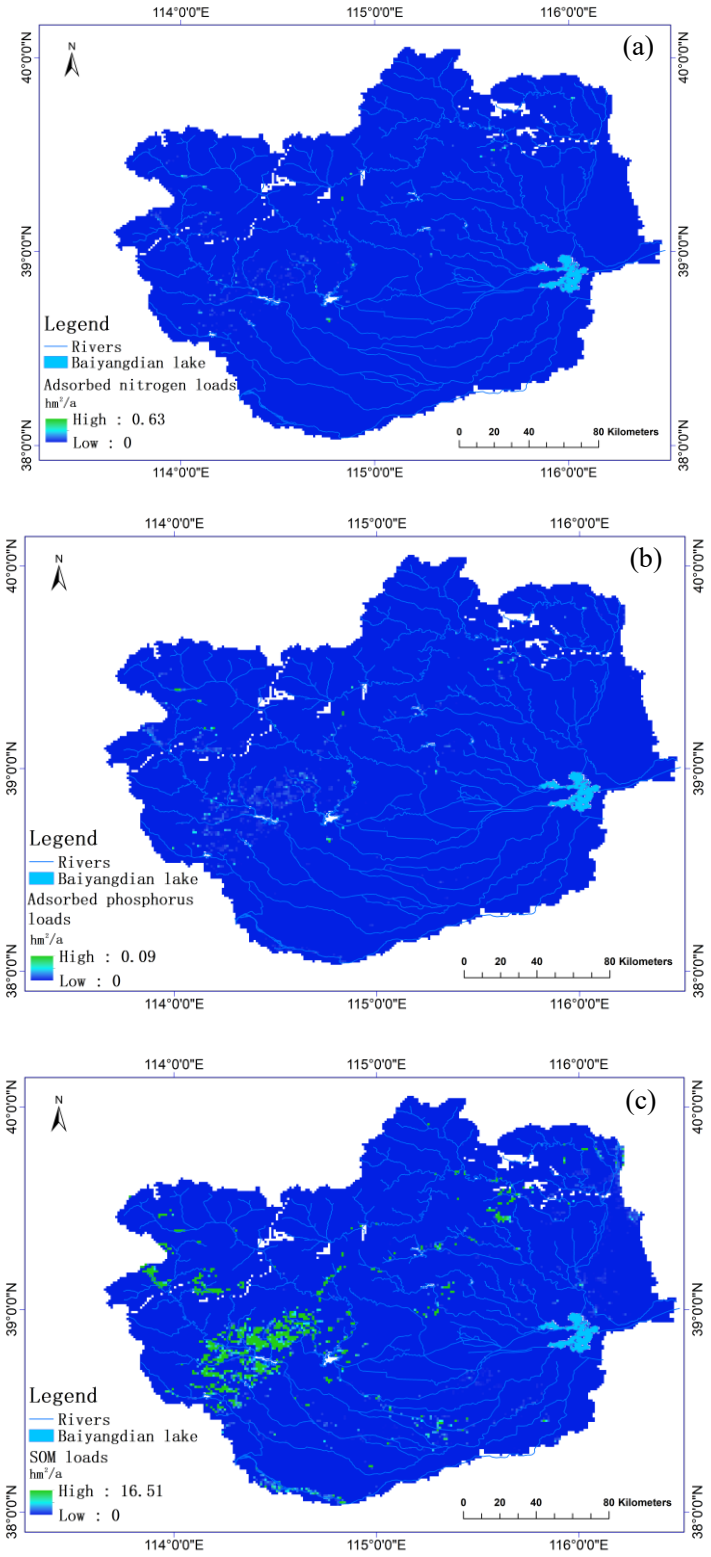

Figure 5.7 Soil erosion sorption pollutant load in the research area. a: Soil organic matter; b: Adsorbed N; c: Adsorbed P.

The results show that the overall non-point source pollution load due to soil erosion in the entire watershed was rather small. The high load areas were focused in the Taihang Mountains,

particularly in the upstream of the Lirong River and the Tang River. The variation among the three pollutant loads of COD, TN and TP is primarily caused by the different contents in the surface soil, generally in the order of TN > TP > COD. Non-point source pollution loads from soil erosion were also adjusted according to the sub-basins of the four major rivers (Fu River, Xiaoyi River, Baigouyin River, Bao River). The results show that the total adsorption load from soil erosion were 94.1 t for TN, 20.8 t for TP and 1262.9 t for COD, as shown in Table 5.3.

Table 5.3 Adsorbed pollutant loads from soil erosion in the corresponding sub-basins of the Baiyangdian watershed

River into Baiyangdian Lake	Load of adsorbed pollutants generated by soil erosion（t/a）		
	COD	TN	TP
Bao River	19.9	1.0	0.6
Fu River	585.9	2 6.5	15.3
Baigouyinyin River	495.4	56.4	2.6
Xiaoyi River	161.7	9.2	2.2
Total	1262.9	94.1	20.8

5.3.3 Pollution load accounting for dissolved pollutants by surface runoff

The primary driver of nonpoint source pollution is surface runoff, and rainfall causes differential generation and transport of nonpoint source loads in agricultural and rural areas; therefore, the agricultural and rural nonpoint source loads computed using the ECM model are inherently loads of dissolved pollutants caused by surface runoff. The dissolved pollutant load calculations performed here are for pollution loads generated by land uses other than agricultural cropland that contribute to surface pollution from precipitation. Combining the mean yield per precipitation and the concentration of pollutants in the production streams of different regions, the final distribution of dissolved pollutant loads due to rainfall erosion per unit of land use other than agricultural cultivation land was computed (Figure 5.8).

Non-point source pollution loads from precipitation in the four inflow river basins were acquired (Table 5.4) based on the sub-basin divisions in Figure 5.7. The results show that the rainfall pollution loads have less impact on pollutant loads than agricultural and rural areas.

Table 5.4 Pollution load from non-point sources produced by precipitation

Pollutant	Fu River	Baigouyinyin River	Bao River	Xiaoyi River	Surrounding	Total
TN (t/a)	37.80	3 6.88	1.72	2 6.20	2.31	10 6.91
TP(t/a)	1.34	1.10	0.04	0.93	0.05	3.47
COD (t/a)	201.74	199.43	7.54	144.14	11.09	564.94

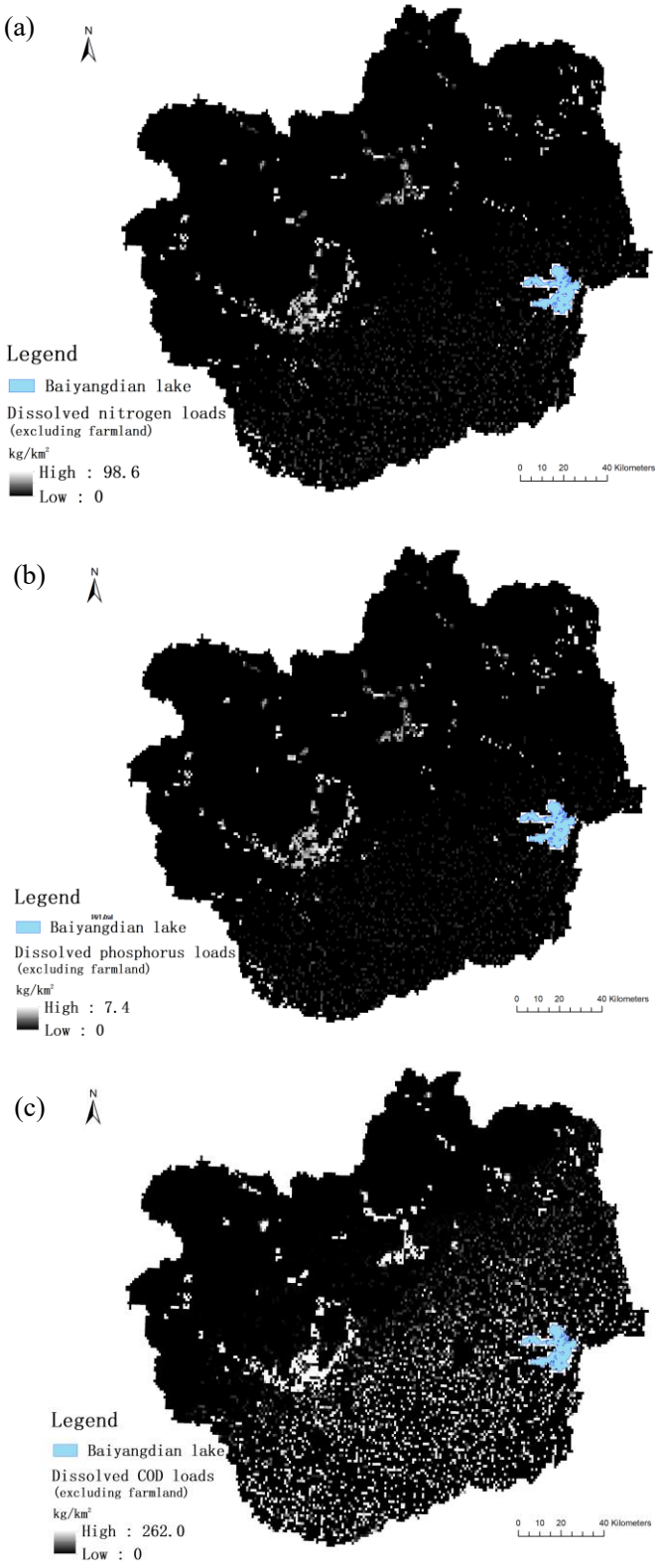

Figure 5.8 Dissolved pollutants loads per unit area of rainfall production streams in the study area. a, TN load; b, TP load; c, COD load.

5.4 Point source pollution load accounting

The point source loads are based entirely on the statistics of the survey of industrial enterprises and wastewater treatment plants in the Baiyangdian watershed. The survey of 682 enterprises in each county of the basin was conducted through a comprehensive compilation of environmental disclosure information of 2021 key discharge units by the Baoding Ecological Environment Bureau. Among these major enterprises, about one-third of them had no significant sewage emissions. As a result, 254 point source emissions were gained. in 2020, the total pollutant emissions computed from point sources were 271.01 t for TN, 21.49 t for TP, and 1302.56 t for COD. since the point source pollution loads were obtained from the collected data, it is important to note that the stability and precision of the released data have a significant impact on the computed results.

5.5 Pollution load contribution ratio

The calculation results of the pollution load of TN, TP and COD in the whole basin were shown in Table 19. The primary pollution source of all three pollutants was non-point source pollution from agriculture and rural cities, accounting for above 50% of the total load of all pollutant sources. Of which, the contribution to the load of TP accounts for 73.37%. In the case of non-point source pollution, as well as runoff and soil erosion, the contribution of runoff to the overall load of all three pollutants was small. The contribution of soil erosion to TN is higher at 22.95%; the contribution of point source pollution to COD is larger at 22.33%; the contribution of TP is more than that of runoff pollution and soil erosion pollution.

Agricultural and rural non-point source pollution and soil erosion are mainly caused by rainfall flushing of surface pollutants into rivers. Usually, due to low vegetation cover and sparse crops, farmland is vulnerable to soil erosion. Therefore, both dissolved pollutants and soil particles from rainfall washing from the soil surface of agricultural fields are part of agricultural surface source pollution, and the non-point sources of TN into the lake are primarily from agriculture and erosion. For the same reason, non-point sources of agriculture also make a considerable contribution to TP. According to the analysis in section 4.2.1, the large amount of phosphorus input into the environment from agricultural production and livelihood activities, and irrational management practices resulting in inefficient nutrient use, make agricultural farming activities a large underlying

source of phosphorus pollution.

Agricultural and rural non-point sources and wastewater from industry are the primary sources of COD. According to section 4.2.1, livestock and poultry farming in agricultural activities is the main source of COD load, accounting for about 73% of the COD load from agricultural non-point sources, which means about 40% of the total COD load into the lake. Therefore, the reduction of COD load into the sediment must focus on the management of discharges from farming activities. Majority of industrial activities generate effluents or wastewater with higher COD concentrations, and the COD concentration of wastewater treatment plant discharge is also usually higher than the surface runoff after rainfall. So, the COD load is smaller than that of agricultural non-point sources (rural domestic sewage, farming effluent, rainfall flushing, etc.) because that the scale of pollution production is smaller than that of agricultural non-point sources.

Table 5.5 Load into the lake of TN、TP and COD

Pollutant Sources	TN		TP		COD	
	Total（t）	Ratio（%）	Total（t）	Ratio（%）	Total（t）	Ratio（%）
Point source	271.01	14.37	21.49	15.35	1302.56	22.33
Agriculture and rural areas	1073.82	56.96	102.72	73.37	3349.33	57.43
Runoff	107.91	5.72	3.47	2.48	564.94	9.69
Soil erosion	432.61	22.95	12.33	8.81	615.52	10.55
Total	1885.35		140.01		5832.35	
Total pollutants of rivers in 2020	1272.5	-	36.5	-	5082.3	-

5.6 Conclusions

The pollutant sources in the Baiyangdian Lake watershed are diverse and complex. The loads of TN, TP and COD from point and non-point sources were calculated in this chapter. The non-point sources include agricultural and rural non-point source loads, dissolved pollutants due to surface runoff and adsorbed pollutant loads. The primary pollution sources of all three pollutants of TN, TP and COD are agricultural and rural surface sources, and the contribution of the loads exceeds 50%. In addition, the contribution of runoff to the total load of all three pollutants was low, and The contribution of soil erosion to TN is greater. The contribution of point source pollution to COD is greater, reaching 22.33%, while the contribution of TP exceeds that of runoff pollution and soil erosion pollution.

The pollution load model used in this chapter has been widely used in the world, and it has been proved by many scholars to be more accurate and reasonable in calculating the pollution load in the basin. Comparing the model calculation results with the load flux calculation results (Table 4.7), it can be found that the calculation results of the two methods are basically within the same order of magnitude, but there is still a considerable gap. The pollution source loads of the three pollutants calculated by the pollution load model are higher than the total amount of pollutants transported by the rivers entering the lake in 2020 (Table 5.5). The TN load is about 1.5 times the amount of that inflowing into the lake, the TP load is about 3.8 times the amount of the lake, and the COD load is about 1.15 times the amount of the lake. An important reason for the gap between load and sediment inflow is pollutant lost during migration and transport. During the process from pollution generation to final arrival at the river inlet, pollutants will gradually separate from the water body through water environment self-purification such as decomposition, adsorption and sedimentation. For example, NH_4-N and small molecule organics can be utilized or mineralized through microbial dominated biochemical processes. The solubility of P containing pollutants is generally lower than that of N containing pollutants, and the proportion of adsorbed P containing pollutants adsorbed on particles in the water body is large in TP. The lower reaches of the four rivers are all in plain area, with meandering channels, small drop and slow flow velocity, resulting in short particle transport distance. The lower reaches of these rivers are developed intensely, and there are multiple gates along the way, which further aggravate the precipitation of particulate matter. At the same time, compared with TN, dissolved P is more easily adsorbed by particles and surface sediments and separated from the water environment. All these factors lead to the high rate of TP loss during long-distance deliveries, so from the calculation results of load and flux, only about one-third of TP finally enters the lake area.

The gap between the model calculation and the actual situation cannot be ignored. The pollution load model used in this paper is based on empirical parameters. On the one hand, accurate parameters need to be measured over a long time period. On the other hand, even if the research object is limited to sub basins, there are spatial differences in its internal model parameters. The use of unified parameter generalization will inevitably lead to a certain gap between model calculation and practice. At the same time, the model itself also has limitations. For example, the socio-

economic statistical data used by the ECM model are based on administrative divisions, and there is a large gap between the model and the natural river basin. Only relying on the river inflow coefficient to describe the process of pollutants entering the natural water body will lead to distortion of the model calculation results.

If the automatic monitoring technology in this study can be used to realize the synchronous daily monitoring of water quality and quantity, the time average dispersion error caused by the asynchronous data frequency can be avoided in the pollutant flux calculation process, and the deviation caused by the time domain discrete sampling of continuous flux can be reduced, thus significantly improving the accuracy of flux calculation. Especially for rivers strongly disturbed by human activities, the pollutant flux into the river fluctuates greatly and the regularity is weak. The results obtained by the flux formula or periodic function regression obtained only by dispersing the integral formula are usually far from the reality.

The automatic monitoring technology can be used to improve the time resolution of the monitoring data. Additionally, the spatial resolution of the calculation process can also be improved through the monitoring of multiple sections. If representative sections are selected for monitoring, the model parameters can be modified according to the river environmental characteristics represented by each section. For example, a separate pollutant inflow coefficient is assigned to each section of the river according to the vegetation type, slope, river hardening, etc. on both banks of the river. It can also determine the attenuation rate of pollutants in the process of being transported by the river and infer the main control factors affecting the attenuation of pollutants. In addition, the sections of the rivers are taken as the nodes to monitor in different areas, which can distinguish the main pollution sources, the amount of pollution entering the river and the change law of each section of the river. It is helpful for the detailed study of the pollution sources entering the river around each section of the river. These improvements can significantly improve the accuracy of the model. It can also help to identify the key pollution producing areas, achieving fine management of rivers and even river basins, and improve management efficiency.

Chapter 6 General conclusions and future works

6.1 General conclusions

Monitoring the parameters of water quality is very important for protecting the water ecological environment. Application of Internet of Things (IoT), Big Data, Cloud Computing, and other latest information technologies helps to achieve environmental monitoring to meet the needs of environmental assessment and management. Real-time monitoring can effectively ensure the safety of water environment protection and water ecological restoration.

In this study, the monitoring system were developed including the NB-IoT in-situ monitoring technology and the patrol monitoring by using the multi-sensor information fusion technology. Using advanced NB-IoT and cloud computing technology, the continuous and real-time monitoring on rivers and lakes were realized and it was applied in monitoring the water level (for calculating water quantity) and water quality of the inflowing river (Fu River) to the Baiyangidan Lake. Additionally, the fluxes of pollutant from the four major inflowing rivers (Xiaoyi River, Baigouyin River, Bao River and Fu River) were computed by using the monitoring data by the government and the LOADESE model were compared. Then the flux of water and pollutants of TN, TP of Fu River by using the above two methods were compared and the advantage of the automatic monitoring system was demonstrated. On the watershed scale, the export efficient method was applied to calculate the pollutants load including the point and non-point source pollutants. As the pollutants flowing into the rivers and the lakes were affected by different factors, the gags between the flux of pollutants into the rivers and the load of pollutants from the watershed were discussed.

The in-situ monitoring system based on the NB-IoT technology were developed in this study. The in-situ automatic monitoring equipment can monitor water quality information continuously and can conduct real-time monitoring. The present automatic monitoring system has been applied in field observation of hydrologic data for water and soil. However, only the data for the fixd point can be obtained, and it is difficult to obtain the spatial status for environment of water and soil due to the expensive cost of the in-situ equipment. Therefore, the patrol monitoring technology provides a view and a practical way to realize the spatial observation for water environments. To satisfy the long-time autonomous work of unmanned vehicle, the reliability of autonomous navigation is

enhanced by using robust multi-sensors information fusion algorithm, which provides a continuous and dynamic patrolling means for observation in large areas. Though the application of patrol monitoring technology was not conducted in the study area, it gives us a view to construct a multiple monitoring system combing the in-situ automatic monitoring technology. The comparison of flux of pollutants in Fu River estimated by LOADEST model with data observed in hydrological stations and by using the in-situ monitoring system based on the NB-IoT technology suggested that the automatic systems provide a high effective method. The estimation of the sources and loads of the pollutants in the watershed scale with the export coefficient models showed the large load in the watershed while there was only small part of pollutants inflowing into the rivers and lakes. In a word, the methods based on 5G Internet of Things (NB-IoT), high precision navigation technology for unmanned inspection robot and LOADEST model was established to provid practical technical solutions for.regional environment monitoring.

The main conclusions of this study are as following:

(1) A 5G (NB-IoT) based environmental monitoring system was designed, deployed and applied in the Baiyangdian Lake and one of its inflowing rivers (Fu River). The system consists of three major parts: sensing terminal, NB-IoT network communication unit and software monitoring platform. The online monitoring software system is designed and developed in PHP programming language, and the retransmission mechanism of NB-IoT network is designed in the network to guarantee the high-quality data transmission, and the data collection and reporting monitoring, storage and real-time display are realized in the cloud platform. The entire system was applied in water environment monitoring of Baiyangdian Lake and the inflowing river of Fu River, which improved monitoring efficiency and reduced costs. This monitoring system resolved the problem of the discontinuous retransmission of monitoring data and the equipment can be established in field, and it decreased the cost by reducing power consumption during the period of the monitoring.

(2) The nonlinear filtering problem of GNSS/INS under non-Gaussian noise is studied, and a modified maximum correlation entropy volume Kalman filter (MCCKF) based on the resampling free sigma point update framework (SUF) is proposed. It turned out that RMCCKF has better performance than MCCKF, which reduces the root mean square error of the land robot heading from 1.77 degrees to 0.26 degrees, in addition to improving the localization results. This method improves

the navigation accuracy through software algorithm, reduces the hardware cost, and ensures the highly reliable automatic navigation of the robot, which can be applied to unmanned ships for patrol monitoring. Unmanned ship patrol monitoring makes up for the limitations of the in-situ monitoring in space, it can realize the dynamic monitoring of the whole region.

(3) The flux of major pollutants of TN, TP and COD in major four rivers inflow into the Baiyangdian Lake were estimated by using the LOADEST model and the hydrological data by the government. Additionally, the results of Fu River were compared with those calculated by using the in-situ automatic monitoring data. The sources and loads of the major pollutants in the whole Baiyangdian Lake basin were also conducted. Based on monitoring data, literature and relevant historical information, the pollutant fluxes of Baiyangdian Lake and major inflowing rivers at different periods were estimated, and the contribution ratios of different inflowing rivers to the lake were derived. The potential pollutant source, pollutant loads and the proportion of contribution to pollutants in Baiyangdian Lake basin were analyzed. Additionally, the pollutant loads from various sources were quantified and the spatial and temporal distribution of pollutants in the lake was analyzed. Although the monitoring frequency of the monitoring data used in this study is relatively low, if we can achieve simultaneous daily monitoring of water quality and quantity using the monitoring techniques in this study, it will significantly improve the accuracy of pollutant flux calculation and the rapid calculation of pollutant fluxes can be achieved. As a result, the analyzing method can be integrated into a cloud platform to achieve automatic environmental assessment and monitoring in practical applications.

6.2 Future works

In this research, the monitoring system including NB-IoT based in-situ monitoring equipment, and monitoring platform have been developed and applied. Although the effectiveness of various parts of the system has been verified, the types and amount of data collected are relatively small at present. The future work will be focused on following aspects:

(1) Complete the compatibility work for more kinds of in-situ monitoring sensors.

(2) Develop efficient and low-cost monitoring UAVs and unmanned ships.

(3) Develop cloud computing function based on assessment model and artificial intelligence to realize monitoring of the whole process.

References

ADU, J. T. & KUMARASAMY, M. V. 2018. Assessing Non-Point Source Pollution Models: a Review. *Polish Journal of Environmental Studies,* 27.

ARASARATNAM, I. & HAYKIN, S. 2009. Cubature kalman filters. *IEEE Transactions on automatic control,* 54**,** 1254-1269.

ARULAMPALAM, M. S., MASKELL, S., GORDON, N. & CLAPP, T. 2002. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing,* 50**,** 174-188.

BOULANGE, J., WATANABE, H., INAO, K., IWAFUNE, T., ZHANG, M., LUO, Y. & ARNOLD, J. 2014. Development and validation of a basin scale model PCPF-1@ SWAT for simulating fate and transport of rice pesticides. *Journal of Hydrology,* 517**,** 146-156.

BUREAU, H. S. 2018. Hebei Economic Yearbook. China Statistics Press: Beijing, China.

CHANG, G., CHEN, C., ZHANG, Q. & ZHANG, S. 2021. Variational Bayesian adaptation of process noise covariance matrix in Kalman filtering. *Journal of the Franklin Institute*.

CHANG, L., LI, K. & HU, B. 2015. Huber's M-estimation-based process uncertainty robust filter for integrated INS/GPS. *IEEE Sensors Journal,* 15**,** 3367-3374.

CHEN, B. & GAN, Z. 2016. Research on Commercial Value and Networking Solution of NB-loT Network. *Mobile Communications,* 40**,** 42-46+52.

CHEN, B., LIU, X., ZHAO, H. & PRíNCIPE, J. 2015. Maximum Correntropy Kalman Filter.

CHEN, H., TENG, Y. & WANG, J. 2013. Load estimation and source apportionment of non-point source nitrogen and phosphorus based on integrated application of SLURP model, ECM, and RUSLE: a case study in the Jinjiang River, China. *Environmental monitoring and assessment,* 185**,** 2009-2021.

CHEN, L., XU, J., WANG, G., LIU, H., ZHAI, L., LI, S., SUN, C. & SHEN, Z. 2018a. Influence of rainfall data scarcity on non-point source pollution prediction: Implications for physically based models. *Journal of Hydrology,* 562**,** 1-16.

CHEN, X.-K., LIU, X.-B., PENG, W.-Q., DONG, F., HUANG, Z.-H., FENG, S.-X. & WANG, R.-N. 2018b. Estimation of and Control Strategies for Pollution Loads from Non-point Sources in the Chenghai Basin. *Huan Jing ke Xue= Huanjing Kexue,* 39**,** 77-88.

CHEN, Y. 2014. *Research and Implementation of LTE-A System Layer Relay Access Link High-Level Protocol.* Guangdong University of Technology.

CHEN, Y., XU, C.-Y., CHEN, X., XU, Y., YIN, Y., GAO, L. & LIU, M. 2019. Uncertainty in simulation of land-use change impacts on catchment runoff with multi-timescales based on the comparison of the HSPF and SWAT models. *Journal of Hydrology,* 573**,** 486-500.

CHENG, Y., TIAN, Y. & CRASSIDIS, J. L. Extension of the sparse grid quadrature filter. 17th International Conference on Information Fusion (FUSION), 2014.

COHN, T. A., DELONG, L. L., GILROY, E. J., HIRSCH, R. M. & WELLS, D. K. 1989. Estimating constituent loads. *Water resources research,* 25**,** 937-942.

COHN, T. A., GILROY, E. J. & BAIER, W. G. 1992. *Estimating fluvial transport of trace constituents using a regression model with data subject to censoring*, The Survey.

CUI, B., CHEN, X. & TANG, X. 2017. Improved Cubature Kalman Filter for GNSS/INS Based on Transformation of Posterior Sigma-Points Error. *IEEE Transactions on Signal Processing,* PP**,**

1-1.

CUI, B., CHEN, X., TANG, X., HUANG, H. & LIU, X. 2018. Robust cubature Kalman filter for GNSS/INS with missing observations and colored measurement noise. *Isa Transactions,* 72**,** 138-146.

CUI, B., WEI, X., CHEN, X., LI, J. & LI, L. 2019. On Sigma-Point Update of Cubature Kalman Filter for GNSS/INS Under GNSS-Challenged Environment. *Vehicular Technology, IEEE Transactions on*.

CUI, B., WEI, X., CHEN, X. & WANG, A. 2021. Improved high-degree cubature Kalman filter based on resampling-free sigma-point update framework and its application for inertial navigation system-based integrated navigation. *Aerospace Science and Technology,* 117**,** 106905.

CUI, H. 2011. Analysis on effect of agricultural non-point source pollution on the water environment in Baiyangdian Basin. *Xiandai Nongye Keji,* 7**,** 298-230.

DENG, Z., SHI, L., YIN, L., XIA, Y. & HUO, B. 2020. UKF Based on Maximum Correntropy Criterion in the Presence of Both Intermittent Observations and Non-Gaussian Noise. *IEEE Sensors Journal,* 20**,** 7766-7773.

DU, Y., LIU, C., CHEN, K., GU, X., HUANG, W. & WEI, Z. 2018. Occurrence and internal loadings of nitrogen and phosphorus in the sediment of Lake Baiyangdian. *J Lake Sci,* 30**,** 1537-1551.

DUAN, W., TAKARA, K., HE, B., LUO, P., NOVER, D. & YAMASHIKI, Y. 2013. Spatial and temporal trends in estimates of nutrient and suspended sediment loads in the Ishikari River, Japan, 1985 to 2010. *Science of the Total Environment,* 461**,** 499-508.

FANG, S. & GANG, X. 2014. Analysis and Study of 4G Mobile CommunicationTechnology. *Applied Mechanics and Materials,* 3468.

GE, W. & ZHAO, C. 2014. Research on the current situation and development countermeasures of agricultural Internet of Things research and application. *Journal of Agricultural Machinery,* 45**,** 222-230+277.

GUO, H., ZHU, J. & YANG, Y. 2008. Research status and development of technologies for controlling agricultural non-point source pollution. *Transactions of the Chinese Society of Agricultural Engineering,* 24**,** 290-295.

HE, F. 2014. *Environmental Risk and SPARROW Simulation of Non-point Source on Land Use System of Basins in Beijing Mountain Area.* PhD dissertation, Chinese Agricultural University.

HILL, C. M. 2001. PLOAD version 3.0-An Arc View GIS Tool to Calculate Non-point Sources of Pollution in Basin and Stormwater Projects: User's Manual. *US Environmental Protection Agency, Washington, DC*.

HU, G., GAO, B., ZHONG, Y. & GU, C. 2020. Unscented kalman filter with process noise covariance estimation for vehicular ins/gps integration system. *Information Fusion,* 64**,** 194-204.

HUANG, W. 2017. NB-IoT low-rate narrowband IoT communication technology status and development trend. *Electronic Test***,** 58+29.

HUANG, Y., ZHANG, Y., WU, Z., LI, N. & CHAMBERS, J. 2018. A Novel Adaptive Kalman Filter With Inaccurate Process and Measurement Noise Covariance Matrices. *IEEE Transactions on Automatic Control*.

HUANG, Y., ZHANG, Y., WU, Z., LI, N. & CHAMBERS, J. A. 2017. A Novel Robust Student's t-Based Kalman Filter. *IEEE Transactions on Aerospace and Electronic Systems,* PP**,** 1-1.

HUANG, Y. L., ZHANG, Y. G., LI, N. & ZHAO, L. 2016. An improved Gaussian approximate filtering method. *Acta Automatica Sinica*.

JI, Z., ZHANG, Y., ZHANG, H., HUANG, C. & PEI, Y. 2019. Fraction spatial distributions and ecological risk assessment of heavy metals in the sediments of Baiyangdian Lake. *Ecotoxicology and Environmental Safety,* 174**,** 417-428.

JIA, B. & XIN, M. 2019. Data-Driven Enhanced Nonlinear Gaussian Filter. *Circuits and Systems II: Express Briefs, IEEE Transactions on,* PP**,** 1-1.

JOHNES, P. J. 1996. Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach. *Journal of hydrology,* 183**,** 323-349.

JUN, X. & YONGYONG, Z. 2017. Water resource and pollution safeguard for Xiongan New Area construction and its sustainable development. *Bulletin of Chinese Academy of Sciences (Chinese Version),* 32**,** 1199-1205.

KANG, J.-H., PARK, M.-H., HA, S. J. & STENSTROM, M. K. 2021. An empirical modeling approach to predicting pollutant loads and developing cost-effective stormwater treatment strategies for a large urban basin. *Science of The Total Environment,* 760**,** 143388.

KANG, M. & YOO, C. 2020. Application of the SCS–CN method to the Hancheon basin on the volcanic Jeju Island, Korea. *Water,* 12**,** 3350.

KARLGAARD, C. D. & SCHAUB, H. 2007. Huber-based divided difference filtering. *Journal of guidance, control, and dynamics,* 30**,** 885-891.

LI, H., MEDINA, D., VILA-VALLS, J. & CLOSAS, P. 2020a. Robust Variational-Based Kalman Filter for Outlier Rejection With Correlated Measurements. *IEEE Transactions on Signal Processing,* PP**,** 1-1.

LI, L., LIU, Y., WANG, K. & ZHANG, D. 2021a. Simulation of Pollution Load at Basin Scale Based on LSTM-BP Spatiotemporal Combination Model. *Water,* 13**,** 516.

LI, N., HAN, W., SHEN, M. & YU, S. 2016. Load evaluation of non-point source pollutants from reservoir based on export coefficient modeling. *Transactions of the Chinese Society of Agricultural Engineering,* 32**,** 224-230.

LI, P., LU, J., HOU, W., PAN, Y., WANG, Y., KHAN, M. R., REN, T., CONG, R. & LI, X. 2017. Reducing nitrogen losses through ammonia volatilization and surface runoff to improve apparent nitrogen recovery of double cropping of late rice using controlled release urea. *Environmental Science and Pollution Research,* 24**,** 11722-11733.

LI, S., XU, B., WANG, L. & RAZZAQI, A. A. 2020b. Improved Maximum Correntropy Cubature Kalman Filter for Cooperative Localization. *IEEE Sensors Journal,* PP**,** 1-1.

LI, Y., CHEN, H. & SUN, W. 2021b. Load estimation and source apportionment of nitrogen, phosphorus and COD in the basin of Lake Baiyang. *China Environmental Science. 2021a,* 41**,** 366-376.

LIU, B., WU, Q. & GAO, Q. 2018a. Application progress of SWAT model in non-point source pollution research in Xiangxi River Basin. *Water Resources Development and Management***,** 28-31.

LIU, D., CHEN, X., XU, Y., LIU, X. & SHI, C. 2019a. Maximum correntropy generalized high-degree cubature Kalman filter with application to the attitude determination system of missile. *Aerospace Science and Technology,* 95**,** 105441-.

LIU, X., HUA, Q., ZHAO, J. & YUE, P. 2018b. Maximum correntropy square-root cubature Kalman filter with application to SINS/GPS integrated systems. *ISA Transactions,* 80**,** S001905781830168X-.

LIU, X., REN, Z., LYU, H., JIANG, Z. & CHEN, B. 2019b. Linear and Nonlinear Regression-Based Maximum Correntropy Extended Kalman Filtering. *IEEE Transactions on Systems, Man, and*

Cybernetics: Systems, PP**,** 1-10.

LIU, Y., YANG, Y. & LI, F. 2011. Estimation of pollution loads from agricultural non-point sources in Beijing region based on export coefficient modeling approach. *Transactions of the Chinese Society of Agricultural Engineering,* 27**,** 7-12.

LUO, S. 2010. Research on the process of establishing RRC connection based on LTE system for terminals. *Information Communication,* 23**,** 19-21+25.

ONGLEY, E. D., XIAOLAN, Z. & TAO, Y. 2010. Current status of agricultural and rural non-point source pollution assessment in China. *Environmental Pollution,* 158**,** 1159-1168.

OUYANG, W., GAO, X., WEI, P., GAO, B., LIN, C. & HAO, F. 2017. A review of diffuse pollution modeling and associated implications for basin management in China. *Journal of Soils and Sediments,* 17**,** 1527-1536.

OUYANG, W., WU, Y. Y., HAO, Z. C., ZHANG, Q., BU, Q. W. & GAO, X. 2018. Combined impacts of land use and soil property changes on soil erosion in a mollisol area under long-term agricultural development. *Science of the Total Environment,* 613**,** 798-809.

PARK, Y. S. & ENGEL, B. A. 2015. Analysis for regression model behavior by sampling strategy for annual pollutant load estimation. *Journal of environmental quality,* 44**,** 1843-1851.

POWELL, J. L. 1984. Least absolute deviations estimation for the censored regression model. *Journal of econometrics,* 25**,** 303-325.

QIAN, Z. & WANG, Y. 2012. Internet of Things Technology and Application Research. *Journal of Electronics,* 40**,** 1023-1029.

REN, W. & GUO, H. 2015. Estimation of pollution load from non-point source in Baoxianghe basin based, Yunnan Province on improved export coefficient model. *China Environmental Science,* 35**,** 2400-2408.

RONG, Q., CAI, Y., CHEN, B., YUE, W. & TAN, Q. 2017. An enhanced export coefficient based optimization model for supporting agricultural non-point source pollution mitigation under uncertainty. *Science of the Total Environment,* 580**,** 1351-1362.

ROTH, M., OZKAN, E. & GUSTAFSSON, F. A Student's t filter for heavy tailed process and measurement noise. International Conference on Acoustics, Speech, and Signal Processing, 2013.

SARKKA, S. & NUMMENMAA, A. 2009. Recursive Noise Adaptive Kalman Filtering by Variational Bayesian Approximations. *IEEE TRANSACTIONS ON AUTOMATIC CONTROL AC*.

SCHON, T., GUSTAFSSON, F. & NORDLUND, P.-J. 2005. Marginalized particle filters for mixed linear/nonlinear state-space models. *IEEE Transactions on signal processing,* 53**,** 2279-2289.

SHEN, Z., HONG, Q., CHU, Z. & GONG, Y. 2011. A framework for priority non-point source area identification and load estimation integrated with APPI and PLOAD model in Fujiang Basin, China. *Agricultural Water Management,* 98**,** 977-989.

SHOEMAKER, L. 1997. Compendium of tools for basin assessment and TMDL development.

STRAKA, O. & DUNIK, J. Resampling-free Stochastic Integration Filter. 2020 IEEE 23rd International Conference on Information Fusion (FUSION), 2020.

TIAN, Y. & CHENG, Y. Novel Measurement Update Method for Quadrature-Based Gaussian Filters. Aiaa Guidance, Navigation, & Control, 2013.

TRONARP, F., KARVONEN, T. & SARKKA, S. 2019. Student's t-Filters for Noise Scale Estimation. *IEEE Signal Processing Letters,* PP**,** 1-1.

URIBE, N., CORZO, G., QUINTERO, M., VAN GRIENSVEN, A. & SOLOMATINE, D. 2018. Impact

of conservation tillage on nitrogen and phosphorus runoff losses in a potato crop system in Fuquene basin, Colombia. *Agricultural Water Management,* 209**,** 62-72.

UTTORMARK, P. D., CHAPIN, J. D. & GREEN, K. M. 1974. *Estimating nutrient loadings of lakes from non-point sources*, US Government Printing Office.

WANG, G., LI, N. & ZHANG, Y. 2017. Maximum correntropy unscented Kalman and information filters for non-Gaussian measurement noise. *Journal of the Franklin Institute,* 354**,** 8659-8677.

WANG, J., GAO, G., PEI, Y.-S. & YANG, Z.-F. 2010a. Sources and transformations of nitrogen in the Fuhe River of the Baiyangdian Lake. *Huan Jing ke Xue= Huanjing Kexue,* 31**,** 2905-2910.

WANG, L. & HU, H. 2011. LTE system air interface overhead analysis. *Automation Technology and Applications,* 30**,** 32-36.

WANG, L., HUANG, Y.-F. & WANG, G.-Q. 2010b. Review of urban non-point source pollution models. *Huan Jing ke Xue= Huanjing Kexue,* 31**,** 2532-2540.

WANG, S., WANG, X. & XU, D. 2007. Advances in the prediction models of agricultural non-point source pollution. *Transactions of the CSAE,* 23**,** 265-271.

WANG, W., CHEN, L. & SHEN, Z. 2020. Dynamic export coefficient model for evaluating the effects of environmental changes on non-point source pollution. *Science of the Total Environment,* 747**,** 141164.

WANG, X., LIN, L., XIAO, Z., WU, H. & ZHAO, C. 2016. Research in NB-loT technology standardization and trend of development. *Modern Science & Technology of Telecommunications,* 46**,** 5-12.

WISCHMEIER, W. H. & SMITH, D. D. 1978. *Predicting rainfall erosion losses: a guide to conservation planning*, Department of Agriculture, Science and Education Administration.

XIAO, H. & SONG, H. 2013. Research on TD-LTE indoor distribution system construction. *Digital Technology & Application***,** 106.

XIE, Y. 2018. NB-IoT standards system evolution and IoT industry development. *Journal of the Internet of Things*.

XU, B., ZHANG, J. & RAZZAQI, A. A. 2020. A novel robust filter for outliers and time-varying delay on an SINS/USBL integrated navigation model. *Measurement Science and Technology,* 32**,** 015903.

XUEMAN, Y., WENXI, L., YONGKAI, A. & WEIHONG, D. 2020. Assessment of parameter uncertainty for non-point source pollution mechanism modeling: A Bayesian-based approach. *Environmental Pollution,* 263**,** 114570.

YANG, Y., SHEN, L., XIE, D., LUO, Y. & NI, J. 2015. Estimation of pollution loads from agricultural non-point sources in Three Gorges Reservoir Area (Chongqing) based on the export coefficient modeling approach. *Journal of Southwest University (Natural Science),* 37**,** 112-119.

YAZDI, M. N., SAMPLE, D. J., SCOTT, D., WANG, X. & KETABCHY, M. 2021. The effects of land use characteristics on urban stormwater quality and basin pollutant loads. *Science of The Total Environment,* 773**,** 145358.

YE, Y., ZHANG, J. & LI, Y. 2013. Spatial-temporal Variation of Agricultural Non-point Source Pollution Based on GIS Technology in Guangdong Province,China. *China,* 32**,** 369-377.

YOUNG, R., ONSTAD, C., BOSCH, D. & ANDERSON, W. 1989. AGNPS: A non-point-source pollution model for evaluating agricultural basins. *Journal of soil and water conservation,* 44**,** 168-173.

YUANXIN, WU, DEWEN, HU, MEIPING, WU, XIAOPING & HU 2005. Unscented Kalman filtering

for additive noise case: augmented versus nonaugmented. *Signal Processing Letters, IEEE,* 12**,** 357-360.

ZEMA, D. A., LUCAS-BORJA, M. E., CARRà, B. G., DENISI, P., RODRIGUES, V. A., RANZINI, M., ARCOVA, F. C. S., DE CICCO, V. & ZIMBONE, S. M. 2018. Simulating the hydrological response of a small tropical forest basin (Mata Atlantica, Brazil) by the AnnAGNPS model. *Science of The Total Environment,* 636**,** 737-750.

ZHANG, G., SUN, Z. & XUN, S. 2013. 4G Evolution in 3GPP and 5G Perspectives. *Telecommunications Technology***,** 12-17.

ZHANG, L., XIANG, B., HU, Y., FANG, G., JIN, X. & MA, G. 2014. Risk assessment of non-point source pollution in Hulan River Basin using an output coefficient model. *J Agro-Environ Sci,* 33**,** 148-154.

ZHANG, X. Y., ZHAO, J., DING, L., LI, Y., LIU, H. X., ZHAO, Y. F. & FU, G. 2022. Eutrophication evolution trajectory influenced by human activities and climate in the shallow Lake Gehu, China. *Ecological Indicators,* 138.

ZHOU, X. 2015. *Research and Implementation of IMS-based VoLTE Technology.* Harbin Institute of Technology.

ZHOU, Z. & LI, J. 2015. The correlation analysis on the landscape pattern index and hydrological processes in the Yanhe basin, China. *Journal of Hydrology,* 524**,** 417-426.

ZHU, H., ZHANG, G., LI, Y. & LEUNG, H. 2021. A novel robust Kalman filter with unknown non-stationary heavy-tailed noise. *Automatica,* 127**,** 109511.

ZHU, Y., JIN, X., MENG, X., ZHANG, C., TANG, W., SHAN, B. & ZHAO, Y. 2018. Tudy on ammonia nitrogen release flux in the sediment-water interface of Baiyangdian Lake. *Huanjing Kexue Xuebao,* 38**,** 2435-2444.

ZOU, Y., CHEN, W., LI, S., WANG, T., YU, L., XU, M., SINGH, R. P. & LIU, C.-Q. 2022. Spatio-Temporal Changes in Vegetation in the Last Two Decades (2001–2020) in the Beijing–Tianjin–Hebei Region. *Remote Sensing,* 14**,** 3958.

ZOU, Y., DING, X. & WANG, Q. 2017. Key Technologies and Application Prospect for NB-loT. *ZTE Technology Journal,* 23**,** 43-46.