ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Non-point Source Pollution Research in China: Progress, Problems and Prospects

Xiaoli Chen[1,*]

[1] College of Water Conservancy Engineering, Guangdong Polytechnic of Water Resources and Electric Engineering, Guangzhou 510635, China

SUMMARY: *Non-point source (NPS) pollution is still a serious issue in the governance of China's water environment because the sources of pollution are scattered, non-continuous in time, and difficult to separate from point-source pollution. This review systematically organizes the progress of NPS pollution research from the beginning stage of exploratory observations to the current era of data-intensive intelligent modelling, and then presents the evidence in three analytical comparisons: research focus at different periods, operating attributes of the main watershed models, and correspondence between monitoring frequency and rainfall-driven pollution peaks. Although China's surface-water quality has improved to some extent according to the review, the control of NPS is still limited by problems such as a lack of event-scale data, inconsistent monitoring standards, poor sharing of hydrometeorological and water-quality data, and insufficient localisation of foreign models. Although there have been recent achievements in high-resolution remote sensing, improved export coefficient refinement, SWAT-based localisation, digital twin watersheds, and hybrid AI-mechanistic models for load estimation and prediction, their application in practical governance still lacks a traceable management indicator. Future research should focus on standardised event-based monitoring, multi-source data fusion, localised and simplified model development, and transparently couple process-based models with machine-learning algorithms. The above directions will help promote better watershed management and more reasonable division of the NPS pollution control responsibility in China.*

KEYWORDS: *Non-Point Source Pollution; Event-Based Monitoring; Multi-Source Data Fusion; Model Localization and Simplification; Hybrid AI-Mechanistic Models*

1 Introduction

Watershed non-point source (NPS) pollution refers to the spread and dispersion of pollutants in receiving water bodies (such as rivers and lakes) due to rainfall-runoff under the action of the watershed. It has a complicated source structure, unevenly distributed routes, and is widely spread in terms of pollution [1]. Due to economic development and climate change/human activities, pollutants are continuously generated in rainfall-runoff processes, spread and transported rapidly, and thus the water quality of rivers, lakes, reservoirs and other water bodies is continuously declining. Therefore, there have been serious problems in human water and food security, as well as ecological safety. China's Second National Census on Pollution Sources Bulletin [2] shows that in 2017, the national water pollutant discharges included Chemical Oxygen Demand (COD) of 21.4398 million tons, Total Nitrogen (TN) of

*13660231668@163.com
https://doi.org/10.65102/is2026852
Published: 30 April, 2026

3.0414 million tons, Ammonia Nitrogen (NH3-N) of 963,400 tons, and Total Phosphorus (TP) of 315,400 tons. The five provinces of Guangdong, Zhejiang, Jiangsu, Shandong and Hebei made up about 52.94 per cent of China's pollution at the top. As the seriousness of NPS pollution in the watershed continues to increase, the damage to the country's production and livelihood will be exacerbated, and a serious threat will be posed to the relationship between people and nature. Consequently, many specialists and other regions have begun studying the issue of the decline in tourism revenue. Over the past ten years, all parties have been making combined efforts to improve the quality of China's surface water continuously. Figure 1 shows the general trend of surface water quality in China published by the State of Ecology and Environment Bulletin over the past decade [3]. As shown above, the proportion of Grade I-III water quality has been rising steadily from 63.1% in 2014 to 89.4% in 2023; this is an outstanding result. Conversely, the proportion of inferior Grade V water quality has been continuously declining from 9.2% in 2014 to 0.7% in 2023 (and remaining below 1% since 2020). The Black and Odourous Water Body have been treated effectively.

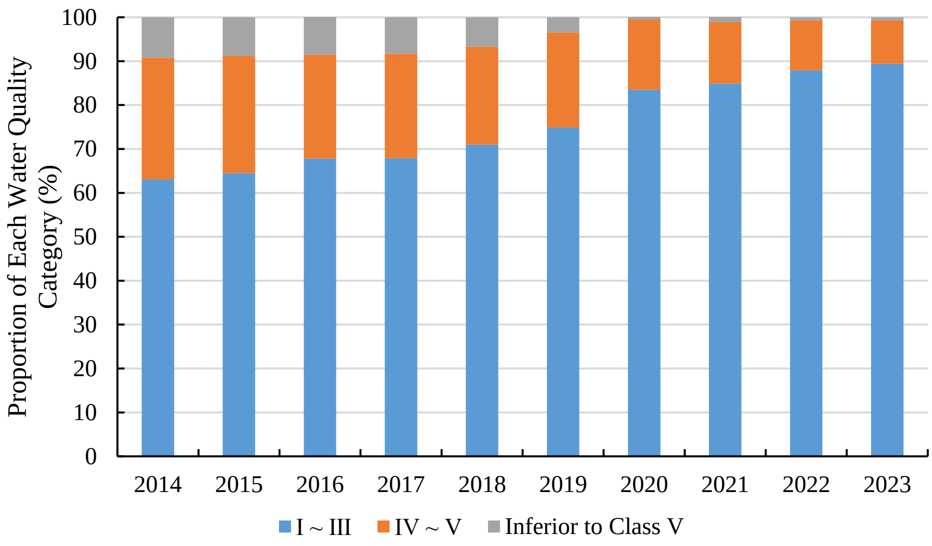

Figure 1: Trends in Surface Water Quality Categories in China (Past 10 Years)

Although China has made good progress in preventing and controlling NPS pollution in watersheds, it is still difficult to monitor, model and mitigate this type of pollution due to the above characteristics: randomness, diffusivity, hysteresis, fuzziness and latency. Therefore, both the academic community and regulatory authorities have maintained a high level of alert and continuously increased research in this area. Systematically review the development and current situation of NPS pollution research in this paper, critically examine existing deficiencies in monitoring, modelling and governance, and propose future research directions. The results will offer scientific support for strengthening the NPS pollution control plan of the Water Management Office.

2 Research Progress of Non-Point Source Pollution

2.1 International Advances

Research on NPS pollution in developed countries such as those in Europe and America can be traced back to the 1930s [4]. Based on the depth and features of many studies, the progress of NPS pollution research in developed countries can be divided into four stages: the exploratory stage (1930s-early 1970s), the development stage (1970s-late 1990s), the refinement stage (21st century to late 2010s), and the intelligent stage (2020s to present).

In the early 1930s, the developed countries began to pay more attention to the study of water pollution in the environment. For example, the United States started to address the problem of widespread pollution from point sources and also at that time started researching farmland soil erosion. Japan has also begun to carry out initial research on the problem of water body eutrophication. In the 1950s, point-source pollution in developed countries had been gradually controlled, but the compliance rate for the water quality of various water bodies was not high enough; therefore, scientists determined that NPS pollution was the main cause of water environmental pollution. In the 1960s and 1970s, advanced nations began to conduct the first studies on NPS pollution, but these were generally limited in scope and mostly employed point-scale observations and qualitative analysis. The well-known Universal Soil Loss Equation was also put forward at this time. At this time, scientists had obtained some insights into the reasons for NPS pollution. They thought that in the process of NPS pollution, pollutants accumulate on the surface of the watershed, are washed away by rainfall, continuously migrate and transform within the watershed with runoff and sediment, and generate pollution loads on the slope surface or in the river channel [5].

In the 1970s and the 1980s, research on NPS pollution was conducted at a fast pace during this time. Countries around the world have carried out in-depth and all-encompassing research on NPS pollution at this time. The contents of this paper are the characteristics and output patterns of NPS pollutants, the accounting of NPS pollution load, and migration and transformation processes of NPS pollutants. In 1996, Johnes put forward the Export Coefficient Model (ECM) and directly established a relationship between land use, soil cover and the annual load of receiving water bodies and NPS pollutants [6]. This way can avoid the complex process of pollutant transport and transformation, thus reducing the need to collect monitoring data during contaminant migration in a watershed and is now used in many other studies. At the same time, NPS pollution models have also developed rapidly by integrating with the 3S technology system (GIS, GPS, RS). The reliable spatial data processing capabilities of these technologies have helped to build and run a model-and-simulation of NPS pollution. Many widely used and influential distributed hydrological models have appeared in recent years, such as the HSPF model [7], AGNPS model [8], MIKE SHE model [9], STORM model [10], SWMM model [11] and SWAT model [12]. As the necessary tools for watershed-scale NPS pollution simulation, these models have different features and functions. At the time of selection, all kinds of conditions for the model should be taken into account, such as its scope of application, local relevance and other requirements for input data.

The first two decades of the 21st century were a time of improvement and application for NPS pollution research. At this time, many scholars have been conducting simulations of pollutant transport and change under various NPS pollution models for different catchments and areas around the world. The first three parts of this paper are as follows: (1) The impact of land-use change on watershed NPS pollution, which explores how changes in NPS pollution are driven by rapid land-use/cover alteration due to urbanization [13]; (2) The effect of climate change on NPS pollution, and analyses are conducted on how modified NPS pollution patterns result from extreme rainfall-runoff fluctuations under a changing climate [14]; and (3) Mitigation measures, which evaluate the effectiveness of various ways, such as farmland-to-forest conversion, terrace farming, reduced fertilizer application, and vegetative buffer strips. Among them, the Best Management Practices (BMPs) in the US are considered to be one of the highly effective means of controlling NPS pollution [15]. In addition, due to various conditions of soil and hydrometeorology, scholars around the world have carried out numerous studies to improve NPS pollution models for the main river basins and adapt them to different areas. Abbaspour and others added a groundwater module to the SWAT model to run many experiments on nitrate infiltration into groundwater across Europe to help evaluate water resource availability and quality [16].

Research on the pollution indices of watershed NPS around the world has been advancing in recent years [17-19]. In terms of data acquisition, the application of satellite remote sensing, drones and the Internet of Things (IoT) for real-time monitoring of water quality has been increasingly studied to expand the area of monitoring and improve the accuracy of pollution data. The EU's Digital Twin Earth project is used to model changes in watershed pollution under multiple sources of remote sensing data. At the level of models, Random Forest (RF) and Long Short-Term Memory (LSTM) networks for machine learning are now being integrated with conventional NPS pollution models to improve the accuracy of pollutant load simulation [20, 21]. Examples include coupling AI with the SWAT model in the Mississippi River Basin (USA), and Deltares' (Netherlands) Delft3D-based virtual watershed model that dynamically integrates monitoring data to simulate pollution dispersion and evaluate mitigation strategies. Artificial intelligence (AI) algorithms are being used to design ecological buffer zones and optimise the layout of constructed wetlands for pollution control to maximise the efficiency of pollution interception, as shown in AI-optimised agricultural cover crop planning in the Chesapeake Bay watershed (USA). At the same time, the world is also actively promoting intelligent and sustainable NPS pollution research and governance by coordinating technologies, policies and society.

The above chronological descriptions can be viewed more clearly by comparing the four stages in the monitoring, modeling and governance dimensions. Therefore, based on the previous research, a synthesis index of 0 to 5 was developed; 0 means the subject matter is not covered in the literature, and 5 indicates a major research problem that is frequently addressed. Figure 2 shows the stage-dimension comparison above.

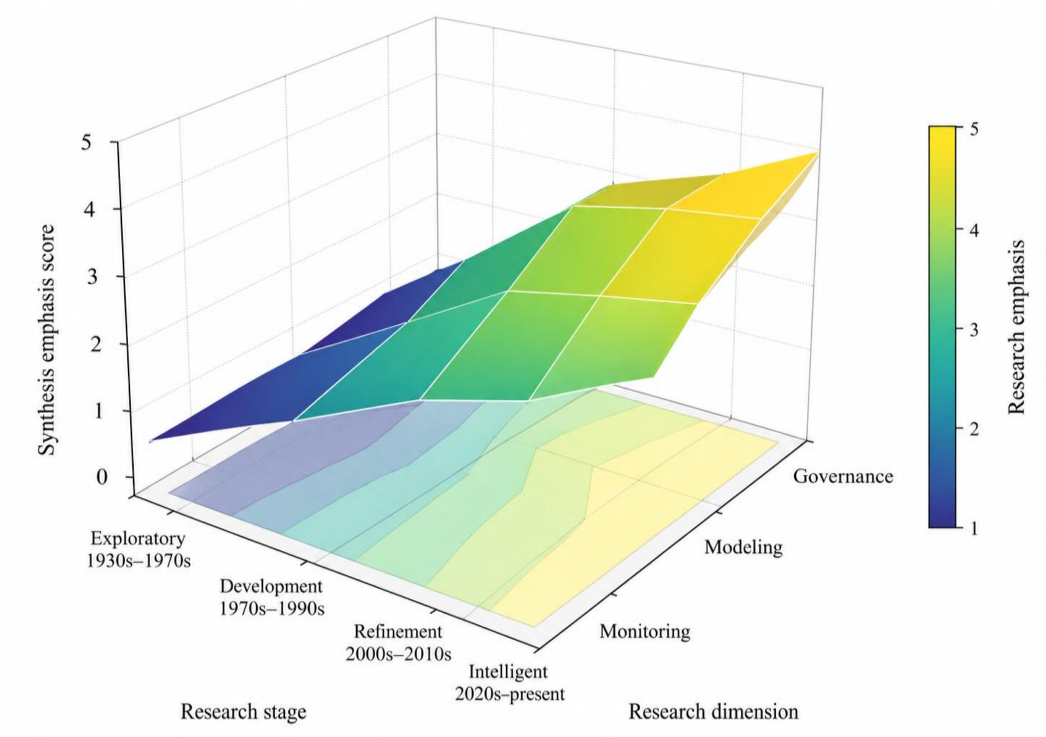

Figure 2: Stage-Dimension Synthesis of NPS Pollution Research Evolution.

Figure 2 shows that the transition from the exploratory to the intelligent stage is not a simple accumulation over time. Model building and improvement capacity are at their highest in the development and optimisation periods; only after the introduction of remote sensing, IoT and high-frequency sensing has monitoring capacity become feasible. Governance maturity is still lower than modelling maturity in the intelligent stage, and the main bottleneck has moved from mechanism identification to data standardisation, operational prediction and management translation. The pattern shows that although advanced models have achieved good results, the problem of NPS pollution governance has not been solved; without event-level monitoring and institutionalized data-sharing mechanisms, the output of the model cannot be used reliably as the target for control measures.

2.2 Chinese Advances

Research on NPS pollution in China began relatively late. Most scholars have agreed that the mechanisms of NPS pollution are complex and diverse, but generally include two main paths: sources and sinks. Before, it was thought that the pollutants accumulated on the surface of the watershed would be washed away by rain and spread as pollution loads down the slopes. The latter refers to the transport and change of state for NPS pollutants in river channels. The cause of NPS pollution is rainfall-runoff processes, and human land-use activities are the sources of these pollutions. Recently, many English-language studies have been modifying load-accounting and process-based models for the Chinese basin. Process-based export-coefficient adjustments have been employed to reduce the monthly TN-load simulation error in semi-arid agricultural areas and enhance the explanatory power over traditional ECM models [22]. In large basins, combined export-instream coefficient approaches have also been put forward to reduce the dependence on fixed empirical coefficients and improve load attribution in spatially complex hydrological regions [23]. At the model-application level, high-resolution remote sensing and SWAT have been used to optimise catchment thresholds and obtain basin-suitable NPS parameters in the Ningxia Yellow River Irrigation District [24]. Based on the above studies, the main trend in China's research has moved from adopting foreign model structures to parameter localisation, process representation and data-driven calibration.

Since 2010, with the continuous progress of China's smart water conservancy development, new information technologies such as digital twins, the Internet of Things (IoT) and artificial intelligence (AI) have developed rapidly in water resource applications [25-27]. The new technology has been introduced in NPS research to improve quality of data, conduct real-time pollution source monitoring, data processing and advanced analysis, and so on. High-resolution national lake and reservoir data have also started to close the gap between sparse field observations and the demand for temporally continuous water-quality information [28]. IoT technology can be used to set up many sensors and monitoring devices at important cross-sections of rivers, lakes and reservoirs to collect real-time data on hydrology and the environment, such as precipitation, water level, flow velocity, and water quality indicators. Apply big data technology to conduct in-depth mining and analysis of the spatial-temporal distribution pattern, migration and transformation mechanism, and driving factors of NPS pollutants. Use the computing and storage capacity of the cloud platform to run NPS pollution models more efficiently, enable cross-regional and cross-departmental data sharing, and carry out joint analysis. Artificial intelligence technology can now be used to forecast high-risk areas and timeframes for NPS pollution, as well as the potential scope of harm, to offer earlier warning information and support decision-making.

Digital Twin Watersheds are now leading-edge technologies in watershed management. Digital mapping, real-time simulation and intelligent decision-making can be used to construct a virtual model of the physical watershed, and thus dynamic simulation and optimised management of all elements and processes can be achieved. The above means to use water resources sustainably, prevent and control pollution in the watershed, and enhance disaster response ability. Typical cases include the Digital Twin Yangtze River project in the Yangtze River Basin, the Smart Yellow River digital twin system, and the flood prevention four-forecasting platform for the Pearl River Basin. The above measures have introduced active intelligent management of NPS pollution instead of passive mechanical control.

3 Current Research Status and Existing Problems in NPS Pollution Studies

3.1 Data Monitoring

Acquisition of water quality section monitoring data serves as the foundation for watershed NPS pollution simulation studies. The first sources of official water quality monitoring data in China are the China National Environmental Monitoring Center (CNEMC) and the Ministry of Water Resources and other local environmental protection offices. In the 14th Five-Year Plan period, a total of 3,641 national surface water assessment sections (monitoring points) were established for the National Surface Water Environmental Quality Monitoring Network and referred to as national assessment sections. Therefore, CNEMC will release a monthly report on the quality of the nation's surface water based on the analysis above. In addition, the assessment sections established by the ecological and environmental departments at the provincial level for surface water are also referred to as provincial assessment sections and serve to monitor and assess the quality of surface water in the province. Provincial assessment sections are also regularly monitoring surface water quality in the same way as the national assessment section. Due to the various reasons in different provinces, the specific monitoring indices, frequencies and evaluation norms will also be different. Both the national and provincial assessment sections have adhered to the principles of stability, convenience and representativeness in their establishment and can generally reflect the water quality situation and pollution features of the region. However, there are still practical problems, such as low monitoring accuracy, an emphasis on quantity over quality, incomplete coverage of monitoring projects, and a single monitoring method; as a result, the work of water environment governance has been severely hindered.

Pollution is widely distributed, and its changes and transformations are complex and take a long time to show. It is not feasible to observe this accurately in the following cases:

(1) Monitoring frequency

As shown above, the CNEMC publishes monthly reports on nationwide surface water quality, which are monthly-scale data, and automatic monitoring of real-time data is released every four hours. NPS pollution in river basins is caused by rainfall-runoff processes. In recent years, with the background of climate change and frequent occurrences of extreme short-duration rainstorm events in southern areas [29], 4-hour-scale monitoring data cannot accurately reflect peak pollutant concentrations, and there is a lack of data for instantaneous pollution events triggered by rainstorms; thus, a real-time response to changes in pollution is not feasible. Generally speaking, a higher monitoring frequency will yield more data, and thus the results of NPS pollution models in simulations will be more accurate.

(2) Monitoring standards

At present, there are no standardised technical specifications, monitoring equipment or measurement methods for NPS pollution monitoring. The reasons for NPS pollution are related to rainfall-runoff processes. Researchers have not yet developed a way to distinguish between the loads of point-source pollution and NPS pollution in a given rainfall-runoff event. Therefore, when monitoring the water quality of a river basin, one must separate and monitor the quality data during flood and dry periods. In addition, the areas of monitoring are different, so the indices for these regions may also be different; therefore, the selection of these indices needs to consider factors such as water environmental quality, functional zoning of the water environment, characteristics of pollution source emissions, etc. Given that most NPS pollution models are simulated at the watershed scale, the deployment of monitoring stations in the monitoring section should be limited to key control points, such as the inlets of major tributaries to rivers, estuaries, reservoirs and lakes.

(3) Data uniformity

Monitoring data are generally distributed among several divisions, such as environmental protection, hydrology, meteorology and agriculture, and do not adhere to the same format and standard; thus, data integration and sharing are difficult. Due to the difference in monitoring methods and ways, there will be a large amount of variation in the quality of data from various places. The above problems will increase the complexity and difficulty of data processing for constructing NPS pollution models, and thus seriously affect the simulation accuracy of the models.

(4) Expensive equipment

The initial purchase and operation costs of high-precision monitoring equipment are very high, such as online water quality mass spectrometers and new hyperspectral remote sensing drones; they also face problems of equipment obsolescence and sunk costs due to rapid technological changes and iterations. At the same time, high-end equipment may be difficult to manage and will not promote technological progress. Small and medium-sized watershed management offices are not equipped for manual sampling and thus have a low monitoring rate.

The main reason for the uncertainty in event-driven NPS load estimation is a low-frequency monitoring cycle. Figure 3 is a simple event-capture metric defined as min(1, event peak duration/monitoring interval), and it is used to compare short rainstorm peaks with common sampling intervals.

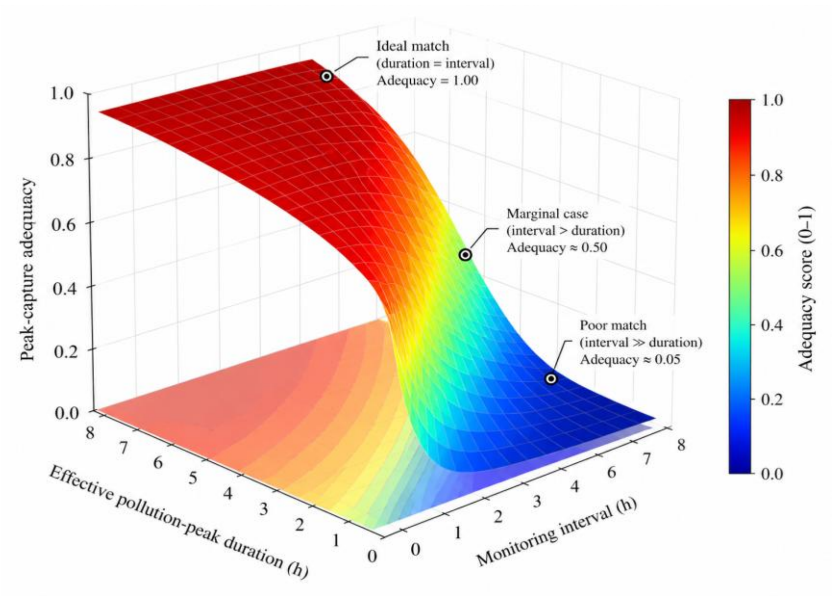

Figure 3: Three-dimensional matching surface of monitoring interval and rainfall-driven NPS pollution peak duration.

As shown in Figure 3, four-hour automatic monitoring can only cover all peaks if the effective pollution-peak duration is four hours or more. If the peak window is one hour long, the capture adequacy is 0.25; if it is two hours long, it is about 0.50. Therefore, it can be concluded that monthly reports and four-hour automatic data are not suitable for first-flush events in urban areas and short-duration rainstorms in southern China. Increase the frequency of monitoring in the critical control area to sub-hourly during a heavy rain; automatically trigger a response based on precipitation or turbidity, and synchronize flow-water quality sampling to improve load assessment accuracy.

3.2 Model Simulation

At present, the main way China's water environment management authorities have dealt with the problem of watershed pollution is to take engineering measures; although they have achieved good results, they have not addressed the fundamental causes of the pollution. The system development of distributed watershed hydrological models for simulation analysis and management prediction of water environment problems is relatively immature. Model simulations can show the path and changes of pollutant migration, as well as where and when they are concentrated in space and time, to identify pollution sources for targeted source control. The above way of thinking offers a scientific basis for the all-encompassing management of watershed water environments. As mentioned above, in the 1970s to the 1990s, many excellent watershed NPS pollution models have appeared, such as the HSPF model, AnnAGNPS model, MIKE SHE model, STORM model, SWMM model and SWAT model, etc. The general information of the model is as follows: Table 1. Among all the models, the SWAT model has a wide range of applications and is used more often than the others. A good tool for studying water resources and nitrogen-phosphorus (NPS) pollution around the world.

Table 1: Introduction to Some NPS Pollution Models

Model Name	The earliest development time	Time scale	Spatial scale
SWMM	1971	Single simulation/Continuous simulation	Regional scale
HSPF	1976	Continuous simulation	Watershed scale
MIKE SHE	1982	Single simulation/Continuous simulation	Watershed scale
SWAT	1994	Continuous simulation	Watershed/Regional scale
PLOAD	2001	Single simulation	Watershed scale
Ann AGNPS	2005	Continuous simulation	Watershed scale
STORM	2011	Single simulation/Continuous simulation	Urban watershed scale

In addition to listing the model name, some of its working conditions must also meet the demands of watershed managers. As shown in Table 1, Figure 4 is a model that has been recoded according to development time, time-scale capacity (1=single-event, 2=continuous, 3=both), spatial applicability (1=regional/urban, 2=watershed, 3=watershed-regional), and relative data demand.

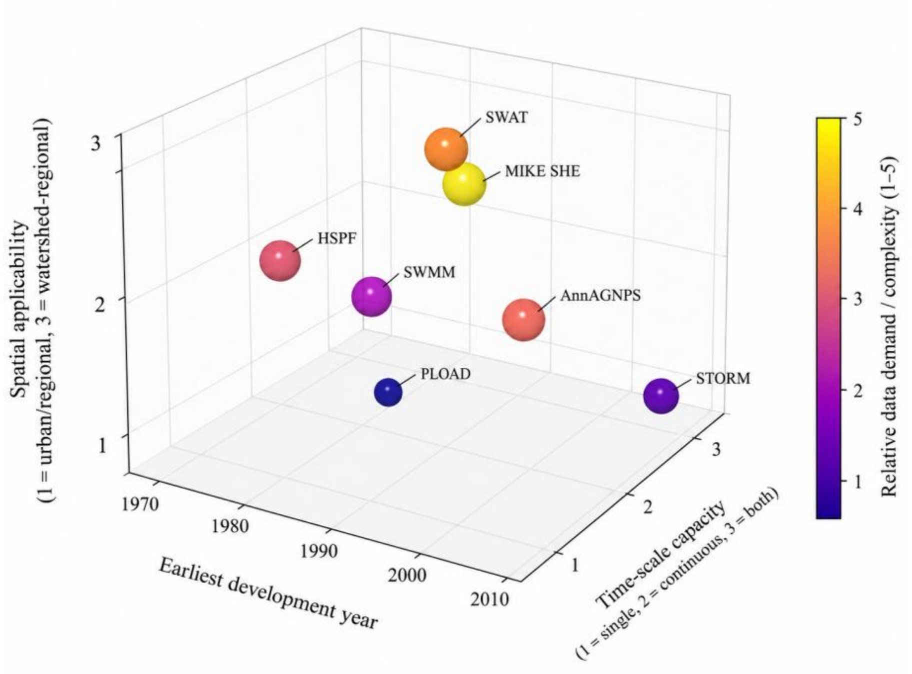

Figure 4: Three-Dimensional Operational Comparison of Common NPS Pollution Models.

Figure 4 shows a relatively small difference. MIKE SHE and SWAT are high-capacity methods for continuous watershed simulation and have richer process representation, but they also require more meteorological, soil, land-use and management data. PLOAD and STORM have a lower data demand and are easier to deploy, but their event or simplified spatial representation limits their utility for refined source apportionment. SWMM is suitable for urban runoff, but HSPF, SWAT and AnnAGNPS are more suitable for watershed-scale agricultural and mixed-source conditions. Therefore, model localisation in China should not focus solely on complexity but also consider the monitoring density, watershed size, target pollutant and decision horizon.

Although some research and applications have been carried out on NPS pollution models in China, there are still some serious defects that prevent their full implementation and promotion in actual environmental management:

(1) Data acquisition and quality limitations

High-resolution meteorological, soil and land-use data, such as hourly rainfall and soil permeability, are not readily available in rural areas of China. The migration parameters of pollutants, such as the loss coefficient for agricultural fertilisers, are based on literature or empirical values; therefore, there has been a lack of local measurement, revision and verification of these parameters. The data are also relatively old and lack timeliness; many models still use static data from years ago, such as land-use maps from 10 years ago, and cannot reflect the changes caused by rapid urbanisation or intensive agricultural production.

(2) Insufficient localization and adaptation of the model

When foreign models such as SWAT and HSPF are introduced in China, they are generally used without detailed calibration of local hydrological processes and pollution mechanisms. Unique agricultural models in China, such as rice field drainage and decentralized farming, are not shown in the model either. In addition, China does not have independent model development, and the application scope of independently developed models (such as Tsinghua University's IHACRES) is limited due to a lack of commercial support.

(3) Scale Conversion and Accuracy Issues

Large-scale simulations of river basins often fail to consider small-scale heterogeneity, such as the interception effect of ditches, ponds and dams; thus, there are considerable errors in pollution load estimation and reduced simulation accuracy. Simulation results at the watershed scale cannot be directly applied to guide pollution control projects in administrative Areas and areas. A long-term average annual or monthly load simulation cannot account for sudden pollution events in severe rainstorm episodes, such as the initial scouring effect of urban runoff that has become more prominent recently.

(4) The model results are disconnected from the management application

The spatial distribution map of pollutant emissions produced by complex models has not been used effectively by grassroots management departments, and decision-making is still based on past experience; therefore, the results of the model cannot be applied. In addition, there are no standardized specifications for applying non-point source pollution models in environmental assessment and river chief system assessment, no unified technical guides, and the research results from different teams vary considerably.

3.3 Prevention-Control Management

China has built a good policy system and technological support for controlling the pollution of urban wastewater in river basins, as well as implemented strict supervision [30]. There are still some problems and deficiencies. A representative number of strengths and weaknesses in this area are listed below.

Several new strengths have been achieved at present in some parts. First, the policy system has been strengthened gradually. A two-tier regulatory system has been established for the "River Chief System" and the "Lake Chief System" to distribute the work of watershed management. Some policy documents have set specific governance goals and issued a "Water Pollution Prevention and Control Action Plan". Ecological compensation mechanisms have also been established in the main river basins, and a cross-provincial pilot program in the Xin'an River Basin is a typical case. Secondly, new applications of technology are being developed continuously. Independent models have also been developed to some extent (e.g., the Time-Variant Gain Model - TVGM). Now, the three S's for pollution monitoring are widely available around the world: RS, GIS and GPS. The promotion of ecological engineering technology has also achieved some good results, such as the construction of artificial wetlands and ecological trenches. Many forms of the government have been put forward. A collaborative governance mechanism of the form "government leadership, enterprise participation and public supervision" has been proposed. A unified management system for "source control, process interception and end-of-pipe treatment" has been built and operates normally. The above efforts have led to the creation of many successful typical cases, such as the governance projects in the Taihu Lake Basin and the Erhai Lake Basin.

However, there are still some serious shortcomings and problems; these can be divided into the following categories:

(1) Fragmented Management System and Insufficient Inter-Departmental Coordination

The problem of "many dragons controlling the water" has not been solved, and there is also a division of labour among departments such as water resources, environmental protection and agriculture. The policies and standards are inconsistent, so the decisions cannot be implemented well. Administrative divisions divide the whole river basin and are the cause of disputes over the areas along the river and among the people on either side of the river. A typical case is the shortage of ecological flow in the middle and lower reaches of the Han River Basin due to development in the upstream area. Ecological Compensation Mechanisms are not yet well-established. The compensation standards for the trans-provincial basin are often vague, and the allocated funds frequently fail to cover the actual governance costs, as shown in the water source areas of the South-to-North Water Diversion Project.

(2) Over-reliance on End-of-Pipe Treatments and Insufficient Source Control

The present strategies tend to address problems late in the process. Most of the measures focus on building wastewater treatment plants and have been unable to address agricultural non-point sources, such as fertilizer reduction and eco-friendly farming practices. The Management of Urban Runoff Pollution is lagging. Road and roof runoff are responsible for as much as 75% of the particulate matter in urban water bodies, and the coverage of stormwater collection and treatment facilities is still low. In addition, ecological restoration technologies that are relatively suitable for small and medium-sized watersheds, such as constructed wetlands and riparian buffer zones, are not widely used.

(3) Inadequate Policy Implementation and Financial Support

Weak Policy Implementation Capacity at the Local Level. Operation and maintenance of rural sewage treatment facilities are generally lacking; for example, 41.9 per cent of such facilities in Jieyang City were reported to be operating abnormally. Funding will not be forthcoming. Funding for lake restoration, such as the 2 billion yuan allocated during the 13th Five-Year Plan period, is far lower than the actual demand to address pollution caused by rapid urbanisation. Market-oriented mechanisms are also missing. Ecological compensation is still mainly provided by government fiscal transfers, and market-based financing options such as Public-Private Partnerships (PPPs) have not been explored.

4 Future Research Directions Exploration

Recently, facing the challenges of climate change and an increasingly intensive human activity, research on NPS pollution in watersheds has begun to explore some more specific directions to make noticeable progress:

4.1 Standardization and Normalization of NPS Pollution Monitoring

NPS Pollution has considerable space-time variation. Data from different research teams or areas are often collected at different times in various forms and are therefore not suitable for space-time comparison. Standardized and normalised monitoring data serve as the "gold standard" for parameter calibration and validation of NPS pollution models, and thus are used to assess the accuracy of model simulation and prediction results. Thus, only a single monitoring protocol and norm need to be set. Standardised monitoring provides a common language for communication in the joint governance of the upstream and downstream areas and offers data support for accountability at the national level. Standardisation and normalisation of NPS pollution monitoring are required to change it from an indefinite source of pollution into an observable, quantifiable and manageable pollutant. Therefore, the above work needs to combine top-down standard setting and bottom-up standardised practice. Finally, it hopes to build a scientific, efficient and reliable network of NPS pollution monitoring in China and provide strong data support for the optimization of water environment management.

4.2 Integration of Multi-Source Data and Open Data Sharing

Data for the research on NPS pollution are relatively comprehensive and include hydrology and climate data, water quality data, land-use type information, soil conditions, fertilizer use patterns, livestock breeding statistics, etc. At present, many sources of data are available to scholars, such as government platforms, scientific databases, commercial data providers and citizen science initiatives. However, the above work needs to ensure the authenticity, reliability and consistency of the data and has significant problems. Therefore, the construction of data-sharing platforms and the search for high-efficiency, high-accuracy methods of data collection need to be strengthened. The first to be released under open-access is the basic data for weather and water resources; sensitive areas such as agricultural nitrogen-five (NPS) pollution will be released later. Pilot watersheds, such as Taihu Lake and Dianchi Lake, can be chosen to build "data sandbox" environments for exploring secure and controllable sharing mechanisms. Build a network of sky-ground monitoring in the main basins and develop an all-encompassing real-time monitoring system using satellite remote sensing, drones and IoT sensors. The above will help to build an ecological and environmentally friendly data governance system for China under a "government-led, research-supported, enterprise-engaged, and public-supervised" model.

4.3 Simplification and Localization of NPS Pollution Models

NPS pollution is diffuse, stochastic, and inherently uncertain; thus, it is difficult to model. Directly applying high-end foreign models (e.g., SWAT, HSPF, AGNPS) often results in problems of incompatibility with local conditions. Therefore, simplification and localisation are required for the good performance of all models. Localisation is not a minor modification; it is an entire engineering project that requires the construction of a full-coverage database for localised parameters. Establish parameter sets suitable for the different soils, crops and management systems in China, and conduct field experiments and extensive literature reviews. Modify some of the algorithms. For example, revising the SCS-CN curve number can improve the accuracy of runoff simulation for paddy fields in the hilly areas of southern China or irrigation districts in the northern plains significantly.

On the other hand, simplification refers to intelligent simplification based on scientific understanding. Reduce the amount of data and the computational cost of the model by using methods such as spatial lumping or statistical regression, and still achieve a good level of accuracy. Given that many of the small and medium-sized watersheds in China lack good monitoring systems, special models with fewer parameters, simpler structures and high computational efficiency should also be developed. Agricultural irrigation districts and small watersheds are especially relevant to this, and recent high-resolution remote sensing and SWAT applications have shown that model thresholds and NPS parameters need to be modified according to basin-specific conditions.

4.4 Hybrid AI-Mechanistic Models

Process-based models are often relatively complicated and compute-intensive, and purely AI-based models are essentially "black boxes" with poor physical interpretability. In the future, both types of models will be integrated more deeply to develop new-generation hybrid models that are both physically reliable and data-driven. Recently, SWAT-XGBoost-LSTM applications have shown that combining mechanistic simulation with interpretable deep learning can improve the accuracy of NH_3-N and TN predictions and retain information on NPS pollution processes [31]. Main applications are as follows:

(1) Replacing Complex Subprocesses: Utilizing machine learning (e.g., LSTM neural networks) to substitute computationally intensive and highly uncertain sub-processes within process-based models (e.g., soil nitrogen-phosphorus transformation, complex infiltration processes). This approach maintains the physical framework while significantly enhancing computational efficiency and accuracy.

(2) Intelligent Parameter Calibration: Developing smart optimization algorithms (e.g., genetic algorithms, particle swarm optimization coupled with AI surrogate models) to

automate and streamline the parameter calibration of complex models, thereby addressing the time-consuming and labor-intensive challenges associated with traditional manual methods.

(3) Embedding Physical Constraints: Integrating physical laws directly as constraints into AI models, forcing neural networks to adhere to fundamental physical principles while learning from data. This significantly enhances the generalization and extrapolation capabilities of AI models.

(4) AI as a "Data Fusion Hub": Deploying AI to act as a "data fusion hub," enabling process-based models to better leverage massive, heterogeneous data from novel sources like remote sensing and the Internet of Things (IoT). For instance, recurrent neural networks can process temporal sensor data for denoising, imputation, and feature extraction. The resulting high-quality data can then be fed into process-based models or used for data assimilation to dynamically correct model states.

5 Conclusions

Based on a systematic review of numerous studies, four periods of development in the research of NPS pollution have been identified: exploration, development, refinement and intelligent sustainable development. The revised analysis shows the following three special cases. First, the surface-water quality improvement in China is considerable: by 2023, the proportion of Grade I-III sections had risen to 89.4% from 63.1% in 2014, and the number of inferior Grade V sections had fallen from 9.2% to 0.7% (3). However, the pressure of NPS is still difficult to manage because pollutant loads are highly event-driven and frequently missed in low-frequency monitoring. Second, according to the results of the model comparison, SWAT, HSPF, MIKE SHE and AnnAGNPS have better process representation; however, simpler tools such as PLOAD and STORM have relatively low data requirements but limited source apportionment capabilities. Third, according to the monitoring-frequency analysis, a four-hour interval can only capture one-hour and two-hour pollution peaks with approximate adequacy of 0.25 and 0.50, respectively; therefore, rainfall-triggered high-frequency monitoring is required for first-flush events. Although there have been some progress, three problems remain: inconsistent monitoring standards, a lack of localised and simplified models for Chinese watersheds, and poor translation of model results into governance measures. Therefore, future work will focus on standardised event-based monitoring, multi-source data sharing, development of localised and transparent model structures, and hybrid AI-mechanistic modelling with physical constraints and interpretable outputs. The above adjustments will enable the calculation of NPS load to be carried out accurately, distribute the responsibilities among administrative units reasonably, and ensure the long-term sustainability of China's watershed water environment.

Funding

This study was supported by the 2024 Guangdong Provincial Ordinary University Youth Innovation Talent Project [grant number 2024KQNCX273]

2026 Guangdong Provincial Special Fund Project for Water Resource Conservation and Protection: Investigation and Assessment of Water Resources and Protection Mechanism Research in the “Two Mountains” Leading Area

References

[1] Blair, A., Lovelace, S., Sanger, D., et al. (2014). Exploring impacts of development and climate change on stormwater runoff. Hydrological Processes, 28(5), 2844-2854.

[2] Ministry of Ecology and Environment of the People's Republic of China, National Bureau of Statistics, & Ministry of Agriculture and Rural Affairs of the People's Republic of China. (2020). Bulletin of the Second National Census of Pollution Sources. Beijing. (Normative source)

[3] Ministry of Ecology and Environment of the People's Republic of China. (2024). 2023 Report on the State of Ecology and Environment in China. Beijing. (Normative source)

[4] Tim, U. S., & Jolly, R. (1994). Evaluating agricultural non-point source pollution using integrated geographic information systems and hydrologic/water quality model. Journal of Environmental Quality, 23(1), 25-35.

[5] Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications, 8(3), 559-568.

[6] 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(3-4), 323-349.

[7] Donigian, A. S., Bicknell, B. R., Imhoff, J. C., et al. (1995). Hydrological Simulation Program-Fortran (HSPF). In V. P. Singh (Ed.), Computer Models of Watershed Hydrology. Water Resources Publications.

[8] Haddock, G., & Jankowski, P. (1993). Integrating nonpoint source pollution modeling with a geographic information system. Computers, Environment and Urban Systems, 17(5), 437-451.

[9] Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986). An introduction to the European Hydrological System - Systeme Hydrologique Europeen, SHE, 1: History and philosophy of a physically-based, distributed modelling system. Journal of Hydrology, 87, 45-59.

[10] Singh, V. P. (1997). Computer models of watershed hydrology. Water Resources Publications.

[11] Liong, S. Y., Chan, W. T., & Law, C. L. (1991). An expert system for storm water management modelling, and its application. Engineering Applications of Artificial Intelligence, 4(5), 367-375.

[12] Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large area hydrologic modeling and assessment - Part 1: Model development. Journal of the American Water Resources Association, 34(1), 73-89.

[13] Lam, Q. D., Schmalz, B., & Fohrer, N. (2010). Modelling point and diffuse source pollution of nitrate in a rural lowland catchment using the SWAT model. Agricultural

Water Management, 97(2), 317-325.

[14] Yang, X. Y., Warren, R., He, Y., et al. (2018). Impacts of climate change on TN load and its control in a river basin with complex pollution sources. Science of the Total Environment, 615, 1155-1163.

[15] Jang, S. S., Ahn, S. R., & Kim, S. J. (2017). Evaluation of executable best management practices in Haean highland agricultural catchment of South Korea using SWAT. Agricultural Water Management, 180, 224-234.

[16] Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., et al. (2015). A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. Journal of Hydrology, 524, 733-752.

[17] Sun, A. Y., & Scanlon, B. R. (2019). How can Big Data and machine learning benefit environment and water management: A survey of methods, applications, and future directions. Environmental Research Letters, 14(7), 073001.

[18] Noori, N., Kalin, L., & Isik, S. (2020). Water quality prediction using SWAT-ANN coupled approach. Journal of Hydrology, 590, 125220.

[19] Aslan, S., Zennaro, F., Furlan, E., & Critto, A. (2022). Recurrent neural networks for water quality assessment in complex coastal lagoon environments: A case study on the Venice Lagoon. Environmental Modelling & Software, 154, 105403.

[20] Cheng, H., Yang, F., Jiang, W., et al. (2025). Assessing the impacts of climate change and anthropogenic activities on non-point source pollutions in shallow lakes by using an OEO-LSTM hybrid model. Process Safety and Environmental Protection, 202, 107760.

[21] Frank, J. K., Suesse, T., & Brenning, A. (2025). An assessment of spatial random forests for environmental mapping: The case of groundwater nitrate concentration. Environmental Modelling & Software, 193, 106626.

[22] Zhang, S., Cheng, G., et al. (2023). An agro-hydrological process-based export coefficient model for estimating monthly non-point source loads in a semiarid agricultural area. Journal of Cleaner Production, 385, 135519.

[23] Wang, X., Wu, L., Luo, Y., Liu, Y., & Wang, R. (2025). Integrated export instream coefficient model for accurate nonpoint source pollution estimation and management in the Yellow River Basin. Scientific Reports, 15, 26302.

[24] Zhang, S., Zhang, L., Meng, Q., Wang, C., Ma, J., Li, H., & Ma, K. (2024). Evaluating agricultural non-point source pollution with high-resolution remote sensing technology and SWAT model: A case study in Ningxia Yellow River Irrigation District, China. Ecological Indicators, 166, 112578.

[25] Li, W., Ma, Z., Li, J., Li, Q., Li, Y., & Yang, J. (2024). Digital Twin Smart Water Conservancy: Status, Challenges, and Prospects. Water, 16(14), 2038.

[26] Zhu, M., Wang, J., Yang, X., Zhang, Y., Zhang, L., Ren, H., Wu, B., & Ye, L. (2022). A review of the application of machine learning in water quality evaluation. Eco-

Environment & Health, 1(2), 107-116.

[27] Wang, F., Wang, Y., Zhang, K., Hu, M., Weng, Q., & Zhang, H. (2021). Spatial heterogeneity modeling of water quality based on random forest regression and model interpretation. Environmental Research, 202, 111660.

[28] Luan, S., Pan, H., Shen, R., Xia, X., Duan, H., Yuan, W., et al. (2025). High Resolution Water Quality Dataset of Chinese Lakes and Reservoirs from 2000 to 2023. Scientific Data, 12, 572.

[29] Chen, X., & Huang, G. (2020). Applicability and hydrologic substitutability of TMPA satellite precipitation product in the Feilaixia Catchment, China. Water, 12, 1803.

[30] Chen, L., Zhong, Y., Wei, G., & Shen, Z. (2014). Upstream to downstream: A multiple-assessment-point approach for targeting non-point-source priority management areas at large watershed scale. Hydrology and Earth System Sciences, 18, 1265-1272.

[31] Huan, J., Fan, Y., Xu, X., Zhou, L., Zhang, H., Zhang, C., Hu, Q., Cai, W., Ju, H., & Gu, S. (2025). Deep learning model based on coupled SWAT and interpretable methods for water quality prediction under the influence of non-point source pollution. Computers and Electronics in Agriculture, 231, 109985.