Development of water basin pollution emission inventory: a preliminary literature review and Its implication for China
Shuchang Zhao, Yi Liu, Yaxin Liu and Manli Gong

ABSTRACT
The water pollution emission inventory is a fundamental decision-making tool that links emissions and water quality. However, accurately quantifying the emissions is challenging, due to the variety of contributing sources, complexity of methods for calibration and validation, and lack of moderate spatio-temporal resolution. Over the last two decades, tremendous efforts have been made to improve the accuracy of emission inventories, and significant improvements have been accomplished. This study summarizes the recent progress on inventory development, by recapping the sector-based configurations and associated coefficient databases. Subsequently, we highlight the calculation, validation, and spatio-temporal resolution of emissions contained in the present inventories. Finally, we suggest future directions for conducting a systematic procedure and further improving the accuracy of emission inventories.

Key words | integrated water basin management, spatio-temporal resolution, validation, water emission inventory

HIGHLIGHTS
- Close attention is paid to the configurations of the multi sources and the coefficient database.
- Calculation and validation, and the spatio-temporal resolution of inventory are highlighted.
- A standardized and systematic inventory procedure scheme is proposed and some recommendations for future investigation are suggested.

INTRODUCTION
Understanding the emission origins of water pollutants is an important prerequisite for designing efficient strategies to improve water quality. Accurately quantifying emissions is challenging, owing to the variety of contributing sources, pathway complexities from the emission sources to the receiving water, and lack of reliable local measurements. The emissions can stem from point sources such as discharges by domestic and industrial sources in the form of direct and indirect discharge via wastewater treatment facilities (Ma et al. 2018), or from diffuse sources such as nutrient surpluses from agriculture and atmospheric re-deposition (Chen et al. 2017; Zhang et al. 2019). The various initiatives to collect data on emissions to water often support specific objectives and do not provide a complete picture. Thus, developing a reliable emission inventory is of great importance for designing water pollution control measures.

A water pollution emission inventory (WPEI) represents the pollutants discharged from various types of discharge sources into water bodies in a certain geographical area within a specified time span. WPEI can provide a scientific
basis to understand the causal relationship between emission and water quality change. During recent years, considerable progress was made to understand and quantify the sources of emissions at the local, regional, and global scale. In the European Union, the Water Framework Directive solicits all member states to set an inventory of the emissions of priority substances for all river basin districts (European Union 2008). Based on such inventories, it is possible to make useful evaluations such as the source apportionment, temporal trends, or pressure indicators. In the USA, Section 303(d) of the Clean Water Act authorizes the Environmental Protection Agency (EPA) to assist states, territories, and authorized tribes in listing impaired waters and developing total maximum daily loads for the water bodies, with an attempt to identify the sources of pollution and allocate pollutant loads in places where water quality goals are not being achieved (Reckhow 2001). The National Waters Inventory in Scotland consists of surface and groundwater samples gathered over two years and analyzed using novel analytical approaches to monitor water quality (Fozzard et al. 1999). The National Pollutant Inventory in Australia explicitly aims to maintain and improve water quality. Industrial facilities (above certain limits/thresholds) estimate their emissions and transfers annually using a variety of techniques and report to EPA Victoria and other jurisdictions (Clarkson et al. 2011). The International Environmental Database system is designed to house an emission inventory for a region. This region could be as big as a country or as small as a single complex source. Its primary intent is to provide a structure for an emission inventory for a metropolitan area (IDE 1995).

Compared with developed countries, China still has several deficiencies and limitations in establishing WPEI. It lacks a unified, appropriately accurate, and standardized inventory estimation method. The majority of the existing WPEI construction methods follow models from developed countries (Bierman et al. 2011), which are of low feasibility to actual Chinese conditions of data and institutional arrangements (Hawkins et al. 2019; Lian et al. 2020). Owing to the lack of investigation on the waterborne pollution-related inventory, the uniform model standardization system has not yet been established in China (Zhou et al. 2008; Wang et al. 2015). The incomplete system limits the full applications of these models to environmental management as there are no references and comparisons among different modeling results. However, there are several possibilities for the spatio-temporal characteristics of pollution sources (Zhang & Chen 2011; Cao et al. 2018). WPEI calculation and validation methods are distinctly different regarding the research objects, application scope, calculation resolution, and application difficulty (Mankin et al. 2003). Therefore, a simplified formalized method is needed to promote the development of WPEI at a water basin scale across the country.

This study summarizes the WPEI development process, focusing on the sector-based configurations and databases of the underlying coefficients. Afterward, the calculation, validation, and spatio-temporal resolution of emissions contained in current inventories are highlighted. Finally, specific suggestions for future WPEI development are presented.

**BASIC CONFIGURATIONS OF POLLUTION SOURCES**

**Industrial emission**

Industrial pollution is the contamination of the environment by factories or plants that dump their water products into the water. Industries discharge wastewater to waterbodies either directly without any treatment, with some preliminary treatment like a septic tank, or via advanced wastewater treatment facilities. Direct and indirect are the two main ways for industrial wastewater discharge (U.S.EPA 1972). According to their discharge patterns, industries with direct discharge belong to the industry inventory, whereas indirect discharge is classified as plant inventory. Industrial wastewater is mainly discharged in the form of point, sometimes through the soil in the form of non-point. The temporal pattern of discharge could be continuous or intermittent if the enterprises discharge while running or after a period of operation. In the current research, the technical manual covers most sub-sectors, and technical processes of industries account for most of the industrial pollutant discharge (MEE 2017). The export coefficient is the crucial parameter to conduct WPEI. The coefficient refers to the statistical average value of pollutants discharged per unit time under a particular land use mode or unit product.
production, and could be determined based on a large number of field experiments or statistical data (Johnes 1996; Zhou et al. 2017). The export coefficient of pollutants could be selected according to sectors and processes in the manual. Meanwhile, on-site monitors have gradually been installed in industries; thus, export coefficients could be obtained directly (Miner & Unwin 1991; Monaghan & Smith 2004; Laatikainen 2005).

**Domestic sewage**

Domestic wastewater is a term used to describe waste material from sewage waste that is collected in towns and urban areas and treated at urban wastewater treatment plants, and also that coming from single houses in the countryside, whether it is treated on-site in either septic tanks or individual wastewater treatment systems, or not treated (U.S. EPA 2020). The emission of domestic wastewater is closely related to the residents and treatment facilities (MEE 2017; U.S. EPA 2020). By monitoring or searching export coefficients and combining with the water discharge data, the load could be roughly estimated (Chen et al. 2006; Yue et al. 2007; Liu et al. 2010). Meanwhile, domestic wastewater covered with pipe network is discharged from sewage plants continuously in point. Various wastewater treatment facilities play a dominant role in the inventory configuration (Kim et al. 2012; Goffin et al. 2018). Therefore, all types of wastewater infrastructures shall be investigated in detail, including centralized or decentralized urban and rural domestic sewage systems (Smith et al. 2018). With the data of pollutants discharge or the removal efficiency of treatment methods, load for treated domestic wastewater could be obtained. However, the serving area of sewage and the amount of treated wastewater in centralized domestic sewage systems could not be identified well in current research. Hence, the load calculation result of domestic wastewater has considerable uncertainty.

**Agricultural outflow**

Agricultural pollution refers to liquid wastes from all types of farming activities, including runoff from pesticide and fertilizer use, erosion and dust from plowing, and crop residues and debris (OECD 1997). It is the dominant contributor of non-point source pollutants into rivers and lakes (U.S. EPA 2001). Management practices and rainfall play crucial roles in the amount and impact of these pollutants (Johnes 1996; Drolc & Zagorc Koncan 2002; Schaffner et al. 2011). In the research of agriculture pollution inventory, crop types and fertilizer application rate are usually taken as activity object to build the inventory, and total nitrogen (TN) and total phosphorus (TP) are usually the representative pollutants (Shi et al. 2002; Ding et al. 2010; Liu et al. 2010). In previous research, the inventories of agriculture have been explored and conducted with export coefficient methods (ECMs) or material flow analysis; however, the spatio-temporal resolution is quite low (Chen et al. 2006; Schaffner et al. 2011). The temporal resolution of agriculture is often year-based in most cases (Chen et al. 2006; Chen & Chen 2007; Liu et al. 2010), whereas the spatial resolution is usually based at a county level (Chen et al. 2006; Liu et al. 2010). Moreover, the route to water is seldom discussed, which could bring uncertainty about the location and quantity of pollutants into the river.

**Urban runoff**

Urban surface runoff pollution mainly refers to the effluent of rainfall on the urban surface. In terms of urban runoff source, research mainly focused on the characteristics of runoff pollution from the underlying surface, such as road and roof, and ribbon types like industrial area or residential area (Tang & Liu 2007; Zhang et al. 2019). Rainfall is a core driver for urban surface runoff pollution (Yue et al. 2007; Tang et al. 2014). Usually, research focuses on the pollution load caused by the rainfall of each session or annual rainfall (Chen et al. 2017). On-site measurement method, ECM, the spatio-temporal process-based accounting method such as the stormwater management model (SWMM), and statistical models are practical tools for calculating loads. In research using on-site measurement and ECMs, the coefficient of different underlying surfaces or ribbon types should be measured or selected from the literature. Except for this, rainfall and runoff coefficient data are also needed. SWMM, Hydrological Simulation Program-FORTRAN, and STORM are commonly used to simulate flow and water quality as process-based models, among which the first two models could simulate a continuous or single
rainfall process. Statistical models, such as EPA models and cumulative-scouring models, could also calculate the pollution loads, although they lack model explanations.

**Livestock and poultry breeding**

Livestock and poultry breeding waste is mainly derived from hosing out by water (or treated effluent) of livestock excreta (wet muck-out), or from washing livestock and livestock premises with water after dry muck-out. Aquaculture pollution is defined as the pollution caused by activities such as the farming of aquatic organisms, including fish, mollusks, crustaceans, and aquatic plants (Birgitta 1994). By the number of animals, breeding is usually divided into free-range and large-scale breeding (Zhou et al. 2017; Xie et al. 2018). According to Chinese environmental regulation, breeding units with less than 50 pigs, 5 dairy cows, 10 beef cattle, 50 sheep, 500 laying hens, or 2,000 broiler slaughters could be called free-large breeding, whereas others belong to large-scale breeding. Considering animal type, growth stage, and the method of cleaning up feces, the coefficients of pollutant discharge related to livestock, poultry, and aquaculture are different (MEE 2017). For large-scale breeding, the pollutants should be treated and discharged at outlets according to certain standards, whereas the pollutants of free-range breeding are usually discharged to farmland in diffuse form. Similar to the agriculture source, the route to water is seldom described in breeding research, which could bring uncertainty about the location and quantity of pollutants entering the river.

**Other sources**

Certain additional sources are gradually added into the inventory list, such as forest land and grassland (Cristan et al. 2016), atmospheric deposition (Yu et al. 2019), endogenous pollution (Zhu et al. 2018), and ship emission inventory (Iduk & Samson 2015). Forest land and grassland pollution is that in which the basin causes pollutants to enter the river owing to surface runoff and land erosion. The effects of land use, rainfall, soil properties, and slope on the forest land and grassland pollution are frequently discussed in the related research, with TN and TP being the two primary pollutants.

Atmospheric deposition sources introduce pollutants to the water environment. Different from the air pollution inventory, atmospheric deposition sources in WPEI mainly study the flux of pollutants deposited on the surface of rivers, lakes, and reservoirs. Nitrogen wet deposition observation research began in the 1850s, whereas nitrogen dry deposition research did not appear until the 1960s (Goulding et al. 1998). In contrast to these nitrogen studies, there are few studies on phosphorus deposition. Zhu et al. (2016) first systematically investigated the regional pattern of atmospheric phosphorus deposition in China based on on-site measurement data.

Endogenous pollution generally refers to the nitrogen and phosphorus nutrients accumulated in the surface layer of lake sediment re-entering the water under certain conditions (Guan et al. 2005). By monitoring water quality and experimentally researching the release mechanism of sediment pollutants and pollutant conversion, the endogenous pollution inventory is built and analyzed to control endogenous pollution (Liang et al. 2014; Pan 2017). The pollutants concerned in endogenous pollution are TN and TP, sometimes chemical oxygen demand (COD), and ammonia nitrogen (NH3-N).

Ship emission inventory typically focuses on domestic sewage produced during shipping, for the pollutants produced when shipping, petroleum, nitrogen, and phosphorus nutrients are the primary concerns (MEE 2018). Increased connectivity and international marine trade have stimulated inland development. The establishment of port-level and regional shipping emission inventories is an urgent priority (Li et al. 2017). The fuel-based method (Yang et al. 2007), trade-based method (Li & He 2011), and automatic identification system-based method (Liu et al. 2016b) have been used in recent studies to obtain emission inventories.

There are several types of pollution sources, and the generation and discharge characteristics of pollutants are quite different. Therefore, rarely any country has ever built a multi-source, full-caliber WPEI.

**Emission coefficients in Chinese contexts**

With more types of emission sources included in WPEI, there are more kinds of emission coefficient databases. In
the early days, owing to the limited data on localized emission factors, WPEI compilation mainly used American and European emission factor databases (U.S. EPA 1985). However, owing to the differences in production technology, weather, and geography, among others, between different regions, the production and emission coefficient could be vastly different. Without localized parameters, it can lead to considerable uncertainty in WPEI. In recent years, researchers have carried out much research on emission coefficient testing.

In China, for industry and plant emission coefficient, the Ministry of Ecology and Environment of the People’s Republic of China conducts a National Pollution Source Census (MEE 2017). The production and emission coefficient manual of industrial pollution source covers 562 sub-sectors, accounting for most of the national industrial pollutant discharge. Among them, the production and emission coefficients of 271 sub-sectors are calculated by on-site measurement, and the other coefficients are obtained by analogy. For industry and plant source, real-time water quality monitoring has been implemented throughout the country, and the emission coefficient databases have continuously been enriched and updated.

For domestic sources, the National Pollution Source Census (MEE 2017) and National Technical Guidelines for Water Environment Capacity Verification (Chinese Academy of Environmental Planning 2003) provide the emission coefficient. The coefficients of the National Pollution Source Census (MEE 2017) are obtained considering different regions, urban or rural areas, treatment, and discharge methods. However, the National Technical Guidelines for Water Environment Capacity Verification (Chinese Academy of Environmental Planning 2003) regulates the coefficients of the standard city, such as 60–100 and 4–8 g/capita·d for COD and NH₄-N, respectively, in urban areas, and 40 and 4 g/capita·d for COD and NH₄-N, respectively, in rural areas. For non-standard cities, the coefficients are corrected according to factors such as geography, population size, built-up area, rainfall, and rainwater collection coverage.

For agriculture, the farmland export coefficient and the fertilizer loss coefficient are the two most critical parameters. According to the National Technical Guidelines for Water Environment Capacity Verification (Chinese Academy of Environmental Planning 2003), the standard farmland export coefficient of COD and NH₄-N are 10 and 2 kg/ha·a, respectively. When the fields would not fully conform to the standard farmland definition, the coefficient should be adjusted considering the impact of the slope, crop types, soil types, fertilizer application, and rainfall. Moreover, in the National Pollution Source Census (MEE 2017), the fertilizer loss coefficients are obtained based on the division of agriculture and dominant agriculture products in China. Through selecting a typical cropping system and landform of farmland, the fertilizer loss coefficients in the databases obtained by field monitoring cover the main agriculture areas, planting methods, farming methods, farmland types, soil types, topography, and main crops in China.

For livestock and poultry breeding and aquaculture pollution inventory, the National Pollution Source Census (MEE 2017) provides the production and emission coefficients of animals in different regions. The main division factors are the location area, types, feeding stages, feeding methods, and feces collection and processing methods. As for urban runoff sources, there is currently no national database, but several studies have conducted on-site monitoring. Table 1 lists the coefficients obtained in previous research.

Meanwhile, there are few relevant studies for the atmospheric decomposition, endogenous, and ship emission inventory. Atmospheric decomposition is mainly based on specific case studies, and few monitoring networks have been formed. The China Nitrogen Deposition Monitoring Network, a national nitrogen deposition monitoring network was established by China Agricultural University in 2004, including the deposition of farmland, grass, forest, and urban areas (Liu & Zhang 2009). The regional pattern of atmospheric phosphorus deposition in China based on measured data was first introduced in 2016 (Zhu et al. 2016). Based on the sedimentation flux data in 2013, the annual average precipitations of soluble nitrogen and phosphorus in atmospheric wet deposition in China are 15.69 and 0.21 kg/hm²·yr, respectively. The average nitrogen to phosphorus ratio (N:P) is 77, and the atmospheric phosphorus deposition flux is quite small.

The coefficients of endogenous inventory are mainly obtained through laboratory experiments. Pan (2017) investigated and evaluated the status of water environment and pollution source discharge of Beiyunhe River Basin, and found that the release rate for NH₄-N and phosphate were
1,136 and 145 mg/m²-d, respectively. Wang (2018) analyzed the pollution sources in the Niugu River in Dingxi, Gansu, and concluded that TN and TP in sediment were 102.77–2,293.94 and 322.20–704.80 mg/kg, respectively. For ship emission inventory, Discharge Standard for Water Pollutants from Ship (GB 3552-2018) (MEE 2018) regulates the export coefficients of sewage treatment devices installed or replaced after 2021, which discharge sewage into inland rivers (Table 1).

The configurations and databases of the underlying coefficients are classified and summarized by sectors. It is of great help for water quality management. Meanwhile, the current pollution emission inventory research focuses on single or several conventional pollution sources and lacks the construction of multi pollution source inventory. It goes against the comprehensive and systematic management of the basin water environment.

**METHODS OF WATER POLLUTION EMISSION INVENTORY**

As the sources in WPEI are quite diverse, there are various corresponding calculation methods and related spatio-temporal resolutions. Additionally, we discussed here the inventory validation.

**Calculation methods of emission inventory**

Owing to the significant differences in the database resolution, major pollutants, and emission characteristics of different pollution sources, the calculation methods for emission inventories are diverse, including the on-site measurement, export coefficient, and material balance method, mechanism model, among others. For the holistic description of the calculation method, the main steps are the emission factors acquisition using methods of investigation or experiment and emission load estimation and activity data of the pollution sources (Kanemoto et al. 2006). Specifically, the data requirements, research objects, applicable scopes, and spatio-temporal resolution of various methods differ. Herein, we introduce the research objects, main processes, and applicable scope of four conventional methods.

**On-site measurement method**

The on-site measurement method uses the pollutant production or emission data obtained from on-site measurements to calculate the amount of pollutant production or emission per unit time. The first category is the manual measurement, which is a standard method for the calculation of emissions from point sources, endogenous pollution, and ships. Kim et al. (2012) investigated the discharge of perfluorinated compounds in 15 large-scale wastewater treatment plants in Korea and calculated an average discharge load of 0.04–0.61 tons/year. Boers & Hese (1988) pioneered the phosphorus release patterns of sediments in Loosdrecht Lakes with a columnar sediment simulation experiment and qualitatively explained the effect of temperature, infiltration, and pH on the phosphorus release rate of sediments. However, manual measurement has the significant drawback of limited monitoring frequency and range.

---

**Table 1 | Summary of runoff coefficients for different types in Chinese studies**

<table>
<thead>
<tr>
<th>Runoff Source</th>
<th>COD (mg/L/a)</th>
<th>NH₄-N (mg/L/a)</th>
<th>TN (mg/L/a)</th>
<th>TP (mg/L/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban runoff</td>
<td>44.5–1420.0</td>
<td>0.42–18.5</td>
<td>6.4–11.2</td>
<td>2.4–78.8</td>
</tr>
<tr>
<td>Roof (mg/L)</td>
<td>16.4–415.7</td>
<td>5.76–31.4</td>
<td>2.8–19.9</td>
<td>0.08–65</td>
</tr>
<tr>
<td>Forestland runoff (kg/m²·a)</td>
<td>–</td>
<td>–</td>
<td>0.03–12.8</td>
<td>0.002–1.5</td>
</tr>
<tr>
<td>Grassland runoff (kg/m²·a)</td>
<td>–</td>
<td>–</td>
<td>0.06–11.5</td>
<td>0.01–1.55</td>
</tr>
<tr>
<td>Limits from ship (mg/L)</td>
<td>60</td>
<td>15</td>
<td>20</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Notes: *for the road runoff coefficients studies refer to Ren et al. 2005; Dong et al. 2008; Zhang & Li 2008; Chen 2011; Chen et al. 2011; Yang et al. 2014; Chen et al. 2017; Wu 2018.


*for the forest export coefficients studies refer to Shi et al. 2002; Ding et al. 2007; Liu et al. 2011; Yang et al. 2015.

*for the grass export coefficients studies refer to Shi et al. 2002; Ding et al. 2007; Liu et al. 2011; Yang et al. 2015; Wang 2016.

*for the coefficients of sewage treatment devices studies refer toShi et al. 2002; Ding et al. 2007; Liu et al. 2011; Yang et al. 2015; Wang 2016.

To improve the inventory accuracy, online monitoring has increasingly been applied which significantly contributes to the continuous temporal analysis and precise calculation of key point sources. Hoppe et al. (2009) used spectrometer probes and mobile flowmeters to monitor industrial drainage continuously, ensuring that countermeasures could be quickly started when a mistake occurs. Goffin et al. (2018) obtained a predictive model of soluble five-day biochemical oxygen demand by using three-dimensional fluorescence spectroscopy combined with parallel factor analysis (PARAFAC). The model applies to wastewater treatment for online monitoring. Although online monitoring data considerably improves inventory accuracy, it has almost no application in the compilation of non-key and non-point source inventories.

**Export coefficient method**

ECM was developed to construct the inventories of pollution sources that cannot be calculated directly by the on-site measurement method (Johnes 1996). The critical step of the method is to select the pollution production coefficient and export coefficient according to the geographical location, scale, and characteristics of pollution sources, and treatment process. Moreover, the production and discharge coefficients are combined with the pollution source activity data to calculate the production and discharge load of pollutants. ECM has been recognized as reliable for modeling non-point source pollution (Ierodiaconou et al. 2005). White & Hammond (2009) used this method to evaluate phosphorous emissions in the Waters of Great Britain and observed that the pollution load in decreasing order corresponded to households, agriculture, and industry. Liu et al. (2008) estimated the TN and TP load discharged from non-point sources in the Yangtze River Basin, showing that Sichuan Province contributed the most to the water basin pollution.

Nevertheless, the classic ECM has some limitations. For example, this method uses the same export coefficient in different hydrological years and regions, without considering the uneven spatio-temporal distribution of precipitation and spatial differences in water basin topography and soil conditions, which influence coefficients. Kaur et al. (2017) found that the emission coefficients of pollution sources in different river basins were significantly different. For different cases, using a unified ECM will lead to large error results. In recent years, some researchers have improved ECM in response to these defects, which are known as Improved Export Coefficient Method (IECM). In the study of non-point source pollution in the Yangtze River, two factors of precipitation and terrain were introduced into the IECM model, and the average export coefficient of the large watershed was replaced by the spatial heterogeneity coefficient of sub-watersheds. Results showed that, compared with the actual pollution flux, the relative error of dissolved nitrogen flux was reduced from 41% to 24% (Ding et al. 2010). Another improved export coefficient model was established to study the influence of terrain on the calculation results in the Haihe River Basin. The plain area presents a higher pollution level than the mountainous area (Qiu et al. 2012). ECM (IECM) is widely used in the construction of non-point source inventories, solving to a certain extent the lack of measured data for pollution sources.

**Mass balance method**

With the updated statistical and data acquisition methods, the available data on the production activities of water pollution sources have significantly increased. The defects of large errors in ECM calculation results are recognized, and it is necessary to develop a more accurate calculation method based on pollution source activity data. The mass balance method can quantitatively analyze the changes of materials used in the production process according to the principle of conservation of material quality. In the production process, the amount of materials entering a system must be equal to the sum of materials discharged and accumulated in the system. Based on this principle, several models for the quantitative analysis of pollutant input in water basins have been developed in recent years. Strokal et al. (2016) developed a MARINA Nutrient Model that assesses the input of river nutrients to the ocean based on the NEWS-2 (Nutrient Export from Watersheds) model. Yang et al. (2009) applied MARINA to lakes and reservoirs and proposed various policy recommendations based on the model calculation results. Li et al. (2009) combined the NEWS-2 model and NUtrient flows in Food chains,
Environment and Resources use the model to assess the impact of agricultural transition on water quality in China’s major provinces.

Mathematical Material Flow Analysis (MMFA) is another typical and commonly used material balance method. Based on the unit and activity balance, MMFA quantitatively calculates the discharge load by analyzing the sources, sinks, and flows of various pollution sources. In existing research, MMFA is used to systematically explore the contributions of various pollution sources in the large-scale water basin, including point sources and non-point sources. Drolc & Zagorc Koncan (2002) presented and applied MMFA to develop a phosphorus balance for simulating and evaluating loads in different pollution reduction scenarios in the Krka River. Schaffner et al. (2011) analyzed the sources and flow paths of nitrogen and phosphorus in the Thachin River Basin and quantitatively identified aquaculture (as a point source) and rice farming (as a non-point source) as vital pollution sources. Moreover, Schmid Neset et al. (2008) analyzed the flow of phosphorus and found that the increased input of chemical fertilizer was a prominent cause of water eutrophication.

Spatio-temporal process-based accounting method

MMFA can be used to link the production, pollution production, and discharge processes of pollution sources, but it is still limited by production data, and only the annual inventory can be calculated generally. For more accurate calculations, spatio-temporal process-based accounting models are used. Using spatio-temporal process-based accounting models, inventory calculation results can be accurate to hydrological response units. Regarding time accuracy, by inputting high-precision hydrological and water quality data, monthly and daily inventories can be constructed. Moreover, the models are often used to simulate the complex process from pollutant production to river entry, which cannot be achieved by other methods.

Spatio-temporal process-based models have been increasingly used as essential tools for calculating non-point source loads, including SWAT (Arnold et al. 1998), HSPF (Becknell et al. 1995), AGNPS (Young et al. 1989), and ANSWERS (Beasley et al. 1980). In recent years, the research progress of China’s watershed-based models has accelerated, and certain models for specific water basins have been developed based on classic models by many Chinese researchers. For instance, Wang et al. (2011) developed a watershed-based model to estimate non-point source pollutant loads in a large-scale water basin (ENPS-LSB), including agricultural fields, rural residential areas, urban areas, and livestock. Lai et al. (2018) modified the non-point source pollution simulation module of the SWAT model, and analyzed the spatio-temporal characteristics of nitrogen and phosphorus pollution in the Poyang Lake area based on the modified model.

Overall, there is a considerable variety of WPEI calculation methods. Table 2 summarizes the data requirements, pollution sources, practical feasibility, spatio-temporal resolution, and practical application costs of each calculation method. Among them, ECM is presently the most widely used method for calculating the amount of pollution emission in a water basin owing to its ease of application, low data requirements, and acceptable resolution of results.

Overall, with the improvement of data quality and measurement technology, calculation methods have gradually improved their result accuracy and increased the number of pollution source categories included in the calculation scope. However, accurate calculation methods usually require sophisticated modeling and calculation processes, and their feasibility is correspondingly low. In the

<table>
<thead>
<tr>
<th>Data requirements</th>
<th>Source variation</th>
<th>Practical feasibility</th>
<th>Spatio-temporal Resolution</th>
<th>Application costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>On-Site Measurement</td>
<td>M</td>
<td>PS; OS</td>
<td>M</td>
<td>H</td>
</tr>
<tr>
<td>ECM/IECM</td>
<td>L</td>
<td>PS; NPS; OS</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>MMFA</td>
<td>H</td>
<td>PS; NPS</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>Process-based Models</td>
<td>H</td>
<td>PS; NPS; OS</td>
<td>L</td>
<td>H</td>
</tr>
</tbody>
</table>

Notes: PS, point sources; NPS, non-point sources; OS, other sources; L, M, and H correspond to low, medium, and high, respectively.
issues of the spatio-temporal resolution

In recent years, WPEI with high spatio-temporal resolution has been increasingly considered. For the atmospheric pollutant emission inventory, near-real-time estimates have been developed (Zhang et al. 2021b), and bulk emission inventories have been constructed based on 1 x 1 km grids (Zhou et al. 2021). Compared with the air pollutant emission inventory, the spatio-temporal resolutions in water pollution-related inventories are less elaborated. Moreover, the spatio-temporal resolution of WPEIs calculated by different monitoring techniques, calculation methods, and allocation methods varies greatly.

Resolution improvement based on advanced monitoring technology

In most current studies, annual emissions inventories are calculated based on statistical data. However, advanced monitoring technologies, including online monitoring and the internet of things, have significantly improved the temporal resolution emission inventories for point source and other pollution sources. For instance, with the application of new monitoring technology, Hoppe et al. (2009) used online monitoring data to obtain real-time emissions of industrial pollution. Wu et al. (2015) developed an online monitoring device combined with GIS and Java for online monitoring of the effluent quality of sewage plants, providing continuous real-time data for the management department. Zhang & Davison (1999) applied techniques such as differential equilibrium and differential gradient in thin-films to the in-situ online monitoring of internal sources in water, which has been an important data source for the study of endogenous pollution inventory. Regarding spatial resolution, accurate latitude and longitude coordinates of point sources can be obtained based on monitoring tools such as online maps and satellite data. Advanced monitoring technologies have improved the spatio-temporal resolution of inventories while making up for the lack of data. However, owing to differences in regional economies and research levels, the maturity and application status of monitoring technologies vary greatly. The quality of data obtained through different data acquisition methods is uneven, resulting in a non-uniform spatio-temporal resolution of inventories in various regions or water basins.

Resolution differences based on multiple calculation methods

It is evident that the calculation accuracy of different methods varies (cf. Table 2), resulting in a significant difference in WPEI spatio-temporal resolution. Owing to the variety of calculation methods, the non-point source pollution source inventories have significant spatio-temporal resolution differences.

Non-point source WPEIs use ECM that can only achieve low spatio-temporal resolution as the conventional calculation method. Sivertun & Prange (2005) calculated the annual non-point source pollution load of large-scale Gisselö watershed. Cai et al. (2004) propose a classic improved ECM considering the impact of rainfall and pollutant migration loss to estimate the load of non-point sources, and only calculate the annual WPEI at the water basin scale. In recent years, different calculation methods and coefficients have brought vast differences in spatio-temporal resolution. Wang et al. (2012) used various multivariate statistical techniques to analyze spatio-temporal characteristics of point and non-point pollution sources in Qinhe River Basin and identified pollution characteristics in different water periods. Yue et al. (2007) used the binary structure model combined with RS and GIS technology to calculate and verify the monthly load of agricultural non-point source pollution in the Songhua River Basin. Barber (2014) used in-situ monitoring equipment to provide high-resolution (50 minutes) data, and calculated agricultural non-point source WPEI (TP) in the River Eden. The development of calculation methods has improved the spatio-temporal resolution of water pollution emission inventories, but the diversified methods have led to an increase in the unevenness of the spatial and temporal resolution of different inventories.

Emission allocation methods

To meet the spatio-temporal resolution requirements of the aquatic environment management in water basins, research
on the spatio-temporal allocation methods for WPEI is increasing. However, there are few studies on the temporal allocation of WPEI. For point sources, two types of methods are used in the present study for temporal allocation of inventories: average allocation method (AAM) and production factor allocation method (PFM). AAM is a rough method, and the monthly inventories are obtained by distributing the calculated annual load equally to each month (Qiao et al. 2010). Owing to the evident advantages of the simple calculation process, AAM is widely used in the inventories of large-scale industrial enterprises, sewage treatment plants, and other point sources with stable emissions throughout the year. PFM refers to the distribution of annual emission loads to months or even days with specific point source emission data as the weighting factor, e.g., to distribute loads to days according to the raw material consumption and product output of industrial enterprises (Vongmahadlek et al. 2009). This method can be used to calculate various high-precision point source inventories with production activity data but is constrained by data and rarely used presently. For the temporal allocation method of non-point sources, in addition to AAM, the driving factor allocation method (DFM) is mainly used for some pollutants with stable emissions such as domestic sewage. For instance, assuming a positive correlation between non-point source pollutant discharge and rainfall, the annual inventory is divided into monthly inventories according to the proportion of monthly rainfall (Yue et al. 2007). Compared with AAM, DFM considers the impact factor of the discharge load, which is more reasonable. However, it ignores the effects of pollution storage and delayed river entry, and the distributed inventory can only achieve monthly resolution.

Additionally, with the increasing demand for grid management of the water environment in recent years, spatial allocation is considerably significant. There are mainly two allocation methods for gridding emission inventories. The spatial allocation method based on raster data refers to the total load allocation based on grid attribute information, such as precipitation, digital elevation, land use, and land-cover. Lai et al. (2020) subdivide the discharge load of agricultural non-point source nitrate-nitrogen (NO3-N) into 1-hectare grids with different slopes. Velthof et al. (2009) propose an agricultural non-point source runoff coefficient calculation model that takes precipitation, slope, soil texture, and soil depth as elements to achieve grid allocation of agricultural non-point sources. Moreover, spatial allocation based on station monitoring values is another downscaling method. Simkin et al. (2016) evaluated the annual flux of atmospheric nitrogen deposition at 15,136 monitoring points in the United States and performed spatial interpolation to explore the effect of nitrogen deposition on species richness. The China Ecological Research Network has established a nationwide atmospheric deposition network based on the measurement data of 41 sites, dividing the atmospheric deposition flux into a grid of 1 × 1 km (Yu et al. 2019). In contrast, the spatial distribution method based on raster data is widely used, fully considering the attributes of each grid, but verifying the accuracy of the distribution results is difficult. The interpolation method based on monitoring values is only suitable for individual pollution sources that are easy to observe. In theory, the distribution result with enough measurement points and data is accurate, but the measurement is costly and difficult to apply practically.

Figure 1 presents the variation of WPEI spatio-temporal resolution in the last ten years by overviewing approximately 80 inventories. In recent years, the spatial resolution of each water pollution emission inventory was still mainly at the water basin scale, and annual inventories account for the vast majority. However, compared with those in 2010–2014, the resolution of inventories in 2015–2019 has significantly improved. The proportion of grid-based or county-based inventories increased from 2.4% and 22.0% to 10.7% and 21.4%, respectively. Although the temporal resolution improvement was not obvious, the proportion of daily inventories still increased from 9.8% to 14.3%. Overall, owing to the large differences in data quality, methods, and parameters selected among different inventories, the current spatio-temporal resolution of water pollution emission inventories is still not uniform.

Multi-dimensional validation of emission inventory

WPEI validation mainly depends on the information about water quality. In recent research, three methods are adopted to validate the WPEI, which include multivariate statistical technologies, the inverse verification of pollution load
using the water quality models, and source apportionment by the receptor models.

**Multivariate statistical technologies**

Multivariate statistical technologies, such as cluster analysis (CA), discriminant analysis, principal component analysis (PCA), and factor analysis (FA), are used for the evaluation of both temporal and spatial variations and the interpretation of large and complex water quality datasets. Through collecting water quality information, the water pollution status and its spatio-temporal variation can be assessed by multivariate statistical analysis (Zhang et al. 2014). By comparing the spatio-temporal characteristics of WPEI and water quality with correlation analyses, the inventory is verified if the relation between the two correlates well. Pekey et al. (2004) used a FA and multiple regression analysis to estimate the contributions from identified sources to the concentration of each parameter. Zhou et al. (2007) evaluated the spatio-temporal patterns and source apportionment of coastal water pollution in eastern Hong Kong based on CA and PCA. Moreover, Song et al. (2019) estimated contributions from identified pollution sources to each water quality variable at each monitoring site with FA and PCA.

**Water quality modeling**

Based on the water quality information and physical and chemical processes of pollutants in the river, water quality models are applied to describe the transportation and transformation processes of pollutants. Certain water quality models (e.g., Qual2E, SIMCAT, WASP, and SALMON_Q)
are specifically intended for the steady-streamflow and the steady-effluent discharge conditions specified in the water quality regulations for waste load apportionment and allocations. After monitoring and collecting water quality data at a specific time, models can trace the main pollutants at different sites. The accuracy of WPEI can be evaluated by comparing the load of WPEI with the pollution load flux calculated by water quality models. Moreover, the contribution of the pollution source to water can be calculated for any river location as well as the contribution to the monitoring water quality at any time. In recent research, water quality models and statistical analysis are often combined for WPEI apportionment and validation. Azzellino et al. (2006) combined a water quality simulation model (Qual2E) and an FA to increase the understanding of the water pollutants source apportionment, and found that non-point sources contributed approximately 80% in the area of extensive agricultural land use. Qiu et al. (2017) applied multivariate statistical analysis and Qual2Kw to investigate the spatio-temporal variation of water quality and pollutant sources in the Huanghai River.

**Receptor-based models**

Receptor-based models, such as absolute principle component score-multiple linear regression (APCS-MLR), and positive matrix factorization (PMF), can be used for pollution source identification. In recent years, the application of these techniques to apportion pollution sources in the water environment has been strengthened. The receptor-based models could infer the contribution of various sources based on physicochemical information acquired by point sampling. Singh et al. (2005) applied an APCS-MLR combined with CA and other methods to explore the main contributors to water pollution. By using multivariate statistical analysis, three catchments regions and seven potential factors were identified based on the similarity between them. The primary sources, i.e., soil weathering, leaching, and runoff, municipal and industrial wastewater, and waste disposal sites leaching, were found by APCS-MLR. Haji Gholizadeh et al. (2016) assessed water quality and identified and quantified the potential pollution sources affecting it by PCA, FA, APCS-MLR, and PMF. Research results showed that the discharge of agriculture waste and domestic industrial wastewater were the significant sources of river water contamination. Jiang et al. (2019) used a combined approach of hierarchal cluster analysis and PMF to identify non-point pollution for the Huaihe River Basin in China. It was found that discharge inputs from urban, agriculture, and industrial land uses are associated with the major causes of Huaihe River water quality deterioration.

In general, multivariate statistical technologies and receptor-based models reveal the spatio-temporal characteristics in specific time and space, in non-continuous scale. In contrast, the water quality modeling could reflect the status of rivers at continuous spatial and temporal scales. Therefore, appropriate validation methods should be chosen according to the different application scenarios and needs. However, in the current basin environment management system, a unified and standardized inverse method selection system or method priority has not yet been formed.

With the diversification and refinement of monitoring methods and calculation methods, a variety of pollution sources is included in the WPEI, the accuracy of the results and spatio-temporal resolution have been improved. In addition, based on diversified validation methods, it is possible to evaluate the reliability of WPEI, and to a certain extent ensure the practical feasibility of WPEI. However, different databases, diverse methods, and inconsistent time-space accuracy have led to large differences in WPEI proposed by different regions and institutions. It is difficult to unify the inventories of different pollution sources completely. Therefore, the meaningful work at the current stage is to build a set of reasonable and standard method selection systems to form a relatively unified WPEI working procedure. At the same time, the method selection system should be restricted by actual water environment management requirements.

**DISCUSSION**

**Development procedure of the water emission inventory**

Currently, there is no systematic and complete procedure for WPEI construction. It is urgent to build a standardized
inventory procedure. According to previous research on WPEI and the construction process of the atmospheric emission inventory, the construction procedure of WPEI should consider the pollution source characteristics. The procedure should contain at least the following steps: identifying and coding pollution sources, determining the proper methods of load calculation for different sources, validating the inventory results, and analyzing their uncertainty. Collecting information and combining with their characteristics, potential local sources could be identified. Subsequently, to improve management, pollution sources could be coded by the pollution source types, modes of generating pollution, pollutant treatment facilities, and pollutant emission paths. Proper methods could be selected based on data availability, calculation ability, and economic situation. As there are several types of pollution sources, a moderate spatio-temporal resolution is of great importance in the whole WPEI. Validation is the crucial step to ensure the reliability of results. During the process of inventory construction, the identification of pollution sources, spatio-temporal description of pollutant emissions, the generalization of emission paths, data input, and simulation process have some differences from the real world, which are then passed to the inventory results, resulting in the inventory uncertainty. Usually, the uncertainties caused by data input and export coefficients are the research hotspot. The data input or parameters that bring more significant uncertainty to the results could be identified. Then, the researcher could improve data quality and reduce result uncertainty by field monitoring and other means. Overall, constructing a comprehensive and standardized procedure is necessary for WPEI development.

**Moderate spatio-temporal resolution for the emission inventory**

The moderate spatio-temporal resolution of WPEI refers to the inventories that could reveal the characteristics of emissions distribution accurately. It is an essential tool for water environmental management. The resolution of WPEI should be consistent with the requirements of management. Basins where water pollution problems are serious and emergencies occur frequently are encouraged to construct a high spatio-temporal resolution WPEI. It is of great benefit to identify pollution sources and control emissions precisely without delay. While for developing areas, data availability is constrained. To reduce the uncertainties of WPEI, WPEI should build connections with the surface water quality based on the available data. From the research perspective, the resolution of WPEI should be selected with the target of emission inventories, such as ‘Large-Scale WPEI’ or ‘Local High Spatio-Temporal Resolution WPEI’. ‘Large-Scale WPEI’ focuses on identifying the general distribution characteristics of emissions. Data can be obtained from statistical departments or associations to establish a parameterized scheme for regional averages. The estimate of total load is generally accurate. In terms of ‘Local High Spatio-Temporal Resolution WPEI’, attention should be paid to parameterize the dynamic factors and simulate the dynamic process of emissions. Data sources are diverse because a single data source hardly meets the requirements of a high-resolution inventory. At this resolution, the emission hot spot is prominent, and the emission gradient around the hot spot is larger. Therefore, the balance of WPEI resolution, technology, data, and costs should be kept considering the combination of management and research needs.

However, the moderate spatio-temporal resolution is not equivalent to high spatio-temporal resolution to some extent. Some research identified that the high spatio-temporal resolution inventories, which were built by untypical inaccurate spatial proxy parameters, would lead to distorted emission inventory (Zheng et al. 2017; Oda et al. 2019). Therefore, the moderate resolution of WPEI should be decided by multi factors, such as the research topics, available data, and construction technologies.

**Validation of the emission inventory**

Validation is an essential part of the whole inventory. However, it mainly focuses on comparing spatio-temporal characteristics between WPEI and water quality through correlation analysis, and the high amount of specific sources could not be validated and adjusted. Pollutant characteristics and concentration from different sources are significantly different. Pollution sources discharge pollutants at various times, from different locations, and in various manners. Therefore, validation could be made from simple reaches to complex rivers, in different periods and at various positions in the same
period. First, the emissions that account for relatively small parts of the pollution load are ignored during the validation. Validation from simple to complex refers to selecting reaches with fewer types of pollution sources or single discharges for verification, and gradually extend to the entire river. The temporal characteristics of emissions are unique for each source, such as pollutants of sewage plants being continuously being discharged throughout the year, whereas those of agriculture are only emitted after irrigation or rainfall. Therefore, validating WPEI by period is a practical method. Similarly, in the same period, the pollutants of sources are poured into the river at different locations, so validation could be made by reaches and locations.

Additionally, to the numerical analysis discussed above, some other emerging techniques in other fields could also be applied in WPEI validation, such as the three-dimensional excitation-emission matrix (3DEEM) (Cheng et al. 2018). EEM has a one-to-one correspondence with the water body, which is also called ‘fluorescence fingerprint.’ The fluorescent characteristics of the EEM are unique by water body types. Some research even established a fingerprint library of water bodies such as sewage wastewater, wastewater of pharmaceutical factories, among others (Wang et al. 2015; Wu et al. 2011, 2016). Comparing the fluorescence spectrum analysis results of the water sample with the fingerprint library could identify the main contributors to water pollution, and compare them with the calculation results of the inventory, which has also been used for inventory validation in recent years.

**CONCLUSIONS**

Based on the efforts made in previous studies, we have increased the knowledge about emission characteristics of certain key sources, including point and diffuse sources. Additional reliable methods of calculation and validation have been used to conduct a water pollution emission inventory.

Further efforts are required to conduct a systematic procedure for the emission inventory, which should contain certain necessary sections, such as identifying and coding pollution sources, determining the proper methods of load calculation for different sources, validating the inventory results, and analyzing the uncertainty of results. Under a unified framework, the inventory results for different regions could be used for comparative analysis.

Moderate spatio-temporal resolution of inventory is promoted in research and management. Research on moderate resolution has been carried out in atmospheric pollutants emission inventories, while there are rare discussions in water pollution emission inventory. It is recommended to carry out relevant research on the relationship between reliability and spatio-temporal resolution of water pollution emission inventory.

A comprehensive validation of the emission inventory is another important issue to be addressed. Except for the traditional methods of comparing the spatio-temporal characteristics of the emission inventory and water quality through correlation analysis, validation of the number of specific sources could also be made from simple reaches to complex rivers, in different periods and at various positions in the same period. Alternative techniques could additionally be applied in the validation of emissions inventory, such as the 3DEEM. With wide application and independent validation, the uncertainties of emission inventories can be identified and further reduced.

**ACKNOWLEDGEMENTS**

The authors would like to thank the Major Science and Technology Program for Water Pollution Control and Treatment for financially supporting this research under Contract No. 2018ZX07111004. We greatly appreciate the helpful comments from the anonymous reviewers.

**DATA AVAILABILITY STATEMENT**

All relevant data are included in the paper or its Supplementary Information.

**REFERENCES**

Chen, M. & Chen, J. 2006


Barber, N. 2014

Designing schemes to mitigate non-point source water pollution from agriculture: the value of high-resolution hydrochemical and hydrophysical data. Agu Fall Meeting. AGU Fall Meeting Abstracts.


Bierman, P., Lewis, M., Ostendorf, B. & Tanner, J. 2011


Birgitta, L. 1994

Three Overviews on Environment and Aquaculture in the Tropics and Sub-Tropics.

Boers, P. C. M. & Hese, O. v. 1988


Cai, M., Li, H., Zhuang, Y. & Wang, Q. 2004


Chen, Y. 2011

Polution Characteristics and Control Technology of Urban Road Runoff in Xi'an City. Chang'an University.

Chen, M., Chen, J. & Lai, S. 2006


Chinese Academy of Environmental Planning 2003

National Technical Guidelines for Water Environment Capacity Verification.

Clarkson, P. M., Overell, M. B. & Chapelle, L. 2021


Ding, X., Shen, Z. & Liu, R. 2007


Ding, X., Shen, Z., Hong, Q., Yang, Z., Wu, X. & Liu, R. 2010


Drolc, A. & Zagorc Koncan, J. 2006


European Union 2008


Fozzard, I., Doughty, R., Ferrier, R. C., Leatherland, T. & Owen, R. 1999


Goffin, A., Guérin, S., Rocher, V. & Varrault, G. 2018


Haji Gholizadeh, M., Melesse, A. M. & Reddi, L. 2016

Water quality assessment and apportionment of pollution sources...


U.S.EPA 2001 Non-point Source Program.


Wang, D. 2016 Non-Point Source Pollution Load Simulation and Water Quality Assessment of the Three Gorges Reservoir. Xinan University.

Wang, S. 2018 Analysis of the Current Situation of River Pollution in Arid and Semi Arid Region Countermeasures-the Niugou River in Dingxi City Gansu Province as an Example. Guanxi University.


Wu, Y. 2018 Characteristics of Runoff Pollution of Different Underlying Surfaces in Cultural and Educational Area of Xi’an. Chang’an University.


Yang, L., Sun, C., Qi, J., Liu, G. & Wang, Y. 2014 Study on the dynamic update of urban surface runoff pollutants load. Environ. Sustainable Dev. https://doi.org/10.19758/j.cnki.issn1673-288x.2014.03.042


Zheng, B., Zhang, Q., Tong, D., Chen, C., Hong, C., Li, M., Geng, G., Lei, Y., Huo, H. & He, K. 2017 Resolution dependence of


First received 13 July 2020; accepted in revised form 5 October 2020. Available online 2 November 2020