RESEARCH ARTICLE

A bi-level multiobjective optimization model for waste load allocation in rivers

Xuan Zhang[1] · Jungang Luo[1] · Jiancang Xie[1]

Abstract

It is often difficult to apply existing waste load allocation (WLA) models to management institutions at all levels of the river basin because the existing WLA models do not consider the principles of fairness and efficiency at each management level of the basin. The implementation of environmental protection tax law has also greatly impacted WLA. This paper proposes the bi-level multiobjective allocation model under an environmental protection tax law to solve the WLA problem for different management levels. The upper allocation targets the minimal environmental Gini coefficient and the minimal unit pollutant emission cost. The impact of the environmental protection tax is also considered. The targets of the lower-level allocation are the maximal industrial output value and the minimal unevenness of reduction rates. The proposed model was applied to the case of the Wei River basin, and the results demonstrated that the bi-level multiobjective allocation model could solve the problem of WLA under an environmental protection tax law. Each level of the bi-level multiobjective allocation model considers the principles of fairness and efficiency to distribute the load in the basin, thereby offering a better reference for decision-makers at both levels.

Introduction

Rivers have recently become the most dominating sewage receivers, and water pollution has become one of the severest worldwide environmental issues (Zhang et al. 2012a). Due to the discharge of industrial wastewater and municipal sewage, water quality in China is inferior and deteriorating (Liu and Diamond 2005). In most countries, regulators set limits on the discharge of wastewater into rivers and other natural water bodies. These limits define the maximum allowable concentration of specified pollutants that could be discharged into the water (Chinyama et al. 2016). The Chinese government has been working on minimizing environmental impacts via control strategies of the total amount of pollution (Li et al. 2010). Waste load allocation (WLA) is an important measure to control the total amount of pollutants and is an important guideline for the process of pollutant discharge.

The vast majority of studies on WLA have focused on uncertainty analysis (Mooselu et al. 2019; Soltani et al. 2016; Nikoo et al. 2016), emission trading (Soltani and Kerachian 2018), and on multiobjective WLA models (Ashtiani et al. 2015; Liu et al. 2014; Cho and Lee 2014; de Andrade et al. 2013; Zewdie and Bhallamudi 2012). The series of objects of WLA include the basins, the water functional

areas, the different administrative levels, the sewage outlets, and the pollution sources. Thus, some studies suggested that watershed has the characteristics of multilevel management. The disparity of the objectives pursued by the stakeholders at different levels may render difficult the implementation of an optimization plan (Liang et al. 2015).To improve the application of the WLA model to managers at different levels and better show the correlation amongst the different watershed management levels, a two-level or even multilevel WLA model could be used. For example, to assist in controlling the total amount of river pollutants, Xu et al. (2017) presented a bi-level optimization WLA programming model which verified the decision feedback relation between the river basin management committee and the regional Environmental Protection Agency (EPA). Rafiee et al. (2017) designed a new simulation optimization framework for the inaccuracy and the target conflict of a pollution control mechanism, which associated Qual2K with a genetic algorithm optimization model. Liang et al. (2015) developed an environmental capacity management system to support the execution of the total maximum daily load (TMDL), which contained a pollutant load grading and a distribution system. Pollutant loads are distributed from top to bottom according to the characteristics of the pollution sources. Hou et al. (2015) focused on the WLA of water functional areas and developed an equilibrium strategy to counterbalance the requirements of water functional areas in the first and second level. Aiming at the problems existing between the two levels of functional areas, a two-level WLA model was proposed to minimize water pollution. Zhang et al. (2012b) established a multilevel WLA method for the water functional areas and administrative units at the regional level. The method considered the region level and sewage point level together to make the WLA more equitable and efficient. Meng et al. (2017) developed a two-stage stochastic programming model to support regional chemical oxygen demand (COD) and ammonia nitrogen (NH$_3$-N) WLA in four main pollution sections (industry, municipal, livestock breeding, and agriculture).

The Gini coefficient was originally proposed by the Italian economist Gini in 1912 to measure income inequality. In recent years, many studies have focused on the Gini coefficient, which aimed at measuring the inequality in the use of environmental resources, including waste load allocation (Wang et al. 2011; Chen et al. 2012; Heon 2013). Based on a multi-criteria system including land area, population, gross domestic product (GDP), and environmental capacity, Sun et al. (2010) adopted the method of environmental Gini coefficient to allocate wastewater discharge permit. Wang et al. (2016) verified the geographical distribution of the pollutant allowable discharge load based on the distribution of water environmental capacity by calculating the Gini coefficient of the economic, social, and environmental efficiency. Cho and Lee (2014) used the environmental Gini coefficient as one type of

inequality measures to achieve the economic goal of reducing waste load while considering the inequality between waste dischargers. Wu et al. (2019) established a multi-index Gini coefficient method to evaluate the fairness of different allocation methods, which combined the Gini coefficient with a linear interactive optimization method in order to realize the fair distribution of ammonia emission permit in the Songhua River basin from two aspects of the basin and the region. These aforementioned studies show very well the widespread use of the Gini coefficient in the field of WLA as a common index.

Besides, various laws and policies have been enacted to protect water environmental. For instance, Sichuan Environmental Protection Agency promulgated a Pigovian tax policy in the Tuojiang and Minjiang Rivers in 2013, which are the most important tributaries of the upper Yangtze River, in order to limit water pollution and avoid conflicts between dischargers. Zhang et al. (2018) designed a WLA water quality management equilibrium strategy under the Pigovian tax policy, which fully considered the Stackelberg-Nash equilibrium between EPA and dischargers. Xu et al. (2016) proposed a hybrid nested particle swarm optimization method to solve the WLA problem of river systems based on the Pigovian tax. As a leader, the responsible EPA determines the pollution tax standards to resolve conflicts among emitters; and as a follower, each emitter decides to eliminate the biological oxygen demand under the prescribed pollution and pollution tax standards to reduce their pollution costs.

China promulgated the environmental protection tax law on December 25, 2016, at the 25th meeting of the standing committee of the 12th National People's Congress. This law states that enterprises, institutions, and other producers, as well as business operators that discharge taxable pollutants directly into the environment in the territory of the People's Republic of China or other sea areas under the jurisdiction of the People's Republic of China must pay an environmental protection tax according to the provisions of this law. The specific amount of the applicable tax is determined and adjusted by the people's governments of the provinces, autonomous regions, and municipalities, within the environmental protection tax range specified in the "Table of items and amounts of the environmental protection tax" attached to this law. This implementation considers the local environmental carrying capacity, the current situation of the pollutant discharge, and the objectives and requirements for economic, social, and ecological development.

These studies have developed many models and methods for WLA. However, not all models consider both fairness and efficiency principles simultaneously, even multiobjective models. Also, the implementation of an environmental protection tax significantly impacts on WLA, while only a few recent studies have considered an environmental protection tax. After the implementation of the environmental protection tax,

the tax amount directly affects the amount of pollution discharged by the pollutant dischargers. Therefore, how to set the environmental protection tax amount is one of the main challenges faced by upper decision-makers.

To address these issues, the bi-level multiobjective allocation model was established in this study. The upper allocation targets the minimal environmental Gini coefficient, as well as the minimal unit pollutant emission cost; and the watershed management committee (provincial people's government) is the initial allocation-level decision-maker. At the same time, the impact of the environmental protection tax is considered as a better reference for decision-makers at all levels and to improve the adaptation to management requirements. The lower level of allocation aims at the maximal industrial output value and the minimal reduction rates unevenness while considering each administrative authority at the lower level as a decision-maker. Each level considers the principles of fairness and efficiency to distribute the load in the basin.

Model construction

Model framework

In the bi-level decision-making problem, the upper-level decision-maker is the watershed management committee or the provincial people's government, and the lower-level decision-makers are the municipal people's governments or the authorities of the subareas.

The watershed management committee or the provincial people's government distributes water rights to each subarea. After subareas obtain their initial water rights, they make WLA decisions according to the waste discharge and utilization. The watershed management committee or the provincial people's government must plan for dischargers with due regard to pollutant handling capacity.

As shown in Fig. 1, the watershed management committee or the provincial people's government selects the minimal environmental Gini coefficient and the minimal discharge cost of per unit pollutant as the most important goals. At the lower level, the objectives are to maximize the industrial output value and to minimize the unevenness of reduction rates under the constraint of the total quantity. The principle of fairness and efficiency should be considered at each level.

Upper level allocation

In this study, the environmental Gini coefficient was adopted for WLA to achieve the fairness principle. The target of unit pollutant emission cost minimum was added to achieve the efficiency principle.

The unit pollutant cost includes two aspects: the sewage treatment cost and the environmental protection tax. After

pollutants are generated, one path is directed towards the sewage treatment plants, which generate sewage treatment costs, while the other path is directed towards the river, which is subjected to the environmental protection tax. Therefore, the target could be unit pollutant emission cost minimum.

Concerning the sewage treatment cost, this study considers the sewage treatment plant as the measure for controlling the point source pollutant and chooses the oxidation ditch as the main treatment technology. By referring to the marginal cost function given by Yan (2003), the national standard of the Environmental Quality Standards for Surface Water (GB3838–2002) was selected as the conversion factor between the volume of water and pollutant reduction, and the marginal cost function of the point source treatment of the pollutant was established as follows(Pang 2010):

$$M_i = a\left(x_{(0)i} - x_i\right)^b \qquad (1)$$

where M_i represents the marginal cost of the pollutants in the administrative region$i$ (10^4 yuan), $x_{(0)i}$ is the current emission, and x_i is the total allocation of pollutants in the ith region, which is the decision variable in the upper level allocation; a and$b$ are constants. In the marginal cost function of the pollutant COD, $a = 59.2445$ and $b = -0.4337$, and the marginal cost function of the pollutant NH$_3$-N, $a = 34.4739$, and $b = -0.4337$.

With regard to the environmental protection tax, the pollution equivalent refers to a comprehensive index or unit of measurement, which measures the environmental pollution caused by different pollutants according to the degree of harmfulness to the environment and the technical economy of the treatment of the pollutants or the pollutant discharge activities. For example, the pollution degree of different pollutants with the same pollution equivalent in the same medium is basically the same. The pollution equivalent value, which is usually expressed in kilogram, is the corresponding value of the pollutant that considers as the benchmark the harmful degree, the toxicity to the organism, and the treatment cost of the specified unit quantity of the main pollutants in the environmental pollution factors. The pollution equivalent amount, which is dimensionless, refers to the quantity of pollution equivalent that is calculated by dividing the discharge amount of pollutant by the pollution equivalent value of the pollutant. The payable tax of the taxable pollutants is equal to the pollution equivalent amount multiplied by the specific applicable tax amount.

The specific pollution equivalent value is subjected to the "Table of taxable pollutants and appropriate values." The items and amounts of the environmental protection tax are subjected to the "Table of items and amounts of the environmental protection tax" annexed to this law under which the range of the water pollutant tax is 1.4 yuan to 14 yuan per pollution equivalent. The specific amount of the applicable tax is determined and adjusted by the people's governments of the provinces, autonomous regions, and municipalities within the

Fig. 1 Frame diagram of the bi-level multiobjective allocation mode
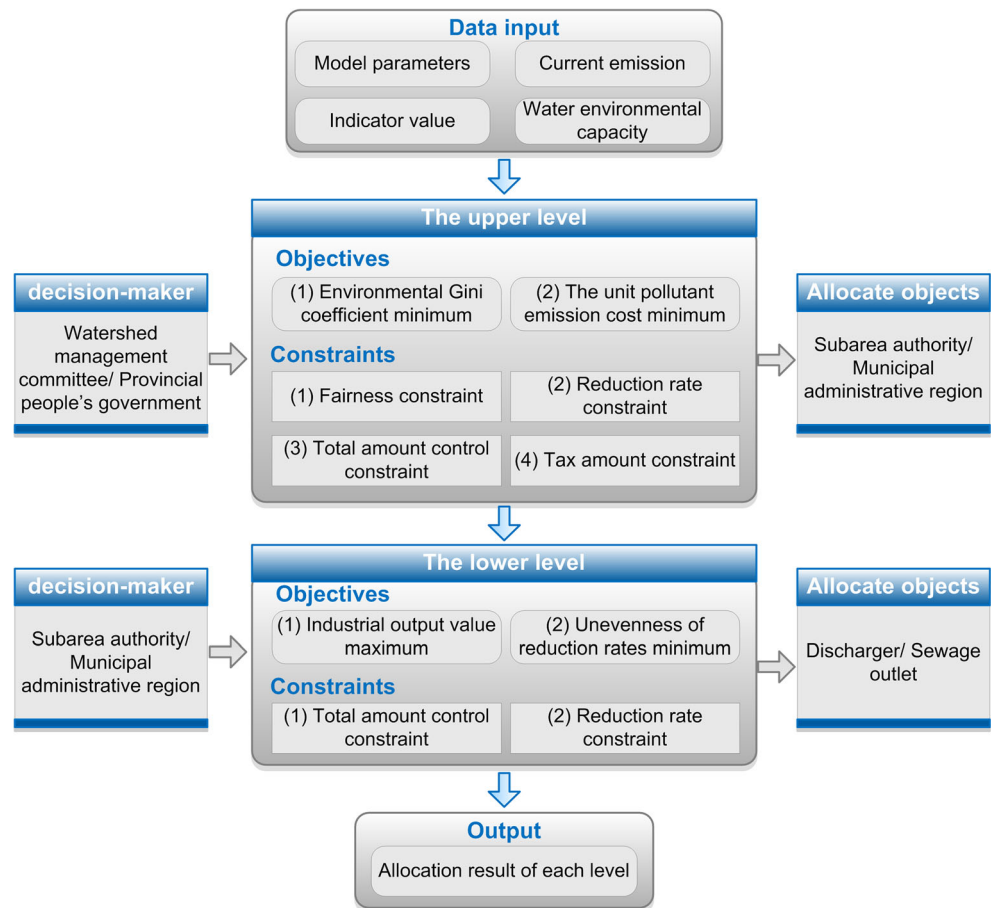
environmental protection tax range specified in the "Table of items and amounts of the environmental protection tax" attached to this law and considers the local environmental carrying capacity, the current situation of pollutant discharge, and the objectives and requirements for economic, social, and ecological development. Therefore, the tax amount can be calculated by the following equation:

$$E_i = \frac{x_i}{E_v}T \tag{2}$$

where E_i is the tax payable by a region on water pollutants; E_v is the pollution equivalent value of the pollutant, according to the "Table of taxable pollutants and appropriate values," where COD is 1.0 kg and NH$_3$-N is 0.8 kg; $\frac{x_i}{E_v}$ is the pollution equivalent amount, which is dimensionless; and T is the tax amount of water pollutants, which is implemented in accordance with the "Table of items and amounts of the environmental protection tax." The range of the tax amount of water pollutants is 1.4–14 yuan per pollution equivalent.

The objectives of the upper level allocation can be written as follows:

Objective 1:   Environmental Gini coefficient minimum

$$\min G = \sum_{j=1}^{m} w_j G_j \tag{3}$$

The Gini coefficient is calculated by the trapezoidal area method as follows:

$$G_j = 1 - \sum_{i=1}^{n} \left(X_{ij} - X_{i-1,j}\right)\left(Y_i + Y_{i-1}\right) \tag{4}$$

where G_j is the Gini coefficient corresponding to each index, G is the environmental Gini coefficient, X_{ij} represents the cumulative value of the jth indicator percentage in the region i, and Y_i represents the cumulative value of the current sewage discharge or the load distribution percentage in the region i.

Objective 2:   Unit pollutant emission cost minimum

$$\min UC = \sum_{i=1}^{n} \left(\frac{M_i}{x_i} + \frac{E_i}{x_i}\right) \tag{5}$$

Constraints

Regarding the fairness constraint, the optimized Gini coefficient is no worse than the current Gini coefficient in order to ensure that fairness is increased. For cases with a small current Gini coefficient, the elastic constraints can be adopted, i.e., the optimized Gini coefficient is a little smaller than the current Gini coefficient.

$$G_j \leq G_{0(j)} \tag{6}$$

where $G_{0(j)}$ is the current Gini coefficient of the jth indicator.

Concerning the reduction rate constraint, the reduction rate P_i of each region is calculated after determining the total control amount. Considering that the reduction capacity of each region is different, the lower limits P_1 and P_2 of the reduction rate are considered. The specific range can be adjusted according to the specific emission situation and the water environmental capacity (WEC) of the region, also known as water assimilative capacity. Generally speaking, the allowable amount of waste load should not exceed the river WEC, which specifies the maximum load quantity of certain pollutants during specific times, under specific design hydrological conditions, in a certain water unit, such that the water meets certain environmental objectives (Wang et al. 1995) and reflects the capacity of the water body to accept pollutants without deteriorating its function (Shu and Ma 2010). When setting P_1 and P_2, it is necessary to ensure the completion of the reduction task and to maintain the continuity of production while considering the bearing capacity of pollution load reduction in different subregions. Generally, the reduction rate could be set slightly larger in the districts that are economically developed and where sewage discharge amount is relatively large.

$$P_1 \leq P_i \leq P_2 \tag{7}$$

where $P_i = \frac{x_{(0)i} - x_i}{x_{(0)i}}$

Regarding the total amount control constraint, the safety margin is one of the basic elements contained in the TMDL of typical pollutants in the USA. It usually refers to the portion of WEC that is reserved based on caution and is necessary because of the uncertainty of the relationship between pollution load and water quality of rivers and lakes and usually accounts for 5–10% of the WEC. The sum of the pollution amount allocated by each region must be less than or equal to the WEC after deducting the safety margin, such that:

$$\sum_{i=1}^{n} x_i \leq \mu W \tag{8}$$

where W is the WEC, and μ is the percentage of the safety margin in the WEC.

With regard to the tax amount constraint, the range of the tax amount of water pollutants is 1.4–14 yuan per pollution equivalent following the "Table of items and amounts of the environmental protection tax", such that:

$$1.4 \leq T \leq 14 \tag{9}$$

Lower level allocation

The allocation of the lower level should also follow the principles of fairness and efficiency. The principle of efficiency is embodied by the maximum industrial output value, while the principle of fairness is embodied by the reduction rates unevenness minimum of the dischargers.

Objective 1:   Industrial output value maximum

$$\max f_i(x) = \sum_{j=1}^{m} C_{ij} x_{ij} \quad j = 1, 2, 3..., m \tag{10}$$

where $f_i(x)$ is the total industrial output value of the ith region, C_{ij} is the GDP generated by the unit quantity of the waste load of the jth discharger into the river, and x_{ij} is the pollutant discharge allocated by the jth discharger in the ith region, which is the decision variable in the lower-level allocation. m is the number of dischargers in the ith region.

Objective 2:   Unevenness of reduction rates minimum

$$\min \sigma_i = \frac{1}{m} \sum_{j=1}^{m} \left(P_{ij} - \overline{P}_{ij} \right)^2 \tag{11}$$

where σ_i is the variance of the deduction rate of all dischargers in the region i, P_{ij} is the deduction rate of discharger j in the region i, and $\overline{P}_{ij}$ is the mean of deduction rate of all dischargers in the region i.

Constraints

Regarding the total amount control constraint, the sum of the amounts allocated to each discharger must be less than or equal to the total amount of pollutants allocated to the region.

$$\sum_{j=1}^{m} x_{ij} L_{ij}(x) \leq x_i \tag{12}$$

where $L_{ij}(x)$ is the loss rate of pollutant of the jth discharger in the ith region, which is determined by the location of the sewage outlet and the distance from the river, and is generally within the range 0.8–1.0.

Regarding the reduction rate constraint, the top and bottom limitations P_1' and P_2' of the reduction rate are also proposed, considering the different reduction capacity of dischargers.

The specific scope can be flexibly adjusted according to the actual sewage discharge situation and the technical conditions. It is necessary to consider the discharge amount and the bearing capacity of the pollution load reduction of the different drain outlets when setting P_1' and P_2'. Usually, the range of the reduction rate could be set higher for the discharger with more sewage discharge amount and better economic development.

$$P_1' \leq P_{ij} \leq P_2' \tag{13}$$

Model solution

The process of using the bi-level multiobjective allocation model to apply the WLA includes three main steps: the determination of the total WLA amount, the upper-level allocation, and the lower-level allocation. The flow chart of the bi-level WLA allocation process is shown in Fig. 2.

   Step 1:   Data collection
             This process includes the model parameters, the current pollution emission situation, and the index values selected for calculating the environmental Gini coefficient, WEC.

Step 2:   Determination of the total waste load allocation amount
             To conduct the study on total pollutant control, it is necessary to determine the total waste load allocation amount in the river (i.e., the total allowable waste load amount), which could be determined by the river WEC and the safety margin. The calculation of the total waste load allocation amount involves the following equation:

$$W_T = W - W_s \tag{14}$$

where W_T is the river total waste load allocation amount, W is the river WEC, and W_s is the safety margin, which generally accounts for 5–10% of the WEC.

Step 3:   Upper-level allocation
             A multiobjective distribution method that considers the Gini coefficient is adopted for the upper distribution. The selection of the indices has the greatest influence on WLA and is also an important factor needed to determine whether the allocation scheme can be accepted. The number of selected indicators is generally within the range 3–6 and should be too many. This study adopted 6 control indices to judge the fairness degree of WLA.
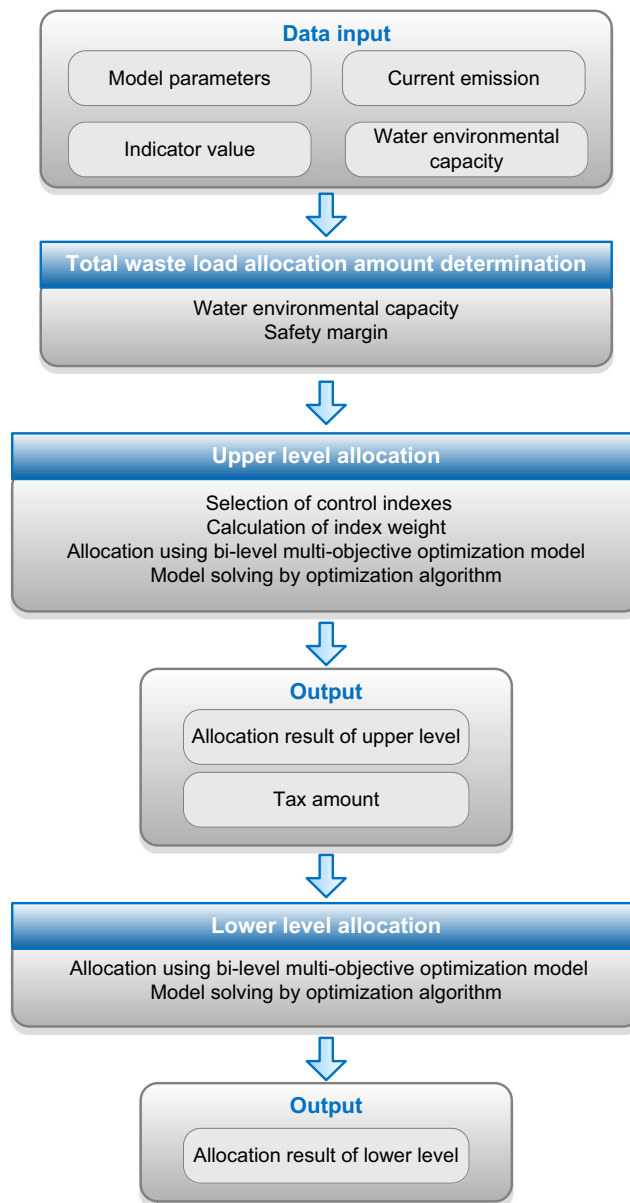   (1) Population: The Gini coefficient of the

Fig. 2 Flow chart of the bi-level WLA allocation

population-sewage discharge reflects the fairness of the per capita amount of sewage discharge.
             (2) GDP: GDP is an important factor used to assess the overall economic status of a region. A higher amount of water pollutants discharged to create per unit GDP implies that to create the economic profit of the same value, the region has a greater impact on the ecological environment and should undertake greater reduction.
             (3) Water consumption: The Gini coefficient of the water consumption-sewage discharge reflects the difference in the water pollutant load per unit water consumption and could better reflect the water

efficiency of a region. The smaller the load intensity of unit water consumption, the more environmentally friendly the region is, and the less impact it has on the ecological environment.

(4) Industrial output value: The allowable discharge volume allocated in this study is primarily targeted at point sources, and the industry is considered to be the main source of point source pollution. Therefore, the Gini coefficient of the industrial output value and the sewage discharge truly reflects the fairness of the load distribution.

(5) Environmental protection investment: To enhance environmental awareness and develop enthusiasm within a region, the region with a large number of pollutant loads in unit environmental protection investment should increase its reduction amount.

(6) WEC: WEC is the maximum amount of sewage that can be borne by a region considering that the water quality target and the water resource quantity. It is reasonable to allocate a greater sewage discharge amount to regions with greater WEC.

Step 4: Calculation of index weight

In this study, the entropy method is used to calculate the weight of each index. In the information theory, entropy is a measure of uncertainty. The smaller the amount of information, the greater the uncertainty and the entropy. Conversely, the more information, the smaller the entropy. From the perspective of WLA, if the difference of the unit pollutant load of an index in a region is greater, it means that the index has a greater influence on the distribution result and that the weight value should be greater. The specific calculation process of the weight is described as follows:

If x_i represents the pollutant emission distributed in the ith region, and z_{ij} represents the value of the jth index in the ith region, then the pollutant quantity load in a unit of the jth index in the ith region is then written as follows:

$$y_{ij} = x_i/z_{ij} \tag{15}$$

Thus, the weight of the jth index in the ith region of all regions is:

$$p_{ij} = y_{ij}/ \sum_{i=1}^{n} y_{ij} \tag{16}$$

The information entropy of the pollutant quantity loaded by a unit of the jth indicator is:

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^{n} \left(p_{ij} * \ln p_{ij} \right) \tag{17}$$

where n is the number of regions.

The corresponding weight of each indicator is:

$$w_j = (1-e_j)/ \sum_{j=1}^{m} (1-e_j) \tag{18}$$

where m is the number of control indicators.

Step 5: Upper level allocation

This process uses the bi-level multiobjective allocation model, which can be solved by an optimization algorithm. The non-dominated sorting genetic algorithm II (NSGA-II) is one of the most widely used and effective algorithms for solving problems of multiobjective optimization. This method offers the advantages of fast running speed and good convergence of the solution set and has been successfully applied in optimal water resource allocation and optimal reservoir operation, among other fields. The NSGA-II algorithm is adopted in this research adopts to solve the model.

Step 6: Lower level allocation

The output of the upper level allocation should be the input data of the lower level allocation. The lower level allocation of the bi-level multiobjective allocation model could also be solved by the optimization algorithm.

Step 7: Output allocation results

Application

Study area

The Shaanxi reach of Wei River mainstream was herein considered as the study area. The Wei River (Zhao et al. 2008) is the largest tributary of the Yellow River, with a total length of 818 km. The Wei River mainstream passes through five cities, namely, Baoji, Yangling, Xianyang, Xi'an, and Weinan in the Shaanxi province (Fig. 3), with a length of 504 km. The Wei River basin in the Shaanxi province has many tributaries; nine of them are large, including the Ba River, the Hei River, and the Jing River. Also, the Wei River contains 71 drain outlets and 5 large water intakes.

In this study, only point source pollutions were considered, while nonpoint source pollutions were not. The Wei River suffers from serious water contamination, which limits the sustainable development in the nearby areas. The main pollutants that exceed the standards in the Wei River are the COD, the NH_3-N, and the volatile phenol. This study exclusively considered COD and NH_3-N as the main research items. For these two pollutants, the WLA results were herein calculated
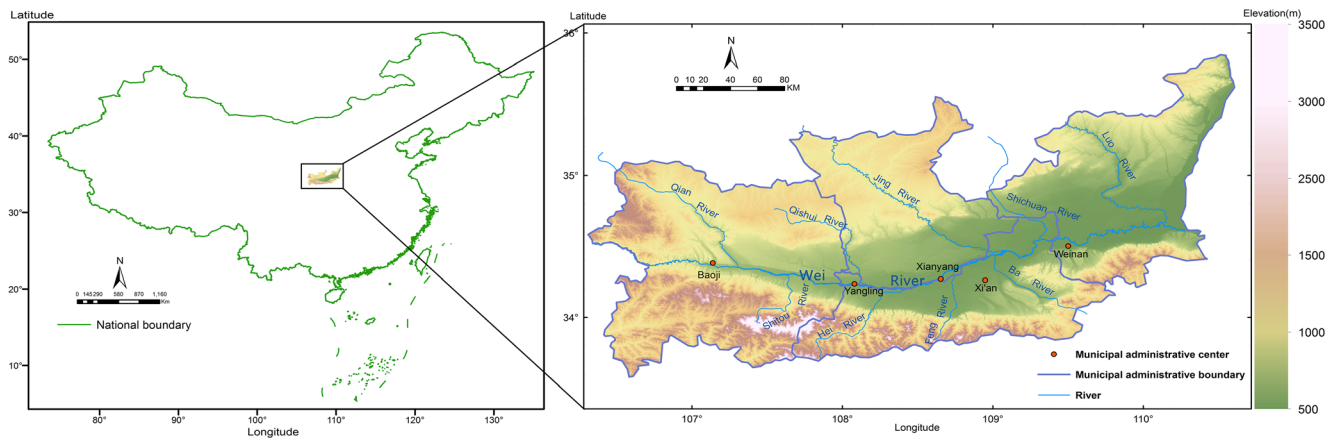
Fig. 3 Schematic diagram of study area

separately. According to the environmental protection tax law, the specific tax amount of water pollutants discharged by sewage dischargers in the Shaanxi province is determined and adjusted by the Shaanxi provincial people's government within the range of the tax amount given in the tax law (1.4–14 yuan per pollution equivalent). The Shaanxi provincial people's government is the decision-maker at the initial allocation level. Each municipal administrative region authority is the decision-maker at the lower level.

Firstly, according to the technical process of the total pollutant control and the principle of the determination of control target, the target amount of COD and NH_3-N in Shaanxi reach of the Wei River mainstream was determined. The bi-level multiobjective allocation model, the traditional equal proportion method, and a traditional multiobjective WLA model were used to allocate the waste load in the Shaanxi reach of the Wei River mainstream, which subsequently resulted in different allocation schemes. Finally, the differences between these three methods were compared and analyzed, and the rationality and practicability of the bi-level multiobjective allocation model were verified. The data were sourced from the statistical yearbook of the Shaanxi province.

Results and analysis

In this study, the safety margin was set to 5% of the WEC, and the results of the WEC were calculated under the 90% design frequency by the section-beginning control model (Zhou et al. 1999), using the flow data from 1971 to 2010 in the Wei River. The annual COD and NH_3-N WEC in the Shaanxi reach of the Wei River mainstream were 77,741.73 tons and 2438.98 tons, respectively. Nonpoint source pollution was not considered in the calculation results. In order to verify the rationality of the model, the bi-level multiobjective allocation model and two traditional load distribution methods were adopted to calculate the load distribution in the Wei River basin. These two traditional load distribution methods included the equal proportion

method and a traditional multiobjective WLA model (Hou et al. 2015; Zhang et al. 2012b). To facilitate the comparison, the indicators selected by the environmental Gini coefficient in the traditional multiobjective WLA model were consistent with the indicators of the bi-level multiobjective allocation model presented herein, with a total of six indices.

Upper-level allocation result

The bi-level multiobjective allocation model was solved by NSGA-II. The Pareto curves of the COD and the NH_3-N allocation were subsequently derived, as shown in Fig. 4:

Each point in the curve corresponds to an allocation scheme, and each allocation scheme is optimal under the same conditions, supplying a reliable scheme set for the decision-makers. According to the Convention, the Pareto optimal solution in the multiobjective optimization problem usually considers the solution at the inflection point as the optimal solution that has achieved the balance of the two objectives. In practice, if there are special decision preferences, the optimal solution could nonetheless be selected according to the preferences.

The results of the two groups with the largest and smallest environmental Gini coefficients in the above schemes were selected and compared with the current Gini coefficients, as shown in Table 1.

The actual Gini coefficient is between 0 and 1. The smaller the Gini coefficient, the fairer the distribution is. Following the provisions of the relevant organizations of the United Nations, the Gini coefficient is lower than 0.2, thereby indicating an absolute average distribution. A Gini coefficient between 0.2 and 0.3 indicates a comparative average distribution, between 0.3 and 0.4 indicates a relatively reasonable distribution, between 0.4 and 0.5 indicates a large gap, and when it exceeds 0.6, it suggests a great disparity. Therefore, the Gini coefficient of 0.4 was usually taken as the warning line of the distribution gap (Li and Shu 2011). This boundary could only be
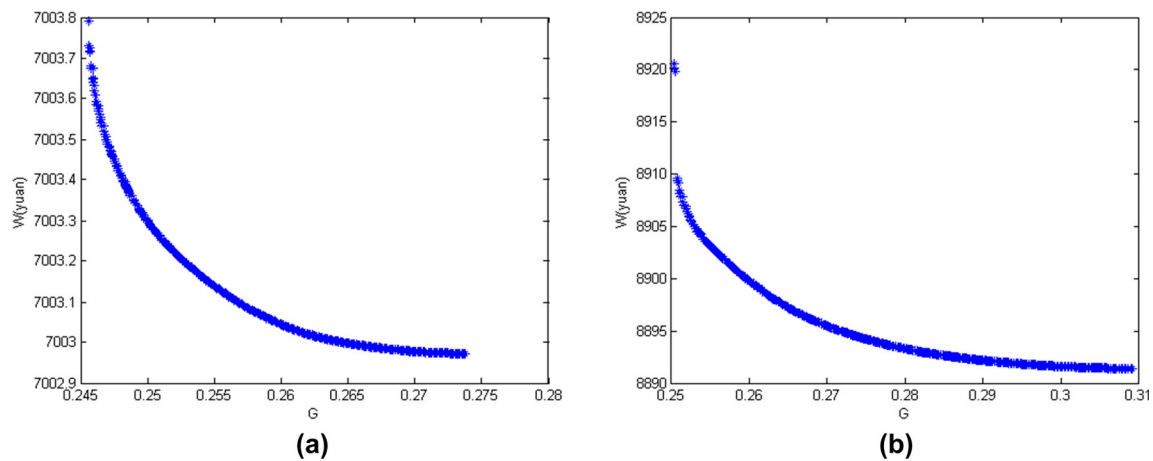
Fig. 4 Pareto curves in the upper level allocation of the bi-level multiobjective allocation model. **a** Pareto curve of COD allocation. **b** Pareto curve of NH$_3$-N allocation

used as a reference. In practice, the actual situation and the application field also must be considered.

As shown in Table 1, after optimization calculation, the Gini coefficient results of all indicators are lower than 0.4, which indicates that the distribution results are generally acceptable to polluters. In the COD optimization scheme, the minimal and maximal environmental Gini coefficients are both smaller than the current value, indicating that the fairness of load distribution has been further improved. In the NH$_3$-N allocation scheme, the situation in which the minimal Gini coefficient was the same as in the current situation, while the maximal environmental Gini coefficient was larger than the current value occurred because the current environmental Gini coefficient was already relatively small, at the comparative average level. While aiming for fairness, one should also consider efficiency. The result of NH$_3$-N was also relatively reasonable in the environmental Gini coefficient maximum scheme.

Because the environmental Gini coefficients were all less than 0.4, which indicate the excellent reliability of the results, the solutions with the minimal unit pollutant emission cost were selected. The tax amount results for the COD and the NH$_3$-N were both 1.4 yuan per pollution equivalent, as calculated by the bi-level multiobjective allocation model. The unit pollutant discharge costs of the COD and the NH$_3$-N were 7002.973 yuan/ton and 8891.388 yuan/ton, respectively. When the tax amount is 1.4 yuan per pollution equivalent, the taxes payable on COD and NH$_3$-N were 1400 yuan/ton and 1750 yuan/ton, respectively. Considering the example of a discharger with average discharge capacity, if all these pollutants are discharged into the water body, the annual tax payable for COD is 1.6572 million yuan, and the tax payable for NH$_3$-N is 97,800 yuan. This is only the tax payable when the tax amount is the minimum value, which may be greater in practice. Thus, the impact of the environmental

Table 1 Comparison table of Gini coefficient changes of each index

Index	COD			NH$_3$-N		
	Current value	Environmental Gini coefficient minimum	Environmental Gini coefficient maximum	Current value	Environmental Gini coefficient minimum	Environmental Gini coefficient maximum
Population	0.116	0.271	0.176	0.168	0.218	0.167
GDP	0.276	0.423	0.339	0.325	0.374	0.310
Water consumption	0.382	0.170	0.297	0.367	0.229	0.374
Industrial output value	0.136	0.308	0.208	0.180	0.253	0.164
Environmental protection investment	0.094	0.270	0.166	0.059	0.219	0.163
Water environmental capacity	0.194	0.079	0.113	0.163	0.032	0.236
Environmental Gini coefficient	0.312	0.246	0.274	0.250	0.250	0.309

protection tax on the cost of the pollutant discharge is significant and cannot be ignored. The WLA schemes of the corresponding COD and NH_3-N are listed as follows (Tables 2 and 3).

According to the allocation results of the COD and the NH_3-N calculated by the equal proportion method, the environmental Gini coefficients were equal to 0.312 and 0.272, respectively. In the traditional multiobjective WLA model, the upper-level allocation aims at the minimal environmental Gini coefficients, which were calculated as 0.246 and 0.250, for COD and NH_3-N, respectively. Considering the tax amount as 1.4 yuan per pollution equivalent calculated by the bi-level multiobjective allocation model to calculate the cost of per unit pollutant discharge for these two pollutants, the results of the COD and the NH_3-N were 7003.727 yuan/ton and 8949.226 yuan/ton, respectively. These were basically equal to the results of the minimal Gini coefficient in the Pareto curves in the bi-level multiobjective allocation model. On the contrary, if the target is to minimize the unit pollutant discharge cost, the calculated results were located at the points of the minimal unit pollutant discharge cost in the Pareto curves in the bi-level multiobjective allocation model. This also verifies the reliability of the results of the bi-level multiobjective allocation model.

By comparing the pollutant reduction ratios of all the municipal administrative regions in the four schemes, namely, the environmental Gini coefficient minimum, the environmental Gini coefficient maximum, the equal proportion method, and the traditional multiobjective WLA model (Fig. 5). One can notice that in the results of the equal proportion method, all the administrative regions must make equal proportion reductions, whereas, in the results of the other models, the reduction rates are different. The allocation results confirm that greater pollution induces a larger reduction, which better reflects the principle of fairness. The bi-level multiobjective allocation model pursues the minimal cost of unit pollutant discharge while considering fairness and solves the problem

of setting the environmental protection tax amount simultaneously.

Lower-level allocation result

In the lower-level allocation, the bi-level multiobjective allocation model, the equal proportion method, and the traditional multiobjective WLA model were also selected for the allocation. The allocation result of the COD with the greatest environmental Gini coefficient in the upper-level allocation results was selected to conduct the lower level WLA in the bi-level multiobjective allocation model. The objectives of the bi-level multiobjective allocation model included the maximal industrial output value and the minimal reduction rate variance, while the objective of the traditional multiobjective WLA model was the industrial output value maximum. When the NSGA-II algorithm was used in the calculation, the two objectives were changed in the same direction, namely, both maximum or both minimum. Thus, the total industrial output value resulted in the minimum value, and the outcome of the bi-level multiobjective allocation model was a set of Pareto optimal curves.

As shown in Fig. 6, the variances are all below 0.06, which not only indicates that the reduction rate of the sewage outlets in the same municipal administrative region was relatively average but also reflects the rationality and the fairness of the calculation results. On the premise of fairness, the result with the greatest benefit could be selected. Therefore, the group of allocation results with the largest industrial output value was selected, and the reduction rates of each sewage outlet calculated by these three methods are shown in Fig. 7.

One can notice that in the COD allocation scheme, the main reduction tasks were undertaken by Baoji, Xianyang, and Xi'an. In the upper-level allocation, these three districts are needed to reduce the discharge amount. Therefore, the reduction rates of the sewage outlets were higher in these three districts, especially in Baoji, where the reduction rate was the greatest. When lower levels were allocated, the other two districts, namely, Yangling and Weinan, were redistributed to maximize efficiency,

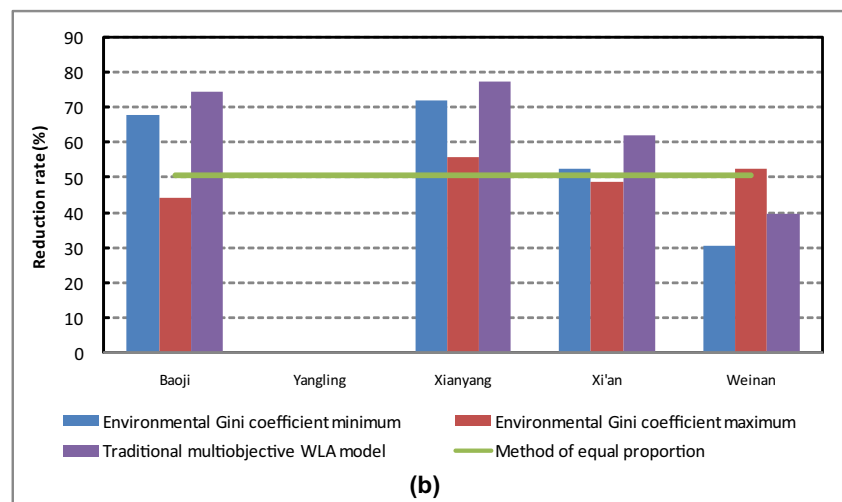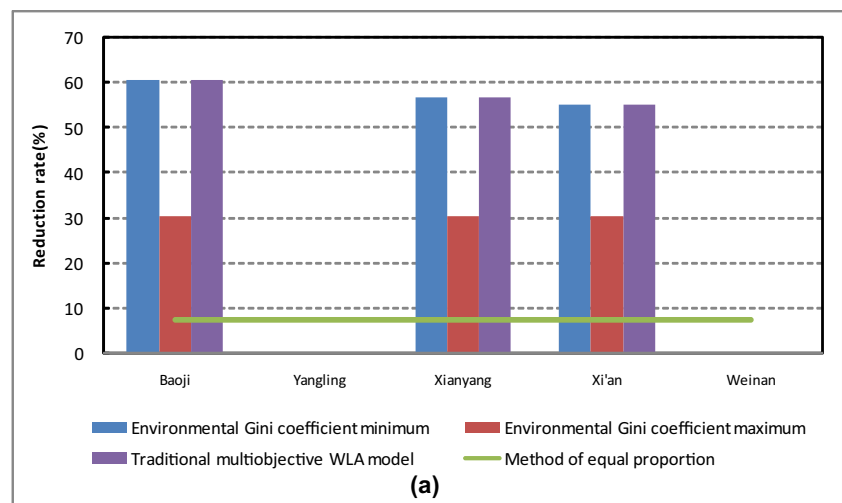**Table 2** Upper-level allocation schemes of COD in the Wei River Shaanxi reach (unit: t)

Order	Administrative region	Environmental Gini coefficient minimum	Environmental Gini coefficient maximum	Method of equal proportion	Traditional multiobjective WLA model
1	Baoji	7012.56	12,456.59	16,546.88	7024.35
2	Yangling	455.92	455.92	421.73	455.92
3	Xianyang	7929.88	12,821.58	17,020.84	7943.67
4	Xi'an	10,889.87	16,896.51	22,474.19	10,908.62
5	Weinan	23,003.22	23,003.22	21,278.08	23,003.22
Total		49,291.45	65,633.82	77,741.73	49,335.78

Table 3 Upper-level allocation schemes of NH$_3$-N in the Wei River Shaanxi reach (unit: t)

Order	Administrative region	Environmental Gini coefficient minimum	Environmental Gini coefficient maximum	Method of equal proportion	Traditional multiobjective WLA model
1	Baoji	329.62	577.00	508.99	261.94
2	Yangling	12.15	12.15	5.97	12.15
3	Xianyang	397.53	627.28	700.09	317.68
4	Xi'an	550.16	590.75	569.83	437.76
5	Weinan	922.36	631.79	654.10	800.30
Total		2211.82	2438.97	2438.98	1829.83

which implied that certain sewage outlets also had to reduce pollutants. In the distribution scheme of the NH$_3$-N, all the districts except Yangling must reduce pollutants. Moreover, the current discharge of NH$_3$-N in the whole basin significantly exceeds the WEC of NH$_3$-N, which subsequently induced the reduction rates of the sewage outlets to be noticeably large. With regard to the trend of the reduction rate of these dischargers, the trend of the bi-level multiobjective allocation model is similar to that of the traditional multiobjective WLA model, which also evidences the reliability and the rationality of the bi-level multiobjective allocation model.

By comparing the industrial output value of each sewage outlet calculated by the bi-level multiobjective allocation model, the equal proportion method, and the traditional multiobjective WLA model (Fig. 8), one may

Fig. 5 Upper-level reduction rate comparison of bi-level multiobjective allocation model and the other two methods. **a** Reduction rate comparison of COD of each municipal administrative region. **b** Reduction rate comparison of NH$_3$-N of each municipal administrative region
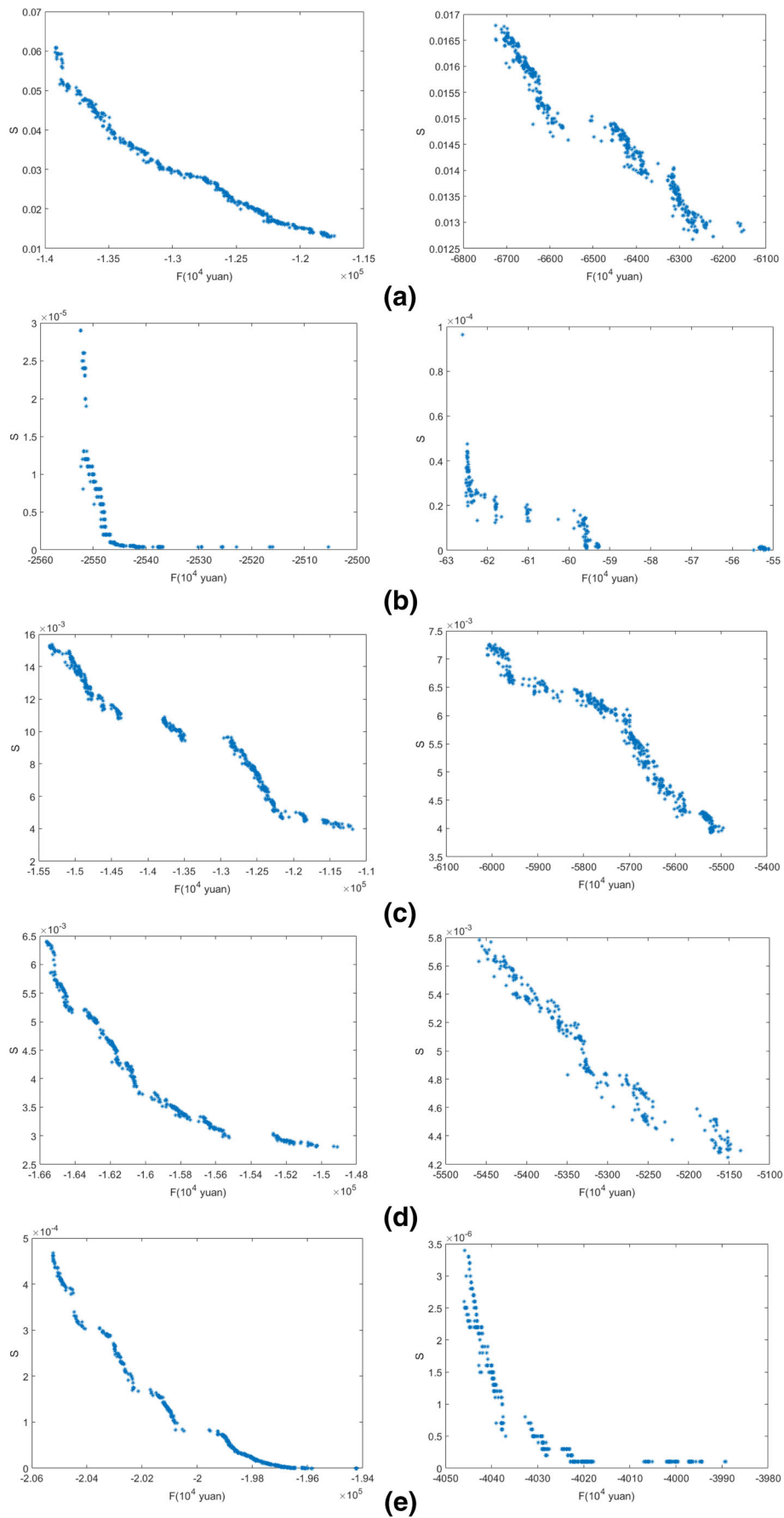
observe that the overall trends of the three methods are generally similar. According to the calculation results of NH₃-N, the resulting industrial output values are 218.52 million yuan, 216.16 million yuan, and 214.55 million yuan, for the bi-level multiobjective allocation model, the equal proportion method and the traditional multiobjective WLA model, respectively. Also, the values of the total allocated NH₃-N discharge amount are 2415.584 tons, 2438.98 tons, and 2067.306 tons for the bi-level multiobjective allocation model, the equal

proportion method, and the traditional multiobjective WLA model, respectively. The bi-level multiobjective allocation model was optimized to obtain a larger industrial output value when the total discharge was less than the equal proportion method. In the calculation results of COD, the total industrial output values in the distribution results are 6702.44 million yuan, 7322.30 million yuan, and 5331.23 million yuan for the bi-level multiobjective allocation model, the equal proportion method, and the traditional multiobjective WLA model, respectively, while the total allocated COD discharge amounts are 70,723.803 tons, 77,741.73 tons, and 52,742.79 tons for the bi-level multiobjective allocation model, the equal proportion method, and the traditional multiobjective WLA model, respectively. The total industrial output value of the bi-level multiobjective allocation model was lower than that
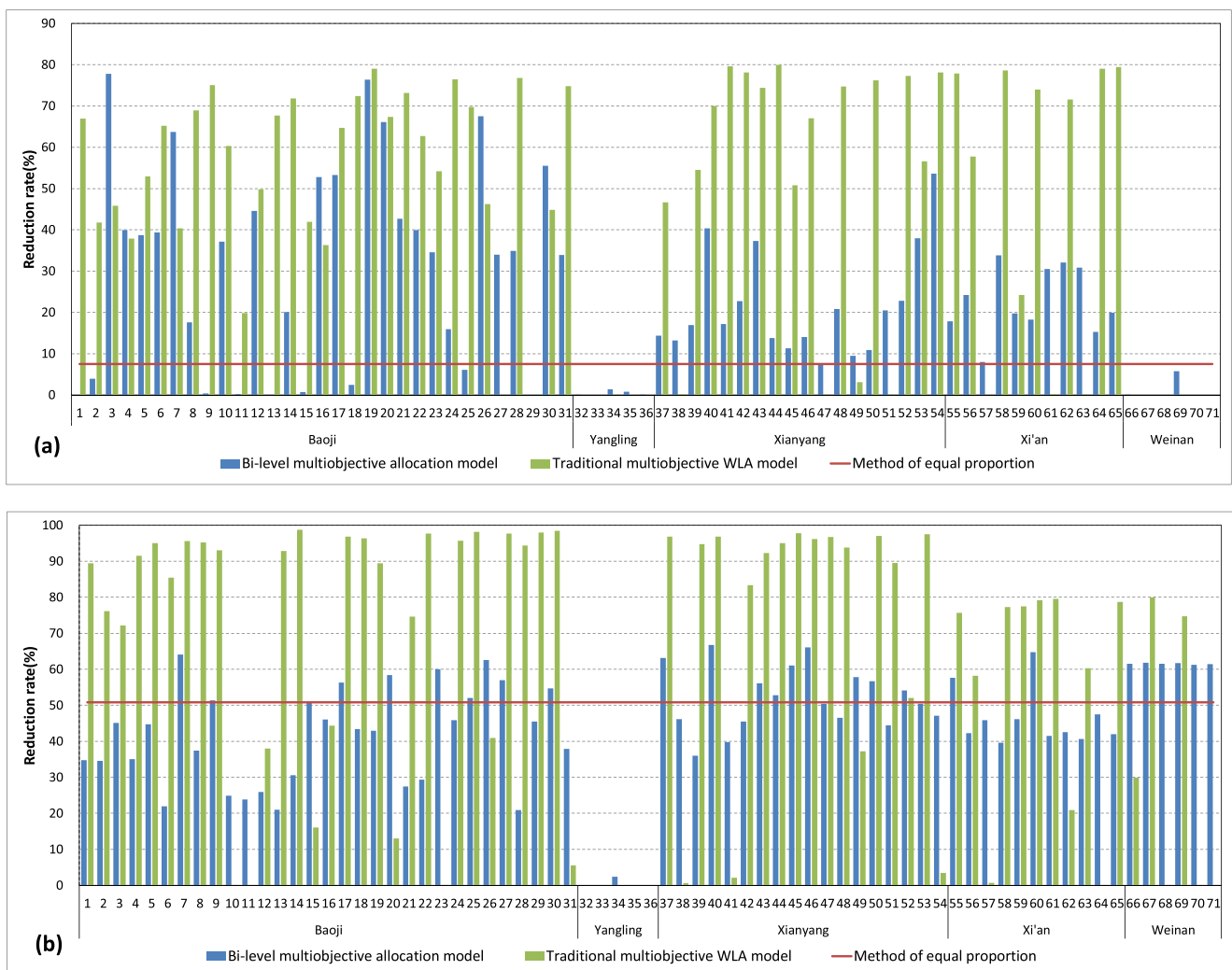

Fig. 7 Lower-level reduction rate comparison of bi-level multiobjective allocation model and other two methods. **a** The reduction rate of COD of each sewage outlet. **b** The reduction rate of NH₃-N of each sewage outlet
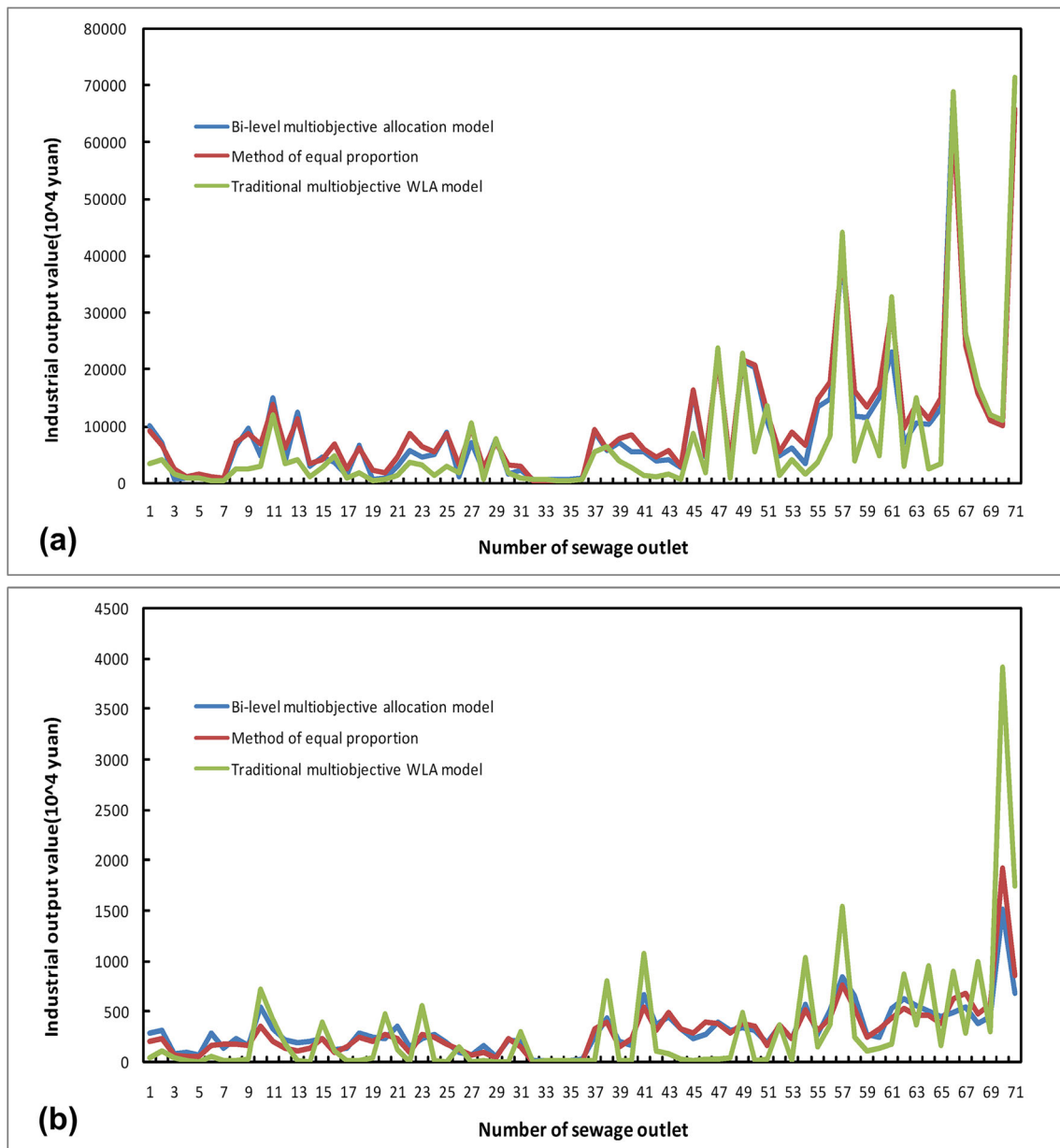
Fig. 8 Industrial output value contrast figure of bi-level multiobjective allocation model and the other two methods. **a** Industrial output values of COD in each discharger. **b** Industrial output values of NH$_3$-N in each discharger

of the equal proportion method. This is mainly because the total pollution amount allocated is significantly lower than that of the equal proportion method. However, the calculated environmental Gini coefficient of the bi-level multiobjective allocation model is less than 0.312 of the equal proportion method.

The comparison of the calculation results of the upper and lower levels shows that the bi-level multiobjective allocation model could yield greater benefits than the traditional multiobjective WLA model while ensuring fairness. This further demonstrates the rationality and reliability of the bi-level multiobjective allocation model.

Conclusion

A compromise exists between cost and equity in the WLA decision-making process under an environmental protection tax law, and the tax rate can significantly affect discharger decisions. It is difficult to consider the principles of fairness and efficiency at the same time in the existing WLA models. This study established a bi-level multiobjective allocation model under an environmental protection tax law to address the WLA problem for different management levels. The application of the proposed model to the Wei River basin demonstrated the following points:

Compared with traditional methods, in the upper allocation and under the premise of considering environmental Gini coefficient to ensure fairness, the bi-level multiobjective allocation model could solve the problem of setting the tax amount for the upper decision-makers, which influenced the pollution reduction behavior of the dischargers and derived the minimal unit pollutant emission cost.

Regarding the results of the lower-level allocation, the bi-level multiobjective allocation model derived a larger industrial output value than the traditional multiobjective WLA model. This approach also reflected that the bi-level multiobjective allocation model pursued the maximal benefit while considering fairness.

The bi-level multiobjective allocation model solves the problem of WLA under an environmental protection tax law. Each level of the bi-level multiobjective allocation model considers the principles of fairness and efficiency to distribute the load in the basin to create a better reference for decision-makers at both levels and to improve the adaptation to management requirements.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Ashtiani EF, Niksokhan MH, Ardestani M (2015) Multi-objective waste load allocation in river system by MOPSO algorithm. Int J Environ Res 9:69–76

Chen B, Gu C, Wang H, Tao J, Wu Y, Liu Q, Xu H (2012) Research on Gini coefficient application in the distribution of water pollution. J Food Agric Environ 10:833–838

Chinyama A, Ncube R, Ela W (2016) Critical pollution levels in Umguza River, Zimbabwe. Phys Chem Earth 93:76–83. https://doi.org/10.1016/j.pce.2016.03.008

Cho JH, Lee JH (2014) Multi-objective waste load allocation model for optimizing waste load abatement and inequality among waste dischargers. Water Air Soil Pollut 225:17

De Andrade LN, Mauri GR, Mendonca ASF (2013) General multiobjective model and simulated annealing algorithm for waste-load allocation. J Water Resour Plan Manage ASCE 139:339–344

Heon CJ (2013) Application of multi-objective genetic algorithm for waste load allocation in a river basin. J Environ Impact Assess 22:713–724

Hou SH, Song XL, Yao LM (2015) A bi-level waste load allocation model based on water function zoning for Sichuan-Neijiang. In: Xu J, Nickel S, Machado VC, Hajiyev A (eds) Proceedings of the ninth international conference on management science and engineering management, vol 362. Springer-Verlag, Berlin, pp 209–220

Li RZ, Shu K (2011) Model for wastewater load allocation based on multi-objective decision making. Acta Sci Circumst 31:2814–2821. https://doi.org/10.13671/j.hjkxxb.2011.12.031 (in Chinese)

Li YX, Qiu RZ, Yang ZF, Li CH, Yu JS (2010) Parameter determination to calculate water environmental capacity in Zhangweinan canal sub-basin in China. J Environ Sci 22:904–907. https://doi.org/10.1016/S1001-0742(09)60196-0

Liang SD, Jia HF, Yang C, Melching C, Yuan YP (2015) A pollutant load hierarchical allocation method integrated in an environmental capacity management system for Zhushan bay, Taihu lake. Sci Total Environ 533:223–237

Liu J, Diamond J (2005) China's environment in a globalizing world. Nature 435:1179–1186

Liu DD, Guo SL, Shao QX, Jiang YZ, Chen XH (2014) Optimal allocation of water quantity and waste load in the northwest pearl river delta, China. Stoch Env Res Risk A 28:1525–1542

Meng C, Wang XL, Li Y (2017) An optimization model for waste load allocation under water carrying capacity improvement management, a case study of the Yitong River, Northeast China. Water 9:16

Mooselu MG, Nikoo MR, Sadegh M (2019) A fuzzy multi-stakeholder socio-optimal model for water and waste load allocation. Environ Monit Assess 191:16

Nikoo MR, Beiglou PHB, Mahjouri N (2016) Optimizing multiple-pollutant waste load allocation in rivers: an interval parameter game theoretic model. Water Resour Manag 30:201–4220

Pang YY (2010) Research on permissible pollution bearing capacity and wasteload allocation in Lakes Basin. Dissertation, Zhengzhou University. (in Chinese)

Rafiee M, Lyon SW, Zahraie B, Destouni G, Jaafarzadeh N (2017) Optimal wastewater loading under conflicting goals and technology limitations in a riverine system. Water Environ Res 89:211–220

Shu SH, Ma HA (2010) Comparison of two models for calculating water environment capacity of songhua river. In: Li K, Jia L, Sun X, Fei M, Irwin GW (eds) Life system modeling and intelligent computing, vol 6330. Springer-Verlag, Berlin, p 683. https://doi.org/10.1007/978-3-642-15615-1_80

Soltani M, Kerachian R (2018) Developing a methodology for real-time trading of water withdrawal and waste load discharge permits in rivers. J Environ Manag 212:311–322

Soltani M, Kerachian R, Nikoo MR, Noory H (2016) A conditional value at risk-based model for planning agricultural water and return flow allocation in river systems. Water Resour Manag 30:427–443

Sun T, Zhang H, Wang Y, Meng X, Wang C (2010) The application of environmental Gini coefficient (egc) in allocating wastewater discharge permit: the case study of watershed total mass control in Tianjin, China. Resour Conserv Recycl 54(9):601–608

Wang HD, Wang SH, Bao QS, Qi Z (1995) On regional differentiation of river water environment capacity and strategies to control water environment pollution in China. Chin Geogr Sci 5:116–124. https://doi.org/10.1007/bf02664322

Wang M, Luo B, Zhou W, Huang Y, Liu L (2011) The application of Gini coefficient in the total load allocation in Jinjiang river, China. IEEE 2:992–995. https://doi.org/10.1109/ISWREP.2011.5893179

Wang F, Li Y, Yang J, Sun Z (2016) Application of wasp model and Gini coefficient in total mass control of water pollutants: a case study in Xicheng canal, China. Desalin Water Treat 57:2903–2916

Wu WJ, Gao PQ, Xu QM (2019) How to allocate discharge permits more fairly in China?-a new perspective from watershed and regional allocation comparison on socio-natural equality. Sci Total Environ 684:390–401

Xu JP, Zhang MX, Zeng ZQ (2016) Hybrid nested particle swarm optimization for a waste load allocation problem in river system. J Water Resour Plan Manag ASCE 142:19

Xu JP, Hou SH, Yao LM, Li CZ (2017) Integrated waste load allocation for river water pollution control under uncertainty: a case study of Tuojiang river, China. Environ Sci Pollut Res 24:17741–17759

Yan HZ (2003) Tehnieal Anaylsis and Eeonomic comparison of the projects to treat the sewage in small towns. Dissertation, Wuhan University of Technology. (in Chinese)

Zewdie M, Bhallamudi SM (2012) Multi-objective management model for waste-load allocation in a tidal river using archive multi-objective simulated annealing algorithm. Civ Eng Environ Syst 29:222–230

Zhang R, Qian X, Li HM, Yuan XC, Ye R (2012a) Selection of optimal river water quality improvement programs using QUAL2K: a case study of Taihu Lake Basin, China. Sci Total Environ 431:278–285

Zhang Y, Wang XY, Zhang ZM, Shen BG (2012b) Multi-level waste load allocation system for xi'an-Xianyang section, weihe river. In: Yang Z, Chen B (eds) 18th biennial ISEM conference on ecological modelling for global change and coupled human and natural system, vol 13. Elsevier Science Bv, Amsterdam, pp 943–953

Zhang MX, Ni JN, Yao LM (2018) Pigovian tax-based equilibrium strategy for waste-load allocation in river system. J Hydrol 563:223–241

Zhao CC, Ma HR, Yang XY, Liu LT (2008) Analysis of water environmental organic pollution load and its capacity in Xianyang section of Weihe River. Environ Sci Technol 31:65–67. https://doi.org/10.19672/j.cnki.1003-6504.2008.08.016 (in Chinese)

Zhou XD, Guo JL, Cheng W, Song C, Cao G (1999) The comparison of the environmental capacity calculation methods. J Xi'an Univ Technol 15:1–6. https://doi.org/10.19322/j.cnki.issn.1006-4710.1999.03.001 (in Chinese)