Environmental Engineering and Management Journal

February 2019, Vol.18, No. 2, 329-348
http://www.eemj.icpm.tuiasi.ro/; http://www.eemj.eu

"Gheorghe Asachi" Technical University of Iasi, Romania

A SUPPORT VECTOR REGRESSION AND MONTE CARLO SIMULATION - BASED INTERVAL TWO-STAGE PROGRAMMING FOR ENVIRONMENTAL SYSTEMS PLANNING IN BEIJING

Chao Dai[1], Yongping Li[2], Yanpeng Cai[3*], Wei Sun[2], Guohe Huang[2]

[1]Department of Environmental Sciences, Peking University, Beijing 100871, China
[2]Institute for Energy, Environment and Sustainable Communities, University of Regina, 120, 2 Research Drive, Regina, Saskatchewan S4S 7H9, Canada
[3]State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing 100875, China

Abstract

In this research, a support vector regression and Monte Carlo simulation-based interval two-stage programming (SVRMC-ITSP) method was developed through integrating support vector regression (SVR) and Monte Carlo simulation into a two-stage interval-stochastic programming (TISP) framework. The developed SVRMC-ITSP method can effectively tackle dynamic, interactive and uncertain characteristics of municipal solid waste (MSW) management systems. It can also be used to simulate waste generation rates to provide relevant their PDFs for the consequent optimization. The method can improve previous studies in terms of uncertainty reflection and simulation. The developed method was applied for planning of Beijing's MSW management system. The results indicated that the SVRMC-ITSP method performed better than the original TISP model in its capability of improving the credibility of computed results. Uncertainties that can be expressed as both intervals and PDFs can be tackled through the introduction of SVR, Monte Carlo simulation, and TISP. The approach was valuable for supporting the adjustment of the existing waste allocation patterns, and the capacity planning of the city's waste management system.

Keywords: Monte Carlo, municipal solid waste, support vector regression, two-stage stochastic programming

Received: April, 2014; Revised final: June, 2015; Accepted: June, 2015; Published in final edited form: February, 2019

1. Introduction

The amount of municipal solid waste (MSW) generated annually in China is increasing rapidly (Yang et al., 2011). Averagely, China is annually producing over 150 million tons of MSW. Also, this amount is still increasing at an annual rate of 8-10%, resulting in many challenges to decision makers for wisely managing MSW (Cheng et al., 2007). However, multiple uncertainties may exist in real-world MSW management systems, such as cost parameters, capacity limits, and waste-generation rates (Chen et al., 2011; Gunalay et al., 2012). Therefore, models in which inherent uncertainties can be effective handled are increasingly being relied upon to inform and support MSW management in China.

Two-stage interval-stochastic programming (TISP) model (Maqsood and Huang, 2003) is effective for handling uncertainties expressed as probability distribution functions (PDFs) and discrete intervals (Li et al., 2008; Sun et al., 2010). The TISP can generate desired alternatives for MSW management system under multiply uncertainties. However, it may encounter difficulties due to the requirements on PDFs of relevant parameters such as waste-generation levels. Comparatively, Monte Carlo (MC) simulation is effective in tackling uncertainties described as PDFs. The major components of a MC simulation

*Author to whom all correspondence should be addressed: e-mail: yanpeng.cai@iseis.org; yanpeng.cai@bnu.edu.cn

include a numerical model, a probability distribution function, a random number generator, a sampling rule, a scoring value, an error estimator, as well as variance reduction techniques (Qin and Huang, 2009). An essential issue to obtain good MC simulation results is that the numerical model has a high computational accuracy.

In general, traditional numerical methods for predicting waste generation mostly counted on demographic factors, which may be taken as fixed over time or might be projected to change with time (Li et al., 2008). This may cause major disparity for large-scale MSW management systems such as the one in Beijing that is characterized by large population, rapid economic development, and large-scale consumptions. Also, normally, three typical factors (i.e., total population (TP) as a direct cause of MSW arises, gross domestic product (GDP) as the widely-available economic parameter indicating the ability to consume and produce goods, and consumption level of individual residents (CLIR) assessing the level of consumerism) can be considered as the major driving factors to the waste generation (Ali Abdoli et al., 2012; Dyson and Chang, 2005). They may fluctuate frequently over the planning time. As a result, the relationship between response variable (i.e., waste generation level) and the explanatory variables (i.e., TP, GDP and CLIR) is quite complicated with nonlinear and dynamic properties.

Intelligence models, such as artificial neural networks (ANN) and support vector regression (SVR), are the potential tools for forecasting the quantity of waste generation as the input of optimization model (Chen et al., 2017; Noori et al., 2009). These data-driven models can model waste generation fluctuations without having to depend on the complete perception of MSW generation process (Abbasi et al., 2012; Ghinea et al., 2016). Kumar et al., (2011) used the ANN model to estimate the quantity of waste generation generated in the Eluru city, India from 2010 to 2026, where results showed good coincidence between observed data and predictions. However, the ANN may suffer from many disadvantages, including the need for a large number of training samples, difficulty in obtaining a stable solution and the danger of over-fitting.

Compared to ANN models, SVR is powerful for dealing with the problems with small quantity of statistic samples, nonlinearity, high dimension, and local minima (Vapnik, 1999). Moreover, it is a novel learning machine based on statistical learning theory, and which adheres to the principle of structural risk minimization seeking to minimize an upper bound of the generalization error, rather than minimize the training error (the principle followed by ANN) (Bouhouche et al., 2012; Chen and Wang, 2007; Wieland et al., 2010). Recently, SVR has achieved great success in nonlinear regression problems due to its many attractive features and promising generalization performance. However, few researches that focused on SVR technique for the prediction of waste generation level were reported.

Therefore, one potential approach for improving the modeling calculation accuracy and reflecting uncertainties expressed as PDFs and discrete intervals for MSW management systems is to introduce the SVR and MC simulation into the TISP framework. This will lead to a SVR and MC interval two-stage programming (SVRMC-ITSP) method. The proposed SVRMC-ITSP method will be used for planning MSW management system in Beijing, China.

2. Methodology

When uncertainties of the model's right-hand sides are expressed as random variables and decisions need to be made periodically over time, the study system can be formulated as the following two-stage stochastic programming model (Li et al., 2008) (Eq. 1):

$$\text{Min} f = \sum_{j=1}^{n_1} c_j x_j + \sum_{j=1}^{n_2}\sum_{h=1}^{v} p_h d_j y_{jh} \tag{1}$$

subject to Eqs. (2-5):

$$\sum_{j=1}^{n_1} a_{rj} x_j \le b_r,\ r = 1,\ 2,\ ...,\ m_1 \tag{2}$$

$$\sum_{j=1}^{n_1} a_{tj} x_j + \sum_{j=1}^{n_2} a'_{tj} y_{jh} \ge \omega_h,\ t = 1,\ 2,\ ...,\ m_2,\ h = 1,\ 2,\ ...,\ v \tag{3}$$

$$x_j \ge 0,\ j = 1,\ 2,\ ...,\ n_1 \tag{4}$$

$$y_{jh} \ge 0,\ j = 1,\ 2,\ ...,\ n_2,\ h = 1,\ 2,\ ...,\ v \tag{5}$$

where x_j and y_{jh} are the first- and second-stage decision variables respectively; ω_h are random variables with probability levels p_h, where h = 1, 2,…, v and $\sum p_h = 1$. Through introducing the interval parameter programming to quantify those uncertainties presented in terms of interval values, the two-stage stochastic programming model can be transformed into the following two-stage interval-stochastic programming (TISP) model (Maqsood and Huang, 2003) (Eq. 6):

$$\text{Min} f^{\pm} = \sum_{j=1}^{n_1} c_j^{\pm} x_j^{\pm} + \sum_{j=1}^{n_2}\sum_{h=1}^{v} p_h d_j^{\pm} y_{jh}^{\pm} \tag{6}$$

subject to Eqs. (7-10):

$$\sum_{j=1}^{n_1} a_{rj}^{\pm} x_j^{\pm} \le b_r^{\pm},\ r = 1,2,...,m_1 \tag{7}$$

$$\sum_{j=1}^{n_1} a_{tj}^{\pm} x_j^{\pm} + \sum_{j=1}^{n_2} a'^{\pm}_{tj} y_{jh}^{\pm} \ge \omega_h^{\pm},\ t = 1,2,...,m_2;\ h = 1,2,...,v \tag{8}$$

$$x_j^{\pm} \ge 0,\ j = 1,2,...,n_1 \tag{9}$$

$$y_{jh}^{\pm} \geq 0,\ j = 1,2,...,n_2;\ h = 1,2,...,v \tag{10}$$

where: superscript "±" means interval-valued feature; the "–" and "+" superscripts represent lower and upper bounds of an interval parameter/variable, respectively. The TISP model has numerical solutions only when the PDFs of the random variable (i.e. $\omega_h^{\pm}$) is known. In MSW management systems, the waste generation rate (denoted as $\omega_h^{\pm}$) is a random variable. The PDFs of waste generation rate can be obtained through MC simulation. Also, SVR is a numerical model and can assist MC simulation in producing hundreds or thousands of possible outcomes of waste generation over the planning horizon. This approach that SVR and MC simulation are incorporated into the TISP framework can be named as SVR and MC based interval two-stage programming (SVRMC-ITSP) method. Fig. 1 illustrates the general framework of the SVRMC-ITSP method. The solution algorithm can be summarized as follows:

Step 1: The topological structure of three inputs to one output for the SVR model is constructed, where three inputs of SVR include TP, GDP and CLIR and one output means the waste generation rate. As a constructive learning procedure based on the structural risk minimization principle with the aims at minimizing the true error, the detailed descriptions of SVR model can be found in Smola and Schölkopf, (2004).

Step 2: The surveyed records of TP, GDP, CLIR and waste generation rate can be used to calibrate and validate the SVR model constructed. The data point from 1979 to 2003 can be regarded as the training sets, while the data point from 2004 to 2008 can be regard as the testing sets. The purpose of calibration and validation is to set the regularization constant C and the kernel parameters. The optimal values of parameter C are determined by employing a grid-search in a n-fold cross validation approach, where the searching range was designed as follows: $C = 2^q$, q=-10, -9, …, 0, …, 9, 10. Additionally, the polynomial kernel can help SVR enhance the accuracy for predicting the waste generation rate in Beijing (Dai et al., 2011).

Step 3: In this study, three random variables (i.e. G_v (v = 1, 2, 3)), including the growth rate of TP (G_1), the growth rate of GDP (G_2) and the growth rate of CLIR (G_3) are assumed to be independent random variables and normally distributed. Therefore, Kolmogorov-Smirnov (KS) test can be used to estimate the probability distribution functions, means, and standard deviations of random variables G_v (Massey, 1951). The data for statistical analysis is the time series of G_v collected from 1980 to 2008. The null and the alternative hypotheses for the KS-test are: H_0: $G_v : N(\mu_v, \sigma_v^2)$, and H_1: not H_0.

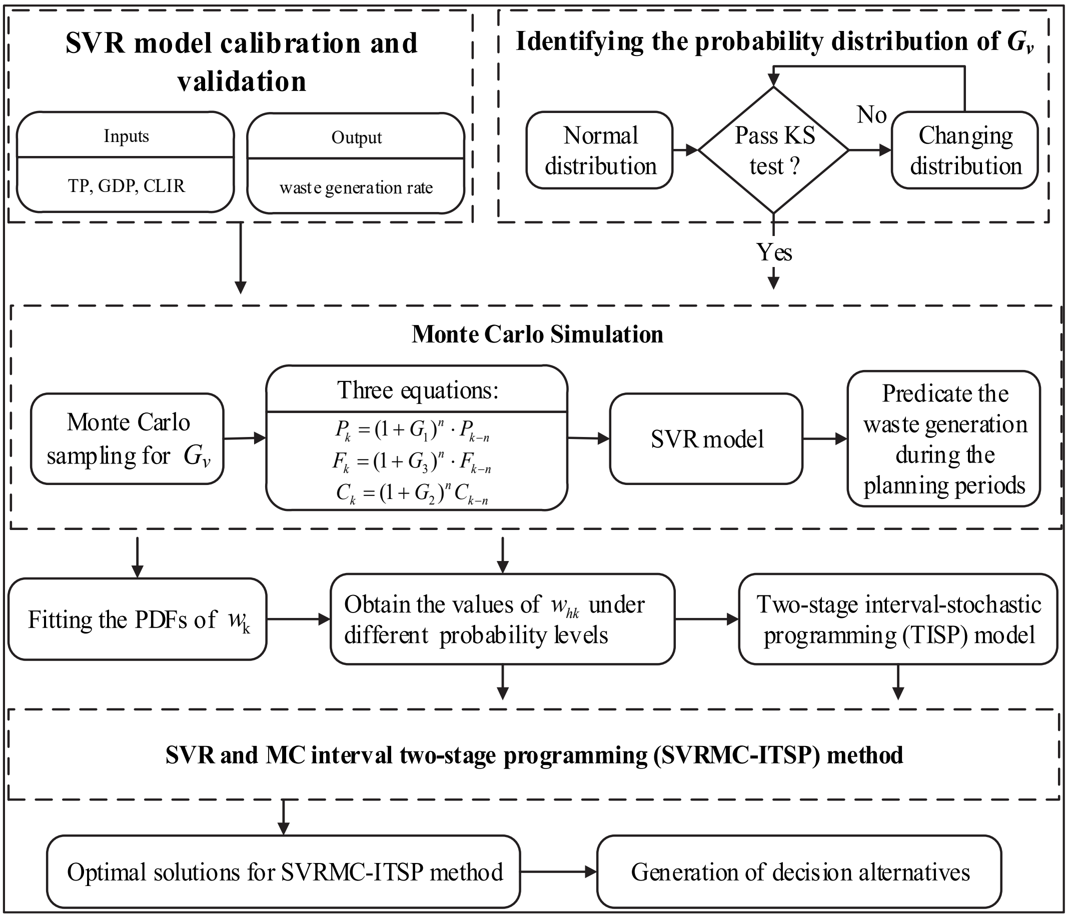


Fig. 1. Framework of SVRMC-ITSP

Step 4: If normality test fails, non-normal (e.g., uniform and gamma) distributions may be assumed and tested for the G_v. If normality test passes, Monte Carlo random samples for each probability distribution of G_v can be undertaken.

Step 5: Data points of TP, GDP and CLIR from 2011 to 2015 as the inputs of SVR model are calculated through the formulas of geometric series: $P_k = (1+G_1)^n \cdot P_{k-n}$, $F_k = (1+G_3)^n \cdot F_{k-n}$ and $C_k = (1+G_2)^n \cdot C_{k-n}$, where P_k, F_k and C_k are the value of TP, GDP and CLIR in period k, respectively; n denotes the interval time; and G_1, G_2 and G_3 (i.e. G_v) are generated from Monte Carlo random samples. The output of the verified SVR model is the waste generation rate from 2011 to 2015, where the training sets are from 1979 to 2008.

Step 6: Steps 4 and 5 are repeated N times. Thus, N samples of annual waste generation rate from 2011 to 2015 can be obtained. KS-test is used to identify the distributions of annual N samples, where three candidate distributions, including Normal and Lognormal distributions, are considered first.

Step 7: If KS-test fails, other distributions (i.e., gamma distribution) may be assumed and tested for fitting the PDFs of waste generation rate in the planning horizon. If passes, make a discretization for the fitted PDFs and calculate the interval values (i.e., waste generation rate, $\omega_{kh}^{\pm}$) under given probability levels. Fig. 2 shows the relationship between PDFs and intervals, where f(x) is the probability distribution function; $w_{k,\min}$ and $w_{k,\max}$ (the minimum and maximum value among the N samples, respectively) are the lower and upper bound of annual waste generation rate, respectively. Suppose the interval $\left[w_{k,\min}, w_{k,\max}\right]$ are divided into v sections with the probability levels $P_h = \int_{w_{kh}^-}^{w_{kh}^+} f(x)dx$, h = 1, 2,…, v. Specially, $P_L = \int_{-\infty}^{w_{k,\min}} f(x)dx$ and $P_R = \int_{w_{k,\min}}^{+\infty} f(x)dx$ are much less than P_h. Because $\sum_{h=1}^{v} P_h = 1$ is a requirement for calculating the two-stage stochastic programming model, P_L and P_R can be approximately integrated into P_1 and P_v, respectively. Then, we have Eqs. (11-13):

$$P_1 = \int_{-\infty}^{w_{k1}^-} f(x)dx + \int_{w_{k1}^-}^{w_{k1}^+} f(x)dx = \int_{-\infty}^{w_{k1}^+} f(x)dx \tag{11}$$

$$P_h = \int_{w_{kh}^-}^{w_{kh}^+} f(x)dx,\ h = 2,3,...v-1 \tag{12}$$

$$P_v = \int_{w_{k,v}^-}^{w_{k,v}^+} f(x)dx + \int_{w_{k,v}^+}^{+\infty} f(x)dx = \int_{w_{k,v}^-}^{+\infty} f(x)dx \tag{13}$$

Assume Z_α is the underside quantile of $f(x)$, and $\alpha = \int_{-\infty}^{Z_\alpha} f(x)dx$. According to the quantile calculation principle, under given probability levels, the interval waste generation rate can be calculated as follows Eqs. (14-16):

$$w_{k1}^- = w_{k,\min} \tag{14}$$

$$w_{k,h}^+ = w_{k,h+1}^- = Z_{\alpha_h},\ a_h = \sum_{h=1}^{h} P_h,\ h = 1,2,...,v-1 \tag{15}$$

$$w_{k,v}^+ = w_{k,\max} \tag{16}$$

Step 8: Import the interval waste generation rate calculated into TISP model. Then, TISP can be transformed into two deterministic sub models that correspond to the lower and upper bounds of desired objective function value (Huang et al., 1992). The objective function value corresponding to f^- is desired first because the objective is to minimize net system costs. The sub model for f^- can be firstly formulated as follows (assume that $b^{\pm} \geq 0$, and $f^{\pm} \geq 0$) (Eq. 17):

$$\text{Min } f^- = \sum_{j=1}^{k_1} c_j^- x_j^- + \sum_{j=k_1+1}^{n_1} c_j^- x_j^+ + \sum_{j=1}^{k_2}\sum_{h=1}^{v} p_{jh} d_j^- y_{jh}^- + \sum_{j=k_2+1}^{n_2}\sum_{h=1}^{v} p_{jh} d_j^- y_{jh}^+ \tag{17}$$

subject to Eqs. (18-23):

$$\sum_{j=1}^{k_1} \left|a_{rj}\right|^+ \text{sign}(a_{rj}^+) x_j^- + \sum_{j=k_1+1}^{n_1} \left|a_{rj}\right|^- \text{sign}(a_{rj}^-) x_j^+ \leq b_r^+,\ \forall\ r \tag{18}$$

$$\sum_{j=1}^{k_1} \left|\text{a}_{tj}\right|^+ \text{sign}(\text{a}_{tj}^+)\text{x}_j^- + \sum_{j=k_1+1}^{n_1} \left|\text{a}_{tj}\right|^- \text{sign}(a_{tj}^-) x_j^+ + \sum_{j=1}^{k_2} \left|a_{tj}'\right|^+ \text{sign}(a_{tj}'^+) y_{jh}^- + \sum_{j=k_2+1}^{n_2} \left|a_{tj}'\right|^- \text{sign}(a_{tj}'^-) y_{jh}^+ \geq \omega_h^-,\ \forall\ t,\ h \tag{19}$$

$$x_j^- \geq 0,\ j = 1,\ 2,\ ...,\ k_1 \tag{20}$$

$$x_j^+ \geq 0,\ j = k_1+1,\ k_1+2,\ ...,\ n_1 \tag{21}$$

$$y_{jh}^- \geq 0,\ \forall\ h,\ j = 1,\ 2,\ ...,\ k_2 \tag{22}$$

$$y_{jh}^+ \geq 0,\ \forall\ h,\ j = k_2+1,\ k_2+2,\ ...,\ n_2 \tag{23}$$

where $x_j^{\pm}$, j = 1, 2,…, k_1, are interval variables with positive coefficients in the objective function; $x_j^{\pm}$, j = k_1 + 1, k_1 + 2,…, n_1 are interval variables with negative coefficients; $y_{jh}^{\pm}$, j = 1, 2,…, k_2 and h = 1, 2,…, v, are random variables with positive coefficients in the objective function; $y_{jh}^{\pm}$, j = k_2 + 1, k_2 + 2,…, n_2 and h = 1, 2,…, v, are random variables with negative coefficients.

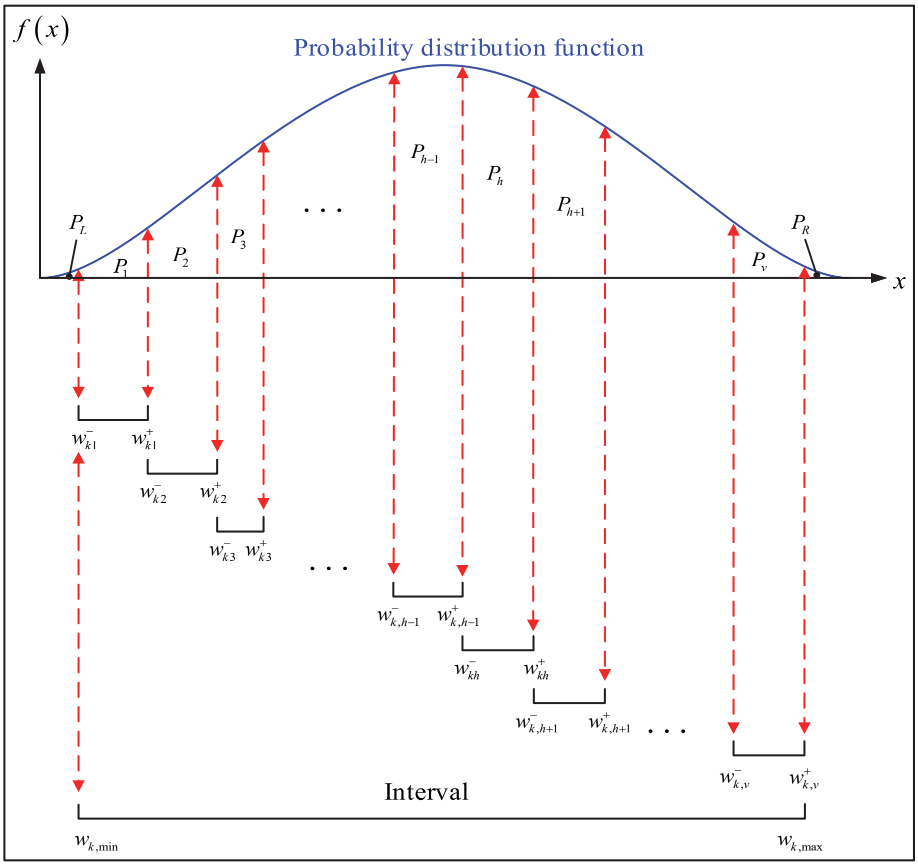


Fig. 2. Relationship between probability distribution function and interval

Solutions of $x^-_{j\,\text{opt}}$ (j = 1, 2,…, k_1), $x^+_{j\,\text{opt}}$ ($j = k_1$ + 1, k_1 + 2,…, n_1), $y^-_{jh\,\text{opt}}$ (j = 1, 2,…, k_2), and $y^+_{jh\,\text{opt}}$ ($j = k_2$ + 1, k_2 + 2,…, n_2) can be obtained through Eqs. (17-23).
Based on the above solutions, the second sub model for f^+ can be formulated as follows (Eq. 24):

$$\text{Min } f^+ = \sum_{j=1}^{k_1} c_j^+ x_j^+ + \sum_{j=k_1+1}^{n_1} c_j^+ x_j^- + \sum_{j=1}^{k_2}\sum_{h=1}^{v} p_{jh} d_j^+ y_{jh}^+ + \sum_{j=k_2+1}^{n_2}\sum_{h=1}^{v} p_{jh} d_j^+ y_{jh}^- \tag{24}$$

subject to Eqs (25-30):

$$\sum_{j=1}^{k_1} |a_{rj}|^- \operatorname{sign}(a_{rj}^-) x_j^+ + \sum_{j=k_1+1}^{n_1} |a_{rj}|^+ \operatorname{sign}(a_{rj}^+) x_j^- \le b_r^-, \ \forall r \tag{25}$$

$$\sum_{j=1}^{k_1} |a_{tj}|^- \operatorname{sign}(a_{tj}^-) x_j^+ + \sum_{j=k_1+1}^{n_1} |a_{tj}|^+ \operatorname{sign}(a_{tj}^+) x_j^- + \sum_{j=1}^{k_2} |a'_{tj}|^- \operatorname{sign}(a'^-_{tj}) y_{jh}^+ + \sum_{j=k_2+1}^{n_2} |a'_{tj}|^+ \operatorname{sign}(a'^+_{tj}) y_{jh}^- \ge \omega_h^+, \ \forall t,\ h \tag{26}$$

$$x_j^+ \ge x^-_{j\,\text{opt}},\ j = 1,\ 2, \ldots,\ k_1 \tag{27}$$

$$0 \le x_j^- \le x^+_{j\,\text{opt}},\ j = k_1 + 1,\ k_1 + 2,\ \ldots,\ n_1 \tag{28}$$

$$y_{jh}^+ \ge y^-_{jh\,\text{opt}},\ \forall\ h,\ j = 1,\ 2,\ \ldots,\ k_2 \tag{29}$$

$$0 \le y_{jh}^- \le y^+_{jh\,\text{opt}},\ \forall\ h,\ j = k_2 + 1,\ k_2 + 2,\ \ldots,\ n_2 \tag{30}$$

Solutions of $x^+_{j\,\text{opt}}$ (j = 1, 2,…, k_1), $x^-_{j\,\text{opt}}$ ($j = k_1$ + 1, k_1 + 2,…, n_1), $y^+_{jh\,\text{opt}}$ (j = 1, 2,…, k_2), and $y^-_{jh\,\text{opt}}$ ($j = k_2$ + 1, k_2 + 2,…, n_2) can be obtained through Eqs. (24-30).
Step 9: Through integrating solutions of the two sub models, solution results for TISP model can be obtained.

3. Case Study

The urban districts of Beijing cover an area of approximately 735 km^2. Its population is approximately 14.39 million. The study area is the urban districts of Beijing, which contains six districts (i.e., Dongcheng, Xicheng, Chaoyang, Haidian, Fengtai, and Shijingshan). The waste generated in the urban districts is collected by residents first and then shipped to transfer stations for pretreatment, and finally it is allocated to waste management facilities. There are many types of trucks for transporting the MSW in the eight urban districts. For example, trucks equipped with a loader crane are used to serve the Chaoyang and Haidian districts after Beijing 2008 Olympic Games, which have the collection ability of [4.8, 5.4] tonnes at every turn. The capacity of the majority trucks in other districts is about [1.8, 2.2] tonnes. Consequently, a general truck capacity of [2.6, 3.0] tonnes will be used to estimate the transportation costs.

The main approaches for waste treatment include landfilling, composting, incineration and comprehensive treatment. There are seven landfills (i.e., Asuwei, Liulitun, Jiaojiapo, Wangzuo, Anding, Dadushe and Gaoantun) around the six urban districts for disposing the MSW. These landfills have an overall cumulative capacity limit. According to the Beijing Environmental Sanitation Development Report, the existing landfills are able to accept waste till 2015. The four composting facilities with the total designing capacity of 4250 tonne/day are used to service the study system. The residue ((23, 28) % of original waste) from composting facilities will be finally disposed by landfills. Besides, revenue can be obtained from compost produced by the adoption of the composting technology. There will be four incinerators used to service the study system with the total designing capacity of 4600 tonne/day in the beginning of 2011. The waste flow can be reduced to 78, 82% by incinerating technology, and the residue will be finally transported to landfill. In addition, the combustion heat could be collected to generate power or provided to local resident for heat supply. Revenue can be obtained from electrical power/heating generated by incinerator heat. Nowadays, there are three comprehensive treatment facilities to service the urban districts with the total designing capacity of 4200 tonne/day. They updated the recycling facilities many years ago due to the two principal functions, such as compaction and recycling the MSW. After the adoption of an integrated approach, perishable goods and combustible materials can be used. Thus, the residue ((10, 15) % of original waste) from comprehensive treatment facilities will be finally disposed by landfills.

According to the Beijing statistical yearbook from 1980 to 2009, the data of the time series of waste generation rate, TP, CLIR and GDP from 1979 to 2008 can be collected. The study time horizon is 5 years (from 2011 to 2015), which is further divided into three periods (i.e., the first period is 2011, the second period includes 2012 and 2013, and the third period is from 2014 to 2015). Tables 1 and 2 contain regular costs for allowable waste flows, operating costs for waste management facilities, and penalty costs for surplus waste flows, as well as revenues from waste management facilities over the three planning periods (Xi et al., 2010).

The capacity-expansion options and relevant capital costs for different waste management facilities are shown in Table 3. It is indicated that the expansion costs decrease along with time. Since the planning problem under consideration is dynamic with multiple stages, discount factors are necessary for each period to obtain a total present value for the object function. In this research, all the cash flows are counted in year 2010 dollars. The customized model can be presented as follows (Eq. 31):

$$
\begin{aligned}
\mathrm{Min}\, f^{\pm} = & \sum_{c=1}^{4}\sum_{k=1}^{3} L_k TUC_{ck}^{\pm}\begin{bmatrix} COC_k^{\pm} + DUC_c^{\pm}TR_k^{\pm} + OC_{ck}^{\pm} - \\ -(1-FC_{ck}^{\pm})RC_{ck}^{\pm} \end{bmatrix} \\
& + \sum_{i=1}^{4}\sum_{k=1}^{3} L_k TUI_{ik}^{\pm}\begin{bmatrix} COC_k^{\pm} + DUI_i^{\pm}TR_k^{\pm} + \\ + OI_{ik}^{\pm} - (1-FI_{ik}^{\pm})RI_{ik}^{\pm} \end{bmatrix} \\
& + \sum_{z=1}^{3}\sum_{k=1}^{3} L_k TUZ_{zk}^{\pm}\begin{bmatrix} COC_k^{\pm} + DUZ_z^{\pm}TR_k^{\pm} + \\ + OZ_{zk}^{\pm} - (1-FZ_{zk}^{\pm})RZ_{zk}^{\pm} \end{bmatrix} \\
& + \sum_{l=1}^{7}\sum_{k=1}^{3} \begin{matrix} L_k TUL_{lk}^{\pm}(COC_k^{\pm} + \\ + DUL_l^{\pm}TR_k^{\pm} + OL_{lk}^{\pm}) \end{matrix} + \\
& + \sum_{c=1}^{4}\sum_{l=1}^{7}\sum_{k=1}^{3} \begin{matrix} L_k TCL_{clk}^{\pm}(DCL_{cl}^{\pm}TR_k^{\pm} + \\ + OL_{lk}^{\pm}) \end{matrix} \\
& + \sum_{i=1}^{4}\sum_{l=1}^{7}\sum_{k=1}^{3} L_k TIL_{ilk}^{\pm}(DIL_{il}^{\pm}TR_k^{\pm} + OL_{lk}^{\pm}) + \\
& + \sum_{z=1}^{3}\sum_{l=1}^{7}\sum_{k=1}^{3} L_k TZL_{zlk}^{\pm}(DZL_{zl}^{\pm}TR_k^{\pm} + OL_{lk}^{\pm}) \\
& + \sum_{c=1}^{4}\sum_{k=1}^{3}\sum_{h=1}^{3} L_k P_h XUC_{ckh}^{\pm}\begin{bmatrix} ECOC_k^{\pm} + DUC_c^{\pm}ETR_k^{\pm} + \\ + EOC_{ck}^{\pm} - (1-FC_{ck}^{\pm})ERC_{ck}^{\pm} \end{bmatrix} \\
& + \sum_{i=1}^{4}\sum_{k=1}^{3}\sum_{h=1}^{3} L_k P_h XUI_{ikh}^{\pm}\begin{bmatrix} ECOC_k^{\pm} + DUI_i^{\pm}ETR_k^{\pm} + \\ + EOI_{ik}^{\pm} - (1-FI_{ik}^{\pm})ERI_{ik}^{\pm} \end{bmatrix} \\
& + \sum_{z=1}^{3}\sum_{k=1}^{3}\sum_{h=1}^{3} L_k P_h XUZ_{zkh}^{\pm}\begin{bmatrix} ECOC_k^{\pm} + DUZ_z^{\pm}ETR_k^{\pm} + \\ + EOZ_{zk}^{\pm} - (1-FZ_{zk}^{\pm})ERZ_{zk}^{\pm} \end{bmatrix} \\
& + \sum_{l=1}^{7}\sum_{k=1}^{3}\sum_{h=1}^{3} \begin{matrix} L_k P_h XUL_{lkh}^{\pm}(ECOC_k^{\pm} + DUL_l^{\pm}ETR_k^{\pm} + \\ + EOL_{lk}^{\pm}) \end{matrix} + \\
& + \sum_{c=1}^{4}\sum_{l=1}^{7}\sum_{k=1}^{3} L_k YCL_{clk}^{\pm}(DCL_{cl}^{\pm}ETR_k^{\pm} + EOL_{lk}^{\pm}) \\
& + \sum_{i=1}^{4}\sum_{l=1}^{7}\sum_{k=1}^{3} L_k YIL_{ilk}^{\pm}(DIL_{il}^{\pm}ETR_k^{\pm} + EOL_{lk}^{\pm}) + \\
& + \sum_{z=1}^{3}\sum_{l=1}^{7}\sum_{k=1}^{3} L_k YZL_{zlk}^{\pm}(DZL_{zl}^{\pm}ETR_k^{\pm} + EOL_{lk}^{\pm}) \\
& + \sum_{c=1}^{4}\sum_{e=1}^{3}\sum_{k=1}^{3} ECE_{cek}^{\pm}BCE_{cek}^{\pm} + \\
& + \sum_{i=1}^{4}\sum_{f=1}^{3}\sum_{k=1}^{3} EIF_{ifk}^{\pm}BIF_{ifk}^{\pm} + \\
& + \sum_{z=1}^{3}\sum_{g=1}^{3}\sum_{k=1}^{3} EZG_{zgk}^{\pm}BZG_{zgk}^{\pm}
\end{aligned}
\tag{31}
$$

subject to (Eq. 32-55):

$$
\sum_{k=1}^{3} L_k \left\{ \begin{matrix} (TUL_{lk}^{\pm} + XUL_{lkh}^{\pm}) + \sum_{c=1}^{4}(TCL_{clk}^{\pm} + YCL_{clk}^{\pm}) + \\ + \sum_{i=1}^{4}(TIL_{ilk}^{\pm} + YIL_{ilk}^{\pm}) \end{matrix} \right. + \left. + \sum_{z=1}^{3}(TZL_{zlk}^{\pm} + YZL_{zlk}^{\pm}) \right\} \le \delta_l^{\pm} LC_l^{\pm};\ \forall\, h,\ l,
\tag{32}
$$

(Landfill capacity constraint)(Eq. 32)

$$
TUC_{ck}^{\pm} + XUC_{ckh}^{\pm} \le \zeta_c^{\pm}\left\{ CC_c^{\pm} + \sum_{e=1}^{3}\sum_{k=1}^{k} \Delta CCE_{cek} BCE_{cek}^{\pm} \right\}, \\ ,\forall\, h,\ c, k,
\tag{33}
$$

(Composting facility capacity constraint) (Eq. 33)

$$TUI_{ik}^{\pm} + XUI_{ikh}^{\pm} \le \kappa_i^{\pm}\left\{IC_i^{\pm} + \sum_{f=1}^{3}\sum_{k=1}^{k}\Delta CIF_{ifk}BIF_{ifk}^{\pm}\right\}, \ \forall\ h,\ i, k, \quad (34)$$

(Incinerator capacity constraint) (Eq. 34)

$$TUZ_{zk}^{\pm} + XUZ_{zkh}^{\pm} \le \psi_z^{\pm}\left\{ZC_z^{\pm} + \sum_{g=1}^{3}\sum_{k=1}^{k}\Delta CZG_{zgk}BZG_{zgk}^{\pm}\right\}, \ \forall\ h,\ z, k, \quad (35)$$

(Comprehensive treatment facility capacity constraint)(Eq 35)

$$\begin{aligned}&\sum_{c=1}^{4}L_k(TUC_{ck}^{\pm} + XUC_{ckh}^{\pm}) + \\ &+\sum_{i=1}^{4}L_k(TUI_{ik}^{\pm} + XUI_{ikh}^{\pm}) \\ &+\sum_{z=1}^{3}L_k(TUZ_{zk}^{\pm} + XUZ_{zkh}^{\pm}) + \\ &+\sum_{l=1}^{7}L_k(TUL_{lk}^{\pm} + XUL_{lkh}^{\pm}) \ge \alpha_k^{\pm}w_{kh}^{\pm},\ \forall\ k,\ h,\end{aligned} \quad (36)$$

(Mass balance constraint) (Eq. 36)

$$FC_{ck}^{\pm}TUC_{ck}^{\pm} = \sum_{l=1}^{7}TCL_{clk}^{\pm},\ \forall\ z,\ k, \quad (37)$$

$$FI_{ik}^{\pm}TUI_{ik}^{\pm} = \sum_{l=1}^{7}TIL_{ilk}^{\pm},\ \forall\ i,\ k, \quad (38)$$

$$FZ_{zk}^{\pm}TUZ_{zk}^{\pm} = \sum_{l=1}^{7}TZL_{zlk}^{\pm},\ \forall\ z,\ k, \quad (39)$$

$$\sum_{l=1}^{7}YCL_{clk}^{\pm} = \sum_{h=1}^{3}P_hFC_{ck}^{\pm}XUC_{ckh}^{\pm},\ \forall\ c,\ k, \quad (40)$$

$$\sum_{l=1}^{7}YIL_{ilk}^{\pm} = \sum_{h=1}^{3}P_hFI_{ik}^{\pm}XUI_{ikh}^{\pm},\ \forall\ i,\ k, \quad (41)$$

$$\sum_{l=1}^{7}YZL_{zlk}^{\pm} = \sum_{h=1}^{3}P_hFZ_{zk}^{\pm}XUZ_{zkh}^{\pm},\ \forall\ z,\ k, \quad (42)$$

(Residue constraint)Eqs. (37-42)

$$\sum_{c=1}^{4}L_k(TUC_{ck}^{\pm} + XUC_{ckh}^{\pm}) \ge\ GC_K^{\pm}\alpha_k^{\pm}w_{kh}^{\pm},\ \forall\ k,\ h, \quad (43)$$

$$\sum_{i=1}^{4}L_k(TUI_{ik}^{\pm} + XUI_{ikh}^{\pm}) \ge\ GI_K^{\pm}\alpha_k^{\pm}w_{kh}^{\pm},\ \forall\ k,\ h, \quad (44)$$

$$\sum_{z=1}^{3}L_k(TUZ_{zk}^{\pm} + XUZ_{zkh}^{\pm}) \ge\ GZ_K^{\pm}\alpha_k^{\pm}w_{kh}^{\pm},\ \forall\ k,\ h, \quad (45)$$

(Diversion rate) Eqs. (43-45)

$$TUC_{ck}^{\pm} \ge XUC_{ckh}^{\pm} \ge 0,\ \forall\ c,\ k,\ h, \quad (46)$$

$$TUI_{ik}^{\pm} \ge XUI_{ikh}^{\pm} \ge 0,\ \forall\ i,\ k,\ h, \quad (47)$$

$$TUZ_{zk}^{\pm} \ge XUZ_{zkh}^{\pm} \ge 0,\ \forall\ z,\ k,\ h, \quad (48)$$

$$TUL_{lk}^{\pm} \ge XUL_{lkh}^{\pm} \ge 0,\ \forall\ l,\ k,\ h, \quad (49)$$

(Non-negativity constraint) Eqs. (46-49)

$$\sum_{e=1}^{3}BCE_{cek}^{\pm} \le 1,\ \forall\ c,\ k, \quad (50)$$

$$\sum_{f=1}^{3}BIF_{ifk}^{\pm} \le 1,\ \forall\ i,\ k, \quad (51)$$

$$\sum_{g=1}^{3}BZG_{zgk}^{\pm} \le 1,\ \forall\ z,\ k, \quad (52)$$

(waste management facilities can be expanded once within each period k) Eqs. (50-52)

$$BCE_{cek}^{\pm}\begin{cases}=1, \text{ if composting facility is expanded}\\ =0, \text{ if otherwise}\end{cases},\ \forall\ c,\ e,\ k \quad (53)$$

$$BIF_{ifk}^{\pm}\begin{cases}=1, \text{ if incinerator is expanded}\\ =0, \text{ if otherwise}\end{cases},\ \forall\ i, f,\ k \quad (54)$$

$$BZG_{zgk}^{\pm}\begin{cases}=1, \text{ if comprehensive treatment facility is expanded}\\ =0, \text{ if otherwise}\end{cases},\ \forall\ z,\ g,\ k \quad (55)$$

(Binary constraints) Eqs. (53-55)

The detailed nomenclatures for the variables and parameters of Eqs. (31-55) are provided in Appendix.

4. Result analysis

In this research, three models, including SVR, ANN and multiple linear regression (MLR), were considered as the alternative destination to extract as much information as possible from the historical data to identify an appropriate regression relationship between TP, CLIR, GDP and waste generation rate.

Table 1. Costs and revenues for allowable waste flows (present values) (Unit: $/tonne)

	Planning period					
	k = 1		***k = 2***		***k = 3***	
	Cost	***Revenue***	***Cost***	***Revenue***	***Cost***	***Revenue***
For landfill						
Asuwei landfill	(9.27, 9.65)	(3.41, 3.26)	(8.47, 8.85)	(2.91, 2.76)	(7.37, 7.75)	(2.11, 1.96)
Liulitun landfill	(9.22, 9.61)	(3.59, 3.42)	(8.42, 8.81)	(3.09, 2.92)	(7.32, 7.71)	(2.29, 2.12)
Jiaojiapo landfill	(9.96, 9.31)	(3.41, 3.53)	(9.16, 8.51)	(2.91, 3.03)	(8.06, 7.41)	(2.11, 2.23)
Wangzuo landfill	(8.75, 10.38)	(3.45, 3.25)	(7.95, 9.58)	(2.95, 2.75)	(6.85, 8.48)	(2.15, 1.95)
Anding landfill	(9.63, 9.11)	(3.33, 3.29)	(8.83, 8.31)	(2.83, 2.79)	(7.73, 7.21)	(2.03, 1.99)
Dadushe landfill	(8.95, 10.3)	(3.52, 3.56)	(8.15, 9.50)	(3.02, 3.06)	(7.05, 8.40)	(2.22, 2.26)
Gaoantun landfill	(8.66, 10.34)	(3.64, 3.35)	(7.86, 9.54)	(3.14, 2.85)	(6.76, 8.44)	(2.34, 2.05)
For composting facility						
Liulitun composting	(20.09, 20.89)	(6.70, 6.54]	(18.29, 19.09)	(5.90, 5.74)	(16.39, 17.19)	(4.80, 4.64)
Nangong composting	(19.15, 19.92)	(6.48, 6.33]	(17.35, 18.12)	(5.68, 5.53)	(15.45, 16.22)	(4.58, 4.43)
Dongcun composting	(19.30, 20.08)	(6.52, 6.36]	(17.50, 18.28)	(5.72, 5.56)	(15.60, 16.38)	(4.62, 4.46)
Gaoantun composting	(19.22, 19.98)	(6.85, 6.68]	(17.42, 18.18)	(6.05, 5.88)	(15.52, 16.28)	(4.95, 4.78)
For incinerator						
Asuwei incinerator	3(0.98, 31.76)	(8.72, 8.51)	(30.38, 31.16)	(7.92, 7.71)	(29.58, 30.36)	(6.82, 6.61)
Liulitun incinerator	(31.33, 32.11)	(7.65, 7.47)	(30.73, 31.51)	(6.85, 6.67)	(29.93, 30.71)	(5.75, 5.57)
Nangong incinerator	(33.94, 34.79)	(8.66, 8.45)	(33.34, 34.19)	(7.86, 7.65)	(32.54, 33.39)	(6.76, 6.55)
Gaoantun incinerator	(31.91, 32.71)	(8.68, 8.48)	(31.31, 32.11)	(7.88, 7.68)	(30.51, 31.31)	(6.78, 6.58)
For comprehensive treatment facility						
Asuwei comprehensive	(37.29, 38.41)	(11.93, 11.56)	(36.69, 37.81)	(11.03, 10.66)	(35.89, 37.01)	(9.83, 9.46)
Fengtai comprehensive	(40.59, 41.81)	(12.99, 12.58)	(39.99, 41.21)	(12.09, 11.68)	(39.19, 40.41)	(10.89, 10.48)
Xitianyang comprehensive	(37.09, 37.18)	(11.87, 11.19)	(36.49, 36.58)	(10.97, 10.29)	(35.69, 35.78)	(9.77, 9.09)

Table 2. Costs and revenues for excess waste flows (present values) (Unit: $/tonne)

	Planning period					
	k = 1		***k = 2***		***k = 3***	
	Cost	***Revenue***	***Cost***	***Revenue***	***Cost***	***Revenue***
For landfill						
Asuwei landfill	(17.89, 18.63)	(3.41, 3.26)	(16.34, 17.09)	(2.91, 2.76)	(14.22, 14.97)	(2.11, 1.96)
Liulitun landfill	(17.72, 18.46)	(3.59, 3.42)	(16.19, 16.93)	(3.09, 2.92)	(14.07, 14.81)	(2.29, 2.12)
Jiaojiapo landfill	(20.67, 19.33)	(3.41, 3.53)	(19.01, 17.67)	(2.91, 3.03)	(16.73, 15.38)	(2.11, 2.23)
Wangzuo landfill	(15.94, 18.91)	(3.45, 3.25)	(14.48, 17.45)	(.95, 2.75)	(12.48, 15.45)	(2.15, 1.95)
Anding landfill	(19.32, 18.28)	(3.33, 3.29)	(17.72, 16.68)	(2.83, 2.79)	(15.51, 14.47)	(2.03, 1.99)
Dadushe landfill	(16.69, 19.21)	(3.52, 3.56)	(15.20, 17.72)	(3.02, 3.06)	(13.15, 15.67)	(2.22, 2.26)
Gaoantun landfill	(15.63, 18.65)	(3.64, 3.35)	(14.19, 17.21)	(3.14, 2.85)	(12.20, 15.22)	(2.34, 2.05)
For composting facility						
Liulitun composting	(32.05, 33.34)	(6.70, 6.54)	(29.18, 30.46)	(5.90, 5.74)	(26.15, 27.43)	(4.80, 4.64)
Nangong composting	(29.55, 30.73)	(6.48, 6.33)	(26.77, 27.96)	(5.68, 5.53)	(23.84, 25.02)	(4.58, 4.43)
Dongcun composting	(29.96, 31.16)	(6.52, 6.36)	(27.17, 28.36)	(5.72, 5.56)	(24.22, 25.41)	(4.62, 4.46)

Gaoantun composting	(31.33, 32.58)	(6.85, 6.68)	(28.39, 29.65)	(6.05, 5.88)	(25.30, 26.55)	(4.95, 4.78)
For incinerator						
Asuwei incinerator	(48.01, 49.20)	(8.72, 8.51)	(47.07, 48.27)	(7.92, 7.71)	(45.83, 47.03)	(6.82, 6.61)
Liulitun incinerator	(49.07, 50.31)	(7.65, 7.47)	(48.13, 49.36)	(6.85, 6.67)	(46.88, 48.11)	(5.75, 5.57)
Nangong incinerator	(57.61, 59.05)	(8.66, 8.45)	(56.59, 58.03)	(7.86, 7.65)	(55.23, 56.67)	(6.76, 6.55)
Gaoantun incinerator	(50.93, 52.20)	(8.68, 8.48)	(49.97, 51.24)	(7.88, 7.68)	(48.69, 49.97)	(6.78, 6.58)
For comprehensive treatment facility						
Asuwei comprehensive	(59.71, 61.50)	(11.93, 11.56)	(58.74, 60.53)	(11.03, 10.66)	(57.46, 59.25)	(9.83, 9.46)
Fengtai comprehensive	(64.36, 66.29)	(12.99, 12.58)	(63.41, 65.34)	(12.09, 11.68)	(62.14, 64.07)	(10.89, 10.48)
Xitianyang comprehensive	(58.36, 58.48)	(11.87, 11.19)	(57.41, 57.54)	(10.97, 10.29)	(56.15, 56.28)	(9.77, 9.09)

Table 3. Capacity expansion options and their capital costs for waste management facilities

Expansion option	*Total capacity*	*Capacity for residential waste*	*Expansion cost ($\$10^6$)*		
			k = 1	*k = 2*	*k = 3*
Composting facility					
Option 1 (e = 1)	282 t/d	200 t/d	(2.93, 2.99)	(2.82, 2.87)	(2.64, 2.69)
Option 2 (e = 2)	1142 t/d	800 t/d	(11.11, 11.17)	(10.66, 10.72)	(10.01, 10.05)
Option 3 (e = 3)	1428 t/d	1000 t/d	(13.59, 13.65)	(13.05, 13.11)	(12.23, 12.29)
Incinerator					
Option 1 (f = 1)	282 t/d	200 t/d	(18.67, 19.87)	(17.55, 18.67)	(15.87, 16.89)
Option 2 (f = 2)	1142 t/d	800 t/d	(71.80, 73.02)	(67.49, 68.62)	(61.03, 62.05)
Option 3 (f = 3)	1428 t/d	100 t/d	(87.95, 89.15)	(82.67, 83.80)	(74.76, 75.78)
Comprehensive treatment facility					
Option 1 (g = 1)	571 t/d	400 t/d	(24.42, 25.54)	(23.21, 24.26)	(21.49, 22.47)
Option 2 (g = 2)	1142 t/d	800 t/d	(47.33, 48.45)	(44.97, 46.03)	(41.65, 42.64)
Option 3 (g = 3)	1714 t/d	1200 t/d	(68.19, 69.31)	(64.78, 65.84)	(60.01, 60.99)

Also, the root-mean-square error (RMSE) was used to assess the fitting and generalization performances for the three models. Fig. 3 shows the comparison results of the observed value and predicted value calculated by SVR, ANN and MLR. For the training sets, the RMSE of SVR, ANN and MLR are 0.01, 0.02, and 0.13 respectively. For the testing sets, the RMSE of SVR, ANN and MLR are 0.05, 3.77, and 0.17 respectively. The results indicate that (i) the order of fitting performance is: SVR > ANN > MLR; and (ii) the order of generalization performance is: SVR > MLR > ANN. Therefore, SVR model has the best predicting performance in comparison with the other models.

Fig. 4 presents description statistics results for the time series of G_1 to G_3 from 1980 to 2008. It is indicated that the data points concentrate on the vicinity of mean, and most of them uniformly distribute in the interval of [mean- variance, mean+ variance]. Through the normality test for the three random variables, G_1 to G_3, the results indicated that the average KS-statistic value is 0.19, with the maximum and minimum values being 0.21 and 0.16, respectively. Given a significance level of 0.05, all of the statistic values are lower than the critical one (0.22). Therefore, it is judged the normality hypothesis can hardly be rejected. By using maximum likelihood estimation method, the distribution of G_1 to G_3 can be estimated as follows: $G_1 \sim N\,(0.024, 0.03^2)$, $G_2 \sim N\,(0.168, 0.061^2)$ and $G_3 \sim N\,(0.153, 0.96^2)$.

Fig. 5 presents the statistic histograms of the simulated waste generation rate in the planning horizon. Since these figures have approximately shapes of normal, gamma, poisson, and lognormal distributions, KS-test is used to find the best distribution. Fig. 6 shows the calculated SK-statistic values. As shown in Fig. 6(a), the maximum and minimum SK-statistic values are 0.099 and 0.045. Given a significance level of 0.05, all of the statistic values are higher than the critical one (0.0431); while given a significance level of 0.01, statistic value (i.e. waste generation in period 1) is lower than the critical one (0.0516). In Fig. 6(b), statistic value (i.e. waste generation in period 3) cannot pass the KS-test under significance level of 0.05. In Fig. 6(c), all of the statistic values cannot pass the KS-test. In Fig. 6(d), all of the statistic values pass the KS-test.

Therefore, it is judged the lognormal distribution hypothesis can hardly be rejected since all KS-statistic values are lower than or very near the critical ones given a significance level of 0.05. Assume that three levels (i.e., low, medium and high) of waste generation have the probability of 0.2, 0.6 and 0.2, respectively, we can obtain the interval waste generation rate based on quantile calculation principle described in the Step 7 of Methodology section. Table 4 presents the probability levels correspond to different interval values of the waste generation rate ($\omega_{kh}^{\pm}$), which are the input data of optimization framework. Table 5 shows the solution of optimized waste-flow allocation for the composting facilities. It is indicated that an excess flow can be generated if the allowable-waste-flow level is exceeded (i.e. excess flow=generated flow−assigned quota). The waste-flow patterns (including allowable and excess flows) would vary dynamically due to temporal and spatial variations of waste-generation/ management conditions. Fig. 7 presents waste flows allocated to the landfills, composting facilities, incinerators and comprehensive treatment facilities during the entire planning horizon. Waste flows allocated to the composting facility, incinerator and comprehensive treatment facility would keep increasing with time. This is because the division rate is increased rapidly and the waste flow is transferred to those MSW management facilities under compulsion.

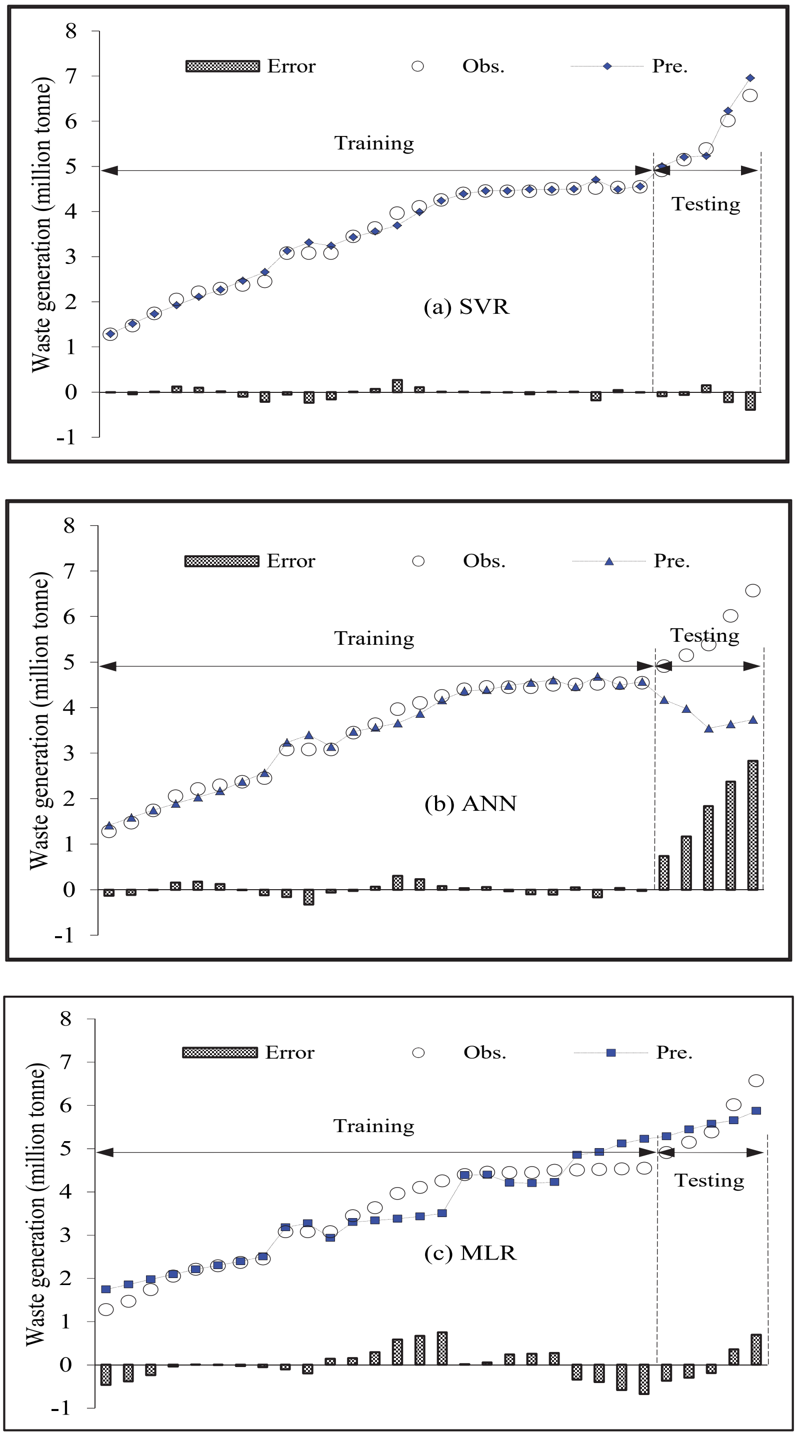


Fig. 3. Comparison of the observed and predicted value with (a) SVR model, (b) ANN model, (c) MLR (note: MLR denotes multiple linear regression; Obs. denotes the observed value; Pre. denotes the predicted value)

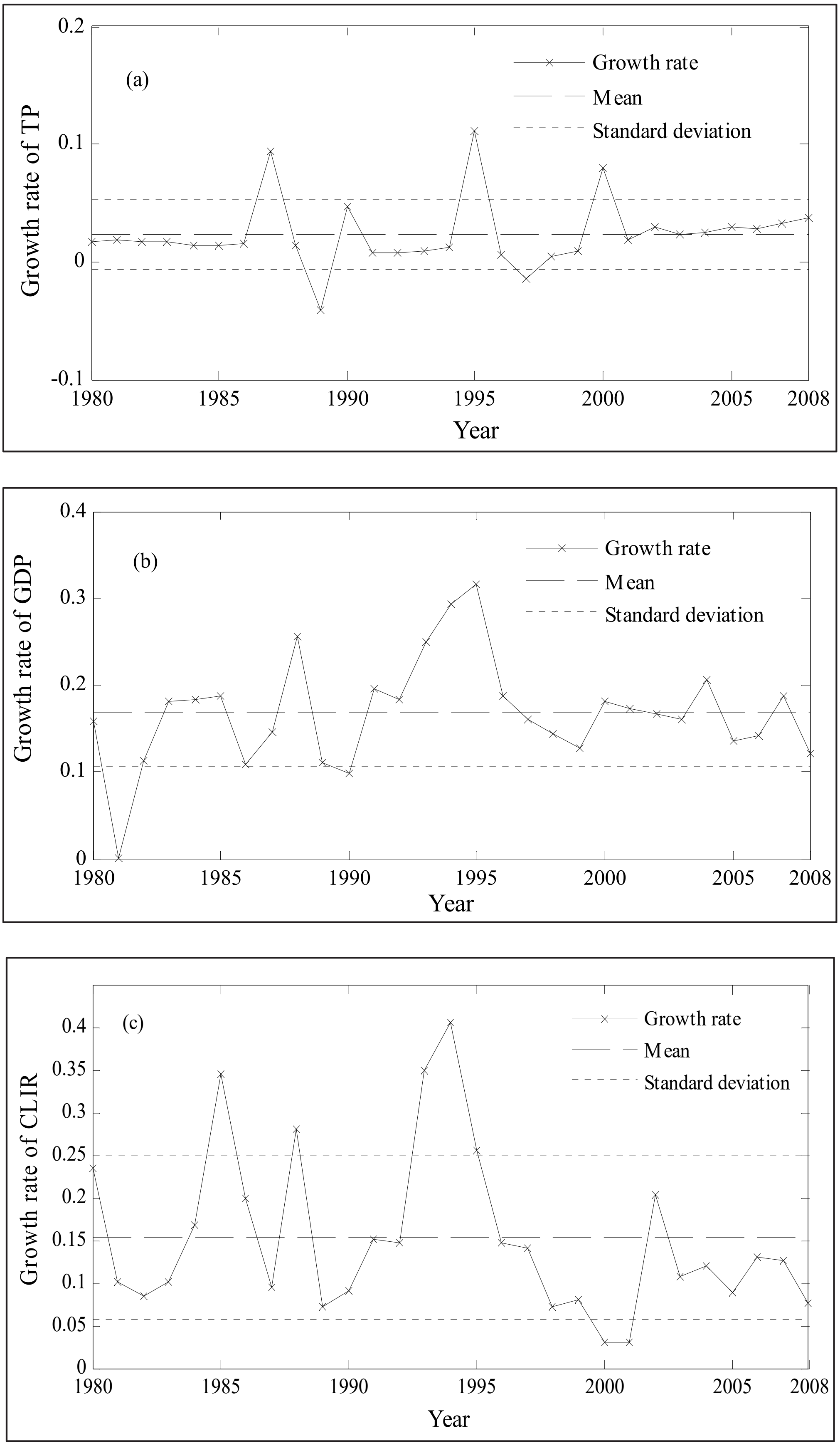


Fig. 4. Description statistics results for the time series of (a) G_1, (b) G_2 and (c) G_3

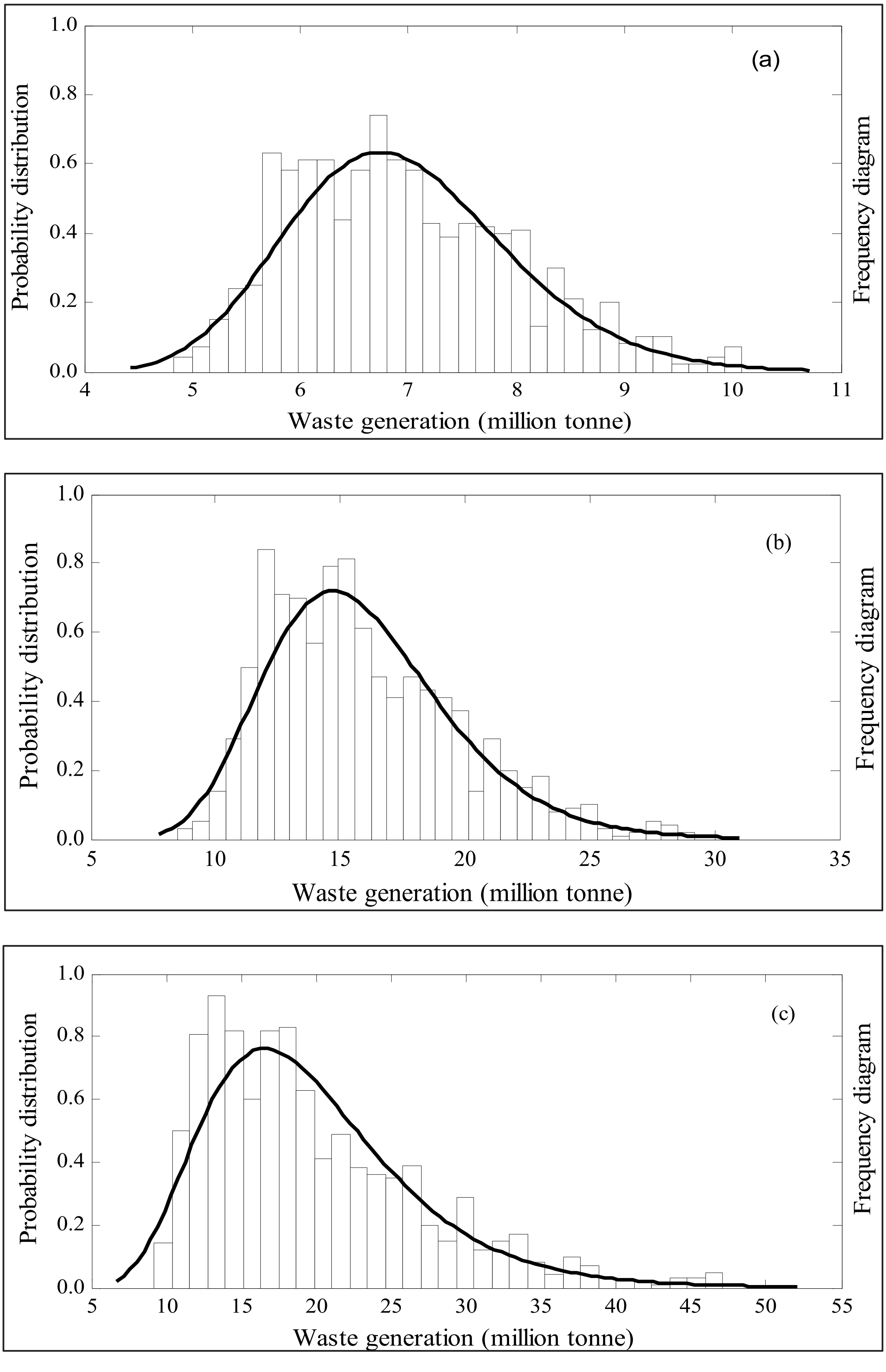


Fig. 5. Statistic histograms for the simulated waste generation rate in (a) period 1, (b) period 2 and (c) period 3

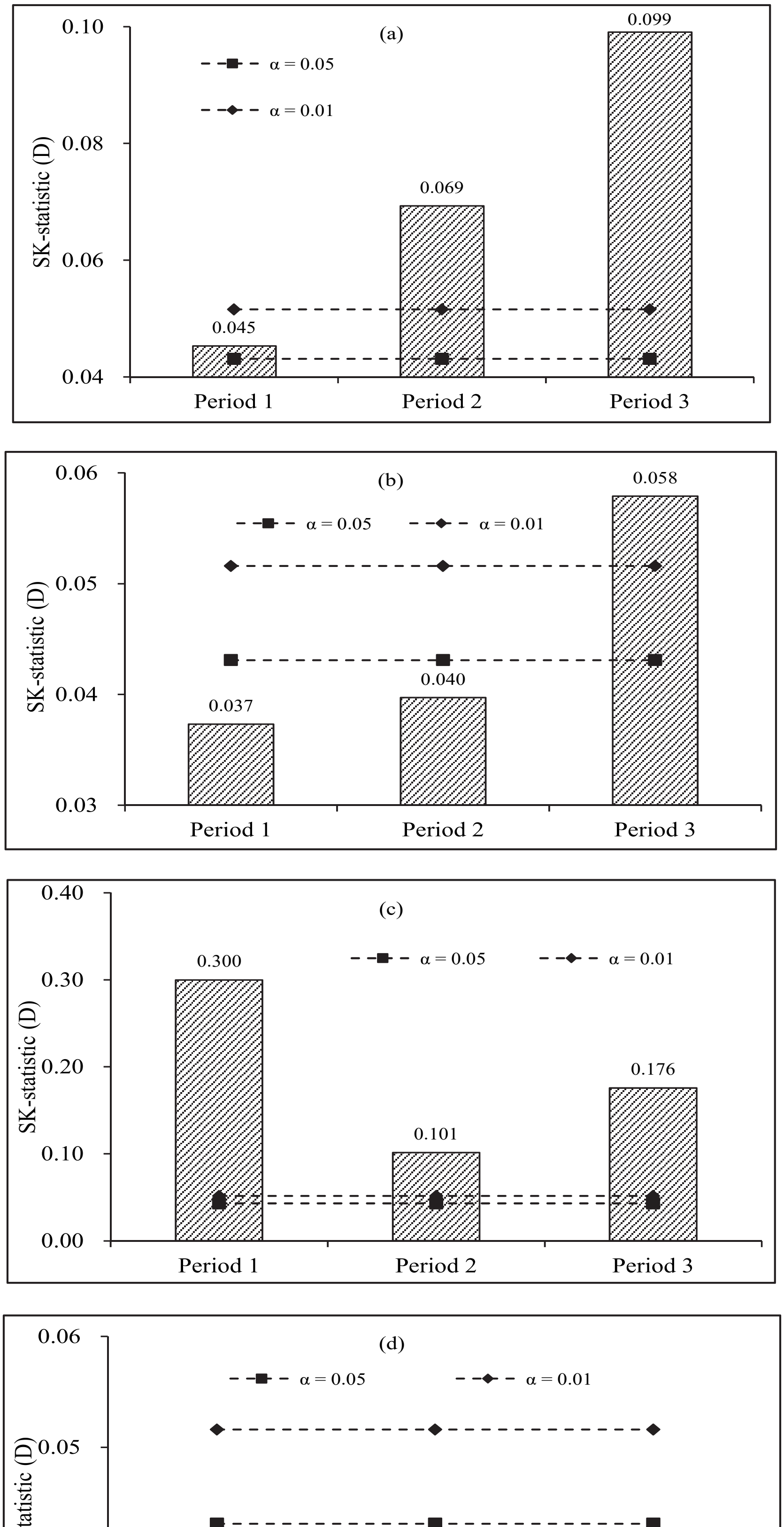


Fig. 6. SK- test results for different alternative distributions. (a) Normal distribution; (b) Gamma distribution; (c) Poisson distribution; (d) Lognormal distribution

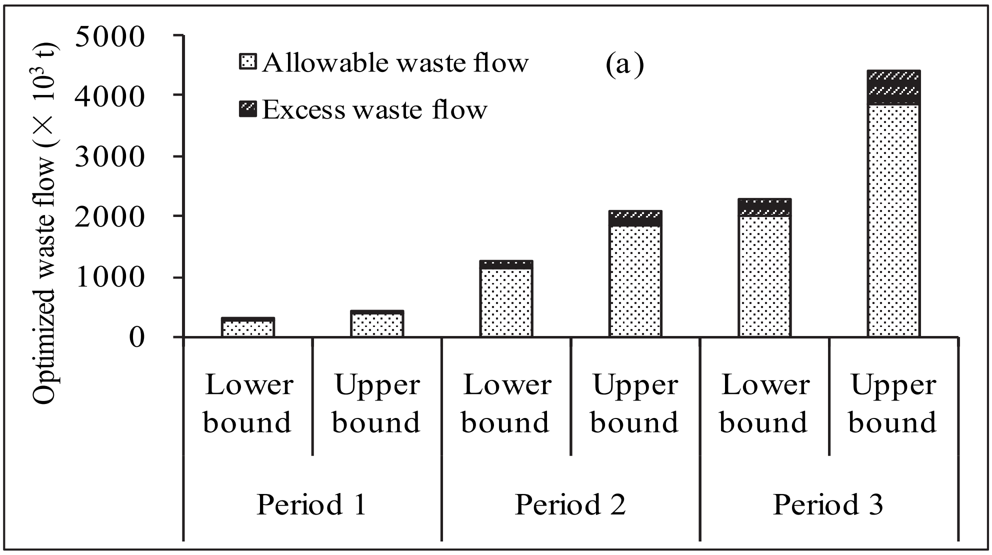
Table 4. Simulation results of waste-generation rates and the associated probabilities

Level of waste generation	*Probability*	*Waste generation rate, $w_{kh}^{\pm}$ (10^6 tonnes)*		
		k = 1	*k = 2*	*k = 3*
h = 1 (low)	0.2	(4.78, 6.11)	(7.95, 12.87)	(9.15, 14.13)
h = 2 (medium)	0.6	(6.11, 7.89)	(12.87, 19.44)	(14.13, 25.74)
h = 3 (high)	0.2	(7.89, 10.11)	(19.44, 29.30)	(25.74, 47.31)

Fig. 8 presents the optimal solution for facility expansions. For example, the Dongcun composting facility should be expanded by an increment of (0, 286) tonne/day at the starts of periods 2 and 3, respectively, with 70% of the capacity (i.e., (0, 200) tonne/day) being dedicated to the residential waste. Thus, the capacity of Liulitun, Nangong, Dongcun and Gaoantun composting facilities for residential wastes would eventually be increased to (3200, 4200), (1000, 1600), (1250, 1650) and (1200, 4200) tonne/day, respectively. However, the expansion capacity of Dongcun composting facility is less than other three composting facilities. The capacity of Asuwei, Liulitun, Nangong and Gaoantun incinerators for residential wastes would eventually be increased to (1000, 2000), (2000, 3000), 1000 and (1600, 3400) tonne/day, respectively.

The expansion capacity of incinerators is less than that of composting facilities. This is because the incinerators have higher disposal costs and allotment of waste flows to the composting facilities would be more economical. The capacity of Fengtai comprehensive treatment facility for residential wastes would remain the status quo (i.e. 1600 tonne/day), while the capacity of Asuwei and Xitianyang comprehensive treatment facilities would eventually be increased to (2800, 5200) and (1800, 3800) tonne/day, respectively.

The optimal value of objective function is \$(1.33, 2.92) × 10^9, which denotes the system cost. The system cost includes expenses for handling fixed allowable waste flows, probabilistic excess flows, and expansions/developments of composting facilities, incinerators, comprehensive treatment facilities and landfills. The results indicate that the cost for facility expansions is \$(207.2, 824.2) million (or (15.6, 28.2)% of the total system cost); the regular cost for disposing/diverting allowable waste flows is \$(918.7, 1715.8) million (or (58.8, 69.2)% of the total system cost); the penalty cost for handling excess waste flows is \$(202.3, 379.3) million (or (13.0, 15.2)% of the total system cost). In addition, the cost for waste landfilling is \$(495.1, 896.6) million (or (30.7, 37.3)% of the total system cost); the costs for composting facilities, incinerators and comprehensive treatment facilities are \$(219.5, 465.5) million (or (15.9, 16.5)% of the total system cost), \$(265.9, 747.1) million (or (20.0, 25.6)% of the total system cost) and \$(347.7, 810.1) million (or (26.2, 27.7)% of the total system cost), respectively.



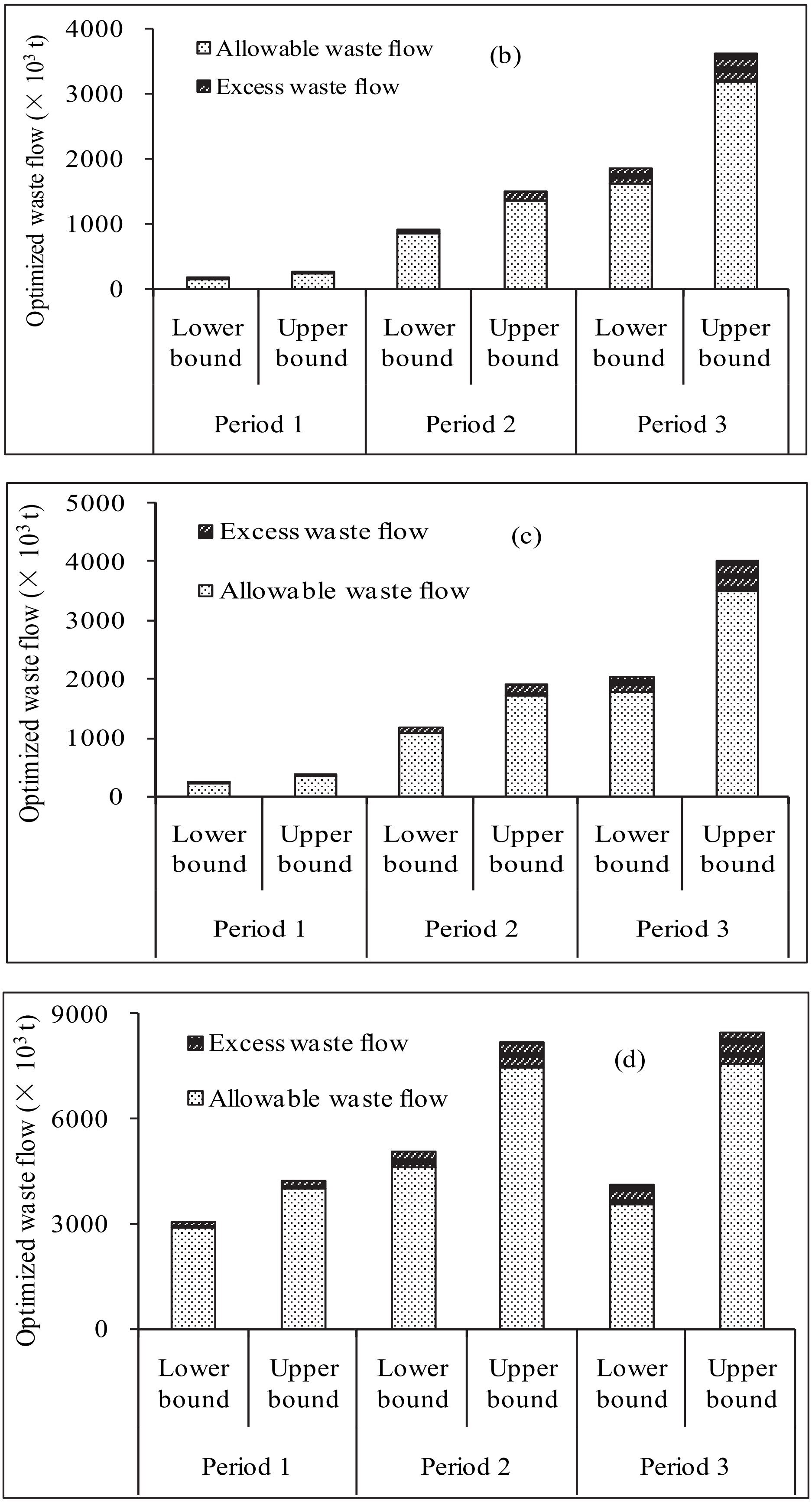


Fig. 7. Optimal waste-flow allocation pattern for (a) composting facility, (b) incinerator, (c) comprehensive treatment facility and (d) landfill

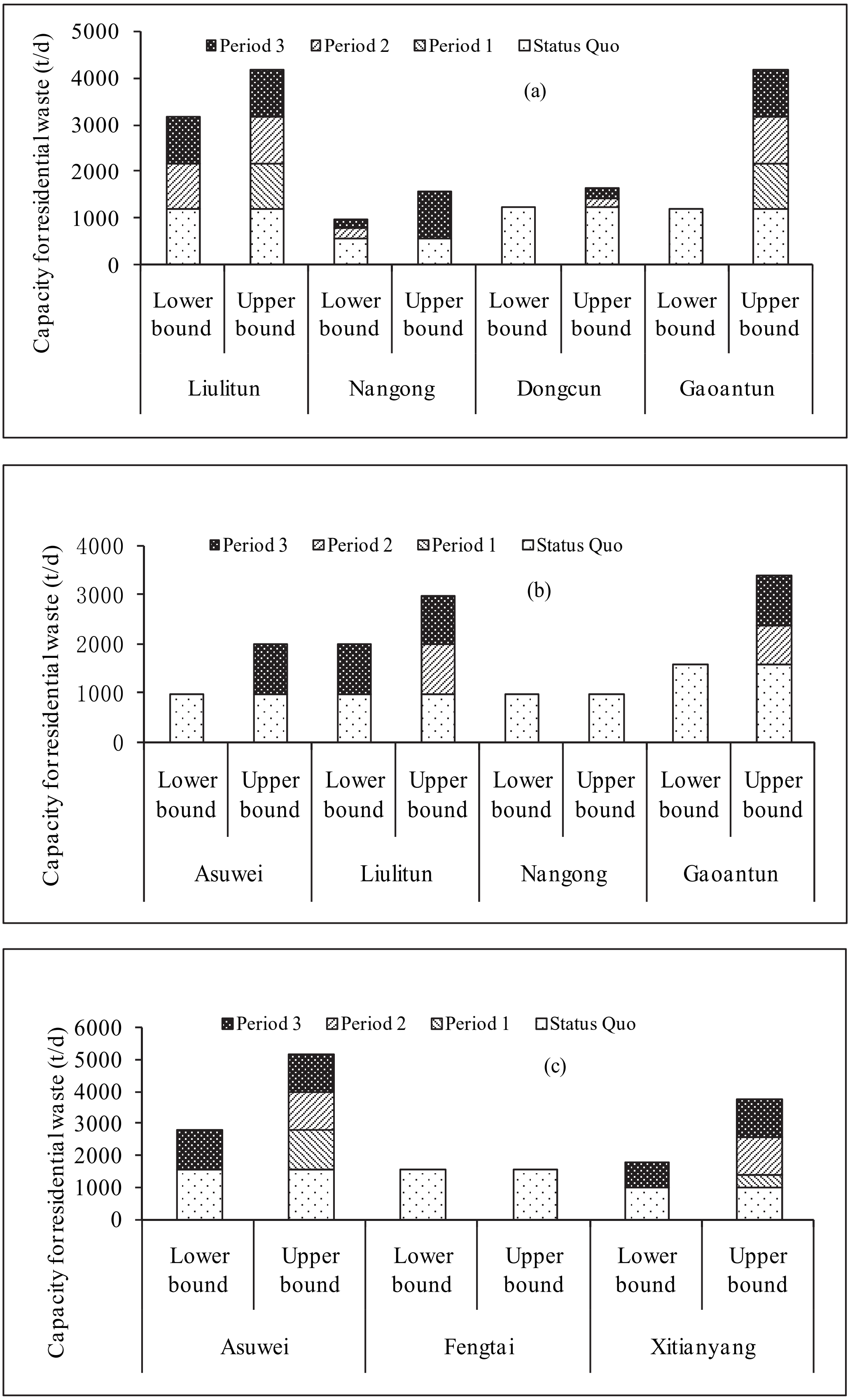


Fig. 8. Capacity planning for (a) composting facility, (b) incinerator and (c) comprehensive treatment facility

Table 5. Results of waste flow allocation for composting facility

Period	*Facility*	*Waste generation rate*	*Allowable waste flow (t/d)*	*Excess waste flow (t/d)*	*Optimized waste flow (t/d)*
1	Liulitun composting	Low	[361.59, 463.33]	0	[361.59, 463.33]
1	Nangong composting	Low	[100.03, 110.16]	0	[100.03, 110.16]
1	Dongcun composting	Low	[91.34, 100.58]	0	[91.34, 100.58]
1	Gaoantun composting	Low	[250.55, 432.69]	0	[250.55, 432.69]
1	Liulitun composting	Medium	[361.59, 463.33]	0	[361.59, 463.33]
1	Nangong composting	Medium	[100.03, 110.16]	0	[100.03, 110.16]
1	Dongcun composting	Medium	[91.34, 100.58]	0	[91.34, 100.58]
1	Gaoantun composting	Medium	[250.55, 432.69]	0	[250.55, 432.69]
1	Liulitun composting	High	[361.59, 463.33]	[234.08, 332.67]	[595.67, 796]
1	Nangong composting	High	[100.03, 110.16]	[0, 0]	[100.03, 110.16]
1	Dongcun composting	High	[91.34, 100.58]	[0, 0]	[91.34, 100.58]
1	Gaoantun composting	High	[250.55, 432.69]	[0, 68.34]	[250.55, 501.03]
2	Liulitun composting	Low	[835.23, 1258.87]	0	[835.23, 1258.87]
2	Nangong composting	Low	[231.07, 254.45]	0	[231.07, 254.45]
2	Dongcun composting	Low	[210.97, 232.32]	0	[210.97, 232.32]
2	Gaoantun composting	Low	[309.44, 810.85]	0	[309.44, 810.85]
2	Liulitun composting	Medium	[835.23, 1258.87]	0	[835.23, 1258.87]
2	Nangong composting	Medium	[231.07, 254.45]	0	[231.07, 254.45]
2	Dongcun composting	Medium	[210.97, 232.32]	0	[210.97, 232.32]
2	Gaoantun composting	Medium	[309.44, 810.85]	0	[309.44, 810.85]
2	Liulitun composting	High	[835.23, 1258.87]	[599.70, 903.87]	[1434.93, 2162.74]
2	Nangong composting	High	[231.07, 254.45]	[0, 0]	[231.07, 254.45]
2	Dongcun composting	High	[210.97, 232.32]	[0, 0]	[210.97, 232.32]
2	Gaoantun composting	High	[309.44, 810.85]	[210.30, 582.19]	[519.74, 1393.04]
3	Liulitun composting	Low	[1338.24, 1911.01]	0	[1338.24, 1911.01]
3	Nangong composting	Low	[448.73, 693.71]	0	[448.73, 693.71]
3	Dongcun composting	Low	[461.99, 771.71]	0	[461.99, 771.71]
3	Gaoantun composting	Low	[497.42, 1934.98]	0	[497.42, 1934.98]
3	Liulitun composting	Medium	[1338.24, 1911.01]	0	[1338.24, 1911.01]
3	Nangong composting	Medium	[448.73, 693.71]	0	[448.73, 693.71]
3	Dongcun composting	Medium	[461.99, 771.71]	0	[461.99, 771.71]
3	Gaoantun composting	Medium	[497.42, 1934.98]	0	[497.42, 1934.98]
3	Liulitun composting	High	[1338.24, 1911.01]	[960.85, 1372.11]	[2299.09, 3283.12]
3	Nangong composting	High	[448.73, 693.71]	[258.28, 498.08]	[707.01, 1191.79]
3	Dongcun composting	High	[461.99, 771.71]	[331.71, 554.09]	[793.7, 1325.8]
3	Gaoantun composting	High	[497.42, 1934.98]	[357.14, 1389.32]	[854.56, 3324.3]

5. Conclusions

The hypothesis test results indicate that the assumptions for simulated waste generation can be accepted statistically. However, this does not conclude that they are also acceptable under all the conditions. Therefore, the assumptions should be re-tested based on simulation results when the model is extended to other cities. If the normality test fails, other distributions may be assumed and tested for the waste generation.

The developed SVRMC-ITSP method is formulated by incorporating the SVR and Monte Carlo simulation within a TISP framework. Three special characteristics make it unique compared with other optimization techniques: (i) it can simulate the PDFs of waste generation rate over the planning horizon; (ii) it can provide a linkage to pre-regulated policies (determined by authorities) that have to be respected when a modeling effort is undertaken; (iii) it is useful for tackling uncertainties presented as both probabilities and intervals. The developed SVRMC-ITSP method represents a new effort to enhance the analysis accuracy in planning the MSW management system in the urban districts of Beijing, China. The Monte Carlo simulation results indicate that the random variable of waste generation rate would follow the lognormal distribution. The results of optimization method identify desired capacity expansion and waste flow-allocation plans.

The main limitations of this research are as follows: (i) for the SVR model, selecting the kernel function is usually based on application-domain knowledge, which may add the disturbance to the prediction accuracy of waste generation rate; and (ii) except Monte Carlo simulation, fuzzy sets theory is an alternative tool to handle uncertainty, thus future studies on integrating the fuzzy two-stage stochastic programming into the SVRMC-ITSP framework are desired.

Acknowledgments

This research was supported by the Major State Basic Research Development Program of MOST (2005CB724200). Also, the authors are grateful to the editors and the anonymous reviewers for their insightful comments and suggestions.

Appendix

List of indexes, decision variables and parameters in Eqs. (31-55) is shown as follows:

Indexes

k - planning period, $k = 1, 2, 3$;
c - name of composting facility, $c = 1, 2, 3, 4$;
l - name of landfill, $l = 1, 2, \ldots 7$;
i - name of incinerator, $i = 1, 2, 3, 4$;
z - name of comprehensive treatment facility, $z = 1, 2, 3$;
e - expansion option for composting facility, $e = 1, 2, 3$;
f - expansion option for incinerator, $f = 1, 2, 3$;
g - expansion option for comprehensive treatment facility, $g = 1, 2, 3$;
h - level of waste generation rate, $h = 1, 2, 3$;

Decision variables

$XUC_{ckh}^{\pm}$ - amount by which the target allowable flow level (i.e. $TUC_{ck}^{\pm}$) is exceeded when the waste-generation rate is $w_{kh}^{\pm}$ with probability P_h (tonne/day);
$XUI_{ikh}^{\pm}$ - amount by which the target allowable flow level (i.e. $TUI_{ik}^{\pm}$) is exceeded when the waste-generation rate is $w_{kh}^{\pm}$ with probability P_h (tonne/day);
$XUZ_{zkh}^{\pm}$ - amount by which the target allowable flow level (i.e. $TUZ_{zk}^{\pm}$) is exceeded when the waste-generation rate is $w_{kh}^{\pm}$ with probability P_h (tonne/day);
$XUL_{lkh}^{\pm}$ - amount by which the target allowable flow level (i.e. $TUL_{lk}^{\pm}$) is exceeded when the waste-generation rate is $w_{kh}^{\pm}$ with probability P_h (tonne/day);
$YCL_{clk}^{\pm}$ - residue for excess waste flow generated by composting facility c during period k (tonne/day);
$YIL_{ilk}^{\pm}$ - residue for excess waste flow generated by incinerator i during period k (tonne/day);
$YCL_{clk}^{\pm}$ - residue for excess waste flow generated by comprehensive treatment facility z during period k (tonne/day);
$BCE_{cek}^{\pm}$ - binary decision variable for composting facility c with expansion option e;
$BIF_{ifk}^{\pm}$ - binary decision variable for incinerator i with expansion option f;
$BZG_{zgk}^{\pm}$ - binary decision variable for comprehensive treatment facility z with expansion option g;

Parameters

$f^{\pm}$ - system expected cost (\$);
L_k - length of period k (day);
$CC_c^{\pm}$ - initial capacity of composting facility c (tonne/day);
$LC_l^{\pm}$ - capacity of landfill l at the beginning of planning period (tonne);
$IC_i^{\pm}$ - initial capacity of incinerator i (tonne/day);
$ZC_z^{\pm}$ - initial capacity of comprehensive treatment facility z (tonne/day);
$DUC_c^{\pm}$ - distances from study area to composting facility c (km);
$DUI_i^{\pm}$ - distances from study area to incinerator i (km);
$DUZ_z^{\pm}$ - distances from study area to comprehensive treatment facility z (km);
$DUL_l^{\pm}$ - distances from study area to landfill l (km);
$DCL_{cl}^{\pm}$ - distances from composting facility c to landfill l (km);
$DIL_{il}^{\pm}$ - distances from incinerator i to landfill l (km);
$DZL_{zl}^{\pm}$ - distances from comprehensive treatment facility z to landfill l (km);
$TR_k^{\pm}$ - unit traffic expense for allowable waste [\$/(km·tonne)];
$ETR_k^{\pm}$ - unit traffic expense for excess waste [\$/(km·tonne)], $ETR_k^{\pm} \geq TR_k^{\pm}$;
$COC_k^{\pm}$ - collection cost for allowable waste (\$/tonne);
$ECOC_k^{\pm}$ - collection cost for excess waste (\$/tonne), $ECOC_k^{\pm} \geq COC_k^{\pm}$;
$OC_{ck}^{\pm}$ - operating costs of composting facility c (\$/tonne);
$OI_{ik}^{\pm}$ - operating costs of incinerator i (\$/tonne);
$OZ_{zk}^{\pm}$ - operating costs of comprehensive treatment facility z (\$/tonne);
$OL_{lk}^{\pm}$ - operating costs of landfill l (\$/tonne);
$EOC_{ck}^{\pm}$ - operating costs of composting facility c for excess waste (\$/tonne), $EOC_{ck}^{\pm} \geq OC_{ck}^{\pm}$;
$EOI_{ik}^{\pm}$ - operating costs of incinerator i for excess waste (\$/tonne), $EOI_{ik}^{\pm} \geq OI_{ik}^{\pm}$;
$EOZ_{zk}^{\pm}$ - operating costs of comprehensive treatment facility z for excess waste (\$/tonne), $EOZ_{zk}^{\pm} \geq OZ_{zk}^{\pm}$;
$EOL_{lk}^{\pm}$ - operating costs of landfill l during period k for excess waste (\$/tonne), $EOL_{lk}^{\pm} \geq OL_{lk}^{\pm}$;
$ECE_{cek}^{\pm}$ - capital cost of expanding composting facility c by option e (\$/tonne);

$EIF_{ifk}^{\pm}$ - capital cost of expanding incinerator *i* by option *f* in period *k* ($/tonne);
$EZG_{zgk}^{\pm}$ - capital cost of expanding comprehensive treatment facility *z* by option *g* ($/tonne);
$RC_{ck}^{\pm}$ - revenue of composting facility *c* ($/tonne);
$RI_{ik}^{\pm}$ - revenue of incinerator facility *i* ($/tonne);
$RZ_{zk}^{\pm}$ - revenue of comprehensive treatment facility *z* ($/tonne);
$ERC_{ck}^{\pm}$ - revenue of composting facility *c* because of excess waste ($/tonne);
$ERI_{ik}^{\pm}$ - revenue of incinerator facility *i* because of excess waste ($/tonne);
$ERZ_{zk}^{\pm}$ - revenue of comprehensive treatment facility *z* because of excess waste ($/tonne);
$FC_{ck}^{\pm}$ - residue MSW rates (% of incoming mass to composting facility c);
$FI_{ik}^{\pm}$ - residue MSW rates (% of incoming mass to incinerator i);
$FZ_{zk}^{\pm}$ - residue MSW rates (% of incoming mass to comprehensive treatment facility z);
$GC_{k}^{\pm}$ - diversion rate of waste flow to composting facility *c*;
$GI_{k}^{\pm}$ - diversion rate of waste flow to incinerator *i*;
$GZ_{k}^{\pm}$ - diversion rate of waste flow to comprehensive treatment facility *z*;
$TUC_{ck}^{\pm}$ - allowable waste flow from urban districts to composting facility *c* (tonne/day);
$TUI_{ik}^{\pm}$ - allowable waste flow from urban districts to incinerator *i* (tonne/day);
$TUZ_{zk}^{\pm}$ - allowable waste flow to comprehensive treatment facility *z* (tonne/day);
$TUL_{lk}^{\pm}$ - allowable waste flow from urban districts to landfill *l* during period *k* (tonne/day);
$TCL_{clk}^{\pm}$ - residue for allowable waste flow generated by composting facility *c* to landfill *l* (tonne/day);
$TIL_{ilk}^{\pm}$ - residue for allowable waste flow generated by incinerator *i* to landfill *l* (tonne/day);
$TZL_{zlk}^{\pm}$ - residue for allowable waste flow generated by comprehensive treatment facility *z* to landfill *l* (tonne/day);
P_h - probability of $w_{kh}^{\pm}$ with waste-generation level *h*;
$w_{kh}^{\pm}$ - residential MSW generated by urban districts with probability P_h in period *k*;
$\kappa_i^{\pm}$ - capacity rate of composting facility *c* for urban districts (% of the whole capacity);
$\zeta_c^{\pm}$ - capacity rate of incinerator *i* for urban districts (% of the whole capacity);
$\psi_z^{\pm}$ - capacity rate of comprehensive treatment facility *z* for urban districts (% of the whole capacity);
$\delta_l^{\pm}$ - capacity rate of landfill *l* for urban districts (% of the whole capacity);
ΔCCE_{cek} - amount of capacity expansion option *e* for composting facility *c* (tonne/day);
ΔCIF_{ifk} - amount of capacity expansion option *f* for incinerator *i* (tonne/day);
ΔCZG_{zgk} - amount of capacity expansion option *g* for comprehensive treatment facility *z* (tonne/day).

References

Abbasi M., Abduli M., Omidvar B., Baghvand A., (2012), Forecasting municipal solid waste generation by hybrid support vector machine and partial least square model, *International Journal of Environmental Research*, **7**, 27-38.

Ali Abdoli M., Falah Nezhad M., Salehi Sede R., Behboudian S., (2012), Longterm forecasting of solid waste generation by the artificial neural networks, *Environmental Progress & Sustainable Energy*, **31**, 628-636.

Bouhouche S., Mentouri Z., Meradi H., (2012), *Combined Use of Support Vector Regression and Monte Carlo Simulation in Quality and Process Control Calibration*, Proc.of the 2012 Int. Conf. on Industrial Engineering and Operation Management, Istanbul, Turkey, July 3-6.

Chen K.Y., Wang C.H., (2007), Support vector regression with genetic algorithms in forecasting tourism demand, *Tourism Management*, **28**, 215-226.

Chen H.B., Zhang L., Zhou J.C., Hu Y., (2011), Development of municipal solid waste management practices in different geographical units and urban systems in China, *Environmental Engineering and Management Journal,* **10**, 945-949.

Chen H.B., Zhou J.C., Zhang H.L., Yang Y., (2017), Optimizing household waste collection through AHP-MEA model: Case study of Kunming, China, *Environmental Engineering and Management Journal,* **16**, 2887-2899.

Cheng H., Zhang Y., Meng A., Li Q., (2007), Municipal solid waste fueled power generation in China: A case study of waste-to-energy in Changchun city, *Environmental Science and Technology-Columbus*, **41**, 7509-7515.

Dai C., Li Y.P., Huang G.H., (2011), A two-stage support-vector-regression optimization model for municipal solid waste management- a case study of Beijing, China, *Journal of Environmental Management*, **92**, 3023-3037.

Dyson B., Chang N.B., (2005), Forecasting municipal solid waste generation in a fast-growing urban region with system dynamics modeling, *Waste Management*, **25**, 669-679.

Ghinea C., Drăgoi E.N., Comăniţă E.D., Gavrilescu M., Câmpean T., Curteanu S., Gavrilescu M., (2016), Forecasting municipal solid waste generation using prognostic tools and regression analysis, *Journal of Environmental Management*, **182**, 80-93.

Gunalay Y., Yeomans J.S., Huang G.H., (2012), Modelling to generate alternative policies in highly uncertain environments: an application to municipal solid waste management planning, *Journal of Environmental Informatics*, **19**, 58-69.

Huang G., Brian W.B., Gilles G.P., (1992), A grey linear programming approach for municipal solid waste management planning under uncertainty, *Civil Engineering Systems*, **9**, 319-335.

Kumar J.S., Subbaiah K.V., Rao P., (2011), Prediction of municipal solid waste with RBF net work-a case study of Eluru, AP, India, *International Journal of Innovation, Management and Technology*, **2**, 238-243.

Li Y., Huang G., Yang Z., Nie S., (2008), An integrated two-stage optimization model for the development of long-term waste-management strategies, *Science of the Total Environment*, **392**, 175-186.

Maqsood I., Huang G.H., (2003), A two-stage interval-stochastic programming model for waste management under uncertainty, *Journal of the Air & Waste Management Association*, **53**, 540-552.

Massey F.J., (1951), The Kolmogorov-Smirnov test for goodness of fit, *Journal of the American statistical Association*, **46**, 68-78.

Noori R., Abdoli M., Ghazizade M.J., Samieifard R., (2009), Comparison of neural network and principal component-regression analysis to predict the solid waste generation in Tehran, *Iranian Journal of Public Health*, **38**, 74-84.

Qin X., Huang G., (2009), Characterizing uncertainties associated with contaminant transport modeling through a coupled fuzzy-stochastic approach, *Water, Air, & Soil Pollution*, **197**, 331-348.

Smola A.J., Schölkopf B., (2004), A tutorial on support vector regression, *Statistics and Computing*, **14**, 199-222.

Sun Y., Huang G. H., Li Y. P., (2010), ICQSWM: An inexact chance-constrained quadratic solid waste management model, *Resources Conservation and Recycling*, **54**, 641-657.

Vapnik V.N., (1999), An overview of statistical learning theory, *IEEE Transactions on Neural Networks*, **10**, 988-999.

Wieland R., Mirschel W., Zbell B., Groth K., Pechenick A., Fukuda K., (2010), A new library to combine artificial neural networks and support vector machines with statistics and a database engine for application in environmental modeling, *Environmental Modelling & Software*, **25**, 412-420.

Xi B., Su J., Huang G., Qin X., Jiang Y., Huo S., Ji D., Yao B., (2010), An integrated optimization approach and multi-criteria decision analysis for supporting the waste-management system of the City of Beijing, China, *Engineering Applications of Artificial Intelligence*, **23**, 620-631.

Yang Z., Li S., Zhang Y., Huang G., (2011), Emergy synthesis for three main industries in Wuyishan City, China, *Journal of Environmental Informatics*, **17**, 23-35.