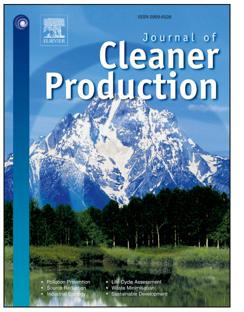
Journal Pre-proof



Efficiency assessment of rural domestic sewage treatment facilities by a slacked-based DEA model

Panpan Cheng, Qiang Jin, Hui Jiang, Ming Hua, Zhen Ye

PII: S0959-6526(20)32158-2

DOI: https://doi.org/10.1016/j.jclepro.2020.122111

Reference: JCLP 122111

To appear in: *Journal of Cleaner Production*

Received Date: 5 March 2020

Revised Date: 30 April 2020

Accepted Date: 6 May 2020

Please cite this article as: Cheng P, Jin Q, Jiang H, Hua M, Ye Z, Efficiency assessment of rural domestic sewage treatment facilities by a slacked-based DEA model, *Journal of Cleaner Production* (2020), doi: https://doi.org/10.1016/j.jclepro.2020.122111.

Credit Author Statement

Panpan Cheng: Writing- Original draft preparation, Methodology, Data curation, Validation, Formal analysis, Investigation

Qiang Jin: Conceptualization, Methodology, Data curation, Writing- Reviewing and Editing, Supervision

Hui Jiang: Investigation, Writing- Reviewing and Editing

Ming Hua: Resources, Project administration

Zhen Ye: Writing- Reviewing and Editing

Efficiency assessment of rural domestic sewage treatment facilities by a slacked-based DEA model

Panpan Cheng[a], Qiang Jin[a,b]* Hui Jiang[a], Ming Hua[c], Zhen Ye[d]

a Shanghai Engineering Research Center of Solid Waste Treatment and Resource Recovery, School of Environmental Science and Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China

b China Institute for Urban Governance, Shanghai Jiao Tong University, Shanghai, 200240, China

c State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Nanjing, 210023, China

d The Bartlett School of Construction and Project Management, University College London, London, WC1E 6BT, UK

***Corresponding author:**

Qiang Jin, Ph. D.

Phone: +86-18621863638

Fax: +86-21-54740825

E-mail: jinqiang@sjtu.edu.cn

Word Count:6243

Abstract

In the context of sustainable development, a number of rural domestic sewage treatment facilities had been built in China to solve the problem of rural domestic sewage pollution. The comprehensive, quantitative and objective efficiency assessment of facilities is urgent. This study used a non-radial slacked-based data envelopment analysis model combined with cluster analysis to construct an index system covering multiple aspects, including three inputs and four outputs to assess 681 facilities. These samples selected from the biggest demonstration area are the most representative for and exceed 2/5 of the running facilities all over the country. The average efficiency score of samples was 0.496 meaning the improvement potential was about 50.4%. Only 27 samples were relatively effective, scoring 1. The remaining 654 facilities had different levels of input excesses or output shortfalls, which should be the key objects to improve overall performance. In addition, there was evidence that output indicators had more room for improvement than input indicators. The analysis of sensitivity on inputs and outputs confirmed that the idleness and poor treatment effects of rural sewage treatment facilities should be concerned. Finally, Kruskal–Wallis non-parametric test verified that technology and load rate of facilities have significant impacts on efficiency. The performance evaluation results could not only provide guidance for the local government to strengthen the supervision and operation of facilities, but also potentially provide reference for the construction, operation and management of rural sewage treatment facilities in China.

Keywords: Data envelopment analysis, Efficiency assessment, Rural domestic sewage, Potential improvement, Sensitivity analysis, Explanatory factors

1 Introduction

In recent years, the water pollution has become a serious challenge to the development of rural areas. By 2015, the direct emission of rural domestic sewage was about 20 million tons every day. The annual chemical oxygen demand (COD) emission was about 10.69 million tons and the annual ammonia nitrogen emission was 0.73 million tons (China Environmental Statistics Annual Report, 2015). Due to economic and geographical factors, the coverage of treatment facilities is extremely low in most rural areas of China. 96 % of rural villages cannot effectively treat sewage (Gu et al., 2016). To control water pollution in rural areas, the central government had proposed an ambitious plan, that the treatment coverage in rural area will reach 33.6% by 2020. A few rural domestic sewage treatment projects have been set up and demonstrated in key river valleys (Chen et al., 2018; Wu et al., 2011). Although certain progress has been made, the existing rural sewage treatment facilities have problems such as scattered locations, jagged technical levels and weak supervision. Thus, it is urgent to evaluate performance of existing facilities and answer which is the best.

The environmental performance evaluation proved to be an effective and suitable environmental management tool to find out the problems existing in rural sewage treatment facilities (Alemany et al., 2005; Benedetti et al., 2008; Gallego et al., 2008). It can help the local governments and sewage companies formulate reasonable policies to promote the effective development of rural sewage treatment facilities, and also to provide targeted improvement recommendations. Kalbar et al. (2012) assessed the applicability of 4 common rural sewage treatment technologies in India based on scenario analysis. Xia et al. (2012) evaluated treatment technologies from the economic and technical aspects by the fuzzy advantages and disadvantages coefficient method in a village of Changzhou. Shen et al. (2014) combined the analytic hierarchy process with the entropy method to select 10 advanced technologies from 15 commonly used rural domestic sewage treatment technologies. The existing research mainly focused on the simple evaluation of the treatment technology. Besides, artificially assigning weights to indicators led to subjective errors. More importantly, these methods failed to distinguish inefficient from efficient facilities and quantify the improvement potential.

Data envelopment assessment (DEA) has been widely used in the performance evaluation of water sector in recent years (Dong et al., 2017; Hu et al., 2019; Jiang et al., 2020). This method obtains relative efficiency of decision-making units (DMUs) with multiple inputs and multiple outputs based on linear programming (Mostashari-Rad et al., 2019). A significant advantage of the DEA method is that it is not necessary to assume a correlation between input and output indicators (Hosseinzadeh-Bandbafha et al., 2018). Thus, the evaluation results are objective. Traditional DEA models are radial, which fail to calculate the theoretical target values of inputs and outputs for inefficient plant (Gómez et al., 2017; Lombardi et al., 2019). The slack-based measure (SBM) model proposed by Tone (2002) perfectly solved this problem. On other hand, SBM-DEA model can be combined with clustering analysis to minimize the impact of scale effect on plants performance.

In this context, this study selected SBM-DEA model based on clustering analysis to evaluate the efficiency scores of 681 facilities in rural area of Wuxi district, Jiangsu Province, located in southeastern China. As the biggest demonstration area, these samples are the most representative for and exceed 2/5 of the running facilities all over the country. The purpose of the study is (1) to evaluate the performance efficiency of 681 rural sewage treatment facilities; (2) to identify the improvement potential of inefficient facilities and provide specific improvement suggestions; (3) to identify implicit factors that affect the facility performance. The results can help select out the state-of-art for the construction, operation and management of rural sewage treatment facilities in China, effectively promoting the sustainable development of rural water resources.

2 Methodology

2.1 SBM-DEA model

DEA is a powerful non-parametric comprehensive evaluation method to measure relative efficiency of a large number of decision-making units (DMUs) (Nabavi-Pelesaraei et al., 2019). This method selects the efficient DMUs as reference benchmark to identify levels and causes of inefficient DMUs. Different DEA models had been proposed for different purpose. At present, conventional radial models, such as Charnes-Cooper-Rhodes (CCR) and Banker-Charnes-Cooper (BCC) have been widely used. However, these models assume changes of inputs or outputs are proportional, failing to consider the slack of indicators (Carvalho and Marques, 2011).

By comparison, non-radial SBM-DEA model is more suitable for assessing samples with vague interconnections inputs (Thrall, 1996). It considers input excesses and output shortfalls of DMUs further, providing target improvement value for each inefficient DMU's input and output separately (Castellet and Molinos-Senante, 2016; Wang et al., 2018). What's more, this method can treat environmental impacts as undesirable outputs in the index system to achieve a multi-dimensional assessment of the environment impacts, resources consumption and service value (Guo et al., 2017; Robaina-Alves et al., 2015). Finally, SBM model can be combined with clustering analysis by grouping samples according to the design treatment capacity to evaluate the sample efficiency based on the group-frontier, so as to reduce the impact of scale effect (Jiang et al., 2020).

Based on the above reasons, this study composed an output-oriented SBM-DEA model based on constant scale return (CRS) combined with cluster analysis to evaluate rural sewage treatment facilities. Suppose the number of DMU_S is n and each DMU has m inputs and s outputs. The matrices are expressed as $X=[\boldsymbol{x_{ij}}]\in R^{m\times n}$ and $Y=[\boldsymbol{y_{ij}}]\in R^{s\times n}$. The fractional programming form of SBM model is shown as follows:

$$min\,\rho* = \frac{1-\frac{1}{m}\sum_{i=1}^{m} {s_i}^{-}/x_{i0}}{1+\frac{1}{s}\sum_{r=1}^{l} {s_r}^{+}/y_{r0}}$$

$$\begin{gathered} s.t. \\ x_0 = X\lambda + s^- \\ y_0 = Y\lambda - s^+ \\ \lambda \geq 0\,, s^- \geq 0, s^+ \geq 0 \end{gathered} \qquad (1)$$

where $s^-$and s^+ represent the input excesse and output shortfall, respectively. λ indicates non-negative weight vector. The value of $\rho*$ ranges from 0 to 1. The higher the value of $\rho*$, the better the efficiency of the DMU. When $\rho*$=1, the DMU is relative efficient means no input excess and output shortfall. Otherwise, the DMU is inefficient. Inefficient DMUs can improve score by decreasing input excesses or making up output shortfalls as follows:

$$x_0 - s^- \rightarrow x_0',\ y_0 + s^+ \rightarrow y_0' \qquad (2)$$

Traditionally, DEA model assumed all samples to have the same or similar characteristics when efficiency is evaluated. Therefore, all DMUs were taken as reference set to construct meta-frontiers. In reality, the DMUs not always are homogeneity, which will affect the accuracy of DEA results (Corton and Berg, 2009). Clustering analysis approach can usefully deal with heterogeneous DMUs (Galar et al., 2014). This method divides DMUs into different groups according to certain attributes, which can maximize the homogeneity of samples in the same cluster to decrease the effect of heterogeneity on efficiencies. Then, every group takes itself as reference set, constructing group-frontier separately.

The definition of meta-frontier and group-frontier according to output sets and output distance functions (BATTESE et al., 2004; O'Donnell et al., 2007) are as follows. Assume $\boldsymbol{y}$ and $\boldsymbol{x}$ are the output and input vectors of dimension $X \times 1$ and $Y \times 1$, respectively. All DMUs make up the meta-technology set:

$T^{meta} = \{(\boldsymbol{x}, \boldsymbol{y})|\boldsymbol{x} \geq \boldsymbol{0}; \boldsymbol{y} \geq \boldsymbol{0}: \boldsymbol{x}\ production\ \boldsymbol{y}\}$

The corresponding output set P for input vector can be defined as:

$P^{meta}(\boldsymbol{x}) = \{\boldsymbol{y}|(\boldsymbol{x}, \boldsymbol{y}) \in T^{meta}\}$

The upper bound of this set is the meta-frontier. At this time, meta-distance function can be expressed as:

$D^{meta}(\boldsymbol{x}, \boldsymbol{y}) = inf_{\theta}\ \{\boldsymbol{\theta} > 0: (\boldsymbol{y}/\boldsymbol{\theta}) \in P^{meta}(\boldsymbol{x})\}$, if and only if $D^{meta}(\boldsymbol{x}, \boldsymbol{y}) = 1$, the DMU is efficient.

Similarly, if all samples are divided into subgroups according to specific criteria, the DMUs in the kth group are contained in the group-specific technology set:

$T^{k} = \{(\boldsymbol{x}, \boldsymbol{y})|\ \boldsymbol{x} \geq \boldsymbol{0}; \boldsymbol{y} \geq \boldsymbol{0}; \boldsymbol{x}\ production\ \boldsymbol{y}\}$

The corresponding output set P for input can be defined as:

$P^{k}(\boldsymbol{x}) = \{\boldsymbol{y}|(\boldsymbol{x}, \boldsymbol{y}) \in T^{k}\}$

The upper bound of this set is the group-frontier. At this time, group-distance function can be expressed as:

$D^{k}(\boldsymbol{x}, \boldsymbol{y}) = inf_{\theta}\ \{\boldsymbol{\theta} > 0: (\boldsymbol{y}/\boldsymbol{\theta}) \in P^{k}(\boldsymbol{x})\}$, if and only if $D^{k}(\boldsymbol{x}, \boldsymbol{y}) = 1$, DMU is efficient.

$D^{meta}(\boldsymbol{x}, \boldsymbol{y}) \leq D^{k}(\boldsymbol{x}, \boldsymbol{y})$, $TE^{meta}(\boldsymbol{x}, \boldsymbol{y}) \leq TE^{k}(\boldsymbol{x}, \boldsymbol{y})$, which means the meta-frontier envelops the group-frontier. The difference between results based on two frontiers can be measured by technical gap rate (TGR):

$$TGR^{k}(\boldsymbol{x}, \boldsymbol{y}) = TE^{meta}(\boldsymbol{x}, \boldsymbol{y})/TE^{k}(\boldsymbol{x}, \boldsymbol{y}) \tag{3}$$

The value of TGR ranges from 0 to 1. Assuming that $\mathrm{TE}^{\mathrm{meta}}$ is 0.6 and TE^{k} is 0.8, the TGR would equal 0.75. This means that if the input vector is determined, the maximum output that could be produced by a form group k is 75% of the output that is feasible when using the meta-frontier as a benchmark. The higher value of TGR, the smaller gap between the meta-frontier and group-frontier and the smaller gap between technology used by the DMU and technology frontier.

2.2 Data collection and variables

2.2.1 Data source

This study investigated 681 rural sewage treatment facilities in Wuxi, Jiangsu Province. All facilities removed contaminants by conventional secondary treatment, ensuring the comparability fundamentally. The electricity consumption and water quality data were sampled once a month. In this study, the monthly average data of 2017 was used as the benchmark. The investment and operational data come from the information system of Wuxi Wastewater Treatment Authority.

2.2.2 Inputs and outputs

DEA is a data-oriented method, thus, selecting appropriate inputs and outputs is the key to accurately evaluate relative performance efficiencies of samples. In order to comprehensively evaluate the performance of rural sewage treatment facilities for construction, operation and management, an index system should be constructed from multiple dimensions such as economy, environment and energy consumption. It should be noted that the more variables, the more difficult to distinguish DMUs performance because the number of efficient DMUs increases. This study referred to the indicators selected by the previous researches (Lorenzo-Toja et al., 2015; Sala-Garrido et al., 2011; Wang et al., 2018) of sewage treatment plants evaluation and takes into account the availability of data and the characteristics of the selected model. The minimum number of indicators was selected to ensure the integrity of the evaluation elements. The units of the input and output variables do not affect the efficiency score.

The necessary inputs had been grouped into three categories: (1) capital cost (x_1, 10^4 CNY); (2) operating cost: mainly including labor cost and maintenance cost (x_2, 10^4 CNY/year); (3) electricity consumption: the largest energy consumption of operation (x_3, 10^4 kWh/year). These indicators really reflected resource consumption of rural sewage treatment facilities.

Four operational indicators had been chosen as outputs: (1) treatment capacity (y_1, 10^4 ton/year); (2) chemical oxygen demand removed (COD, ton/year) (y_2); (3) ammonia nitrogen removed (NH_3-N) (y_3, ton/year); (4) total phosphorus removed (TP) (y_4, ton/year). The selection of outputs reflected the service value of rural sewage treatment facilities to improve the quality of rural water environment by treating sewage discharged.

2.2.3 Implicit explanatory factors

In addition to the selected three input factors and four output factors, the performance of the DMUs may also be affected by many other implicit factors. To further determine the best operating conditions, the next step is to identify the implicit factors. Based on the reported studies and the available statistical information, another three factors were considered (Molinos-Senante et al., 2013; Teklehaimanot et al., 2015; Zeng et al., 2017) : (i) technology, (ii) load rate: expressed as the ratio of the actual treatment capacity to the designed treatment capacity, (iii) standard of discharge.

3 Results and discussion

3.1 Sample description

Previous studies confirmed that scale has significant impacts on the efficiency scores of sewage treatment facilities: the plants with larger size operate more effectively (Dong et al., 2017; Hernández-Sancho and Sala-Garrido, 2009). To minimize scale effect, the DMUs were divided into five groups according to design treatment scale of facilities: group 1 ([0, 5) t/d), group 2 ([5, 10) t/d), group 3 ([10, 20) t/d), group 4 ([20, 30) t/d) and group 5 ([30, 80) t/d). A brief description of the inputs and outputs was listed in Table 1. With the increase of the treatment scale, the average values of three inputs and four outputs also gradually increased. The degree of data dispersion (standard deviation) did not show obvious rules.

Table 1 The descriptive statistics of the variables for five groups.

group	variables	x_1 (10^4 CNY)	x_2(10^4 CNY/year)	x_3(10^4 kWh/year)	y_1(10^4 t/year)	y_2(t/year)	y_3(t/year)	y_4(t/year)
1	average	4.316	0.979	0.079	0.109	0.181	0.073	0.004
	stdev	1.390	0.027	0.0270	0.039	0.143	0.034	0.003
2	average	10.000	1.074	0.175	0.254	0.3854	0.166	0.008
	stdev	0.000	0.018	0.018	0.033	0.183	0.058	0.003
3	average	17.592	1.335	0.278	0.446	0.632	0.279	0.016
	stdev	2.567	0.181	0.041	0.088	0.326	0.1189	0.020
4	average	26.775	1.620	0.420	0.754	1.192	0.488	0.026
	stdev	1.143	0.047	0.047	0.111	0.731	0.177	0.014
5	average	51.500	2.101	0.901	1.661	2.491	1.246	0.058
	stdev	20.809	0.398	0.398	0.960	1.441	0.720	0.034

3.2 Efficiency analysis and potential improvement

3.2.1 Efficiency scores

In this study, The SBM model based on CRS and group-frontier was established by MaxDEA Ultra 8 (No 812-182) software. Detailed data and results could be found in Table S1 in Appendix. The average TGRs of the five groups ranged from 0.477 to 0.898, indicating that the gap between the two frontiers was obvious.

Fig. 1 compared the efficiency scores of 681 DMUs under group-frontier with the scores based on the meta-frontier. Based on the group-frontier, the number of DMUs with high scores (> 0.5) increased significantly and the number of efficient facilities (score equals to 1) increased from 10 to 27. This result verified the necessity of evaluating operating performance of rural sewage treatment facilities under different scale frontiers. Therefore, the following analysis in the study was all based on group-frontier.

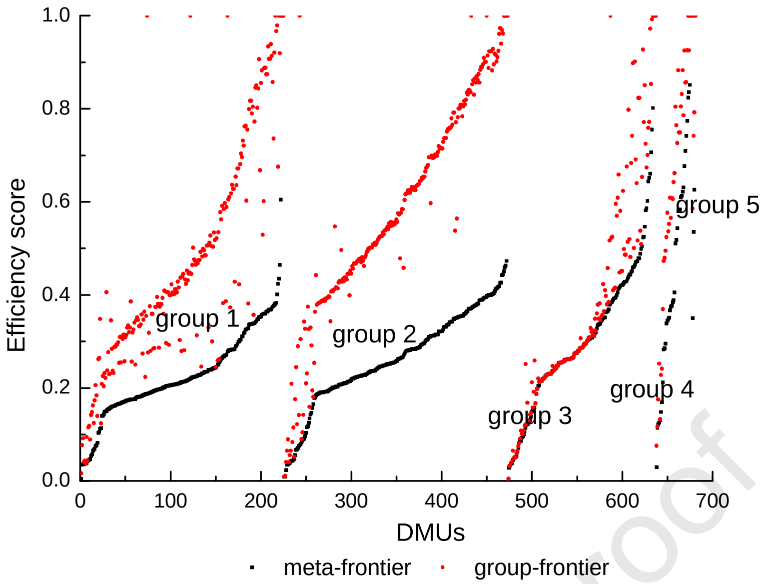


Fig. 1. Efficiency scores of 681 treatment facilities based on meta-frontier and group-frontier respectively.

Fig. 2 showed the number of facilities at different subintervals of efficiency scores based on group-frontier. 27 treatment facilities were relatively efficient, meaning that less than 4% of DMUs located on the optimal production frontier, i.e., maximizing outputs. Considering these treatment facilities as the best benchmark, nearly half of samples (305 out of 681) scored less than or equal to 0.5, which meant that there was great room for improvement in the inefficient facilities. Fig. 3 showed that the average score of the samples was 0.496, so the inefficient DMUs had about 50.4% improvement potential. Thus, how to optimize the allocation of inputs and outputs of inefficient DMUs should be the focus to improve the overall efficiency scores of treatment facilities.

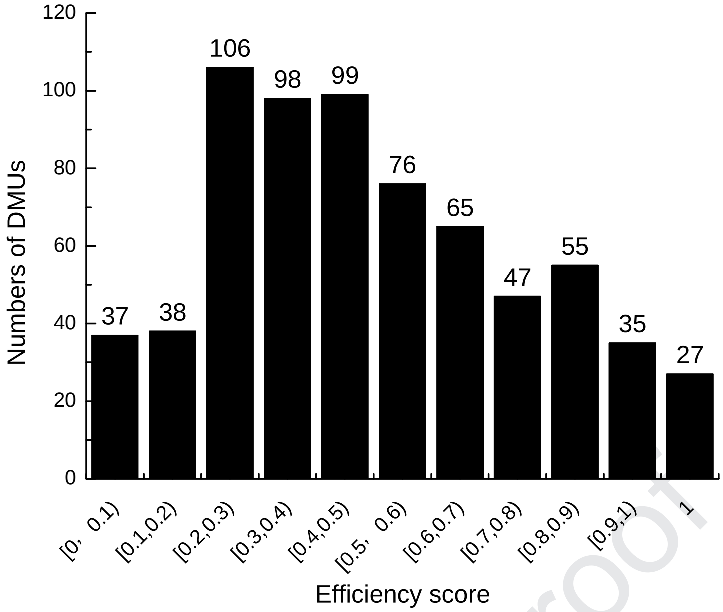


Fig. 2. The number of treatment facilities at different subintervals of efficiency scores based on group-frontier.

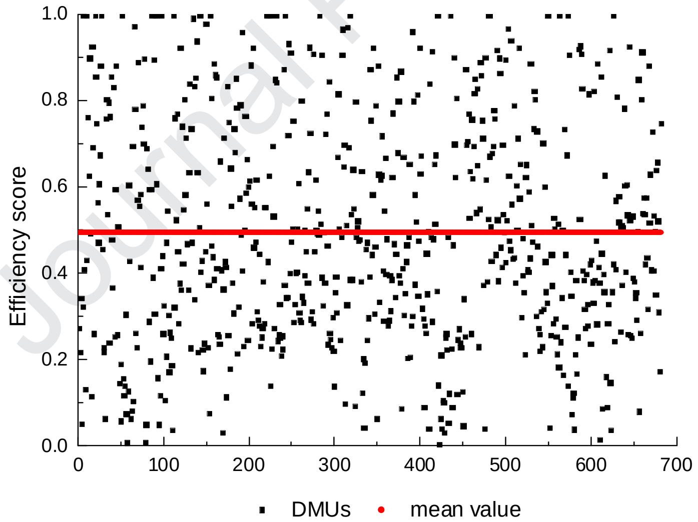


Fig. 3. Efficiency scores of 681 DMUs based on group-frontier.

3.2.2 Potential improvement

As shown in Fig. 4, the difference in the capital and operational costs between inefficient DMUs and efficient DMUs were not significant, showing that the investment of construction and operation for all facilities was overall reasonable. The mean electricity of inefficient DMUs (2379.181 kWh/year) was higher than that of inefficient DMUs (1843.836 kWh/year). Significant output shortfalls existed in inefficient samples. The average values of four output variables for the efficient DMUs were obviously higher than those for the inefficient DMUs. The average annual treatment capacity of efficient plants was 4,541.963 tons, while that of inefficient plants was only 2,831.661 tons. Furthermore, the pollutants removal of an efficient treatment facility was 2 to 3 times that of an inefficient facility.

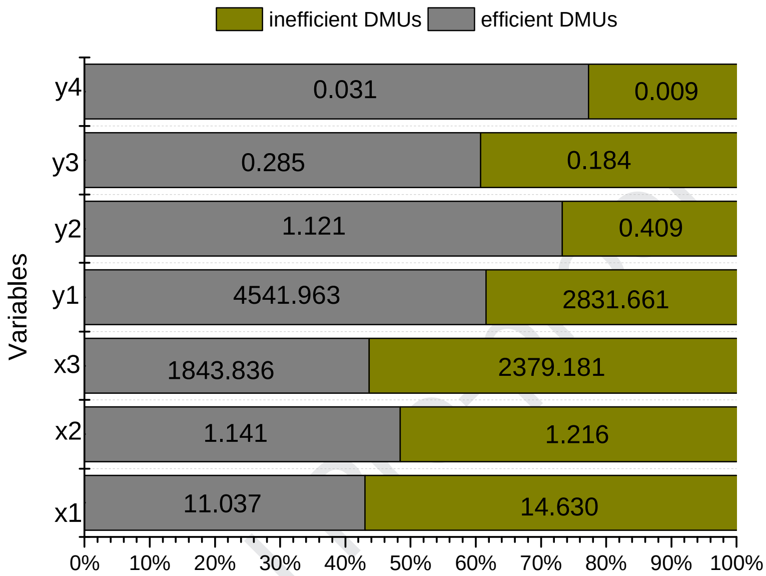


Fig. 4. Comparison of the inputs and outputs for the efficient and inefficient DMUs.

SBM model directly constructs slack variables in the objective function to take the slack of the inputs and the outputs into account. In other words, taking efficient samples as benchmark, it can quantify potential improvement of each item for inefficient DMUs to improve scores of inefficient facilities. The results were shown in Fig. 5 and Table 2. The level of output shortfall in 654 inefficient treatment facilities was serious. For these samples, the treatment capacity (y_1) had the greatest improvement potential, which could improve about 92.45% (1.61×10^5 ton/year) under the current input level. Moreover, the potential improvement for the removal of COD, NH_3-N and TP was 45.49% (357 ton/year), 91.97% (11 ton/year) and 25.33% (20 ton/year) respectively. Under the current output level, the capital cost, the operating cost and electricity consumption could be respectively reduced by 1.74% (130×10^4 CNY), 4.67% (35×10^4 CNY/year) and 8.60% (1.06×10^5 kWh/year). There was almost no input excess. Therefore, the manager of the plants should focus on solving problems of low load operation and poor removal of pollutants.

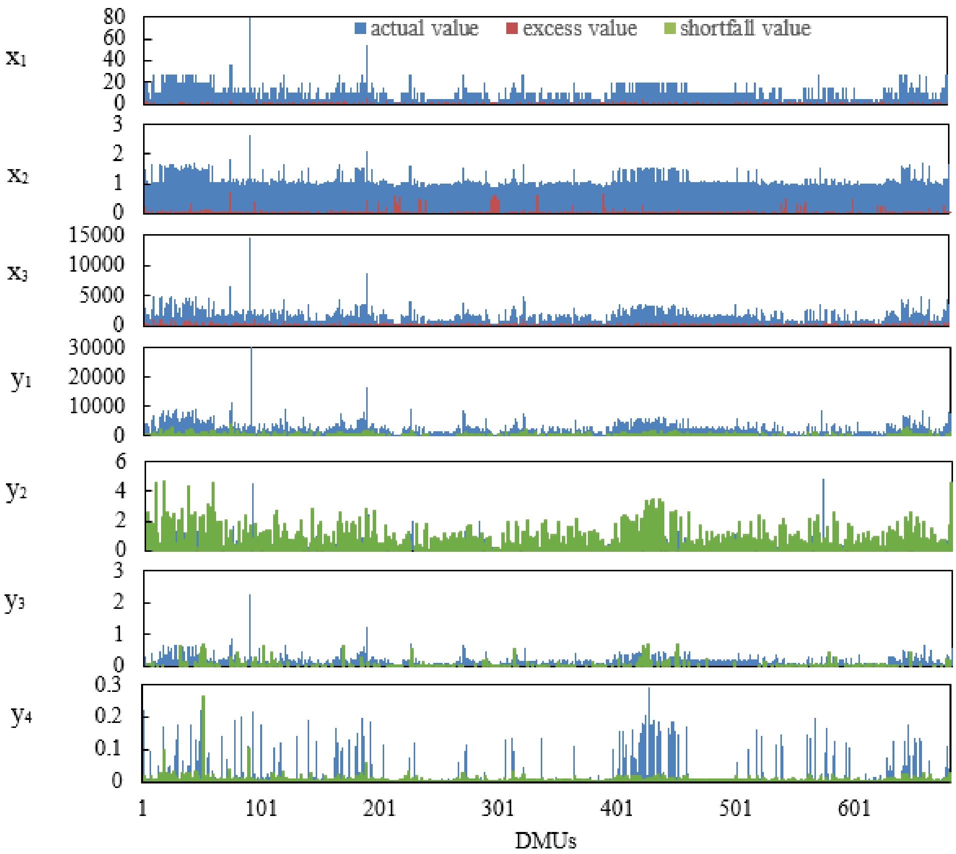


Fig. 5. Potential improvement of each item for every DMU.

Table 2 The mean improvement potential of 681 DMUs.

	capital cost (10^4 CNY)	operating cost (10^4 CNY /year)	electricity consumption (10^4 kWh/year)	treatment capacity (10^4 ton/year)	COD removed (ton/year)	NH_3-N removed (ton/year)	TP removed (ton/year)
origin value	7483	744	117	0.214	0.654	0.139	0.027
target value	7613	779	127	0.197	0.298	0.128	0.007
slack movement	-130	-35	-10	0.017	0.357	0.011	0.020
potential for improvement	-1.74%	-4.67%	-8.60%	92.45%	45.49%	91.97%	25.33%

3.3 Sensitivity analysis of inputs and outputs

The efficiency scores of DMUs are influenced directly by the change of inputs and outputs because each vector introduced uncertainty into DEA model (Castellet and Molinos-Senante, 2016). Changing the input or output of the DMUs to observe the changes in efficiency is the main sensitivity analysis method (Hu et al., 2019). SBM model, as a non-parameter model, the efficiency score has no specific quantitative relations with the number of inputs and outputs (Guo et al., 2017). Thus, omitting one input or one output variable once time to examine degree of change in efficiency score is an effective approach for sensitivity analysis. Fitting the scatters to calculate slope and coefficient of correlation (R^2) of proportional function. Then, the sensitivity of the variables can be identified by the gap between 1 and slope of the function (Hu et al., 2019). The greater the gap, the higher the sensitivity. Fig. 6 and Table 3 showed the result of sensitivity analysis of seven variables. Omitting the variable (y_3), the highest value of |1-slope| (0.167) occurred, indicating the removal of NH_3-N (f) was the most sensitive factor. Other significant factors include TP removed (g), treatment capacity (d) and operating cost (b). The electricity consumption was the least sensitive factor mainly because of the small difference in power consumption of treatment facilities at the same scale. Overall, the outputs were more sensitive than the inputs. Therefore, improving the removal rate of nitrogen and phosphorus and increasing the treatment capacity are the key to the efficient operation of rural sewage treatment facilities.

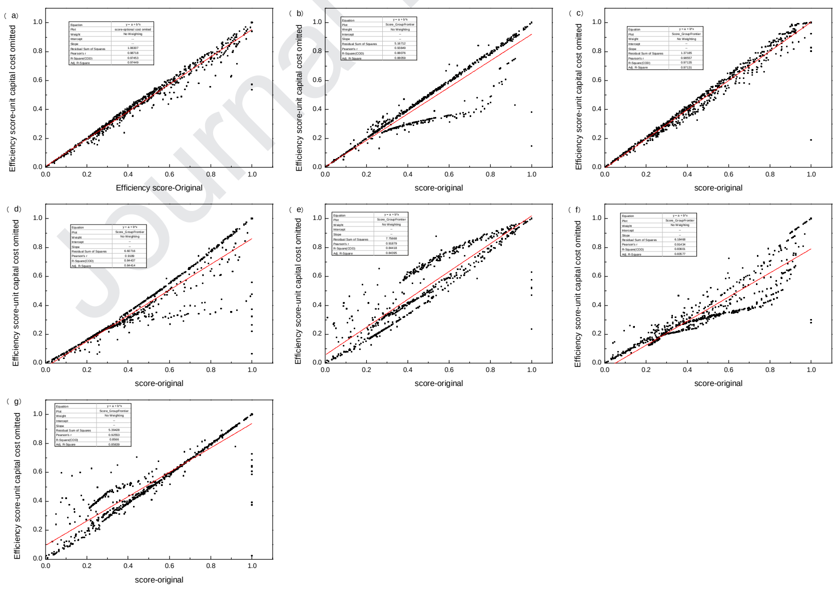


Fig. 6. Sensitivity analysis for capital cost (a), operating cost (b), electricity consumption (c), treatment capacity (d), COD removed (e), NH_3-N removed (f) and TP removed (g).

Table 3 Sensitivity analysis results and variable sensitivity rankings.

Ranking	Variables	Slope	|1-slope|	R^2	Classification
1	operating cost (10^4 CNY /year)	0.916	0.084	0.881	
2	capital cost (10^4 CNY)	0.946	0.054	0.975	Input
3	electricity consumption (10^4 kWh/year)	1.012	0.012	0.971	
4	NH_3-N removed (t/year)	0.833	0.167	0.836	
5	TP removed (t/year)	0.842	0.158	0.857	Output
6	treatment capacity (10^4 t/year)	0.888	0.112	0.844	
7	COD removed (t/year)	0.962	0.039	0.844	

3.4 Implicit explanatory factors

DMUs were grouped according to three selected explanatory factors. The characteristics of efficiency scores were shown in Fig. 7.

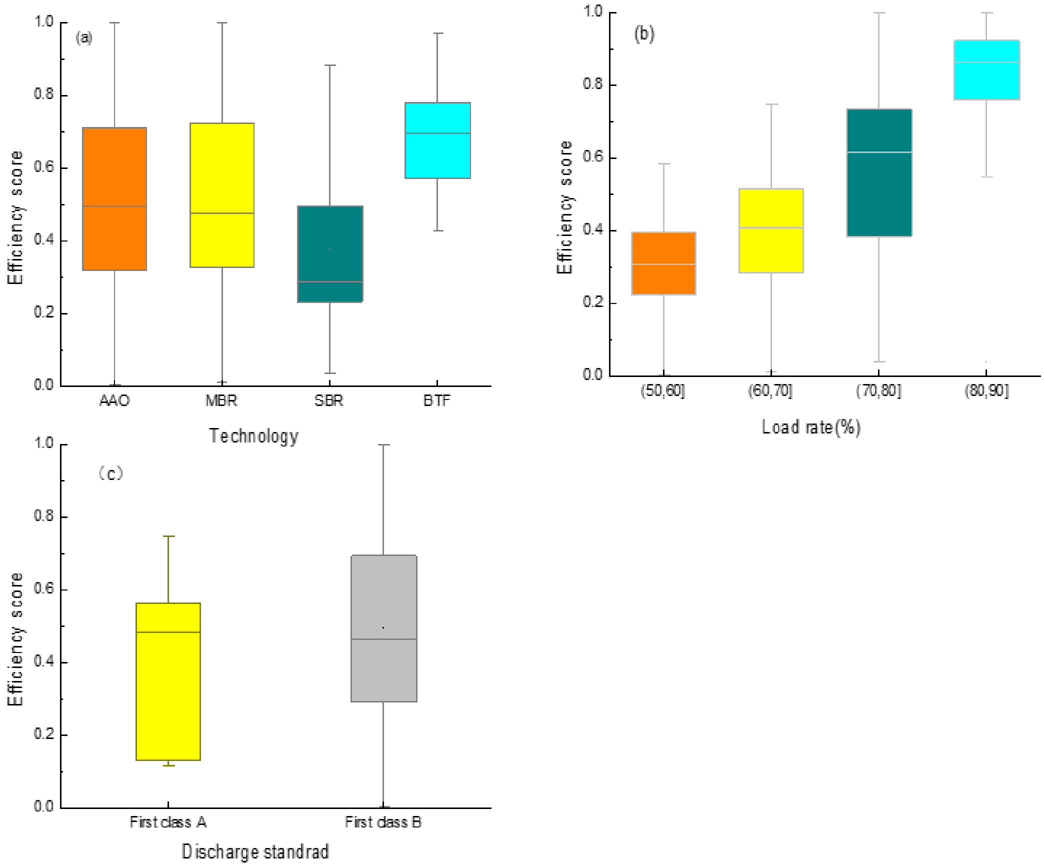


Fig. 7. Box charts of the explanatory factors.

Due to non-normal distribution of analyzed samples, the Kruskal–Wallis non-parametric test, as the most suited way, had been taken to verify significant differences among different groups in this study (Kruskal and Wallis, 1952; Sueyoshi and Aoki, 2001). The statistical significance (p) value is equal or less than 0.05 meaning the explanatory factor significantly impact efficiency scores of samples. Otherwise, the explanatory factor has no significant impact on efficiency score of samples. Table 4 displayed detailed results.

Table 4 Kruskal–Wallis test statistics for explanatory factors.

explanatory factors	total DMUs	mean	std.dev.	P-value	Chi-sq.
Technology				0	38.251
AAO	334	0.519	0.257		
MBR	222	0.515	0.259		
SBR	116	0.382	0.235		
BF	9	0.681	0.162		
Load rate				0	258.706
(50,60]	157	0.316	0.160		
(60,70]	218	0.404	0.178		
(70,80]	187	0.564	0.237		
(80,90]	119	0.799	0.216		
Discharge standard				0.589	0.293
First class A	6	0.422	0.250		
First class B	675	0.497	0.259		

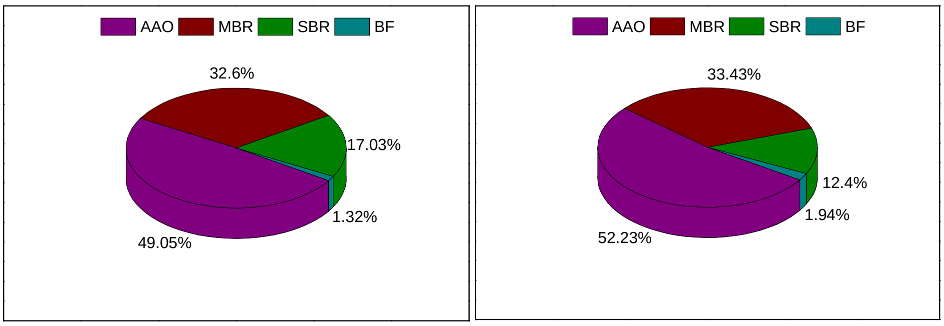
3.4.1 Technology

The sewage treatment technology means removing pollutants in wastewater through physical, chemical and biological processes, directly influencing the removal of pollutants. The process is generally divided into three levels: primary, secondary and tertiary treatment (Jin et al., 2014). Based on the classification of secondary treatment (Rodriguez-Garcia et al., 2011), the 681 samples were divided into four categories: (i) anaerobic-anoxic-oxic process (AAO), (ii) membrane bio-reactor process (MBR), (iii) sequencing batch reactor process (SBR) and (iv) Bio-trickling Filter (BTF).

According to the K-W test results (Table 4), the difference in performance of DMUs across the categories of technology was significant ($p < 0.05$). Hence, selecting efficient and economical technology can improve the performance of facilities. The boxplot for the four technologies efficiency scores were shown in Fig. 7

(a). The average score of AAO, MBR, SBR and BTF was 0.519, 0.515, 0.382 and 0.681, respectively. SBR and MBR had lower scores mainly resulting from their low efficiency in removing contaminants. In addition, SBR and MBR required aerators to provide oxygen source, which increased operation costs and electricity consumption. This conclusion was consistent with previous views (Tolkou and Zouboulis, 2016). BTF had the highest score. When operating cost and energy consumption were similar, BTF had an advantage in pollutant removal, especially for the removal rate of COD (75%) and NH_3-N (94%). Therefore, BTF was suitable for underdeveloped rural areas effectively dealing with small-scale domestic sewage to improve rural water environment. This result agreed with the conclusion of Yang (2011).

The percentages of the number and total treatment capacity of facilities adopting different technologies in WUXI city were shown in Fig. 8. At present, 556 sewage treatment facilities had adopted AAO and MBR and the total treatment capacity was 1.61×10^6 T/A. Only 9 facilities adopted the BTF, accounting for 1.94% of the total treatment capacity. Assuming that all facilities adopt BTF, when the treatment capacity and effect are the same, the average annual operating cost and power consumption of each facility will be reduced by 4,400 CNY and 49.54 kWh, respectively, and the capital cost will also be reduced by 6,400 CNY. Therefore, it is necessary for local government to upgrade of rural domestic sewage treatment facilities and to promote appropriate technology (BTF).



(a) Number of facilities (b) Total treatment capacity

Fig. 8. Percentage of different treatment technologies.

3.4.2 Load rate

A common problem in treatment facilities of rural areas is that the design treatment capacity is significantly higher than the actual treatment water volume, resulting in the idleness of the facilities (Li and Xu, 2015; Yang et al., 2016). In other words, the operating condition of facilities can be affected by not only the design capacity but also actual capacity. Thus, this paper selected load rate as the second implicit factor. DMUs had been divided into four groups based on load rate: (i) 50%-60%, (ii) 60%-70%, (iii) 70%-80% and (iv) 80%-90%. As shown in Table 4 and Fig. 7 (b), the impact of the capacity load rate was significant ($p<0.05$). The average efficiency scores of four groups was 0.316, 0.404, 0.564 and 0.799, respectively. The performance efficiency of DMUs with a high load rate operate relatively better than that of those with a low load rate. Our result was consistent with the finding of Hu et al. (2019). As shown in Fig. 9, the load rates of facilities were all less than 100% also confirmed that phenomenon of idle facilities mentioned above. Therefore, it is essential to design the scale of treatment facilities reasonably to ensure the high load operation of the facilities.

There were also a few DMUs that do not obey this rule: despite the relatively lager scale and higher load rate, the scores of them were very unsatisfactory. This phenomenon had also appeared in M.Molinos-Senante’s study (2013). For example, DMU 31 processed 5694 tons sewage in 2017, with a load rate of 78%, but efficiency score of this plant was only 0.06. Studies showed that the component and concentration of influent influence sewage treatment performance (Dong et al., 2017; Hu et al., 2019). These abnormal inefficient DMUs had low concentration of pollutant inflows, resulting in a poor removal of pollutants. Serious shortfalls of outputs were considered to be the main explanation for the phenomenon. Besides, the relatively higher energy consumption and operation costs also were reasons for low score. Therefore, increasing the concentration of influent by a certain pretreatment while taking the reduction of inputs and the increase of the capacity load rate into account can be a good way to improve the performance of treatment facilities.

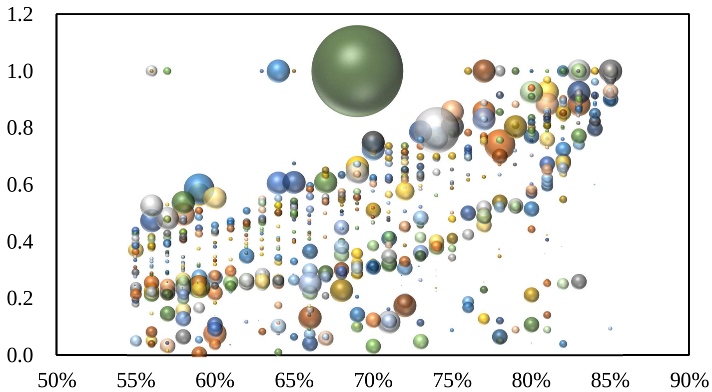


Fig. 9. Efficiency scores of DMUs in WUXI. Bubble size represents the actual capacity of the facilities, and every color represents one facility.

3.4.3 Discharge standard

The discharge standards directly affect the construction, operation and management of rural domestic sewage treatment facilities. According to "Discharge Standard of Pollutants for Municipal Wastewater Treatment Plant (GB18918-2002)" currently implemented in Wuxi rural areas, the samples were divided into two categories: (i) the first class A, (ii) the first class B. As shown in Fig. 7 (d), with the discharge standard more stringent, the efficiency score of samples became lower. The average score decreased from 0.497 to 0.422. At present, the effluent quality of 681 rural sewage treatment facilities all met the first class B standard and 6 (0.89%) facilities met the first class A. Compared with the DMUs that met class B standard, the DMUs meeting the class A can increase the removal of COD, NH_3-N and TP by 0.292, 0.150, and 0.012 tons equally each year, but the operating cost and electricity consumption will equally increase by 5000 CNY and 2587 kWh, respectively. The result of the K-W test showed that discharge standards had no significant effect on performance scores of DMUs. Therefore, upgrading the standard seems not an ideal measure to improve performance scores of rural sewage treatment facilities. Considering the effluent water quality was good, the tail water reuse should be the focus of the local government, which will not only improve the reuse rate of water resources, but also greatly reduce the cost of rural water environmental pollution treatment.

4 Conclusion

With the number and capacity of rural sewage treatment facilities increasing, a comprehensive, quantitative and objective evaluation of them is becoming urgent. DEA is considered to be an effective performance evaluation tool to solve this problem. In this paper, 681 rural sewage treatment facilities were evaluated by SBM-DEA model based on group-frontier from multiple dimensions including economy, environment and society. The main results are as follows: (1) the average efficiency score of samples was 0.496, of which only 27 facilities were operating effectively; (2) compared with efficient DMUs, the inefficient DMUs had significant shortfalls in the outputs, especially in treatment capacity and NH_3-N removal, respectively with the improvement potential of 92.45% and 91.97%; (3) the removal of nitrogen and phosphorus and treatment capacity are the sensitive factors to the efficiencies of rural sewage treatment facilities; (4) technology and capacity load rate had significant impacts on the performance of facilities.

Based on the results above, the targeted recommendations presented as follows to improve the performance of rural sewage treatment infrastructures in China: (1) upgrade and optimize treatment technologies: applying technologies which achieve the trade-off between pollutant removal and cost inputs, such as BTF process; (2) adjust operating conditions: increasing the operating load to avoid facilities idleness and increasing the concentration of influent by pretreatment; (3) encourage reuse of reclaimed water: reusing reclaimed water in various ways to achieve environment benefits and reduce the cost of rural water pollution treatment.

The SBM model selected in this paper identifies the efficient DMUs as the best practices, calculating slack improvement value of inputs and outputs to maximize the efficiencies of the inefficient facilities. It can help government and mangers of water companies to evaluate the operation performance of a large number of sewage treatment facilities and realize the effective supervision and management of local facilities. On the other hand, this method obtains the relative efficiency of the evaluation object, its absolute environmental impact being unknown yet. Besides, this article has not given the quantitative suggestion of improving the performance score.

Thus, further research can combine DEA with quantitative analysis methods such as life cycle assessment or cost-effectiveness analysis to evaluate efficiency of facilities more accurately and provide quantitative improvement measures.

Acknowledgments

This work was supported by Major Science and Technology Program for Water Pollution Control and Treatment (2017ZX07204001) from Ministry of Science and Technology of the People's Republic of China. The authors thank Wuxi wastewater treatment authority for providing data collecting help during the preparation of this paper.

Appendix A. Supplementary data

Analytical data related to this article can be found at online version and the initial data that support the finding of this study are available from the corresponding author upon request.

References

Alemany, J., Comas, J., Turon, C., Balaguer, M.D., Poch, M., Puig, M.A., Bou, J., 2005. Evaluating the application of a decision support system in identifying adequate wastewater treatment for small communities. A case study: the Fluvia River Basin. Water Science & Technology 51(10), 179-186.

BATTESE, G.E., RAO, D.S.P., O'DONNELL, C.J., 2004. A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. Journal of Productivity Analysis(21), 91-103.

Benedetti, L., Dirckx, G., Bixio, D., Thoeye, C., Vanrolleghem, P.A., 2008. Environmental and economic performance assessment of the integrated urban wastewater system. J Environ Manage 88(4), 1262-1272.

Carvalho, P., Marques, R.C., 2011. The influence of the operational environment on the efficiency of water utilities. Journal Of Environmental Management 92(10), 2698-2707.

Castellet, L., Molinos-Senante, M., 2016. Efficiency assessment of wastewater treatment plants: A data envelopment analysis approach integrating technical, economic, and environmental issues. J Environ Manage 167, 160-166.

Chen, C., LeGates, R., Zhao, M., Fang, C., 2018. The changing rural-urban divide in China's megacities. Cities 81, 81-90.

Corton, M.L., Berg, S.V., 2009. Benchmarking Central American water utilities. Util Policy 17(3-4), 267-275.

Dong, X., Zhang, X., Zeng, S., 2017. Measuring and explaining eco-efficiencies of wastewater treatment plants in China: An uncertainty analysis perspective. Water Res 112, 195-207.

Galar, M., Fernández, A., Barrenechea, E., Herrera, F., 2014. Empowering difficult classes with a similarity-based aggregation in multi-class classification problems. Inform Sciences 264, 135-157.

Gallego, A., Hospido, A., Moreira, M.T., Feijoo, G., 2008. Environmental performance of wastewater treatment plants for small populations. Resources, Conservation and Recycling 52(6), 931-940.

Gómez, T., Gémar, G., Molinos-Senante, M., Sala-Garrido, R., Caballero, R., 2017. Assessing the efficiency of wastewater treatment plants: A double-bootstrap approach. Journal of Cleaner Production 164, 315-324.

Gu, B., Fan, L., Ying, Z., Xu, Q., Luo, W., Ge, Y., Scott, S., Chang, J., 2016. Socioeconomic constraints on the technological choices in rural sewage treatment. Environ Sci Pollut Res Int 23(20), 20360-20367.

Guo, Y., Liu, W., Tian, J., He, R., Chen, L., 2017. Eco-efficiency assessment of coal-fired combined heat and power plants in Chinese eco-industrial parks. Journal of Cleaner Production 168, 963-972.

Hernández-Sancho, F., Sala-Garrido, R., 2009. Technical efficiency and cost analysis in wastewater treatment processes: A DEA approach. Desalination 249(1), 230-234.

Hosseinzadeh-Bandbafha, H., Safarzadeh, D., Ahmadi, E., Nabavi-Pelesaraei, A., 2018. Optimization of energy consumption of dairy farms using data envelopment analysis – A case study: Qazvin city of Iran. Journal of the Saudi Society of Agricultural Sciences 17(3), 217-228.

Hu, W., Guo, Y., Tian, J., Chen, L., 2019. Eco-efficiency of centralized wastewater treatment plants in industrial parks: A slack-based data envelopment analysis. Resources, Conservation and Recycling 141, 176-186.

Jiang, H., Hua, M., Zhang, J., Cheng, P., Ye, Z., Huang, M., Jin, Q., 2020. Sustainability efficiency assessment of wastewater treatment plants in China: A data envelopment analysis based on cluster benchmarking. Journal of Cleaner Production 244.

Jin, L., Zhang, G., Tian, H., 2014. Current state of sewage treatment in China. Water Res 66, 85-98.

Kalbar, P.P., Karmakar, S., Asolekar, S.R., 2012. Selection of an appropriate wastewater treatment technology: a scenario-based multiple-attribute decision-making approach. J Environ Manage 113, 158-169.

Kruskal, W.H., Wallis, W.A., 1952. Use of Ranks in One-Criterion Variance Analysis.

Journal of the American Statistical Association 47(260), 583-621.

Li, X., Xu, J., 2015. Reflection on the general idleness of rural sewage treatment facilities in Beijing. Water Supply and Sewerage 41(6), 48-50.

Lombardi, G.V., Stefani, G., Paci, A., Becagli, C., Miliacca, M., Gastaldi, M., Giannetti, B.F., Almeida, C.M.V.B., 2019. The sustainability of the Italian water sector: An empirical analysis by DEA. Journal of Cleaner Production 227, 1035-1043.

Lorenzo-Toja, Y., Vazquez-Rowe, I., Chenel, S., Marin-Navarro, D., Moreira, M.T., Feijoo, G., 2015. Eco-efficiency analysis of Spanish WWTPs using the LCA plus DEA method. Water Research 68, 651-666.

Molinos-Senante, M., Hernandez-Sancho, F., Sala-Garrido, R., 2013. Benchmarking in wastewater treatment plants: a tool to save operational costs. Clean Technologies and Environmental Policy 16(1), 149-161.

Mostashari-Rad, F., Nabavi-Pelesaraei, A., Soheilifard, F., Hosseini-Fashami, F., Chau, K.-w., 2019. Energy optimization and greenhouse gas emissions mitigation for agricultural and horticultural systems in Northern Iran. Energy 186.

Nabavi-Pelesaraei, A., Rafiee, S., Mohtasebi, S.S., Hosseinzadeh-Bandbafha, H., Chau, K.-w., 2019. Assessment of optimized pattern in milling factories of rice production based on energy, environmental and economic objectives. Energy 169, 1259-1273.

O’Donnell, C.J., Rao, D.S.P., Battese, G.E., 2007. Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. Empirical Economics 34(2), 231-255.

Robaina-Alves, M., Moutinho, V., Macedo, P., 2015. A new frontier approach to model the eco-efficiency in European countries. Journal of Cleaner Production 103, 562-573.

Rodriguez-Garcia, G., Molinos-Senante, M., Hospido, A., Hernandez-Sancho, F., Moreira, M.T., Feijoo, G., 2011. Environmental and economic profile of six typologies of wastewater treatment plants. Water Research 45(18), 5997-6010.

Sala-Garrido, R., Molinos-Senante, M., Hernández-Sancho, F., 2011. Comparing the efficiency of wastewater treatment technologies through a DEA metafrontier model.

Chemical Engineering Journal 173(3), 766-772.

Shen, F., Zhang, K., Li, J., 2014. Evaluation method for engineering technology of rural domestic sewage treatment based on fuzzy integral model. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE) 30(15), 272-280.

Sueyoshi, T., Aoki, S., 2001. A use of a nonparametric statistic for DEA frontier shift: the Kruskal and Wallis rank test. Omega-Int J Manage S 29(1), 1-18.

Teklehaimanot, G.Z., Kamika, I., Coetzee, M.A., Momba, M.N., 2015. Population Growth and Its Impact on the Design Capacity and Performance of the Wastewater Treatment Plants in Sedibeng and Soshanguve, South Africa. Environ Manage 56(4), 984-997.

Thrall, R.M., 1996. Chapter 5 Duality, classification and slacks in DEA. Annals of Operations Research 66(2), 109-138.

Tolkou, A.K., Zouboulis, A.I., 2016. Effect of climate change in WWTPs with a focus on MBR infrastructure. Desalin Water Treat 57(5), 2344-2354.

Tone, K., 2002. A slacks-based measure of super-efficiency in data envelopment analysis. Eur J Oper Res 143(1), 32-41.

Wang, S., Qiu, S., Ge, S., Liu, J., Peng, Z., 2018. Benchmarking Toronto wastewater treatment plants using DEA window and Tobit regression analysis with a dynamic efficiency perspective. Environ Sci Pollut Res Int 25(32), 32649-32659.

Wu, S., Austin, D., Liu, L., Dong, R., 2011. Performance of integrated household constructed wetland for domestic wastewater treatment in rural areas. Ecological Engineering 37(6), 948-954 (in Chinese).

Xia, X., Wang, M., Ming, H., 2012. Rural sewage treatment technology evaluation method based on fuzzy advantages and disadvantages coefficient. Journal of Environmental Sciences 32(9), 2287-2293 (in Chinese).

Yang, F., 2011. Application of Compound Biological Drip Filter Technology in Domestic Wastewater Treatment in Small Towns. Environmental protection and recycling economy 31(2), 59-61(in Chinese).

Yang, X., Yuan, J., Yao, M., Huang, W., Wang, D., 2016. Current Status and

Development Countermeasures of Domestic Sewage Treatment in China-Taking the Rural Areas of Southern Jiangsu as an Example. Journal of Fudan University (Natural Science) 55(2) (in Chinese).

Zeng, S., Chen, X., Dong, X., Liu, Y., 2017. Efficiency assessment of urban wastewater treatment plants in China: Considering greenhouse gas emissions. Resources, Conservation and Recycling 120, 157-165.

Highlights

- Efficiency scores of 681 rural sewage treatment facilities were assessed by SBM-DEA model based on group-frontier.
- The improvement potential for samples was about 50.4%.
- 27 treatment facilities were regarded as best practices.
- Explicit factors affecting the performance of treatment facilities were discussed.
- Suggestions to improve efficiency of facilities in rural areas of China were proposed.

Declaration of interests

☑ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: