Benchmarking sustainability of urban water infrastructure systems in China

Xin Dong a, b, Xinning Du a, Ke Li c, Siyu Zeng a, b, *, Brian P. Bledsoe c

a School of Environment, Tsinghua University, Beijing, 100084, China
b Environmental Simulation and Pollution Control State Key Joint Laboratory, School of Environment, Tsinghua University, Beijing 100084, China
c College of Engineering, University of Georgia, GA, USA, 30602

A R T I C L E I N F O

Article history:
Received 6 April 2017
Received in revised form 31 August 2017
Accepted 5 September 2017
Handling Editor: Yutao Wang

Keywords:
Data envelopment analysis
Urban water infrastructure system
Sustainability
Benchmarking
China

A B S T R A C T

The broad scope and definition of sustainability has perplexed assessment of water infrastructure systems, especially for the purpose of directing engineering practices when quantified criteria are desired. An input-oriented data envelopment analysis (DEA) was improved to benchmark the relative sustainability of the water infrastructure of 157 cities in China. The DEA calculates a single sustainability score using seven inputs and five outputs that represent the economic, resource and environmental dimensions of sustainability. Overall, 69 out of the 157 sampled systems obtained high sustainability scores. Eight specific efficiency indicators based on individual DEA input to output ratio were evaluated to identify the causes of performance differences. Compared to water supply systems, the performance of wastewater treatment plants has greater influence on the sustainability score of the overall system. For all systems, the sustainability scores are more sensitive to sludge production and electricity consumption than capital investment and removal efficiency of treatment processes. The DEA provides guidelines to cities for setting priorities in order to meet specific sustainability criteria. Statistical analysis indicates that the overall sustainability score primarily depends on the system scale, meteorological conditions such as air temperature and rainfall, and source water quality.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Reliable and secure water systems are a pre-requisite for the health, prosperity and security of a nation. Assessments of water infrastructure sustainability are complicated due to the various interconnected factors driving its construction, maintenance and operation. Sustainable and resilient water infrastructure development requires long-term economic inputs that are controlled by decision making at a variety of governance levels (Hossain et al., 2015). The extensive water-energy-pollutant nexus in urban water systems increases the complexity of sustainable water infrastructure management (Han et al., 2015). Variation of facility scale, technology adopted, and geographical and meteorological conditions further complicate sustainability assessment at regional and national scales. This calls for a thorough and systematic benchmark to facilitate large-scale comparisons and guide water facilities towards sustainable practices.

Loucks proposed a sustainability index based on reliability, resilience and vulnerability to facilitate the evaluation and comparison of water management policies (Loucks, 1997). This index is based on the concept of life cycle assessment (LCA), which has been applied to identify the main environmental impacts of water infrastructure systems (Emmerson et al., 1995; Hospido et al., 2004; Pasqualino et al., 2009; Garcia et al., 2011). Cost-benefit analysis (CBA) has also been applied to compare the economic feasibility of implementing water infrastructures by defining environmental benefits as positive externalities and assigning monetary valuation (Chen and Wang, 2009; Godfrey et al., 2009; Senante et al., 2010, 2011; Seguí et al., 2009). However, given the current lack of consensus on the metrics and definitions of sustainability, a more quantitative and comprehensive benchmarking approach is necessary for assessing systems with varied size and system properties.

Originated from production theory in economics, data envelopment analysis (DEA) is a nonparametric method to empirically measure the productive efficiency of decision making units (DMUs). Unlike other benchmark methods that assume particular
functional forms, DEA seeks a balanced best practice benchmark based on a set of selected metrics, which can be customized based on local priorities (Sherman and Zhu, 2013). By forming a “composite” system that produces the most output at specific input levels, DEA allows the calculation of an efficient solution for any level of input and output, and therefore provides a criterion to evaluate the system performance. The DEA approach has been validated for the comparison of water and wastewater plant efficiency under a variety of input variables (Lambert et al., 1993; Thanassous, 2000). More recently, DEA has been applied worldwide to analyze the relationship between treatment efficiency and a variety of factors such as labor costs, operational and capital costs, energy and water consumption, water quality, treatment technology, network length, and so on (Thanassous, 2002; Anwandt and Ozuna, 2002; Resende and Tupper, 2009; Garcia, 2006; Kirkpatrick et al., 2006; Tadeo et al., 2008, 2009; Renzetti and Dupont, 2009; Sancho and Garrido, 2009; Byrnes et al., 2010; Garrido et al., 2012; Munisamy and Arabi, 2015; Toja et al., 2015). The factors ranged across social, environmental, and economic perspectives and suggested applicability of DEA as a sustainability benchmarking approach.

Three DEA studies have been performed for water and wastewater treatment plants in China at the province level (Hu et al., 2006; Bian and Yang, 2010; Bian et al., 2014). The results are inconsistent because the efficiency of water infrastructure system is often determined by constraints at city scale. In order to obtained a benchmark for the water infrastructure at city scale in China, in this paper, a novel DEA approach was improved by selecting additional factors covering three aspects of sustainability and collecting a comprehensive dataset to estimate the sustainability of urban water systems. An inventory of water infrastructures was compiled for 657 cities in China. Data from 157 cities were chosen to construct decision-making units for data envelopment analysis that evaluates the sustainability of urban water infrastructure systems in China at the city scale. In contrast to the classic DEA method that only evaluates efficiency, the outputs of the novel DEA assessment approach are sustainability scores for each city. The results can be used to compare relative sustainability among different systems and shed light on system and process bottlenecks for achieving sustainability. The large dataset allowed statistical analysis to uncover the factors that most influence system sustainability.

2. Methods

2.1. System conceptualization and interpretation of sustainability

Urban water infrastructure systems are complex with multiple operational components and corresponding functions. Four basic elements were identified to represent a typical system and its fundamental utility, i.e., water treatment plant (WTP), water distribution network, drainage network (combined or separate sewer), and wastewater treatment plant (WWTP). In this study, urban water facilities in a city were conceptually integrated into one system with four types of elements, which was treated as one decision-making unit (DMU) for DEA.

Although the definition of sustainability is still debatable, it is commonly agreed that it should include three dimensions: economy, environment, and society. The metrics for sustainability assessment applied in this study are: life cycle cost for capital and operation/maintenance, resource/energy consumption and recovery, and environmental impacts. Water infrastructure can be both a fundamental utility, i.e., water treatment plant (WTP), and a source of energy, resources and environmental assets. Sewer network was quantified for the energy consumption and environmental impacts, as shown in Figure 2.1. Sources with energy and materials input, it also contains abundant energy and has the potential to produce sludge for soil improvement (Shen et al., 2008; Heubeck et al., 2011; McCarty et al., 2011).

2.2. DEA input/output selection and DMUs for assessment

In accordance with the selected sustainability metrics, seven inputs were chosen as inputs for DEA to represent economic costs, energy consumption and environmental impacts, as shown in Table 1. The economic costs of the water distribution network and sewer network were quantified based on the total length of pipes and unit cost due to data availability. Since electricity was the main energy source during the operation of the infrastructure system and one of the largest proportions of operation cost, the corresponding variables \( x_{ep} \) and \( x_{en} \) were selected as the key inputs. Sludge production \( x_{sp} \) was selected to represent the environmental performance of the system.

Five outputs were selected in the assessment to summarize the function of water infrastructure, including two outputs for water resource production and three outputs for pollutant removal (Table 1). Three major water quality parameters, the loads of chemical oxygen demand (COD), suspended solids (SS), and total nitrogen (TN) removed from wastewater, were chosen to depict the environmental impacts of the system.

Information was collected and compiled for water infrastructure systems in 657 cities from different geographical regions in China. Statistical data for the inputs and outputs of the year 2014 were retrieved from the Chinese Urban Construction Yearbook, the Urban Water Supply Yearbook, and the Urban Drainage Yearbook (Ministry of Housing and Urban-Rural Development, 2015; China Urban Water Association, Urban Water Supply Yearbook, 2014; China Urban Water Association. Urban Drainage Yearbook, 2014). Among them, 157 cities were selected for the DEA due to the comprehensiveness and representativeness of the available data. The geographical distribution of numbers of DMUs were 22, 15, 54, 23, 23, and 20 in the North, Northeast, East, Central, South, and Southwest of China, respectively, as listed in Table S1 in Appendix. Sample cities are distributed all over the country except for Northwest China; however, 109 of 157 cities are located in the central and eastern China because most of the Chinese population, GDP output and infrastructure construction are concentrated.

2.3. DEA model selection and sensitivity analysis

There are several types of DEA models. Charnes et al. first established the Charnes-Cooper-Rhodes (CCR) model in which the efficient frontier was assumed to be a constant return to scale (CRS) (Charnes et al., 1978). Banker et al. then removed the CRS assumption and proposed the Banker-Charnes-Cooper (BCC) model, which is better suited to assessment of complex systems for which the function that governs input and output relationship is unknown. (Banker et al., 1984). In this study, the BCC model was selected to perform the assessment, in order to take into account various returns to scale in different cities.

When a DMU reaches the maximum output given a vector of inputs (output-oriented DEA), or uses a minimum of inputs to produce a given output (input-oriented DEA), it is placed on the production frontier and thus deemed to be sustainable. The input-oriented model provided a series of target values for the minimized inputs that deliver appropriate benchmarks to calculate the target theoretical performance of unsustainable DMUs. This method was adopted in this study.

The inputs and outputs for the \( k \)th DMU are represented by column vectors \( x_k \) and \( y_k \), respectively. The dataset consists of the
input matrix $X_{157}$ and output matrix $Y_{157}$. The basic linear equations for an input-oriented BCC model are as follows.

$$
\begin{align*}
\min_{\theta, \lambda} & \quad \theta \\
\text{s.t.} & \quad -y_k + Y\lambda \geq 0 \\
& \quad \theta x_k - X\lambda \geq 0 \\
& \quad NE\lambda = 1 \\
& \quad \lambda \geq 0 \\
& \quad k = 1, 2, \ldots, 157
\end{align*}
$$

where $\theta$ is the objective function, and operates once for each DMU in the sample; $\lambda$ is a $157 \times 1$ vector of constants that locates points on the frontier. $NE$ is a $157 \times 1$ vector of ones. In this study, the sustainability level is given by scores obtained in $\theta$ relative to $\lambda$. The model attempts to decrease all inputs proportionally when measuring the urban water infrastructure system’s sustainability. The sustainability score $\theta$ is greater than zero but not more than 1. In this study, the model was solved using the software MaxDEA. With data entered and BCC model established, MaxDEA can provide overall scores and as well as improvement targets of inputs and outputs.

Due to propagation of uncertainties in the collected data, we performed a sensitivity analysis to identify critical external factors and evaluate the validity of the DEA (Sin et al., 2011). The sensitivity analysis followed the approach proposed by Charnes et al. which was performed by changing the input/output variables in increments of 10% for sustainable DMUs while keeping those for unsustainable DMUs constant (Charnes et al., 1992). The seven input variables were decreased by 10% yet the five outputs were increased by 10% for each iteration. The change of sustainability score was examined to evaluate the sensitivity of the inputs and outputs.

### 2.4. Exploration of influential factors of system sustainability

To characterize and explain the sustainability disparities of water infrastructure systems, eight efficiency indicators were defined by ratios of a specific input to an output to depict the system from the different angels as below.

$$
\begin{align*}
IS &= x_{de}/y_{wc} \quad IT &= x_{dw}/y_{ww} \\
ES &= x_{ec}/y_{wc} \quad ET &= x_{ew}/y_{ww} \\
RC &= x_{ew}/y_{pc} \quad RS &= x_{ew}/y_{ps} \\
RN &= x_{ew}/y_{pn} \quad SW &= x_{sp}/y_{ww}
\end{align*}
$$

The impacts of a variety of socio-economic, meteorological, and environmental constraints on the variations of system sustainability scores and specific efficiency indicators were explored using K-W test, multiple nonlinear regression, K-S test, and M-W test with IBM SPSS 20.0.

### 3. Results

#### 3.1. Characteristics of sample systems

Fig. 1 shows a summary of the range and distribution of the input/output of the 157 samples. In general, fixed assets investment for WTP was higher than that for WWTP, while the length of the distribution network was larger than that of the sewer network. The level of electricity consumption in water supply and wastewater treatment was quite similar to each other.

The geographical property of the mean values of the systems followed the economic development pattern in China. Systems in the East China had the highest mean values of the investment, energy consumption, and quantity of water supply and wastewater treatment while the systems in Southwest China had the lowest means.

Table 2 lists statistics of the eight efficiency indicators. The large variance demonstrates technical differences among cities and possible improvement through technology exchange or promoting best available practices. The variance of energy related efficiencies including $ES$, $ET$, $RC$, $RS$, and $RN$ is larger than that of other indicators, and the indicator $IT$ and $SW$ had the least coefficient of variation (44% and 58%, respectively). It clearly shows that energy conservation is a bottleneck for the sustainability of wastewater infrastructure in many cities.

#### 3.2. Scores and gaps of system sustainability

The DEA model calculates a sustainability score for the water infrastructure system of each city. Statistical results of all cities in detail are shown in Table S1 in Appendix. Fig. 2 shows a histogram of sustainability score. A total of 69 water infrastructure systems had the score of 1.0 and were deemed the best practice, which account for approximately 44.0% of all the samples. The mean value of the sustainability scores was 0.883 and the standard deviation of the system sustainability was 0.142, indicating a relatively even
distribution of system performance. Ninety-two percent of the underperformed systems scored between 0.6 and 1.0. Nonetheless, distinguished disparities existed among the systems, and more than half of water infrastructure systems in China show high potential for improving their overall relative performance. It should be noted that, due to the nature of DEA method, i.e. multi-objective evaluation, as the number of evaluation indicators increases, the discrimination of the systems under evaluation is reduced. Therefore, a system with a score of 1 is not an optimal solution, but a non-inferior solution. This means that all performances of this system is

<table>
<thead>
<tr>
<th>Efficiency Indicator</th>
<th>Mean</th>
<th>Median</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS ($/m^3$)</td>
<td>0.87</td>
<td>0.67</td>
<td>0.11</td>
</tr>
<tr>
<td>IT ($/m^3$)</td>
<td>0.64</td>
<td>0.61</td>
<td>0.04</td>
</tr>
<tr>
<td>ES (kWh/m^3)</td>
<td>0.24</td>
<td>0.20</td>
<td>0.21</td>
</tr>
<tr>
<td>ET (kWh/m^3)</td>
<td>0.24</td>
<td>0.22</td>
<td>0.18</td>
</tr>
<tr>
<td>RC (kWh/kg COD)</td>
<td>1.15</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>RS (kWh/kg SS)</td>
<td>2.15</td>
<td>1.62</td>
<td>3.45</td>
</tr>
<tr>
<td>RN (kWh/kg TN)</td>
<td>15.86</td>
<td>12.30</td>
<td>15.53</td>
</tr>
<tr>
<td>SW (kg dry sludge/m^3)</td>
<td>0.15</td>
<td>0.14</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Fig. 1. Inputs and outputs of sample systems.

Fig. 2. Distribution of sustainability scores.
not the best, and this system can be improved in some aspects.

To make comparisons with water infrastructure systems outside China, three new DMUs were added for Spain, German, and U.S, respectively, as shown in Table 3. The data were from United Nation’s Data, World Bank, US EPA, and literature (UN Statistics Division, 2015; World Bank, 2015; U.S. EPA, 2014; Li and Wu, 2006; Jia et al., 2009; Tang and Zhang, 2015). According to the data, there was a big difference on the energy, economic and environmental performance of water infrastructure system among these four countries. To be specific, annual clean water supplied ($y_{ucc}$) and wastewater treated ($y_{uw}$) in China were much higher than that in other three countries, which was directly related to China’s vast territory, large population and high demand for water supply and drainage. Besides, the unit investment in water supply ($x_{lc}$) and drainage ($x_{lw}$) in China were also obviously higher than that in other countries. It seems that China now attatches great importance to the construction and economic input of infrastructure, and there is much room and possibility for the improvement of China’s urban water infrastructure systems.

The DEA indicated that the sustainability scores of the 157 cities remained almost the same with or without three extra DMUs, while Spain, German, and U.S all got a sustainability score of 1.0. This suggests that, despite the technical, management and policies differences, national averages of the three countries are at the same level of the best practices in China. Meanwhile, the water infrastructure systems with the score of 1.0 are comparable to these developed countries. Overall, the best practices in China can be used as a benchmark for sustainability. However, as we discussed previously, more than half (56%) of the systems need enhancements to achieve sustainability goal.

3.3. Sensitive inputs and outputs for sustainability assessment

Fig. 3 shows the result of sensitivity analysis. The “baseline” stands for the result of DEA without any inputs or outputs variation. Overall, the benchmark is not significantly sensitive to any of the variables alone. The number of systems meet the benchmark is more sensitive to the change of inputs than the average sustainability score. Variation of sludge production ($x_{up}$) is the most sensitive factor of which a 10% reduction could lead to the number of best practice systems varies from 69 to 74, an increase of approximately factors of which a 10% reduction could lead to the number of best practice systems. Meanwhile, the water infrastructure systems were iteratively ranked as the input values were changed proportionately. The first 25 percent of systems that need the most input reduction were defined to be the “worst performance” in that specific category, based on how the necessary improvements were prioritized.

The result of the improvement priority for each city is listed in Table S2 in Appendix. Eighteen out of the 88 underperforming cities need to prioritize improvement in two categories, while the remaining cities need only one. The number of systems with improvement priorities in fixed assets investment, network scale, and electricity was 39, 34, and 40, respectively. For existing water infrastructures, great efforts should be taken to improve and better maintain assets and lowering their depreciation to improve the system performance since the fixed asset investment and network scale could not be easily changed due to the lock-in effect. For the 40 systems with a priority of reducing electricity input, both the technical measures for the operational optimization of the facilities and the incentive policies focused on energy conservation should be implemented.

4. Discussion

4.1. Improvement priorities for unsustainable systems

To screen the most important factors for unsustainable performances, the variables related to resource inputs for the construction and operation of water infrastructures were selected and classified into three categories, which are: fixed assets investment (i.e. $x_{lc}$ and $x_{lw}$), network scale (i.e. $x_{ke}$ and $x_{kp}$), and electricity consumption (i.e. $x_{ec}$ and $x_{ep}$). For each category, water infrastructures were iteratively ranked as the input values were changed proportionately. The first 25 percent of systems that need the most input reduction were defined to be the “worst performance” in that specific category, based on how the necessary improvements were prioritized.

Table 3

<table>
<thead>
<tr>
<th>DMU Variable</th>
<th>Unit</th>
<th>China</th>
<th>Spain</th>
<th>German</th>
<th>U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inputs $x_{lc}$</td>
<td>10^14 $</td>
<td>3.24</td>
<td>0.11</td>
<td>0.32</td>
<td>1.39</td>
</tr>
<tr>
<td>$x_{ke}$</td>
<td>10^3 km</td>
<td>6.76</td>
<td>1.57</td>
<td>6.95</td>
<td>18.92</td>
</tr>
<tr>
<td>$x_{ki}$</td>
<td>10^5 $</td>
<td>5.20</td>
<td>0.92</td>
<td>4.64</td>
<td>12.61</td>
</tr>
<tr>
<td>$x_{kp}$</td>
<td>10^10 $</td>
<td>16.15</td>
<td>0.54</td>
<td>1.62</td>
<td>6.92</td>
</tr>
<tr>
<td>$x_{ec}$</td>
<td>10^2 kWh/a</td>
<td>1.10</td>
<td>3.07</td>
<td>1.14</td>
<td>6.43</td>
</tr>
<tr>
<td>$x_{ep}$</td>
<td>10^3 kWh/a</td>
<td>7.36</td>
<td>36.35</td>
<td>27.37</td>
<td>54.79</td>
</tr>
<tr>
<td>$x_{ep}$</td>
<td>10^6 ton/a</td>
<td>6.92</td>
<td>0.07</td>
<td>0.20</td>
<td>8.20</td>
</tr>
<tr>
<td>Outputs $y_{ucc}$</td>
<td>10^6 m^3/a</td>
<td>255.06</td>
<td>3.45</td>
<td>4.54</td>
<td>64.78</td>
</tr>
<tr>
<td>$y_{uw}$</td>
<td>10^6 m^3/a</td>
<td>3.80</td>
<td>0.41</td>
<td>1.10</td>
<td>5.52</td>
</tr>
<tr>
<td>$y_{wp}$</td>
<td>10^6 ton/a</td>
<td>1.01</td>
<td>0.19</td>
<td>0.60</td>
<td>0.31</td>
</tr>
<tr>
<td>$y_{wp}$</td>
<td>10^6 ton/a</td>
<td>4.91</td>
<td>91.05</td>
<td>1.41</td>
<td>7.13</td>
</tr>
<tr>
<td>$y_{wp}$</td>
<td>10^6 ton/a</td>
<td>9.09</td>
<td>1.58</td>
<td>6.59</td>
<td>3.13</td>
</tr>
</tbody>
</table>
treated) showed higher average scores. Despite the ongoing discussion about the centralized versus distributed systems, the scaling effect seems to be dominant for the best practices in China. This is mostly due to the beneficial scaling effect on capital investment and energy efficiency (Chen et al., 2015; Wang et al., 2014; Adetutu, 2014; Camioto et al., 2016). In addition, larger cities have more resources and public awareness to ensure utility performance and maintenance. In contrast, smaller cities tend to have undesirable performances in their water infrastructure systems. More emphasis should be laid on sufficient investment, timely maintenance, and consciousness raise in medium and small cities.

### 4.2.2. Regional meteorological conditions

Previous studies found that meteorological factors could influence the performance of water system, and temperature and rainfall are considered as the common factors with great impacts on the performance (Toja et al., 2015). Therefore, it is expected that systems in different regions of China would behave differently. The statistics of the sustainability scores of systems in six geographic regions are shown in Table 5 (National Bureau of Statistics, 2015). It is shown that the natural conditions in these six regions have great differences. Because of the humid climate, the average annual rainfall of east and south regions are over 1000 mm, while the value in north region is only 557 mm. Distinction also exists in the air temperature and the largest gap among six regions is 12.5 °C. Various meteorological conditions would have effect on the construction and operation of urban water infrastructure, and thus give rise to distinctive system performances among cities.

K-W test was performed to compare the scores among six regions and the result indicated that the spatial difference in the sustainability level of urban water infrastructures in China was significant (p < 0.10). Sample systems in Southwest China had the highest mean score (0.955) and lowest proportion of under-performed system (35%), while those in Northeast China performed worst on average. Although water infrastructures in East China tended to adopt more advanced technologies due to the higher economic affordability, their sustainability scores were comparatively lower. It was indicated that overall economic condition of a city is not directly tied to its sustainable development. At the provincial level, Jiangsu Province and Anhui Province in East China and Liaoning Province in Northeast China had the relatively higher percentages of the unsustainable systems, which were 80.0%, 77.8%, and 77.8%, respectively. Sichuan Province in Southwest China had a quite high ratio of the sustainable systems, which was 72.7%.

To characterize the regional difference of the system sustainability in China, a multiple nonlinear regression model (p < 0.10) was generated, with all the unsustainable systems of 88 cities as samples, shown as below. The reciprocal of sustainability score (1/q) was treated as the dependent variable, while the average annual air temperature (T, °C) and the annual rainfall (P, mm) of the corresponding city were selected as independent variables. At the significance level of 0.10, both of the explanatory variables were significant. As the representative variables of meteorological conditions, the statistics of average air temperature and annual rainfall of cities with unsustainable systems are given in Table 5 as well.

---

**Table 4**

Statistics of sustainability scores of sample cities with different population size.

<table>
<thead>
<tr>
<th>Population Group</th>
<th>Sample Size</th>
<th>Population (10^6)</th>
<th>Water Supplied (10^6 m³/a)</th>
<th>Wastewater Treated (10^6 m³/a)</th>
<th>Sustainability Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2.5</td>
<td>27</td>
<td>1.72</td>
<td>0.50</td>
<td>87.9</td>
<td>71.0</td>
</tr>
<tr>
<td>2.5–4.0</td>
<td>45</td>
<td>3.21</td>
<td>0.45</td>
<td>97.0</td>
<td>69.4</td>
</tr>
<tr>
<td>4.0–7.0</td>
<td>47</td>
<td>5.32</td>
<td>0.83</td>
<td>197.9</td>
<td>138.3</td>
</tr>
<tr>
<td>7.0–10.0</td>
<td>26</td>
<td>7.91</td>
<td>0.73</td>
<td>363.2</td>
<td>332.2</td>
</tr>
<tr>
<td>&gt;10.0</td>
<td>12</td>
<td>14.48</td>
<td>6.11</td>
<td>1076.5</td>
<td>949.3</td>
</tr>
</tbody>
</table>

---

*Fig. 3.* Change of the number of sustainable DMUs (left) and average sustainability score (right) with inputs or outputs variation by 10%. Results for baseline are marked with the red squares. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).
The estimated coefficient of the temperature was positive, while that of the annual rainfall was negative. It seems that cities with higher temperature and lower rainfall tend to obtain higher systems sustainability, which is in accordance with the better performance in Southwest and Central China. Low temperature may interfere the biological treatment process in WWTP, which is a significant factor for sustainability score as discovered in the previous section.

\[
\frac{1}{g} = -1.433 \times 10^{-5} \times T^3 + 0.138 \times \ln(P) + 0.433
\]  

(3)

Frequent heavy rainfall may create surges of quantity and drastic change in quality of water entering into the treatment facility and increase the difficulty of system design and management.

4.2.3. Environmental constraints

Apart from socio-economic and meteorological factors, environmental constraints such as the quality of drinking water source and the discharge standards of WWTP effluent may also be potential influential factors for the system sustainability.

To investigate the impact of the quality of source water, urban

<table>
<thead>
<tr>
<th>Region</th>
<th>Sample Size</th>
<th>Sustainability Score</th>
<th>Percentage of Underperformed System</th>
<th>City with unsustainable system</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td></td>
<td>Average Annual Air Temperature (°C)</td>
</tr>
<tr>
<td>North</td>
<td>22</td>
<td>0.862</td>
<td>0.152</td>
<td>11.9</td>
</tr>
<tr>
<td>Northeast</td>
<td>15</td>
<td>0.841</td>
<td>0.181</td>
<td>8.6</td>
</tr>
<tr>
<td>East</td>
<td>54</td>
<td>0.862</td>
<td>0.154</td>
<td>15.7</td>
</tr>
<tr>
<td>Central</td>
<td>23</td>
<td>0.917</td>
<td>0.098</td>
<td>16.0</td>
</tr>
<tr>
<td>South</td>
<td>23</td>
<td>0.886</td>
<td>0.137</td>
<td>21.1</td>
</tr>
<tr>
<td>Southwest</td>
<td>20</td>
<td>0.955</td>
<td>0.079</td>
<td>15.9</td>
</tr>
</tbody>
</table>

Table 5: Statistics of sustainability scores and meteorological conditions of different regions.
infrastructure systems were classified into two groups based on the quality compliance of the corresponding city water source to the National Environmental Quality Standards for Surface Water (GB3838-2002) and Quality Standard for Ground Water (GB/T14848-93) in China (NEPA, 2002; NEPA, 1993). There were 23 cities with substantial source water, of which the mean value and the standard deviation of the sustainability score were 0.837 and 0.146, respectively. The other 134 cities with good quality of source water had an average score of 0.892 with the standard deviation of 0.140. By both K-S test and M-W test, it was found that the system sustainability was significantly different between these two groups ($p < 0.10$). The better quality of source water, the higher the sustainability score. This analysis agreed with the importance of source water management on the sustainability of the overall water system (Bai et al., 2006; Li et al., 2011).

When it comes to the impact of the discharge standards of WWTP, hypothesis tests were also conducted to explore whether the standard can be a guide and supervision tool to improve system sustainability. The results showed that no significant difference in the sustainability scores among systems complying with different requirements on WWTP effluents. It may be due to the less determinant role of the output indicators of pollutants removed in this sustainability quantification, as is shown in the result of sensitivity analysis.

4.3. Sustainability-based benchmarking for water infrastructures

Benchmarking is a useful management tool to support the system improvement in a cost-effective way. Hunt and Rogers, 2014; Molinos et al., 2014; Taillard, 1993 The distributions of eight efficiency indicators of the sustainable systems (i.e. score = 1.0) and the unsustainable systems (i.e. score<1.0) were shown in Fig. 4, indicating the detailed system performance gaps. To further demonstrate the utility of benchmarking, K-S tests were conducted. All efficiency indicators were significantly different between two groups ($p < 0.10$), making themselves quite useful for the deep understanding and further benchmarking of the system behavior in terms of the sustainability enhancement.

Based on the 25th percentile of each indicator of sustainable systems as an example, a series of efficiency thresholds could be formulated: IS, 0.33 $$/m^3$; IT, 0.37 $$/m^3$; ES, 0.08 kWh/m$^3$; ET, 0.16 kWh/m$^3$; RC, 0.60 kWh/kgCOD; RS, 0.81 kWh/kgSS; RN, 6.96 kWh/kgTN; and SW, 0.10 kg Sludge/m$^3$. This means that to construct and operate such an ideal example system in a city to supply clean water of $1.0 \times 10^8$ m$^3$ and treat wastewater of $7.5 \times 10^7$ m$^3$ per year (the average ratio of wastewater treated to clean water supplied in China was 0.75 of the case study year) would require the investment of $6.0 \times 10^7$ $/$ for the fixed assets in total and the electricity of $2.0 \times 10^8$ kWh per year, while the system would remove $2.0 \times 10^4$, $1.5 \times 10^4$, and $1.7 \times 10^3$ ton COD, SS, and TN and produce $7.5 \times 10^3$ ton dry sludge per year. This hypothetical system scored 1.0 with the DEA model.

5. Conclusion

In the framework of data envelopment analysis, this paper evaluated the relative sustainability of urban water infrastructure systems in China to serve as a benchmarking. The assessment at the city scale was unique and meaningful, which supports the exploration of more system-dependent information for further improvement. Although the development of urban water infrastructure in China is not extremely unbalanced, disparities exist among cities. The most determinate inputs for the sustainability differentiation include sludge production and electricity consumption, which revealed common bottlenecks in the performance improvement of urban water infrastructure systems in China. More cost effective utilization of energy sources and cautious handling of sludge produced are priorities for the continuous enhancement of system sustainability, which is highlighted in this study and meaningful for other developing cities in assessing the current and potential performance of their urban water infrastructures.

It was confirmed that the system scale, meteorological conditions of the city, and environmental constraints on the system operation regulate the system sustainability. Although some natural constraints cannot be easily changed such as the ambient temperature, the results are applicable for policy makers to identify management weaknesses and facilitate the system design and operation on the basis of a comprehensive investigation. Better designs that fully utilize the scale effect of infrastructure systems, improved water source protection, and more effective operation of the system including energy conservation should be encouraged, as required to improve the system sustainability.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (51308320 & 71473148).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclepro.2017.09.048.

References

China Urban Water Association, Beijing, China, 2015.
