Consistency in efficiency benchmarking: urban water utility regulation with performance improvement

Dickson K. Gidion, Jin Hong, Bojun Hou, Samson S. Shillamkwese, Magdalene Z. A. Adams and Mohammad Khoveyni

ABSTRACT

For more than 20 years, urban water utility (UWU) regulators have been using key performance indicators to monitor water supply services. In many circumstances, the empirical methods used to rank UWU performance are different between regulators of different countries, although the benchmarking basics are the same. The diversity of benchmarking methods limits the sharing of management strategies between countries. Using data envelopment analysis (DEA), this paper presents a consistency analysis of the performance score method used by the Energy and Water Utilities Regulatory Authorities of Tanzania. DEA is appropriate for this purpose because of its demonstrated flexibility in applications with diverse production environments. Thus, this paper proposes methods for benchmarking strategies to assess UWU efficiency- and management-specific goals within and between countries. The assessment shows that network DEA (NDEA) outperforms empirical methods; regulators using the proposed NDEA technique will benchmark UWU efficiency under a yardstick competition regime and, at the same time, identify the most efficient and weak UWUs using pure variable values. This innovation monitors UWU performance progress and promotes sharing of quality management strategies between UWUs and countries.

Key words | efficiency assessment and performance improvement, network DEA, performance score, utility regulation

Dickson K. Gidion (corresponding author)
Jin Hong
Bojun Hou
Samson S. Shillamkwese
School of Management,
University of Science and Technology of China,
Hefei, 230026 Anhui,
China
E-mail: *dgidion@gmail.com*

Dickson K. Gidion
Ministry of Water,
NBC Mazengo Branch,
Kuu Street, P.O. Box 456, Dodoma,
Tanzania

Magdalene Z. A. Adams
School of Public Affairs,
University of Science and Technology of China,
Hefei, 230026 Anhui,
China

Mohammad Khoveyni
Department of Applied Mathematics,
Yadegar-e-Imam Khomeini (RAH) Shahre Rey Branch,
Islamic Azad University,
Tehran,
Iran

INTRODUCTION

The International Water Association (IWA) has specified performance indicators (PIs) used to assess, monitor, and improve the performance of urban water utilities (UWUs) (Pinto *et al.* 2017). Researchers have developed many methods that utilise a limited number of PIs to analyse UWU efficiency. Advances in research, management, and computer technology have resulted in many water supply regulators using quantifiable key performance indicators (KPIs), well-developed benchmarking methodologies, and tools to benchmark the efficiency of a regulated UWU. Recently, studies have applied advanced methods to measure UWU performance and proposed adjustments to some management incentives. For example, Singh *et al.* (2014) presented a comparative analysis involving the use of PIs and data envelopment analysis (DEA) to benchmark UWUs; the authors concluded that efficiency benchmarking using PIs is 'effective'. However, the current study found that the consistency analysis followed by Singh *et al.* (2014) used different sets of variables in their PI-based efficiency assessment and DEA efficiency assessment. Thus, we view the PIs and standard DEA benchmarked efficiency results as insufficient for consistency investigation, requiring further details. Moreover, assessing UWU efficiency- and management-specific behaviour using DEA has many advantages over PI-based analysis, thus the

doi: 10.2166/ws.2019.072

المنارة للاستشارات
www.manaraa.com
Downloaded from https://iwaponline.com/ws/article-pdf/19/7/1955/607776/ws019071955.pdf
by PROQUEST user

PI-based approach cannot be more effective in UWU management than standard DEA (Thanassoulis & Silva 2018).

UWU efficiency benchmarking has many advantages including identifying management gaps and providing service transparency. The efficiency assessment industry uses standard DEA as an extended KPI benchmarking approach (Thanassoulis & Silva 2018). Both standard DEA and KPI methods use absolute numerical values to produce a UWU efficiency score and the differences between the scores consider inputs, analysis, and output efficiencies. Empirical methods have been developed in many forms. There are empirical methods that utilise absolute values as inputs to rank a UWU, and there are empirical methods that utilise KPIs as inputs to rank a UWU. Likewise, there are many forms of DEA. For example, the standard DEA methods developed by Charnes *et al.* (1978) and Banker *et al.* (1984) utilise absolute numerical values to produce a UWU efficiency score and an alternative DEA developed by Emrouznejad & Amin (2009) utilises ratio variables (probably KPIs) to generate UWU efficiency scores (Gidion *et al.* 2019). According to Emrouznejad & Amin (2009), the presence of a ratio variable in an efficiency assessment using DEA developed by Charnes *et al.* (1978) or Banker *et al.* (1984) provides an incorrect efficiency result.

A study of Singh *et al.* (2014) used a mixture of single and ratio variables to generate UWU efficiency values using PI and DEA methods. In addition, between the two methods, the variables selected to analyse UWU efficiency were not the same. The study of Emrouznejad & Yang (2017) concludes that the DEA approach has broad global applicability in efficiency benchmarking, due to the flexibility that can be developed in many forms depending on the working environment of the decision-making unit (DMU) (Li & Reeves 1999). Thus, based on flexibility, the current study compares efficiency results generated by the performance score (PS), alternative DEA model and network DEA (NDEA) model methods that utilise KPIs as input and output variables and the results are used to extend efficiency assessment using various DEA applications in the water supply industry. In the water supply industry, it is well known that UWUs operate under variable returns to scale (VRS) rather than constant returns to scale (CRS) and the production direction aims to increase outputs while minimising inputs (Brettenny & Sharp 2016). Thus, the VRS input-oriented production was used in the current study utilising the same data published by EWURA (2014) and (2016). Here, PS stands for the majority of empirical methods and DEA stands for non-parametric methods.

Moreover, this study prioritises methods' robustness and advantages, or disadvantages, in managing UWU services. Real-life regulators concentrate on improving available benchmarking methods to improve UWU performances and this is also the goal of the current research. Recently, regulators have used performance targets to manage UWU performance and UWU efficiency is presented as a public, yardstick competition (YC) regime. Considering the benefits of using DEA to improve UWU efficiency and management techniques, this study seeks to extend the possibilities of using DEA and other DEA applications as a UWU regulation tool. The investigation shows that NDEA not only outperforms PS methods in efficiency assessment, but the proposed method also identifies a UWU with the best management strategies. In addition, network DEA allows the sharing of quality management techniques between UWUs and the country's regulators, thereby improving UWU efficiency and management.

EFFICIENCY ANALYSIS USING PERFORMANCE SCORE AND DEA METHODS

Generation of UWU efficiency using a performance score method

This section details the efficiency analysis and results of the PS method, where this study uses element $PS_{i,x}$ to denote the performance score of a UWU_x in terms of a KPI_i. The $PS_{i,x}$ consists of four components, namely: Score Based on Best Performer ($\text{SBP}_{i,x}$), Score Based on Attaining Performance Target ($\text{SPT}_{i,x}$), Score Based on Confidence Grading ($\text{SCG}_{i,x}$), and Score Based on Attaining Service Level Benchmark ($\text{SSLB}_{i,x}$). The components were analysed separately, and later summed up and scaled by a KPI weight ($W_{i,x}$) to generate a weighted $PS_{i,x}$ (see Equation (3)).

Score based on the best performer ($SBP_{i,x}$)

The $\text{SBP}_{i,x}$ measure is calculated based on attaining an accepted service level, and it is considered to be of high

www.manaraa.com
Downloaded from https://iwaponline.com/ws/article-pdf/19/7/1955/607776/ws019071955.pdf
by PROQUEST user

impact in the $PS_{i,x}$ analysis approach. The maximum efficiency score for any UWU in any KPI is 1, which can also be expressed as 100% or 100 points (i.e., percentage points). The $SBP_{i,x}$ is considered to contribute 70 points of the $PS_{i,x}$ total and the other 30 points are equally distributed among $SPT_{i,x}$, $SCG_{i,x}$, and $SSLB_{i,x}$. In calculating the $SBP_{i,x}$, an outperforming UWU attains the highest performance score of 70 points, a medium performer scores 50 points, and an underperforming UWU scores 0 points. A medium performer is identified as one whose performance value is equal to the average value of an assessed KPI. Equation (1) specifies how to calculate the $SBP_{i,x}$ based on a KPI average. A detailed analysis is provided in the linked material, Appendix B, available with the online version of this paper. Element x_i in Equation (1) denotes a performance attained by UWU_x in a KPI_i. Thus, based on the average value of a KPI, the $SBP_{i,x}$ is calculated using the following formula:

$$\text{If } x_i \leq x_{\text{Average}},\ \text{then } SBP_{i,x} = 50* \left[\frac{x_i - x_{\min}}{x_{\text{Average}} - x_{\min}}\right] \text{ else if}$$
$$x_i > x_{\text{Average}},\ \text{then } SBP_{i,x} = 50 + 20* \left[\frac{x_i - x_{\text{Average}}}{x_{\text{Max}} - x_{\text{Average}}}\right] \tag{1}$$

Score based on attaining the performance target ($SPT_{i,x}$)

Using $SPT_{i,x}$ the ranking score is calculated based on attaining performance targets, and since the performance target can change from one year to another, $SPT_{i,x}$ differs from one year to another. During analysis, a UWU with a current performance (P_n) at least reaching the performance target (Pt_n) scores 10 points, while the intermediate performers score a linear interpolation in the range (0,10). Any UWU underperforming its previous year's performance P_{n-1} in an assessed KPI scores 0 (see Equation (2), analysis in linked material). The $SPT_{i,x}$ assessment analysis uses the previous year's performance, the performance target, and current year's performance to develop an $SPT_{i,x}$ combined with a PS for UWU efficiency benchmarking. This approach is currently beyond the standard DEA and the PI approach presented by Singh et al. (2014) because both of these previous methods benchmark DMU efficiency without incorporating performance targets. This study recommends assessing UWU efficiency by incorporating performance targets determined using the DEA.

$$\textit{If } P_n \geq Pt_n,\ \textit{then } SPT = 10\ \text{ else,}$$
$$\textit{If } P_{n-1} < P_n < Pt_n\ \textit{then } SPT = \left[\frac{P_n - P_{n-1}}{Pt_n - P_{n-1}}\right]*10 \text{ else,}$$
$$\textit{If } P_n < P_{n-1}\ \textit{then } SPT = 0 \tag{2}$$

Score based on confidence grading ($SCG_{i,x}$)

The $SCG_{i,x}$ measure is assessed based on KPI reliability associated with accuracy ranges given in the EWURA performance benchmarking guideline and the IWA Manual of Best Practice. Here, the term 'accuracy' for a KPI refers to how close the figures for that KPI are to the accepted service level, which means the level that a UWU's customers will accept without complaint. EWURA (2014) and Alegre et al. (2016) grouped the accuracy into reliability ranges of 0–5%, 5–20%, 20–50%, and >50%, labelling these as A, B, C, and D respectively. Band A corresponds to the UWU being highly reliable in providing the accepted service level and D corresponds to high unreliability. Due to the poor performances of the studied utilities, this study uses a target accuracy range of 5–20%. Thus, it is considered that the assessed UWU must meet a minimum of band B reliability (±20% of the accepted level) in delivering acceptable services. Therefore, for the assessed UWU, each KPI with a performance value falling within ±20% of its accepted service level scores 10 points while those outside ±20% of its accepted service level score 0. For example, consider using ±20% to analyse $SCG_{i,x}$ for KPI_1 with 20% or KPI_{10} with $\geq 98\%$ accepted service levels. A UWU with accepted service level $\leq 20\% + 20\%*0.2$ in KPI_1 or $\geq 98\% - 98\%*0.2$ in KPI_{10} will score 10 points and score 0 otherwise (detailed analysis in the linked material).

Score based on attaining the service level benchmark ($SSLB_{i,x}$)

For $SSLB_{i,x}$, a utility can only score 0 or 10 points within a KPI; no CG, no intermediate value, and no linear interpolation is allowed for a UWU failing to meet or outperform the acceptable service level. Therefore, a utility that attains

المنارة للاستشارات
www.manaraa.com
Downloaded from https://iwaponline.com/ws/article-pdf/19/7/1955/607776/ws019071955.pdf
by PROQUEST user

or outperforms the acceptable service level for a KPI scores 10 points and 0 points are given to a utility not attaining an acceptable service level for that KPI (EWURA 2014).

UWU efficiency analysis using a PS

Using Equation (3), EWURA (2014) analysed $PS_{i,x}$ as 70% of the weighted sum of $\mathrm{SBP}_{i,x}$, $\mathrm{SPT}_{i,x}$, $\mathrm{SCG}_{i,x}$, and $\mathrm{SSLB}_{i,x}$ whereas a 30% weight was given to monthly timely reporting. This study does not include monthly timely reporting in the analysis because it is not among the variables used for the consistency analysis. Therefore, the study estimates $PS_{i,x}$ as 100% of the weighted sum of $\mathrm{SBP}_{i,x}$, $\mathrm{SPT}_{i,x}$, $\mathrm{SCG}_{i,x}$ and $\mathrm{SSLB}_{i,x}$ and PS as $\sum_{x=1}^{n} PS_{i,x}$ where $n = 10$ and

$$PS_{i,x} = W_{i,x}(SBP_{i,x} + SPT_{i,x} + SCG_{i,x} + SSLB_{i,x}) \tag{3}$$

Moreover, the efficiency scores were generated using Equation (4) as in Table A1 in Appendix A (available with the online version of this paper), an equation that is not used by EWURA for UWU efficiency benchmarking.

$$\text{Efficiency score} = \frac{\left(\sum_{i=1}^{n} PS_i\right)_x}{\left(\sum_{i=1}^{n} PS_i\right)_{\max}}; \quad x = 1, \ldots, v; \quad i = 1, \ldots, n \tag{4}$$

Generation of UWU efficiency using DEA

DEA is a non-parametric method used to measure efficiency by enveloping production sets. The method was introduced by Charnes *et al.* (1978), and it is popular for benchmarking the efficiency of DMUs. One of the reasons for its popularity is its ability to handle situations involving multiple inputs and multiple outputs with multiple DMUs, situations that were difficult or impossible to analyse using other benchmarking methods. DEA produces effective results without any assumptions about functional forms that relate inputs to outputs (See 2015). DEA also does not need subjective allocations of weights to variables; the weights are allocated endogenously. However, DMU efficiency results are definitely influenced by the ratio of the number of variables to the number of DMUs and DEA reliability requires the number of DMUs to be at least three times the total number of inputs and outputs. The number of efficient DMUs will increase to an unreliable figure when using a small number of DMUs compared with the number of variables used (Toloo & Tichý 2015).

Furthermore, DEA has been developed into an envelopment form and a multiplier form. The envelopment form is also known as the 'primal form' and is the DEA form that develops DMU efficiency based on production possibility sets. The multiplier form, also known as the 'dual form', is the DEA form that generates DMU efficiency through the use of relative weights. In the normal assessment of DMU efficiency, envelopment forms yield the same results as multiplier forms (Boloori *et al.* 2016). This study uses the alternative DEA model (5) shown below, which was extended by Emrouznejad & Amin (2009) from the standard VRS DEA model, to overcome the effect of input-ratios and output-ratios and benchmark UWU efficiency while avoiding the false convexity caused by input/output-ratios. Thus, consider a set of j UWUs, each consuming a certain amount of i input-ratios to generate r output-ratios under the condition that $n \geq 3(m+s)$. Let $x_j = (x_{ij} \ldots\ldots, x_{mj})^T$ and $y_j = (y_{rj} \ldots\ldots, y_{mj})^T$ respectively represent the vectors of m consumed input-ratios and s output-ratios, $\bar{x}_{pj}$ and $\underline{x}_{pj}$ represent a numerator and denominator of the p^{th} input (x_{pj}), and $\bar{y}_{kj}$ and $\underline{y}_{kj}$ represent a numerator and denominator of the k^{th} output (y_{kj}) for UWU_j. Thus,

$$E_0 = \min \theta_0$$

s.t.

$$\sum_{j=1}^{z} \lambda_j x_{ij} - \theta_0 x_{i0} \leq 0; \quad i = 1, \ldots, m; \quad i \neq p$$

$$\sum_{j=1}^{z} \lambda_j y_{rj} \geq y_{r0}; \quad r = 1, \ldots, s; \quad r \neq k$$

$$\sum_{j=1}^{z} \lambda_j \bar{x}_{pj} - \theta_0 x_{p0} \sum_{j=1}^{z} \lambda_j \underline{x}_{pj} \leq 0; \qquad i = p$$

$$\sum_{j=1}^{z} \lambda_j \bar{y}_{kj} - y_{k0} \sum_{j=1}^{z} \lambda_j \underline{y}_{kj} \geq 0; \qquad r = k$$

$$x_{p0} = \frac{\bar{x}_{pj}}{\underline{x}_{pj}}, \qquad y_{k0} = \frac{\bar{y}_{kj}}{\underline{y}_{kj}}$$

$$\sum_{j=1}^{z} \lambda_j = 1; \quad \lambda_j \geq 0; \quad j = 1, \ldots, z \tag{5}$$

Downloaded from https://iwaponline.com/ws/article-pdf/19/7/1955/607776/ws019071955.pdf
by PROQUEST user
المنارة للاستشارات
www.manaraa.com

Here, the x_{i0} compose the known input-ratio vector (data) of the target UWU_0, the y_{r0} compose the known output-ratio vector (data) of the target UWU_0, λ is a vector describing the percentages of other producers used to construct the virtual producer, and θ is the producer's efficiency score. Considering a UWU scoring zero values in some KPIs due to underperforming the previous year's performance as analysed in the PS method, the standard DEA analysis technique does not run the analysis of any UWU with zero value in any variable and deleting a UWU or a variable is the standard approach to run the analysis when zero values are encountered (Gidion *et al.* 2019). This difficult analysis environment can be solved using a network DEA model. Thus, we recommend that after generating information for UWUs that outperformed the previous year's performance, users should arrange the UWUs in groups (demonstrated in the linked material, Appendix B) that allow the DEA analysis technique to run. The first group should contain only UWUs which outperformed the previous year's performance score, the second group should contain the UWUs analysed in the first group plus any UWU observed to have a minimum number of zeros. This means that the second and subsequent groups should have redundant variables observed to contain zeros. Refer to G_1, G_2,…G_6 examples in the linked material where all G_1 UWUs outperformed their previous year's performance in all KPIs, whereas for UWUs in G_2 (1) this study treats KPI_3 and UWU_{19}, UWU_{21}, UWU_{24}, UWU_{26}, UWU_{28}, UWU_{29}, UWU_{32}, UWU_{34} and UWU_{35} as redundant. G_3 has one more variable redundant as compared with G_2, which means it must be analysed separately and G_2 (1) has a different variable with zero score compared with G_2 (2) which means they are analysed separately too. Note that there will be UWU efficiency shifts in some groups as the ratio of number of input/output variables to number of UWUs increases or decreases.

The average of the efficiencies of the groups (efficiencies generated using model (5)) in separate groups forms a system efficiency and this will absolutely benchmark a UWU in a YC regime as does the PS method. Standard DEA benchmarks UWU efficiency through the use of a multi-input and multi-output production technology where the internal structure of production UWUs is ignored or treated as a black box to model inputs and outputs through several interconnected subprocesses or divisions (Boloori *et al.* 2016). Network DEA models allow users to look into a black box and format a black box technology to benchmark DMU efficiency (Gidion *et al.* 2019). Thus, averaging groups' efficiencies is beyond the black box technique and is fit for a network DEA model, to be achieved as:

$$E_c^{VRS} = \frac{1}{q}\sum_{n=1}^{q} E_0^n$$

$$n = 1, \ldots, q \qquad (6)$$

where E_0 is as in model (5).

AN ILLUSTRATIVE APPLICATION

Data for UWU efficiency assessment

The main data content and information used in this study were retrieved from EWURA (2014, 2016). EWURA (2016) was used for the UWU performance information and EWURA (2014) was used as the source of the guidelines and equations used to draw the consistency conclusions about the best efficiency benchmarking technology. Table 1 lists the variables: KPI_1 to KPI_5 were used as input variables, and KPI_6 to KPI_{10} were employed as output variables to generate UWU efficiency using the alternative standard DEA. The sum of weights presented in Table 1 is 100%. Some KPIs are given high weights indicating that these KPIs have high priority. It is envisaged that a UWU scoring well on a high-weighted KPI will score sufficiently in the low-weighted KPIs as well. The weights were distributed based on promoting the efficiency of quality of services; this distribution makes the empirical analysis in this study different from the Singh *et al.* (2014) analysis. The two approaches also differ in their subprocess analysis because the PS method uses performance-target data to control the efficiency analysis of a UWU based on 2 years of performance data, while Singh *et al.* (2014) employ only a single year of data and do not consider performance targets. Also, the PS method consists of the four components used to develop a utility ranking score, while Singh *et al.* (2014) relied on direct KPI weight scaling (single-measured gap analysis).

www.manaraa.com
المنارة للاستشارات
Downloaded from https://iwaponline.com/ws/article-pdf/19/7/1955/607776/ws019071955.pdf
by PROQUEST user

Table 1 | Proposed service levels and weight scores

KPI no.	Variable name	PI	Acceptable service level[a]	Weight score
KPI_1	Non-revenue water (NRW)	%	≤20	15%
KPI_2	Working ratio	–	≤0.67	10%
KPI_3	Operating ratio	–	≤0.1	10%
KPI_4	Personnel expenditure	%	≤30	10%
KPI_5	Staff/1000 connections	FTE/1000 Conn.	≤5	5%
KPI_6	Proportion of population served with water	%	100	5%
KPI_7	Average hours of supply	hrs	24	5%
KPI_8	Metering ratio	%	100	10%
KPI_9	Revenue collection efficiency	%	≥95	15%
KPI_{10}	Water quality compliance	%	≥98	15%

Sources: EWURA (2014) and Alegre *et al.* (2016).
[a]Accepted service level is a default/standard value used to indicate a constant/minimum acceptable performance of a UWU in the delivery of viable water services (see Hastak *et al.* (2017) and González-Gómez *et al.* (2011)).

Consistency validation under efficiency assessment

Efficiency benchmarking provides a tool for a DMU to improve performance and management practices through learning from competitors (Kyrö 2004). The method used for efficiency benchmarking should be descriptive enough to evaluate the DMU based on the nature of the underlying industry while minimising application complexities. The benchmarking processes and results generated should reflect changes in the firms and thereby direct underperformers towards the best practices. The three key methods used in this study utilise the same variables to benchmark UWU efficiency, and the main differences are in the aggregation of variables, benchmarking techniques, and efficiency results. The PS approach uses a gap analysis approach with assumed weights, incorporating benchmarked acceptable service levels and assumed scores based on the regulator's priorities to generate the utility's rank score.

This study's consistency analysis approach hinges on the fact that the PS method has a component to assess UWU performance using the previous year, current year, and performance targets at the same time. The method includes targets in the analysis that have not been incorporated previously into the DEA method and benchmarks UWU efficiency in a competitive scenario which is difficult using standard DEA. Overcoming this difficulty the proposed network DEA model benchmarks UWU efficiency including the advantages of the PS method. Delivering viable water services requires regular monitoring of UWU efficiency goals and management goals, through performance targets and utility benchmarking in a YC, and the PS method fails in this area. Figure 1 presents the underlying UWU efficiency benchmarking concept using the PS method and the standard DEA.

The differences in efficiency results presented in Figure 1 show the quality-specific nature of the methods (see the data and efficiency results in the linked material, Appendix B, available with the online version of this paper). Here, the PS method outperforms standard DEA in the task of benchmarking a UWU in a YC regime (see efficiency results in Table A1 in Appendix A, available online). However, the alternative DEA model outperforms the PS method when considering UWU efficiency and management-specific goals. The DEA outperforms the PS and other empirical methods because of its unique ability to utilise multi-input and multi-output variables without assumptions and, at the same time, analyse a UWU to produce a global optimal solution. By 'global optimal solution', this study means the method does not rely on many solutions to generate DMU efficiency (Gidion *et al.* 2019). The global optimality is very important when it comes to sharing management techniques and experience between utilities and regulators in various countries. Figure 2 presents the comparison between efficiency results generated by a network DEA model and the PS method, when the network DEA model benchmarks a UWU like the PS method. Note that the UWU efficiency scores and ranking differ between the two methods.

المنارة للاستشارات
www.manaraa.com
Downloaded from https://iwaponline.com/ws/article-pdf/19/7/1955/607776/ws019071955.pdf
by PROQUEST user

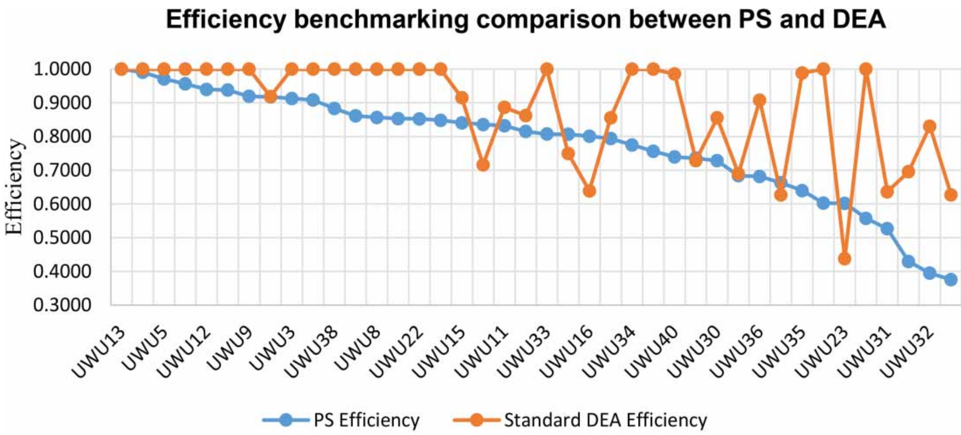

Figure 1 | Comparison of efficiency results generated using the PS method and standard DEA model.

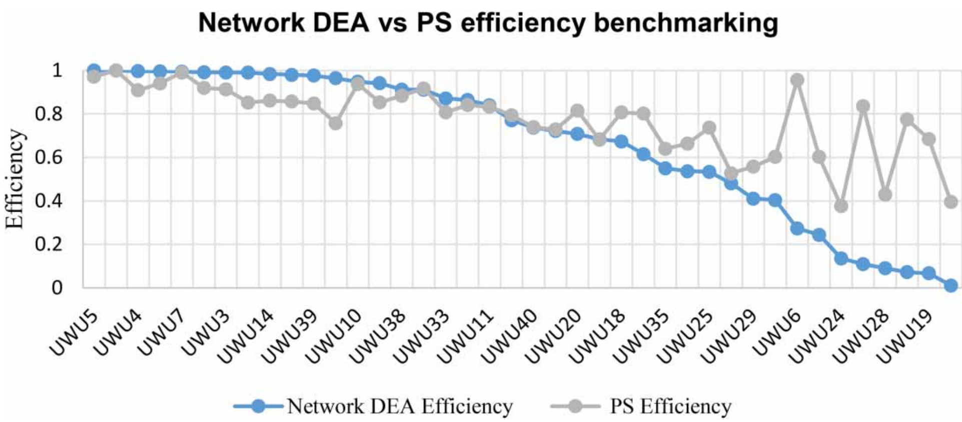

Figure 2 | Comparison of efficiency results generated using the proposed network DEA and the PS method.

Figure 2 presents a comparison of efficiencies generated between the proposed network DEA model and the PS method. In Figure 1, the PS method analyses UWUs under the YC regime and the curve is smooth like network DEA model curve in Figure 2. Figure 2 shows the PS curve is not as smooth as in Figure 1, which is due to the differences in quality and benchmarking technologies employed by the two methods. The PS method uses many assumptions to generate a UWU efficiency score while the network DEA model does not use any assumptions. The results in Figure 2 show that the use of assumptions can promote a poorly performing UWU to the best position in the ranking. For example, Tables A1 and A2 in Appendix A present DEA and PS efficiency results used to generate curves in Figures 1 and 2, and many UWUs ranked in high positions using the PS method are ranked in low positions using the network DEA method. This indicates that when regulators employ empirical methods to manage UWU performance, a weakly performing UWU can adopt management strategies from another weak performing UWU rather than a strong UWU rather than a strong UWU, thus making many UWUs underperform in the next performance assessment. The efficiency consistency analysis proves that empirical methods that employ assumptions are not effective for assessing UWU efficiency- and management-specific performance, as concluded by Singh *et al.* (2014). To promote a UWU working on targets, a network DEA model benchmarks any UWU that underperforms in a variable in a low-rank position (e.g. UWU_6 underperformed KPI_3 and was ranked in 33 position). The action of suspending a UWU below UWUs outperformed their previous year's performance help a UWU simulate the best management strategies and deliver viable water services.

المنارة للاستشارات
www.manaraa.com
Downloaded from https://iwaponline.com/ws/article-pdf/19/7/1955/607776/ws019071955.pdf
by PROQUEST user

PS consistency with DEA integrated applications

The DEA model has been extended into different areas and modified by integrating various mathematical models to simplify its application for analysing DMU efficiency and providing recommendations for management improvement. The current analysis demonstrates an efficiency assessment using DEA and the PS methods. Table 2 presents the theoretical consistency analysis between the PS and DEA together with modified applications. In similar environments, this study aims to emphasise the importance of incorporating performance targets into DEA, and DEA integrated applications that benchmark UWU under the YC. Such considerations result in management improvement. A detailed analysis accompanies the summary in Table 2 and specifies requirements for extending the methods into UWU efficiency benchmarking while considering service improvement.

This study encourages the inclusion of performance targets in the analysis while benchmarking UWUs. The efficiency assessment technique of the PS method can be extended into any form of DEA discussed above and yet allows DEA benchmarks for UWUs under the YC regime, just as the PS method does. For example, the PS method assigns a high-weight score to a UWU outperforming a variable target, an intermediate-weight score for a UWU underperforming a variable target, and zero weight for a UWU underperforming its previous year's performance for a variable (see linked material for detailed analysis, Appendix B). The same methodology can be followed in DEA using the endogenous weights; a UWU that outperformed a target should be analysed using high-weight settings, and those UWUs underperforming targets should be analysed using low-weight settings.

CONCLUSIONS

This paper presents some recommendations for extending DEA to apply it to UWU efficiency assessment and benchmarking. The benchmarking recommendations were developed from examining literature studies on efficiency analysis for water supply utilities, and a consistency analysis between the PS and DEA methods. Recently, UWU performance benchmarking has become an essential tool to promote efficient water supply services. This study found that most of the UWU regulators use empirical methods to benchmark a UWU performance which does not allow full

Table 2 | Validation of efficiency benchmarking methods

		Working environment compatibility		Best practice benchmark	
Method	**Efficiency stability over time**	**Ability to handle variables without limitation**	**Variable descriptions**	**Cross-utility**	**Internationally**
PS	NO	YES	YES[2]	YES	NO
DEA	YES[1]	NO	YES	YES[3]	YES[4]

Notes:

YES[1]: The Malmquist Productivity Index (MPI) developed in DEA allows DEA to benchmark the relative efficiency of utilities over time by employing a base period technology developed in MPI (Chen & Iqbal Ali 2004) using absolute variables. The approach is important to evaluate the production change in a DMU. However, the MPI technical change technology developed in DEA cannot combine the previous year's performance, performance targets, and current year's performance to form a single variable which can be used to analyse the technical and frontier change in the UWU over time. The PS has an SPT component that combines the previous year's performance, current year's performance, and performance target variables to form a single ranked or scaled variable value used to generate a utility performance efficiency using KPIs. This component suggests an improvement of MPI-DEA to incorporate performance targets in the analysis, to allow UWU benchmarking over time using KPIs while considering water service improvement.

YES[2]: PS has no proper descriptions of production variables before, during, and after UWU benchmarking as in DEA. DEA benchmarking utilises multiple defined input and output variables (Thanassoulis & Silva 2018). However, PS compares two periods of data under the control of performance targets and compares the variable values of a UWU with those of other UWUs that are outperforming the previous year's performance or performance targets.

YES[3]: Using standard DEA, 50% of UWUs were found to be efficient in comparison with 3% benchmarked using PS and NDEA models. PS and NDEA models outperformed in a cross-utility benchmarking following the ranking of UWU under a YC regime. Thus, the chance of a UWU to underperform in the next assessment is reduced when compared with standard DEA efficiency benchmarking. Using standard DEA, many UWUs are benchmarked as efficient, and some of the UWUs designated by PS and network DEA as inefficient were benchmarked efficient. This implies that underperforming UWUs might learn incorrect management strategies from a utility wrongly identified as efficient and continue to underperform (Gidion *et al.* 2019). However, the benchmarking technology of the NDEA is more efficient compared with the PS method.

YES[4]: Standard DEA in its conventional model cannot produce the efficiency results which are 100% accepted by UWU stakeholders because the utilities operate in different environments with different technologies. The use of DEA with the meta-frontier technique allows DEA to be employed internationally for UWU efficiency analysis (Molinos-Senante & Sala-Garrido 2016). Although the meta-frontier analysis developed in DEA is able to benchmark UWUs internationally, in the current model, it cannot benchmark UWU efficiency using KPIs and at the same time incorporate performance targets. This inability necessitates improvement of the meta-frontier model.

المنارة للاستشارات
www.manaraa.com
Downloaded from https://iwaponline.com/ws/article-pdf/19/7/1955/607776/ws019071955.pdf
by PROQUEST user

efficiency comparison between countries and the employed methods use assumptions that rank a weakly performing UWU in a high position, thus leading to sharing of poor management strategies, which in turn contributes to the inefficiency of UWU services. Through a consistency analysis, this study shows that network DEA identifies UWUs with the best management strategies compared with empirical methods, because the proposed network DEA techniques benchmark UWU efficiency without assumptions while considering the nature of the water supply industry and remove the limitation of sharing management strategies between countries to improve urban water delivery. Thus, performing UWU analysis based on variables outperforming the previous year's targets lets utilities access information pointing out the UWUs with the best management strategies, and the UWUs will learn together to improve water supply services.

This study recommends that future researchers should develop an efficiency assessment approach that uses DEA endogenous weights in place of PS assumed weights to promote UWUs that outperform targets into a good position as compared with UWUs that underperform those targets. The efficiency should be analysed using a network DEA model when a few UWUs are observed to underperfom their previous year's scores; otherwise a standard DEA is acceptable. The network DEA model will allow benchmarking even of UWUs scoring zero values during analysis and expand the scope of UWU efficiency benchmarking in a YC (Gidion *et al.* (2019) detailed network DEA analysis). Table 2 of this study provides a theoretical analysis summary illustrating the differences between the PS and DEA model results. We expect that the proposed efficiency analysis approach will overcome the negative aspects of using DEA for UWU efficiency analysis. Moreover, UWU efficiency will improve because managers of a UWU which fails in some aspect will learn new ideas used by other UWUs which succeed in that aspect, allowing the weaker UWU to improve its performance. Because this study analysis relies on KPIs used by the EWURA to rank a UWU, we further propose a future study on a few KPIs that improve utility performance in the dimension of economic, environment and social trust that will be used by regulators between countries to report UWUs' efficiency.

ACKNOWLEDGEMENTS

The authors would like to extend their thanks to the four reviewers and the journal editors for the manuscript improvement achieved through the revisions process. The authors acknowledge the role of the Energy and Water Utilities Regulatory Authority of Tanzania, which provided data and guidelines for the empirical method analysis. The authors also acknowledge the work of Professor Ali Emrouznejad and Professor Gholam R. Amin, as this research extends the application of the standard DEA model to the water industry based on their related work. However, the interpretations and conclusions presented in this research work originate from the authors.

REFERENCES

Alegre, H., Baptista, J. M., Cabrera, E., Cubillo, F., Duarte, P., Hirner, W., Merkel, W. & Parena, R. 2016 *Performance Indicators for Water Supply Services*. IWA Publishing, London, UK. https://doi.org/10.2166/9781780406336.

Banker, R. D., Charnes, A. & Cooper, W. W. 1984 Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* **30** (9), 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078.

Boloori, F., Afsharian, M. & Pourmahmoud, J. 2016 Equivalent multiplier and envelopment DEA models for measuring efficiency under general network structures. *Measurement* **80** (8), 259–269. http://dx.doi.org/10.1016/j.measurement.2015.11.012.

Brettenny, W. & Sharp, G. 2016 Efficiency evaluation of urban and rural municipal water service authorities in South Africa: a data envelopment analysis approach. *Water SA* **42** (1), 11–19. http://dx.doi.org/10.4314/wsa.v42i1.02.

Charnes, A., Cooper, W. W. & Rhodes, E. 1978 Measuring the efficiency of decision making units. *European Journal of Operational Research* **2** (6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8.

Chen, Y. & Iqbal Ali, A. 2004 DEA Malmquist productivity measure: new insights with an application to computer industry. *European Journal of Operational Research* **159** (1), 239–249. https://doi.org/10.1016/S0377-2217(03)00406-5.

Emrouznejad, A. & Amin, G. R. 2009 DEA models for ratio data: convexity consideration. *Applied Mathematical Modelling* **33** (1), 486–498. https://doi.org/10.1016/j.apm.2007.11.018.

Emrouznejad, A. & Yang, G.-l. 2017 A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences* **61**, 4–8. https://doi.org/10.1016/j.seps.2017.01.008.

المنارة للاستشارات
www.manaraa.com
Downloaded from https://iwaponline.com/ws/article-pdf/19/7/1955/607776/ws019071955.pdf
by PROQUEST user

EWURA 2014 *Performance Benchmarking Guidelines for Water Supply and Sanitation Authorities*. http://www.ewura.go.tz/wp-content/uploads/2015/03/Performance-Benchmarking-Guidelines-for-WSSAs-2014.pdf (accessed 9 April 2018).

EWURA 2016 *Water Utilities Performance Review Report for the Year 2015/2016: Regional and National Project Water Utilities*. EWURA, Dar es Salaam, Tanzania. http://www.ewura.go.tz/wp-content/uploads/2015/04/EWURA-REGIONAL-WATER-REPORT-2015-16.pdf (accessed 13 April 2018).

Gidion, D. K., Hong, J., Adams, M. Z. A. & Khoveyni, M. 2019 Network DEA models for assessing urban water utility efficiency. *Utilities Policy* **57**, 48–58. https://doi.org/10.1016/j.jup.2019.02.001.

González-Gómez, F., García-Rubio, M. A. & Guardiola, J. 2011 Why is non-revenue water so high in so many cities? *International Journal of Water Resources Development* **27** (2), 345–360. https://doi.org/10.1080/07900627.2010.548317.

Hastak, S., Labhasetwar, P., Kundley, P. & Gupta, R. 2017 Changing from intermittent to continuous water supply and its influence on service level benchmarks: a case study in the demonstration zone of Nagpur, India. *Urban Water Journal* **14** (7), 768–772. https://doi.org/10.1080/1573062X.2016.1240808.

Kyrö, P. 2004 Benchmarking as an action research process. *Benchmarking: An International Journal* **11** (1), 52–73. https://doi.org/10.1108/14635770410520302.

Li, X.-B. & Reeves, G. R. 1999 A multiple criteria approach to data envelopment analysis. *European Journal of Operational Research* **115** (3), 507–517. https://doi.org/10.1016/S0377-2217(98)00130-1.

Molinos-Senante, M. & Sala-Garrido, R. 2016 Cross-national comparison of efficiency for water utilities: a metafrontier approach. *Clean Technologies and Environmental Policy* **18** (5), 1611–1619. https://doi.org/10.1007/s10098-016-1133-z.

Pinto, F. S., Simões, P. & Marques, R. C. 2017 Raising the bar: the role of governance in performance assessments. *Utilities Policy* **49**, 38–47. https://doi.org/10.1016/j.jup.2017.09.001.

See, K. F. 2015 Exploring and analysing sources of technical efficiency in water supply services: some evidence from Southeast Asian public water utilities. *Water Resources and Economics* **9**, 23–44. https://doi.org/10.1016/j.wre.2014.11.002.

Singh, M., Mittal, A. K. & Upadhyay, V. 2014 Efficient water utilities: use of performance indicator system and data envelopment analysis. *Water Science and Technology: Water Supply* **14** (5), 787–794. https://doi.org/10.2166/ws.2014.036.

Thanassoulis, E. & Silva, M. C. A. 2018 Measuring efficiency through data envelopment analysis. *Impact* **2018** (1), 37–41. https://doi.org/10.1080/2058802X.2018.1440814.

Toloo, M. & Tichý, T. 2015 Two alternative approaches for selecting performance measures in data envelopment analysis. *Measurement* **65** (2), 29–40. http://dx.doi.org/10.1016/j.measurement.2014.12.043.

First received 8 October 2018; accepted in revised form 2 May 2019. Available online 15 May 2019

المنارة للاستشارات
Downloaded from https://iwaponline.com/ws/article-pdf/19/7/1955/607776/ws019071955.pdf
by PROQUEST user
www.manaraa.com

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.

المنارة للاستشارات

www.manaraa.com