Applied Economics

ISSN: 0003-6846 (Print) 1466-4283 (Online) Journal homepage: www.tandfonline.com/journals/raec20

Benchmarking South African water utilities using three efficiency analysis methods

Genius Murwirapachena, Johane Dikgang & Richard Mulwa

To cite this article: Genius Murwirapachena, Johane Dikgang & Richard Mulwa (2025) Benchmarking South African water utilities using three efficiency analysis methods, Applied Economics, 57:31, 4472-4485, DOI: 10.1080/00036846.2024.2364074

To link to this article: https://doi.org/10.1080/00036846.2024.2364074

Published online: 15 Jul 2024.

APPLIED ECONOMICS
2025, VOL. 57, NO. 31, 4472–4485
https://doi.org/10.1080/00036846.2024.2364074

Routledge
Taylor & Francis Group

Check for updates

Benchmarking South African water utilities using three efficiency analysis methods

Genius Murwirapachena[a], Johane Dikgang[b] and Richard Mulwa[c]

[a]Department of Public Management and Economics, Durban University of Technology, Durban, & School of Economics and Development Studies, Nelson Mandela University, Port Elizabeth, South Africa; [b]Department of Economics and Finance & The Water School, Florida Gulf Coast University, Fort Myers, FL, USA; [c]Centre for Advanced Studies in Environmental Law and Policy (CASELAP), University of Nairobi, Nairobi, Kenya

ABSTRACT

This paper compares efficiency scores from the data envelopment analysis (DEA), stochastic frontier analysis (SFA), and stochastic non-parametric envelopment of data (StoNED) methods in the context of regulating South African water utilities. We estimate cost efficiency based on cross-sectional data from 102 South African water utilities in the period 2013/14. We compare the impact of methodological choices on the efficiency scores. For StoNED, we compare scores from two different estimation techniques (method of moments and pseudolikelihood). We also use the naïve method of averaging (NMA) and compare its results to those generated by the other methods. Results show that the choice of method has an impact on efficiency estimates and monetary cost reduction targets. When water sector regulators in the developing countries intend to benchmark utilities based on each of these methods, heterogeneity and the operating environments of utilities should be considered before deciding on the estimation method.

KEYWORDS

Data envelopment analysis; stochastic frontier analysis; stochastic non-parametric envelopment of data; water utilities

JEL CLASSIFICATION

D24; H41; Q25; Q28

I. Introduction

In the management of water supply utilities, there is usually debate on which method would be most preferred to guide in the choice of level of inputs or outputs to achieve the utilities' goals. The application of different approaches can bring different policy implications. Water utility efficiency estimation in the literature is commonly done using the stochastic frontier analysis (Amor and Mellah 2023; Nyathikala et al. 2023) or data envelopment analysis (Amaral, Martins, and Dias 2022; Sala-Garrido et al. 2023). While SFA and DEA are the commonly used tools, the Malmquist Productivity Index has also been used to measure efficiency and regulate the water distribution sector (Tourinho et al. 2023). Equally, contemporary literature shows the emergence of StoNED as an alternative efficiency analysis tool (Maziotis et al. 2023; Molinos-Senante and Maziotis 2021). This paper compares the three alternative efficiency analysis approaches, namely: DEA, SFA and StoNED.

The study is an extension of existing studies that use SFA, DEA, or other indicator-based approaches in the regulation of water utility operations (see Maziotis et al. 2023; Molinos-Senante, Maziotis, and Sala-Garrido 2022). This study is one of the few that examines the efficiency of water distribution utilities in a natural monopoly setting using three techniques (DEA, SFA and StoNED). In addition, this study provides general information on the performance of South African water distributing water utilities, which happens to be local municipalities. The provision of this information using different techniques is useful to water policymakers and is a step ahead of earlier studies conducted in the South African sector using either DEA or SFA (see Brettenny and Sharp 2016; Ngobeni and Breitenbach 2021; Nithammer, Mahabir, and Dikgang 2022).

The rest of the paper is organized into seven sections. Section II discusses the benchmarking methods used in the study. Section III gives an overview of the South African water sector. Section IV reviews some empirical literature. Section V presents the empirical approach. Section VI discusses the data used in the study. Section 7 presents and discusses the results. Section 8 concludes the study.

CONTACT Johane Dikgang jdikgang@fgcu.edu Department of Economics and Finance & The Water School, Florida Gulf Coast University, 10501 FGCU Blvd, Fort Myers FL 33965, USA

This article has been corrected with minor changes. These changes do not impact the academic content of the article.
Present affiliation for Genius Murwirapachena is Department of Economics and Finance, University of the Free State, South Africa.

II. The frontier approaches for benchmarking

Data envelopment analysis (DEA)

DEA constructs a non-parametric envelopment frontier over given data points, such that all observed points are on or below the frontier. In the context of this study, the idea is to obtain a measure of the ratio of all outputs over all inputs, such as $u'y_i/v'x_i$, where u is an $Mx1$ vector of output weight and v is a $Kx1$ vector of input weights; given that M represents outputs and K represents inputs for each water utility. The process involves obtaining values for u and v such that the efficiency measure of the i^{th} water utility is maximized (subject to the constraint that all efficiency measures are equal to or less than one). DEA can be output orientated, which aims at expanding outputs while keeping the inputs constant, or input oriented, which aims at reducing the inputs while holding the outputs constant. In both orientations, the researcher can estimate constant returns to scale (CRS) or variable returns to scale (VRS) model. Depending on the depth of the analysis, one can obtain technical efficiency, allocative efficiency or cost efficiency scores. Equally, one can exploit the returns to scale and estimate the scale efficiency scores (see Charnes, Cooper, and Rhodes 1978).

The DEA model as developed by Charnes, Cooper, and Rhodes (1978) had an input orientation that assumed CRS. In this study, we adopt the input-oriented assumption as originally but with the VRS assumption to estimate cost efficiency in South Africa's water distribution water utilities. Thus, we estimate an input-oriented DEA that assumes VRS and produce cost efficiency scores for South Africa's water providing municipalities. An input-oriented measure is adopted because it can quantify the cost reduction necessary for water utilities to become efficient holding the output constant (Luo, Bi, and Liang 2012; Reinhard, Lovell, and Thijssen 2000). This is essential in the context of this study where the model emphasizes cost reduction while holding the volume of water, the supply network, and other contextual variables constant. Further, the literature provides that when decision making units are of varying sizes, the VRS model is more appropriate (Lim and Zhu 2015). This suggests that a VRS model is ideal in this study since South Africa's water service providing municipalities varies in sizes. South Africa distinguishes municipalities according to size and groups them into three broader categories (A, B and C). Category A refers to metropolitan municipalities which are highly urbanized and populated, while Category C refers to district municipalities made up of various local municipalities (usually 3–5 local municipalities). Category B refers to local municipalities and this category is further classified into B1, B2, B3 and B4, depending on each municipality's size and operating environments.[1] While district and metropolitan municipalities are authorized water service providers, the former often delegate the water service provision function to local municipalities under their jurisdictions. This is the case when the local authority has some financial ability to manage water services provision which is common with B1 and B2 municipalities (Mahabir 2014; Nithammer, Mahabir, and Dikgang 2022).

Stochastic frontier analysis (SFA)

SFA is a parametric efficiency analysis technique that assumes a Cobb-Douglas, a log-linear or a translog functional form (Aigner, Lovell, and Schmidt 1977; Meeusen and van den Broeck 1977). The efficiency of DMUs is determined based on the specified functional form. The original formulation that is the foundation of SFA is:

$$y = \boldsymbol{\beta}'\mathbf{x} + v - u, \tag{1}$$

where y is the total cost, $\boldsymbol{\beta}\mathbf{x} + v$ is the optimal frontier goal pursued by the DMU (e.g. minimum cost), $\boldsymbol{\beta}\mathbf{x}$ is the deterministic part of the frontier, and $v\tilde{\mathrm{N}}\left[0, \sigma_{\mathrm{v}}^2\right]$ is the stochastic part. The two parts together constitute the stochastic frontier. The amount by which the observed DMU fails to reach the optimum (i.e. the frontier) is μ, where $u = |U|$ and $U\tilde{\mathrm{N}}\left[0, \sigma_{\mathrm{u}}^2\right]$. The stochastic cost frontier then changes to $v + u$, where u represents inefficiency.

[1]B1 is a local with a large town or city as urban core, B2 is a local with a medium town as urban core, B3 is a local with a small town as urban core, and B4 is a local without urban core.

Different specifications of the terms u and v distinguish stochastic frontier models. According to Aigner, Lovell, and Schmidt (1977), the normal-half normal model is the basic form of the stochastic frontier model. It assumes u to be independently half-normally $\left[\mathrm{N}+\left(0,\sigma_u^2\right)\right]$ distributed, with the idiosyncratic component v independently normally $[\mathrm{N}(0,\sigma_v)]$ distributed over the observation. Other SFA model specifications are the normal-exponential model (where u is independently exponentially distributed with variance (σ_u^2)), and the truncated-normal model (where u is independently $\left[\mathrm{N}+\left(\mu,\sigma_u^2\right)\right]$ distributed with truncation point at 0). Our study adopts an SFA model that assumes a half normal distribution of the inefficiency term to estimate cost efficiency in South African water utilities. One advantage of the half normal SFA model specification is that it can fit models with heteroskedastic error components, conditional on a set of covariates (Kumbhakar, Parmeter, and Zelenyuk 2020; Makieła and Mazur 2022). As such, most studies adopting SFA to estimate efficiency and regulate water utilities use half normal distribution as the traditional assumption for the inefficiency term (Amor and Mellah 2023; Nyathikala et al. 2023). Contemporary studies like Maziotis et al. (2023) that compare efficiency analysis techniques also use SFA with a half normal distribution assumption. Thus, for comparative and easiness of our study to fit in the contemporary literature, we also adopt an SFA with half normal distribution. The half normal distribution is also preferred because it provides a closed form of likelihood function for a stochastic frontier model which is essential given that this study uses cross sectional data (Centorrino and Pérez-Urdiales 2023).

Stochastic non-parametric envelopment of data (StoNED)

This method combines the two methodologies developed earlier into a unified model. Developed by Kuosmanen and Kortelainen (2012), the StoNED has two main stages. The first stage estimates the shape of the total cost function using the convex non-parametric least squares (CNLS) regression, which belongs to the set of continuous, monotonic increasing and globally concave functions whose disturbances satisfy the Gauss-Markov assumptions. The second stage estimates the expected inefficiency (μ), variance parameters $\left(\sigma_u^2,\sigma_v^2\right)$, and DMU-specific inefficiencies. A composite error term ($\varepsilon_i = u_i + v_i$) is included and the cost frontier function is linearized by taking the natural logs of both sides:

$$lnTC_i = lnf\left(\boldsymbol{y}_i\right) + \varepsilon_i = lnf\left(\boldsymbol{y}_i\right) - u_i + v_i \tag{2}$$

In this case, $lnTC_i$ refers to the natural logarithm of total cost of water providing water services by each municipality. The main challenge in the least squares equation above is that the expected value of the composite error term is negative, due to the inefficiency term $u > 0$; that is, $E(u_i) = \mu > 0$. The composite error term in the model violates the Gauss-Markov properties, and that it can be restored by rewriting Equation 2 as follows:

$$lnTC_i = \left(lnf\left(\boldsymbol{y}_i\right) - \mu\right) + (\varepsilon_i + \mu) = lng\left(\boldsymbol{y}_i\right) + v_i \tag{3}$$

$$\hat{\varepsilon}_i = \hat{v}_i - \hat{\sigma}_u\sqrt{2/\pi}$$

where $lng\left(\boldsymbol{y}_i\right) = lnf\left(\boldsymbol{y}_i\right) - \mu$ is the average practice cost function which can be contrasted with the best practice cost frontier $lnf\left(\boldsymbol{y}_i\right)$, while $v_i = \varepsilon_i + \mu$ is the modified composite error term. Since μ is a constant, the average practice function $lng\left(\boldsymbol{y}_i\right)$ inherits concavity and monotonicity properties from the best practice function$lnf\left(\boldsymbol{y}_i\right)$. The modified error term v_i satisfies the Gauss-Markov assumptions. The average practice frontier function can be estimated by a non-parametric regression technique such as StoNED. In the StoNED model, the assumption is that the total cost of providing water (TC) by a utility depends on a vector of outputs $\boldsymbol{y}$. Therefore, for each water utility, the CNLS problem is to find$g \in F_2$ that minimizes the sum of square deviations of the average practice function, given as:

$$\min_{f,v}\sum_{i=1}^{n} v_i^2 \left| lnTC_i = lng\left(\boldsymbol{y}_i\right) + v_i \quad \forall_i = 1,\ldots,n \right. \tag{4}$$

$$\text{s.t.} \quad g \in F_2(4)$$

The CNLS estimator for the water utilities cost function is obtained as the optimal solution to the following least squares problem, which can be solved by convex programming algorithms and solvers:

$$\min_{y,\boldsymbol{\beta},\nu} \sum_{i=1}^{n} \left(\nu_i^{CNLS}\right)^2$$

$$s.t.\ lnTC_i = \alpha_i + \boldsymbol{\beta}_i' ln\boldsymbol{y}_i + \nu_i^{CNLS} \qquad i = 1, \ldots, n$$

$$\alpha_i + \boldsymbol{\beta}_i' \boldsymbol{y}_i \leq \alpha_h + \boldsymbol{\beta}_h' \boldsymbol{y}_h \qquad \forall_i, \forall_h = 1, \ldots, n$$

$$\boldsymbol{\beta}_i \geq 0 \qquad \forall_i = 1, \ldots, n$$

$$g \in F_2 \tag{5}$$

where α_i is the intercept and $\boldsymbol{\beta_i}$ represents the coefficient of the tangent hyperplanes, which can also be interpreted as the marginal costs of output variables. These coefficients are analogous to the multiplier weights in DEA; and in contrast to the linear regression model, they are specific to each DMU (Kuosmanen 2012). Parameter ν_i^{CNLS} is the CNLS residual, the CNLS estimator of g is monotonic increasing and concave, while ε_i^{CNLS}does not need to be identically and independently distributed but is uncorrelated with outputs $\boldsymbol{y}$ (see Kuosmanen and Johnson 2010).

After the estimation of the CNLS residuals ($\hat{\nu}_i^{CNLS}$), the next step disentangles inefficiency from noise by imposing more specific distributional assumptions. Following the basic SFA (Aigner, Lovell, and Schmidt 1977), we assume the half-normal distribution for the inefficiency term, and a normally distributed noise term. Usually, the noise term is symmetrically distributed, and any skewing in the CNLS residual estimates can be attributed to inefficiency. To the best of our knowledge, there currently does not exist empirical comparison of the performance of these three methods in the water literature. We show evidence that combining various efficiency estimation methods improves the accuracy of efficiency estimates for water utilities. Our results show that comparison of the various approaches is a valuable way of determining best suited tool following evaluation of the performance of various tools.

III. Empirical approach

Water provision is a process involving several operational costs. To characterize the process, it is essential to assume the existence of a mathematical relationship between water supply inputs and outputs. The empirical model estimated in this study relates each water utility's total operation cost (TC_i) to output variables that influence its cost structure. In this context, TC_i includes costs from bulk water purchases, labour, interest on capital, depreciation of fixed assets and other general expenditures, such as fuel and oil, printing and stationery, and hiring of plant equipment incurred by each water utility. On the other hand, output variables used in this study are the volume of water supplied by each utility (Q_i), length of water pipes for each utility ($MAINS_i$) and the number of connections served by each utility (CON_i). The models also include the population served by each utility (POP_i), as a contextual variable. Therefore, our study estimates the following total cost frontier:

$$TC_i = f\left(Q_i^{\alpha} MAINS_i^{\boldsymbol{\beta}} CON_i^{\gamma} POP_i^{\Omega}\right) - u_i + \nu_i \tag{6}$$

where u_i is a random variable representing the cost inefficiency of the water utility i, ν_i is a stochastic noise term that captures the effects of measurement errors, omitted variables and other random disturbances, and $\alpha, \boldsymbol{\beta}, \gamma, \Omega$ are parameters to be estimated. When the vector $\boldsymbol{y}$ is used to represent the output variables $Q, MAINS$, CON and exogenous variable POP, Equation 6 can be written as:

$$TC_i = f\left(\mathbf{y}_i\right) - u_i + \nu_i \tag{7}$$

If a composite error term$\varepsilon_i = u_i + \nu_i$, which consists of an inefficiency term$u > 0$, and a random parameter term $E(\nu) = 0$ is introduced, and the cost function is linearized in logs, then Equation 7 will be rewritten as:

$$lnTC_i = lnf\left(\mathbf{y}_i\right) + \varepsilon_i = lnf\left(\mathbf{y}_i\right) - u_i + \nu_i \tag{8}$$

The cost function presented in Equation 8 is estimated using DEA, SFA and StoNED. For DEA, we estimate an input-oriented DEA that assumes VRS. We also estimate an SFA function where total operation cost of providing water services (TC_i) is expressed as a function of the volume of water supplied (Q_i), the length of the water pipes

($MAINS_i$), and the number of customers connected(CON_i)[2] Therefore, the SFA functional form assumes:

$$lnTC_i = \alpha_i + lnQ_i + lnMAINS_i + lnCON_i - u_i + v_i \tag{9}$$

where v_i is the noise term assumed to be in normal distribution $v_i\tilde{N}[0, \sigma_v^2]$. The u_i notation is the non-negative inefficient term (which is the distance from the observed cost to the cost on the frontier). The assumptions on the error term require both u_i and v_i to be homoscedastic. However, South African water utilities are diverse in size and operating environment. Such differences are likely to be captured in v_i, resulting in heteroscedasticity. But heteroscedasticity in v_i could lead to biased estimates, while heteroscedasticity inu_i leads to deceptive efficiency scores (Kumbhakar and Lovell 2000). To address heteroscedasticity in an SFA function, one can account for the key drivers of the variation when estimating the efficiency term. This is done by estimating a simultaneous regression on the SFA function, the inefficiency term, and the random noise term. In this study, the variation in the inefficiency term u_i driven by heteroscedasticity is controlled for by regressing u_i on the total number of water connections served by each utility (total connections in this regard is the sum of metered and unmetered water connections).

$$u_i = \alpha_1 + \alpha_2 lnCON_i + \delta_i \tag{10a}$$

This will ensure that the size of each utility is accounted for, minimizing the impact of size on the efficiency estimates. Likewise, we also control for heteroscedasticity in the noise termv_i by regressing v_i on the total population served by each water utility.

$$v_i = \alpha_1 + \alpha_2 lnPOP_i + \delta_i \tag{10b}$$

Despite earlier studies raising concerns over modelling both u_i and v_i as being heteroskedastic, literature provides that if heteroskedasticity is utility specific, one can express the heteroskedasticity in u_i and v_i as functions of utility specific variables (Kumbhakar and Lovell 2000). In doing this, one controls for bias estimates that could have emanated from heteroscedasticity in the noise term, and biased efficiency scores because of heteroscedasticity in the inefficiency term. We present results from both the method of moments (MM) and pseudolikelihood (PSL) estimators of StoNED.

IV. Data

The sample in our study comprises cross-sectional data for the 2013/14 period for 102 water utilities. We could not include all 152 water utilities, due to missing data; nor could we use panel data for other periods, due to too many gaps in the dataset. The period we are using for our analysis had the most complete data. As is the case in Dong, Hamilton and Tippett (2014), where input and output variables are defined, treating Chinese banks as multiproduct firms that employ inputs X_i at given prices W_i that minimize total costs (TC_i) to produce outputs Q_i, this study treats South African water utilities in exactly the same manner. Our study uses a single input (TC_i) with three outputs ($Q_i, MAINS_i, CON_i$) and an environmental variable (POP_i), used to control for heterogeneity in the operating environments of the utilities. Kuosmanen (2012) also uses the same variables in the context of electricity distribution utilities.

Total cost (TC) is the total water-related operating cost for each water utility and is expressed in South African Rands,[3] and comprises both direct and indirect costs resulting from bulk water purchases, labour, interest on capital, depreciation of fixed assets, and other general expenditure. In the context of this study, total cost is used as the only input variable, and water utilities are expected to minimize cost, given output variables.

Water output (Q) is the total quantity of water supplied by each water utility. The authorized consumption expressed in kilolitres (kl) per annum is used to account for water output. Total connections (CON) are the total number of metred and non-metered water connections for each utility.

[2]Justifications for the selection of these variables are given in section VI of the study.
[3]The Rand is the South African currency. As of 31 March 2020, US$1 = ZAR17.80. For easy understanding by international readers, monetary figures in this study will be presented in US$.

The connections variable shows the number of water consumer units for a water utility and is used as one of the three output variables. Length of mains (MAINS) is the total length in kilometres of the water pipes owned by each water utility during the year. This shows the distance that the water moves, from point of extraction to the last consumer for each water utility. The variable is used as one of the three output variables.

Population (POP) is the number of people served by each utility. The hypothesis is that the total cost of providing water services is likely to be higher for utilities with higher population figures. As such, this variable is used as an exogenous variable necessary to control for heterogeneity. In the literature, population is used extensively to control for heterogeneity (see Baranzini, Faust, and Maradan 2008; Filippini, Hrovatin, and Zorić 2008). Table 1 presents the summary statistics of the variables used in the analysis.

Table 1 reveals that water distribution in South Africa is heterogeneous. The distinctions in size are evident from the statistics provided. Our sample contains all the categories of water utilities (i.e. city, big-town, small-town and rural water utilities). City and big-town water utilities serve huge populations, because they have urban cores that are highly populated due to urbanization, which is prevalent in South Africa. On the other hand, utilities serving small towns and rural areas are relatively poor and have low densities in terms of population distribution. For such utilities, population levels may be relatively less and the number of water connections relatively few. However, the length of the mains and the total cost of providing water services may be relatively large because water is distributed across widely spaced household units. Total costs are higher in large water utilities, due to larger population sizes and many connections.

The implication of these variations is that they affect the selection of efficiency-analysis tools, and how the selected tools are used. The most applicable estimation tools in such heterogeneous samples are SFA and StoNED because of their ability to separate and control for noise (Andor and Hesse 2014; Kuosmanen, Saastamoinen, and Sipiläinen 2013). If one decides to use DEA, as Brettenny and Sharp (2016), did, one needs to carefully separate water utilities according to their sizes and operating environments. This is because DEA is more susceptible to the influence of outliers (see Banker and Chang 2006) and is likely to make utilities serving lower population numbers appear relatively more efficient, while utilities serving higher population numbers are deemed relatively less efficient. This is probably due to smaller utilities with smaller inputs being compared to larger utilities. Such results highlight the importance of accounting for utility size, given the heterogeneous nature of South African water utilities.

V. Results and discussion

This section presents results estimated using StoNED, SFA, DEA and the naïve method of averaging (NMA), which averages scores from SFA and DEA. Table 2 presents the utility-specific estimates of the average marginal values for the six utility categories. The average marginal values based on StoNED MM are presented at the top and average marginal values based on StoNED PSL are presented at the bottom.

The table shows the estimated beta coefficients which are interpreted to mean that a unit increase in the quantity of water supplied by a 'category A' municipality increases total cost by about 0.4. Thus, a 10% increase in the quantity of water supplied increases the total cost by about 4%. This relationship is consistent in both StoNED MM and StoNED PSL. Slight variations are only noted across utility categories where the impact of the quantity of water on total cost is relatively higher

Table 1. Descriptive statistics.

Variable	Description	Mean	Std. Dev	Min	Max
TC	Total cost (thousand US$)	16 067	45 674	56	302 247
Q	Quantity of water (thousand kilolitres)	22 700	56 700	350	352 000
CON	Total number of connections (number)	73 543	130 269	2 306	713 143
MAINS	Length of mains (kilometres)	1 485	2 593	46	12 479
POP	Population served in thousands (number)	413	801	11	4 500

Ample size (N) =102 water utilities.

Table 2. StoNED marginal values by utility category.

Category	N	Quantity	Mains	Connections	Population
StoNED MM					
A	8	0.409	0.318	0.133	0.506
C	12	0.479	0.369	0.198	0.417
B1	17	0.456	0.237	0.183	0.512
B2	17	0.418	0.231	0.189	0.589
B3	46	0.576	0.173	0.178	0.729
Average		**0.505**	**0.235**	**0.180**	**0.617**
StoNED PSL					
A	8	0.409	0.133	0.318	0.506
C	12	0.479	0.198	0.369	0.417
B1	17	0.456	0.183	0.237	0.512
B2	17	0.418	0.189	0.231	0.589
B3	46	0.576	0.178	0.772	0.729
Average		**0.505**	**0.180**	**0.534**	**0.617**

in smaller utilities (categories B3 and B4). In general, beta coefficients are similar across the two StoNED estimators.

Subsequently, we present the marginal values estimated from the SFA and compare whether the estimates are consistent with those reported in StoNED. Unlike in StoNED where utility specific coefficients are given for each variable, SFA gives aggregated marginal values. As explained earlier, we applied a half normal SFA model specification and controlled for heteroscedasticity, in both the noise term and the inefficiency term. The marginal values estimated by this method are presented in Table 3.

SFA results show quantity and length of mains and population to be significant variables that are positively related to total cost. This relationship suggests that an increase in any of these variables will result in an increase in the total operating cost. For example, a unit increase in ‘quantity’ increases the total cost by about 0.5. Thus, a 10% increase in ‘quantity’ leads to about 5% increase in the total cost of providing water services. Except for the number of connections which has a negative (though insignificant) coefficient, the marginal values reported by SFA are within the same range of magnitude as those reported by the two StoNED estimators. This implies some consistency in the frontiers estimated by both StoNED and SFA.

Table 3. SFA marginal values.

Total Cost	Coefficient	Std. Err.	p-value
Quantity	0.528***	0.137	0.000
Connections	−0.150	0.193	0.436
Length of mains	0.457***	0.189	0.016
Population	0.505***	0.164	0.002
_cons	1.663	1.267	0.189
Noise term			
Population	−0.543***	0.143	0.000
_cons	5.510	1.646	0.001
Inefficiency term			
Connections	−0.656***	0.264	0.013
_cons	5.357	2.737	0.050

Observation (N) = 102. Std. Err = standard errors. *** = significance at 1% level (i.e. $p < 0.01$).

After the frontier estimates, we compare the efficiency scores generated by each technique. This done with the intention to provide a deeper understanding of the performance of each technique based on the efficiency scores it generated. Table 4 provides the summary statistics of the efficiency scores generated by each technique.

Table 4 shows that the average efficiency scores for the 102 water utilities range from 0.447 in DEA to 0.680 in StoNED MM. For StoNED MM, the average efficiency score means that on average, water utilities are 68% efficient (i.e. 32% inefficient). This implies that utilities could reduce their operating costs by 32% and still afford to supply the same quantity of water, serve both the same population and number of connections. StoNED MM recorded considerably higher scores in terms of the mean, median and the minimum relative to the other techniques. This could be because StoNED takes the noise term explicitly into account and captures heterogeneity of firms and their operating environments using the contextual variable z, which is omitted in SFA and DEA (Johnson and Kuosmanen 2015; Kuosmanen 2012; Kuosmanen, Saastamoinen, and Sipiläinen 2013).

Another key statistic Table 4 that can shed light on the performance of each model is the standard deviation. This expresses how much the utility-specific efficiency scores vary from the mean

Table 4. Summary statistics of efficiency scores based on each estimation technique.

	Mean	Std. Dev	Median	Minimum	Maximum
StoNED (MM)	0.680	0.080	0.705	0.392	0.762
StoNED (PSL)	0.529	0.146	0.563	0.016	0.725
SFA	0.662	0.139	0.659	0.223	0.896
DEA	0.447	0.280	0.355	0.095	1.000

score. StoNED MM reported the lowest standard deviation of 0.080, while SFA reported a standard deviation lower than the one reported in StoNED PSL. If standard deviations were to be used to indicate each model's ability to control for heterogeneity in the sample, it would be argued that StoNED MM controlled heterogeneity better than the other techniques. To give details on how the efficiency scores for all utilities in the sample were distributed around the mean, we use histograms to illustrate utility-specific efficiency scores. Figure 1 presents the frequency distribution of efficiency scores generated by each technique for all water utilities.

Figure 1 shows that efficiency scores generated by StoNED MM had less variation compared to scores from the other techniques. Only five utilities reported efficiency scores of less than 0.5 under StoNED MM. The five utilities are smaller local municipalities that are predominantly rural. All the other utilities reported efficiency scores closer to the mean of 0.680 reported under StoNED MM. On the other hand, StoNED PSL had thirty-two (32) utilities with scores below 0.5 with greater variations from the mean of 0.529 noted in the sample. The 32 utilities that reported efficiency scores of less than 0.5 include one metropolitan and four district municipalities. In terms of SFA, eleven utilities had efficiency scores below 0.5, and these were predominantly smaller municipalities and two district municipalities. SFA also reported efficiency scores that considerably varied from the technique's mean of 0.662.

Meanwhile, DEA reported sixty-seven (67) utilities with efficiency scores of less than 0.5 and very high variations of scores from the reported mean of 0.447. DEA identifies utilities on the frontier and benchmarks the rest of the sample relative to those on the frontier (Lim and Zhu 2015; Luo, Bi, and Liang 2012). In this context, 7 utilities were identified as on the frontier with an efficiency score of 1 (i.e. they were 100% efficient). Thus, while the VRS property of DEA managed to control for heterogeneity in the sample, there could be other factors determining efficiency under the DEA technique.

Although the standard deviation can shed light into the distribution of efficiency scores under each model, Cheng, Bjorndal, Lien and Bjorndal (2015) argue that the statistic is not a proven scientific indicator of performance. To understand the true performance of each method, other less subjective approaches may be adopted. Since efficiency benchmarking using these methods is not yet adopted in the South African water sector, we follow Kuosmanen, Saastamoinen and Sipiläinen (2013) and use scatter plots to compare the performance of the techniques used in this study. Figure 2 provides different panels where the efficiency scores generated by each model are compared.

Each panel shows the pair of efficiency scores obtained from two technique. Almost all points in panel A are above the 45° line, indicating that StoNED MM produced efficiency scores that were greater than those reported from StoNED PSL. Panel B shows that StoNED MM produced more scores higher than those reported from DEA, while panel C shows some closer distribution of scores between StoNED MM and SFA even though StoNED MM seemed to have slightly more scores that were greater than those reported in SFA. Panel D shows that SFA produced more scores greater than those reported from StoNED PSL, while panel E shows StoNED PSL having more scores greater than scores from DEA. Lastly, panel F indicates that SFA had more scores greater than those reported from DEA, although DEA had some scores that were greater than those reported in SFA. The major implication of these comparison lies within the cost reduction targets for utilities given the efficiency scores generated by each model. Techniques that produced higher efficiency

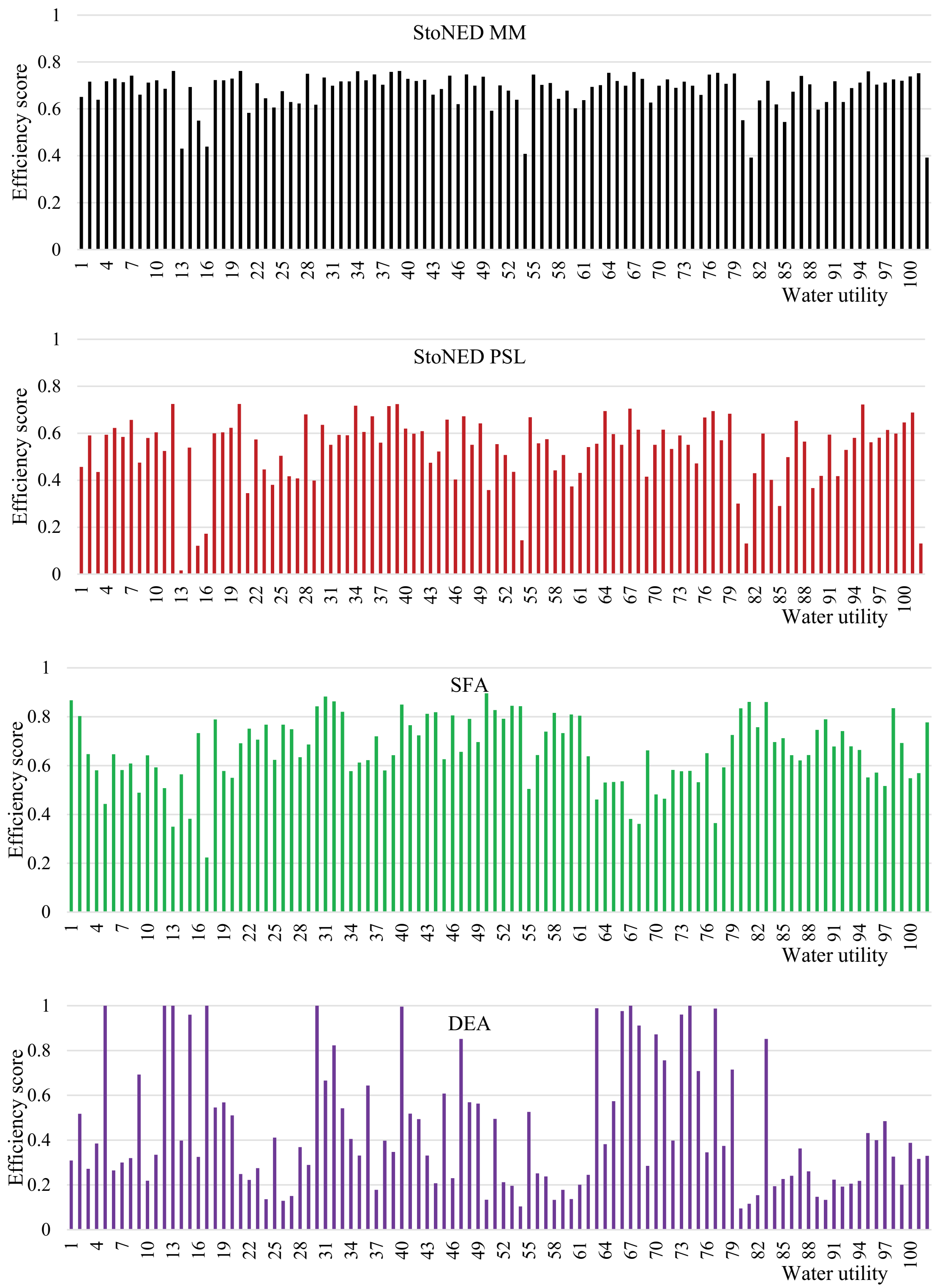

Figure 1. Frequency distribution of efficiency scores per estimation technique.

scores imply less cost reduction targets for utilities compared to those that produced lower scores.

Since South African municipalities are grouped into different categories, depending on the population dynamics served by each municipality, it is important to also understand how each municipal category performed in terms of efficiency scores based on each estimation technique. The average efficiency scores for each municipal category (except for B4 local

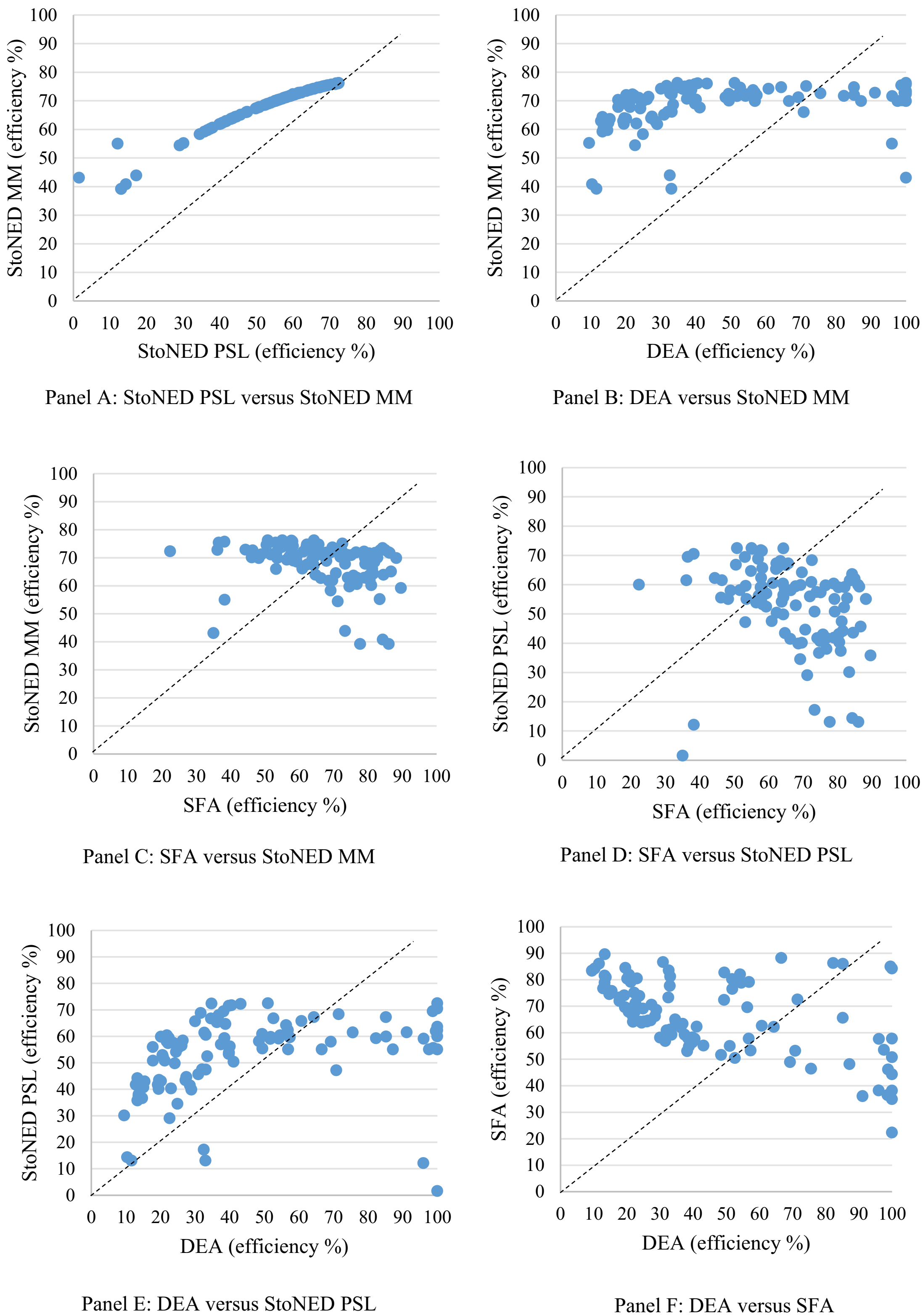

Figure 2. Comparison of utility-specific efficiency scores generated by each technique.

municipalities) under each given method are presented in Table 5.

The general summary of these results is fourfold. First, StoNED PSL generates lower efficiency scores for bigger utilities serving larger population sizes and reports moderate scores for smaller utilities compared to the other techniques. Second, SFA generates the highest efficiency scores for bigger

Table 5. Average efficiency scores for each municipal category in each estimation technique.

	A	C	B1	B2	B3
StoNED MM	0.712	0.676	0.703	0.707	0.662
StoNED PSL	0.580	0.498	0.565	0.587	0.497
SFA	0.845	0.698	0.728	0.660	0.596
DEA	0.714	0.517	0.359	0.335	0.464
N	*8*	*12*	*17*	*17*	*47*

Table 6. Kruskal-Wallis equality rank test results.

		DEA	SFA	StoNED MM	StoNED PSL
Category	Obs.	Rank Sum	Rank Sum	Rank Sum	Rank Sum
A	8	629.00	746.00	493.00	494.50
B1	17	690.00	1 118.00	982.50	983.00
B2	17	698.00	820.00	1 105.50	1 104.00
B3	47	2 532.00	1 705.00	2 122.50	2 129.00
C	12	695.00	779.00	537.50	530.50
Chi^2 with ties		*14.086*	*36.276*	*9.820*	*9.834*
Probability		*0.015*	*0.0001*	*0.081*	*0.080*

utilities and reports moderate scores for relatively smaller utilities compared to the other techniques. Third, DEA generates the least efficiency scores for smaller utilities and reports moderate scores for bigger utilities relative to the other techniques. Finally, StoNED MM reports moderate efficiency scores for bigger utilities and generates the highest efficiency scores for smaller utilities compared to the other techniques. The general implication of these findings is that the operating environment is an important factor to consider before selecting the best technique for water utility efficiency analysis. Efficiency analysis techniques that reported moderate scores under each utility category are preferred to those that either reported the highest or the least scores.

We further performed hypothesis tests to complement the varied results produced by each technique under each municipal category. To do this, the Kruskal-Wallis equality rank test (Kruskal and Wallis 1952) which tests the hypothesis that several samples are from the same population was used. Table 6 shows the Kruskal-Wallis test results.

Results from DEA and SFA confirm the existence of heterogeneity in the sample and the inability of the two methods to sufficiently control it. Meanwhile, the two StoNED techniques were robust in controlling heterogeneity across municipal categories as evidenced by the statistically insignificant p-values which suggest that there are no significant differences among the efficiency scores reported by the two StoNED techniques across the different municipal categories.

Since one of the main reasons for efficiency analysis is to establish monetary cost reduction targets for each utility, we examine the impact of each model on monetary cost reduction targets. To do this, we adopt the approach used in Kuosmanen, Saastamoinen, and Sipiläinen (2013) where firm-specific efficiency scores were converted to monetary cost targets through multiplying the total cost for each utility by the inefficiency score, that is, $TC_i(1 - TE_i)$. In this context, TC_i is the total cost for each utility, while TE_i is the cost efficiency score reported for each utility. Summary statistics on the monetary cost reduction targets are presented in Table 7.

A comparison of the average cost reduction targets generated by each method shows that each method (or estimation technique in the case of StoNED) yields different targets. StoNED PSL yielded the highest average cost reduction target of \$6 597 000, implying that if StoNED PSL is used, utilities in the sample should reduce their total costs by about \$6 597 000. SFA reported the least average cost reduction target of \$3 198 000, which is almost half the cost reduction target reported in StoNED PSL. Revelations in Table 7 are consistent with the correlation coefficients estimates presented earlier which showed no correlation in estimates generated by the different methods. This implies that the choice of method and technique has an impact on the cost reduction targets. Such a result was also reported in

Table 7. Monetary cost reduction targets (thousand US$).

	Mean	Std. Dev	Median	Minimum	Maximum
StoNED (MM)	4 605	12 560	693	17	80 697
StoNED (PSL)	6 597	17 715	1 056	25	110 427
SFA	3 198	6 871	830	13	47 780
DEA	4 183	7 127	1 634	0	46 910

Kuosmanen, Saastamoinen, and Sipiläinen (2013). In the context of our study, using SFA gives less average cost reduction target compared to using DEA and StoNED.

The empirical estimation performed in this study gives some indication on the efficiency scores generated by each method as well as the impact of model selection on both efficiency scores and cost reduction targets. However, literature suggests that the true performance of each method may not be sufficiently known from such estimations. To understand these true performances, studies such as Andor and Hesse (2014) and Kuosmanen, Saastamoinen, and Sipiläinen (2013) perform Monte Carlo simulations in addition to the empirical analysis. Although these are plausible activities on measuring the real performance of each model, this study only focused on a comparison of efficiency scores based on empirical analysis.

VI. Conclusion

In this paper, we employ parametric (SFA), non-parametric frontier (DEA) and semi-non-parametric (StoNED) approaches on a sample of South African water utilities, for methodological cross-checking purposes. SFA generated the highest efficiency scores for bigger utilities and moderate scores for relatively smaller utilities. DEA generated the least efficiency scores for local municipalities. The study estimated two forms of StoNED models (i.e. StoNED PSL and StoNED MM). The former generated the least average efficiency score of 0.497 under B3 local municipalities. On the other hand, StoNED MM reported moderate scores across all municipal categories. All techniques found metropolitan municipalities to be the most efficient utilities, suggesting that highly urbanized municipalities were relatively more efficient.

Considering the above, we consider StoNED MM as the most consistent technique. Unlike the other techniques that report very distinct results across utility categories, StoNED MM moderated efficiency scores, managing to sufficiently control for heterogeneity. When average efficiency scores generated from StoNED MM are rounded off to one decimal place, all municipal categories produced an average efficiency of 0.7. This suggests that South African municipalities are 70% efficient or can reduce their cost by 30% and still be able to provide the same quantity of water across the same network while serving the same population if they become efficient. As a result, we consider StoNED MM as the most suitable technique for a water distribution sector that is diverse and characterized with intermittent water supply and indigent consumers who also rely on free basic water services. Thus, StoNED MM is considered the most robust technique when benchmarking water utilities in the context of developing countries where utilities operate in diverse environments, with distinct budgets, and divergent intermittent water supply.

The major limitation of our study is that it only used the half normal distribution of the inefficiency term for SFA. We recommend future research wishing to compare DEA, SFA and StoNED in the water sector to include other forms of distribution for SFA, e.g. exponential, and truncated distributions. Another limitation is that our study only focused on an input-oriented DEA that assumed VRS because South African municipalities are different and are expected to be at different levels of production. Therefore, we recommend that future studies also test other forms of returns to scale. We further recommend future studies to also model heteroskedasticity in DEA. Although literature tends to use conventional DEA, its input/output data may contain random errors which may result in distorted efficiency scores due to statistical noise. Studies like Nithammer, Mahabir and Dikgang (2022) and Pereira and Marques (2021) use the double-bootstrap DEA which adjusts efficiency estimates downwards because DEA estimates may be upwardly biased under the no noise assumption.

Acknowledgements

*Funding from South African Water Research Commission (WRC) is gratefully acknowledged. This paper has been published as an ERSA Working Paper 780.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Water Research Commission.

ORCID

Genius Murwirapachena http://orcid.org/0000-0003-0767-932X
Johane Dikgang http://orcid.org/0000-0002-8076-4443

References

Aigner, D., C. K. Lovell, and P. Schmidt. 1977. "Formulation and Estimation of Stochastic Frontier Production Function Models." *Journal of Econometrics* 6 (1): 21–37. https://doi.org/10.1016/0304-4076(77)90052-5.

Amaral, A. L., R. Martins, and L. C. Dias. 2022. "Efficiency Benchmarking of Wastewater Service Providers: An Analysis Based on the Portuguese Case." *Journal of Environmental Management* 321:115914. https://doi.org/10.1016/j.jenvman.2022.115914.

Amor, T. B., and T. Mellah. 2023. "Cost Efficiency of Tunisian Water Utility Districts: Does Heterogeneity Matter?" *Utilities Policy* 84:101616. https://doi.org/10.1016/j.jup.2023.101616.

Andor, M., and F. Hesse. 2014. "The StoNED Age: The Departure into a New Era of Efficiency Analysis? A Monte Carlo Comparison of StoNED and the "Oldies" (SFA and DEA)." *Journal of Productivity Analysis* 41 (1): 85–109. https://doi.org/10.1007/s11123-013-0354-y.

Banker, R. D., and H. Chang. 2006. "The Super-Efficiency Procedure for Outlier Identification, Not for Ranking Efficient Units." *European Journal of Operational Research* 175 (2): 1311–1320. https://doi.org/10.1016/j.ejor.2005.06.028.

Baranzini, A., A. K. Faust, and D. Maradan. 2008 September. "Water Supply: Costs and Performance of Water Utilities: Evidence from Switzerland." Paper presented at the13th International Water Resources Association World Water Congress, Montpellier, 1–4 September 2008.

Brettenny, W., and G. Sharp. 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. https://doi.org/10.4314/wsa.v42i1.02.

Centorrino, S., and M. Pérez-Urdiales. 2023. "Maximum Likelihood Estimation of Stochastic Frontier Models with Endogeneity." *Journal of Econometrics* 234 (1): 82–105. https://doi.org/10.1016/j.jeconom.2021.09.019.

Charnes, A., W. W. Cooper, and E. Rhodes. 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.

Cheng, X., E. Bjorndal, G. Lien, and M. H. Bjorndal. 2015. "Productivity Development for Norwegian Electricity Distribution Companies 2004-2013." *Discussion Paper*, Norwegian School of Economics.

Coelli, T., D. S. P. Rao, and G. E. Battese. 2005. *An Introduction to Efficiency and Productivity Analysis*. Boston/Dordrecht/London, Kluwer Academic Publishers.

Dong, Y., R. Hamilton, and M. Tippett. 2014. "Cost Efficiency of the Chinese Banking Sector: A Comparison of Stochastic Frontier Analysis and Data Envelopment Analysis." *Economic Modelling* 36:298–308. https://doi.org/10.1016/j.econmod.2013.09.042.

Filippini, M., N. Hrovatin, and J. Zorić. 2008. "Cost Efficiency of Slovenian Water Distribution Utilities: An Application of Stochastic Frontier Methods." *Journal of Productivity Analysis* 29 (2): 169–182. https://doi.org/10.1007/s11123-007-0069-z.

Johnson, A. L., and T. Kuosmanen. 2015. *An Introduction to CNLS and StoNED Methods for Efficiency Analysis: Economic Insights and Computational Aspects, in Benchmarking for Performance Evaluation*, 117–186. New Delhi: Springer.

Kruskal, W. H., and W. A. Wallis. 1952. "Use of Ranks in One-Criterion Variance Analysis." *Journal of the American Statistical Association* 47 (260): 583–621. https://doi.org/10.1080/01621459.1952.10483441.

Kumbhakar, S. C., and C. A. K. Lovell. 2000. *Stochastic Frontier Analysis*. New York: Oxford University Press.

Kumbhakar, S. C., C. F. Parmeter, and V. Zelenyuk. 2020. "Stochastic Frontier Analysis: Foundations and Advances I." *Handbook of Production Economics* 4:1–40.

Kuosmanen, T. 2012. "Stochastic Semi-Nonparametric Frontier Estimation of Electricity Distribution Networks: Application of the StoNED Method in the Finnish Regulatory Model." *Energy Economics* 34 (6): 2189–2199. https://doi.org/10.1016/j.eneco.2012.03.005.

Kuosmanen, T., and A. L. Johnson. 2010. "Data Envelopment Analysis As Nonparametric Least Squares Regression." *Operations Research* 58 (1): 149–160. https://doi.org/10.1287/opre.1090.0722.

Kuosmanen, T., and M. Kortelainen. 2012. "Stochastic Non-Smooth Envelopment of Data: Semi-Parametric

Frontier Estimation Subject to Shape Constraints." *Journal of Productivity Analysis* 38 (1): 11–28. https://doi.org/10.1007/s11123-010-0201-3.

Kuosmanen, T., A. Saastamoinen, and T. Sipiläinen. 2013. "What Is the Best Practice for Benchmark Regulation of Electricity Distribution? Comparison of DEA, SFA and StoNED Methods." *Energy Policy* 61:740–750. https://doi.org/10.1016/j.enpol.2013.05.091.

Lim, S., and J. Zhu. 2015. "DEA Cross-Efficiency Evaluation Under Variable Returns to Scale." *Journal of the Operational Research Society* 66 (3): 476–487. https://doi.org/10.1057/jors.2014.13.

Luo, Y., G. Bi, and L. Liang. 2012. "Input/Output Indicator Selection for DEA Efficiency Evaluation: An Empirical Study of Chinese Commercial Banks." *Expert Systems with Applications* 39 (1): 1118–1123. https://doi.org/10.1016/j.eswa.2011.07.111.

Mahabir, J. 2014. "Quantifying Inefficient Expenditure in Local Government: A Free Disposable Hull Analysis of a Sample of South African Municipalities." *South African Journal of Economics* 82 (4): 493–517. https://doi.org/10.1111/saje.12050.

Makieła, K., and B. Mazur. 2022. "Model Uncertainty and Efficiency Measurement in Stochastic Frontier Analysis with Generalized Errors." *Journal of Productivity Analysis* 58 (1): 35–54. https://doi.org/10.1007/s11123-022-00639-y .

Maziotis, A., R. Sala-Garrido, M. Mocholi-Arce, and M. Molinos-Senante. 2023. "Cost and Quality of Service Performance in the Chilean Water Industry: A Comparison of Stochastic Approaches." *Structural Change and Economic Dynamics* 67:211–219. https://doi.org/10.1016/j.strueco.2023.07.011.

Meeusen, W., and J. van den Broeck. 1977. "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error." *International Economic Review* 18 (2): 435–444. https://doi.org/10.2307/2525757.

Molinos-Senante, M., and A. Maziotis. 2021. "Benchmarking the Efficiency of Water and Sewerage Companies: Application of the Stochastic Non-Parametric Envelopment of Data (Stoned) Method." *Expert Systems with Applications* 186:115711. https://doi.org/10.1016/j.eswa.2021.115711.

Molinos-Senante, M., A. Maziotis, and R. Sala-Garrido. 2022. "Benchmarking the Economic and Environmental Performance of Water Utilities: A Comparison of Frontier Techniques." *Benchmarking: An International Journal* 29 (10): 3176–3193. https://doi.org/10.1108/BIJ-08-2021-0481.

Ngobeni, V., and M. C. Breitenbach. 2021. "Production and Scale Efficiency of South African Water Utilities: The Case of Water Boards." *Water Policy* 23 (4): 862–879. https://doi.org/10.2166/wp.2021.055.

Nithammer, C. M., J. Mahabir, and J. Dikgang. 2022. "Efficiency of South African Water Utilities: A Double Bootstrap DEA Analysis." *Applied Economics* 54 (26): 3055–3073. https://doi.org/10.1080/00036846.2021.2002802.

Nyathikala, S. A., T. Jamasb, M. Llorca, and M. Kulshrestha. 2023. "Utility Governance, Incentives, and Performance: Evidence from India's Urban Water Sector." *Utilities Policy* 82:101534. https://doi.org/10.1016/j.jup.2023.101534.

Pereira, M. A., and R. C. Marques. 2021. "Technical and Scale Efficiency of the Brazilian municipalities' Water and Sanitation Services: A Two-Stage Data Envelopment Analysis." *Sustainability* 14 (1): 199. https://doi.org/10.3390/su14010199.

Reinhard, S., C. K. Lovell, and G. J. Thijssen. 2000. "Environmental Efficiency with Multiple Environmentally Detrimental Variables; Estimated with SFA and DEA." *European Journal of Operational Research* 121 (2): 287–303. https://doi.org/10.1016/S0377-2217(99)00218-0.

Sala-Garrido, R., M. Mocholi-Arce, M. Molinos-Senante, M. Smyrnakis, and A. Maziotis. 2023. "Benchmarking the Performance of Productive Units Using Cross-Efficiency Techniques: An Empirical Approach for Water Companies." *Water Resources Management* 37 (14): 5459–5476. https://doi.org/10.1007/s11269-023-03614-w.

Tourinho, M., F. Barbosa, P. R. Santos, F. T. Pinto, and A. S. Camanho. 2023. "Productivity Change in Brazilian Water Services: A Benchmarking Study of National and Regional Trends." *Socio-Economic Planning Sciences* 86:101491. https://doi.org/10.1016/j.seps.2022.101491.