(Initial page layout)

Benchmarking energy use in wastewater treatment plants

S. Longo*, M. Mauricio-Iglesias*, J.M. Lema*, A. Hospido*

*Department of Chemical Engineering, Institute of Technology, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain (E-mail: stefano.longo@usc.es)

Abstract

Academic research on benchmarking energy efficiency predominately focuses on two main streams of alternatives techniques: parametric and nonparametric. Since there is no agreement on the consistency of their estimates, the purpose of this paper is to investigate whether parametric ordinary least squares (OLS) and nonparametric data envelopment analysis (DEA) generate consistent wastewater treatment plant (WWTP) efficiency estimates. Our findings confirm that energy efficiency is function of several variables, including operational and exogenous factors. For it, a two-stage DEA process was followed in order to observe the overall effect produced by operation size and environmental conditions on efficiency index estimated by DEA. Based on the results of our analysis, the simply DEA in the variable return to scale assumption may not reflect the real efficiency of a WWTP. On the contrary a good consistency between OLS and 2-stage DEA was found. Conclusion of this study is that WWTP efficiency estimates may be sensitive to the method employed, and the use of multiple approaches for robustness checking is recommended.

Keywords

Energy efficiency; ordinary least squares (OLS); data envelopment analysis (DEA); Cross-methodological validation; wastewater treatment plant (WWTP)

INTRODUCTION

Minimizing energy consumption is one of the main challenges in wastewater treatment sector. To prevent environmental threats and the increased operation costs, the wastewater treatment industry has been subject to government policies aimed at improving its energy performance. In this context, water agencies and wastewater treatment plants (WWTPs) operators are manifesting a growing interest in the use of tools and methodologies to save energy, such as benchmarking procedures.

The involved stakeholders - which include federal agencies, state of regional agencies, local political leaderships, water utilities, plant managers, and operators – need to have fairly accurate information about the likely effects of their decisions on the performance of the WWTPs they direct or operate. If not, then their decisions may have the unintended consequences of rising the cost of providing sanitation service to the public or reducing the quality of this service. However, defining and measuring energy efficiency in WWTPs is still a challenge. Energy efficiency is typically approximated by energy intensity, despite several shortcomings related to this measure (Longo et al. 2016). Energy intensity is defined as the amount of energy use per unit of activity (e.g. volume of treated wastewater). Changes in energy intensity are just approximate indicators for changes in energy efficiency since they are affected by external (exogenous) factors such as influent characteristics, climate factors, scale effect of plant size and other construction parameters. Furthermore, considering that WWTPs perform different functions, i.e. removing of COD, removing of N and/or P, resource recovery, producing an effluent free of pathogens, general energy intensity indicators (i.e. kWh/m3 or kWh/PE) have limited value, as they do not provide enough information of the WWTPs operation.

These considerations imply that a good energy benchmarking methodology should be able to measure how well a WWTP performs in relation to the best firms in the industry holding constant a number of exogenous factors such as climate factors faced in the region or area where it operates. Others criteria for assessing a possible candidate for benchmarking energy efficiency are: the ability to integrate multiple outputs, sensitivity to outliers, data requirement, end user characteristics (public or internal benchmarking system), transparency and robustness.

Academic research on benchmarking energy efficiency has predominately focused on two main streams of alternatives techniques: parametric and nonparametric. The parametric approach employs econometric techniques to estimate efficiency scores. In contrast, the nonparametric approach employs mathematical programming techniques to obtain relative efficiency scores.

Despite intense research efforts in the employment of both parametric and nonparametric approaches in various goods or service provision (banks, waste collection, police services or hospitals), there is no agreement on the consistency of their estimates; i.e. the efficiency estimates derived from the different approaches should be consistent in their efficiency levels, rankings, and identification of best and worst plants. Moreover, only few studies (Longo et al., 2016) have tried to assess the energy efficiency of WWTPs using the aforementioned approaches so far.

This information gap motivates the following research question: do various benchmarking approaches generate consistent energy efficiency assessment for WWTPs? To this end, this paper cross-examines benchmarking techniques, using the same dataset. Specifically, the estimation methods used are ordinary least squares (OLS) and data envelopment analysis (DEA) in the variable return to scale model, and the bootstrapping correction for exogenous variables in DEA analysis proposed by Simar and Wilson (2007), which has never been included in for the analysis of WWTPs energy performance before. To the best of our knowledge, this is the first methodological cross-examination study to focus WWTPs energy efficiency in the spirit of Charnes et al. (1978), who advocated the methodological cross-checking of results to increase robustness of conclusions.

METHODS

Data and variables

The data about energy consumption and operation of WWTP were gathered i) by web-search engines with keywords: 'wastewater', 'WWTP', 'energy', 'energy consumption', 'energy performance', 'energy efficiency assessment', 'energy benchmarking', 'life cycle assessment', and ii) collecting energy data from regional water agencies (in particular from Germany, Spain and Switzerland) by private communications. A total of 415 WWTPs from different countries were inventoried. The data included and calculated for the analysis are summarized in Table 1.

Table 1. Summary of variables considered for the exploratory analysis.

Variable name	Abbreviation	Range	Units
Total energy consumption (input)	*E*	10 – 36,563	kWh/d
kg of COD removed (output)	*CODrem*	2.7 – 158,318	kgCOD/d
kg of N removed (output)	*Nrem*	0.01 – 4,098	kgN/d
Tertiary treatment	*TerTreat*	'YES' , 'NO'	-
Plant size	*SIZE*	23 – 507,511	Person equivalent
Plant load factor	*PLF*	4 - 512	%
Dilution factor	*DIL*	61 – 1,565	L/(PE·d)
Outdoor Temperature	*TEMP*	9.5 – 19.5	°C

E is total energy demand. *CODrem* and *Nrem*, are respectively chemical oxygen demand and nitrogen removed daily by the WWTP. *SIZE* is plant size expressed as COD-base person equivalent (PE). *TEMP* is outdoor temperature. *TertTreat* is categorical variable for presence or not of tertiary treatment. *PLF* and *DIL* are two indices defined as plant load factor and dilution factor, calculated as follow:

$$PLF = \frac{served\ PE}{design\ PE}\ 100\ [\%] \qquad (1)$$

$$DIL = \frac{daily\ influent\ flowrate}{served\ PE}\ \ [L/PE \cdot d] \qquad (2)$$

Estimation models

Parametric approach. In this study, a linear regression approach is used to control for aspects that systematically influence the energy use at WWTPs. Regression models describe the relationship between a *depended variable*, Y, and *independent variables*, X. The dependent variable is also called the *response variable*. Independent variables are also called *explanatory* or *predictor variables*. Preliminary data analysis has shown that energy efficiency has a nonlinear dependency of operational variables (Longo et al., 2016). Therefore, using a log-log functional form it is possible to determine a linear model describing the energy use in function of the explanatory variables. The log-log estimated equation is given by:

$$lnY = \beta_o + lnX\,\beta + \varepsilon \qquad (3)$$

After estimating the log-log model, the coefficients can be used to determine the impact of independent variables (*X*) on dependent variable (*Y*). In fact the coefficients in a log-log model represent the elasticity of *Y* variable with respect to *X* variable. In other words, the coefficient is the estimated percent change in your dependent variable for a percent change in the independent variables. The sign of the coefficient gives the direction of the effect. Moreover, ε, which represents the difference between actual and predicted average energy use, defines the relative energy inefficiency versus an equivalent plant with average performance.

The models obtained were refined and checked for outliers, multicollinearity, leverage and whether improper functional forms were used. The regression model yields a prediction of energy use based on the plant's operating constraints. The predicted energy use of each observation in the training dataset is divided by its actual energy use to calculate an energy efficiency ratio. A higher efficiency ratio indicates that a plant uses less energy then predicted, and consequently is more efficient. The ratios were sorted from smallest to largest and the cumulative percent of the population at each ratio was computed. For example, the ratio on the cumulative curve at 1% corresponds to a rating of 99; only 1% of the population has a ratio this small or smaller. The ratio on the curve at the value of 25% will correspond to the ratio for a rating of 75; only 25% of the population has ratios this small or smaller.

Nonparametric approach. DEA (Charnes et al., 1978) is based on the idea that the best practice frontier envelopes the data as tightly as possible. This envelopment is achieved by solving a sequence of linear programmes, one for each decision making unit (DMU) (in this study WWTPs). There are several variants of DEA concerning the behavioural objective of the DMUs (input- or output-orientation) and the returns to scale. The input- orientation is applied since the wastewater treatment sector usually has to provide a given amount of output (e.g. in our case removal of contaminants such as COD, N and P) and, based on this output, an optimal amount of input(s) (e.g. in our case electricity) will be used (and not vice versa). The input-oriented program with variable returns to scale (VRS) (Banker et al. 1984) can be written in as the following linear programming problem:

$$
\begin{aligned}
\text{Min} \quad & \theta_0 \\
\text{Subject to} \quad & \theta_0 x_{k,0} - \sum_{i=1}^{n} \lambda_i \, x_{k,i} \geq 0 \ \text{ with } k = 1, \dots, m \\
& \sum_{=1}^{n} \lambda_i \, y_{r,i} \geq \ y_{r,0} \ \text{with } r = 1, \dots, s \\
& \sum_{i=1}^{n} \lambda_i = 0 \\
& \theta_0 > 0, \qquad (4)
\end{aligned}
$$

where x_{ij} and y_{rj} are respectively the vectors of the inputs and outputs for DMU *i*. Furthermore, k (r) equals the number of inputs (outputs) employed, n represents the number of DMUs, λ_i are the weightings and θ_0 is the efficiency score. If the $\theta = 1$ the plant is fully efficient, on the other hand if $\theta < 1$ the plant evaluated is inefficient since it is capable of reducing its input(s) without affecting the amount of outputs(s).

One crucial problem with the efficiency estimates derived from equation (4) is that all DMUs are treated as they operate on the same footing. However, in practice, the efficiency level of WWTPs also depends on the effect of other factors, which are beyond the operator's control. Therefore, the only way the energy efficiency of WWTPs can correctly be evaluated is by taking into account the influence of those exogenous variables on efficiency scores, to that they can be interpreted as an appropriate measure of their performance.

The most common approach to incorporate these exogenous variables into efficiency analyses for nonparametric approach is the use of multi-stage models. One of its variants consists in a two-stage procedure: in the first-stage the linear programming problem (4) is solved; in the second stage, the inefficiency derived from the first stage are regressed on the exogenous variables using a bootstrap procedure which provides bias-correcting efficiency estimates Simar and Wilson (2007).

Procedures for cross-examination of efficiency

The efficiency scores from the different measurement models are compared using a cross-examination procedure. As one would expect, parametric and nonparametric benchmarking techniques contain different information. Thus, it is not necessary to achieve consensus on which is the single best approach for measuring energy efficiency. However, the efficiency estimates derived from the various approaches should be consistent in their efficiency levels, rankings, and identification of best and worst firms. Thus, following Bauer et al. (1998), this study compares the efficiency models by checking three consistency conditions for the efficiency measures:

(i) the efficiency scores generated by the different approaches should have comparable means, standard deviations, and other distributional properties;

(ii) the different approaches should rank the WWTPs in approximately the same order;

(iii) the different approaches should identify mostly the same WWTPs as "best-practices" and as "worst-practices".

Summary statistics of the various methods are compared to analyse consistency condition (i) in order to analyse central tendency (i.e. arithmetic mean), statistical dispersion (i.e. standard deviation) and shape of the distribution (i.e. skewness). Consistency condition (ii) is tested using

Spearman's rank order correlation coefficients for the efficiency scores generated by different methods. Spearman's ρ is defined in equation (5):

$$\rho = \frac{1 - 6 \cdot \sum d^2}{n \cdot (n^2 - 1)} \quad (5)$$

Where n is the number of rank pairs, and d is the difference between paired ranks. Consistency condition (iii) evaluates whether efficiency scores generated by different benchmarking methods identify mostly the same "best-practices" and "worst-practices" WWTPs. Following Bauer et al. (1998), the "best-practices" are defined as the firms within the top 25 percent when arranged in the descending order of efficiency scores calculated using each method. The "worst-practices" firms are those in the lowest 25 percent.

RESULTS AND DISCUSSION

Summary of efficiency scores by method

The summary of efficiency scores obtained from the different models is presented in Table 2. The mean efficiency scores vary for the three models, with the two-stage DEA model scoring the lowest efficiency. Our observations are compatible with the argument that nonparametric efficiency scores should be lower on average with comparison to parametric methods (Bauer et al., 1998). However, the comparison of these results is infeasible due to absence of previous evidence for WWTPs. As expected, due to the method employed for calculation of efficiency score, no skewness of efficiency scores was observed for OLS. On the contrary, the efficiency scores obtained from nonparametric methods are positively skewed, and it is even more so for two-stage DEA after application of the bootstrap bias correction.

Table 2. Descriptive statistics of the efficiency scores by methods

	OLS	DEA-VRS	two-stage DEA
Mean	0.5012	0.3554	0.2522
Median	0.5012	0.3207	0.2350
Minimum	0.0024	0.0668	0.0432
Maximum	1	1	1
Standard deviation	0.289	0.1997	0.1364
Skewness	1.76E-16	1.3917	2.0411
Num. Plants	413	413	413

Effect of exogenous variability on energy efficiency at WWTPs

Table 3 contains the results of the linear regression analysis (left column), and the estimated effect of exogenous variables on two-stage DEA (right column). The estimated coefficients are statistically significant in the two models and show that the effect of the covariates on the dependent variable has the expected sign.

Table 3. Determinants of WWTPs energy efficiency.

Variable	OLS	two-stage DEA
Intercept	0.015 (-1.159)	5.001** (36.61)
kg of COD removed, log	0.641** (22.78)	-
kg of N removed, log	0.335** (12.54)	-
Operational factors		
Tertiary treatment (YES)	0.140 **(3.273)	1.593** (3.476)
Size, log	-	-1.800** (-12.257)
Plant load factor, log	-0.093** (-6.526)	-0.772** (-4.739)
Exogenous factors		
Temperature, log	0.08* (3.532)	0.759** (4.214)
Dilution, log	0.07* (2.439)	0.343* (1.809)
R-squared	0.952	-
Adjusted R-squared	0.940	-
Sigma-squared	0.051	6.495
Log-likelihood	-	-929.5

Note: N = 413. Dependent variable of linear regression is total energy consumption; Dependent variable of Tobit regression is inverse of efficiency scores from DEA in first stage. t-values are given in parentheses. **(*) denotes significance at 5% (10%) level.

It has been reported that the size of WWTPs influences its energy efficiency (Longo et al., 2016), which is confirmed by this study. Plant size has the largest effect in the two-stages model, Likewise, as size increases the inverse of DEA efficiency (delta) decrease, leading to an improvement of energy efficiency.

The estimated impact of PLF of is positive and highly significant in the two models. Plants receiving lower loads compared to design values present a significantly worse energy performance, since energy consumption decreases when approaching values of 100% and keeps decreasing for overloaded plants. The results suggest that plant oversize should be as much as possible reduced the design phase and/or by revamping operation with division in two or more treatment lines in order to adapt process operation to seasonal variation of pollution load.

An increasing number of WWTPs include today tertiary treatment to prevent the discharge of effluent wastewater containing residual pathogens. The presence of a tertiary treatment increased the predicted energy use by about 15% in the OLS model. Also the efficiency obtained with the two-stage DEA model results affected by this variable. As a consequence, in a benchmark exercise it results fundamental to control for this variable if the specified output does not reflect the additional effluent improvement (as for example in this case) by disinfection. On the contrary, those WWTPs who include tertiary treatment will results penalized due to the additional energy consumption.

Another factor that impact negatively energy use at WWTPs is influent dilution. From the analysis of our dataset energy consumption increases when increasing dilution. This could be due to additional energy consumption for influent pumping.

Temperature has a complex effect on a WWTP operation. On the one hand increasing the temperature increases the biological activity, both the substrate uptake rate as the endogenous respiration. On the other hand, oxygen solubility decreases sharply when increasing temperature, leading to a higher energy demand for aeration. It is difficult to conclude which of these effects prevail. Based on our results temperature has a positive effect on the energy use (i.e. increasing temperature corresponds to an increase of the energy use).

Analysis of consistency of efficiency estimation methods

In a benchmarking exercise, if two different methods rank firms in the about the same order, than plant operators or managers would generally get the same answers when evaluating the efficiency of a set of WWTPs. If different models rank WWTPs completely differently, it is doubtful that any generalized conclusion can be drawn. The three Spearman's rank-order correlations coefficients are shown in Table 4. These coefficients capture the similarity in the efficiency rankings across the various model used.

Table 4. Spearman rank-order correlation among the efficiency scores created by various methods

	OLS	DEA-VRS	two-stage DEA
OLS	1	0.489	0.689
DEA-VRS		1	0.634
two-stage DEA			1

DEA-VRS and OLS give weakly consistent rankings with each other. The average rank-order correlation between DEA-VRS and OLS was only 0.489. On the contrary Spearman's correlation is clearly increased up to 0.689 when two-stage DEA is compared with OLS. This suggest that including exogenous variables in the efficiency estimation leads to an increase in the rank correlation. Our data suggests that, unless the correction for environmental factors is carried out, DEA-VRS and OLS are not mutually consistent, and so any conclusion based only on efficiency estimation deriving from DEA-VRS may give rise to misconceptions.

Even if the methods do not always rank WWTPs similarly, the may still be useful if they are consistent in identifying which are the most and the least efficient WWTPs. In Table 5 the upper triangle of the matrix correspond to the proportion of WWTPs that were identified by one method as having efficiency scores in the most efficient 25% of plants that were also identified in the most efficient 25% by other methods. On the contrary, in the lower triangle are reported the proportion of WWTPs that were identified in the lowest 25% of plants. For example, 53% and 54% of WWTPs identified as best-practices and worst practices by OLS, respectively, were also identified as best-practices and worst practices by DEA-VRS. Interestingly, after DEA scores were corrected by the bootstrap procedure in two-stage DEA these proportion increase to 69% and 65%. Therefore, provided that efficiency scores are corrected for environmental factors, these results indicate that there is consistency between parametric and nonparametric methods.

Table 5. Correspondence of best-practices and worst-practices plants across methods (upper triangle shows best practices, and lower triangle shows worst practices).

	Regression	DEA-VRS	two-stage DEA
OLS	1	0.535	0.687
DEA-VRS	0.545	1	0.535
two-stage DEA	0.646	0.687	1

CONCLUSIONS

In summary, the contribution of this paper to the literature is threefold:

(i) To investigate the main determinants of WWTPs energy efficiency using various estimation approaches;
(ii) To improve the existing method using DEA by bootstrap the DEA scores with Tobit regression to better explain efficiency levels; and to enrich the broader relevance of this type of analysis;
(iii) To compare efficiency estimate deriving from different benchmarking techniques by analysing consistency condition.

Moreover, as far as we are aware, this is the first methodological cross-examination study on energy efficiency of wastewater treatment sector. To be sure that the applications were comparable, the different techniques have the same efficiency concept, the same sample of WWTPs, and same specification of input and outputs.

Our findings confirm that energy efficiency is function of several factors including operational (i.e. plant load factor) and exogenous (i.e. a the grade of dilution of influent wastewater). As a consequence, those factors need to be taken into consideration when the energy efficiency of a WWTP is estimated. Thus, based on the results of our analysis, the simply DEA in the variable return to scale assumption may not reflects the real efficiency of a WWTP. The DEA-VRS method yield much lower average efficiency, ranked WWTPs differently and identified best and worst WWTPs differently from OLS. To overcome the latter, this study proposes the two-stages procedure (DEA – Tobit regression).

Although we found a good consistency between OLS and two-stage DEA, conclusion of this study is that WWTP efficiency estimates may be sensitive to the method employed, and the use of multiple approaches for robustness checking is recommended.

Acknowledgements
This project is carried out with financial support from the H2020 Coordinated Support Action ENERWATER (grant agreement number 649819): www.enerwater.eu.

REFERENCES

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.

Bauer, P. W., Berger, A. N., Ferrier, G. D., & Humphrey, D. B. 1998 Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency methods. *Journal of Economics and business*, 50(2), 85-114.

Charnes, A., Cooper, W. W., & Rhodes, E. 1978 Measuring the efficiency of decision making units. *European journal of operational research*, 2(6), 429-444.

Longo, S., d'Antoni, B. M., Bongards, M., Chaparro, A., Cronrath, A., Fatone, F., ... & Hospido, A. 2016 Monitoring and diagnosis of energy consumption in wastewater treatment plants. A state of the art and proposals for improvement. *Applied Energy,* 179, 1251-1268.

Simar, L., & Wilson, P. W. 2007 Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics,* 136(1), 31-64.