

Water pollution in wastewater treatment plants: an efficiency analysis with undesirable output

Giovanna D'Inverno^{a,b}, Laura Carosi^{c,*}, Giulia Romano^c, Andrea Guerrini^d

^a*IMT School for Advanced Studies Lucca, Italy*

^b*Faculty of Economics and Business, KU Leuven, Belgium*

^c*Department of Economics and Management, University of Pisa, Italy*

^d*Department of Business Administration, University of Verona, Italy*

Abstract

The environmental efficiency of 96 Tuscan (Italian) wastewater treatment plants (WWTPs) is investigated taking into account the quality of the outgoing water in terms of pollutant. In this regard, the presence of the residual nitrogen in the outgoing treated water is considered as undesirable output. The efficiency analysis is performed by applying a novel integrated Analytic Hierarchy Process/Non-radial Directional Distance Function (AHP/NDDF) approach, combining the benefits of the two techniques. Similarly to the standard NDDF approach, the suggested model allows to include simultaneously inputs, desirable and undesirable outputs and not to overestimate the efficiency scores. At the same time, the AHP inclusion gives the possibility to directly take into account the decision maker preferences in the weighting system and to encompass some existing directional distance function models as special cases.

The obtained results are then used to identify the efficiency explanatory variables: among them, the facilities' capacity, the percentage of wastewater discharged by the industrial and agricultural activities and the level of compliance with the pollutant concentration threshold set by the legislator have a significant impact on the WWTP performance. The integrated performance assessment allows the water authorities to combine the WWTP efficiency together with the environmental sustainability issue and it has the potential for further promising environmental inspections.

*Corresponding author

Email addresses: giovanna.dinverno@imtlucca.it (Giovanna D'Inverno),
laura.carosi@unipi.it (Laura Carosi), giulia.romano@unipi.it (Giulia Romano),
andrea.guerrini@univr.it (Andrea Guerrini)

Keywords: Data envelopment analysis (DEA), Wastewater treatment, Non-radial Directional Distance Function (NDDF), Undesirable output, Analytic Hierarchy Process (AHP)

1. Introduction

Over the last decades, the wastewater treatment has received growing attention worldwide as one of the relevant activities to ensure environmental sustainability. Referring to Goodland (1995, p. 3), environmental sustainability “seeks to improve human welfare by protecting the sources of raw materials [food, water, air and energy] used for human needs and ensuring that the sinks for human wastes are not exceeded, in order to prevent harm to humans”. More precisely, environmental sustainability is defined as “a set of constraints on the four major activities regulating the scales of the human economic subsystem: the use of renewable and non-renewable resources on the source side, and pollution and waste assimilation on the sink side” (ibidem, p. 10). The same concept has been proposed by the Organisation for Economic Co-operation and Development (OECD) in the OECD Environmental Strategy for the First Decade of the 21st Century (OECD, 2001, p. 6). The document considered the “assimilation” one of the four specific criteria for environmental sustainability, as “the releases of hazardous or polluting substances into the environment shall not exceed their assimilative capacity”. Obviously, the stated criterion is intrinsically linked to the wastewater treatment and the water and nitrogen cycles are directly involved in the definition of environmental sustainability: since the main goal of the wastewater treatment is to remove nitrogen and other pollutants from the ingoing water, this activity is extremely relevant on the sink side.

It is worth recalling that Environmental Sustainability is one of the three pillars of Sustainable Development and its key role is universally acknowledged¹. Even the 2030 Agenda for Sustainable Development (UN General Assembly, 2015) confirms the triple bottom line (social, environmental and economic) approach

¹During the World Summit on Sustainable Development in 2002 (UN General Assembly, 2002) the three pillars of Sustainable Development were identified by the words People (social sustainability), Planet (environmental sustainability) and Prosperity (economic sustainability) (see also Moldan et al., 2012).

and it defines 17 Sustainable Development Goals (SDGs) to be implemented and achieved by 2030. In this context, the improvement of water quality is considered a necessary step to ensure availability and sustainable management of water and sanitation for all (Goal 6, *ibidem*): the water quality has to be improved “by reducing pollution, eliminating dumping and minimizing release of hazardous chemicals and materials, halving the proportion of untreated wastewater and substantially increasing recycling and safe reuse globally” (Target 6.3., *ibidem*). Among the pollutants to be removed in the water after the treatments, the nitrogen is considered the most relevant one. Eutrophication, reduction of crop quality, pollution of groundwater and death of aquatic life are part of the fallout of an excessive presence of nitrogen.

The main objective of this paper is to assess the environmental efficiency of the wastewater treatment plants (WWTPs) by including in the analysis the residual nitrogen in the outgoing water. Looking at the wastewater treatment plant as a production process, the presence of nitrogen in the treated water can be considered an undesirable output. In the water sector efficiency literature, few papers deal with undesirable outputs (Picazo-Tadeo et al., 2008; De Witte and Marques, 2010; Hernández-Sancho et al., 2012; Molinos-Senante et al., 2014a, 2015b) and none of them are related to the quality of the outgoing water in terms of leftover pollutants. In line with part of the existing literature on undesirable output, the WWTP performance is evaluated using a Non-radial Directional Distance Function (NDDF) approach. As the WWTPs under analysis exhibit variable returns to scale, the Kuosmanen technology is considered (Kuosmanen, 2005) and a vector directional distance function is proposed. The efficiency scores are computed by solving a DEA-like program whose objective function is the weighted sum of the vector function components and the corresponding weights are determined by the Analytic Hierarchy Process (AHP). The integration between NDDF and AHP allows to generalize the framework proposed by Zhou et al. (2012) and by Adler and Volta (2016).

The described methodology is used to evaluate 96 WWTPs located in Tuscany (Italy). According to the different parameter specifications of the NDDF model, different efficiency scores are computed. The obtained results are then used to identify the variables affecting the efficiency. On the basis of the WWTP

features, it is evident that the facilities' capacity, the percentage of wastewater discharged by the industrial and agricultural activities and the level of compliance with the pollutant concentration threshold set by the Italian legislator exhibit an incontrovertible impact on the WWTP performance. The policy implications of the findings are mainly twofold: firstly, the WWTPs should exploit larger scale economies. Secondly, the water utilities, the environmental agencies and the regulators should promote inspection activities to stimulate a better functioning in particular of those plants that treat only domestic sewage and do not respect the nitrogen concentration regulatory limit.

The remainder of the paper is organized as follows: in Section 2 a review of the related literature is presented and in Section 3 the adopted methodology is described. Then, the next three sections are devoted to the empirical analysis. More precisely, data choice can be found in Section 4, while the critical discussion of the performed WWTPs efficiency analysis is in Section 5 and in Section 6. In Section 7 the main findings of the analysis are summarized, together with some concluding remarks. The appendix includes a short description of the wastewater treatment process.

2. Related literature

Among the huge amount of quantitative studies on the water sector (Berg and Marques, 2011; Worthington, 2014), the wastewater treatment plant efficiency analysis has gained growing attention in the recent years, i.e. starting from the 2000s (for a review, Fuentes et al., 2015). In this strand of literature, Data Envelopment Analysis (DEA) is the most used technique: it can manage a multiplicity of inputs and outputs and it does not require the selection of a specific functional form, thus resulting useful to estimate different model specifications, e.g. non-radial DEA (Hernández-Sancho et al., 2011a; Molinos-Senante et al., 2014b), DEA with uncertainty (Sala-Garrido et al., 2012a), DEA metafrontier approach (Sala-Garrido et al., 2011), Malmquist Productivity Index (Hernández-Sancho et al., 2011b; Molinos-Senante et al., 2015a). Moreover, several studies propose a further assessment of the WWTP environmental impact by means of a second stage analysis, to detect the effects of specific WWTP features on the efficiency. The most common practice is to perform

non-parametric tests such as the Mann-Whitney U test and the Kruskal-Wallis test, since they do not need the normal distribution assumption of the efficiency scores (e.g. Hernández-Sancho and Sala-Garrido, 2009; Sala-Garrido et al., 2012a; Molinos-Senante et al., 2014a). In line with this strand of literature, the present paper provides new evidences on the WWTP efficiency assessment, providing also a second stage analysis. Despite the great environmental impact of the wastewater treatment process, few papers take into account the sustainability aspects. This lack of consideration could lead to biased estimates in the performance assessment: those utilities that devote more resources to increase their environmental sustainability are penalized and turn out to be less efficient compared to those that *ceteris paribus* spend less. To address this issue in the production efficiency analysis, most of the studies introduce the undesirable output: it refers to those outputs whose increase may not be desirable. In the water sector performance assessment, few DEA papers deal with the undesirable output. With respect to other fields of application such as energy and cement sector, the notion of undesirable output has been conceived with a broader meaning. Looking at the various contributions, the undesirable outputs encompass unintended bad consequences (or negative externalities) which can be largely attributed to the production process, given the fact that producing good outputs is accompanied by the production of bads (Färe et al., 2014). More precisely Picazo-Tadeo et al. (2008) consider as non-desirable output the unaccounted-for water losses, De Witte and Marques (2010) and Hernández-Sancho et al. (2012) use the water losses, Molinos-Senante et al. (2015b, 2016) and Romano et al. (2017) introduce variables representing the lack of service quality such as the value of the penalties, the number of complaints, the number of unplanned interruptions and the number of connected water service properties with water pressure below a reference level. Concerning the WWTPs, only Molinos-Senante et al. (2014a) deal with undesirable output by considering the CO₂ emission resulting from the WWTP activity. Except for this study, the environmental impact issue has not been addressed: as far as the authors know, none of the studies on the efficiency analysis considers outgoing water pollution. This paper contributes to the literature addressing this gap and considering the nitrogen left in the water after the treatment as undesirable output (for more

details Section 4 and the Appendix).

From a methodological point of view, there are several ways to model undesirable outputs, depending on how the production technology process has been formalized. Following Dakpo et al. (2016), a first approach considers the undesirable outputs as inputs and it implicitly assumes its strong disposability. To mitigate this unrealistic assumption, a second strand of approaches considers undesirable outputs under the null-jointness and weak disposability assumption (see for all Färe et al. (1989) and for more recent contributions Färe et al. (2014), Adler and Volta (2016) and Färe et al. (2016)). However, in the attempt to better understand the role of undesirable outputs for some production processes, alternative approaches have been recently developed. They basically rely on the presence of some inputs directly responsible for pollution and to the possibility of identifying two distinct technologies, one related to the production of desirable outputs and the other specifically taking into account how certain inputs generate pollution. In this line, Murty (2010) defines undesirable output as a by-product *incidental output*; Hampf and Rødseth (2015) assume the weak-G disposability on inputs and outputs; Sueyoshi and Goto (2012a,b) identify two different notions of disposability, the natural and the managerial one, to describe the managers' response to the environmental regulations by exploiting the presence of two sub-technologies². The presence of two different technologies appears particularly suitable for the energy, petroleum and cement sector. By contrast, looking at the WWTPs' activity (see also Section 3), there is no input which can be directly associated with the chosen undesirable output, the nitrogen left in the water, and it is not possible to separate the production process into two distinct technologies. Therefore, in compliance with the above mentioned second strand of the literature, the undesirable output is modelled by assuming null-jointness and weak disposability. Regarding the efficiency assessment with undesirable output, different types of models have been used (see for all the recent surveys by Dakpo et al. (2016) and Liu et al. (2016)). Among them, it is worth mentioning the standard DEA model, the slack-based

²For a detailed discussion on the role of the undesirable output in the production process and how it can be included in a non-parametric efficiency analysis see for all Førsund (2008), Dakpo et al. (2016) and references therein.

DEA model and its extension, the model based on Russell index, the network DEA model and the Directional Distance Function (DDF) model. This latter one occupies a prominent role since it allows simultaneously the desirable output expansion and the input/undesirable output contraction (see for example Picazo-Tadeo et al. (2005), Zhang and Choi (2014) and the references therein): it is referred as radial DDF if there is a proportional adjustment of the variables (Chambers et al., 1996; Chung et al., 1997; Chambers et al., 1998; Färe and Grosskopf, 2004). Referring to the undesirable output specifically in the water sector, the DDF is considered a very suitable approach and hence it is the most developed method (e.g. Picazo-Tadeo et al., 2008, 2011; Molinos-Senante et al., 2014a, 2015b, 2016). As Fukuyama and Weber (2009) underline, the radial DDF may overestimate the efficiency when there exist non-zero slacks. To overcome this problem, several authors propose a non-radial DDF approach where slacks are directly incorporated in the efficiency measures (e.g. Cheng and Zervopoulos, 2014; Fukuyama and Weber, 2009; Färe and Grosskopf, 2010; Barros et al., 2012). A formal definition of the non-radial Directional Distance Function method (NDDF), together with several environmental indexes, is given in Zhou et al. (2012). With a similar approach, Adler and Volta (2016) suggest an economic environmental directional distance function with variable return of scale. Going further with the NDDF approach, in the present paper the non-radial directional distance function is conceived as a vector function whose components are the scaling factors associated with the reduction of inputs, the good output expansion and the reduction of the undesirable output. With a standard technique of vector optimization, a solution is found by choosing a proper scalarization (see for example Pomerol and Barba-Romero, 2000). More precisely, the vector objective function is replaced by the normalized weighted sum of its components and the weights are analytically defined, according to the Analytic Hierarchy Process (AHP) (Saaty, 1977, 1990). A NDDF approach is then integrated by the AHP; the model proposed by Zhou et al. (2012) can be seen as a particular case. There, the normalized weight vector is chosen by “assigning the same importance” to the set of inputs, the set of outputs and the undesirable output; the same happens among the input and good output variables (see also Section 3). It is worth underling that with respect to the current

analysis, although several integrated DEA/AHP models have been proposed in the recent literature (see for example Pakkar (2015) and references therein), they address different issues and they serve different purposes.

3. Methodology

3.1. The WWTP production technology

The overall production process of a WWTP is characterized by a very high environmental impact and therefore the efficiency assessment of a plant cannot be separated from its sustainability performance evaluation. Therefore, the environmental production technology has to take into account inputs, good (desirable) outputs and bad (undesirable) outputs. Inputs are described by vector $x = (x_1, \dots, x_N) \in \mathbb{R}_+^N$, while good and undesirable outputs are represented by $y = (y_1, \dots, y_M) \in \mathbb{R}_+^M$ and $b = (b_1, \dots, b_J) \in \mathbb{R}_+^J$ respectively. The environmental production technology is characterized by the following set $T = \{(x, y, b) : x \text{ can produce } (y, b)\}$ or alternatively $P(x) = \{(y, b) : (x, y, b) \in T\}$. Regarding inputs and good outputs, the environmental production technology satisfies the standard axioms of the production theory (for further details see Färe and Grosskopf, 2003): i) inactivity is always possible, i.e., $(0, 0, 0) \in T$; ii) finite amount of inputs can produce only finite amount of outputs; iii) T is convex; iv) good outputs are strongly disposable, i.e., if a given amount of inputs can produce a certain level of outputs, even a smaller quantity of outputs can be produced.

According to a very standard approach (see for all Färe et al., 1989, 2014), in the present paper, the undesirable output is considered as an unintended “by-product” of the production process and the technology is assumed to verify the following assumptions: (i) null-jointness, i.e., if $(x, y, b) \in T$ and $b = 0$, then $y = 0$; (ii) weak disposability of undesirable outputs, i.e., if $(x, y, b) \in T$, then $(x, \theta y, \theta b) \in T$, with $\theta \in [0, 1]$. Roughly speaking, null-jointness implies that there is no possibility to eliminate the undesirable outputs without stopping the good output production. Weak disposability states that a reduction of undesirable outputs is possible only if it is accompanied by a corresponding reduction of good outputs (for further discussion, see Section 4).

Looking at the wastewater treatment production, there are no empirical evidences allowing the description of the process by means of a specific functional form; therefore, non-parametric approaches are the most developed in this framework and, among them, DEA models occupy a prominent position. The technology set is then described by means of inequality and equality constraints which characterize DEA models. Moreover, preliminary analysis on the present data set show that the production processes of the wastewater treatment plants exhibit variable returns to scale³. Following Kuosmanen (2005), the production technology can be then described as follows

$$T_l = \{(x, y, b) : \sum_{k=1}^K \lambda^k y_m^k \geq y_m, \forall m \quad \sum_{k=1}^K \lambda^k b_j^k = b_j, \forall j \quad \sum_{k=1}^K (\lambda^k + \mu^k) x_n^k \leq x_n, \forall n \quad \sum_{k=1}^K \lambda^k + \mu^k = 1 \quad \lambda^k, \mu^k \geq 0, \forall k\} \quad (1)$$

where k is the number of observed production units, i.e. DMUs (in the specific context of the present paper, WWTPs); y_m^k is the m -th output produced by the k -th DMU and y_m is the m -th desirable output produced by the evaluated DMU. Similarly b_j^k (x_n^k) is the j -th undesirable output (n -th input) produced by the k -th DMU and b_j (x_n) is the j -th undesirable output (n -th input) associated with the evaluated DMU; $\lambda^k + \mu^k$ represents the intensity weights for constructing the convex combinations of the observed DMUs. Strong disposability of inputs and good outputs are formalized by the inequality constraints and weak disposability of bad outputs is described by the equality constraints related to b .

3.2. The Analytic Hierarchy Process/Non-radial Directional Distance Function approach

Taking into account the Kuosmanen technology, the environmental efficiency analysis of WWTP is performed by introducing the following vector non-radial distance function:

$$\vec{V}D(x, y, b) = \sup\{(\beta_x, \beta_y, \beta_b) : (x, y, b) + g \text{diag}(\beta_x, \beta_y, \beta_b) \in T_l\} \quad (2)$$

where $\beta = (\beta_x, \beta_y, \beta_b) = ((\beta_{x_n})_{n=1}^N, (\beta_{y_m})_{m=1}^M, (\beta_{b_j})_{j=1}^J)$ is the scaling vector function. More precisely, β_{y_m} is the scaling factor of output m , β_{b_j} and β_{x_n}

³Regarding the scientific debate about the proper way of modelling weak disposability and variable returns to scale see also Chen and Ang (2016).

represent the scaling factor of the j -th undesirable output and n -th input respectively; $g = (g_x, g_y, g_b) = ((g_{x_n})_{n=1}^N, (g_{y_m})_{m=1}^M, (g_{b_j})_{j=1}^J)$ is the explicit directional vector in which the input-output combination will be scaled (Zhou et al., 2012). For each DMU, the value of the vector non-radial distance function can be obtained by solving the following DEA-like vector maximization problem

$$\begin{aligned}
& \max \quad \beta = (\beta_x, \beta_y, \beta_b) \\
& s.t. \quad \sum_{k=1}^K (\lambda^k + \mu^k) x_n^k \leq x_n + g_{x_n} \beta_{x_n}, \quad \forall n \\
& \quad \quad \sum_{k=1}^K \lambda^k y_m^k \geq y_m + g_{y_m} \beta_{y_m}, \quad \forall m \\
& \quad \quad \sum_{k=1}^K \lambda^k b_j^k = b_j + g_{b_j} \beta_{b_j}, \quad \forall j \\
& \quad \quad \sum_{k=1}^K \lambda^k + \mu^k = 1 \\
& \quad \quad \lambda^k, \mu^k \geq 0, \quad \forall k \\
& \quad \quad \beta_{x_n} \geq 0, \beta_{y_m} \geq 0, \beta_{b_j} \geq 0 \quad \forall n, \forall m, \forall j
\end{aligned} \tag{3}$$

The above general formalization can be differently specified according to the form of β and g and hence Problem 3 can be seen as a general framework encompassing some relevant directional distance function models. In the present analysis, the following specifications are considered: i) $\beta = (\beta_x)$ and $g = (-x, 0, 0)$, ii) $\beta = (\beta_x, \beta_y)$ and $g = (-x, y, 0)$, iii) $\beta = (\beta_x, \beta_y, \beta_b)$ and $g = (-x, y, -b)$. In the first case, the model allows for the reduction of inputs, while in the second case it deals with the simultaneous input reduction and output expansion. The third specification of β and g takes into account the input and bad output reduction together with the output expansion. The vector formulation emphasizes that improvements for inefficient DMUs can be suggested through different directions; the non-radial approach is therefore taken to its extreme. However, from a mathematical point of view, a vector maximization problem has multiple non-dominated solutions and, among them, it is necessary to identify the most suitable ones with respect to the environmental efficiency analysis of the DMUs ⁴. Following a very standard approach (see for all Pomerol and Barba-

⁴DEA vector optimization problems have been introduced even in other different contexts. Among them, an interesting contribution is the one by An et al. (2016) where a Nash bar-

Romero, 2000), the vector maximization problem can be solved by maximizing a scalar function of the type $w^T \beta$ where $w = ((w_{x_n})_{n=1}^N, (w_{y_m})_{m=1}^M, (w_{b_j})_{j=1}^J)$ is the normalized weight vector, that is $\sum_{n=1}^N w_{x_n} + \sum_{m=1}^M w_{y_m} + \sum_{j=1}^J w_{b_j} = 1$ and $w_{x_n} \geq 0, w_{y_m} \geq 0, w_{b_j} \geq 0, \forall n, m, j$. Once the set of weights is chosen, the environmental efficiency score of each DMU is obtained by solving the following scalar problem

$$\begin{aligned}
& \max && w^T \beta \\
& s.t. && \sum_{k=1}^K (\lambda^k + \mu^k) x_n^k \leq x_n + g_{x_n} \beta_{x_n}, \quad \forall n \\
& && \sum_{k=1}^K \lambda^k y_m^k \geq y_m + g_{y_m} \beta_{y_m}, \quad \forall m \\
& && \sum_{k=1}^K \lambda^k b_j^k = b_j + g_{b_j} \beta_{b_j}, \quad \forall j \\
& && \sum_{k=1}^K \lambda^k + \mu^k = 1 \\
& && \lambda^k, \mu^k \geq 0, \quad \forall k \\
& && \beta_{x_n} \geq 0, \beta_{y_m} \geq 0, \beta_{b_j} \geq 0 \quad \forall n, \forall m, \forall j
\end{aligned} \tag{4}$$

In this regard, the non-radial directional function model proposed by Zhou et al. (2012) (see Problem 12 p. 629) can be seen as a particular case of Problem 4⁵. As regards the weights' assignment, in the authors' opinion, the decision maker (in the present analysis the water utility) should be allowed to determine a set of weights according to the specific features of the evaluated production process. To address this issue, the Analytic Hierarchy Process appears a valuable tool. AHP has been developed by Saaty at the end of the seventies (Saaty, 1977) and later widely applied in different fields and contexts. In the present case, the AHP model is able to define a set of weights w which takes into account the preferences of the decision maker (DM) on the relative importance of inputs, good and bad outputs. The DM is asked to establish if inputs are more important than good outputs and to define the intensity of such importance. The

gaining approach is proposed.

⁵Zhou et al. (2012) deeply investigate the relationship among their models and the most relevant distance function models (radial and non-radial) in the recent literature. Their considerations still apply for the present paper.

comparison yields a number which is determined on the basis of the standard AHP scale, from 1 (Equal Importance) to 9 (Extreme Importance).

Similar judgements have to be given between inputs and bad outputs and between good and bad outputs. Therefore, the following order 3 square matrix A is obtained

$$\begin{array}{c} x \quad y \quad b \\ x \begin{pmatrix} a_{xx} & a_{xy} & a_{xb} \\ y \begin{pmatrix} a_{yx} & a_{yy} & a_{yb} \\ b \begin{pmatrix} a_{bx} & a_{by} & a_{bb} \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{array}$$

Clearly $a_{ij} > 0$, $a_{ii} = 1$ and $a_{ij} = \frac{1}{a_{ji}} \forall i, j$. A is called pairwise comparison matrix; the set of global weights $w = (w_x, w_y, w_b)$ is the normalized eigenvector associated with the dominant eigenvalue of matrix A (see for all Saaty, 1990).

Going on with the pairwise comparisons among inputs, the relative weights of input variables are determined. More precisely, according to the DM preferences a pairwise comparison matrix I is constructed. The normalized eigenvector r associated with the dominant eigenvalue of matrix I represents the relative weights of input variables and hence the weight associated with the n -th input is defined as $w_{x_n} = w_x r_n$. The weights for desirable and undesirable output are computed following the same procedure. Furthermore, to detect the inconsistency of the DM preferences, the following Inconsistency Ratio is computed: $IR = \frac{\alpha_{\max} - n}{(n - 1)CRI}$, where α_{\max} is the dominant eigenvalue of the pairwise comparison matrix, n is the dimension of the matrix and CRI represents the coefficient of random inconsistency which is computed by calculating $\frac{\alpha_{\max} - n}{(n - 1)}$ for randomly filled reciprocal matrices (see for all Saaty, 1990). If $IR > 0.1$, then the inconsistency occurs and the DM has to revise her/his judgement.

Following the AHP approach, the system of weights suggested by Zhou et al. (2012) can be obtained by considering pairwise comparison matrices whose entries are all ones.

Once the set of weights are determined, Problem 4 is solved for each DMU; the higher the optimal value, the lower the efficiency level of the evaluated unit. In line with Zhou et al. (2012), the obtained scores are then used to construct normalized efficiency indexes, where 1 corresponds to the best performance and 0 to the worst. Obviously, according to the specification of g and β different

models can be considered and then different water efficiency performance indexes (WPI) can be constructed. The present analysis deals with the indexes presented in Table 1.

| Table 1: Indexes | | | |
|------------------|--|---|---|
| | Input model | Input/Good Out. model | Input/Good/Bad Out. model |
| β | $\beta = (\beta_x)$ | $\beta = (\beta_x, \beta_y)$ | $\beta = (\beta_x, \beta_y, \beta_b)$ |
| g | $g = (-x, 0, 0)$ | $g = (-x, y, 0)$ | $g = (-x, y, -b)$ |
| Index | $WPI_1 = 1 - \sum_{n=1}^N w_{x_n} \beta_{x_n}$ | $WPI_2 = \frac{1 - \sum_{n=1}^N w_{x_n} \beta_{x_n}}{1 + \sum_{m=1}^M w_{y_m} \beta_{y_m}}$ | $WPI_3 = \frac{1 - \left(\sum_{n=1}^N w_{x_n} \beta_{x_n} + \sum_{j=1}^J w_{b_j} \beta_{b_j} \right)}{1 + \sum_{m=1}^M w_{y_m} \beta_{y_m}}$ |

3.3. Identifying WWTP efficiency explanatory variables

In line with an increasing part of the literature, a second-stage analysis is performed to identify whether there is a relationship between some WWTP features and the efficiency scores obtained as described in the previous section. The regression analysis is one of the most common methodological approach, but its application presents few drawbacks: among them, there are the misspecification of the model because of omitted variables that should have been introduced rather in the first stage (Hernández-Sancho et al., 2011b) and inaccurate results that might arise from serial correlation between the error term and the covariates in the second stage (Simar and Wilson, 2007). Therefore, in this paper a different approach is preferred and it is applied into two steps: first of all, the WWTPs have to be categorized into groups by different operational factors that could affect the WWTP performance; then, a test is performed to assess whether there is or not statistically significant difference between/among groups according to the explanatory factor under scrutiny. As the WWTP sample does not satisfy all the necessary assumptions to apply parametric and statistical tests such as the *t-test* or the *analysis of variance-ANOVA* (Hernández-Sancho et al., 2011b; Molinos-Senante et al., 2014b), the corresponding non-parametric test is performed: the Mann-Whitney U test applies for two groups, while the Kruskal-Wallis test for three groups or more (see e.g. Kruskal and Wallis (1952) and Ruxton and Beauchamp (2008) for further details). The null hypothesis states that the groups/samples originate from the same population, while the alternative hypothesis asserts that they originate from other populations. The

null hypothesis is rejected for a p lower than or equal to 0.05: if this is the testing result, it is possible to conclude that the factor under investigation does affect the WWTP efficiency.

4. Data

The empirical analysis involves 96 wastewater treatment plants located in Tuscany and controlled by Acque SpA, a public-private utility entrusted in 2002 with water services in the so called “Basso Valdarno” river basin in the Pisa province. Data are provided by Acque Spa and refer to 2014. The data grid for this study has been constructed with the support of the Tuscan water authority staff and the technical staff of Acque SpA and Ingegnerie Toscane. The data have been gathered by a team of engineers and their consistency has been double-checked by Acque management and researchers. In compliance with the basic DEA requisites, the sample consists of a group of homogeneous WWTPs to be compared: the units under analysis refer to those plants that have costs both for the water treatment and for the sludge process and they have been refined by means of a preliminary outlier detection analysis.

As a fundamental step in the efficiency assessment, the variables have been selected not only according to the related literature, but also according to the opinion of the engineers and the data availability (e.g. Fuentes et al., 2015). Before going to define the variables, it is worth pointing out that as concerns the input and output choice, a selection screening process has been executed as proposed in Golany et al. (1994): in particular, the correlation analysis between pairs of factors turns out to be useful to identify redundant variables and then to increase the discriminatory power of the DEA method. Table 2 presents the descriptive statistics of the variables introduced as follows.

Input. According to the mainstream literature, costs for the wastewater treatment functioning are considered as inputs. They can be taken both at an aggregate level as total costs (e.g. Da Cruz et al., 2012; Molinos-Senante et al., 2014a) or at a disaggregate level (e.g. Hernández-Sancho and Sala-Garrido, 2009; Hernández-Sancho et al., 2011a,b; De Witte and Marques, 2012; Sala-Garrido et al., 2012b; Molinos-Senante et al., 2014b). In this analysis three different cost items have been identified: (x_1) materials and energy costs; (x_2) staff and main-

tenance costs; (x_3) sludge transport and disposal costs: they all are expressed as €/year.

Desirable output. Looking at the literature on the WWTP efficiency evaluation, basically two approaches can be identified for the output choice. In the first one, the volume of treated/delivered water and/or the population served are considered as output. The papers by e.g. De Witte and Marques (2010), Picazo-Tadeo et al. (2011), Da Cruz et al. (2012), De Witte and Marques (2012) belong to this first strand of literature. Alternatively, outputs can be chosen among the eliminated contaminants and the quantity of pollutants removed to value the production of a plant or as the difference between the pollution level in the influent and effluent, namely Net Environmental Benefits (see for example Hsiao et al., 2007; Hernández-Sancho et al., 2011a,b; Sala-Garrido et al., 2011, 2012a,b; Fuentes et al., 2015; Molinos-Senante et al., 2015a). After the preliminary screening process and in compliance with the first strand of the literature, the treated water, expressed in m^3 , is chosen as output (y_1). Considering in addition the main WWTP competences and the engineering expertise, a second output (y_2) has been selected, the Kg of removed sludge; actually it is by far the largest removed constituent (for further details see the Appendix).

Undesirable output. As already pointed out in Section 2, there are no efficiency analysis papers dealing with the WWTP water pollution as an undesirable output. Once the wastewater enters in the treatment plant, it is characterized by the presence of several constituents. Even though one of the main objective of a WWTP should be the removal of as many contaminants as possible and to get the water purified for further reuse, it is almost impossible to remove them completely and so they are still present in the ongoing wastewater. The higher the pollutants in the outgoing wastewater, the higher the negative impact on the environment. Among the constituents, the nitrogen is one of the most preeminent pollutant and its relevance is widely acknowledged in the related literature (see e.g. Lorenzo-Toja et al., 2015): accordingly, this paper introduces as undesirable output the quantity of nitrogen which remains in the outgoing wastewater. It is worth pointing out that the chosen undesirable output can be seen as a bad externality and as an unintended by-product of the production process in the sense of Färe et al. (2014), but it is far from the definition of by-product given

by Murty (2010). Actually, looking at the data and at the treatment process, the assumptions of null-jointness and weak disposability are fulfilled, ruling out other approaches proposed in the literature to model the undesirable outputs⁶.

Table 2: Descriptive statistics

| | Inputs | | | Desirable outputs | | Undesirable output |
|-----------|------------|-------------|-------------------|-------------------|--------------|--------------------|
| | Materials+ | Staff+ | Sludge transport+ | Treated | Removed | Residual |
| | Energy | Maintenance | disposal | water | sludge | Nitrogen |
| | € | € | € | m^3 | Kg | Kg |
| Mean | 41,894.25 | 13,556.09 | 33,190.35 | 402,233.20 | 503,315.50 | 5,557.39 |
| Std. Dev. | 86,448.89 | 23,710.57 | 67,347.96 | 942,951.40 | 757,611.10 | 11,745.93 |
| Min | 1,270.81 | 375.19 | 316.90 | 1,515.00 | 2,000.00 | 70.75 |
| Max | 529,684.70 | 145,919.00 | 447,154.00 | 6,234,272.00 | 4,659,130.00 | 78,551.83 |

5. The WWTPs performance assessment

5.1. Model set-up

As explained in Section 3.2, the choice of the normalized weight vector w is a key element of the analysis: in the following, two sets of weights are used. The first one is constructed by assigning the same importance to the three groups of variables (inputs, good and bad outputs) and the same applies inside each group (see also Section 3.2). The obtained weight vector coincides with the one proposed by Zhou et al. (2012). Referring to the second set, different pairwise comparison matrices are taken, following a discussion with the water utility staff. More precisely, the importance of the undesirable output is judged “very strong” with respect to good output (input). On the input side, the first two inputs share the same level of importance and they are strongly more important than the third input. Finally, with respect to the kg of removed sludge, the importance of treated water is judged very strong. Table 3 describes the chosen pairwise comparison matrices, the associated inconsistency ratio (IR) and the corresponding generated weights.

⁶ In the presented WWTP framework, there are no inputs which are specifically related to the “production” of undesirable output and therefore one of the five attributes for the by-product technology “à la Murty” fails to be verified. Therefore, the technology proposed in Murty et al. (2012) cannot be used in the present context. The authors are grateful to an anonymous referee for giving the opportunity to better clarify this important and debated aspect (for a broader overview, see Dakpo et al., 2016).

Table 3: AHP-non-radial set of weights

| | Global comparison | Inputs | Good Outputs | Bad Output |
|-----------------|---|---|--|------------|
| Matrix | $\begin{pmatrix} 1 & 3 & 1/7 \\ 1/3 & 1 & 1/9 \\ 7 & 9 & 1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 1 & 5 \\ 1 & 1 & 5 \\ 1/5 & 1/5 & 1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 7 \\ 1/7 & 1 \end{pmatrix}$ | |
| IR | 0.0692 | 0 | 0 | |
| WPI_1 Weights | (1, 0, 0) | (0.455, 0.455, 0.09) | | |
| WPI_2 Weights | (0.75, 0.25, 0) | 0.75 * (0.455, 0.455, 0.09) | 0.25 * (0.875, 0.125) | |
| WPI_3 Weights | (0.149, 0.066, 0.785) | 0.149 * (0.455, 0.455, 0.09) | 0.066 * (0.875, 0.125) | 0.785 |

Before showing the obtained results, it is worth pointing out that a further index WPI_0 has been computed in addition to the three indexes described in Section 3.2 and listed in Table 1: for the sake of comparison, WPI_0 has been constructed by considering a non-radial Directional Distance Function model where the undesirable output is completely ignored. Referring to the parameter specification, the form of β and g has been set as $\beta = (\beta_x, \beta_y)$ and $g = (-x, y)$, while the equality constraint associated with the weak disposability of the undesirable output has been cancelled. The set of weights of WPI_0 coincides with the one for WPI_2 .

5.2. Results

For both sets of weights and for the 96 WWTPs, the efficiency indexes have been computed. Table 4 presents in a synthetic way the main descriptive statistics of the results: the full list of the results are available upon request from the authors. The first part of the table is denoted as “Non-radial” and it refers to the set of weights proposed by Zhou et al. (2012). The second part is referred as “AHP-non-radial” and it is related to the set of weights coming from the pairwise comparison matrices of Table 3⁷.

⁷All models have been implemented using MATLAB 9.0 R2016a.

Table 4: Efficiency score for each Wastewater Performance Index

| | WPI_0 | WPI_1 | WPI_2 | WPI_3 |
|-----------------------|---------|---------|---------|---------|
| <i>Non-radial</i> | | | | |
| Mean | 0.430 | 0.654 | 0.556 | 0.564 |
| Std. Dev. | 0.322 | 0.277 | 0.349 | 0.338 |
| <i>AHP-non-radial</i> | | | | |
| Mean | 0.419 | 0.647 | 0.571 | 0.623 |
| Std. Dev. | 0.333 | 0.290 | 0.338 | 0.306 |
| N° efficient WWTPs | 19 | 33 | 33 | 31 |

Not surprisingly, the number of efficient units increases as the undesirable output enters in the analysis. In fact, the wastewater treatment plants might face higher costs because of their water quality concern: the more their effort, the more efficient the wastewater treatment process and the lower the quantity of dangerous nitrogen in the outgoing water. If this aspect is not considered in the WWTP performance assessment, the “environmentally oriented” WWTPs will be penalized.

By comparing the results across the three models with the undesirable output (WPI_1 , WPI_2 and WPI_3), a more discriminating power is shown by the one where the contraction of both inputs and bad output and the expansion of outputs are simultaneously considered (WPI_3). Looking inside the same model specification, the number of efficient units does not change even if the normalized weight vectors varies. However, the different weighting scheme affects the WWTP performance assessment in terms of efficiency scores: as it will be clarified in the next section, it influences also the explanatory variable investigation. Not surprisingly, the WPI_2 and WPI_3 efficiency scores in the AHP-non-radial case are higher than in the non-radial one: those plants who are more “environmentally” focused are valued more for keeping their water quality commitment when the undesirable output is taken into account and its importance is considered way more important than the other variables in the assessment.

6. A step further in the WWTPs performance assessment

Once the wastewater environmental performance indexes have been obtained, a further analysis has been performed. Firstly, the variables affecting the

WWTP environmental impact and their efficiency assessment have been detected. Then, the main evidences have been discussed so to give the water authorities and the decision makers additional insights for the WWTP management.

6.1. External factor choice

In compliance with the related literature (e.g. Hernández-Sancho et al., 2011a,b; Molinos-Senante et al., 2014a,b, 2015a) and according to the opinion of the engineers, the data availability and the WWTP main activities (for more technical details see the Appendix), the following external factors have been put under scrutiny. (1) *Age*: two groups of WWTPs have formed depending on whether the buildings are under 30 years old (that is after or before 1985). The thresholds have been set considering that the useful life of a WWTP is averagely 25/30 years, so that a plant over 30 years old is outdated. (2) *Plant capacity* expressed as *Population Equivalent*⁸: the facility size can be expressed in terms of per capita and per day pollution load; three groups have been defined in accordance with the Decision on Implementation Programmes (European Commission, 2007): i) less than 2000 PE, ii) between 2000 and 10,000, and iii) greater than 10,000. (3) *Sewage system*: WWTPs have been split into two groups, as they can have either a separate or a combined system. (4) *Kind of treatment*: two groups have been defined as the WWTPs in the sample use either secondary or tertiary treatment. (5) *Technologies*: referring to the wastewater treatment technology, WWTPs have been divided into two groups, as they can use or not activated sludge process. (6) *Estimated dry weather flow*: it refers to the wastewater flow occurring during the dry season when groundwater infiltration and surface runoff have a minimum influence: i) less than 100,000 m^3 /year, ii) between 100,000 and 500,000 m^3 /year, and iii) greater than 500,000 m^3 /year. (7) *Wastewater discharged by industrial and agricultural activities* (expressed as %): the WWTP sample has been clustered in three different groups, considering separately i) those with no wastewater discharged by these activities at all, ii) those with a percentage lower than 10, and iii) those with one higher than 10.

⁸“1 P.E. (Population Equivalent)” means the organic biodegradable load having a five-day biochemical oxygen demand (BOD5) of 60 g of oxygen per day (directive 91/271/EEC).

(8) *N Concentration regulatory limit*: since the detrimental effect of nitrogen in terms of eutrophication and environmental impact on plant and aquatic life is acknowledged also by the national legislator, two groups of WWTPs have been distinguished depending on whether the outgoing nitrogen concentration is below or above the limit (30 mg/L) set by the legislative decree 152/2006.

6.2. Results and policy remarks

Table 5 presents the Mann-Whitney U and Kruskal-Wallis test results and the average efficiency scores, together with the efficient WWTP percentages and the standard deviations, for each Water Performance Index and for both the weighting schemes over the WWTPs under analysis, grouped according to the described explanatory variables⁹. Before going into detail, the results suggest two preliminary considerations. First of all, when the undesirable output is considered in the efficiency assessment, there can be different test outcomes: this is the case between, on the one hand, WPI_1 and, on the other hand, WPI_2 and WPI_3 . As already highlighted in Section 5.2, the inclusion of the undesirable output in the WWTP technology process allows those more “environmentally focused” plants to be evaluated in a fairer way and, more broadly speaking, it better depicts the overall WWTP process framework. Accordingly, the factors that might affect the performance assessment are better captured when also the undesirable output is included in the analysis. Moreover, it can be observed that there are not remarkable differences across the two different weighting scheme proposed in the current analysis. However, this might not have been the case if the water utilities would have assigned different level of importance to the inputs, good and bad outputs at the efficiency analysis stage.

Among the examined factors, the facility capacity seems to play an important role. In fact, considering as its proxy both the Population Equivalent and the dry weather flow, the efficiency means are greater for WWTPs with bigger capacity than for smaller plants and the Kruskal-Wallis test results lead to reject the equality of means hypothesis especially in the WPI_3 model. These evidences are consistent to each other and in line with other empirical applications (e.g. Hernández-Sancho et al., 2011b; Molinos-Senante et al., 2014a,b).

⁹Descriptive statistics and test results have been obtained using Stata 13.

This suggests the water utilities room for improvement in terms of unexploited economies of scale: large plants show a good hydraulic performance and their pumps achieve a high productivity. Furthermore, a higher scale of operations allows the adoption of more advanced technology and therefore a higher removal rate is obtained.

A different reasoning applies for the aspects related to the adopted technologies: the Mann-Whitney U test results do not lead to reject the null hypothesis. In fact, the distinction between secondary or tertiary, combined or separate, activated sludge process usage or not usage, does not suggest statistically significant differences and therefore these variables cannot be considered as a determinant explanatory factor for the efficiency assessment in this context. For example, the lack of significance between secondary and tertiary treatment might rely on the opposite effects exerted by high advanced treatment process (as tertiary) on costs and quality: as a matter of fact, costs grow up for the relevant capital expenditure as well as the water quality is improved. In general, these evidences are in line with other applied studies: one explanation can be related to the WWTPs sample choice. In fact, to perform an efficiency analysis, the units have to be rather homogeneous in the treatment process: looking at the specific features of the sample under analysis, this requirement is fulfilled despite the different classification. However, exploring these factors has at least two advantages: firstly, it is useful to double check the selected sample in the first stage of the analysis. Secondly, it is possible to observe the characteristics of the most efficient WWTPs: on average, there are higher efficiency scores and higher presence of efficient plants in the group that shares a tertiary, separate technology and does not use activated sludge process.

Then, an interesting consideration stems from the feature related to the wastewater discharged by the industrial and agricultural activities: the efficiency score averages increase as the percentage of the efficient WWTPs increases, but the rejection of the test null hypothesis depends on the model specification. In fact, the null hypothesis is rejected for both the weighting schemes only when the undesirable output is considered: this might suggest that the efficiency assessment can be conditional upon the main target of the WWTP operators. The result is in conflict with prior literature that shows a poor performance among

the plants treating the sewage from factories and farms (Guerrini et al., 2016). The novelty of the result obtained in the current paper can be attributed to the measurement of the nitrogen as undesirable output in the efficiency model, mainly for two reasons. First of all, the sewage produced by some farms and factories (mainly paper mills in the Pisa area) is poorer of nitrogen than the domestic wastewater: accordingly, the residual amount of this pollutant after the treatment process is rather low, making the plants performance better. Moreover, the plants treating sewage from factories generally turn out to be more “environmentally” oriented.

Instead, referring to the year of the plant, the Mann-Whitney U test results do not enable to reject the null hypothesis in most of the cases. Consistently with other studies (e.g. Hernández-Sancho et al., 2011b; Molinos-Senante et al., 2014b,a), age cannot be considered as a determinant factor in the efficiency assessment.

Lastly, it is possible to observe that the two groups obtained following the concentration limit set by the Italian legislator are statistically different whenever the undesirable output is taken into account. This information provides useful insights: in fact, even if few efficient WWTPs are found among those above the set threshold, the efficiency score average is way larger in the case the plants manage to keep the outgoing concentration below the limit. Accordingly, this evidence suggests that the “environmentally focused” WWTPs benefit in terms of performance assessment rather than being damaged, despite their water quality commitment and the incurred high costs to develop a good treatment with high pollutant removal rate.

In terms of policy implications, the obtained results could provide useful suggestions for the water utilities, the environmental agencies and the regulators. For example, since “big is better” for the wastewater treatment, the utilities’ managers should plan to exploit larger scale economies: this would imply higher cost savings, but at the same time higher environmental standards achievement. From the point of view of the environmental agencies, the highest efficiency scores obtained by the plants serving factories and farms suggest to perform an inspection activity on small plants treating only domestic sewage. Moreover, the environmental controls concerning the nitrogen concentration regulatory

limit should also be increased, so to stimulate a better functioning of those plants that do not respect the set threshold and show a lower performance level. Finally, the evidences show the water authorities the benefits that could arise from an integrated performance assessment that penalizes WWTPs aiming only at getting cost savings and achieving poor environmental standards. The results obtained by the adoption of the “environmental performance index” might suggest the water regulators a benchmarking model for the WWTPs and a yardstick competition to water utilities regulation.

Table 5: WPIs by explanatory factors and Mann-Whitney U/Kruskal-Wallis test results

| Explanatory factor | Total WWTPs | WPI ₁ | | | | WPI ₂ | | | | WPI ₃ | | | |
|---|-------------|------------------|-------|-----------|--------------|------------------|-------|-----------|---------------|------------------|-------|-----------|---------------|
| | | % Eff. | Mean | Std. Dev. | Test | % Eff. | Mean | Std. Dev. | Test | % Eff. | Mean | Std. Dev. | Test |
| Non-radial | | | | | | | | | | | | | |
| Year | | | | | | | | | | | | | |
| <1985 | 44 | 36% | 0.677 | 0.278 | 0.5185 | 36% | 0.611 | 0.322 | 0.0637 | 34% | 0.626 | 0.311 | 0.0584 |
| ≥ 1985 | 52 | 33% | 0.634 | 0.277 | | 33% | 0.510 | 0.366 | | 31% | 0.511 | 0.353 | |
| PE | | | | | | | | | | | | | |
| <2,000 | 57 | 32% | 0.644 | 0.269 | 0.0536 | 32% | 0.510 | 0.355 | 0.0577 | 28% | 0.497 | 0.343 | 0.0138 |
| 2,000 - 10,000 | 29 | 31% | 0.606 | 0.288 | | 31% | 0.560 | 0.324 | | 31% | 0.601 | 0.303 | |
| 10,000 - 150,000 | 10 | 60% | 0.852 | 0.220 | | 60% | 0.812 | 0.289 | | 60% | 0.837 | 0.271 | |
| Estimated Dry Weather Flow | | | | | | | | | | | | | |
| <100,000 | 63 | 29% | 0.625 | 0.264 | 0.045 | 29% | 0.497 | 0.342 | 0.0207 | 25% | 0.486 | 0.329 | 0.0027 |
| 100,000 - 500,000 | 25 | 40% | 0.651 | 0.307 | | 40% | 0.610 | 0.348 | | 40% | 0.656 | 0.318 | |
| >500,000 | 8 | 63% | 0.889 | 0.170 | | 63% | 0.861 | 0.227 | | 63% | 0.888 | 0.200 | |
| Sewage System | | | | | | | | | | | | | |
| Combined | 39 | 28% | 0.608 | 0.280 | 0.2103 | 28% | 0.487 | 0.356 | 0.0741 | 26% | 0.491 | 0.340 | 0.0657 |
| Separate | 57 | 39% | 0.686 | 0.273 | | 39% | 0.604 | 0.339 | | 37% | 0.613 | 0.330 | |
| Level of Treatment | | | | | | | | | | | | | |
| Secondary treatment | 92 | 34% | 0.644 | 0.277 | 0.0885 | 34% | 0.542 | 0.348 | 0.0885 | 32% | 0.548 | 0.336 | 0.0704 |
| Tertiary treatment | 4 | 50% | 0.892 | 0.125 | | 50% | 0.881 | 0.139 | | 50% | 0.918 | 0.106 | |
| Technologies | | | | | | | | | | | | | |
| Others | 6 | 50% | 0.778 | 0.276 | 0.3619 | 50% | 0.731 | 0.347 | 0.2726 | 33% | 0.642 | 0.356 | 0.7348 |
| Activated sludge | 90 | 33% | 0.646 | 0.277 | | 33% | 0.545 | 0.348 | | 32% | 0.558 | 0.338 | |
| % industrial WW | | | | | | | | | | | | | |
| No activity | 60 | 30% | 0.628 | 0.267 | 0.1144 | 30% | 0.509 | 0.341 | 0.0366 | 27% | 0.508 | 0.330 | 0.0102 |
| ≤ 10% | 26 | 40% | 0.650 | 0.317 | | 40% | 0.599 | 0.366 | | 40% | 0.636 | 0.339 | |
| >10% | 10 | 60% | 0.845 | 0.231 | | 60% | 0.826 | 0.262 | | 60% | 0.852 | 0.238 | |
| N Concentration regulatory limit | | | | | | | | | | | | | |
| Below (≤30 mg/L) | 69 | 39% | 0.668 | 0.291 | 0.6805 | 39% | 0.601 | 0.349 | 0.0419 | 39% | 0.631 | 0.333 | 0.0015 |
| Above (>30 mg/L) | 27 | 22% | 0.619 | 0.239 | | 22% | 0.443 | 0.326 | | 15% | 0.391 | 0.289 | |
| AHP-non-radial | | | | | | | | | | | | | |
| Explanatory factor | Total WWTPs | WPI ₁ | | | | WPI ₂ | | | | WPI ₃ | | | |
| | | % Eff. | Mean | Std. Dev. | Test | % Eff. | Mean | Std. Dev. | Test | % Eff. | Mean | Std. Dev. | Test |
| Year | | | | | | | | | | | | | |
| <1985 | 44 | 36% | 0.692 | 0.277 | 0.1495 | 36% | 0.628 | 0.312 | 0.0434 | 34% | 0.677 | 0.278 | 0.0953 |
| ≥ 1985 | 52 | 33% | 0.609 | 0.297 | | 33% | 0.524 | 0.354 | | 31% | 0.577 | 0.323 | |
| PE | | | | | | | | | | | | | |
| <2,000 | 57 | 32% | 0.634 | 0.286 | 0.1142 | 32% | 0.518 | 0.347 | 0.0242 | 28% | 0.540 | 0.320 | 0.0009 |
| 2,000 - 10,000 | 29 | 31% | 0.605 | 0.291 | | 31% | 0.586 | 0.306 | | 31% | 0.694 | 0.240 | |
| 10,000 - 150,000 | 10 | 80% | 0.840 | 0.251 | | 80% | 0.833 | 0.266 | | 60% | 0.889 | 0.180 | |
| Estimated Dry Weather Flow | | | | | | | | | | | | | |
| <100,000 | 63 | 29% | 0.615 | 0.281 | 0.0777 | 29% | 0.506 | 0.334 | 0.0071 | 25% | 0.539 | 0.308 | 0.0003 |
| 100,000 - 500,000 | 25 | 40% | 0.651 | 0.308 | | 40% | 0.637 | 0.321 | | 40% | 0.741 | 0.240 | |
| >500,000 | 8 | 63% | 0.883 | 0.200 | | 63% | 0.878 | 0.209 | | 63% | 0.916 | 0.156 | |
| Sewage System | | | | | | | | | | | | | |
| Combined | 39 | 28% | 0.581 | 0.295 | 0.0659 | 28% | 0.513 | 0.339 | 0.114 | 26% | 0.575 | 0.308 | 0.2261 |
| Separate | 57 | 39% | 0.691 | 0.280 | | 39% | 0.611 | 0.334 | | 37% | 0.656 | 0.303 | |
| Level of Treatment | | | | | | | | | | | | | |
| Secondary treatment | 92 | 34% | 0.636 | 0.290 | 0.1074 | 34% | 0.557 | 0.337 | 0.092 | 32% | 0.609 | 0.305 | 0.0734 |
| Tertiary treatment | 4 | 50% | 0.899 | 0.144 | | 50% | 0.898 | 0.147 | | 50% | 0.938 | 0.099 | |
| Technologies | | | | | | | | | | | | | |
| Others | 6 | 50% | 0.788 | 0.304 | 0.3078 | 50% | 0.751 | 0.350 | 0.2794 | 33% | 0.674 | 0.358 | 0.8534 |
| Activated sludge | 90 | 33% | 0.637 | 0.288 | | 33% | 0.560 | 0.336 | | 32% | 0.620 | 0.304 | |
| % industrial WW | | | | | | | | | | | | | |
| No activity | 60 | 30% | 0.619 | 0.281 | 0.0727 | 30% | 0.527 | 0.330 | 0.0631 | 27% | 0.569 | 0.307 | 0.0073 |
| ≤ 10% | 26 | 40% | 0.637 | 0.327 | | 40% | 0.615 | 0.345 | | 40% | 0.716 | 0.253 | |
| >10% | 10 | 60% | 0.857 | 0.220 | | 60% | 0.821 | 0.286 | | 60% | 0.870 | 0.222 | |
| N Concentration regulatory limit | | | | | | | | | | | | | |
| Below (≤ 30 mg/L) | 69 | 39% | 0.660 | 0.301 | 0.5191 | 39% | 0.625 | 0.329 | 0.005 | 39% | 0.719 | 0.262 | 0.0001 |
| Above (>30 mg/L) | 27 | 22% | 0.612 | 0.261 | | 22% | 0.434 | 0.327 | | 15% | 0.378 | 0.277 | |

7. Conclusion

In this paper 96 wastewater treatment plants located in Tuscany (Italy) are evaluated through a novel integrated AHP/NDDF approach. The wastewater treatment plants production process is described by the following variables: material and energy costs, staff and maintenance costs and sludge transport and disposal costs are chosen as inputs; treated water and kg of removed sludge are the desirable outputs while the undesirable output is represented by the quantity of nitrogen in the outgoing wastewater. The selection of nitrogen as undesirable output is related to the environmental quality of the outgoing water and adds a new dimension to the literature on the WWTP efficiency analysis.

From a methodological point of view, the paper goes a further step along the path traced by Zhou et al. (2012). The vector directional distance function allows a new formulation for the simultaneous reduction of inputs and bad outputs together with the expansion of good outputs. According to the specification of the non-radial distance function and the explicit directional vector, different combinations of inputs/outputs can be analysed and thus different efficiency indicators can be constructed. The normalized weight vector is selected by taking into account the decision makers' preferences (water utility managers, water authorities) and following the AHP methodology. In this regard, the suggested model encompasses the one proposed by Zhou et al. (2012). In the empirical analysis, two different sets of weights are specified presenting thus two models. In the first case (Non-radial model), the associated weight vector coincides with the one in Zhou et al. (2012). In the second case (AHP-non-radial model), the set of weights is constructed starting from the water utility staff suggestions. The computed environmental indexes differ across the two models, although the efficient units are the same.

The environmental efficiency is explained by means of several variables related to the technical features of the WWTP. Irrespective of the model specifications, the population equivalent size and the estimated dry weather flow have a significant impact on the WWTP performance. This evidence represents a clear indication for water utilities in term of WWTPs' size. On the other hand, whenever the expansion of outputs and the contraction of undesirable output are allowed, the efficiency scores for both the Non-radial and the AHP-non-radial

model are affected by the percentage of the discharged industrial wastewater and by the plants' ability to respect the legal nitrogen concentration threshold. From the environmental agency side, the introduced performance indexes suggest inspection activities to control those plants that treat only domestic sewage and/or do not respect the nitrogen concentration regulatory limit.

As a concluding remark, in this paper the WWTP efficiency is addressed together with the environmental sustainability issue which is specifically related to the quantity of nitrogen in the outgoing water. The analysis of the environmental quality of the treated water might take into account other relevant residuals such as phosphorus, pharmaceutical pollutants, toxic metals and therefore further undesirable outputs might be chosen. In this context, the methodology of this paper might be very promising for further inspections on the environmental efficiency of the wastewater treatment plants.

Acknowledgement

The authors would like to thank the anonymous referees for their valuable comments and suggestions on the earlier versions of the manuscript, Acque SpA and Ingegnerie Toscane Srl for their precious help in collecting data and for their insightful remarks. This research was partially supported by the University of Pisa (PRA 2016 Project).

References

- Adler, N., Volta, N., 2016. Accounting for externalities and disposability: A directional economic environmental distance function. *European Journal of Operational Research* 250, 314–327.
- An, Q., Chen, H., Xiong, B., Wu, J., Liang, L., 2016. Target intermediate products setting in a two-stage system with fairness concern. *Omega* .
- Barros, C.P., Managi, S., Matousek, R., 2012. The technical efficiency of the Japanese banks: non-radial directional performance measurement with undesirable output. *Omega* 40, 1–8.
- Berg, S., Marques, R., 2011. Quantitative studies of water and sanitation utilities: a benchmarking literature survey. *Water Policy* 13, 591–606.

- Chambers, R.G., Chung, Y., Färe, R., 1996. Benefit and distance functions. *Journal of economic theory* 70, 407–419.
- Chambers, R.G., Chung, Y., Färe, R., 1998. Profit, directional distance functions, and Nerlovian efficiency. *Journal of Optimization Theory and Applications* 98, 351–364.
- Chen, C.M., Ang, S., 2016. Measuring Environmental Efficiency: An Application to US Electric Utilities, in: Zhu, J. (Ed.), *Data Envelopment Analysis*. Springer, New York, pp. 345–366.
- Cheng, G., Zervopoulos, P.D., 2014. Estimating the technical efficiency of health care systems: A cross-country comparison using the directional distance function. *European Journal of Operational Research* 238, 899–910.
- Chung, Y.H., Färe, R., Grosskopf, S., 1997. Productivity and undesirable outputs: a directional distance function approach. *Journal of Environmental Management* 51, 229–240.
- Da Cruz, N.F., Marques, R.C., Romano, G., Guerrini, A., 2012. Measuring the efficiency of water utilities: a cross-national comparison between Portugal and Italy. *Water Policy* 14, 841–853.
- Dakpo, K.H., Jeanneaux, P., Latruffe, L., 2016. Modelling pollution-generating technologies in performance benchmarking: recent developments, limits and future prospects in the nonparametric framework. *European Journal of Operational Research* 250, 347–359.
- De Witte, K., Marques, R.C., 2010. Influential observations in frontier models, a robust non-oriented approach to the water sector. *Annals of Operations Research* 181, 377–392.
- De Witte, K., Marques, R.C., 2012. Gaming in a benchmarking environment. a non-parametric analysis of benchmarking in the water sector. *Water Policy* 14, 45–66.
- European Commission, 2007. *Terms and Definitions of the Urban Waste Water Treatment Directive 91/271/EEC*, Brussels.

- Färe, R., Grosskopf, S., 2003. *New directions: efficiency and productivity*. Kluwer Academic Publishers, Boston.
- Färe, R., Grosskopf, S., 2004. Modeling undesirable factors in efficiency evaluation: comment. *European Journal of Operational Research* 157, 242–245.
- Färe, R., Grosskopf, S., 2010. Directional distance functions and slacks-based measures of efficiency. *European Journal of Operational Research* 200, 320–322.
- Färe, R., Grosskopf, S., Lovell, C.K., Pasurka, C., 1989. Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. *The review of Economics and Statistics* , 90–98.
- Färe, R., Grosskopf, S., Pasurka, C., 2016. Technical change and pollution abatement costs. *European Journal of Operational Research* 248, 715–724.
- Färe, R., Grosskopf, S., Pasurka, C.A., 2014. Potential gains from trading bad outputs: The case of us electric power plants. *Resource and Energy Economics* 36, 99–112.
- Førsund, F.R., 2008. Good modelling of bad outputs: pollution and multiple-output production. Technical Report. Memorandum, Department of Economics, University of Oslo.
- Fuentes, R., Torregrosa, T., Ballenilla, E., 2015. Conditional order-m efficiency of wastewater treatment plants: The role of environmental factors. *Water* 7, 5503–5524.
- Fukuyama, H., Weber, W.L., 2009. A directional slacks-based measure of technical inefficiency. *Socio-Economic Planning Sciences* 43, 274–287.
- Golany, B., Roll, Y., Rybak, D., 1994. Measuring efficiency of power plants in Israel by data envelopment analysis. *IEEE Transactions on Engineering Management* 41, 291–301.
- Goodland, R., 1995. The concept of environmental sustainability. *Annual review of ecology and systematics* 26, 1–24.

- Guerrini, A., Romano, G., Mancuso, F., Carosi, L., 2016. Identifying the performance drivers of wastewater treatment plants through conditional order-m efficiency analysis. *Utilities Policy* 42, 20–31.
- Hampf, B., Rødseth, K.L., 2015. Carbon dioxide emission standards for us power plants: An efficiency analysis perspective. *Energy Economics* 50, 140–153.
- Hernández-Sancho, F., Molinos-Senante, M., Sala-Garrido, R., 2011a. Energy efficiency in Spanish wastewater treatment plants: a non-radial DEA approach. *Science of the Total Environment* 409, 2693–2699.
- Hernández-Sancho, F., Molinos-Senante, M., Sala-Garrido, R., 2011b. Techno-economical efficiency and productivity change of wastewater treatment plants: the role of internal and external factors. *Journal of Environmental Monitoring* 13, 3448–3459.
- Hernández-Sancho, F., Molinos-Senante, M., Sala-Garrido, R., Del Saz-Salazar, S., 2012. Tariffs and efficient performance by water suppliers: an empirical approach. *Water Policy* 14, 854–864.
- Hernández-Sancho, F., Sala-Garrido, R., 2009. Technical efficiency and cost analysis in wastewater treatment processes: A DEA approach. *Desalination* 249, 230–234.
- Hsiao, C., Yang, C., Bjornlund, H., 2007. Performance measurement in wastewater control-pig farms in Taiwan. *WIT Transactions on Ecology and the Environment* 103, 467–474.
- Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association* 47, 583–621.
- Kuosmanen, T., 2005. Weak disposability in nonparametric production analysis with undesirable outputs. *American Journal of Agricultural Economics* 87, 1077–1082.
- Liu, J.S., Lu, L.Y., Lu, W.M., 2016. Research fronts in data envelopment analysis. *Omega* 58, 33–45.

- Lorenzo-Toja, Y., Vázquez-Rowe, I., Chenel, S., Marín-Navarro, D., Moreira, M.T., Feijoo, G., 2015. Eco-efficiency analysis of Spanish WWTPs using the LCA+ DEA method. *Water Research* 68, 651–666.
- Metcalf & Eddy, I., Tchobanoglous, G., Burton, F., Stensel, H.D., 2003. *Wastewater Engineering: Treatment and Reuse*. 4 ed., McGraw-Hill Education, New York.
- Moldan, B., Janoušková, S., Hák, T., 2012. How to understand and measure environmental sustainability: Indicators and targets. *Ecological Indicators* 17, 4–13.
- Molinos-Senante, M., Hernández-Sancho, F., Mocholí-Arce, M., Sala-Garrido, R., 2014a. Economic and environmental performance of wastewater treatment plants: potential reductions in greenhouse gases emissions. *Resource and Energy Economics* 38, 125–140.
- Molinos-Senante, M., Hernandez-Sancho, F., Sala-Garrido, R., 2014b. Benchmarking in wastewater treatment plants: a tool to save operational costs. *Clean Technologies and Environmental Policy* 16, 149–161.
- Molinos-Senante, M., Hernández-Sancho, F., Sala-Garrido, R., 2015a. Comparing the dynamic performance of wastewater treatment systems: A metafrontier Malmquist productivity index approach. *Journal of environmental management* 161, 309–316.
- Molinos-Senante, M., Maziotis, A., Sala-Garrido, R., 2016. Estimating the cost of improving service quality in water supply: a shadow price approach for England and Wales. *Science of The Total Environment* 539, 470–477.
- Molinos-Senante, M., Sala-Garrido, R., Lafuente, M., 2015b. The role of environmental variables on the efficiency of water and sewerage companies: a case study of Chile. *Environmental Science and Pollution Research* 22, 10242–10253.
- Murty, S., 2010. On the theory of a firm: the case of by-production of emissions. Technical Report. University of Warwick. Dept. of Economics.

- Murty, S., Russell, R.R., Levkoff, S.B., 2012. On modeling pollution-generating technologies. *Journal of Environmental Economics and Management* 64, 117–135.
- OECD, 2001. OECD environmental strategy for the first decade of the 21st century. Paris .
- Pakkar, M.S., 2015. An integrated approach based on DEA and AHP. *Computational Management Science* 12, 153–169.
- Picazo-Tadeo, A.J., Reig-Martinez, E., Hernandez-Sancho, F., 2005. Directional distance functions and environmental regulation. *Resource and Energy Economics* 27, 131–142.
- Picazo-Tadeo, A.J., Sáez-Fernández, F.J., González-Gómez, F., 2008. Does service quality matter in measuring the performance of water utilities? *Utilities Policy* 16, 30–38.
- Picazo-Tadeo, A.J., Sáez-Fernández, F.J., González-Gómez, F., 2011. Assessing performance in the management of the urban water cycle. *Water Policy* 13, 782–796.
- Pomerol, J.C., Barba-Romero, S., 2000. *Multicriterion decision in management: principles and practice*. Kluwer Academic Publishers, New York.
- Romano, G., Molinos-Senante, M., Guerrini, A., 2017. Water utility efficiency assessment in Italy by accounting for service quality: An empirical investigation. *Utilities Policy* 45, 97–108.
- Ruxton, G.D., Beauchamp, G., 2008. Some suggestions about appropriate use of the Kruskal–Wallis test. *Animal Behaviour* 76, 1083–1087.
- Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology* 15, 234–281.
- Saaty, T.L., 1990. How to make a decision: the analytic hierarchy process. *European Journal of Operational Research* 48, 9–26.
- Sala-Garrido, R., Hernández-Sancho, F., Molinos-Senante, M., 2012a. Assessing the efficiency of wastewater treatment plants in an uncertain context: a DEA with tolerances approach. *Environmental Science & Policy* 18, 34–44.

- Sala-Garrido, R., Molinos-Senante, M., Hernández-Sancho, F., 2011. Comparing the efficiency of wastewater treatment technologies through a DEA metafrontier model. *Chemical Engineering Journal* 173, 766–772.
- Sala-Garrido, R., Molinos-Senante, M., Hernández-Sancho, F., 2012b. How does seasonality affect water reuse possibilities? An efficiency and cost analysis. *Resources, Conservation and Recycling* 58, 125–131.
- Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics* 136, 31–64.
- Sueyoshi, T., Goto, M., 2012a. Data envelopment analysis for environmental assessment: comparison between public and private ownership in petroleum industry. *European journal of operational research* 216, 668–678.
- Sueyoshi, T., Goto, M., 2012b. Dea environmental assessment of coal fired power plants: Methodological comparison between radial and non-radial models. *Energy Economics* 34, 1854–1863.
- UN General Assembly, 2002. Report of the world summit on sustainable development, johannesburg, south africa, 26 august–4 september 2002. New York: United Nations .
- UN General Assembly, 2015. Transforming our world: the 2030 agenda for sustainable development. New York: United Nations .
- Worthington, A.C., 2014. A review of frontier approaches to efficiency and productivity measurement in urban water utilities. *Urban Water Journal* 11, 55–73.
- Zhang, N., Choi, Y., 2014. A note on the evolution of directional distance function and its development in energy and environmental studies 1997–2013. *Renewable and Sustainable Energy Reviews* 33, 50–59.
- Zhou, P., Ang, B., Wang, H., 2012. Energy and CO₂ emission performance in electricity generation: a non-radial directional distance function approach. *European Journal of Operational Research* 221, 625–635.

Appendix - Some essential notions of wastewater treatment

To understand the background of the present analysis, a short description of the wastewater treatment is provided. Due to the great extent of the subject, only the information strictly related to those technical features, explicitly recalled in the previous sections is provided. For a comprehensive and detailed presentation the reader can refer for example to Metcalf & Eddy et al. (2003). Wastewater may be defined as a mix of liquid or water-borne wastes discharged by houses, commercial properties, factories, farms and public institutions. Ground, surface and stormwaters may be also included as components of the wastewater flow. As a preliminary step, the wastewater is collected in sewers and then conveyed to treatment or disposal facilities. Regarding the sewerage, two different systems can be identified: combined or separate sewerage. The first one transports both stormwater and wastewater, while separate sewerage is designed either to convey wastewater (sanitary sewers), or to drain surface runoff (storm sewers). For both combined and separate sewerage, the estimated dry weather flow is a relevant value. In fact, during the period of dry weather, infiltration and surface runoff have a minimum influence in combined sewerage. Therefore the estimated dry weather flow provides a basis for works design. The wastewater entering in the treatment plant is called influent or ingoing water and it is characterized by the presence of several physical, chemical and biological constituents such as suspended solids (Ss), nitrogen (N), phosphorus (P), carbon (C) and biodegradable organics; the latter is usually measured in terms of biochemical oxygen demand (BOD) and of chemical oxygen demand (COD). Together with constituents, pathogens and priority pollutants (e.g. heavy metals, pharmaceutical molecules) can be found in the influent wastewater. One of the main objective of a wastewater treatment plant (WWTP) should be the removal of as many contaminants as possible and the purification of the water for further reuse. A suitable wastewater treatment should guarantee an acceptable level of overall water quality and this can be done by using different methods. The treatment methods which exploit physical phenomena are referred as Unit Operations (UO); on the other hand, if the removal of contaminants is based on chemical reactions, then the process is called Unit Process (UP). To ensure a certain level of contaminant removal, unit operations and processes

are jointly performed according to different wastewater treatment technologies:

- preliminary treatment removes gross solids (large objects, rags, grit);
- primary treatment eliminates floating and settable materials;
- advanced primary treatment allows the removal of suspended solids;
- secondary treatment eliminates organic contaminants through biological and chemical processes which are carried on;
- tertiary treatment eliminates other pollutants that cannot be removed by means of primary and secondary treatments.

The secondary treatment removes also nitrogen and phosphorus (in this context they are also referred as nutrients) together with other pathogens and some heavy metals (Metcalf & Eddy et al., 2003). Among the different technologies that can be used in the secondary treatment, it is worth mentioning the activated sludge process. This can be considered as the main biological process in secondary treatment and it basically refers to a mass of microorganisms metabolizing the suspended and soluble matter in an aeration basin. Solids and in particular biosolids can be considered by-products of the wastewater treatment; they are often referred to as “sludge” and they are by far the largest removed constituents. After a primary treatment, solids can be further biologically, chemically or by heat treated (e.g. stabilization, composting, dewatering, drying, thickening) so to get them suitable for further reuse (agriculture, home gardens...). The term biosolids indicates that solids are further treated and an important distinction between class A biosolids and class B biosolids has to be made (Metcalf & Eddy et al., 2003). Biosolids belonging to the first class are also known as “clean sludge” while Class B biosolids have a reduced concentration of pathogens and other unhealthy contaminants (mainly metals), but they do not satisfy specific legal requirements and therefore their application to land is strictly regulated. Sludge which is not eligible for further use, is then transported to either landfill or incinerators. Moreover, among the other constituents, nitrogen occupies a preeminent position in wastewater treatment activities; excessive concentrations of nitrogen can actually be harmful to humans and wildlife. Nitrogen can be found in the wastewater under various forms, namely organic nitrogen and inorganic nitrogen which is in turn divided into ammonia nitrogen, nitrite nitrogen and nitrate nitrogen. Ammonia con-

centrations can affect hatching and growth rates of fish. If excessive amounts of nitrates are discharged into the aquatic environment, it can lead to the growth of undesirable aquatic life and then to the eutrophication. Nitrate can even affect human health if it is present in drinking water. Moreover, a great discharged of total nitrogen onto land can lead to the pollution of groundwater, causing excessive vegetative growth and a reduction of crop quality. Due to this, many alternative technologies have been designed to remove total nitrogen from wastewater (suspended growth nitrification and denitrification variations, attached growth nitrification and denitrification variations, biological nutrient removal variations). On the other hand, it is almost impossible to completely remove nitrogen from wastewater and it is still present in the effluent flow.