ALDA ALMENDRA HENRIQUES

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

A sustainable energy system approach for the wastewater sector

Thesis submitted to Faculdade de Engenharia da Universidade do Porto in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Sustainable Energy Systems, supervised by Ana Camanho, Associate Professor of Faculdade de Engenharia da Universidade do Porto, and Pedro Amorim, Associate Professor of Faculdade de Engenharia da Universidade do Porto

Faculdade de Engenharia da Universidade do Porto

2021

This PhD thesis was supported by FCT – Fundação para a Ciência e a Tecnologia (Portuguese national funding agency for science, research and technology), with a funding allocated through the Grant PD/BD/142815/2018 respecting the Doctoral Program on Sustainable Energy Systems by MIT-Portugal. It was also supported by the Centre of Industrial Engineering and Management of INESC TEC (https://www.inesctec.pt/en), and by the Portuguese water company AdCL – Águas do Centro Litoral (https://www.aguasdocentrolitoral.pt/).

This PhD thesis was developed under the framework of the MIT Portugal program (https://www.mitportugal.org/).

“Grow with discipline. Balance intuition with rigor. Innovate around the core. Don’t embrace the status quo. Find new ways to see. Never expect a silver bullet. Get your hands dirty. Listen with empathy and overcommunicate with transparency. Tell your story, refusing to let others define you. Use authentic experiences to inspire. Stick to your values, they are your foundation. Hold people accountable, but give them the tools to succeed. Make the tough choices; it’s how you execute that counts. Be decisive in times of crisis. Be nimble. Find truth in trials and lessons in mistakes. Be responsible for what you see, hear, and do. Believe.”

Howard Schultz

Acknowledgements

In this section, I would like to acknowledge the assistance and support of any kind that each and every person mentioned below gave me at some point or throughout the entire project. Without them, I wouldn't have made it the way it is or even had started it. Regardless of their names will appear again in this document, I believe they know exactly the way they contributed to this project so it could have reached a tangible form!

Nevertheless, I owe a very special thanks to my supervisor, Professor Ana Camanho, and to my co-supervisor, Professor Pedro Amorim. It was a very long journey, filled with ups and downs, as well as challenges of all sorts. They were always doing their job in the best way they could, sharing their knowledge, suggesting, and inspiring me to keep doing what I was supposed to do, even though in some moments it has not been straightforward the way we met our wills halfway. Regardless of any less positive note, I must say that I learned so much with them both, and I deeply appreciate all the support that they gave me! Thank you so much for everything and I apologize for anything less than desirable on my end!

I apologize to those people who might have also helped me in some way related to the thesis but who I eventually have forgotten!

Thank you so much to all of you!

Professor Ana Camanho

Professor Pedro Amorim

Professor Vítor Leal

Engineer Jaime Gabriel Silva

Engineer Milton Fontes

Professor Manuel Matos

Sara Martins

Maria João Santos

Giovanna D'Inverno

Behrouz Arabi

Célia Couto

Hugo Santos

Zenaida Mourão

Sofia Cruz Gomes

Flávia Barbosa

My parents.

Abstract

This thesis project is composed of five research papers under the broad topic of wastewater organizational structures in what respects some of their underlying management aspects. The main objective is to contribute to increased sustainability in the wastewater sector through the improvement of energy management. This research was motivated by the general gaps identified in the literature review and presented in the introduction section of this thesis. More specific literature reviews are also embedded in the papers that constitute the Part II of this thesis.

The first paper (presented in Chapter 3) concerns the development of a composite indicator (CI) to aggregate all the individual performance indicators used by the regulator to assess the quality of service provided by the wastewater operators to users. The CI gives both the regulator and the operators relevant information about operators' performance that the individual indicators cannot provide. This methodology can be useful for acting at strategic and operational levels in the areas of sustainability that are more in need of improvement, as well as motivating policy developments.

The second paper (presented in Chapter 4) refers to the development of a logistics network to promote sludge transfers from wastewater treatment plants (WWTPs) without anaerobic digestion infrastructure to WWTPs with anaerobic digesters. The viability of the logistics network is confirmed through a positive and optimal economic benefit obtained from the overall energy balance. This occurs because the sludge transfers are expected to contribute both for a reduction in the electricity consumption at the first-type plants and for a significant increase in the overall biogas production at the second-type plants.

The third paper (presented in Chapter 5) consists of a quantitative benchmarking approach that is applied to a set of WWTPs to derive insights on the relative performance of the plants, on the overall potential for resource consumption reduction and/or reallocation, and the individual targets for improvement based on the identified benchmarks for each individual plant. This paper also considers an analysis of the contextual factors that may be impacting the plants' performance and thus contributing to explain some of the observed differences among plants. The major particularity is that this sample of plants is problematic, which means that it does not have the ideal characteristics for the application of benchmarking techniques.

The fourth paper (presented in Chapter 6) consists in the assessment of productivity changes over a period of four years. It is also made an analysis of potential factors, particularly rain, that may be influencing the productivity variations observed at the WWTPs.

The fifth paper (presented in Chapter 7) concerns the development of an optimization model to search for potential energy efficiency improvements and simultaneously consider the desirable treatment effectiveness attainment. It adopts an innovative perspective of a system-level optimization that allows for emission quotas trade-offs among WWTPs.

In order to assist the readers that are less familiar with the most common methodologic approaches used in the research papers, the thesis also provides, in the Introduction section, a brief introductory overview of frontier techniques for performance assessment, with specific focus on data envelopment analysis. Also, this thesis adopts the paradigm "think globally, act locally". This is specifically translated in the thesis structure that goes from very broad and general views upon the wastewater sector, which are mainly drawn in the Introduction section, toward case-study applications of the methodologic approaches and rationales envisioned in each research paper. Thus, each paper explores in detail a specific theme to propose solutions to real-world problems that are aligned with the broader research contexts previously identified.

Keywords: Wastewater treatment, Data Envelopment Analysis (DEA), Performance measurement, Mixed Integer Linear Programming (MILP), Energy efficiency, Energy recovery from sewage sludge

Resumo

A presente tese é composta por cinco artigos de investigação no âmbito do tema das estruturas organizacionais de águas residuais, particularmente no que se refere a aspetos associados à gestão. O objetivo principal é desenvolver uma contribuição para a melhoria da sustentabilidade no setor das águas residuais através da melhoria da gestão de energia. Para isto, serão consideradas as principais lacunas de investigação identificadas, quer na revisão de literatura que está presente na secção de introdução da tese, quer integrada nos artigos de investigação apresentados na Parte II da tese.

O primeiro artigo (apresentado no Capítulo 3), refere-se ao desenvolvimento de um indicador compósito (IC) que permite agregar todos os indicadores individuais de desempenho utilizados pelo regulador para avaliar a qualidade do serviço prestado pelos operadores aos utilizadores. O IC transmite informação relevante sobre o desempenho dos operadores que é útil quer para o regulador quer para os próprios operadores, sendo que esta informação não é possível de obter quando considerados somente os indicadores individuais de desempenho. A informação proporcionada pelo IC pode ser utilizada tanto ao nível do planeamento estratégico, como ao nível do planeamento operacional, principalmente nas áreas da sustentabilidade onde se observar maior necessidade de melhoria, assim como motivar novos desenvolvimentos das políticas do setor.

O segundo artigo (apresentado no Capítulo 4) refere-se à modelação de uma rede logística para promover transferências de lama proveniente de estações de tratamento de águas residuais (ETARs) que não têm a estrutura necessária à utilização da digestão anaeróbia para as ETARs com digestores anaeróbios. A viabilidade da rede logística é garantida através do benefício económico positivo e ótimo que se obtém a partir do balanço energético total associado à operação da rede logística, uma vez que é expectável que as transferências de lama contribuam tanto para a redução do consumo de eletricidade nas ETARs do primeiro tipo, como para o aumento significativo da produção de biogás nas ETARs do segundo tipo.

O terceiro artigo (apresentado no Capítulo 5), consiste numa abordagem de benchmarking quantitativo que é aplicado a um conjunto de ETARs para a obtenção de informação relevante, tanto acerca do desempenho relativo dessas ETARs, como acerca do potencial total para melhoria na redução do consumo de recursos e/ou realocação de recursos, e ainda acerca da definição de objetivos individuais de melhoria tendo como base a identificação das melhores práticas para cada ETAR individual. Este artigo também considera uma análise dos fatores contextuais para se determinarem possíveis impactos no desempenho das ETARs e, desse modo, auxiliar na justificação de algumas diferenças de desempenho observadas entre as ETARs. A maior particularidade deste artigo é a questão de a amostra de ETARs utilizada como caso de estudo ser considerada problemática, o que significa que a mesma não reúne todas as condições ideais para a aplicação de métodos de benchmarking quantitativo.

O quarto artigo (apresentado no Capítulo 6), consiste na medição de variações na produtividade observada ao longo de um período de quatro anos. Também é feita uma análise de fatores, nomeadamente a chuva, que poderão estar na origem de algumas variações de produtividade observadas nas ETARs.

O quinto artigo (apresentado no Capítulo 7), refere-se ao desenvolvimento de um modelo de otimização que permite encontrar potenciais melhorias que conduzam a uma utilização de energia mais eficiente e que simultaneamente considera que todas as ETARs garantem a eficácia do tratamento desejada. Neste artigo, considera-se uma perspetiva inovadora de otimização ao nível do sistema que permite a existência de *trade-offs* de quotas de emissão entre as ETARs.

De modo a auxiliar os leitores menos familiarizados com as técnicas metodológicas utilizadas nos artigos, é também facilitada uma breve introdução a métodos de fronteira para medição de desempenho, enfatizando concretamente a técnica de *data envelopment analysis*. Para além disso, a tese foi desenvolvida com base no paradigma "pensar globalmente para agir localmente". Este aspeto é particularmente visível através da gradação observada a partir de uma visão mais globalizante sobre os principais assuntos tratados na tese, e que são tratados principalmente na secção da introdução, até à aplicação mais concreta das metodologias e estruturas de investigação aos casos de estudo em cada artigo de investigação. Por sua vez, os temas mais concretos referentes a problemas específicos tratados nos artigos, encontram-se alinhados com os contextos de investigação mais alargados e previamente identificados.

Palavras-chave: Tratamento de águas residuais, *Data Envelopment Analysis* (DEA), Medição de desempenho, *Mixed Integer Linear Programming* (MILP), Eficiência energética, Recuperação de energia contida nas lamas de ETAR

Contents

PART I

INTRODUCTION

Chapter 1 –

Motivation and Research Rationale

Chapter 1

Motivation and research rationale

This chapter is structured into three sub-chapters to bring clarity into the overall context of this thesis project. First, it will be presented a general overview of the thesis' topic so as to identify the major problem to tackle, explain the motivation for delivering the study, and thus be able to derive the thesis' objectives. Second, it will be conducted a motivational literature review to better understand how to position this thesis' contribution. Third, from the information previously gathered and discussed, it will be delineated a thesis rationale, then formulated the research questions, and finally, it will be explained the overall thesis structure.

1.1.Motivation and objectives

The purpose of this thesis is to develop a contribution for the wastewater sector rooted in a sustainable energy system perspective through management approaches. The topic of energy is inherent to all sectors' functioning, but it has particular importance in the case of the wastewater sector due not only to the fact that the industrial activity carried out is energy intensive (Gurung et al., 2018; Daw et al., 2012; Silva et al., 2015; U.S. EPA, 2013) but also because of the high amounts of embedded energy in the raw wastewater (Jenicek et al., 2013; Qadir et al., 2020). While the former fact calls up to an energy efficiency perspective, the latter one, as it considers the raw wastewater a renewable energy source (Nowak et al., 2015), calls to an appropriate exploration of the associated recovery potential. This is therefore useful to balance out at least part of the energy demand required for wastewater treatment. In reality, the potential for energy recovery from wastewater can go even beyond the total energy demand for the treatment processes (Gikas, 2017; Nowak et al., 2015). For these reasons, the energy topic will be the primary focus of this thesis project. However, this will be included in a broader context within the wastewater sector's functioning improvement. Therefore, the problem we aim to tackle is the following:

How to expand the frontier of knowledge to further propel the current research state-of-the-art toward the successful transformation of the wastewater sector into a fully sustainable part of the society's existence, organization, and activity?

The wastewater sector is very complex in many ways. Particularly, this complexity can be disaggregated into several layers, such as: (i) the numerous utilities in direct relationship with other entities in the context of business models along with the interconnectedness with other sectors and industrial activities (Garrido-Baserba et al., 2020; Grigg, 2016); (ii) the large infrastructure that interconnects strongly with the urban environment (Ampe et al., 2020; Brown et al., 2009); (iii) the technology diversity and complexity (Maktabifard et al., 2018); (iv) the uncertainty regarding the volumes of wastewater to treat alongside its composition (i.e., type of pollutants, the share of each pollutant, and the proportion in terms of the amounts that are present in the influent wastewater), and in relation to the amounts of rainfall (Eggimann et al., 2017); (v) the considerable diversity of stakeholders, such as regulators, suppliers, managers, collaborators, service users and other customers, who tend to have conflicting interests (Ampe et al., 2020; Burkhard et al., 2000; Iribarnegaray et al., 2012); (vi) the interaction with the environment and society in a broader sense, namely toward the fulfillment of sustainable development goals determined by the United Nations (Qadir et al., 2020); (vii) the interconnectedness of the wastewater sector with the water sector as a whole (Burkhard et al., 2000; Grigg, 2016; Nguyen et al., 2019; Van der Brugge and Rotmans, 2007); and (viii) the numerous priorities to give response to while fulfilling the service provision expected at different dimensions, for example, energy-related functions, equipment verification and maintenance, operations optimization to reduce costs, fulfilling regulator's collaborative requirements, activity assessment and planning, among others (Daw et al., 2012).

Furthermore, the wastewater treatment process itself is of a multidisciplinary nature (Garrido-Baserba et al., 2020). Such multidisciplinary nature stems primarily from the physical, chemical and biochemical characteristics of the raw sewage, which in turn demands for complex engineering approaches to carry out the corresponding treatment processes in parallel with adequate infrastructure and management expertise. As we enter the new era of digitalization and use of artificial intelligence (Sharma, 2019), the multidisciplinary nature of the wastewater sector broadens even more due to the computer science and data analytics areas of knowledge that need to be brought into the equation (Garrido-Baserba et al., 2020).

Therefore, the first necessary step in this thesis work is to clarify the boundaries within which the research developed is focused. For that, let us consider an initial observation position upon all dimensions and layers that build up into the complex topic of the wastewater sector.

The first boundary identified refers to the definition of the thesis' study-object (Figure 1.1). In other words, I will identify the constituent parts of the wastewater sector, either physical or organizational, that will be specifically focused throughout this thesis. The thesis' study-object thus comprehends the regulator and the utilities, particularly in what concerns the regulation of the utilities' service provision. It also comprehends the assets, particularly the wastewater treatment plants' (WWTPs) infrastructure. In this latter case, the focus will be specifically on strategic and operational management.

In regards to the first boundary just defined, which denote a certain hierarchical culture within the wastewater sector, the approaches taken along this thesis will assume the perspective from the stakeholders that are directly involved. This will also imply to acknowledge the scope of action and mission of those stakeholders.

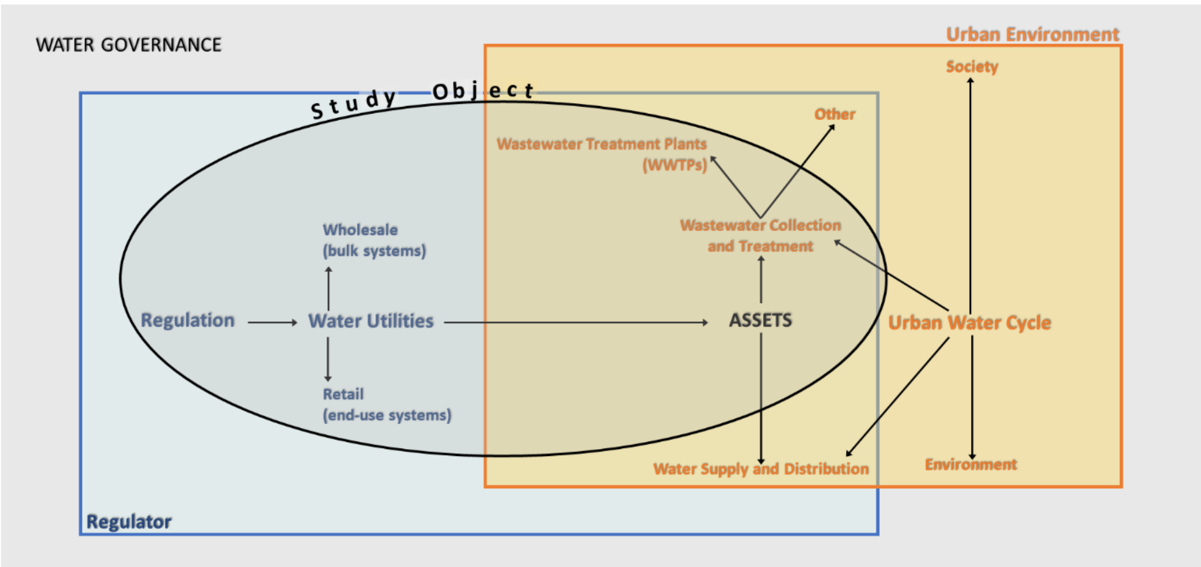


***Figure 1.1** – Thesis' study object.*

As previously discussed, there is a considerable complexity that radiates dispersedly from the concept of *wastewater sector* through many branches of research fields (Figure 1.2). Therefore, in addition to the hierarchical culture that underpins the wastewater sector in the current era, as well as its connection to the urban environment through the inherent impact on the environment and society, an overview of the discussion related to the wastewater sector is the position taken to define the second boundary for the thesis research. In this regard, the following topics were found to be of meaningful interest in the scope of the wastewater sector, which will be further explored in the motivational literature review section to allow for the design of the research questions: (i) sustainable development, sustainability, and circular economy; (ii) water-energy nexus; (iii) management, performance measurement, and regulation; (iv) technology; (v) fourth water revolution and artificial intelligence environment.

Although this thesis work may be small and easily overlooked, it will still be attempted a humble contribution for the wastewater sector improvement and evolution. The final goal is to enable improvements in resource utilization and in process effectiveness, along with providing alternative perspectives and overviews at the issues studied that hopefully might serve as a starting point for future research works.

In this context, the thesis contribution will relate with research already developed within those two boundaries above identified. I bear in mind the interconnectedness that ideally should encompass the wastewater sector throughout all its complexity (Burkhard et al., 2000; Eggimann et al., 2017; Qadir et al., 2020), and, therefore, a system-wide vision will be adopted as the first step toward integrated approaches and unification among complexity.

Taking the information above exposed as a motivation and starting point, the objectives of this thesis project are the following:

(i) To conduct a motivational literature review and analyze the current research trends within the wastewater sector, particularly concerning the available options leading to improved sustainability;
(ii) To assert alternative thinking perspectives worth considering to tackle the identified problem;
(iii) To develop alternative, innovative, and complementary approaches to those already existent;
(iv) To promote and contribute to increased transparency within the wastewater sector;
(v) To promote increased engagement of all parties involved in the wastewater sector;
(vi) To derive information that can be useful for policy improvements.

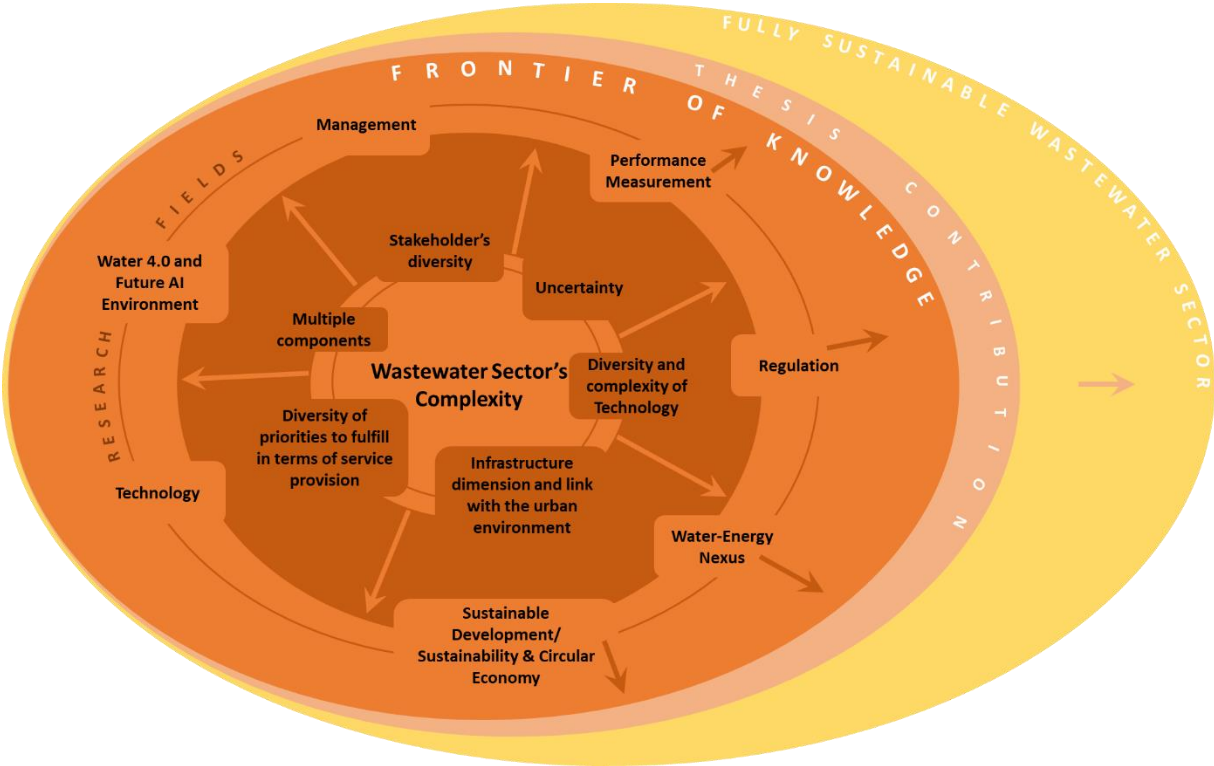


Figure 1.2 *- Shape of the frontier of knowledge stemming from the complexity of the wastewater sector, as a starting point to define the scope for the thesis contribution.*

1.2. Motivational literature review

The wastewater sector is a research field that justifies placing attention on the inherent challenges and the discussion emergent from and surrounding those challenges. In parallel to other sectors, such as the energy sector (Rotmans, 2001), the construction sector (Killip et al., 2018), and the water sector (Van der Brugge and Rotmans, 2007), to give some examples, the wastewater sector is currently undertaking a transition phase into a more sustainable one (Ampe et al., 2020; Eggimann et al., 2017; Moy de Vitry et al. 2019; Ujang and Buckley, 2002).

As highlighted by Loorbach (2010) and Van der Brugge and Rotmans (2007), a transition phase is what is necessary for a sector to undergo in order to cope with persistent problems arising from the outward conditions. These authors also defend that when persistent problems challenge the current societal systems' way of functioning and responding, a transition management needs to be adopted as a set of integrative and complex approaches able to propel them into a new and sustainable form of the previous ones within a long-term and irreversible process that can last for more than two decades. The persistent problems faced nowadays, such as water scarcity, increased energy costs, public health, urban development and climate change (Cantor et al., 2021; Eggimann et al., 2017; Ujang and Buckley, 2002), stem mainly from the same root, which is the world population growth. However, the population growth only highlighted that the past, and, to some extent, still current approaches of exploring energy, water and other resources that are necessary to guarantee the basic quality of life of human populations, were not sustainable. Therefore, new strategies underpinning broader and integrative perspectives upon existing activities throughout societal organizational sectors are already under development and starting to become structural in society in general. International drivers such as the sustainable development goals (UN General Assembly, 2015) steer the change and transition that countries all over the globe, organizations, institutions, and individuals are coming to willingly, collaboratively, and creatively deploy within the horizon of 2030.

As discussed by Van der Brugge and Rotmans (2007), a set of transition principles should be considered and put into practice in order for an effective change of paradigm to take place and a desirable goal be achieved. In other words, when put into practice, a set of transition principles allow challenges to be overcome to the best interest of all society representative members relating to those problems at hand. A multi-level governance across the related sectors, with mutual reinforcement of objectives and instruments as well as the collaborative participation of stakeholders for the same end-goal are some of the advocated management principles underpinning a transition phase. Transparency and information fluidity are ingredients that promote participation and collaboration of stakeholders and should therefore be promoted along with the instruments used. Another relevant feature is the natural promotion of decentralized decision-making, so that an increased involvement, participation and collaboration of stakeholders may directly align toward the same end-goal along with coherent policy developments (Loorbach, 2010). This comes along with increased diversity, creativity and experimentation, being all of those welcome drivers of disequilibrium that characterizes the dynamics of a transition phase (Van der Brugge and Rotmans, 2007).

In regards to the wastewater sector, the challenges faced during this transition phase are different across countries (Grigg, 2016). Accordingly, in many developing countries, the current focus is stronger at the design and implementation of new sewage systems along with policy development (Marcos von Sperling and Chernicharo, 2002; Murray et al., 2011; Nguyen et al., 2019; Ujang and Buckley, 2002), whereas in the majority of developed countries, the current effort is predominantly concentrated in the maintenance and retrofitting of aging infrastructure, technology updates, as well as increasing the effluent water quality requirements (Cantor et al., 2021; Gandiglio et al., 2017; Gurung et al., 2018; Zayed et al., 2015). Moreover, as pointed out by Abbott and Cohen (2009), challenges related to environmental management activities, such as water conservation, the role of wastewater as a potential source of reusable water (alternative water supply), and the relationship between water supply and urban planning, are yet to be fully overcome. Also, according to the authors, a good approach taken in tackling the referred challenges may contribute to efficiency improvements in the water and wastewater systems, although with the need of considering underlying structural implications.

One major requirement, that constitutes another challenge for the wastewater sector, is the development of reliable computing and data management systems to extract the information needed as the basis for more integrative decision-making and policy improvements and redesign. As discussed in Eggimann et al. (2017), data-driven approaches are a current trend in face of the exponential data availability in the wastewater sector. Big data in the wastewater sector is a consequence of the advancements in sensoring technology that is applicable with the purpose of process effectiveness and efficiency improvements as well as monitoring functions. Furthermore, as summarized by Garrido-Baserba et al. (2020), big data analytics and artificial intelligence constitute a doorway to a new and comprehensive approach to cope with the challenges faced within the wastewater sector at different levels, such as operation, management, and rehabilitation, and even promote interrelations for new business development (related to resource recovering but also to technology development) as well as circularly bonding up with the water and health sectors. In this regard, one new interesting aspect enabled by the use of big data analytics and artificial intelligence is the ability to obtain information about pathogen monitoring and prevention of epidemics, human health and life-style characteristics (Eggimann et al., 2017; Garrido-Baserba et al., 2020).

Big data analytics and artificial intelligence are tools to wisely use toward increased organic behavior among the components and dimensions of the wastewater sector and a more integrated vision and comprehensive decision making. Certainly, those are tools that will be known in association with the current transition phase of the wastewater sector in the pursuit of full sustainability worldwide (Garrido-Baserba, 2020). But the inertia of the wastewater sector towards overcoming the transitioning process worldwide is too big, and leveraging the transformation process to the actual materialization of the desired results in every corner of the world is complex and it takes time for the culture, societal structures and roles, as well as public acceptance and financial institutional support to be changed for good. As highlighted by Loorbach (2010), transitions are built up of activities occurring at different levels namely, strategic, tactical, operational, and reflexive. The strategic level is mainly characterized by vision development, strategic discussions and long-term goal formulation. The tactical level relates to structures, such as regulation, institutions, organizations, and infrastructure. The operational level is characterized by experiments within innovation projects or entrepreneurial skills of groups or individuals that are in alignment with the strategic and tactical levels' goals and mission. The reflexive level comprehends monitoring, assessing and evaluating ongoing policies and related actions to assert their appropriateness and to enable the exploration of untapped knowledge pathways. This reflexive level is conducted by researchers and public opinion through the media and internet.

Considering the ideas above, every effort that aims to contribute to achieve the desired goal of a fully sustainable wastewater sector is welcome in this transition phase. Particularly, there are many studies developed with a high level of integrated knowledge that is a required ingredient for the sustainability goal (Garrido-Baserba et al., 2020), although with the intent of tackling, understanding or systematizing different types of problems. Those studies should be valuably considered, such as, for example, those of Admiraal and Jan van Helden (2003), Ampe et al. (2020), Cantor et al. (2021), Eggimann et al. (2017), Garrido-Baserba et al. (2020), Grigg (2016), Iribarnegaray et al. (2012), Nguyen et al. (2019), and Picazo-Tadeo et al. (2011). It could be argued that these studies fall into the category of the operational and reflexive levels above mentioned.

Another important aspect to keep in mind is the concept of circular economy (CE) as a model of consumption and production at a micro-level, but which can be scaled up into a form of socio-economical structure, organization and functioning to accomplish sustainable development (Sauvé et al., 2016). A CE paradigm requires integrative perspectives upon systems, which is in alignment with the transition phase requirements and can be implemented at different levels (i.e., micro, meso and macro), depending among other aspects, on the number of participants (Suárez-Eiroa et al., 2019). As already displayed above, the wastewater sector does not operate in isolation. One example of integrative perspective among sectors at a meso-level is the one of industrial symbiosis (Lehtoranta et al., 2011). The concept of industrial symbiosis corresponds to frameworks built up with the objective of enabling mutual economic benefits among industrial partners across different sectors. This concept can be used within a framework of CE to amplify the life-cycle of products, extract additional value from end-life products, and reduce the overall negative environmental impact of the overall production activities. One example of this type of industrial symbiosis that can be further explored that relates to the wastewater sector is co-digestion of organic by-products and sewage sludge (Gandiglio et al., 2017; Gu et al., 2017) to take place in adequate infrastructure (anaerobic digesters) existing at some wastewater treatment plants. Another example is the use of thermal energy from the wastewater treatment by nearby buildings for heating purposes (Gu et al., 2017), and yet another example is the use of reclaimed water (Abbott and Cohen, 2009; Voulvoulis, 2018).

Qadir et al. (2020) refer to the wastewater produced worldwide, emphasizing it as a source for resource recovery, including water, nutrients and energy. In this context, the authors point out the situation of the untreated wastewater in many developing countries as a matter deserving resolution so that economic and financial benefits can be derived, specifically when considered integrative perspectives to favor the paradigm of wastewater as a source of resources worth exploring. Equally important are the data collecting mechanisms at different level or aggregation (e.g., regional, national) to better guide decision- and policy-makers in the development of adequate frameworks and measures with the purpose of improving resource recovery from wastewater. The authors place the wastewater resource exploration as a strategy to cope with the challenges associated with attaining some of the sustainable development goals (UN General Assembly, 2015). Actually, one fundamental idea that seems to be gaining acceptance concerns the so-called wastewater treatment plants (WWTPs). More specifically, as water is becoming increasingly scarce, even the treated wastewater is being considered for specific end-uses (the so-called reclaimed water, Voulvoulis, 2018). As such, and also because WWTPs are progressively managed to recover nutrients and energy for self-sufficiency and sustainability purposes, this type of wastewater sector core facilities is equally progressively being instead known as water resource recovery facilities (WRRFs) (Garrido-Baserba, 2020; WEF, 2013).

As mentioned previously in this chapter, energy is a core element to be taken into consideration in the wastewater sector. In the words of Ibrahim Dincer (Dincer, 2000), "energy is the convertible currency of technology". In fact, energy is abundant either in the Sun as well as the surface and internal layers of the Earth, not to mention the energy that exists in the entire galaxy and the universe. It is the available technology that lends the ability to capture the energy in its useful form. As the technology is the hardest part to develop and to economically sustainably deploy (Dincer, 2000), it is attached to policies, governance and public acceptance. Nowadays, it is widely recognized that the renewable energy forms are more suitable to cope with the persistent problems related to the security of energy supply along with an environmentally friendly footprint in their use than the widespread adoption of fossil fuels (Dincer, 2000). The wastewater itself is an energy carrier, both thermal and chemical energy, and it is a source of renewable energy that is worth being explored, more so if considered in the scope of the water-energy nexus (IEA, 2018). The water-energy nexus refers to the close relationship that exists between water and energy, which derives from the exploration and use of both natural resources (Hamiche et al., 2016). This interrelation extends to the wastewater. In fact, in order for the wastewater to be collected and transported through kilometers of pipes to the treatment facilities and from there to the receiving natural watercourses, energy is required for pumping purposes. Similarly, in some of the more common wastewater treatment technologies, such as the conventional activated sludge, high amounts of electricity are needed for the aeration phase of the biological treatment, not to mention the electricity requirements for dehydration of sludge, among many other overall energy requirements associated with the treatment processes (Gu et al., 2017). On the other hand, as already mentioned, the wastewater is an energy carrier that requires adequate recovery processes and technology.

Two main pathways derive from this water-energy nexus attached in particular to the wastewater. These are the pathway of efficiency improvements, either through treatment processes optimization (Wett et al., 2007; Solon et al., 2019) or in combining treatment process optimization with more efficient technology (Gandiglio et al., 2017; Gikas, 2017), and the pathway of energy recovery (Gandiglio et al., 2017; Gu et al., 2017), and, of course, the combination of both. In the recent years, another trend is gaining momentum, which refers to the utilization of more conventional renewable sources, such as solar, wind, and even water flows (Chae and Kang, 2013; Gu et al., 2017) in parallel with the recovery of embedded chemical energy, mainly in the form of biogas. This new trend is helping achieve the target of energy neutrality of wastewater treatment plants and even help turn this type of facilities into energy producers (Maktabifard et al., 2018). Furthermore, as also mentioned, the exploration of the thermal energy embedded in the wastewater, both locally for heating and cooling purposes through heat pumps (Chae and Kang, 2013) or for heat injection in district heating systems through heat exchangers and heat pumps (Gu et al., 2017) is another form of energy recovery gaining increased attention in some countries.

Another relevant aspect that helps to grasp the potential for both energy efficiency improvements and energy recovery that are associated with exploring the water-energy nexus within the wastewater treatment activity is the use of benchmarking and performance assessment strategies. These are practices highlighted in several studies dedicated to the research toward energy neutral and energy positive WWTPs, such as those of Gu et al. (2017), Maktabifard et al. (2018), Wett et al. (2007), and Gurung et al. (2018).

Although the efforts made in the scope of the water-energy nexus regarding the wastewater, are as many and diverse as they are valuable, there is a common trend associated to the majority of them, which is the isolated approach into the wastewater treatment plants, although some exceptions exist such as those derived from industrial symbiosis to enable co-digestion (Gandiglio et al., 2017), some already in practice urban integration of the WW thermal energy (Neugebauer et al., 2015), and the studies that focus on benchmarking activities (Gurung et al., 2018). In other words, the concept of system is, in general, being explored at very micro-levels. However, as argued by (Maktabifard et al., 2018), it is necessary to broaden the horizons beyond the limits of the individual plants in the pursuit of sustainability. Therefore, this is an important gap worth being explored, i.e., taking a system-wide vision at meso- and even macro-levels in regards to the energy efficiency and energy recovery, as well as energy production associated with the wastewater treatment. This type of broader perspective upon the systems is a call for integrated thinking and multi-disciplinary approaches. An example of an integrated thinking upon the wastewater treatment is the one defended by Silva et al. (2015) in the scope of energy management systems such as those framed by the ISO 50001 that can be perfectly applied to the wastewater treatment systems. In particular, the authors recommend to consider a unit-process as a basis for energy management rather than individual assets. The authors gave the example of the network-based infrastructure of the WW system and the need to tackle some major problems, such as overflows, yet the same can also be applied to the energy production processes as an example in the scope of the water-energy nexus.

More examples of integrated thinking combined with notably required multi-disciplinary approaches are those explored by Nguyen et al. (2019) regarding the concept of sponge cities, and by Garrido-Baserba et al. (2020) regarding the role of big data analytics and artificial intelligence in the new era of WW management systems.

Another aspect that is worth highlighting within the transition phase undertaken by the wastewater sector is the issue of governance alongside with organizational and management systems, and the role of performance measurement in the pathway toward full sustainability. In this context, we have found three topics that seem to have significant relevance in the pursuit of an improved, more functional, as well as more sustainable wastewater sector worldwide. The first one is benchmarking and performance measurement at a regulatory and utility level as a way of either deriving meaningful information to assist decision-making and policy development or to provide an artificial market competition to boost service quality and efficiency in a sector that is typically characterized by a natural monopoly structure. Some studies that dedicate to this particular topic are those from Abbott and Cohen (2009), Admiraal and Helden (2003), Marques et al. (2011), Picazo-Tadeo et al. (2011), Seppälä (2015), Tupper and Resende (2004), and Walker et al. (2020). The second topic refers to the concept of integrated management at a very comprehensive scale, either in terms of integration of sectors, activities, or even in a broader sense, considering urban planning, local-, regional-, national-level approaches, and even the so-called river-basin approaches. Some studies that focus on this topic are those from Burkhard et al. (2000), Grigg (2016), Nguyen et al. (2019), and Ujang and Buckley (2002). The third topic concerns the involvement of the private sector as a crucial participator on the development, expansion and innovation of the wastewater sector with special emphasis in developing countries. Examples of studies that discuss this topic are those of Murray et al. (2011), Sarvari et al. (2021), and Tecco (2008).

For a more detailed analysis, Table I.1 (see Annex I) provides a more in-depth and systematized overview of some studies considered relevant in terms of the focus that is given to the challenges and the discussion around the topic of sustainability in the wastewater sector. More precisely, these studies address how sustainability can be more consistently attained within the wastewater sector worldwide under the call for transition, which is being triggered by the great drivers of climate change and population growth.

The main general ideas that can be retrieved from this motivational literature review, which is centered on the challenges and discussion currently related to the wastewater sector, are the following:

- The wastewater sector is undertaking a transition phase toward fully sustainability all over the world, with developing and developed countries facing different needs to cope with the major challenges at hand and in alignment with the sustainability development goals determined by the United Nations for the horizon of 2030.
- Increased integrated and multidisciplinary approaches to management and operation of wastewater systems are sought to be key toward sustainability given the complex nature of the wastewater sector. Circular economy is a core example of an integrated approach to the function of systems across sectors that requires the use of a multidisciplinary perspective.
- A system-wide perspective upon entities within the wastewater sector that are usually considered individually is a form of integration that has room for research exploration.
- Good data management systems for enhanced information extraction and ready to use in data-based approaches, such as benchmarking, performance measurement, and general optimization applications, is key to support improved decision-making and policy developments.
- Big data, data analytics, and artificial intelligence are research areas gaining momentum within the wastewater sector, and that will be key for highly advanced integrative approaches and full sustainability.
- The water-energy nexus concerning the wastewater sector refers to how energy is used along the process of wastewater collection, treatment, and disposal, specifically in terms of efficiency but also in terms of the type of energy sources used, as well as how the wastewater-embedded energy, either in the thermal or chemical forms, is recovered.

Given the information displayed above, I found it interesting to derive a schematized overview that corresponds to a critical and relational perspective centered on the topic of the transitional phase undertaken by the wastewater sector since, in my perspective, it corresponds to the main ground for the overarching research development in what concerns the wastewater sector (see Figure 1.3). In this schematic representation, I would like to highlight the general idea, based on the literature reviewed, that corresponds to a vision of what a fully sustainable wastewater sector should look like.

Figure 1.3 *– Summary perspective upon the current research trends toward a more sustainable wastewater sector worldwide.*

1.3. Research rationale and thesis structure

In the above sections, it was identified the problem to tackle, the research boundaries to guide the thesis work, and the objectives. Also, from the motivational literature review, some of the current research trends and research gaps were identified. These are in need to be covered with increasing studies, consistent deployment of innovative practices worldwide, as well as increased diversity and experimentation.

Here, I wish to recall the transition phase discussion. As it was pointed out, a transition phase takes long time to lead a complex system from point A to point B in terms of the changes that are necessary to occur, considering all the interconnectedness among all structures, dimensions, contexts, and intervening parties. As such, if considering the wastewater sector as a worldwide concept, its constituent parts will be the wastewater sector as it is experienced at a country or region level. And, from this point of view, the transition phase develops at different paces and at the need to cope with different challenges that are context-dependent (e.g., Iribarnegaray et al., 2012).

Nonetheless, there is a common denominator that lies beyond the same mission of treating the wastewater, and that is the need for good governance of the complex structures at different levels and dimensions (e.g., Akhmouch, 2016; OECD, 2011), which can be accomplished, in some way, through good organizational management.

Figure 1.4 gives a summarized perspective upon how a water (wastewater) organization is able to undertake a good management in order to cope with their multiple dimensions within an evolving context, particularly in terms of information technology that is able to provide numerous management benefits but at the cost of the need to adapt to the inherent challenges derived from the increasing amount of data that so become available (Eggimann et al., 2017; Moy de Vitry).

Here, it is relevant to highlight that data is the key to information extraction and information is the key to good management. Therefore, as the transition phase is slow and uneven throughout different countries, or even in the same country or region, it means that there is still scope for diversity in terms of strategies and approaches to undertake so that the change that is demanded, and in accordance with the existing drivers, technology, and amount of available data, may effectively be accomplished.

In this context, I will select two sub-areas of knowledge and place dedicated attention and research efforts into completing the contribution of this thesis to the scientific research.

The first one is benchmarking and performance measurement. As reported in the motivational literature review, benchmarking and performance measurement can be applied at different hierarchic levels within the wastewater sector and with different objectives. In this thesis project, the aim is to develop benchmarking and performance measurement studies based on the data envelopment analysis (DEA) technique that are able to cope with challenges both at the more aggregated and at a less aggregated level. At the more aggregated level, I will refer to the regulation of utilities' activity that can lead to improved and more committed performance and engagement, as well as to promote increased transparency and collaboration within the sector, along with hopefully providing some assistance to decision-making as well as triggering policy improvements. At the less aggregated level, I will refer to the assets directly managed by the utilities, in this case with the main goal of resource efficiency improvements, particularly energy, along with the promotion of a more effective treatment process and also the corporate social responsibility of utilities through the good management of their assets to an improved service quality provision.

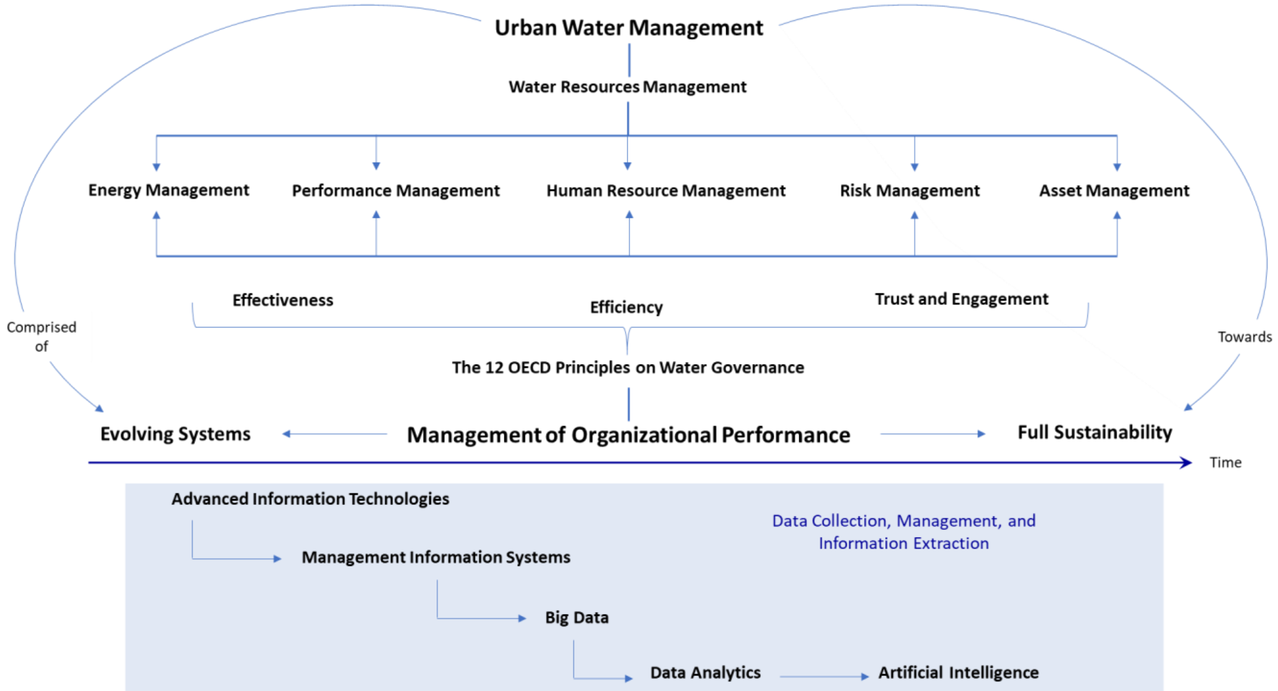


Figure 1.4 *– A perspective upon the urban water management under a data evolving environment.*

The research questions to be answered within the context of benchmarking and performance measurement are identified below, along with the correspondent papers that will be developed to answer them.

Note that the research questions will not be here discriminated in their natural order and according to their relative relevance for the research problem, considering my personal perspective, but rather they are grouped according to the main area of knowledge of the research conducted to answer them, which in this case is benchmarking and performance measurement as previously stated.

- **Research Question 1** (**RQ1**):
  How to contribute for enhancements in the wastewater sector governance so as to promote improved corporate social responsibility and a wiser use of resources, including energy, at a utility level?
  Paper:
  "Performance benchmarking using composite indicators to support regulation of the Portuguese wastewater sector"

- **Research Question 3** (**RQ3**):
  In what ways can a quantitative benchmarking assessment framework be explored from a water company perspective given the very particular characteristics of their assets' design and operation, with the aim of reducing resource consumption and increasing the sustainability of the company?
  Paper:
  "The use of order-*m* for a benchmarking analysis of a wastewater treatment plants' problematic sample".

- **Research Question 4** (**RQ4**):

What additional insights and guidelines can be drawn from a productivity analysis over time respecting a set of WWTPs' activity in a problematic sample?
Paper:
"Measuring productivity changes of Portuguese wastewater treatment plants using an order-*m* Global Malmquist Index".

- **Research Question 5** (**RQ5**):
How to conduct a performance measurement of a set of WWTPs aiming to integrate improvements both in resource's use efficiency and treatment effectiveness? What insights and guidelines can be obtained from a system-level assessment allowing for trade-offs in emission quotas?
Paper:
"A system-level optimization framework for efficiency and effectiveness improvement of wastewater treatment plants".

The second sub-area of knowledge is the water-energy nexus, which will be approached by giving special emphasis to the energy recovery potential improvements that can be considered using an integrated system vision of WWTPs' activity.

The research question to be answered in this context is indicated below, along with the research paper that will answer it:

- **Research Question 2** (**RQ2**):
Is it economically viable to optimize WWTPs' sludge management and biogas production through an interconnected and logistics approach that links different types of wastewater treatments at a system level, considering more than two plants?
Paper:
"Leveraging logistics flows to improve the sludge management process of wastewater treatment plants".

Both sub-areas of knowledge selected for this research work allow the development of the main thesis contribution, i.e., a set of decision-support tools that I believe can be further developed so as to be part of more integrated systems within wastewater organizations for improved management, particularly energy management, and improved governance. Eventually, the proposed tools can be directly deployed or adapted to local or regional conditions at the will and availability of interested stakeholders and where such tools could derive the major benefits.

Regarding the validation of the proposed decision-support tools, Portuguese data will be used, either provided by a water company that had the internal willingness of working towards the betterment of its assets' performance, or data that were publicly available on the regulator's webpage.

To better explore the selected sub-areas of knowledge above identified, and given all the context previously discussed, I identify below the main points of the research rationale adopted so to derive the overall thesis contribution:

- System-wide vision
- Multi-decision level intervention
- Decision-support tools
- Multimethodological approach
- Management perspective

From the above identified papers, two have already been published in international and peer review journals, and other two are currently under revision. Further details will be given in the correspondent sections.

I would also like to note that the research papers already published, as well as those that are currently under revision, were developed by a team of several people. Therefore, previous to displaying the content related to each research paper, my contribution to the research will also be due identified.

Moreover, before presenting the research papers, in Chapter 2 it will be made a brief introduction to performance measurement and the data envelopment analysis technique to facilitate the understanding of the thesis since these areas of knowledge play a major role in the overall research conducted (i.e., four out of the five research papers are DEA-based).

References

Abbott, M., Cohen, B. (2009). Productivity and efficiency in the water industry. *Utilities Policy*, 17, 233-244.

Admiraal, R. J., van Helden, G. J. (2003). Benchmarking in the Dutch wastewater treatment sector. *Public Money and Management*, 23(2), 113-118.

Akhmouch, A. (2016). The 12 OECD principles on water governance – When science meets policy. *Utilities Policy*, 43, 14-20.

Ampe, K., Paredis, E., Asveld, L., Osseweijer, P., Block, T. (2020). A transition in the Dutch wastewater system? The struggle between discourses and with lock-ins. *Journal of Environmental Policy & Planning*, 22(2), 155-169.

Brown, R. R., Keath, N., Wong, T. H. (2009). Urban water management in cities: historical, current and future regimes. *Water Science and Technology*, 29(5), 847-855.

Burkhard, R., Deletic, A., Craig, A. (2000). Techniques for water and wastewater management: a review of techniques and their integration in planning. *Urban Water*, 2, 197-221.

Cantor, A., Sherman, L., Milman, A., Kiparsky, M. (2021). Regulators and utility managers agree about barriers and opportunities for innovation in the municipal wastewater sector. *Environmental Research Communications*, 3, 031001.

Chae, K. J., Kang, J. (2013). Estimating the energy independence of a municipal wastewater treatment plant incorporating green energy resources. *Energy Conversion and Management*, 75, 664-672.

Daw, J., Hallet, K., DeWolfe, J., Venner, I. (2012). Energy efficiency strategies for municipal wastewater treatment facilities. Technical Report, National Renewable Energy Laboratory (NREL), Golden, Co, USA.

Dincer, I. (2000). Renewable energy and sustainable development: a crucial review. *Renewable & Sustainable Energy Reviews*, 4, 157-175.

Eggimann, S., Mutzner, L., Wani, O., Schneider, Y., Spuhler, D., de Vitry, M. M., Beutler, P., Maurer, M. (2017). The potential of knowing more: a review of data-driven urban water management. *Environmental Science & Technology*, 51, 2538-2553.

Foladori, P., Vaccari, M., Vitali, F. (2015). Energy audit in small wastewater treatment plants: methodology, energy consumption indicators, and lessons learned. *Water Science & Technology*, 72(6), 1007-1015.

Gandiglio, M., Lanzini, A., Soto, A., Leone, P., Santarelli, M. (2017). Enhancing the energy efficiency of wastewater treatment plants through co-digestion and fuel cell systems. *Frontiers in Environmental Science*, 5(70).

Garrido-Baserba, M., Corominas, L., Cortés, U., Rosso, D., Poch, M. (2020). The fourth revolution in the water sector encounters the digital revolution. *Environmental Science & Technology*, 54, 4698-4705.

Gikas, P. (2017). Towards energy positive wastewater treatment plants. *Journal of Environmental Management*, 203, 621-629.

Grigg, N. S. (2016). The business of water in a changing world: organizations, connectors and support sector. *International Journal of Water Resources Development*, 32(5), 708-720.

Gu, Y., Li, Y., Li, X., Luo, P., Wang, H., Robinson, Z. P., Wang, X., Wu, J. (2017). The feasibility and challenges of energy self-sufficient wastewater. *Applied Energy*, 204, 1463-1475.

Gurung, K., Tang, W. Z., Sillanpää, M. (2018). Unit energy consumption as benchmark to select energy positive retrofitting strategies for Finnish wastewater treatment plants (WWTPs): a case study of Mikkeli WWTP. *Environmental Processes*, 5, 667-681.

Hamiche, A. M., Stambouli, A. B., Flazi, S. (2016). A review of the water-energy nexus. *Renewable and Sustainable Energy Reviews*, 65, 319-331.

IEA (International Energy Agency) (2018). World Energy Outlook.

Iribarnegaray, M. A., Copa, F. R., D'Andrea, M. L. G., Arredondo, M. F., Cabral, J. D., Correa, J. J., Liberal, V. I., Seghezzo, L. (2012). A comprehensive index to assess the sustainability of water and sanitation management systems. *Journal of Water, Sanitation and Hygiene Development*, 2, 205-222.

Jenicek, P., Kutil, J., Benes, O., Todt, V., Zabranska, J., Dohanyos, M. (2013). Energy self-sufficient sewage wastewater treatment plants: is optimized anaerobic sludge digestion the key? *Water Science & Technology*, 68(8), 1739-1744.

Killip, G., Owen, A., Morgan, E., Topouzi, M. (2018). A co-evolutionary approach to understanding construction industry innovation in renovation practices for low-carbon outcomes. *The International Journal of Entrepreneurship and Innovation*, 19(1), 9-20.

Lehtoranta, S., Nissinen, A., Mattila, T., Melanen, M. (2011). Industrial symbiosis and the policy instruments of sustainable consumption and production. *Journal of Cleaner Production*, 19, 1865-1875.

Loorbach, D. (2010). Transition management for sustainable development: a prescriptive, complexity-based governance framework. Governance: *An International Journal of Policy, Administration, and Institutions*, 23(1), 161-183.

Maktabifard, M., Zaborowska, E., Makinia, J. (2018). Achieving energy neutrality in wastewater treatment plants through energy savings and enhancing renewable energy production. *Reviews in Environmental Science and Biotechnology*, 17, 655-689.

Marcos von Sperling, Chernicharo, C. A. L. (2002). Urban wastewater treatment technologies and the implementation of discharge standards in developing countries. *Urban Water*, 4, 105-114.

Marques, R. C., da Cruz, N. F., Pires, J. (2015). Measuring the sustainability of urban water services. *Environmental Science & Policy*, 54, 142-151.

Marques, R.C., Simões, P., Pires, J.S. (2011). Performance benchmarking in utility regulation: the worldwide experience. *Polish Journal of Environmental Studies*, 20(1), 125-132.

Moy de Vitry, M., Schneider, M. Y., Wani, O., Manny, L., Leitão, J.P., Eggimann, S. (2019). Smart urban water systems: what could possibly go wrong? *Environmental Research Letters*, 14.

Murray, A., Mekala, G. D., Chen, X. (2011). Evolving policies and the roles of public and private stakeholders in wastewater and faecal-sludge management in India, China and Ghana. *Water International*, 36(4), 491-504.

Neugebauer, G., Kretschmer, F., Kollmann, R., Narodoslawsky, M., Ertl, T., Stoeglehner, G. (2015). Mapping thermal energy resource potentials from wastewater treatment plants. *Sustainability*, 7, 12988-13010.

Nguyen, T. T., Ngo, H. H., Guo, W., Wang, X., C., Ren, N., Li, G., Ding, J., Liang, H. (2019). Implementation of a specific urban water management – Sponge City. *Science of the Total Environment*, 652, 147-162.

Nowak, O., Enderle, P., Varbanov, P. (2015). Ways to optimize the energy balance of municipal wastewater systems: lessons learned from Austrian applications. *Journal of Cleaner Production*, 88, 125-131.

OECD (2011). Water in OECD countries: a multi-level approach – highlights.

Picazo-Tadeo, A., Sáez-Fernández, F. J., González-Gómez, F. (2011). Assessing performance in the management of the urban water cycle. *Water Policy*, 13, 782-796.

Qadir, M., Drechsel, P., Cisneros, B. J., Kim, Y., Pramanik, A., Mehta, P., Olaniyan, O. (2020). Global and regional potential of wastewater as a water, nutrient and energy source. *Natural Resources Forum*, 44, 40-51.

Rotmans, J., Kemp, R., Marjolein van Asselt (2001). More evolution than revolution: transition management in public policy. *Foresight: The Journal of Futures Studies, Strategic Thinking and Policy*, 3(1), 15-31.

Sarvari. H., Chan, D.W.M., Banaitiene, N., Noor, N.M., Beer, M. (2021). Barriers to development of private sector investment in water and sewage industry. *Built Environment Project and Asset Management*, 11(1), 52-70.

Sauvé, S., Bernard, S., Sloan, P. (2016). Environmental sciences, sustainable development and circular economy: alternative concepts for trans-disciplinary research. *Environmental Development*, 17, 48-56.

Scott, C. A., Pierce, S. A., Pasqualetti, M. J., Jones, A. L., Montz, B. E., Hoover, J. H. (2011). Policy and institutional dimensions of the water-energy nexus. *Energy Policy*, 39, 6622-6630.

Seppälä, O. T. (2015). Performance benchmarking in Nordic water utilities. *Procedia Economics and Finance*, 21, 399-406.

Sharma, P. (2019). Digital revolution of education 4.0. *International Journal of Engineering and Advanced Technology*, 9(2), 3558-3564.

Shen, Y., Linville, J. L., Urgun-Demirtas, M., Mintz, M. M., Snyder, S. W. (2015). An overview of biogas production and utilization at full-scale wastewater treatment plants (WWTPs) in the United States: challenges and opportunities towards energy-neutral WWTPs. *Renewable and Sustainable Energy Reviews*, 50, 346-362.

Silva, C., Alegre, H., Rosa, M.J. (2015). Introduction to energy management in wastewater treatment plants. In: Sewage treatment plants – economic evaluation of innovative technologies for energy efficiency. Stamatelatou, K. and Tsagarakis, K. P. (Eds). *IWA Publishing*.

Solon, K., Jia, M., Volckle, E. I. P. (2019). Process schemes for future energy-positive water resource recovery facilities. *Water Science & Technology*, 79(9), 1808-1820.

Spinosa, L., Ayol, A., Baudez, J.C., Canziani, R., Jenicek, P., Leonard, A., Rulkens, W., Xu, G., Lex van Dijk (2011). Sustainable and innovative solutions for sewage sludge management. *Water*, 3, 702-717.

Suárez-Eiroa, B., Fernández, E., Méndez-Martínez, G. (2019). Operational principles of circular economy for sustainable development: linking theory and practice. *Journal of Cleaner Production*, 214, 952-961.

Tecco, N. (2008). Financially sustainable investments in developing countries water sectors: what conditions could promote private sector involvement? *International Environmental Agreements*, 8, 129-142.

Tupper, H.C., Resende, M. (2004). Efficiency and regulatory issues in the Brazilian water and sewage sector: an empirical study. *Utilities Policy*, 12, 29-40.

U.S. EPA. (2013). Energy efficiency in water and wastewater facilities. Available: https://www.epa.gov/sites/production/files/2015-08/documents/wastewater-guide.pdf. Accessed 6/07/2021

Ujang, Z., Buckley, C. (2002). Water and wastewater in developing countries: present reality and strategy for the future. *Water Science and Technology*, 46(9), 1-9.

UN General Assembly (2015). Transforming our world: the 2030 Agenda for Sustainable Development. A/RES/70/1, available at: https://www.refworld.org/docid/57b6e3e44.html [accessed 12 July 2021].

Van der Brugge, R., Rotmans, J. (2007). Towards transition management of European water resources. In: Integrated assessment of water resources and global change. Craswell, E., Bonell, M., Bossio, D., Demuth, S., van de Giesen, N. (Eds.), *Springer*.

Voulvoulis, N. (2018). Water reuse from a circular economy perspective and potential risks from an unregulated approach. *Current Opinion in Environmental Science & Health*, 2, 32-45.

Walker, N.L., Williams, A.P., Styles, D. (2020). Key performance indicators to explain energy & economic efficiency across water utilities, and identifying suitable proxies. *Journal of Environmental Management*, 269, 110810.

Walsh, B. P., Murray, S. N., O'Sullivan, D. T. J. (2015). The water energy nexus, an ISO50001 water case study and the need for a water value system. *Water Resources and Industry*, 10, 15-28.

WEF (2013). Biogas production and use at water resource recovery facilities in the United States. *Water Environment Federation*. Alexandria, VA, USA.

Wett, Bernhard & Buchauer, Konrad & Fimml, Christian. (2007). Energy self-sufficiency as a feasible concept for wastewater treatment systems. *IWA Leading Edge Technology Conference*.

Zayed, T., Chen, Z., Qasem, A. (2015). Infrastructure performance rating models for wastewater treatment plants. *Structure and Infrastructure Engineering*, 11(5), 668-682.

Chapter 2 –

Performance Measurement and Data Envelopment Analysis (DEA)

Chapter 2

Performance measurement and Data Envelopment Analysis

This chapter makes a brief introduction to performance measurement and the data envelopment analysis (DEA) technique, as these areas of knowledge play a core role throughout the thesis.

2.1. A brief note on performance measurement

Performance measurement enables transforming data into useful information regarding the efficiency, effectiveness, and productivity of organizations. Particularly, performance measurement usually intends to provide information derived from the observation of systems' functioning so that improvements can be sought and implemented in a practical way. Also, when considering organizations, it can be conducted at different levels, e.g., from the core production process (e.g., the wastewater treatment itself, in the case of wastewater organizations), going through each of the inherent systems and functions that give structure to the organization, to the entire organization and the governance organisms above the organization. In the case of the wastewater sector, performance measurement is a relevant mechanism that is recommended by OECD for effective management (OECD, 2011).

Departing from one of the more aggregated level, e.g., evaluating the management of organizational performance, one recognized performance measurement methodology is the Balanced Scorecard (BSC) proposed by Kaplan and Norton (1992), which acknowledges the fact that an organization shouldn´t be evaluated only through its financial performance but rather from an integrative perspective of at least four dimensions (i.e., financial, customer, internal business, and innovation and growth) through relevant indicators. One good example of the application of the BSC is the work conducted by the Dutch water boards on the performance measurement and evaluation of the wastewater treatment (Admiraal and Helden, 2003). In this particular case, a fifth dimension is actually added, namely the environmental dimension. Another application of the BSC is the one proposed by Otley (1999) as a framework to translate an organization's vision and strategy into operational terms. More recently, it is interesting to note that the BSC tool is being used in combination with other methodologies and, as such, providing an enlargement of the integration capability to cope with more data and to derive even more meaningful information. Some examples of this type of combinations are identified in the work of Amado et al. (2012), particularly through a literature review that is followed by the development of a combination of the BSC methodology with DEA. This study in particular proposes a framework where a sequence of DEA models (according to the concept of network DEA proposed by Färe and Grosskopf, 2000) are integrated in a way that the four perspectives underpinning the BSC are accounted for so as to derive greater insights from the organization's performance evaluation.

Before entering into the in-depth DEA structural details, we will start by recognizing DEA as a performance measurement technique (Cook et al., 2014) that derives a summary score to each individual unit within a sample of units under analysis. DEA has played a very prominent role in the fields of banking, health care, agriculture and farm, transportation, and education (Liu et al., 2013). It has also been relevant in the fields of manufacturing and production research as a data-driven methodology (Kuo and Kusiak, 2019). DEA continues to play a relevant role throughout a panoply of application fields, including sustainability (Zhou et al., 2018), and also undertaking methodologic developments (Liu et al., 2016). DEA will continue to be relevant even within the context of big data and data analytics due to its capability for valuable information extraction (Zhu, 2020).

As mentioned, DEA is a recognized technique for performance measurement in production economics. Nonetheless, it can also be used as a multicriteria decision-making (MCDM) tool and as a benchmarking methodology (Cook et al., 2014). In case of production economics, where the decision-making units (DMUs) make use of resources (inputs) to produce outcomes (outputs), DEA allows the construction of a production function and then the score obtained for each DMU corresponds to its production efficiency or technical efficiency. In the case of MCDM, the DMUs are alternatives, each of which is characterized by sets of different performance criteria (e.g., with minimization and maximization objectives), and therefore the scores derived will allow to choose the best alternatives (Cook et al., 2014). In case of benchmarking, the objective is to derive insights into the overall performance of a system, or part of a system, and to construct a best-practice frontier to be the reference set for the remaining DMUs (for more information upon benchmarking we refer the reader to Bogetoft and Otto, 2011). In this case, the DEA score obtained will represent the performance measure of each DMU. In summary, the DEA score obtained from an analysis will be coined as an aggregated (or composite) measure that translates the relative performance of DMUs, which is to be interpreted according to the context and purpose of the assessment (Cook et al., 2014).

2.1. An introductory approach to the basics of DEA

The following paragraphs correspond to an introduction to the DEA functioning. For more details, we refer the reader to the work of Cooper et al. (2007).

DEA was developed by Charnes et al. (1978) as a linear programming technique to evaluate the performance of decision-making units (DMUs). DEA uses the concept of radial efficiency, as presented by Farrell, (1957), as a measure of technical efficiency. Under a conventional DEA framework, the DMUs are similar units in terms of the types of inputs and outputs (i.e., individual indicators) that characterize them. These indicators can be translated by different measurement units. DEA is a nonparametric technique since it relies completely on the data set of indicators, without any a-priori imposition of a functional form for the production function, in what resembles a black box (Färe et al., 1997) of efficiency evaluation (see Figure 2.1). In other words, for an effective performance evaluation there is no need to know how the technology (i.e., machinery, components, processes, know-how) exactly transforms inputs into outputs. Rather, it is only required a good knowledge of the overall performance problem to best define the purpose of the evaluation and to identify the best set of indicators that are the most suitable for representing the performance measurement problem at hand (Cook et al., 2014).

Figure 2.1 *– Schematic representation of a DMU's activity for a DEA evaluation.*

The rationale behind the original DEA model (Charnes et al., 1978) is the comparison of a given DMU's productivity (i.e., the ratio of the weighted outputs produced to the weighted inputs used), with the productivity of other DMUs. This comparison involves finding weights that maximize the ratio of the aggregated output over the aggregated input, provided that the remaining DMUs' productivity is always less than or equal to one when evaluating using the same weights. Model (2.1) corresponds to the mathematical programming formulation of such model.

$$\max e_{j_0} = \frac{\sum_{r=1}^{s} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij_0}} \tag{2.1}$$

subject to:

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \qquad j = 1, \dots, n$$

$$u_r, v_i \geq 0 \qquad r = 1, \dots, s; i = 1, \dots, m$$

where, y_{rj} and x_{ij} are the outputs and inputs observed at DMU j, respectively, u_r and v_i are the corresponding weights to be determined (i.e., the decision variables of the optimization problem), j stands for each of the n DMUs in the sample, and j_0 stands for the DMU under evaluation.

One of the major advantages of DEA as a performance measurement technique is that there is no need for information about each indicator's relative importance, i.e., the indicators' weights, since the weights are actually the decision variables to be computed with the model, by giving each DMU the possibility of weighting more those indicators where its performance is better, which is the natural consequence obtained from the model's rationale indicated above. In production economics, those weights are best known as prices, which are not always easily available. Model (2.1)'s formulation is also what is best suited for MCDM purposes, one of the three types of methods above indicated that DEA can be used for (Cook et al., 2014). However, its correspondent linear programming model, expressed with formulation (2.2), is the one that is known as a DEA output-oriented model in its "weights formulation". It is also known by CCR, which stands for the initials "Charnes, Cooper, and Rhodes" – the authors of Charnes et al., 1978. This model also corresponds to the constant returns to scale (CRS) formulation reflecting how the outputs vary with varying amounts of inputs.

$$\min g_{j_0} = \sum_{i=1}^{m} v_i x_{ij_0} \tag{2.2}$$

subject to:

$$-\sum_{r=1}^{s} u_r y_{rj} + \sum_{i=1}^{m} v_i x_{ij} \geq 0 \qquad j = 1, \dots, n$$

$$\sum_{r=1}^{s} u_r y_{rj_0} = 1$$

$$u_r, v_i \geq 0 \qquad r = 1, \dots, s; i = 1, \dots, m$$

where v_i and u_r are the weights correspondent to the inputs, x_i, and outputs, y_r, respectively, to be determined for each DMU j_0 at the optimal solution of (2.2).

The dual of the linear programming model (2.2) is given by (2.3), which corresponds to the envelopment formulation of the conventional DEA output-oriented model under the CRS assumption, as developed in Charnes et al. (1978).

$$\max \beta_{j_0} \tag{2.3}$$

subject to:

$$-\sum_{j=1}^{n} y_{rj} l_j + y_{rj_0}\beta_{j_0} \leq 0 \qquad r = 1, \ldots, s$$

$$\sum_{j=1}^{n} x_{ij} l_j \leq x_{ij_0} \qquad i = 1, \ldots, m$$

$$l_j \geq 0 \qquad j = 1, \ldots, n$$

where l_j is an intensity vector that gives the set of best-practices in the efficient frontier to be considered as peers for the DMU j_0, which is the one under evaluation. When the model (2.3) is run for each of all DMUs j in the sample, then all the best-practices in the sample will be identified and so the efficient frontier will be composed of all those DMUs. In order to be considered as a best-practice (i.e., benchmark), a DMU must have assigned an efficiency score of 1 (i.e., $\beta_{j_0} = 1$) at the optimal solution of model (2.3). All the remaining DMUs will score an overall efficiency measure that is positive and less than 1 (i.e., $0 < \left(1/\beta_{j_0}\right) < 1$). The greater the efficiency score the closest is the DMU to the best-practice frontier.

In the following, we will present a small numerical example to illustrate the application of model (2.3). We considered a sample of 8 DMUs (A to H). Table 2.1 shows the data regarding the indicators that are considered relevant to translate the hypothetical performance measurement problem. The results of the DEA assessment after applying model (2.3) to the indicated sample are displayed in Table 2.2. From the sample of 8 DMUs, only 3 were considered efficient, i.e., they were assigned an efficiency score of one. Figure 2.2 gives a graphic representation of the efficient frontier and the DMUs that compose it (i.e., DMU B, DMU F, and DMU E). These DMUs, as they are on the best-practices frontier, they are being compared with themselves, i.e., in case of DMU B, for example, lambda B = 1 and all the remaining lambdas are zero. All the remaining DMUs are inefficient and, therefore, as they are radially projected towards the frontier, they are shown the best practices closest to them so to know how to make improvements in their own performance. For example, DMU A is being compared only with DMU B (lambda B = 1, lambda F = 0, lambda E = 0), and DMU D is being compared with DMU B and with DMU F (lambda B = 0.524, lambda E = 0, lambda F = 0.476).

Table 2.1 – Data regarding a sample of 8 DMUs (numerical example number 1).

DMU	A	B	C	D	E	F	G	H
Input 1	1	1	1	1	1	1	1	1
Output 1	4	6	8	6	12	11	9	3
Output 2	12	14	8	9	8	11	7	10

Table 2.2 – Output-oriented, radial, CRS-DEA performance measurement results (numerical example).

DMU	Efficiency score $(1/\beta_{j_0})$	Lambdas (l_j)							
		A	B	C	D	E	F	G	H
A	0.857		1						
B	**1**		1						
C	0.727						1		
D	0.716		0.524				0.476		
E	**1**					1			
F	**1**						1		
G	0.773					0.647	0.353		
H	0.714		1						

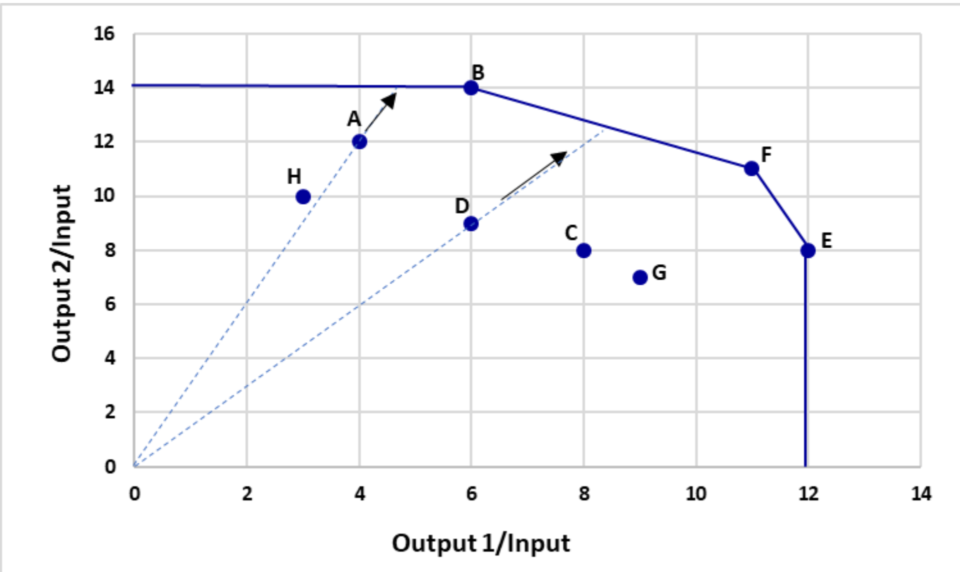


Figure 2.2 *– Graphic representation of the output-oriented and CRS best-practices frontier according with the DEA performance measurement technique.*

Regarding the inefficient DMUs, the targets for improvement (see Table 2.3) were calculated as a linear combination of the output indicators observed in the peers (best-practices), which were identified by the DEA model through the lambda values (see equation (2.4)).

$$y_{rj_0}^{T} = \sum_{j=1}^{n} \lambda_{jj_0}^{*} y_{rj} \qquad r = 1, \dots, s \qquad \textbf{(2.4)}$$

Table 2.3 – Targets for improvement.

DMU	Observed Indicators		Targets for improvement			
	Output 1	**Output 2**	**Output 1**	**(% of improvement)**	**Output 2**	**% of improvement**
A	4	12	6	(+50%)	14	(+17.7%)
B	6	14	---	---	---	---
C	8	8	11	(+37.5%)	11	(+37.5%)
D	6	9	8.4	(+39.7%)	11.6	(+29.1%)
E	12	8	---	---	---	---
F	11	11	---	---	---	---
G	9	7	11.6	(+29.4%)	9.1	(+29.4%)
H	3	10	6	(+100%)	14	(+40%)

The input-oriented DEA radial CRS model is shown in (2.5).

$$\min \theta_{j_0} \qquad \textbf{(2.5)}$$

subject to:

$$\sum_{j=1}^{n} y_{rj} l_j \geq y_{rj_0} \qquad r = 1, \dots, s$$

$$\sum_{j=1}^{n} x_{ij} l_j \leq x_{ij_0} \theta_{j_0} \qquad i = 1, \dots, m$$

$$l_j \geq 0 \qquad j = 1, \dots, n$$

An important note is that, under the CRS assumption, the efficiency score obtained for a given DMU is independent of the model orientation selected. In other words, the input- and output-oriented efficiency scores are identical under the CRS assumption (i.e., $\theta_{j_0} = 1/\beta_{j_0}$). However, if the purpose of the DEA assessment is to obtain the peers, or best-practices, so that those can be used as references for performance improvement, then it is important to understand which orientation is best suited to the problem at hand (Cook et al., 2014). In other words, the peers will vary according to the model orientation and, therefore, the targets for each indicator improvement that can be computed through the identification of peers will also vary accordingly.

Another note is that for benchmarking purposes, the model formulation that is best suited, either output- or input-oriented, is the envelopment formulation as indicated in (2.3) and (2.6). The envelopment formulation gives the possibility of directly identifying the best-practices in the sample of n DMUs under assessment, which is an information derived from the l_j vector computed for each $DMU_j, j = 1, \dots, n$. As previously demonstrated, the vector of lambdas is the information used for the computation of the targets for improvement of each inefficient DMU, which is one major purpose of a benchmarking assessment.

However, there are some limitations to the original DEA models above presented. The first limitation is that the radial efficiency concept that is implicit may allow the identification of efficient DMUs (i.e., efficiency score equal to 1) when, in fact, they are not. First of all, the concept of efficiency in the context of performance measurement through the use of frontier methods, such as DEA, is commonly understood through the concept of dominance. In other words, a DMU is considered efficient if there is no other DMU in the setting that dominates it in any way, i.e., there is no other DMU that uses simultaneously less inputs to obtain more outputs than itself. This is in fact the concept of efficiency in the Pareto-Koopman's sense. The absence of the Pareto-Koopman's efficiency that can eventually be obtained through the basic DEA models (2.3) and (2.5) derive from the fact that the "efficient" frontier is an isoquant (Färe and Lovell, 1978) since it relies on the concept of Farrell's radial efficiency, which can give rise to the existence of slacks either in the input or output side. Therefore, as recognized by Charnes et al. (1978), it is necessary to consider the eventual existence of slacks so to analyze if a DMU is in fact efficient regardless of being considered at the frontier (i.e., efficient according to what is known as Pareto-Koopmans efficiency). Therefore, they introduced a slight modification to models (2.3) and (2.5). The DEA modified models are given through the sets of equations (2.6) and (2.7), which correspond respectively to the output- and input-orientations. According to these models, a DMU is truly efficient (i.e., in the Pareto-Koopman's sense) if and only if the efficiency score is one and all slacks are zero.

$$\max(\beta_{j_0} + \varepsilon \sum_{r=1}^{s} S_{rj_0} + \varepsilon \sum_{i=1}^{m} S_{ij_0}) \qquad \textbf{(2.6)}$$

subject to:

$$\sum_{j=1}^{n} y_{rj} l_j = y_{rj_0} \beta_{j_0} + S_{rj_0} \qquad r = 1, \dots, s$$

$$\sum_{j=1}^{n} x_{ij} l_j = x_{ij_0} - S_{ij_0} \qquad i = 1, \dots, m$$

$$l_j \geq 0 \qquad j = 1, \dots, n$$

$$S_{rj_0}, S_{ij_0} \geq 0 \qquad r = 1, \dots, s; \; i = 1, \dots, m$$

where S_{rj_0} are the slacks associated to outputs, also known as output shortfalls (i.e., it is possible to produce more outputs with the same amount of inputs), and S_{ij_0} are the slacks associated to inputs, also known as input excesses (i.e., it is possible to produce the same amount of outputs with less inputs).

$$\min\left(\theta_{j_0} - \varepsilon\sum_{r=1}^{s} S_{rj_0} - \varepsilon\sum_{i=1}^{m} S_{ij_0}\right) \quad \textbf{(2.7)}$$

subject to:

$$\sum_{j=1}^{n} y_{rj} l_j = y_{rj_0} + S_{rj_0} \qquad r = 1, \dots, s$$

$$\sum_{j=1}^{n} x_{ij} l_j = x_{ij_0}\theta_{j_0} - S_{ij_0} \qquad i = 1, \dots, m$$

$$l_j \geq 0 \qquad j = 1, \dots, n$$

$$S_{rj_0}, S_{ij_0} \geq 0 \qquad r = 1, \dots, s;\ i = 1, \dots, m$$

Another limitation may arise when using DEA models for performance measurement in small samples, or when the ratio of the number of DMUs to the total number of input and output indicators is quite small, even though there is no common sense in regards to what "small" means. In fact, according to Cook et al. (2014), the concept of "small" in regards to a sample's size can be completely meaningless since it depends on the problem's characteristics as well as the purpose of the analysis that needs to be considered. Nonetheless, there is a real implication on the number of efficient DMUs that can be identified through a DEA assessment when the said ratio is relatively "small", which is a matter usually known as a lack of discretionary capacity, i.e., there will be relatively too many DMUs considered efficient. Therefore, in the cases where there are only few DMUs in the sample and the number of indicators needed to translate the performance of those DMUs is relatively not small, or otherwise, i.e., in cases where despite a considerable good number of DMUs in the sample there is the need for a greater number of indicators, it may be advisable the use of weight restrictions (for more information please refer to Cooper et al., 2007). This is a solution that is also valid in cases when the decision-makers are willing to ensure that all the indicators (or some specific indicators) are considered despite the eventual underachievement that may be observed by some DMUs in those indicators. Despite the accountability for the problem of lack of discretionary capacity in some instances, the use of weight restrictions will always worsen (or maintain) the efficiency score obtained when those type of weight restrictions are not considered since by construction conventional DEA models allow each DMU to be seen at its "best light".

There is one further limitation of DEA, which stems from its pure empirical nature. Considering the context of a benchmarking application, the best-practice frontier will be constructed from the DMUs that are a reference in their relative performance, which was observed in a given past moment. In a benchmarking context, i.e., with no inference intentions, variations in these observations, either through measurement errors that may be embedded in the data used, or by the addition or removal of some observations (i.e., DMUs) from the sample, may originate a best-practices frontier that does not correspond to the real best-practices frontier, or originate different frontier shifts that will imply different efficiency measures, respectively. Also, since DEA is completely based on observations, the construction of the best-practice frontier is very sensitive to the presence of outliers. Therefore, from the first aspect pointed out, it is necessary to look at the DEA efficiency results with a certain amount of caution and, therefore, in its pure deterministic or conventional form DEA is not advisable for direct inference (Cook et al., 2014). Nonetheless, this should not constitute an impediment for DEA analyses since in many cases the insights derived can still provide a good indication of the system's performance while also provide guidance for performance improvements. In relation to the second aspect, regarding the sensitivity of the frontier to the presence of outliers, there can be applied sensitivity analysis (e.g., Bogetoft and Otto, 2011) to test if there may be any outlier DMUs in the sample under analysis, in which case they should be removed from the sample to avoid setting too ambitious and impractical targets for improvement for the DMUs considered inefficient.

In other words, the fact that DEA is both a deterministic and nonparametric frontier technique, leads to uncertainty that stems from the data sets used for the analyses, which is not accounted for. This main limitation to the DEA analyses is explored in Dyson and Shale (2010). The authors identify the several ways that are covered in the literature to tackle such limitation, and select imprecise DEA (IDEA) as a simple and useful method for uncertainty accountability in DEA and identification of robust efficient DMUs, and the Monte Carlo simulation as the most flexible and comprehensive approach for the same goal. The order-*m* method (Cazals et al., 2002) is another alternative and comprehensive approach that is able to account for the data uncertainty and simultaneously mitigate the impact of outliers on the computation of efficiency scores.

2.2. Technology, nonparametric frontier, and efficiency measurement

In the previous section, we presented the mathematical programming formulation that allows to assign a measure of performance to each DMU in a DEA assessment. However, in order for the complete understanding of the extent of a DEA application or, otherwise, for the complete adequacy of the DEA measurement to the given sample's performance, we need to keep in mind the assumptions underlying the determination of such performance measures, or efficiency scores, which should be in accordance with the nature of the problem at hand as well as the purpose and objectives of the analysis. In the previously indicated mathematical programming formulations, we already integrated some of those assumptions that in this section we will make clearer. Figure 2.3 provides an overview of a DEA framework for an adequate performance measurement, including the need for the identification of a reference technology.

In a very general way, the technology indicates the frontier of the production possibility set that transforming inputs into outputs. In this sense, the concept of technology is better understood in the context of the production manufacturing where resources are used, or transformed, to originate the outcomes, whether products or services. However, the technology has the more general purpose of providing boundaries within which a certain amount of multiple inputs (i.e., those indicators with the inherent objective of using the less of them as possible), can be combined so to originate combinations of multiple outputs (i.e., the indicators that are desirable to be maximized as much as possible). In other words, the technology allows to determine the frontier for the so-called production possibility set or, more generally, the set of all observations (i.e., all observed combinations of multiple inputs and multiple outputs), to allow for those observations having a reference to compare themselves with and so to be assigned an overall performance measure. This performance measure is no more than the distance that the observation must go during the process of its own improvement (i.e., eliminating all sources of inefficiency) to also become part of the best-practice frontier.

Mathematically, the technology can be generally represented by:

$$T = \{(x, y) \in \mathbb{R}_+^m \times \mathbb{R}_+^s \,|\, x \; can \; produce \; y\} \quad \textbf{(2.8)}$$

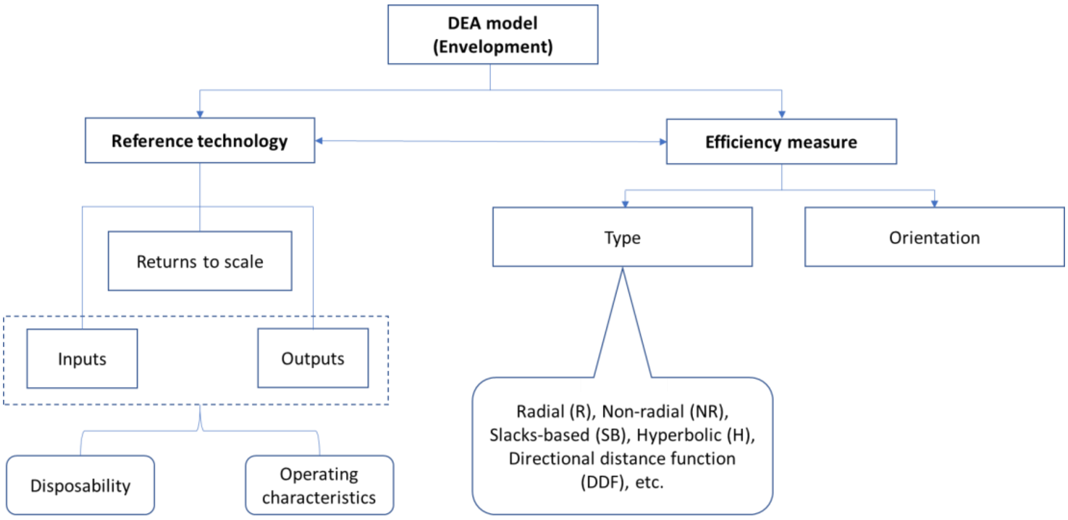


***Figure 2.3** – The general structure of a DEA model (envelopment form). Source: Zhou et al. (2008).*

Using the equation (2.8) in combination with the information displayed, for example, in the set of equations (2.7), we have:

$$T = \{(x, y) \in \mathbb{R}_+^m \times \mathbb{R}_+^s \,|\, x = \textstyle\sum_{j=1}^n x_j l_j + S^-,\ y = \sum_{j=1}^n y_j l_j - S^+, \\ such\ that\ l_j \geq 0,\ j = 1, \dots, n,\ S^- \geq 0,\ S^+ \geq 0\}, \tag{2.9}$$

where S^- represents the input slacks (or inputs excess) and S^+represents the output slacks (or outputs shortfalls).

Expression (2.9) defines the boundary of the production possibility set, against which the technical efficiency of each DMU j will be measured, according to the linear programming model presented in (2.7). For more information on the properties of the production possibility set identified by T, we refer the reader to Cooper et al. (2007). Formulation (2.9) indicates that the technical efficiency of each DMU j accounts for the eventual input and output slacks. It also assumes free disposability of inputs and outputs, which is an assumption also implicit in the models (2.3), (2.5) – (2.7). The free disposability assumption means that if with a given amount of inputs it is possible to produce a certain amount of outputs, then it is also true that with the same amount of inputs it is possible to produce a lower amount of outputs, and vice-versa, i.e., the same amount of outputs can be obtained by using a higher amount of inputs (Bogetoft and Otto, 2011). Moreover, in formulation (2.9) it is assumed that the efficient behavior of any DMU in the sample analyzed is that of a constant-returns to scale (CRS), i.e., departing from the identified behavior of the set of best-practices, all DMUs are assumed to present an input usage that should be equivalently proportional to the outputs attained by those considered best practices, regardless of the scale of operation so to derive their efficiency score. A less restrictive version of the CRS technology, is a variable-returns to scale (VRS) technology, i.e., a frontier that is closer to the observations, enveloping them more tightly. A VRS technology is given by equation (2.10).

$$T = \{(x, y) \in \mathbb{R}_+^m \times \mathbb{R}_+^s | x = \sum_{j=1}^n x_j l_j + S^-,\ y = \sum_{j=1}^n y_j l_j - S^+, \\ such\ that\ l_j \geq 0, j = 1, \dots, n,\ \sum_{j=1}^n l_j = 1,\ \ S^- \geq 0,\ S^+ \geq 0\}, \quad \textbf{(2.10)}$$

The VRS technology translated by expression (2.11) differs from the one described in (2.10) through the addition of the condition $\sum_{j=1}^n l_j = 1$, which limits the comparison of a given DMU j with those in the best-practices frontier that are closest to it in terms of scale of operation. This condition is also known as the convexity condition. Graphically, the VRS DEA frontier can be represented in the two-dimensional space (when the number of indicators so allow) by piece-wise linear segments connecting all efficient DMUs in an enveloping way of the remaining, and inefficient, DMUs (see Figure 2.4 for a graphic representation of the CRS versus the VRS DEA technologies).

The VRS technology of DEA distinguishes from the free disposable hull (FDH), which is also a nonparametric frontier technology for performance measurement (for more information on the FDH we refer the reader to Cooper et al., 2007) through the continuity of the frontier, i.e., any linear combinations among DMUs in the frontier can also be referred to as best-practices, which in the case of the FDH technology it is not possible. The FDH technology is therefore more adequate to use in some specific situations, i.e., situations where it does not make much sense to make comparisons with best-practices rather than those that can be effectively observed (Cooper et al., 2007), or simply because that is the intention of the decision-maker. Mathematically, the FDH technology is represented very similarly to the CRS DEA technology indicated by the equation (2.9). The only difference is that the intensity vector of l_j can only assume integer values, either of 1, if the DMU belongs to the frontier, or 0 if otherwise. Therefore, the efficiency scores under the FDH technology will be computed using a mixed integer linear programming (MILP) model rather than a linear programming (LP) model as in the case of DEA, and the corresponding schematic representation (in the cases where the number of indicators so allow) will remind that of a staircase.

The relationship between the VRS and CRS performance measures obtained for a given DMU gives an indication about its scale efficiency (in case it is of interest to obtain such knowledge), or, in other words, the adequacy of its scale of operation, i.e., how much is potentially being lost (either in terms of resources or outcomes) due to the current scale of operation. The scale efficiency (SE) is given by the equation (2.11), as follows.

$$SE = \frac{CRS\ efficiency}{VRS\ efficiency} \quad \textbf{(2.11)}$$

The CRS efficiency is also known as overall efficiency (or overall performance measure) and the VRS efficiency is known as the pure technical efficiency. If the scores obtained under both the CRS and VRS technologies are equal, then the DMU is operating at the most productive scale size and, as such, all inefficiency eventually observed derives from technical aspects and cannot be attributed to the scale of operation.

In the following we present another numerical example to illustrate the issue of the CRS versus the VRS technologies and the issue of scale efficiency (see Table 2.4 for the data used and Table 2.5 for the DEA application results).

Table 2.4 – Data regarding a sample of 12 DMUs (numerical example number 2).

DMU	A	B	C	D	E	F	G	H	I	J	K	L
Input 1	20	19	25	27	22	55	33	31	30	50	53	38
Output 1	100	150	160	180	94	230	220	152	190	250	260	250

Table 2.5 – Input-oriented, radial, CRS- and VRS-DEA performance measures and scale efficiency (SE).

DMU	A	B	C	D	E	F	G	H	I	J	K	L
CRS_Eff	0.633	1	0.811	0.844	0.541	0.530	0.844	0.621	0.802	0.633	0.621	0.833
VRS_Eff	0.95	1	0.836	0.915	0.864	0.622	0.979	0.625	0.887	0.76	1	1
SE	0.667	1	0.970	0.923	0.627	0.852	0.863	0.993	0.905	0.833	0.621	0.833

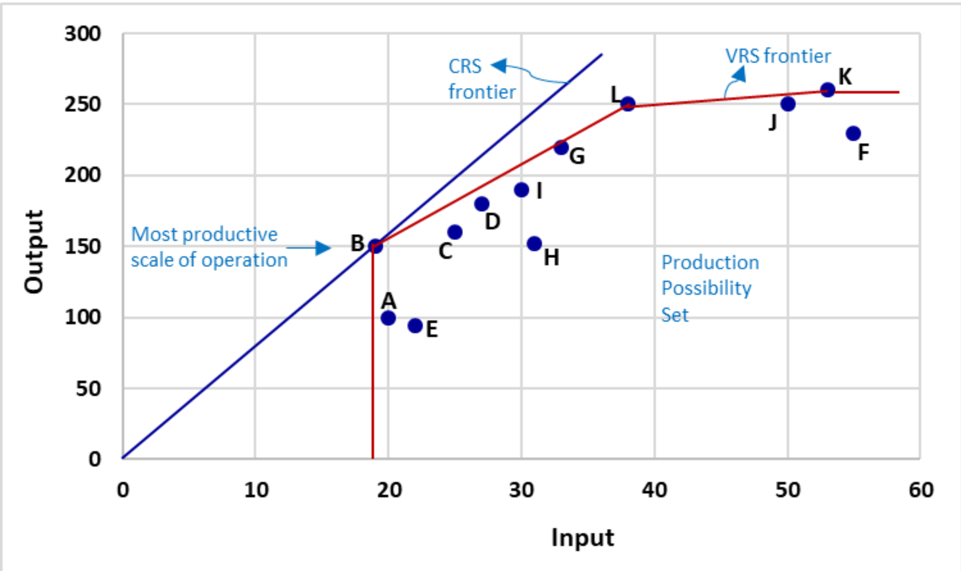


Figure 2.4 *– Graphic representation of the CRS versus the VRS DEA technologies given the example of 1input-1output indicators.*

From both Table 2.5 and Figure 2.4 we can observe that only **DMU B** is operating at the best scale of operation since its performance is equal under a VRS and a CRS technology assumption. Also, we can observe that with the less restrictive VRS technology, either **DMU L** and **DMU K** are considered efficient. Or, we can simply say that these DMUs are efficient in a pure technical sense. In this case they are operating at a scale of operation that does not give them the opportunity to perform at their best potential. Particularly, they are operating under decreasing returns to scale, which means that the amount of outputs that can possibly be produced (or attained) increase very little (if at all) in spite of a potential considerable increase in the input usage. These DMUs would gain in performance if they could reduce their scale of operation, so that the amount of inputs could give rise to a proportional amount of outputs rather than less than proportional.

As previously mentioned, the above presentation leaves out many DEA models, DEA model derivations, and potential applications. For example, in Chapter 3, it will be explored the use of a composite indicator, which in a general way corresponds to a DEA output-oriented model but with a so-called dummy input, i.e., an input that has the value of 1 for all DMUs. However, this composite indicator will be rather constructed on the basis of a bit more complex technology, which has not been explored here, i.e., a directional distance function (DDF), so that the output indicators used can be either maximized or minimized according to their nature and how they can improve (i.e., either "the more the better" or "the less the better"). The composite indicator model will integrate weight restrictions as developed by Zanella et al. (2015a,b). The use of weight restrictions was briefly mentioned in this small introduction. In Chapter 5, it will be explored an input-oriented VRS model under a robust framework that was first developed by Cazals et al. (2002). In Chapter 6, it will be used a DEA-based model in its envelopment formulation where there will be considered restrictions to the selection of the most adequate peers for each given DMU and therefore derive the efficiency scores, in alignment with the study of Golany and Thore (1997). In Chapter 7, it will be conducted a performance measurement over time. More precisely, the methodologies developed by Färe et al. (1992), Wheelock and Wilson (2003), and Pastor and Lovell (2005) will be used to measure productivity change over time.

Hopefully, this introductory section to performance measurement and data envelopment analysis may provide some assistance to the successful reading of the remaining chapters of the thesis, particularly those that make direct use of such techniques.

References

Admiraal, R. J., van Helden, G. J. (2003). Benchmarking in the Dutch wastewater treatment sector. *Public Money and Management*, 23(2), 113-118.

Amado, C.A.F., Santos, S.P., Marques, P.M. (2012). Integrating the Data Envelopment Analysis and the Balanced Scorecard approaches for enhanced performance assessment. *Omega*, 40, 390-403.

Bogetoft, P. & Otto, L. (2011). Benchmarking with DEA, SFA and R. *Springer*.

Cazals, C., Florens, J. P., and Simar, L. (2002). Nonparametric frontier estimation: A robust approach. *Journal of Econometrics*, 106 (1), 1-25.

Charnes, A., Cooper, W.W., Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444.

Cook, W., Tone, K., Zhu, J. (2014). Data envelopment analysis: prior to choosing a model. (Technical note), *Omega*, 44, 1-4.

Cooper, W.W., Seiford, L.M., Kaoru, T. (2007). Data Envelopment Analysis. A comprehensive text with models, applications, references and DEA-solver software. Second Edition. *Springer*.

Dyson, R. G., Shale, E. A. (2010). Data Envelopment Analysis, operational research and uncertainty. *Journal of the Operational Research Society*, 61, 25-34.

Färe, R., Grabowski, R., Grosskopf, S., Kraft, S. (1997). Efficiency of a fixed but allocatable input: a non-parametric approach. *Economic Letters*, 56, 187-193.

Färe, R., Grosskopf, S. (2000). Network DEA. *Socio-Economic Planning Sciences*, 34, 35-49.

Färe, R., Grosskopf, S., Lindgren, B., Roos, P. (1992). Productivity changes in Swedish Pharmacies 1980-1989: A non-parametric Malmquist approach. *The Journal of Productivity Analysis*, 3, 85-101.

Färe, R., Lovell, C.A.K. (1978). Measuring the technical efficiency of production. *Journal of Economic Theory*, 19, 150-162.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*. Series A (General), 120(3), 253-290.

Golany, B., Thore, S. (1997). Restricted best practice selection in DEA: An overview with a case study evaluating the socio-economic performance of nations. *Annals of Operations Research*, 73, 117-140.

Kaplan, R.S., Norton, D.P. (1992). The balanced scorecard – measures that drive performance. *Harvard Business Review*, 70(1), 71-79.

Kuo, Y.H., Kusiak, A. (2019). From data to big data in production research: the past and future trends. *International Journal of Production Research*, 57(15-16), 4828-4853.

Liu, J.S., Lu, L.Y.Y., Lu, W.M. (2016). Research fronts in data envelopment analysis. *Omega*, 58, 33-45.

Liu, J.S., Lu, L.Y.Y., Lu, W.M., Lin, B.J.Y. (2013). A survey of DEA applications. *Omega*, 41, 893-902.

OECD (2011). Water in OECD countries: a multi-level approach – highlights.

Otley, D. (1999). Performance management: a framework for management control systems research. *Management Accounting Research*, 10, 363-382.

Wheelock, D. C., Wilson, P. W. (2003). Robust nonparametric estimation of efficiency and technical change in U.S. commercial banking. *Unpublished working paper. The John E. Walker Department of Economics, 222 Sirrine Hall, Clemson University, Clemson, South Carolina 29634, USA.*

Zanella, A., Camanho, A.S., Dias, T.G. (2015a). The assessment of cities livability integrating human wellbeing and environmental impacts. *Annals of Operations Research*, 226, 1, 695-726.

Zanella, A., Camanho, A.S., Dias, T.G. (2015b). Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. *European Journal of Operational Research*, 245, 2, 517-530.

Zhou, H., Yang, Y., Chen, Y., Zhu, J. (2018). Data envelopment analysis application in sustainability: the origins, development and future directions. (Invited Review), *European Journal of Operational Research*, 264, 1-16.

Zhou, P., Ang, B.W., Poh, K.L. (2008). A survey of data envelopment analysis in energy and environmental studies. (Invited Review), *European Journal of Operational Research*, 189, 1-18.

Zhu, J. (2020). DEA under big data: data enabled analytics and network data envelopment analysis. Annals of Operations Research, https://doi.org/10.1007/s10479-020-03668-8.

PART II

SCIENTIFIC PAPERS

Chapter 3 –

Performance Benchmarking using Composite Indicators to support Regulation of the Portuguese Wastewater Sector

Chapter 3

Performance benchmarking using composite indicators to support regulation of Portuguese wastewater sector

This chapter refers to the following paper:

Henriques, Alda A.; Camanho, Ana S.; Amorim, Pedro; Silva, Jaime G. (2020). Performance benchmarking using composite indicators to support regulation of the Portuguese wastewater sector. *Utilities Policy*, 66, 101082.

Overarching research question to be explored with this paper:

RQ1: How to contribute for enhancements in the wastewater sector governance so as to promote improved corporate social responsibility and a wiser use of resources, including energy, at a utility level?

Authors:

Alda A. Henriques, Ana S. Camanho, Pedro Amorim, Jaime G. Silva

Thesis' author contribution:

Conceptualization (collaboration), data curation, software usage (except the model implementation), formal analysis (main collaboration), investigation, writing – original draft (except section 3.2.1), writing – review and editing (collaboration), funding acquisition (collaboration).

Performance benchmarking using composite indicators to support regulation of the Portuguese wastewater sector

Abstract

This paper develops a benchmarking framework to support sunshine regulation in the water sector. It uses benefit-of-the-doubt composite indicators formulated with a directional distance function. Weight restrictions are incorporated in the model to account for different perspectives in the performance assessment. The framework is tested using data of the Portuguese regulation authority concerning the activity of wastewater operators. The information obtained using this framework reflects overall performance at the firm level and complements traditional evaluations of regulatory authorities based on the analysis of individual indicators. The results can be used to disseminate best practices, motivate continuous improvement, and foster enhancements in the governance of regulated utilities.

Keywords: Wastewater operators; Composite indicators; Sunshine regulation.

3.1. Introduction

Efficiency and effectiveness, alongside with trust and engagement, constitute the pillars of the twelve OECD principles on water governance. They are essential to cope with water challenges worldwide (OECD, 2015). One of the challenges of governance in the water sector is the integrated management of various aspects of services (e.g., the interface between operators and users, as well as between operators' activity and public health/environmental issues). Other concerns are for controlling the cost of service and keeping prices at an affordable level for all users. Asset management is also a key challenge for this sector, requiring the prioritization of systems' interventions to ensure service quality despite the aging of the infrastructures. Finally, there is a growing need for the efficient use of resources, particularly energy, in response to more demanding environmental restrictions and the increase in energy prices (Baptista, 2014).

At the utility level, efficiency concerns the use of the resources required to deliver the level of service expected by consumers. Effectiveness relates to the level of satisfaction of users' needs (Vilanova et al., 2015). At the regulatory governance level, efficiency depends on "how" regulatory governance is structured, and effectiveness involves "what" regulatory governance does (Marques & Pinto, 2018). In the context of a regulatory system, the "how" corresponds to the legal settings that are designed to frame the "what", and the "what" corresponds to the concrete decisions and actions that are put in place to achieve the regulatory goals (Marques & Pinto, 2018). Once the "how" is adequately established, the "what" must ensure high-performance levels in the sector.

Regulatory authorities play a significant role in the governance of the water sector. They are responsible for the supervision of operators' activities, including water supply and wastewater collection and treatment, and the deployment of policies to balance the needs and expectations of all stakeholders of water-related activities (Akhmouch & Correia, 2016). The role of the regulator is also to ensure the overall quality of service that is provided to users, which requires monitoring all dimensions of the sectors' activity to ensure global sustainability, independently of the management model of the operator. The sustainability of the sector includes Social Sustainability, i.e., accounting for the protection of users' interests, Economic Sustainability, including the management of infrastructures and human resources, and Environmental Sustainability, concerning the use of natural resources and the control of impacts in terms of pollution.

Benchmarking has been recognized as an essential strategy to foster improvements within the water sector. These are sought both in terms of its functioning and sustainability (for an overview of performance assessment and benchmarking in the water sector, please refer to Appendix 3A). The public disclosure of benchmarking results is a means for increased transparency within the sector. Furthermore, the use of benchmarking can also help managing conflict in the design and implementation of policies by redirecting the attention of stakeholders to performance improvements (Berg, 2007).

Particularly in the context of regulatory governance, the role of benchmarking can be seen from two different perspectives. First, it provides an overview of performance within the sector, empowering the regulatory authority to guide decision making. Second, when there is a public disclosure of results, which is inherent to a sunshine regulation strategy, it also introduces artificial competition in sectors characterized by a natural monopoly nature, providing incentives for improvements (Pinto et al., 2017). However, the impact of the disclosure of benchmarking results depends on how the information is shared with the public. Passing a clear and straightforward message to the utilities and the general public is essential to achieve the expected effects of public disclosure of benchmarking results.

Utilities and regulatory authorities commonly use performance indicators (PIs) as a baseline for monitoring performance. They can also be used as decision support tools to prioritize actions and analyze the impact of measures previously implemented (Vilanova et al., 2015). Despite the generalized acceptance of individual indicators as measures of performance, there are a few problems associated with their use. First, each ratio examines only a part of the utilities' activity, and it is difficult to gain an overall view of performance when the number of PIs collected for each utility is large. The ranking of utilities is also impaired, particularly when not all ratios indicate a similar level of performance for each utility. Another limitation of individual indicators is that they cannot be used straightforwardly to set performance targets. Although any particularly poor value of an indicator identifies an aspect of the activity in particular need for improvement, the target performance levels cannot be estimated with confidence, as a target for one indicator may have implications for others.

A meaningful comparison of the multiple-dimensional activities of regulated utilities requires the analysis of a set of PIs. Although the collection of data and computation of a comparable set of indicators for regulated utilities poses some challenges, the aggregation of those indicators into an overall measure of performance may be an even more complicated process, involving creativity and experienced judgment. Nevertheless, it is instrumental in informing public opinion and promoting discussion on the potential for improvements (Vilanova et al., 2015).

According to Mehta and Mehta (2013), the full positive potential impact that benchmarking tools can bring to regulated sectors such as urban water supply and sanitation is not yet realized. Systematic performance measurement efforts may contribute to improve the quality and reliability of the information, to strengthen the link between monitoring and improvement, and ensure the systems' sustainability in the long run. In this context, the development of enhanced performance assessment tools to monitor water-related functions, activities, or institutions can be very relevant to help managing the water challenges described in OECD (2012).

The studies of Marques et al. (2015) and Molinos-Senante et al. (2016b) are two examples of studies focusing on a holistic analysis of the sustainability of water utilities using Multi-Criteria Decision Analysis (MCDA). They estimated an aggregated measure of sustainability, such that a single sustainability level is assigned to each utility based on the information provided by several PIs. Both studies were applied to the Portuguese water sector.

In the quest for improved sustainability in the water sector, the quality of service provided to users is an aspect of utilities' performance that has gained increased attention in recent years. Giannakis et al. (2005) demonstrate the importance of integrating the quality of service in regulatory benchmarking procedures rather than focusing only on cost efficiency. Baptista (2014) also argues that the quality of service cannot be dissociated from the economic component of utilities' activity. Synergies between these components should be explored to improve utilities' performance.

Pinto et al. (2017) developed an innovative approach to provide a holistic view of the quality of service provided by water utilities. The approach consists of the construction of a quality of service index (QSI), which aggregates all the information embedded in a panoply of performance indicators. The QSI is obtained recurring to MCDA, involving the active participation of decision-makers. The method enables assigning categories representing the level of quality of service provided to each utility. This method is illustrated using data on water supply services publicly disclosed by the Portuguese Water and Waste Sector Regulatory Authority (ERSAR).

Stemming from the importance that quality of service has in sustainability improvements within the water sector, and building from the reported interest in enhanced performance assessment tools, this paper proposes an alternative method to that developed by Pinto et al. (2017) to holistically evaluate the quality of service provided to users within the water sector. However, the focus of our study is on wastewater operators rather than in water supply operators. Furthermore, the method proposed is based on Data Envelopment Analysis (DEA) instead of MCDA. One core methodological difference is that our method is data-driven and does not require strong interaction with stakeholders to decide upon the relative importance of the indicators. From a modeling perspective, this paper is based on a composite indicator framework using the directional distance function benefit-of-the-doubt model of Zanella et al. (2015a,b).

Two distinctive features of the approach developed by Zanella et al. (2015a,b) determined its adoption for the analysis of wastewater operators' performance: the ability to consider both desirable and undesirable factors measured in their original scales, and the possibility to incorporate weight restrictions to reflect minimum bounds for the relative importance of individual indicators or groups of indicators from a given dimension. These features enabled conducting an exploratory study of alternative perspectives of wastewater operators' activity, corresponding to "Adequacy of the Interaction with Users", "Service Management Sustainability" and "Environmental Sustainability". These perspectives are aligned with the three pillars of sustainability (social, economic and environmental), allowing a comprehensive multidimensional evaluation of performance. Through a judicious choice of weight bounds, decision makers / regulatory authorities can specify the relative importance of the different sustainability dimensions and of individual indicators within each dimension, while still granting some flexibility in the weights underlying the efficiency assessment. This allows highlighting the strengths of individual operators and account for the adoption of specific strategies.

This approach was first applied to regulated utilities (Brasilian hydropower plants) in Calabria et al. (2018), but it has never been applied in the context of water sector regulation. Its applicability is demonstrated using the Portuguese sunshine regulation context. Furthermore, as the performance of utilities is influenced by the context of the operation, our study includes a second-stage analysis to gain insights regarding the impact of contextual factors on the quality of service provided by the utilities.

The approach proposed aims to contribute to improved governance and diffusion of best practices by complementing the insights provided by the analysis of individual PIs in a sunshine regulation context. It allows flexibility in the assessment of different objectives under a formative regulatory framework and provides insights concerning the key performance drivers among contextual factors. This approach can also be used in other regulatory settings, as it can be seen as a complementary formative tool promoting continuous improvement in the context of normative regulation.

The outcomes of the approach are three-fold: (i) identification of benchmarks (i.e., efficient performers in at least one perspective of analysis); (ii) information regarding each operator's distance to the efficient frontier, and (iii) identification of the contextual factors whose impact on efficiency is statistically significant.

The framework proposed in this paper goes one step forward in terms of the information provided to support formative performance evaluations. It is focused on the assessment of performance at the operator level, contributing to increased transparency and accountability in the water sector since it allows obtaining an aggregate (and thus simplified) vision of the information contained in a large set of indicators. According to OECD (2015), transparency is a significant asset in what concerns the trust and engagement of stakeholders within the water sector, including users, operators, investors, and policy-makers. The framework proposed also strengthens the link between performance monitoring and continuous improvement.

The structure of this paper is as follows. Section 3.2 describes the methodology used in this study. Section 3.3 describes the case study, providing an overview of the Portuguese water sector regulatory model and data available. Section 3.4 presents the results and discusses their managerial implications, giving particular emphasis to the potential use of the information for regulatory purposes. The paper finishes with the presentation of the core conclusions, also discussing the main limitations of the study and suggesting avenues for further research.

3.2. Methodology

The performance assessment framework developed in this paper is structured in two stages. The first stage involves the construction of a composite indicator (CI) using a directional distance function (DDF) to obtain an overall measure of operators' performance. The second stage is a contextual analysis, based on the use of nonparametric hypothesis tests, to identify the factors that significantly affect performance.

3.2.1 Construction of the composite indicator

This section presents the formulation of the CI model used to assess the quality of service provided by wastewater operators. The model considers a dummy input, identical for all decision-making units (equal to one), and a set of outputs. The outputs correspond to the performance indicators that are used by the regulator to evaluate the quality of service provided by the wastewater operators. These include both desirable and undesirable indicators. For desirable indicators, larger values correspond to better performance, whereas for undesirable indicators lower values correspond to better performance.

Given the objective of aggregating several PIs into a summary measure of performance, we used the Directional benefit-of-the-doubt (BoD) CI formulation of Zanella et al. (2015a). The advantage of this formulation in comparison with the standard DEA formulation of BoD composite indicators (Cherchye et al., 2007) is to allow simultaneous adjustments to desirable and undesirable factors, without requiring prior adjustments to the measurement scales. Also, the output orientation of this formulation allows direct incorporation of weights restrictions to reflect the relative importance of the output indicators considered in the assessment.

The weights formulation of the Directional BoD CI model is shown in (3.1).

Weights formulation of the directional BoD CI

$$CI_k = min - \sum_{r=1}^{s} y_{rk}u_r + \sum_{i=1}^{m} x_{ik}v_i + w \quad \textbf{(3.1)}$$

$$s.t. \sum_{r=1}^{s} g_y u_r + \sum_{i=1}^{m} g_x v_i = 1 \quad \textbf{(3.1.1)}$$

$$-\sum_{r=1}^{s} y_{rj}u_r + \sum_{i=1}^{m} x_{ij}v_i + w \geq 0, j = 1, \dots, n \quad \textbf{(3.1.2)}$$

$$u_r \geq 0, \quad r = 1, \dots, s \quad \textbf{(3.1.3)}$$

$$v_i \geq 0, \quad i = 1, \dots . m \quad \textbf{(3.1.4)}$$

$$w \geq 0 \quad \textbf{(3.1.5)}$$

In formulation (3.1), the decision variables are the desirable and undesirable output weights (u_r and v_i), respectively, and the variable (w) associated with the dummy input. y_{rj} are the desirable indicators (r=1,…s) and x_{ij} are the undesirable indicators (i=1,…,m) associated with operator j (j =1,…,n). The index k signals the DMU under assessment.

The direction in which improvements may be sought is given by the directional vector (g_x, g_y). In the case of the assessment presented in the empirical part of this paper, this vector was specified as being equal to the values of the desirable and undesirable outputs observed in DMU k under evaluation, i.e., $(g_x, g_y) = (x_{ik}, y_{rk})$. Note that according to the formulation presented in (3.1), the directional vector seeks for expansion of desirable indicators and contraction of undesirable indicators. The objective function value CI_k reflects the relative magnitude of inefficiency for operator k in comparison with the peers. It also corresponds to the maximal feasible proportional expansion of desirable indicators and contraction of undesirable indicators that can be achieved simultaneously. Therefore, a value equal to zero corresponds to the best performance level of the sample, whereas values greater than zero signal potential for improvement.

Model (3.1) allows the incorporation of weight restrictions (WR) to reflect the relative importance of the output indicators considered. The aim of applying WR in a formative evaluation of the quality of service provided by wastewater treatment operators is to ensure that all indicators in all dimensions of performance are taken into account in the computation of the overall measure of performance. Moreover, the use of the WR allows exploring alternative perspectives of performance, attributing different emphasis to specific groups of indicators (i.e., corresponding to different dimensions of performance). This allows revealing the dimensions of performance in which each operator excels as well as those in need of improvement.

The formulation of the weight restrictions used in the empirical part of this paper is presented in (3.2), (3.3), and (3.4). These WRs are applied only to an artificial operator, whose indicators $(\bar{x}_i, \bar{y}_r)$ are equal to the average of the values observed for each indicator (x_{ij}, y_{rj}) in the set of operators under assessment ($j = 1, \dots, n$) (see Zanella et al., 2015a). The weight restrictions can be imposed to groups of desirable indicators that belong to mutually exclusive categories $C_d, d = 1, \dots l$, and groups of undesirable indicators that belong to mutually exclusive categories $G_u, u = 1, \dots q$, corresponding to a given subset of restricted indicators g_0 as shown in (3.2).

Alternatively, the restrictions can be imposed for a specific desirable output indicator (r_0), as shown in (3.3), or for a specific undesirable output indicator (i_0), as shown in (3.4).

$$\frac{\sum_{r\in C_d} u_r\bar{y}_r + \sum_{i\in G_u} v_i\bar{x}_i}{\sum_{r=1}^{s} u_r\bar{y}_r + \sum_{i=1}^{m} v_i\bar{x}_i} \geq \delta_{g_0}, \qquad g_0 = 1, \dots, \mathrm{p} \tag{3.2}$$

$$\frac{u_{r_0}\bar{y}_{r_0}}{\sum_{r=1}^{s} u_r\bar{y}_r + \sum_{i=1}^{m} v_i\bar{x}_i} \geq w_{r_0}, \qquad r_0 = 1, \dots, s \tag{3.3}$$

$$\frac{v_{i_0}\bar{x}_{i_0}}{\sum_{r=1}^{s} u_r\bar{y}_r + \sum_{i=1}^{m} v_i\bar{x}_i} \geq w_{i_0}, \qquad i_0 = 1, \dots, m \tag{3.4}$$

For computational ease, equations (3.2), (3.3) and (3.4) can be equivalently written in the format of assurance regions type I (ARI) (see Thompson et al., 1990). The general formulation of this type of linear restrictions is shown in (3.5). In this formulation c_{rz} and k_{iz} are the parameters associated to the desirable output weights and undesirable output weights, respectively. One restriction of this type is required for each output dimension aggregating several indicators ($d = 1, \dots, l$), as well as for each desirable output ($r_0 = 1, \dots, s$) and undesirable output ($i_0 = 1, \dots, m$).

$$\sum_{r=1}^{s} c_{rz}u_r + \sum_{i=1}^{m} k_{iz}v_i \geq 0, \qquad z = 1, \dots. (l + s + m) \tag{3.5}$$

The incorporation of these weight restrictions in the weights formulation of the directional BoD CI involves adding restrictions (3.5) to model (3.1).

To identify the peers underlying the benchmarking exercise, the envelopment (dual) formulation of the directional BoD CI model with WR is formulated as shown in (3.6).

$$CI_k = \max \beta_k \tag{3.6}$$

$$s.t. \sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{rk} + \beta_k g_y + \sum_{z=1}^{s+m+l} t_z c_{rz}, \qquad r = 1, \dots, s \tag{3.6.1}$$

$$\sum_{j=1}^{n} x_{ij}\lambda_j \leq x_{ik} - \beta_k g_x - \sum_{z=1}^{s+m+l} t_z k_{iz}, \qquad i = 1, \dots, m \tag{3.6.2}$$

$$\sum_{j=1}^{n} \lambda_j \leq 1 \tag{3.6.3}$$

$$\lambda_j \geq 0, \qquad j = 1, \dots, n$$

$$t_z \geq 0, \qquad z = 1, \dots. (l + s + m)$$

Note that in this formulation the decision variables (t_z) are originated by the weight restrictions that exist in the primal formulation (i.e., the restrictions shown in (3.5)). The decision variable β_k reflects the relative magnitude of inefficiency of operator k in comparison with the peers. λ_j are intensity variables that allow obtaining a point on the frontier of the Production Possibility Set against which the operator under assessment (k) is compared to determine its distance to the frontier. This point is obtained as a linear combination of the values of the indicators y_{rj} and x_{ij} observed in peers (see left-hand side of restrictions (3.6.1) and (3.6.2)).

3.2.2 Treatment of outliers

Given the deterministic nature of nonparametric models for the assessment of performance, the construction of a best-practice frontier is very sensitive to the presence of outliers. Extreme observations are a reason of concern in any empirical methods, particularly in Data Envelopment Analysis and Directional Distance Functions, because an atypical observation may shift the position of the frontier and have a significant impact on the relative efficiency score estimated for other firms. Given the importance of estimating robust frontiers (i.e., well populated and not too distant from the bulk of firms in the analysis) in regulatory settings, our methodology involved a preliminary stage dedicated to the identification of potential outliers. If the presence of outliers is confirmed, the outperforming operators are removed from the sample and the overall measure of operators' performance is obtained considering the reduced sample. This procedure aims to prevent the misevaluation of the majority of operators due to the impact that a few operators may have on the position of the frontier.

The mitigation of the effect of outliers in DEA could be done using order-*m* methods, order-α methods, or robust approaches (Cazals et al., 2002; Aragon et al., 2005; Simar, 2003; Daraio and Simar, 2005,2007). However, the bootstrapping procedures underlying these approaches are more difficult to be accepted in regulatory settings, as it is not possible to identify in detail the set of operators underlying the construction of the best-practice frontier of the benchmarking exercise. Therefore, in this study we follow the procedure proposed by Wilson (1995) to detect influential observations in DEA, with a few adaptations to take into account the specificities of our empirical study.

Our method for the identification of outliers consists of two steps. First, we identify all candidate outliers for each PI in the sample analyzed. The candidate outliers correspond to the operators whose value in that PI is higher than the third quartile plus 1.5 times the interquartile range. Then, a sensitivity analysis is conducted by removing, one by one, the candidate outliers that are located in the frontier of the production possibility set for the composite indicator model estimated. The CI estimates for the original sample are compared with the estimates for the sample without the candidate outlier operator using the criteria described below. The thresholds defined for the removal of potential outlier operators from the sample are also described.

Criterion 1: Percentage of operators that change the CI score with the removal of the candidate outlier, calculated as shown in (3.7). $\boldsymbol{n_{OR}}$ is the number of operators for which the CI has a different value in the original sample and in the sample with the candidate outlier removed. $\boldsymbol{N_{OR}}$ is the total number of operators in the reduced sample (with the candidate outlier removed).

$$\boldsymbol{\alpha = \frac{n_{OR}}{N_{OR}}} \tag{3.7}$$

Criterion 2: Average change in CI score for the operators that changed score due to the removal of the candidate outlier operator from the sample, measured as shown in (3.8). $\boldsymbol{CI_i^{OS}}$ and $\boldsymbol{CI_i^{OR}}$ are the CI scores of operator i ($i = 1, \dots, N_{OR}$) in the original sample (OS) and in the sample with the candidate outlier removed (OR), respectively.

$$\boldsymbol{\delta = \frac{\sum_i^{N_{OR}} CI_i^{OS} - CI_i^{OR}}{n_{OR}}} \tag{3.8}$$

Criterion 3: Variation in the percentage of efficient operators that define the frontier, calculated using expression (3.9). $\boldsymbol{EF_{OR}}$ and $\boldsymbol{EF_{OS}}$ are the number of efficient operators in the original sample and in the sample with the candidate outlier removed, respectively.

$$\boldsymbol{\phi = \frac{EF_{OR}}{EF_{OS}}} \tag{3.9}$$

In order for a candidate outlier to be removed from the original sample of operators, it must have high scores in the three criteria described above. The thresholds considered in the specific case of this assessment are as follows: an operator should have a minimum value of 60% on criterion 1, a minimum value of 2.5% on criterion 2 and a minimum value of 1.2 on criterion 3 to be considered an outlier with an influential effect on the performance assessment (see expression (3.10)).

$$\boldsymbol{\alpha} \geq \mathbf{60\%} \wedge \boldsymbol{\delta} \geq \mathbf{2.5\%} \wedge \boldsymbol{\phi} \geq \mathbf{1.2} \qquad \textbf{(3.10)}$$

3.2.3 Contextual analysis

The second stage of the framework aims to provide insights upon the impact of the operating context on the performance of the operators. The differences observed in the performance levels of the operators may be partially explained by the influence of these contextual conditions.

The analysis of the factors contributing to efficiency differences among decision-making units is a relevant area of research in Data Envelopment Analysis (DEA). In this context, Badin et al. (2014) discuss the state of the art of the possible approaches that include contextual variables in the nonparametric efficiency measurement techniques such as DEA (i.e., the one-stage, the two-stage, and the nonparametric conditional approaches). The advantages of the nonparametric conditional approach proposed by Daraio and Simar (2005), often implemented with the order-*m* approach for the construction of the frontier, is particularly emphasized. The order-*m* method has been used in several performance assessment studies within the water sector (e.g., Carvalho and Marques, 2011; De Witte and Marques, 2010; Fuentes et al., 2015; Guerrini et al., 2016; Marques et al., 2014). Similarly, two-stage approaches using nonparametric tests in the second stage have also been extensively used in this sector (e.g., Dong et al., 2017; D'Inverno et al., 2018; Molinos-Senante et al., 2015b).

As in our study the focus of the analysis is on the estimation of a composite indicator under alternative perspectives, based on regulators' data for 14 benchmark indicators, we conducted the contextual analysis adopting a two stage-approach. Following this procedure, the first stage would be more comparable to the sunshine regulation approach currently adopted by ERSAR, which doesn´t take into account contextual conditions in the direct comparison of individual indicators among operators.

In the second-stage, if only two groups of operators are considered in terms of a contextual variable, then the Mann-Whitney hypothesis test is used. Else, when the operators can be subdivided into three or more groups in terms of the contextual variable under analysis, the Kruskal-Wallis test is used (Kruskal and Wallis, 1952; Ruxton and Beauchamp, 2008). This approach follows the procedure used in Molinos-Senante et al. (2014c). Accordingly, the null hypothesis is that the groups of operators under evaluation have the same distribution for the value of the composite indicator, and the alternative hypothesis is that some groups of operators have different distributions for the value of the composite indicator. All tests were performed with a level of significance of 5% (i.e., the alternative hypothesis is accepted in case the p-value is lower than or equal to 0.05; otherwise, the null hypothesis is retained).

3.3. Case study

The formative performance assessment framework proposed in this paper was applied to the Portuguese wastewater regulatory context. This is part of a broader regulatory context that covers both water (including water and wastewater) and waste services.

3.3.1 Portuguese regulatory model for water and waste sectors

Two important aspects that dictate the role and activity of the regulatory authority are the challenges faced by the sector, as described in the Introduction, and the objectives that are defined for tackling those challenges.

The model used by the Portuguese regulatory authority for water and waste sectors (ERSAR) consists of two interdependent approaches, under an integrated regulatory perspective known as ARIT-ERSAR (Baptista, 2014). The first approach concerns structural regulation and is sector oriented, i.e., the "how" according to the definition of Marques & Pinto (2018). It includes the organization of the sector, legal aspects, information, technical assistance and training. The second approach concerns behavioral regulation and is directed to each utility. It comprises several dimensions, including legal and contract-related aspects, economic issues, quality of service, quality of water for human consumption and interaction with users. These two approaches are expected to interact and evolve over time (Baptista, 2014). The regulatory approach adopted by ERSAR is characterized by collaborative mediation rather than a restrictive and sanctioning intervention. In the regulatory approach followed by ERSAR, behavioral regulation, adopting a Sunshine regulation model, has an important impact in terms of the operators' activity. The benchmarking exercises are conducted periodically (every year) and the results are publicly displayed. The data is provided to the regulatory authority by each operator. It is then used to compute performance indicators for the evaluation of the operators' quality of service. The regulatory strategy of ERSAR consists of formative supervision of the operators, including the provision of specific guidance whenever the operators are more vulnerable in terms of management activities (Baptista, 2014). Thus, ERSAR assumes a relevant role in the development of the sector, rather than only supervision and control.

In terms of wastewater operators, which are the focus of the analysis conducted in this paper, the Portuguese market structure is subdivided into wholesale (bulk systems) and retail (end-user/drainage systems) sectors. Currently, the wholesale sector has 12 operators. From these, 10 operators correspond to concessions, which cover 96% of the population (ERSAR, 2019). Regarding the retail sector, there are 257 operators. In terms of ownership, direct management covers 59% of the population, management by concessionary companies covers 17% of the population and management by delegation covers 24% of the population.

Although our study is focused on the wastewater sector, the quality of service is evaluated both in terms of the water transmission and distribution, as well as the wastewater collection, treatment and disposal, thus covering the different stages of the urban water cycle (ERSAR, 2019). Fourteen performance indicators are currently used to comparatively evaluate the performance of each of these water-related activities (ERSAR, 2019).

3.3.2 Data collected

The assessment conducted in this paper used the fourteen performance indicators (PIs) collected by ERSAR to evaluate the quality of service provided to users (ERSAR, 2019, p.218). The data is publicly available on the ERSAR website[1]. It corresponds to the year 2018, whose analysis is described in the annual report on water and waste services in Portugal for the year of 2019 (ERSAR, 2019). Our analysis is focused on wastewater services, both for the wholesale and retail sectors.

Table 3.1 presents the PIs used for the construction of the composite indicator. These are the indicators used by ERSAR for the comparative evaluation of retail and wholesale operators, through benchmarking exercises conducted at the indicator level. The indicators are classified

[1] http://www.ersar.pt/pt/site-publicacoes/Paginas/edicoes-anuais-do-RASARP.aspx

by ERSAR in three dimensions of quality of service: Adequacy of the Interaction with Users (AIU), Service Management Sustainability (SMS), and Environmental Sustainability (ES). In this study, the indicators are further distinguished between desirable (y) and undesirable (x), depending on whether they should be increased or decreased to enhance the quality of service. Further information regarding each PI is available in ERSAR (2019).

Table 3.1 – Indicators used to evaluate wholesale and retail wastewater operators (ERSAR, 2019).

Dimension	Indicators
Adequacy of the Interaction with Users (AIU)	AR01 (y) – Service physical accessibility (%)
	AR02 (y) – Service economical accessibility (%)
	AR03 (x) – Floods occurrence [W: no./100 km sewer.year; R: no./1000 branches.year]
	AR04 (y) – Reply to suggestions and complaints (%)
Service Management Sustainability (SMS)	AR05 (y) – Coverage of total costs (%)
	AR06 (y) – Service subscription (%)
	AR07 (y) – Sewers rehabilitation (%/year)
	AR08 (y) – Sewers failures (number/100 km.year)
	AR09 (y) – Adequacy of staff (W: number/millions m^3.year; R: number/100 km.year)
Environmental Sustainability (ES)	AR10 (x) – Energy efficiency of pumping stations (kWh/m^3.100m)
	AR11 (y) – Treatment physical accessibility (%)
	AR12 (y) – Control of emergency discharges (%)
	AR13 (y) – Effectiveness in accomplishing legal parameters of wastewater discharge (%)
	AR14 (y) – Final disposal of sludge (%)

This study used in the second-stage data on the contextual factors characterizing the conditions faced by retail operators. The selection of the contextual factors to be included in the second-stage analysis was based on the following criteria: (i) the studies available in the literature have demonstrated that these factors have a potential influence on wastewater operators' performance; (ii) the factors are independent of the 14 indicators used for the construction of the CI in the first-stage; (iii) the contextual factors are non-discretionary to managers, at least in the short-term; (iv) the data for the contextual factors are collected by the regulator and publicly available in the ERSAR website.

Table 3.2 shows the contextual variables selected for the second stage analysis of the wastewater operators.

Table 3.2 – List of contextual variables (ERSAR, 2019).

	Type of variable		Indicator (ERSAR database)
	Quantitative	Categorical	
z_1 – Wastewater collected (m^3/year)	x		dAR50a (wholesale) / dAR50b (retail)
z_2 – Energy production (MWh/year)	x		dAR59a (wholesale) / dAR59b (retail)
z_3 – Investment subsidies (€/year)	x		dAR81a (wholesale) / dAR81b (retail)
z_4 – Operator's management system		x	dAR02a (wholesale) / dAR02b (retail)
z_5 – Type of geographical area		x	dAR19a (wholesale) / dAR19b (retail)

Variable z_1 (wastewater collected, expressed in m^3/year) depicts the scale size of the operators. The scale size has been widely studied in terms of its impact on performance within the water sector (Abbott and Cohen, 2009). The findings support that, in general, larger scales of operation are beneficial, although from a certain level of dimension the benefits can become exhausted. The operators were subdivided into four groups according to the values observed for this variable: (i) less than the value of the first quartile, (ii) between the value of the first quartile and the value of the second quartile, (iii) between the value of the second quartile and the value of the third quartile, and (iv) greater than the value of the third quartile.

Variable z_2 (energy production, expressed in kWh/year) is related to the structural conditions of the operators' assets, and with energy self-sufficiency. Since energy is an important input for the wastewater activity (Renzetti and Kushner, 2004), its production can contribute significantly for the operators' sustainability. The operators were subdivided into two groups according to the values of this variable as follows: (i) less than the average, and (ii) greater than the average.

Variable z_3 is the amount of investment subsidies attributed to the operators (expressed in €/year). The selection of this variable is straightforward as it intends to acknowledge the significance that the investment subsidies may have on the overall quality of service delivered to the customers of the wastewater operators. Operators were also subdivided into two groups according to variable z_3: (i) less than the average, and (ii) greater than the average.

Variable z_4 corresponds to the operators' management system. The selection of this variable is justifiable in the context of the Portuguese market structure for the water sector (Carvalho and Marques, 2011). Accordingly, operators were subdivided into three groups according to this variable: (i) Management by Concessionary Companies, (ii) Management by Delegation, and (iii) Direct management by the State or Local Authorities.

Variable z_5 is the type of geographical area (predominantly urban, medium urban or rural), which represents both the population density and the type of residuals treated. The population density has been previously analyzed in the literature (e.g., Carvalho and Marques, 2011), showing that low population densities have a negative influence on the performance of both water and wastewater utilities. Variable z_5 was subdivided into three groups: (i) predominantly urban areas, (ii) medium urban areas, and (iii) predominantly rural areas.

3.3.3 Treatment of missing data

The datasets available in ERSAR website are not complete for all PIs and contextual variables. In order to ensure soundness and practical relevance of the results obtained, the operators included in the sample analysed in this study were only those that had reported data for at least two-thirds of the indicators considered in the regulatory framework. The removal of operators due to problems with missing values (MV) was only necessary in the case of the retail sample, and occurred for 57 retail operators. Accordingly, the sample size was reduced from 257 operators to 200 operators. Table 3.3 describes the data relative to the wholesale sample and to the reduced sample of retail operators.

Table 3.3 – Summary information of indicators for the wholesale (W) and retail (R) operators (ERSAR, 2019).

Dimension	Indicators	Unit	Wholesale (12 operators)				Retail (200 operators)			
			Mean	St.Dev.	Best value observed**	% MV	Mean	St.Dev.	Best value observed**	% MV
AIU	AR01	%	96.2	6.0	100	0.0%	77.7	19.6	100	0.0%
	AR02	%	0.20	0.10	0.44	0.0%	0.29	0.12	0.64	0.0%
	AR03*	W: no./(100 km sewer.year); R: no./(1000 branches.year)	32.4	84.0	0	0.0%	3.7	5.9	0	3.0%
	AR04	%	95.1	9.1	100	16.7%	77.6	30.1	100	17.5%

SMS	AR05	%	116.3	30.3	157	66.7%	76.1	31.6	156	6.5%
	AR06	%	93.6	7.4	100	0.0%	86.9	11.6	100	0.0%
	AR07	%/year	0.42	0.85	3	0.0%	0.38	1.19	14.7	0.0%
	AR08*	Number/(100 km.year)	1.84	2.0	0	0.0%	3.25	16.42	0	0.5%
	AR09	W: number/(millions m^3.year); R: number/(100 km.year)	3.6	1.5	6.9	0.0%	8.6	6.1	34.1	0.0%
ES	AR10*	kWh/(m^3.100m)	0.65	0.20	0.35	8.3%	0.92	0.67	0.3	62.0%
	AR11	%	99.6	1.4	100	0.0%	99.0	4.3	100	0.0%
	AR12	%	68.1	32.2	100	0.0%	34.1	43.9	100	53.5%
	AR13	%	94.7	7.8	100	0.0%	72.7	27.9	100	52.5%
	AR14	%	100	0	100	0.0%	94.3	22.8	100	63.0%

* Undesirable indicator

** In the case of undesirable indicators, it corresponds to the minimum value observed; in the case of desirable indicators, it corresponds to the maximum value observed.

After the removal of operators with a high proportion of missing data, a few operators still presented a problem of MV for some indicators. Therefore, we adopted the approach proposed by Kuosmanen (2002) for the treatment of missing data in the DEA analysis, both for the retail and wholesale operators' sample. For the desirable output indicators with missing data, we assigned a small value equal to the minimum value observed in the sample for each indicator. For undesirable indicators with missing data (which are treated as inputs in the directional distance function formulation), we assigned a very large value, equal to the maximum value observed in the sample for each indicator.

By modeling missing data in this way, one is, in fact, assuming that an operator will not weight the factors that are missing (unless by the amount strictly imposed), and therefore its radial distance to the frontier is equal to that obtained from an assessment without considering the missing factors (see also Thanassoulis et al., 2007, for details). This ensures that an operator is not judged as being better than what it is in fact due to not having made available to the regulator data on certain indicators. However, the results obtained for these operators need to be interpreted with caution since these operators are evaluated on a restricted set of indicators.

Concerning the data of the contextual variables, in the case of the retail sample there were also problems with MV for a total of 31 operators (out of 200). To better understand how the MV could be treated in the contextual analysis, we tested if missing data is Missing Completely At Random (MCAR), using the Little's MCAR test (Little, 1988) available in the Statistical Package for the Social Sciences (SPSS) software. The results of the test demonstrated that the data are MCAR (Chi-Square = 15.656; DF= 12, p_value = 0.200). Therefore, the listwise deletion of cases was adopted since a sample of 169 operators (84.5% of the sample) is still acceptable for conducting the hypothesis tests. Tables 3.4 and 3.5 describe the contextual information regarding the wholesale and retail operators.

Table 3.4 – Summary statistics of quantitative contextual variables (ERSAR, 2019).

Quantitative variables	Wholesale (12 operators)				Retail (169 operators)			
	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max
z_1 – Wastewater collected (m^3/year)	45 546 232	53 349 106	4 644 092	198 158 332	3 352 240	6 870 932	31 392	63 531 216
z_2 – Energy production (MWh/year)	2 305 307	5 037 486	0	17 727 717	23 800	162 767	0	1 804 062
z_3 – Investment subsidies* (€/year)	----	----	----	----	160 959	370 327	0	3 152 686

*This variable was not considered in the case of wholesale operators due to the high percentage of missing values (83.3%).

Table 3.5 – Summary information of categorical contextual variables (ERSAR, 2019).

Categorical variables		no. of operators (%) Wholesale (12 in total)	Retail (169 in total)
z_4 – Operator's management system	Management by Concessionary Companies	10 (83.3%)	23 (13.6%)
	Management by Delegation	1 (8.3%)	20 (11.8%)
	Direct management by the State or Local Authorities	1 (8.3%)	126 (74.6%)
z_5 – Type of geographical area	Predominantly urban areas	3 (25%)	18 (10.7%)
	Medium urban areas	6 (50%)	46 (27.2%)
	Predominantly rural areas	3 (25%)	105 (62.1%)

3.4. Results and discussion

This section presents the results obtained in the performance assessment of wholesale and retail wastewater operators. It includes both the results of the composite indicator reflecting different perspectives in the evaluation of operators' activity, as well as the second-stage analysis to ascertain the influence of contextual factors on performance.

3.4.1 Composite indicator for operators' performance evaluation

The performance assessment model incorporates WR to allow the specification of bounds for the importance of different groups of indicators, as well as the importance of individual PI. In this study we consider three different perspectives of analysis, emphasizing three different groups of indicators corresponding to the dimensions of Adequacy of the Interaction with Users (AIU), Service Management Sustainability (SMS), and Environmental Sustainability (ES).

The lower bound attributed to the weight associated with the group of indicators emphasized in each perspective of analysis is $\delta_d = 50\%$ (see expression (3.2)). The lower bound attributed to the weight associated with each individual indicator is $w_{r_0} = w_{i_0} = 3.5\%$ (see expressions (3.3) and (3.4)).

It is noteworthy to mention that the lower bound imposed to each individual indicator (w_{r_0} and w_{i_0}) has a determinant effect on the value of the CI score obtained for each operator. The value of 3.5% considered in this study implies that, in total, the percentage of free weight to distribute among all indicators is 15% in the AIU perspective, and 18.5% in the SMS and ES perspectives.

Initially, the performance assessment model was run to evaluate the impact of outliers in the efficiency results. Following the procedure described in section 3.2.2, outlier operators that significantly affect the results of the efficiency model are removed from the sample. This procedure was conducted separately for wholesale and retail operators, in order to remove operators in case they fulfill all criteria specified in expression (10) at least in one perspective of analysis (AIU, SMS or ES). In the first step, three candidate outliers were identified in the sample of wholesale operators, and seven candidate outliers were identified among retail operators (i.e., these operators had values higher than the third quartile plus 1.5 times de interquartile range for at least one PI). In the second step, two wholesale operators (Águas da Serra and TRATAVE) and one retail operator (INOVA) were removed from the sample since they fulfilled all criteria for the removal of outliers, as specified in expression (10), in at least one of the perspectives of analysis (AIU in the case of Águas da Serra, SMS in the case of TRATAVE, and AIU, SMS and ES in the case of INOVA). Therefore, the final sample analyzed consisted of 199 retail operators and 10 wholesale operators, corresponding to 77% and 83% of the total Portuguese wastewater companies, respectively.

The results of the CI obtained for the wholesale operators are summarized in Table 3.6.

Table 3.6 – Results of the performance assessment of wholesale operators (10 operators, year 2018).

Perspective of analysis	Average CI	St.dev CI	No. (%) of efficient operators
Emphasis on Adequacy of the Interaction with Users	0.0272	0.0500	7 (70%)
Emphasis on Service Management Sustainability	0.0216	0.0522	7 (70%)
Emphasis on Environmental Sustainability	0.0252	0.0533	8 (80%)

These results show that wholesale operators demonstrate similar levels of performance in all perspectives considered. The relatively low average values of the CI score in all perspectives considered is indicative of high homogeneity among wholesale operators in terms of performance levels.

Table 3.7 shows the benchmark operators with a CI value equal to zero at least in one perspective of analysis. Note that a value of the CI equal to zero means that the operator is located on the frontier of the production possibility set when that perspective is adopted. Only one wholesale operator (Águas do Vale do Tejo) was considered inefficient in all perspectives of analysis (see Appendix 3B for more information about the CI values obtained by this operator). One possible reason for the relative inefficiency of this operator is that it started operating in the market very recently, after the sector's reorganization in the year of 2017. More specifically, in that reorganization, this operator was the result of the merger of several operators in the central interior region of Portugal, characterised by low density population, high infrastructure investments per capita, and other business constraints. In addition, in 2015, some of these operators had been previously merged with profitable operators from the littoral region. Thus, the services provided by Águas do Vale do Tejo were eventually impacted by this sector's reorganization, which might have prevented the stabilization of management processes.

Table 3.7 – Benchmark wholesale operators at least in one alternative perspective of analysis.

	CI Score		
	Emphasis on Adequacy of the Interaction with Users	*Emphasis on Service Management Sustainability*	*Emphasis on Environmental Sustainability*
Águas do Algarve	0	0	0
Águas do Centro Litoral	0	0	0
Águas do Norte	0	0	0
Águas do Tejo Atlântico	0	0	0
Águas Públicas do Alentejo	0	0.0094	0
SIMARSUL	0	0	0
SIMDOURO	0	0	0
Águas de Santo André	0.0317	0.0408	0
Associação de Municípios de Terras de Santa Maria	0.1214	0	0.1181

A visual representation of wholesale operator's performance regarding the three perspectives is provided in Figure 3.1. For example, there are six operators considered efficient in the three perspectives (Águas do Algarve, Águas do Centro Litoral, Águas do Norte, Águas do Tejo Atlântico, SIMARSUL, and SIMDOURO), clearly demonstrating best performance in all areas.

The regulator can signal these operators as lighthouses, representing references that can be followed by peers to gain further insights upon best practices leading to superior performance. It is also noteworthy mentioning that these six benchmark operators correspond to 60% of the wholesale sample. This high proportion of benchmarks may be due to the lack of discretionary power of the DDF model, as the number of operators assessed (10 in total) is relatively small in comparison with the number of indicators considered in the assessment (14 PIs). Although the weight restrictions imposed to our model allow overcoming to some extent the curse of dimensionalty of nonparametric models, the flexibility allowed in the choice of weights is intended to enable the assessment of operators under a light that favours their strengths (and hence the name "benefit-of-the doubt" assessment).

It is worth noting that the six operators that are signaled as benchmarks in all perspectives of analysis are all part of the same group (Águas de Portugal, AdP). This result is not surprising, given that they share similar managerial conditions (e.g., operation and maintenance procedures, procurement policies, centralized acquisitions, asset management policies, investment funding, clients' management procedures, among others). Note that the results for Associação de Municípios Terras de Santa Maria must be interpreted with caution, as this is an atypical operator, whose activity covers other functions beyond the water sector. Furthermore, the wastewater collection and treatment for some municipalities included in this association are provided by other operators.

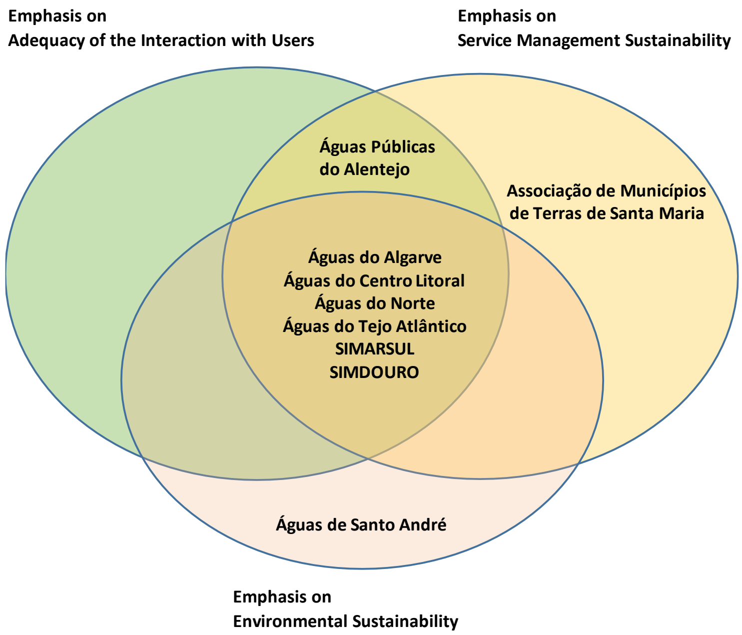


***Figure 3.1** – Overview of benchmark wholesale operators regarding the three perspectives of analysis.*

Concerning the retail operators, the performance assessment results are reported in Table 3.8.

Table 3.8 – Results of the performance assessment of retail operators (199 operators, year 2018).

Perspective of Analysis	Average CI	St.dev	No. (%) of efficient operators
Emphasis on Adequacy of the Interaction with Users	0.4010	0.2923	15 (7.5%)
Emphasis on Service Management Sustainability	0.3863	0.2542	18 (9.0%)
Emphasis on Environmental Sustainability	0.3592	0.2477	14 (7.0%)

Retail operators exhibit relatively large heterogeneity in terms of efficiency levels. In all perspectives considered, the frontier is spanned by less than 10% of the operators. The higher average value of the CI (representing the level of inefficiency) corresponds to the AIU dimension, which indicates that this dimension requires more attention regarding the implementation of actions to promote performance gains and convergence within the sector. Nonetheless, performance gains are also possible in the other dimensions, as the average CI values are still quite far from zero.

The benchmark retail operators, with a CI score equal to zero in at least one perspective of analysis, are listed in Table 3.9.

Table 3.9 – Benchmark retail operators for the alternative perspectives of analysis.

	CI score		
	Emphasis on Adequacy of the Interaction with Users	*Emphasis on Service Management Sustainability*	*Emphasis on Environmental Sustainability*
AGERE	0	0	0
Águas da Figueira	0	0	0
Águas de Coimbra	0.0289	0	0.0125
Águas de Gondomar	0	0	0
Águas de Valongo	0	0	0.0046
Águas do Porto	0.0193	0	0
CM de Alvaiázere	0	0.0720	0.0952
CM de Mealhada	0	0	0
CM de Miranda do Corvo	0	0	0
CM de Pombal	0	0	0
CM de Seia	0	0	0
CM de Vila de Rei	0	0	0
EMAS de Beja	0.1240	0	0.1423
INFRALOBO	0	0	0
INFRAQUINTA	0	0	0
SMAS de Almada	0	0	0
SMAS de Caldas da Rainha	0.0187	0	0.0034
SMAS de Peniche	0	0	0
SMAS de Sintra	0	0	0

Figure 3.2 provides a pictorial representation of the efficient operators listed in Table 9. We can observe that 13 operators (AGERE, Águas da Figueira, Águas de Gondomar, CM de

Mealhada, CM de Miranda do Corvo, CM de Pombal, CM de Seia, CM de Vila de Rei, INFRALOBO, INFRAQUINTA, SMAS de Almada, SMAS de Peniche, and SMAS de Sintra) show the best observed performance in all perspectives considered.

It is noteworthy mentioning that although the operators INFRALOBO and INFRAQUINTA are considered excellent based on the analysis of the 14 indicators provided to ERSAR, they are rather atypical operators, exploring luxury resorts in the south of Portugal. The benchmarking model does not acknowledge the exceptional positive conditions in which they operate, presumably running relatively modern infrastructures and operating in luxurious areas mainly constructed as holidays destinations, in Algarve. Hence, in those systems, the return of any investment is expected to be rapidly achieved. Therefore, the use of these operators for benchmarking purposes must be done with due caution.

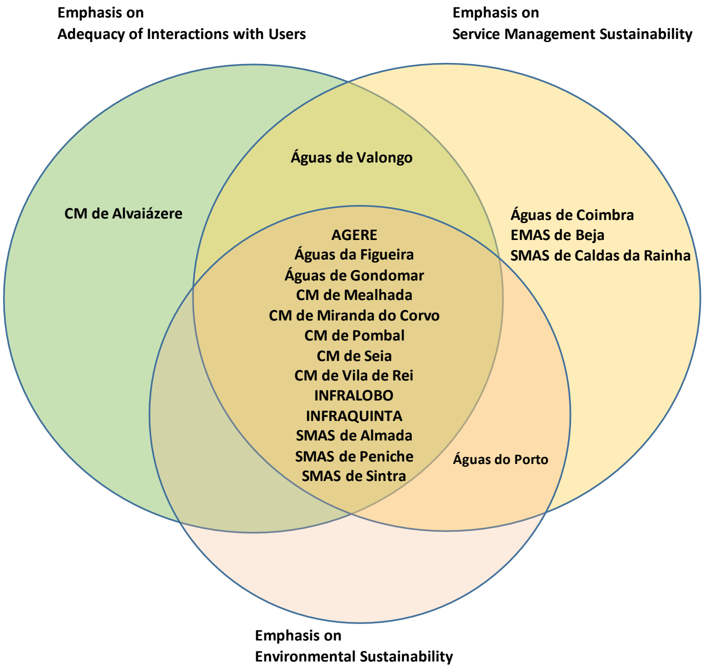


Figure 3.2 – *Overview of benchmark retail operators regarding the three perspectives of analysis.*

Given that the disclosure of performance results can trigger behavioral changes in the wastewater sector, we also included the list of inefficient retail operators in all perspectives of analysis in Appendix 3B (see Table 3B.1). The CI score reported for each perspective allows the operators to identify the dimensions of quality of service with greater score for performance improvement.

We also illustrate how the results of the approach proposed in this paper can be used by individual operators to guide improvements based on the observation of best practices among peer operators. Accordingly, Table 3.10 describes the peers identified for the operator "CM de Vidigueira" in the three perspectives of analysis. The CI score obtained for the AIU perspective is CI=0.3856, for the SMS perspective is CI=0.3440, and for the ES perspective is CI=0.2417.

Table 3.10 – Peers of CM de Vidigueira.

		Operator	Peers				
		CM de Vidigueira	AGERE	Águas da Figueira	Águas de Gondomar	CM de Mealhada	CM de Pombal
AIU	AR01	89	**99**	**93**	88	87	57
	AR02	0.27	0.26	**0.44**	**0.4**	0.18	0.2
	AR03*	0	2.11	0.64	0	4.44	1.67
	AR04	MV	100	100	99	88	94
SMS	AR05	69	**127**	**97**	**96**	**90**	**88**
	AR06	95.7	92.9	86.3	95.5	**96.5**	86
	AR07	0	**0.1**	0	**0.1**	**3.8**	**1.7**
	AR08*	0	0	1.7	0	0	0
	AR09	5.5	**12.4**	**10.9**	**8.7**	3.3	2.7
ES	AR10	MV	0.62	0.46	0.66	MV	0.63
	AR11	100	100	100	100	100	100
	AR12	MV	97	100	100	0	67
	AR13	18	**76**	**100**	**72**	**89**	**100**
	AR14	100	100	100	100	100	100
		AIU perspective	$\lambda = 0.639$	$\lambda = 0.361$	$\lambda = 0$	$\lambda = 0$	$\lambda = 0$
		SMS perspective	$\lambda = 0$	$\lambda = 0$	$\lambda = 0.168$	$\lambda = 0.832$	$\lambda = 0$
		ES perspective	$\lambda = 0$	$\lambda = 0$	$\lambda = 0$	$\lambda = 0$	$\lambda = 1$

* undesirable output; MV – missing value

From the analysis of Table 3.10, we conclude that the operator CM de Vidigueira should look for examples of practices in AGERE, Águas da Figueira, Águas de Gondomar, CM de Mealhada, and CM de Pombal. The importance of each of these peers in the different perspectives in signaled by the values of λ. For example, when considering the perspective of emphasis on the AIU dimension, the peers are AGERE ($\lambda = 0.639$) and Águas da Figueira ($\lambda = 0.361$).

In particular, the results show that CM de Vidigueira can improve its performance in 7 indicators (AR01, AR02, AR05, AR06, AR07, AR09, AR13), since there is at least one of the benchmarks identified with better values in these indicators. For example, to enhance the performance of indicator AR01, it should be looking at the practices implemented by AGERE and Águas da Figueira, whereas for AR02 it should observe the practices of Águas da Figueira e Águas de Gondomar. All cases corresponding to best practices observed in peers are signaled in bold in Table 3.10.

Despite the insights that can be attained from the peer analysis, it is essential to mention that there are non-manageable factors that may have influenced the values of the indicators reported by each operator. Therefore, all results here presented should be considered with due caution. Just to give an example, the indicator AR03 (floods occurrence) is related both to the weather conditions and to geographical features. Although Portugal is a small country, it is still characterized by different climatic conditions along some areas (e.g., coastal areas vs. interior areas) and different geographical features (e.g., mountain areas that are more typical in the North vs. plain areas that are more typical in the South). Therefore, some of the underperformance or good performance in this indicator may be attributed to these differences.

Nonetheless, it is still very difficult to assert the significance of those differences in explaining variations in the values of the indicator, since such detailed information for each operator is not available for quantitative consideration.

The analysis conducted so far was based on three perspectives, emphasizing each of the three dimensions of performance specified by the regulator (ERSAR). Given that the indicators framework reflects the activity of wastewater operators, we asked a stakeholder from this sector to express his opinion regarding the most relevant and reliable indicators among the 14 indicators collected by ERSAR. Hence, a total of four performance indicators (AR01, AR05, AR09, and AR13) were selected. Building from this, we explored a fourth perspective to emphasize these four indicators. This implied imposing an overall virtual weight of 50% for this group of 4 indicators, and a minimum virtual weight of 3.5% for all individual indicators.

Concerning wholesale operators, the results obtained were very aligned with the results previously obtained for the three perspectives considered. Accordingly, four operators were considered inefficient (Águas do Centro Litoral, Águas de Santo André, Águas do Vale do Tejo, and Associação de Municípios de Terras de Santa Maria). From these four operators, Águas do Vale do Tejo was also considered inefficient in all three perspectives of analysis previously considered, demonstrating that this operator has clear room for improvement in quality of service. Moreover, Águas do Centro Litoral, despite being considered efficient in the three perspectives of analysis (AIU, SMS and ES), reveals some potential for improvement in quality of service when the importance of indicators AR01, AR05, AR09, and AR13 is emphasized.

Under this perspective, the average value of the composite indicator was 0.0324, signaling more heterogeneous performance compared with the results obtained for the other perspectives. These results indicate that there is some scope for improvement regarding these four core indicators.

Concerning retail operators, the results of this alternative perspective revelead a total of 15 efficient operators (AGERE, Águas da Figueira, Águas de Valongo, Águas do Porto, CM de Mealhada, CM de Pombal, CM de Seia, CM de Vila de Rei, EMAS de Beja, INFRALOBO, INFRAQUINTA, SMAS de Almada, SMAS de Caldas da Rainha, SMAS de Peniche, and SMAS de Sintra). Comparing these results with those previously obtained, it is observed that Águas de Gondomar and CM de Miranda do Corvo dropped out from the subset of operators with best observed performance. Also, the average CI results are higher in this fourth perspective of analysis (0.4301) than for the other perspectives previously considered, leading to the same conclusions obtained for wholesale operators, signaling stability of the results and the existence of potential for improvement regarding the core indicators.

3.4.2 Contextual analysis

The results of the hypothesis tests conducted are displayed in Tables 3.11 and 3.12, for wholesale and retail operators, respectively.

Table 3.11 – Results of the contextual analysis for the wholesale operators in each perspective of analysis.

Contextual variable		Number of operators (10 in total)	Composite Indicator in each Perspective of analysis								
			"Emphasis on AIU"			"Emphasis on SMS"			"Emphasis on ES"		
			% Eff	Mean	Test *p*_value	% Eff	Mean	Test *p*_value	% Eff	Mean	Test *p*_value
Wastewater collected (m^3/year)											
G1:	<1stQ (13 986 848)	3	10%	0.0510		10%	0.0167		20%	.0394	
G2:	1stQ (13 986 848) – 2ndQ (35 478 615)	2	20%	0	0.278	20%	0	0.294	20%	0	0.451
G3:	2ndQ (35 478 615) – 3rdQ (54 150 038)	2	10%	0.0595		10%	0.082		10%	0.0671	
G4:	>3rdQ (54 150 038)	3	30%	0		30%	0		30%	0	
Energy production (kWh/year)											
G1	<Average (2 765 396)	8	50%	0.0340	0.335	50%	0.027	0.335	60%	0.0315	0.456
G2	>Average	2	20%	0		20%	0		20%	0	
Operators management system											
G1	Concession	8	60%	0.0188		60%	0.025		70%	0.0168	
G2	Delegation	1	10%	0	0.142	0%	0.009	0.491	10%	0	0.210
G3	Direct management	1	0%	0.121		10%	0		0%	0.1181	
Type of geographical area											
G1	Predominantly urban	3	30%	0		30%	0		30%	0	
G2	Medium urban	5	30%	0.0297	0.449	40%	0.033	0.125	30%	0.0504	0.329
G3	Predominantly rural	2	10%	0.0158		0%	0.025		20%	0	

The results displayed in Table 3.11 show that there is no statistical evidence that the wholesale operators belong to different populations within each contextual variable tested since the *p*-value of all hypothesis tests conducted was higher than 0.05. This result may be partially explained by the small number of operators comprising the sample. Nonetheless, this result also means that the dispersion observed in the values of the composite indicator cannot be explained with the contextual variables that were considered in the second stage assessment.

Table 3.12 – Results of the contextual analysis for the retail operators in each perspective of analysis.

Contextual variable		Number of operators (168 in total)	Composite Indicator in each Perspective of analysis								
			"Emphasis on AIU"			"Emphasis on SMS"			"Emphasis on ES"		
			% Eff	Mean	Test	% Eff	Mean	Test	% Eff	Mean	Test
Wastewater collected (m^3/year)											
G1	<1stQ (403 568)	44	1.2%	0.4860	**0.000**	0.6%	0.4750	**0.000**	0.6%	0.4276	**0.000**
G2	1stQ (403 568) – 2ndQ (1 110 100)	41	1.2%	0.4469		1.2%	0.4254		1.2%	0.3937	
G3	2ndQ (1 110 100) – 3rdQ (3 253 791)	41	1.8%	0.3752		2.4%	0.357		1.8%	0.3281	
G4	>3rdQ (3 253 791)	42	4.2%	0.2097		5.4%	0.212		4.2%	0.2062	
Energy production (kWh/year)											
G1	<Average (24 230)	163	7.7%	0.389	**0.010**	8.9%	0.3771	**0.012**	7.1%	0.3480	**0.008**
G2	>Average	5	0.6%	0.097		0..6%	0.0949		0.6%	0.0693	
Investment subsidies (€/year)											
G1	<Average (161 192)	131	6.0%	0.403	0.055	6.0%	0.3876	0.076	4.8%	0.3590	**0.048**
G2	>Average	37	2.4%	0.297		3.6%	0.3019		3.0%	0.2715	
Operators management system											
G1: Concession		23	1.8%	0.187	**0.000**	1.8%	0.1994	**0.000**	1.2%	0.2039	**0.000**
G2: Delegation		19	1.2%	0.1920		2.4%	0.2006		1.8%	0.2013	
G3: Direct management		126	5.4%	0.4444		5.4%	0.4249		4.8%	0.3854	
Type of geographical area											
G1	Predominantly urban	18	3.0%	0.1561		3.6%	0.1411		3.0%	0.1520	
G2	Medium urban	46	1.8%	0.297	**0.000**	3.0%	0.2808	**0.000**	1.8%	0.2671	**0.000**
G3	Predominantly rural	104	3.6%	0.455		3.0%	0.4468		3.0%	0.4043	

Conversely, in the case of the retail operators, the results displayed in Table 3.12 demonstrate that the hypothesis test conducted for each contextual variable is statistically significant with a 5% level of confidence for at least one perspective of analysis.

Accordingly, regarding the volume of wastewater collected, groups G2 and G4 were found to be significantly different in all perspectives of analysis, as well as groups G1 and G4. Given that the volume of wastewater collected represents the scale of operation, this result shows that in the case of retail operators larger sizes significantly favour better performances than smaller sizes. This result is in alignment with the results found in literature.

In terms of the energy production, the results demonstrate, in all perspectives of analysis, that this factor has a positive impact in retail operators' performance, which confirms the initial expectations.

The two groups of operators respecting the factor investment subsidies were considered significantly different only when emphasizing the ES dimension. Therefore, this factor can only help explaining the dispersion of results observed in that perspective of analysis. As expected, higher investment subsidies favor a better quality of service.

In terms of the operator's management system, it was found in all perspectives of analysis that direct management is significantly worse than delegation and significantly worse than concessions. This result is aligned with the vision of the sector within the Portuguese market structure. However, there is no statistical evidence supporting that concessions are better than delegations.

In what concerns the type of geographical area, it was found that predominantly urban areas significantly favor better performances compared to predominantly rural areas, and that medium urban areas also significantly favor better performances compared to predominantly rural areas. This result is aligned with literature. However, there is no statistical evidence supporting that predominantly urban areas are better than medium urban areas.

In face of these results, it is interesting to analyze whether the retail operators that were assigned a zero value for the composite indicator in the three perspectives of analysis have favorable or unfavorable contextual conditions. The aim is to help both the regulator and the wastewater operators to better understand the results obtained and how they can be used in practice. As such, the position of each efficient operator is compared with a reference value for each variable that is signaled to have a favorable impact on performance in the contextual analysis previously conducted (see Table 3.13). The symbol √ reflects that the value of the variable is favorable.

Table 3.13 – Contextual variables of the retail operators that are on the efficient frontier in the three perspectives of analysis.

	Wastewater collected (m^3/year)		Energy production (kWh/year)		Investment subsidies (€/year)		Operator management system		Type of geographical area		**Total √**
AGERE	13 767 587	√	14 362		764 448	√	Delegation	√	Pred. Urban	√	**4**
Águas da Figueira	3 591 895	√	8 584		0		Concession	√	Medium Urban	√	**3**
Águas de Gondomar	9 097 655	√	0		154 546		Concession	√	Pred. Urban	√	**3**
CM de Mealhada	1 093 037		0		0		Direct management		Pred. Rural		**0**
CM de Miranda do Corvo	846 589		0		50 980		Direct management		Pred. Rural		**0**
CM de Pombal	3 325 066	√	0		957 529	√	Direct management		Pred. Rural		**2**
CM de Seia	2 192 160		0		111 793		Direct management		Pred. Rural		**0**
CM de Vila de Rei	106 936		0		4 975		Direct management		Pred. Rural		**0**
INFRALOBO	278 418		MV		23 284		Delegation		Medium Urban	√	**2**
INFRAQUINTA	1 131 087		0		0		Delegation	√	Medium Urban	√	**2**
SMAS de Almada	13 313 145	√	1 804 062	√	549 070	√	Direct management		Pred. Urban	√	**4**
SMAS de Peniche	3 106 057		0		130 954		Direct management		Medium Urban	√	**1**
SMAS de Sintra	31 241 628	√	17 517		336 912	√	Direct management		Pred. Urban	√	**3**

It can be observed that AGERE and SMAS de Almada have an extremely favorable context of operation that can help explain their excellent performance. Conversely, CM de Mealhada, CM de Miranda do Corvo, CM de Seia, and CM de Vila de Rei face a considerably unfavorable situation regarding the contextual conditions. Thus, in the year of 2018, these operators demonstrate an acknowledgeable effort in providing an exemplar quality of service to their users.

3.5. Conclusions

A well-functioning water sector is a fundamental element for sustainable development. Water services are a basic requirement for human wellbeing and a driver of the economy. Nonetheless, the prospect of a worldwide water crisis is a core concern to administrative authorities. As asserted by OECD, the water crisis is rooted in a governance crisis.

The study developed in this paper constitutes an attempt to contribute for improvements in water governance, through the empowerment of water sector regulators. This empowerment stems from the simplicity of the information that can be achieved with the approach proposed in this paper. A simple and summarized information allows for more objectivity to be attained.

The key baseline of the framework proposed is a comprehensive benchmarking approach that takes full advantage of the use of individual performance indicators to estimate a composite indicator reflecting the performance of each operator. Additionally, the computation of the composite indicator is flexible in terms of the importance that can be given to different dimensions of performance. This allows accounting for several perspectives concerning the evaluation of the quality of service. The imposition of weight restrictions in the benchmarking model, based on a Directional Distance Function, is crucial to achieving this purpose.

The CI models imposing a minimum weight for each indicator as well as to groups of indicators belonging to the dimensions specified by the regulator enables obtaining a detailed and clear picture of the performance of each operator by comparison with its peers. This benchmarking exercise allows an objective evaluation of the gap in quality of service among operators, facilitating the identification of the efforts required for improvement in this sector.

The methodology proposed particularly enhance sunshine regulation by contributing with a holistic view of performance at the operator level. Nonetheless, other regulatory settings can also benefit from its application. This methodology facilitates the identification of benchmark operators according to different perspectives that can be adopted to highlight different dimensions of the quality of service provided.

The framework developed in this paper also integrates a second-stage contextual analysis to explore the differences in performance that can be attributed to the contextual factors faced by the operators. This constitutes valuable information to regulators that can complement the knowledge obtained through the composite indicator. The overall information gathered constitutes an overview of performance of the regulated wastewater operators, which yields simplicity and transparency into the water sector in each evaluation period.

The formative evaluation framework developed in this paper was tested in the Portuguese water sector regulatory context. We carried out an analysis based on the information collected by the regulator on individual performance indicators. As such, the main objective of this benchmarking framework was to complement the information that is currently available to the regulator.

The results show that in the case of wholesale operators, six operators (Águas do Algarve, Águas do Centro Litoral, Águas do Norte, Águas do Tejo Atlântico, SIMARSUL, and SIMDOURO) were found to be excellent in the three perspectives of analysis considered, based on ERSAR dimensions of AIU, SMS and ES. In terms of the contextual analysis, given the small number of operators (only twelve), none of the variables considered was found to be significant to explain any dispersion observed in the values of the composite indicator.

In the case of retail operators, thirteen operators (AGERE, Águas da Figueira, Águas de Gondomar, CM de Mealhada, CM de Miranda do Corvo, CM de Pombal, CM de Seia, CM de Vila de Rei, INFRALOBO, INFRAQUINTA, SMAS de Almada, SMAS de Peniche, and SMAS de Sintra) were considered excellent in the three perspectives of analysis aligned with ERSAR dimensions. Regarding the contextual analysis, it was found that larger sizes, larger investment subsidies, higher energy production, concessions and predominantly urban areas positively impact the quality of service delivered to customers. This information can help to better understand part of the dispersion obtained in the values of the composite indicator.

In terms of policy implications, we highlight that no extrapolations can be done regarding the ranking of the operators in other periods of time, as the analysis was only based on data for the year 2018. However, the approach proposed in this paper can be very useful to complement sunshine regulation and promote continuous improvement in the wastewater sector. The disclosure of benchmarking results done at operators' level is very important to informing public opinion and promote continuous improvement in the sector. The analysis of a large set of performance indicators considered individually does not allow grasping essential information regarding the potential for improvement that may exist in complex systems, such that the systematic use of innovative frameworks such as the one presented in this paper facilitates the extraction of relevant information for the management of utilities and for the design of effective regulatory policies.

The representativeness of some indicators currently in use by ERSAR in terms of the real performance of operators is one of the potential topics that a holistic benchmarking exercise, as the one proposed in this paper, can trigger further discussion. Unexpected results can raise awareness of the need to eventually redefine the indicators, to ensure their correct specification and the reliability of the data collected and made available to the regulator. Also, some indicators can benefit from the aggregation of data for several years (i.e., using the average for a range of years considered representative of the operators' activity), instead of considering information referring to a single year. For example, the indicator AR07, which refers to sewers' rehabilitation, expressed in percentage per year, may lead to a biased evaluation of the operators in a given year, since some of the operators may have invested in sewers rehabilitation in previous years or plan to start a substantial investment program in the following years.

Furthermore, the application and release of results of a comprehensive benchmarking exercise as the one illustrated in this paper, can contribute to increase transparency and engagement within the wastewater sector. This can be achieved due to both the public disclosure of the results obtained and the discussion that can be promoted between the regulator and stakeholders of the sector to balance the importance given to particular groups of indicators.

The scope of application of the framework proposed in this paper is vast, as it can be easily adapted to other sectors and geographical contexts worldwide.

Acknowledgements

The authors wish to acknowledge the financial support of Project "TEC4Growth - Pervasive Intelligence, Enhancers and Proofs of Concept with Industrial Impact/NORTE-01-0145-FEDER-000020." This project is financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF). The first author also wishes to acknowledge the financial support from FCT - Fundação para a Ciência e a Tecnologia (Portuguese national funding agency for science, research and technology), given through the Grant PD/BD/142815/2018 respecting the Doctoral Program on Sustainable Energy Systems by MIT-Portugal. The authors also gratefully acknowledge the computational assistance provided by Behrouz Arabi in the initial modeling phases of this work.

Appendix 3A

Overview of performance assessment and benchmarking in the water sector

Performance assessment and benchmarking are recognized worldwide as positively impacting the functioning and sustainability of the water sector. In particular, they may lead to enhanced knowledge upon the sector that can support the design of managerial policies and foster changes.

In a review of international studies on performance measurement and benchmarking in the water sector, Vilanova et al. (2015) reinforced the need to integrate data analysis techniques as part of any performance assessment framework. Nudurupati et al. (2011) also noted that any performance measurement system should include the integration, accessibility, and visibility of information within a dynamic system, to support quick decision making and promote proactive management.

According to Marques & De Witte (2010) and Marques & Pinto (2018), the assessment of performance can be conducted at different levels, from upstream to downstream in the system. These levels can also be seen hierarchically, from a more general level, focusing on regulation, to a more specific level, focusing on individual activities (see the schematic representation in Figure 3A.1).

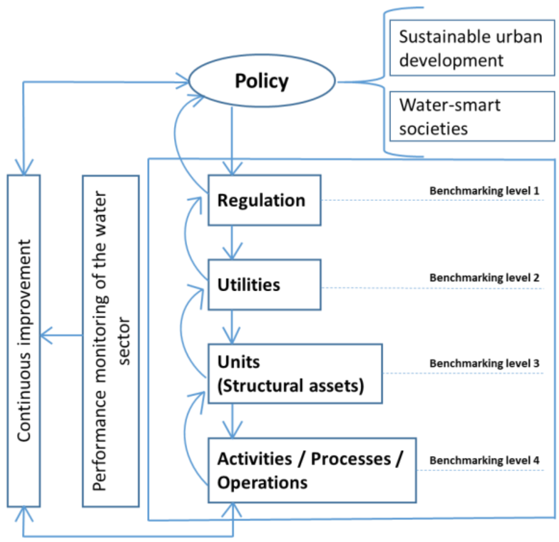


***Figure 3A.1** – Benchmarking levels in the structural organization of the water sector system. Adapted from Marques & De Witte (2010, p. 43).*

In the pictorial representation of Figure 3A.1, the policy level is indicated outside the scope of the benchmarking application fields since it can be seen as either the mentor of changes to be triggered by the benchmarking exercise, or the actor whose practices are changed according to the results of the benchmarking process. Accordingly, it corresponds to the governance level, under the responsibility of national authorities. Considering the benchmarking box in the same schematic representation, the top benchmarking levels (Regulation and Utilities) correspond to the regulatory activity upon the sector's utilities. It involves the assessment of utilities' efficiency, productivity change and quality of service provided to users. Water utilities must follow government policies and regulatory guidelines in their development of water-related activities. The lower levels (Structural assets / activities / processes /operations) correspond to the assets managed by each utility that should allow the effective conduction of processes at an operational level.

Each of these levels should have associated a specific performance assessment system to foster performance improvements within the sector. The actions and results obtained in each level have implications in the remaining levels and consequently, in the sector as a global system (Marques & De Witte, 2010).

Concerning the lower levels of the water sector (Structural assets / Activities / Processes / Operations), research efforts on performance assessment and benchmarking attempt to tackle the sustainability of the treatment processes, the increase of effluent quality, and the reduction of resource consumption and operational costs. The aim of the majority of studies is the quantification of potential improvements in technical efficiency and productivity (e.g. Gómez et al., 2017; Molinos-Senante et al., 2014c; Hernández-Sancho and Sala-Garrido, 2009; Sala-Garrido et al., 2012; Hernández-Sancho et al. 2011b; Molinos-Senante el al., 2015a). These studies used yearly data. The study of Torregrossa et al. (2016) focused on daily benchmarking and presented a methodology to deal with uncertainty using data collected on a daily basis. Some studies also dedicate special attention to environmental issues (e.g., D'Inverno et al., 2018; Hernández-Sancho et al., 2011a; Castellet and Molinos-Senante, 2016; Dong et al., 2017; Molinos-Senante et al., 2016a; Molinos-Senante et al., 2014b; Lorenzo-Toja et al., 2016; Molinos-Senante and Guzmán, 2018).

In the lower levels of the water sector, and from a methodological perspective, the majority of studies reviewed used Data Envelopment Analysis (DEA). A few studies combined DEA with other decision support techniques. For example, D'Inverno et al. (2018) uses Analytical Hierarchical Process (AHP) combined with Non-Radial Direction Distance Functions (NRDDF). Molinos-Senante et al. (2014a) used the AHP to reflect the preferences of experts in the assignment of the weights to each sustainability dimension and indicators considered in the performance assessment of wastewater treatment systems. This study involved the computation of a composite indicator for each sustainability dimension, as well as the construction of a composite indicator to reflect the global sustainability of each type of wastewater treatment system analyzed (i.e., constructed wetlands, pond systems, extended aeration, membrane bioreactor, rotating biological contactor, trickling filter and sequencing batch reactor). Other methodological approaches, such as index numbers (e.g., the Malmquist index or Luemberger indicators) and parametric techniques (e.g., cost functions estimated using Stochastic Frontier Analysis) have also been explored in literature. See the study by Guerrini et al., 2016, for a literature review. In contrast with most benchmarking studies that analyse the entire productive process carried out at each plant, the study of Torregrossa et al. (2016) focused on the benchmarking of individual Key Performance Indicators (KPIs) on a daily basis, focusing on specific parameters related to energy consumption and pollutant load.

Focusing on the middle level of the hierarchical organization of the water sector, Abbot and Cohen (2009) review the studies that concern the measurement of productivity and efficiency in the water sector (including water and sewerage companies). It comprises studies that range from 1969 till 2007 and relate to more than 20 different geographical areas, including countries, states and regions in the five continents. This study emphasizes the input and output variables, as well as the methodologies used. Cost functions, Stochastic Frontier Techniques and Data Envelopment Analysis are among the methods more frequently used.

Still in the middle level of the water sector, Dong et al. (2018) used a comprehensive benchmarking approach to assess the sustainability of the water sector by focusing on water infrastructure systems rather than individual companies or individual assets. These systems include the main elements of the urban water system in a city (i.e., water treatment plants, water distribution networks, drainage networks, and wastewater treatment plants). The study applied DEA to a sample of 157 Chinese water infrastructure systems.

Furthermore, Le Lannier & Porcher (2014), Molinos-Senante et al. (2015b) and Romano et al. (2017) assessed the performance of water utilities using an input-output framework to describe utilities' operation. Quality of service was also a central aspect considered in the evaluation of efficiency. Le Lannier & Porcher (2014) also considered a quality of service indicator as an output of the DEA model used to assess the performance of French water utilities. The quality variable reflects both environmental performance and network quality to analyze whether the water companies comply with quality standards. Molinos-Senante et al. (2015b) use the lack of service quality as an undesirable output of the Data Envelopment Analysis model used to evaluate the efficiency of water and sewerage companies from Chile. In this study, the variables used to translate the lack of service quality were the water and sewerage quality and the users' perception of the quality of the service provided. The importance of integrating quality of service in regulatory benchmarking was also noted by Giannakis et al. (2005). Romano et al. (2017) used DEA to analyze the effect of ownership (public ownership versus public-private partnerships) on the efficiency of the water utilities from two Italian regions. The results of these studies highlighted the importance of including quality of service in the assessments to ensure equity in the utilities' evaluation. The economic objectives of water utilities (i.e., cost minimization and profit maximization) may disregard the quality of service dimension, so a comprehensive evaluation should include all dimensions of activity simultaneously.

Focusing on the middle level of the water sector, there is often heterogeneity of environmental conditions faced by water utilities, which may have a considerable impact on the production process (Carvalho & Marques, 2011). Therefore, several studies complemented the efficiency assessment with the analysis of the effect of environmental conditions on performance. A research design involving a two-stage procedure combining DEA with regression allows obtaining a more comprehensive perspective of utilities' performance. Le Lannier & Porcher (2014) and Molinos-Senante et al. (2015b) are examples of studies resorting to multi-stage procedures of this type. In particular, Le Lannier & Porcher (2014), used a three-stage DEA approach to compute performance scores separately for public and private owned French utilities. Molinos-Senante et al. (2015b) used nonparametric statistical tests (Mann-Whitney U and Kruskal-Wallis) in a second stage analysis, to test the impact of the contextual factors on the DEA efficiency scores of water and sewerage companies from Chile. Another trend of studies includes the contextual factors directly in the efficiency assessment by using partial frontier methods based on the probabilistic approach developed by Cazals et al. (2002). One example is the study by Marques et al. (2014) that evaluate the performance of Japanese water utilities.

For the top level of the hierarchy of the water system, as depicted in Figure 3A.1, Marques & Pinto (2018) focused on the performance assessment of regulatory governance. In this case, the regulatory activity itself, as conducted by different regulators, is analyzed. The authors developed a scorecard approach using MCDA to assign a category to each regulator according to its level of performance in comparison with a reference profile. This study is aligned with the proposal of Berg (2016), who argue that the performance of regulatory governance cannot be completely coupled to the efficiency and effectiveness of the regulatory process, which is correlated with the sector's performance. Focusing on the impact that regulatory activity has on the performance of the water sector, the study Marques et al. (2011) discussed the example of three countries (Portugal, Australia and UK) where the use of benchmarking in sunshine regulation for several years has proven to contribute to good results in the sector. In what concerns quality of service regulation, the use of performance indicators is widespread. The authors also noted that despite the differences that exist among regulatory approaches worldwide, the use of benchmarking in the context of quality of service regulation and tariff setting is growing. Also, the study by Abbott & Cohen (2009) explored the role of regulation on productivity and efficiency levels of the US and UK water sectors. It concluded that regulation has a positive impact on the performance levels observed in these two regulatory settings.

Marques & De Witte (2010) also highlighted the role of regulatory incentives to enhance the efficiency levels in the water sector.

On the top level of the water sector, quality of service is often monitored by regulatory authorities to ensure the protection of users interests under a sunshine regulation model. In this context, Pinto et al. (2017) developed an innovative approach to provide a holistic view of the quality of service provided by water utilities. The approach consists of the construction of a quality of service index (QSI), which aggregates all the information embedded in a panoply of performance indicators. The QSI is obtained recurring to MCDA, involving the active participation of decision-makers. The method enables assigning each utility to categories representing the level of quality of service provided to customers. This method is illustrated using data on water supply services publicly disclosed by the Portuguese water and waste sector regulatory authority (ERSAR).

Combining all the information obtained from this review of studies on performance assessment and benchmarking in the water sector, the most relevant outcomes can be summarized as follows. There are studies on performance measurement at all hierarchical levels of the water sector. At the lower level of the water sector system, the focus of benchmarking studies is mainly the search for improvements in technical and economic efficiency, enabling resource or cost reductions. There is also a growing concern regarding the minimization of the negative impacts on the environment of treatment technologies associated with plants activity. At the middle level of the water sector system, there is a growing concern with the integration of quality of service indicators in the efficiency evaluation of water utilities. Sustainability issues and the analysis of the influence of environmental factors on performance are also topics that deserved considerable attention in the literature. DEA is the technique most frequently used in the analyzed studies at the lower and middle layers of the water system. In the governance and regulatory levels, the focus is on the provision of an overview of performance levels, which involves the estimation of aggregate measures of performance. The methodology most often applied in these studies is MCDA. However, the studies analyzed only represent a small part of the literature upon benchmarking within the water sector. For a meta-analysis review of studies in the water sector, including the overview of the quantitative techniques applied for benchmarking purposes, the study by Berg and Marques (2011) is recommended.

Appendix 3B

Table 3B.1 – Performance results for inefficient wholesale and retail operators in the three alternative perspectives of analysis considered.

	Emphasis on Adequacy of the Interaction with Users	*Emphasis on Service Management Sustainability*	*Emphasis on Environmental Sustainability*
Wholesale operators			
Águas do Vale do Tejo	0.1190	0.1657	0.1341
Retail operators			
Abrantaqua	0.1300	0.0931	0.0962
Águas da Azambuja	0.4230	0.4675	0.5090
Águas da Covilhã	0.1391	0.3273	0.3197
Águas da Região de Aveiro	0.1289	0.1372	0.1178
Águas da Teja	0.4549	0.6523	0.5433
Águas de Alenquer	0.2354	0.3828	0.3810
Águas de Barcelos	0.2061	0.2564	0.1614
Águas de Carrazeda	0.2318	0.2377	0.1773
Águas de Cascais	0.0718	0.0525	0.0348
Águas de Gaia	0.2451	0.2702	0.3197
Águas de Mafra	0.1672	0.1238	0.1306
Águas de Paços de Ferreira	0.3119	0.3238	0.2091
Águas de Paredes	0.0635	0.1009	0.0497
Águas de S. João	0.1134	0.0708	0.0531
Águas de Santarém	0.1785	0.1100	0.0836
Águas do Marco	0.3103	0.3767	0.2825
Águas do Norte (Parceria Estado/municípios)	0.1530	0.1971	0.1374
Águas do Ribatejo	0.0320	0.0679	0.0331
Águas do Sado	0.0208	0.0014	0.0048
AMBIOLHÃO	0.0790	0.0596	0.0401
Aquaelvas	0.2684	0.2788	0.3961
Aquafundalia	0.1950	0.1139	0.1586
Aquamaior	0.2702	0.1585	0.3226
CARTÁGUA	0.0476	0.0324	0.0245
CM de Alandroal	0.4654	0.5526	0.5541
CM de Albufeira	0.1344	0.1057	0.0701
CM de Alcácer do Sal	0.5340	0.6362	0.5877
CM de Alcanena	0.3229	0.2757	0.1783
CM de Alcochete	0.4414	0.3746	0.4245
CM de Alcoutim	0.3519	0.5076	0.4452
CM de Alfândega da Fé	0.4727	0.5275	0.5338
CM de Aljustrel	0.5360	0.5683	0.5814
CM de Almeida	0.6085	0.7247	0.6822
CM de Almodôvar	0.3770	0.3024	0.2008
CM de Amares	0.9515	0.7037	0.7501
CM de Ansião	0.9158	0.7859	0.7959
CM de Arcos de Valdevez	0.2913	0.4681	0.3612
CM de Arganil	0.5153	0.3075	0.2098
CM de Arronches	0.7881	0.7509	0.7779

CM de Arruda dos Vinhos	0.4718	0.5441	0.5875
CM de Avis	0.8055	0.8038	0.8435
CM de Barrancos	0.1359	0.0052	0.1354
CM de Barreiro	0.1929	0.1603	0.1989
CM de Batalha	0.5272	0.7327	0.6423
CM de Bombarral	0.7413	0.6428	0.7060
CM de Borba	0.7000	0.7092	0.7652
CM de Boticas	0.5151	0.4967	0.5091
CM de Bragança	0.2513	0.2177	0.1728
CM de Cabeceiras de Basto	0.5174	0.4122	0.5417
CM de Cadaval	0.4403	0.5817	0.5248
CM de Caminha	0.7298	0.6975	0.6652
CM de Carregal do Sal	0.4797	0.3663	0.2651
CM de Castro Daire	0.4920	0.8523	0.5833
CM de Castro Marim	0.6843	0.6642	0.6642
CM de Celorico da Beira	0.5829	0.3811	0.4896
CM de Chaves	0.5931	0.6708	0.5631
CM de Condeixa-a-Nova	0.2911	0.3862	0.2418
CM de Constância	0.5758	0.5755	0.6297
CM de Cuba	0.4708	0.4648	0.4724
CM de Entroncamento	0.4644	0.5166	0.5536
CM de Espinho	0.6169	0.5623	0.6097
CM de Estremoz	0.3236	0.2248	0.1721
CM de Felgueiras	1.8716	0.9465	0.9313
CM de Ferreira do Zêzere	0.4367	0.5894	0.5409
CM de Fornos de Algodres	0.3754	0.4955	0.2641
CM de Gavião	0.5744	0.4911	0.5192
CM de Góis	0.0862	0.1524	0.0734
CM de Gouveia	0.6699	0.7229	0.6319
CM de Grândola	0.4899	0.3714	0.2847
CM de Guarda	0.4780	0.4655	0.4766
CM de Lagoa	0.1686	0.0520	0.0943
CM de Lagos	0.0940	0.0583	0.0341
CM de Lamego	0.7234	0.8245	0.7072
CM de Lisboa	0.5309	0.5162	0.5840
CM de Loulé	0.7026	0.6299	0.6881
CM de Lourinhã	0.5215	0.5181	0.5841
CM de Lousã	0.4692	0.4846	0.4790
CM de Lousada	0.3555	0.4012	0.2525
CM de Mação	0.8219	0.6350	0.7341
CM de Macedo de Cavaleiros	0.3169	0.3537	0.4152
CM de Mangualde	0.4235	0.3825	0.2401
CM de Marinha Grande	0.5670	0.5906	0.5666
CM de Marvão	0.6210	0.4636	0.3445
CM de Melgaço	0.0753	0.1837	0.1441
CM de Mértola	0.7566	0.5566	0.5267
CM de Mira	1.1768	0.8273	0.7987
CM de Miranda do Douro	0.6265	0.6195	0.5341

CM de Mirandela	0.3428	0.3066	0.2278
CM de Mogadouro	0.2857	0.3278	0.2102
CM de Moimenta da Beira	0.5238	0.5839	0.6336
CM de Moita	0.2264	0.2037	0.1312
CM de Monção	0.5493	0.7371	0.5785
CM de Mondim de Basto	0.4348	0.4149	0.5146
CM de Montemor-o-Novo	0.3771	0.4275	0.5055
CM de Montemor-o-Velho	0.2207	0.2677	0.1433
CM de Mora	0.6507	0.6983	0.6660
CM de Moura	0.5630	0.5660	0.5720
CM de Nelas	0.5660	0.6000	0.5598
CM de Nisa	0.4380	0.4328	0.3624
CM de Odemira	0.3007	0.2580	0.1995
CM de Oleiros	0.3074	0.3100	0.2321
CM de Oliveira do Hospital	0.2644	0.3206	0.1984
CM de Ourém	0.6816	0.4508	0.2893
CM de Palmela	0.1377	0.1347	0.0612
CM de Paredes de Coura	0.5172	0.4158	0.2824
CM de Pedrógão Grande	1.4365	0.9860	0.9302
CM de Penacova	0.3469	0.3785	0.4157
CM de Penedono	0.4218	0.3747	0.2656
CM de Peso da Régua	0.6200	0.6896	0.5888
CM de Pinhel	0.3028	0.4263	0.2479
CM de Ponte da Barca	0.5375	0.4205	0.2554
CM de Ponte de Lima	0.4682	0.4911	0.4967
CM de Ponte de Sor	0.3838	0.3922	0.3653
CM de Póvoa de Lanhoso	0.3892	0.5350	0.4727
CM de Póvoa de Varzim	0.4607	0.4130	0.4631
CM de Proença-a-Nova	0.5153	0.6443	0.6205
CM de Redondo	0.6361	0.6450	0.6320
CM de Reguengos de Monsaraz	0.4775	0.5076	0.5701
CM de Resende	1.0323	1.1778	0.8972
CM de Ribeira de Pena	0.5228	0.2882	0.2237
CM de Rio Maior	0.9181	0.7733	0.7361
CM de Sabrosa	0.4950	0.5631	0.5933
CM de Sabugal	0.4905	0.4678	0.3282
CM de Santa Comba Dão	0.5074	0.3981	0.2956
CM de Santiago do Cacém	0.4088	0.3197	0.2043
CM de São João da Pesqueira	0.3832	0.3675	0.2424
CM de São Pedro do Sul	0.3868	0.1156	0.0803
CM de Sardoal	0.7907	0.6686	0.6552
CM de Sátão	0.2266	0.2739	0.1733
CM de Seixal	0.4750	0.4679	0.5108
CM de Sernancelhe	0.3629	0.3788	0.3764
CM de Serpa	0.3114	0.3185	0.2333
CM de Sertã	1.1272	0.6732	0.7234
CM de Sesimbra	0.4322	0.4762	0.4696
CM de Silves	0.3255	0.2488	0.1373

CM de Sines	0.2646	0.2476	0.2554
CM de Sobral de Monte Agraço	0.5761	0.5969	0.6435
CM de Soure	0.2329	0.2270	0.1593
CM de Tábua	0.0624	0.0877	0.0492
CM de Terras de Bouro	1.3374	0.8765	0.8809
CM de Tondela	0.2902	0.4064	0.2217
CM de Torre de Moncorvo	0.4815	0.5096	0.5669
CM de Vale de Cambra	0.9205	0.8054	0.6908
CM de Vidigueira	0.3856	0.3440	0.2417
CM de Vila Nova de Cerveira	0.2475	0.2122	0.0957
CM de Vila Nova de Famalicão	0.0521	0.0856	0.0644
CM de Vila Nova de Foz Coa	0.4494	0.4478	0.3148
CM de Vila Nova de Paiva	0.6329	0.7886	0.6725
CM de Vila Pouca de Aguiar	0.4649	0.6214	0.5560
CM de Vila Velha de Ródão	0.6020	0.6781	0.6125
CM de Vila Verde	0.5977	0.6345	0.6155
CM de Vimioso	0.2639	0.2389	0.1836
CM de Vinhais	0.4545	0.4705	0.2776
CM de Vouzela	1.5192	1.2274	0.8308
EMAR de Portimão	0.3810	0.3329	0.4364
EMAR de Vila Real	0.4455	0.4378	0.4423
Esposende Ambiente	0.6279	0.6871	0.6261
FAGAR - Faro	0.4141	0.4963	0.4924
Indaqua Feira	0.0942	0.0488	0.0639
Indaqua Matosinhos	0.0313	0.0280	0.0190
Indaqua Oliveira de Azeméis	0.4189	0.4367	0.6965
Indaqua Vila do Conde	0.3494	0.4202	0.4232
INFRAMOURA	0.2849	0.2700	0.3431
INFRATRÓIA	0.4902	0.5048	0.5656
Penafiel Verde	0.1943	0.2059	0.1561
SIMAR de Loures e Odivelas	0.3226	0.3178	0.3703
SIMAS de Oeiras e Amadora	0.3023	0.2211	0.3268
SM de Alcobaça	0.6134	0.5658	0.6412
SM de Castelo Branco	0.3591	0.4390	0.4293
SMAS de Leiria	0.5818	0.5567	0.5948
SMAS de Montijo	0.3021	0.3262	0.3658
SMAS de Tomar	0.3828	0.3478	0.3868
SMAS de Torres Vedras	0.2260	0.3017	0.2316
SMAS de Vila Franca de Xira	0.1571	0.0694	0.1706
SMAS de Viseu	0.2666	0.3106	0.2002
SMAT de Portalegre	0.4027	0.3352	0.2360
SMEAS de Maia	0.1687	0.1346	0.1224
SMSB de Viana do Castelo	0.1050	0.0923	0.0567
Taviraverde	0.0910	0.0514	0.0665
VIMÁGUA	0.4745	0.5422	0.5254

References

Abbott, M. & Cohen, B. (2009). Productivity and efficiency in the water industry. *Utilities Policy*, 17, 233-244.

Akhmouch, A., Correia, F.N. (2016). The 12 OECD principles on water governance – When science meets policy. *Utilities Policy*, 43, 14-20.

Aragon, Y., Daouia, A., Thomas-Agnan, C. (2005). Nonparametric frontier estimation: a conditional quantile-based approach. *Econometric Theory*, 21(2), 358-389.

Badin, L., Daraio, C., Simar, L. (2014). Explaining inefficiency in nonparametric production models: the state of the art. *Annals of Operation Research*, 214, 5-30.

Baptista, J.M. (2014). Uma abordagem regulatória integrada (ARIT-ERSAR) para os serviços de águas e resíduos (An integrated regulatory approach (ARIT-ERSAR) for the water and waste services). *Entidade Reguladora dos Serviços de Águas e Resíduos (ERSAR)*. Lisboa, setembro de 2014. http://www.ersar.pt/pt/publicacoes/estudos-ersar

Berg, S.V. (2007). Conflict resolution: benchmarking water utility performance. *Public Administration and Development*, 27, 1-11.

Berg, S.V. (2016). Seven elements affecting governance and performance in the water sector. *Utilities Policy*, 43, 4-13.

Berg, S. V., Marques, R. C. (2011). Quantitative studies of water and sanitation utilities: a benchmarking literature survey. *Water Policy*, 13(5), 591-606.

Calabria, F.A., Camanho, A.S., Zanella, A. (2018). The use of composite indicators to evaluate the performance of Brazilian hydropower plants. *International Transactions in Operations Research*, 25(4), 1323-1343.

Carvalho, P., Marques, R.C. (2011). The influence of the operational environment on the efficiency of water utilities. *Journal of Environmental Management*, 92, 2698-2707.

Castellet, L. and Molinos-Senante, M. (2016). Efficiency assessment of wastewater treatment plants: A data envelopment analysis approach integrating technical, economic and environmental issues. *Journal of Environmental Management*, 167, 160-166.

Cazals, C., Florens, J., Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106, 1-25.

Cherchye, L., Moesen, W., Rodgge, N., Puyenbroeck, T.V. (2007). An Introduction to 'Benefit of the Doubt' Composite Indicators. *Social Indicators Research*, 82, 111-145.

Daraio, C. and Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: a probabilistic approach. *Journal of Productivity Analysis*, 24, 93-121.

Daraio, C. and Simar, L. (2007). Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. *Journal of Productivity Analysis*, 28, 13-32.

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.

D'Inverno, G., Carosi, L., Romano, G., Guerrini, A. (2018). Water pollution in wastewater treatment plants: An efficiency analysis with undesirable output. *European Journal of Operational Research*, 269, 24-34.

Dong, X., Du, X., Li, K., Zeng, S., Bledsoe, B.P. (2018). Benchmarking sustainability of urban water infrastructure systems in China. *Journal of Cleaner Production*, 170, 330-338.

Dong, X., Zhang, X., Zeng, S. (2017). Measuring and explaining eco-efficiencies of wastewater treatment plants in China: An uncertainty analysis perspective. *Water Research*, 112, 195-207.

ERSAR (2019). Caracterização do setor de águas e resíduos; vol. 1 of Relatório Anual dos Serviços de

Águas e Resíduos em Portugal (Annual report on water and waste services in Portugal; 2019). *Entidade Reguladora dos Serviços de Águas e Resíduos* (Regulatory Authority for the Water and Waste Services)

Fuentes, R., Torregrosa, T., Ballenilla, E. (2015). Conditional order-*m* efficiency of wastewater treatment plants: the role of environmental factors. *Water*, 7, 5503-5524.

Giannakis, D., Jamasb, T., Pollitt, M. (2005). Benchmarking and incentive regulation of quality of service: an application to the UK electricity distribution networks. *Energy Policy*, 33, 2256-2271.

Gómez, T., Gémar, G., Molinos-Senante, M. Sala-Garrido, R., Caballero, R. (2017). Assessing the efficiency of wastewater treatment plants: a double bootstrap approach. *Journal of Cleaner Production*, 164, 315-324.

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.

Hernández-Sancho, F. & Sala-Garrido, R. (2009). Technical efficiency and cost analysis in wastewater treatment process: A DEA approach. *Desalination*, 249, 230-234.

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.

Kruskal, W.H., Wallis, W.A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583-621.

Kuosmanen T. (2002). Modelling blank data entries in data envelopment analysis. EconWPA working paper at WUSTL, no. 0210001 (Econometrics).

Le Lannier, A. & Porcher, S. (2014). Efficiency in the public and private French water utilities: prospects for benchmarking. *Applied Economics*, 46(5), 556-572.

Little, R., 1988. A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association,* 83(404), 1198-202.

Lorenzo-Toja, Y., Vásquez-Rowe, I., Amores, M.J., Termes-Rifé, M., Marín-Navarro, D., Moreira, M.T., Feijoo, G. (2016). Benchmarking wastewater treatment plants under an eco-efficiency perspective. *Science of the Total Environment*, 566-567, 468-479.

Marques, R.C., Berg, S.V., Yane, S. (2014). Nonparametric benchmarking of Japanese water utilities: institutional and environmental factors affecting efficiency. *Journal of Water Resources Planning and Management*, 140(5), 562-571.

Marques, R.C., da Cruz, N.F., Pires, J. (2015). Measuring the sustainability of urban water services. *Environmental Science & Policy*, 54, 142-151.

Marques, R.C., De Witte, K. (2010). Towards a benchmarking paradigm in European water utilities. *Public Money & Management*, 30:1, 42-48.

Marques, R.C. & Pinto, F.S. (2018). How to watch the watchmen? The role and measurement of regulatory governance. *Utilities Policy*, 51, 73-81.

Marques, R.C., Simões, P., Pires, J.S. (2011). Performance benchmarking in utility regulation: The worldwide experience. *Polish Journal of Environmental Studies*, 20(1), 125-132.

Mehta, M., Mehta, D. (2013). A review of performance benchmarking – Urban water supply and sanitation. *CEPT University*. Available at: https://pas.org.in/web/ceptpas/reportsandpapers

Molinos-Senante, M., Gémar, G., Gómez, T., Caballero, R. (2016a). Eco-efficiency assessment of wastewater treatment plants using a weighted Russell directional distance model. *Journal of Cleaner Production*, 137, 1066-1075.

Molinos-Senante, M., Gómez, T., Garrido-Baserba, M., Caballero, R., Sala-Garrido, R. (2014a). Assessing the sustainability of small wastewater treatment systems: A composite indicator approach. *Science of the Total Environment*, 497-498, 607-617.

Molinos-Senante, M. & Guzmán, C. (2018). Benchmarking energy efficiency in drinking water treatment plants: Quantification of potential savings. *Journal of Cleaner Production*, 176, 417-425.

Molinos-Senante, M., Hernández-Sancho, F., Mocholí-Arce, M., Sala-Garrido, R. (2014b). Economic and environmental performance of wastewater treatment plants: Potential reductions in greenhouse gases emissions. *Resource and Energy Economics*, 38, 125-140.

Molinos-Senante, M., Hernández-Sancho, F., Sala-Garrido, R. (2014c). 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., Marques, R.C., Pérez, F., Gómez, T., Sala-Garrido, R., Caballero, R. (2016b). Assessing the sustainability of water companies: A synthetic indicator approach. *Ecological Indicators*, 61, 577-587.

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.

Nudurupati, S.S., Bititci, U.S., Kumar, V., Chan, F.T.S. (2011). State of the art literature review on performance measurement. *Computers & Industrial Engineering*, 60(2), 279-290.

OECD (2012). OECD Multilingual Summaries. OECD Environmental Outlook to 2050. The consequences of inaction. http://www.oecd.org/env/indicators-modelling-outlooks/oecd-environmental-outlook-1999155x.htm

OECD (2015). OECD Principles on Water Governance Welcomed by Ministers at the OECD Ministerial Council Meeting on 4 June 2015. *Directorate for Public Governance and Territorial Development, Organization for Economic Cooperation and Development, Paris.* http://www.oecd.org/gov/regional-policy/OECD-Principles-on-Water-Governance-brochure.pdf

Pinto, F.S., Costa, A.S., Figueira, J.R., Marques, R.C. (2017). The quality of service: An overall performance assessment for water utilities. *Omega*, 69, 115-125.

Renzetti, S., Kushner, J. (2004). Full cost accounting for water supply and sewage treatment: concepts and case application. *Canadian Water Resources Journal*, 29(1), 13-22.

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(3), 1083-1087.

Sala-Garrido, R., Hernández-Sancho, F., Molinos-Senante, M. (2012). Assessing the efficiency of wastewater treatment plants in an uncertain context: a DEA with tolerances approach. *Environmental Science & Policy*, 18, 34-44.

Simar, L. (2003). Detecting outliers in frontier models: a simple approach. *Journal of Productivity Analysis*, 20, 391-424.

Thanassoulis, E., Portela M., Despic, O. (2007). DEA—The mathematical programming approach to efficiency analysis. In: Fried HO, Lovell CAK and Schmidt SS (eds). The Measurement of Productive Efficiency and Productivity Growth. *Oxford University Press: New York, Oxford*.

Thompson, R.G., Langemeier, L.N., Lee, C.T., Lee, E., Thrall, R.M. (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics*, 46, 93-108.

Torregrossa, D., Schutz, G., Cornelissen, A., Hernández-Sancho, F., Hansen, J. (2016). Energy saving in WWTP: daily benchmarking under uncertainty and data availability limitations. *Environmental Research*, 148, 330-337.

Vilanova, M.R.N., Filho, P.M., Balestieri, J.A.P. (2015). Performance measurement and indicators for water supply management: Review and international cases. *Renewable and Sustainable Energy Reviews*, 43, 1-12.

Wilson, P. W. (1995). Detecting influential observations in Data Envelopment Analysis. *The Journal of Productivity Analysis*, 6, 27-45.

Zanella, A., Camanho, A.S., Dias, T.G. (2015a). The assessment of cities livability integrating human wellbeing and environmental impacts. *Annals of Operations Research*, 226, 1, 695-726.

Zanella, A., Camanho, A.S., Dias, T.G. (2015b). Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. *European Journal of Operational Research*, 245, 2, 517-530.

Chapter 4 –

Leveraging Logistics Flows to improve the Sludge Management Process of Wastewater Treatment Plants

Chapter 4

Leveraging logistics flows to improve the sludge management process of wastewater treatment plants

This chapter refers to the following paper:

Henriques, Alda A.; Fontes, Milton; Camanho, Ana; Silva, J. Gabriel; Amorim, Pedro (2020). Leveraging logistics flows to improve the sludge management process of wastewater treatment plants. *Journal of Cleaner Production*, 276, 122720.

Overarching research question to be explored with this paper:

RQ2: Is it economically viable to optimize WWTPs' sludge management and biogas production through an interconnected and logistics approach that links different types of wastewater treatments at a system level with more than two plants?

Authors:

Alda A. Henriques, Milton Fontes, Ana Camanho, J. Gabriel Silva, Pedro Amorim

Thesis' author contribution:

Conceptualization (collaboration), methodology (collaboration), data curation (collaboration), software usage (collaboration), formal analysis (main collaboration), investigation, writing – original draft, writing – review and editing (collaboration), funding acquisition (collaboration).

Leveraging logistics flows to improve the sludge management process of wastewater treatment plants

Abstract

In this paper, we propose an innovative approach to integrate the wastewater treatment (WWT) that is carried out in a set of plants and the anaerobic digestion (AD) that is carried out in a different set of plants, all from the same system. The aim is to contribute to sustainability improvements (e.g., cost reduction and environmental and social benefits) within the system. The main drivers that support this approach are the existence of unutilized capacity in some digesters and the possibility of reducing the process sludge age to maintain its biogas potential during the aerobic treatment in the plants without AD facilities. Furthermore, the effect of co-digestion as a means of enhancing the amount of biogas that can be produced through the sludge is also considered. This innovative approach leads to a mixed-integer linear programming (MILP) model to optimize the sludge flows that connect the plants without AD facilities to the ones with anaerobic digesters. The developed approach was applied to a Portuguese water utility and the results provide new insights for an optimized sludge management process of WWTPs, also indicating that the economic benefit obtained from the proposed strategy for sludge management can lead to a reduction of up to 23.5% in terms of current energy costs.

Keywords: Wastewater treatment plants; logistics network design; MILP model; biogas production; sewage energy recovery

4.1. Introduction

Wastewater collection and treatment accounts for about 1% of global electricity consumption (IEA, 2018). In addition, the energy needs for wastewater treatment are expected to increase by 60% until 2040, due to increased demand (IEA, 2018). However, although the wastewater treatment plants (WWTPs) are energy-intensive facilities, it is widely documented that there is also a high potential for energy efficiency and energy recovery improvements. Moreover, on-site alternative green energy production strategies, like the use of photovoltaic panels (Chae et al., 2013), could also be a valid option.

The main goal is to achieve energy self-sustainability. In order to accomplish it, and particularly considering the treatment systems, there are two complementary objectives, as extensively described by Gu et al., (2017): (i) improving energy efficiency in the treatment processes, and (ii) recovering more energy from the generated waste sludge. In the literature, there are several approaches to meet these objectives. From a management perspective, performance assessment studies constitute an example of approaches to facilitate energy efficiency improvements (e.g., Hernández-Sancho et al., 2011). From an alternative perspective, Gikas (2017) proposes an innovative combination of technical and technological measures to improve both the energy efficiency of sludge production during the treatment of wastewater and the energy recovery from sludge. One of the most common technologies that is currently applied for energy recovery from sludge is anaerobic digestion (AD) (e.g., Madsen et al., 2011). In this context, Gretzschel et al. (2014) argue that the AD process should be considered even in the case of small WWTPs (with design capacities between 10 000 and 50 000 Population Equivalent). On the other hand, the sewage sludge is one of the primary wastes generated during the wastewater treatment process (Kwasny and Balcerzak, 2017). This waste is organic and must be treated accordingly, so that it can be safely disposed. As highlighted in Molinos-Senante et al. (2010), sludge management represents a significant share of the overall wastewater treatment costs.

So far, the literature has mostly adopted an isolated perspective over each WWTP, without acknowledging the possibility of connecting several facilities to address both the target of energy self-sustainability and the question of sludge management. Few studies consider the interconnections that may exist among different treatment processes or even among various facilities managed by the same entity.

Wett et al. (2007) describe a case study of a single WWTP in Austria that reached a positive energy balance through a set of strategies that included, among others, a maximization of sludge transfers from the aerobic treatment phase to the anaerobic digesters, for which only a previous minimum aeration time is required.

Liu and Wang (2015) focus on the sludge production and treatment, by referring to the connection that may exist between the aerobic and anaerobic treatment of sludge, to increase the biogas production as an output of AD. In this study, the authors discuss two alternative ways of sludge treatment optimization (i.e., reduction of sludge retention time or reduction of dissolved oxygen), both leading to either the reduction of energy consumption or the maximization of sludge production during the aerobic treatment. The final aim was to increase biogas production in the anaerobic treatment. However, this study was developed in a laboratory context, without testing the implications in a full scale WWTP or a set of connected WWTPs.

Mitchell and Beasley (2011) focused on sludge treatment by describing a linear programming model for optimizing the flows of sludge within a system with three different types of plants (i.e. plants that produce sludge without any kind of treatment, plants that conduct an intermediate type of treatment, and plants that perform a final treatment). The objective was to minimize logistics and treatment costs (fixed and variable). The developed model is a simplification of what is being applied in a water utility in the UK. In this study, however, maximizing the energy recovery from sludge was not a priority.

Hobus et al. (2010) introduced the concept of sludge flows from WWTPs that merely carry out aerobic treatment to WWTPs that have anaerobic digesters with available capacity. This was done by reducing the sludge retention time in the first type of plants, which decreases the electricity demand that is usually required for the complete stabilization of sludge using aerobic treatment. The sludge of the first type of plants is then used to increase the biogas-to-energy recovery from sludge in the second type of plants, since said sludge increases the organic loading in the digesters, as well as the feasibility in the use of the combined heat and power (CHP) systems - that otherwise may not be sustainable. However, this study addressed flows between two plants, leaving room for additional research on how this interconnected approach could be implemented, considering a more extensive system of plants. This constitutes a problem with two types of significant challenges: (i) the logistics challenge and (ii) the definition of the boundaries for the sludge quantities that each WWTP with a digester should ideally receive.

According to the work of Hobus et al. (2010), and taking into account the claim of Gretzschel et al. (2014) and the studies of Wett et al. (2007) and Liu and Wang (2015) - plus the possibility of sludge flows among different WWTPs, as described in Mitchell and Beasley (2011) - the main contribution of the research described in this paper is to test the hypothesis of achieving a more sustainable system in terms of energy consumption and sludge treatment, by turning several WWTPs that are currently managed as individual entities into elements of an integrated system managed accordingly.

This integrated approach considers logistics connections among plants, in order to enable the linkage of treatments through sludge flows, and the attainment of economic benefit through the energy balance achieved. To the best of our knowledge, this concept was never explored with a system of more than two WWTPs. In terms of sustainability gains, they ought to represent economic, social, and environmental benefits. The economic benefits are those that will drive the design of the modeling approach presented in the following sections of this paper. The environmental benefits will be obtained as a consequence of the energy sustainability improvements in WWTPs achieved with the proposed approach (e.g., energy efficiency and energy recovery advances). These can be considered as positive environmental externalities related to a more efficient wastewater treatment process achieved with the proposed approach. Furthermore, if the transportation vehicles are fueled with biogas (c.f., Forsberg, 2014; Shah et al., 2017), the environmental impact associated with the required logistics can be considered negligible. The social benefits correspond to improvements in the quality of service provided to customers, which is a desirable and viable result of the economic and environmental benefits attained.

To explore the hypothesis of this paper, a MILP model was developed to optimize the network design of existent WWTPs, and to leverage the sewage sludge flows among them. This model follows a comprehensive analysis of the aspects that should be considered when planning and designing waste biomass supply chains for energy recovery, presented in Iakovou et al. (2010). The different levels of decision making, the logistics aspects, and the energy conversion technologies are some of the topics that are discussed, based on the extensive literature review available in that paper. Accordingly, our model can be perceived as a forward, single product (sludge) supply chain design problem with two echelons (facilities that only carry out aerobic treatment and facilities that have anaerobic digesters with available capacity). A further in-depth review on supply chain design, specifically considering MILP models, can be conducted by analyzing, among others, the papers of Ahmadi-Javid et al. (2018), Balaman and Selim (2014), Entezaminia et al. (2016), Galvez et al. (2015), Görgüner et al. (2014), Lashine et al. (2006), Mota et al. (2015), Mota et al. (2018), Özceylan and Paksoy (2013), Pourjavad and Mayorga (2019), Saffar et al. (2014), Rabbani et al. (2018), Ramudhin et al. (2010), Santibañez-Aguilar et al. (2014) and Zhu et al. (2011). Furthermore, Barbosa-Póvoa et al. (2018) conducted an extensive review upon the topic of sustainable supply chain design, identifying the greatest challenges and opportunities in this research area.

In this paper, we address the operational and technical specifications required to improve the sludge treatment and stabilization, namely in determining the amount of sludge to be transported, so that the logistics costs can be well compensated. The effect of the incorporation of co-digestion is further analyzed, which enables treating another type of organic waste without the spending of additional resources (Pan et al., 2015). Co-digestion may have a significant impact on AD performance by improving the biogas yield (Mao et al., 2015), generally through the increase of the ratio of volatile solids over total solids and through the improvement of the organic loading (Pavan et al. 2007). Moreover, co-digestion also allows the provision of nutrients that register deficient amounts in the sewage sludge (Hallaji et al., 2019). This factor has a catalytic effect on biogas production. However, integrating co-digestion into a conventional AD system leads to an increase in the system's complexity (Hagos et al., 2017). This complexity may derive from the required operating conditions for optimal outcomes, such as the temperature and pH that may be dependent on the chemical composition of co-substrates. In this study, the main concern is modeling the logistics network when considering the benefits of co-digestion in terms of the enhancements of biogas production. Thus, the needed requirements in the

digesters for optimal outcomes are assumed to be met, which includes the selection of a type of co-substrate that is compatible with the sludge at AD facilities.

This paper contributes to the literature with the development of a network design model, presented in Section 4.2. Sludge flows between two types of facilities already existent (those with and without AD) are leveraged by the model developed. The objective is to improve the system in terms of sludge management, by increasing the energy efficiency of the treatment processes, as well as the energy recovery from sludge. Additionally, according to the network design approach, there's been a research and modelling of some of the key parameters to integrate core operational and structural aspects of sludge treatment and management. The critical challenge is to integrate the functioning of anaerobic digestion into the model. A real-world case study comprehending a set of WWTPs managed by the Portuguese water utility Águas do Centro Litoral – described in Section 4.3 – is used to validate the feasibility of the approach developed. Moreover, Section 4.4 presents the main conclusions and suggestions for future developments.

Nomenclature

Model Specification

Indices

i	Export plant
j	Destination plant
k	Potential number of digesters at a destination plant that can be selected

Parameters

V_{jk}	Monetary value that is accrued by selling the electricity generated from the biogas, which is produced through the anaerobic digestion of sludge transported to the k digesters at each destination plant j $(€/\text{ton})$
P_i	Amount of sludge produced at the export plant i (ton)
c_{ij}	Logistics costs associated to arc (i, j) $(€/\text{ton})$
SC_i	Saved costs at export plant i (€)
SC_j	Saved costs at destination plant j (€)
β_i	Minimum percentage of sludge that must be transported from the export plant i in case it is selected to make part of the logistics network (%)
LB_{jk}	Lower bound of the sludge quantity that can be incorporated in k digesters at destination plant j (ton)
UB_{jk}	

Auxiliary Parameters

%S	Percentage of TS in the sludge
TS	Total solids (potentially biodegradable solids + non-degradable solids)
VS	Volatile solids (potentially biodegradable solids)
RatioVT	Percentage of VS in TS in the sludge
V	Volume of the anaerobic digester (m^3)
HRT	Hydraulic retention time, which corresponds to the number of days the sludge stays inside the anaerobic digester undertaking the process of anaerobic stabilization
μ	Electricity conversion factor of the CHP system (%)
η	Rate of degradation of volatile solids in the anaerobic digestion $(\text{VS}_{\text{destroyed}}/\text{VS})$
Q	Amount of sludge at the entry of each digester (ton/day)
M	Unitary price that is legally contracted for selling the electricity produced $(€/\text{kWh})$
R	Rate of biogas production $(\text{m}^3_{\text{biogas}}/\text{kg VS}_{\text{destroyed}})$
P_j	Amount of sludge produced at the destination plant j (ton)
x	Amount of co-substrate to be considered (ton)

	Upper bound of the sludge quantity that can be incorporated in k digesters at destination plant j (ton)	E	Energy content of the biogas produced (kWh)
Decision Variables		t_l	Estimated time for loading and unloading the sludge (h)
q_{ijk}	Continuous variable that represents the sludge flows (amount of sludge to be transported from each export plant i to k digesters at each destination plant j)	C	Cost of truck (€/(ton.h))
		d_{ij}	Length of the arc (i, j) (km)
y_i	Binary variable that assumes the value of 1 if the export plant i is selected, 0 otherwise	v	Truck's average speed (km/h)
α_{jk}	Binary variable that assumes the value of 1 if k digesters are selected at destination plant j, 0 otherwise		

4.2. Methodology

This section provides the description of the problem to be modeled, followed by the corresponding mathematical formulation.

4.2.1 Problem statement

Let us consider a system composed of two types of sludge production and treatment WWTPs: the ones for conventional treatment (aerobic treatment), with no structural means for conducting AD, and the ones conducting AD where there is potentially unutilized capacity in the digesters. The same entity manages all the WWTPs.

The optimization opportunity addressed in this study is based on the possibility of obtaining an economic benefit through a positive energy balance in the overall system, when both types of WWTPs are connected through flows of sludge, i.e., a logistics network. From now on, WWTPs that currently treat sludge in a conventional way (i.e., aerobic stabilization) are designated by export plants, since they constitute the origin of sludge flows. The WWTPs that currently conduct anaerobic stabilization of sludge (that have anaerobic digesters) are hereby identified as destination plants, because they constitute the endpoint of sludge flows. The problem is to find the best logistics flows to connect export and destination plants (see Figure 4.1). Moreover, the optimal structural operating conditions at the anaerobic digesters (i.e., HRT and mass capacity) also need to be determined. In short, the focus is finding a balance between the plants of the designed logistics system, in order to obtain the maximum economic benefit, considering:

- The current existing structure of the system of WWTPs, i.e., number and location of export plants, number and location of destination plants, and number and type of anaerobic digesters at each destination plant.
- The current operational parameters of the WWTPs, i.e., sludge production at export plants, sludge production at destination plants, and capacity utilization of anaerobic digesters.

- The operational restrictions of export plants, and the operational restrictions of the anaerobic digesters.
- The possibility of incorporating co-digestion[2].

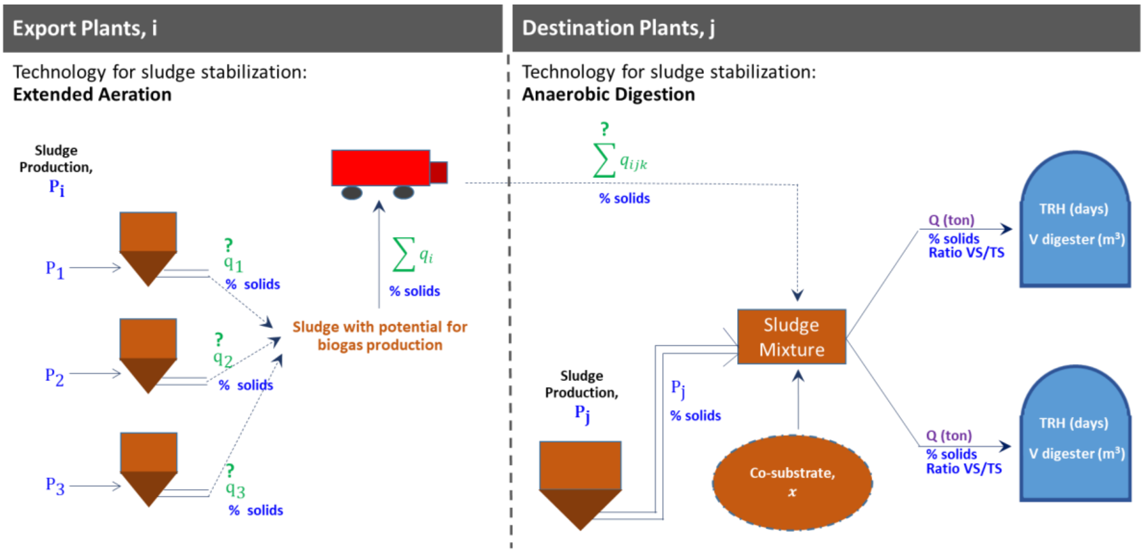


Figure 4.1 *– Schematic representation of the problem of sludge flows between two different types of facilities.*

Simultaneously, there is the possibility of adopting an optimized sludge production during the aerobic treatment, particularly in terms of the electricity that is currently used. This can be done by reducing the sludge retention time as supported by Ge et al. (2013), and/or by reducing the oxygen demand, which according to Liu and Wang (2015) are the two ways to maximize the sludge biogas potential and to increase the amount of biogas that is produced in plants with AD.

In this sense, Figure 4.2 provides a systematized overview of the methodological framework developed in this paper, aimed at tackling the issue of identifying the optimized sludge flows (or other organic wastes) between two types of facilities – in order to obtain an economic benefit from the enabled positive energy balance.

[2] In this study, we did not consider the logistics flows between the companies that will provide the co-substrate and the destination WWTPs. Usually, destination WWTPs are relatively close to industries that could provide organic co-substrates.

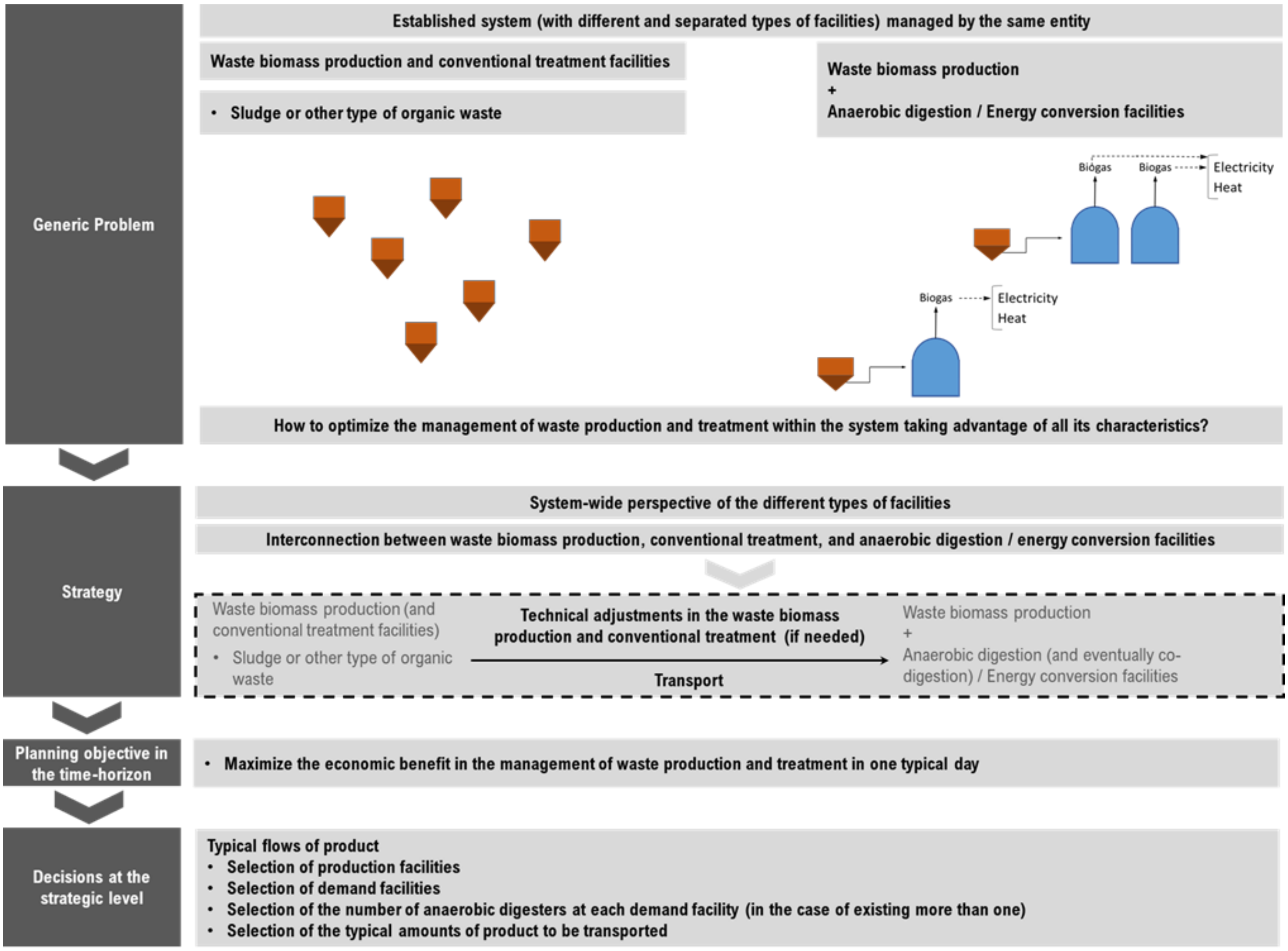


Figure 4.2 *– Methodological framework for the improvement of sludge management (or other organic waste).*

4.2.2 Model development

This section includes the development of the core mathematical model that defines the optimal network design into a mixed-integer linear programming (MILP) model, and, secondly, the description of the approach for the estimation of the model parameters.

4.2.2.1 Core model mathematical formulation

Usually, AD requires a constant flow of sludge to operate and maintain its daily balance. Hence, for modeling the aforementioned strategic problem, one typical day of operation is considered as planning horizon. Additionally, the following assumptions are considered:

- The sludge demand in each destination plant may be satisfied using sludge from several export plants, notwithstanding the type of sludge.
- Each export plant may supply sludge for more than one destination plant.
- There is a finite number of export and destination plants in the system targeted for optimization.
- It is not mandatory to meet the demand at all destination plants.
- It is not mandatory to use all the export plants to supply sludge.

Two sets are considered: $I = \{1, \dots, n\}$ represents all export plants and $J = \{1, \dots, m\}$ represents all destination plants. For each plant $j \in J$, there is a set $K = \{1, \dots, s\}$, which represents the different number of digesters that can be selected in the optimization process at each destination plant. The index i will be used to designate an export plant, the index j will identify a destination plant and the index k will be used to designate the potential number of digesters that can be selected at a destination plant. Each pair of plants, i and j, is separated by a distance that corresponds to the length of the correspondent arc (i, j). Each arc (i, j) has an associated logistics cost of transporting one ton of sludge, represented by c_{ij}.

Each plant i produces an amount of sludge, P_i , in the time horizon, and each plant j has a certain available capacity to contain the transported sludge, which depends on the number of digesters selected, k. Additionally, the available capacity also depends on other parameters of each anaerobic digester at a single plant j. Therefore, the lower and upper bound (LB_{jk} and UB_{jk}) are defined for the available capacity of the number of digesters k - relative to a specific plant j -, in order to ensure the model's freedom to select the best amount of sludge that can eventually be transported to each plant j.

There are three decision variables to be determined: q_{ijk} is a continuous variable that corresponds to the amount of sludge to be transported from plant i to k digesters at plant j; y_i is a binary variable that assumes the value of 1 if plant i is selected and 0 otherwise, and α_{jk} is a binary variable that assumes the value of 1 if k digesters are selected at plant j and 0 otherwise.

The optimization problem is formulated as follows:

$$max \sum_i \sum_j \sum_k V_{jk} \cdot q_{ijk} - \sum_i \sum_j \sum_k c_{ij} \cdot q_{ijk} + \sum_i SC_i \cdot y_i + \sum_j \sum_k SC_j \cdot \alpha_{jk} \quad \textbf{(4.1)}$$

$$s.t. \quad \sum_j \sum_k q_{ijk} \leq P_i \cdot y_i\,, \ \forall i \in I \quad \textbf{(4.2)}$$

$$\sum_j \sum_k q_{ijk} \geq \beta_i \cdot P_i \cdot y_i, \ \forall i \in I \quad \textbf{(4.3)}$$

$$\sum_i q_{ijk} \geq LB_{jk} \cdot \alpha_{jk}, \ \forall j \in J, k \in K \quad \textbf{(4.4)}$$

$$\sum_i q_{ijk} \leq UB_{jk} \cdot \alpha_{jk}, \ \forall j \in J, k \in K \quad \textbf{(4.5)}$$

$$\sum_k \alpha_{jk} \leq 1, \forall j \in J \quad \textbf{(4.6)}$$

$$y_{i,} \alpha_{jk} \in \{0,1\}, \forall i \in I, k \in K \quad \textbf{(4.7)}$$

$$q_{ijk} \geq 0, \forall i \in I, j \in J, k \in K \quad \textbf{(4.8)}$$

Equation (4.1) is the objective function of the optimization problem and corresponds to the maximization of the economic benefit obtained from the overall sludge management. This equation is composed of four terms: i) the monetary value (V_{jk}) that is accrued by selling the electricity generated from the biogas, which is produced through the anaerobic digestion of the total amount of sludge transported to the digesters, ii) the logistics costs (c_{ij}), iii) the saved costs (SC_i) at plant i, and iv) the saved costs (SC_j) at plant j.

Equation (4.2) sets the limit to the total amount of sludge transported to final destination plants, according to the amount of sludge produced at each export plant i (P_i). Equation (4.3) imposes a minimum value ($\beta_i \cdot P_i$) for the amount of sludge that is necessary to transport from each plant i, namely in cases where said plant is selected to be part of the network. The reason for this constraint is that the sludge will be stabilized in the final destination plant, making it necessary to transport the maximum amount of sludge produced to a destination plant - otherwise, the sludge cannot be stabilized. The remaining amount of sludge ought to be kept at the export plant as inventory.

Equations (4.4) and (4.5) provide the lower and upper boundaries (LB_{jk} and UB_{jk}) concerning the sludge quantities that can be incorporated in the k digesters at each plant j, which are related to the structural parameters of each digester. These boundaries account for the sludge that is produced at each plant j, which has priority over the transported sludge to be incorporated in the digesters. These boundaries also account for the amount of co-substrate that is determined to be added to the sludge produced at each plant j.

Equation (4.6) shows that only a set of digesters at each plant j is chosen. Equation (4.7) sets the binary nature of variables y_i and α_{jk}. Finally, equation (4.8) is the non-negativity constraint.

4.2.2.2 Model parameters

The model depends heavily on a set of parameters, which are either part of the objective function or the constraints. These parameters mostly relate to the specificities of the system's structure and operation. This section comprehends the analysis of the approach to obtain the monetary value, associated with the sludge (V_{jk}), the logistics costs (c_{ij}), the saved costs (SC_i, SC_j), the sludge production (P_i, P_j), and the boundaries that define the amount of sludge that can be transported (LB_{jk}, UB_{jk}).

For the estimation of some of the parameters, the following assumptions are considered:

- The sludge transported, as well as the sludge produced in all plants j, is entirely used to produce biogas.
- At each plant j, all digesters have the same structural features, namely the same volumetric capacity, and the same operating characteristics.
- All sludge ought to be mixed before entering the digesters, thus enabling the same percentage of total solids and the same percentage of volatile solids over total solids (ratioVT) in the sludge at the entry of all digesters.
- All anaerobic digestions, in all destination plants, have the same rate of degradation of VS (h) and the same rate of biogas production (R).
- All biogas produced is used for electricity and heat generation in a CHP system.
- The heat generated in the co-generation process is enough to heat all the sludge that enters the digesters.
- All the electricity generated is sold to the electricity market under a fixed cost per unit, regardless of the contract agreements for electricity sale or the direct use of electricity in the WWTP.

4.2.2.2.1 Monetary value associated with the sludge *(V_{jk})*

For the estimation of the monetary value associated with each ton of sludge entering each digester (V_{jk}), the starting point is to consider the total amount of volatile solids (VS) existing in that amount of sludge. That will be the sludge's component decomposing through the AD process to produce biogas. Therefore, it is also necessary to multiply the VS by the terms h and R. Additionally, it is necessary to consider the energy content, E, of the total amount of biogas that is expected to be produced, multiplied by the electricity conversion factor, μ, to know the electricity that is potentially produced. Moreover, this value should be further multiplied by the cost per unit that is legally contracted, M. Finally, the maximum amount of sludge expected (Q_e)

entering each digester also needs to be considered. Then, equation (9) (adapted from Andreoli et al., 2007, pp. 60-63) is used to estimate V_{jk}.

$$V_{jk} = \frac{VS \times \text{h} \times R \times E \times \mu \times M}{Q_e} \tag{4.9}$$

The value of Q_e is calculated according to the minimum HRT and the volumetric capacity of the digester (V), as demonstrated by equation (4.10).

$$Q_e = \frac{V}{HRT_{min}} \tag{4.10}$$

The value of Q_e, along with the percentage of total solids (%S) and the ratio between volatile over total solids ($RatioVT$) that characterizes it, is then used to obtain the total amount of VS, using equation (4.11).

$$VS = Q_e \times 1000 \times \%S \times RatioVT \tag{4.11}$$

Figure 4.3 shows a schematic representation of the steps followed in the estimation of V_{jk}.

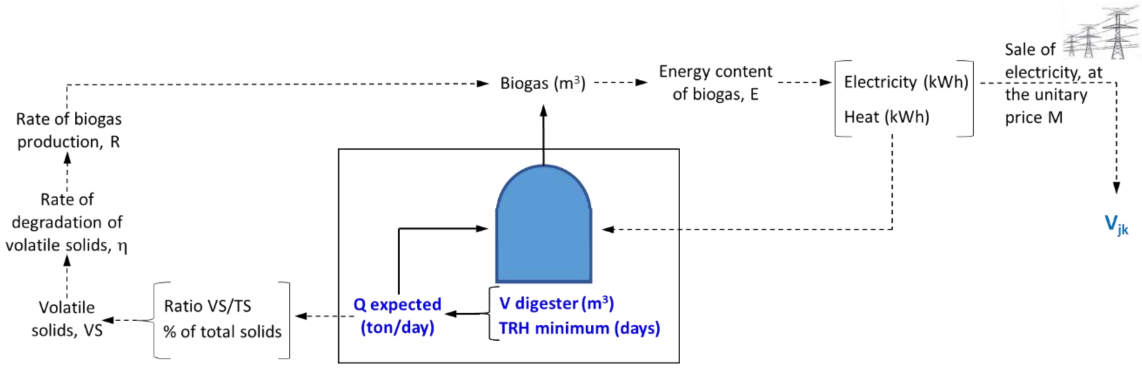


Figure 4.3 *– Step by step schematic representation of the estimation of V_{jk}.*

The difference between the amount of sludge that enters each digester, Q, as determined on the basis of the optimal solution of the model, $\sum_i q_{ijk}$, and the amount of sludge that is expected at each digester, Q_e, is assumed to be relatively small (it basically depends on the range of values between HRT_{min} and HRT_{max}, as it will be later explored). Therefore, the monetary value associated with Q is expected to be significantly similar to the one estimated for Q_e in the same digester.

4.2.2.2.2 Logistics costs (c_{ij})

In order to calculate the logistics cost associated to each arc (i, j), several factors are considered: i) the length of the arc (i, j), d_{ij}, ii) the average truck's speed, v, iii) the time that is necessary to load and unload the sludge, t_l, iv) the hourly cost of the truck, related to its type and capacity, C, and v) the dilution factor, $\frac{\%S_{digester}}{\%S_{transport}}$, in the case of export plants that conduct dehydration. Equation (4.12) gives the logistics cost corresponding to an arc (i, j), where the export plant conducts a dehydration process of sludge. In the case of the plants where this process does not occur, the dilution factor is omitted.

$$c_{ij} = \left(\frac{d_{ij}}{v} + t_l\right) \times C \times \frac{\%S_{digester}}{\%S_{transport}} \tag{4.12}$$

4.2.2.2.3 Saved costs at export plants (SC_i)

The saved costs at export plants, SC_i, relate mainly to the decrease in electricity consumption. This is a result of the reduction of the aeration time so the sludge can maintain its potential for biogas production. Therefore, the saved costs will be proportional to the reduction of electricity that is possible to achieve, which will directly relate to the minimum aeration time at which it is possible to guarantee the wastewater treatment.

4.2.2.2.4 Saved costs at destination plants (SC_j)

The saved costs at destination plants, SC_j, relate to a potential increase in h, which will yield a lower volume on the total sludge digested and, consequently, lower dehydration and transportation costs for the final disposal of sludge - even if there is a higher amount of sludge feeding the digester. There are two reasons for this expected outcome: i) higher use of the organic capacity at each digester and ii) the incorporation of co-digestion. Nonetheless, the increase of maintenance costs related to a higher use of the CHP system, due to an increase in the number of hours required to generate the electricity from the additional biogas produced, may, eventually, offset (at least partially) the saved costs at destination plants.

4.2.2.2.5 Sludge production at export and destination plants (P)

The sludge produced at each plant, after the thickening process, P, is estimated based on the sludge produced annually ($SP_{annually}$), according to equation (4.13).

$$P = \frac{SP_{annually}}{\%S \times 365} \tag{4.13}$$

The term $SP_{annually}$ is estimated according to the observed percentage of utilization of the installed capacity (in Population Equivalent), and the typical amount of sludge that is produced per inhabitant.

4.2.2.2.6 Boundaries to the amount of sludge transported to each destination plants (LB_{jk} , UB_{jk}): Sludge demand

The boundaries to the amount of sludge transported are given by Equations (4.14) and (4.15) i.e. the lower and upper bounds, correspondingly[3]. These boundaries define the limits for the sludge demand at each plant j. For its computation, both the structural conditions of each digester (e.g., volumetric capacity and number of digesters) and operational conditions (e.g., minimum and maximum possible HRT and the sludge that is produced at each plant j, P_j) are accounted for. P_j has priority over the potentially transported sludge, to be used entirely. This ultimately means that there may be no available capacity at plant j to accommodate transported sludge.

$$LB_{jk} = \frac{V}{HRT_{max}} \times k - P_j \tag{4.14}$$

$$UB_{jk} = \frac{V}{HRT_{min}} \times k - P_j \tag{4.15}$$

where $k \in \{1, \ldots, s\}$, and s is the total number of digesters at each plant j. Whenever co-digestion is considered for the design of the logistics network, the factor P_j in equations (4.14) and (4.15) will be substituted by the factor $(P_j + x)$, where x is the total amount of co-substrate

[3] It is considered that the density of sludge with up to 6% solids is equal to the density of water, i.e., $1m^3$ = 1ton (Andreoli et al., 2007).

to be considered (for more details about how the computation of x is modeled, see Appendix 4A). As one can observe, the addition of co-substrate to the sludge results in an increase in the ratio VS over TS, compared to the situation of the single substrate (sewage sludge). Furthermore, when the co-substrate is added to the sludge, it increases its percentage of total solids. Therefore, we consider the addition of co-substrate to achieve the percentage of total solids in the sludge mixture expected at the entry of the anaerobic digester[4]. This replaces the need for a mechanical thickener to obtain the desired percentage of total solids, while preventing the corresponding electricity consumption requirement. Figure 4.4 shows how the sludge demand at each destination plant is set as a variable parameter of the MILP model.

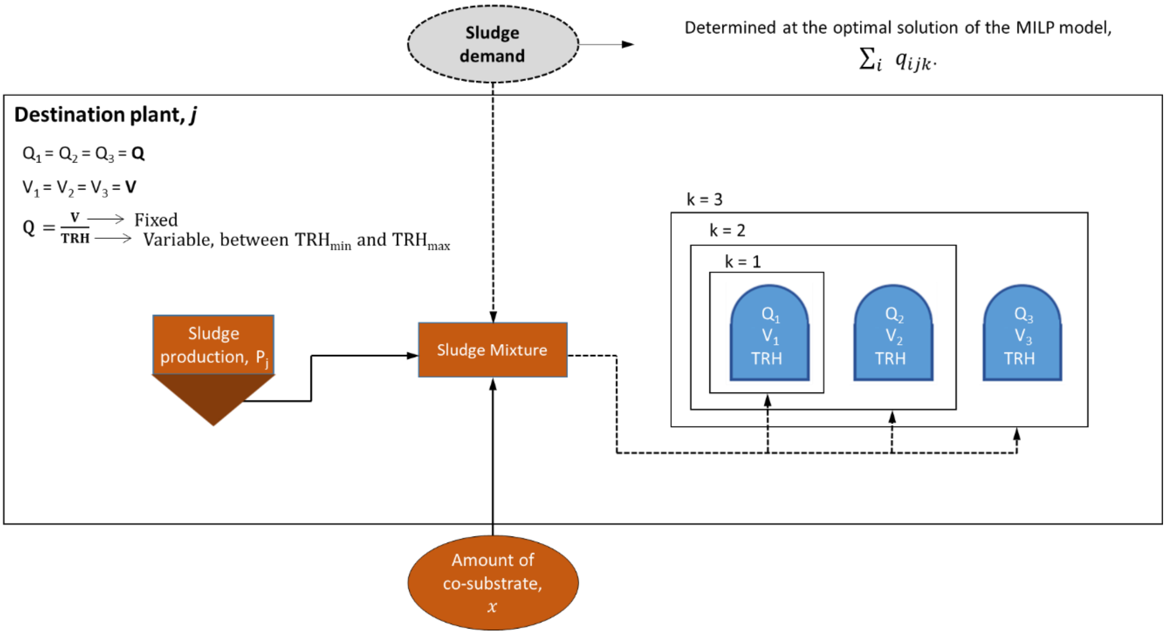


Figure 4.4 *– Schematic representation of the sludge demand at a destination plant j with three anaerobic digesters.*

A practical implication deriving from the demand being determined by the model, rather than fixed a priori, is that the HRT of each digester is determined as a consequence of the optimal solution of the MILP model ($\sum_i q_{ijk}$), which may have a significant impact on the current operating conditions of the system (e.g., h and energy sustainability). For example, if a digester is currently operating at a considerably high HRT, which indicates a low amount of sludge at the entry of the digester, the current amount of biogas produced will not be enough to sustain the anaerobic digestion in terms of energy. In this case, the logistics model may determine the transportation of enough sludge to this digester, in order to achieve a situation of energy sustainability. To avoid worsening the operating conditions of each digester in terms of the AD viability and h, a value for the HRT_{min} equal or higher than that currently observed may be set, whereas a plausible value for the HRT_{max} (not higher than 25 days) should also be used. This will prevent setting a digester to function with insufficient amounts of sludge.

[4] We assume that the amount of sludge transported from export plants to each destination plant is significantly lower than the amount of sludge produced at the destination plant. In this sense, it will not have a significant impact on the features of the mixture obtained by adding the co-substrate to the sludge produced at the destination plant.

4.3. Results and discussion

The logistics model developed in the previous section was tested in a real-world case study. The model was implemented in IBM ILOG CPLEX Optimization Studio, version 12.6. All instances run in less than 1 second in an Intel®Core™I7-6500 processing units, at 2.50 GHz and 16.0 GB memory. The total number of constraints is 76, while the variables are 342, from which 35 are binary.

This section describes the selected case study and the results obtained from the model application. The results concern both the model application to the case study conditions and the sensitivity analysis conducted.

4.3.1 Case study description

The case study considered is a set of 17 export plants and 6 destination plants in Portugal. These plants are located in the coastal center region of the country, and are managed by the same water utility Águas do Centro Litoral (AdCL) (see Figure 4.5).

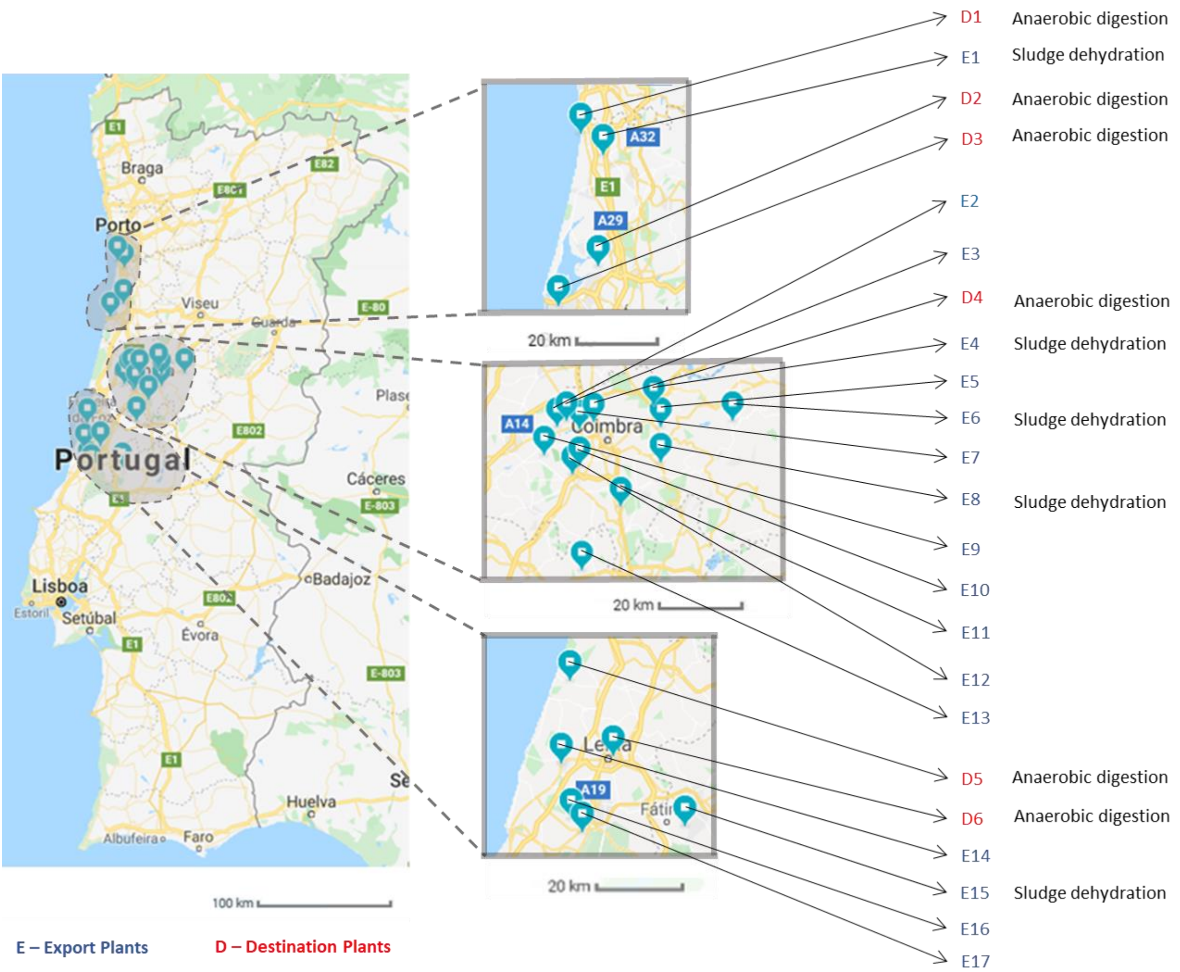


Figure 4.5 *– Geographical location of the WWTPs of the case study.*

The average distance between export and destination plants is 85.6 km. The biogas that is produced in destination plants is stored before being converted into electricity and heat through a CHP system. The storage allows the optimization of the energy production, according to the end-use of the electricity produced, the need for heat within the plant, and the monetary return that can be accrued when selling the electricity produced. For the sake of modeling simplicity, and as described in section 4.2.2.2, we will consider that all the electricity generated is sold to the electricity market according to a cost per unit. Table 4.1 shows the main characteristics of all plants considered in the case study.

Table 4.1 – Overview of data regarding the WWTPs of the Portuguese case study (year 2015).

		Export Plants	**Destination Plants**
Technology for sludge stabilization		Extended aeration	Anaerobic digestion
Technology for sludge treatment after stabilization	Gravity Thickeners	17	6
	Dehydration	5	6
Population Equivalent (PE)	Average	12 572	186 202
	Min	2640	59 596
	Max	45 000	272 000
% of utilization of the installed capacity of the plant	Average	54%	64%
	Min	6.7%	48%
	Max	155%	76%
Sludge production (after thickening)	Average	17.2 ton	299.1 ton
	Min	1.2 ton	49.7 ton
	Max	40.9 ton	821.6 ton
Number of digesters	Average	---	2
	Min	---	1
	Max	---	3
Technology of all anaerobic digesters		---	High-rate digester
Operating temperature of anaerobic digesters		---	35ºC
% of utilization of the volumetric capacity of digesters	Average	---	75%
	Min	---	48%
	Max	---	130%
% of utilization of the mass capacity of digesters	Average	---	36%
	Min	---	15%
	Max	---	45%

4.3.2 Model implementation and results

The model implementation comprehended the testing of six alternatives (A1 to A6). Each alternative corresponds to different combinations of the percentage of total solids in the sludge entering the digesters and the ratio VS over TS (see Table 4.2). Both the percentage of total solids and the ratio VS over TS are characteristics of the sludge that can be increased by introducing co-digestion. The alternatives are tested under two different scenarios (Sc1 and Sc2), each corresponding to different values of h, which is a proxy for the anaerobic digestion performance that is likely to occur. The use of scenarios regarding the value of h aims to acknowledge the unpredictability of the anaerobic digestion performance, given the complexity of such biological process that is usually referred as a black box (Premier et al., 1999; Beltramo et al., 2019). Though the use of two different scenarios can provide insights into the influence that different values of h may have on the economic viability of the logistics network, the same value of h is considered for all digesters in each scenario, for the sake of simplicity. However, it is important to point out that, from the model solution, different HRT may be set for the digesters at each plant j. This particularity, alongside with the integration of co-digestion in some of the alternatives tested, is likely to originate different values of h in each digester, even if the operating temperature is quite similar. Nonetheless, it is expected that the values considered are lower bounds of h, which in practical terms, means that better results for the economic benefit can be even further achieved. In scenario 1 (Sc1), the value considered for h is 35%, and in scenario 2 (Sc2) is 45%.

Table 4.2 – Characterization of the alternatives for analysis.

Alternative	Characteristics of sludge during transport	Characteristics of the sludge mixture at the entry of digesters	Co-Digestion
A1	2%S; 70% VS/TS	2%S; 70% VS/TS	No
A2	6%S; 70% VS/TS	6%S; 70% VS/TS	No
A3	6%S; 70% VS/TS	3%S; 80% VS/TS	Yes
A4	6%S; 70% VS/TS	4%S; 80% VS/TS	Yes
A5	6%S; 70% VS/TS	5%S; 80% VS/TS	Yes
A6	6%S; 70% VS/TS	6%S; 80% VS/TS	Yes

In addition, for the application of the logistics model to the case study, the values of the parameters considered are indicated in Table 4.3.

Table 4.3 – Values of the logistics model's auxiliary parameters used in the case study application.

Auxiliary parameter	Notation	Value considered in the case study	Unit
Inventory capacity in terms of the sludge production[5]	$1-\beta_i$	30	%
Minimum hydraulic retention time	HRT_{min}	20	days
Maximum hydraulic retention time	HRT_{max}	24	days
Rate of biogas production (Andreoli, 2007)	R	0.8	m^3_{biogas} / kg $\mathrm{VS}_{\mathrm{destroyed}}$
Energy content of the biogas produced, considering 70% CH_4 (Andreoli, 2007)	E	23.3	MJ/m^3_{biogas}
Electricity conversion factor of the CHP system	μ	35	%
Price per unit that is legally contracted for selling the electricity produced	M	0.14	€/kWh
Truck's average speed	v	70	km/h
Estimated time for loading and unloading the sludge	t_l	2	h
Cost of truck (dehydrated sludge)	C	2	€/(ton.h)
Cost of truck (thickened sludge)	C	3.44	€/(ton.h)

4.3.2.1 Economic feasibility of the model and impact on the total biogas in the system

The model implementation allows different perspectives of analysis of the logistics network regarding the management of sludge. The economic feasibility of the model is the first angle worth of assessment. Secondly, the results of the model implementation - in terms of the impact of the logistics network on the total amount of biogas that can be produced in the system - will be further analyzed and compared with the total amount of biogas that is currently produced in the system of WWTPs of the case study. The reason for this second approach is that by improving the operational conditions of digesters, there is a positive catalytic effect that improves the total biogas production (i.e., not only produced through the sludge transported, but through the complete mixture of sludge that enters each digester). This is due to a higher amount of sludge that feeds the digesters and the effect of co-digestion considered in many alternatives.

A comprehensive overview of the results obtained in terms of the economic feasibility, the structure of the network, the operational implications in the anaerobic digesters in each setting, and the total biogas produced in the system for each alternative and scenario considered are provided in Table 4.4 For a visual analysis of the network configuration (respecting Sc1), the reader can refer to the schematic representations in Appendix 4B.

[5] The same value of β is considered for all export plants.

Table 4.4 – Results of the logistics model implementation.

Alternative	Scenario	Economic benefit (€/year)	Export plants	Destination plants	Total amount of sludge to be transported (ton/day)	Number of digesters selected	Total amount of sludge at the entry of each digester (ton/day)	HRT (days)	Organic loading rate	Differential in biogas production in the system[6] (m³/day)	
										Relative to the correspondent theoretical conditions without the logistics network	Relative to the observation
A1	Sc1	74 789	E1	D1	20.5	1	129.3	20	0.7	481	712
			E1, E4, E6, E8, E15	D4	76.7	2	136.1	22	0.64		
			E15	D5	20.4	2	225	20	0.7		
			E15	D6	5.3	1	55	20	0.7		
	Sc2	90 812	E1	D1	20.5	1	129.3	20	0.7	615	2 856
			E1, E4, E6, E8, E15	D4	75.8	2	135.7	22	0.63		
			E15	D5	20.4	2	225	20	0.7		
			E15	D6	5.3	1	55	20	0.7		
A2	Sc1	76 358	E1, E3, E6	D2	13.2	1	287.1	22	1.88	499	731
			E4, E6, E8, E15, E17	D6	29.3	1	45.8	24	1.75		
	Sc2	93 002	E1, E3, E4, E6	D2	15.6	1	289.5	22	1.89	677	2 920
			E4, E8, E10, E15, E17	D6	29.3	1	45.8	24	1.75		
A3	Sc1	51 976	E1, E3, E4, E6, E8, E10, E15, E17	D4	44.7	2	125	24	1	301	6 157
			E17	D6	0.2	1	52.4	21	1.14		
	Sc2	61 941	E1, E3, E4, E6, E8, E10, E15, E17	D4	44.7	2	125	24	1	387	9 860
			E17	D6	0.2	1	52.4	21	1.14		
A4	Sc1	65 810	E1	D1	8.1	1	127	20	1.57	403	11 177
			E3, E4, E6, E8, E10, E15	D4	36.2	2	125	24	1.33		
			E15, E17	D6	0.7	1	55	20	1.6		
	Sc2	79 242	E1	D1	8.1	1	127	20	1.57	518	16 315
			E3, E4, E6, E8, E10, E15	D4	36.2	2	125	24	1.33		
			E15, E17	D6	0.7	1	55	20	1.6		
A5	Sc1	76 285	E1	D1	5.2	1	129.3	20	2	504	16 625
			E1, E3, E4, E6, E8, E10, E15, E17	D4	39.7	2	131.4	23	1.75		
	Sc2	93 586	E1	D1	5.2	1	129.3	20	2	685	23 356

[6] All destination plants excluded from the logistics network are considered to be producing biogas according to the theoretical modeling (c.f. Section 4.3.2.1).

			E1, E2, E3, E4, E6, E8, E10, E15, E17	D4	42.4	2	132.7	23	1.77		
A6	Sc1	87 089	E1, E2, E3, E4, E6, E8, E10, E15, E17	D4	47.5	2	140.3	21	2.25	637	22 586
	Sc2	108 077	E1, E2, E3, E4, E6, E8, E10, E11, E15, E17	D4	49.8	2	141.5	21	2.26	861	31 024

All cases tested proved that the model is economically feasible, even when considering a low percentage of total solids in the sludge. This means that it is possible to manage the sludge that is produced in the set of WWTPs selected for the logistics network in a profitable way. When only considering the criterion of economic benefit, the best alternative allows the achievement of 108077 €/year, which corresponds to approximately 23.5% of the current energy costs in the same set of plants.

Generally, it can be observed that the design of the logistics network is independent of the performance of AD, since for each alternative, the export and destination plants remain the same under the two scenarios (alternatives A1, A3, and A4), or display only slight differences (alternatives A2, A5 and A6). One of the consequences, as expected, is the growth in the economic benefit with the increase in AD performance (higher level of performance in Sc2 than in Sc1), since the total amount of sludge transported is roughly the same. Regarding the structural operating parameter TRH, it is independent of the alternative tested, which means that the amount of sludge transported is distributed through the digesters regardless of its characteristics or the characteristics of the sludge mixture entering the digesters. Moreover, it is worth mentioning that the export plants that conduct dehydration were selected to be part the network (i.e., plants E1, E4, E6, E8, and E15) in all cases.

When the effect of co-digestion is not considered, the logistics network has a positive impact on the total biogas produced in the system. This impact corresponds to a minimum of 6.8% higher than the total amount of biogas expected under the correspondent theoretical conditions, without the logistics network (in alternative A1 under both scenarios), and to a minimum of 10.5% higher than the total amount of biogas currently produced in the system (in the same alternative). When the effect of co-digestion is considered, the minimum impact is 2.2% higher than the total amount of biogas expected under the correspondent theoretical conditions, without the logistics network (in alternatives A5 and A6 under scenario Sc1), and 90.5% higher than the total amount of biogas currently produced in the system (in alternative A3 under scenario Sc1). The effect of co-digestion allows achieving a maximum of 456% higher than the total amount of biogas observed in the system (in alternative A6 under scenario Sc2).

4.3.2.2 Sensitivity analysis

This section includes a sensitivity analysis to verify the impact on the objective function, due to variations in the following parameters: percentage of volatile solids in the sludge caused by potential biogas loss during sludge dehydration, distances between export and destination plants, and amount of sludge available at export plants. The impact on the selection of export plants is also analyzed, according to variations in the value of TRH maximum. Furthermore, the possibility of a centralized system of WWTPs, rather than a decentralized one, is also analyzed. For the sake of simplicity, some of these analyses will only be conducted upon one alternative from the set indicated in Table 4.2.

4.3.2.2.1 Potential biogas loss during sludge dehydration

Since the selection of plants focused mainly on plants that carry out dehydration, it becomes important to run a sensitivity analysis to the amount of VS of the dehydrated sludge generated in said plants. This is because the process of sludge dehydration adds some delay to the transportation of sludge, which may eventually cause the loss of a considerable amount of VS converted into biogas before the sludge reaches the digesters. Since dehydration of non-stabilized sludge is an element of our approach, we provide the values of 15% and 50% of VS that can be lost just to give insights on the magnitude of the impact that could be observed on the value of the objective function. The percentage of VS reduction was included in the modeling of the V_{jk} parameter (equation (9)), and the results are displayed in Table 4.5. From there, one can assume that, despite the eventual loss of biogas, it is essential to consider dehydrated sludge - since in all tested cases the model delivers an economically feasible solution.

Table 4.5 – Results on the effect of biogas loss (translated in terms of the economic benefit) during the dehydration of sludge due to VS reduction.

		Economic benefit (€/year)		% Reduction in the economic benefit (comparison with the situation where no reduction of VS is considered)	
Alternative	Scenario	15% VS reduction	50% VS reduction	15% VS reduction	50% VS reduction
A1	Sc1	66 686	47 304	10.8%	36.7%
	Sc2	80 118	55 188	11.8%	39.2%
A2	Sc1	68 109	48 691	10.8%	36.2%
	Sc2	82 308	57 378	11.5%	38.3%
A3	Sc1	43 691	24 309	15.9%	53.2%
	Sc2	51 319	26 390	17.1%	57.4%
A4	Sc1	57 561	38 179	12.5%	42.0%
	Sc2	68 474	43 581	13.6%	45.0%
A5	Sc1	68 073	48 655	10.8%	36.2%
	Sc2	82 928	57 999	11.4%	38.0%
A6	Sc1	78 694	59 313	9.6%	31.9%
	Sc2	97 492	72 562	9.8%	32.9%

4.3.2.2.2 Distances between export and destination plants

Given that the model could be implemented in broader contexts, the impact of variations in the distances between export and destination plants on the objective function is also tested. The goal is to verify what would be the distance boundaries from which the model could not find a positive optimal solution. In this sense, we present an analysis conducted in such regard. Let us focus on alternative A5 under scenario Sc1, whose results are displayed in the graphic of Figure 4.6. The distance boundaries from which there is no positive optimal solution stay at 13.5 times the real distances of the case study, which corresponds to an average distance of 1155 km between export and destination plants.

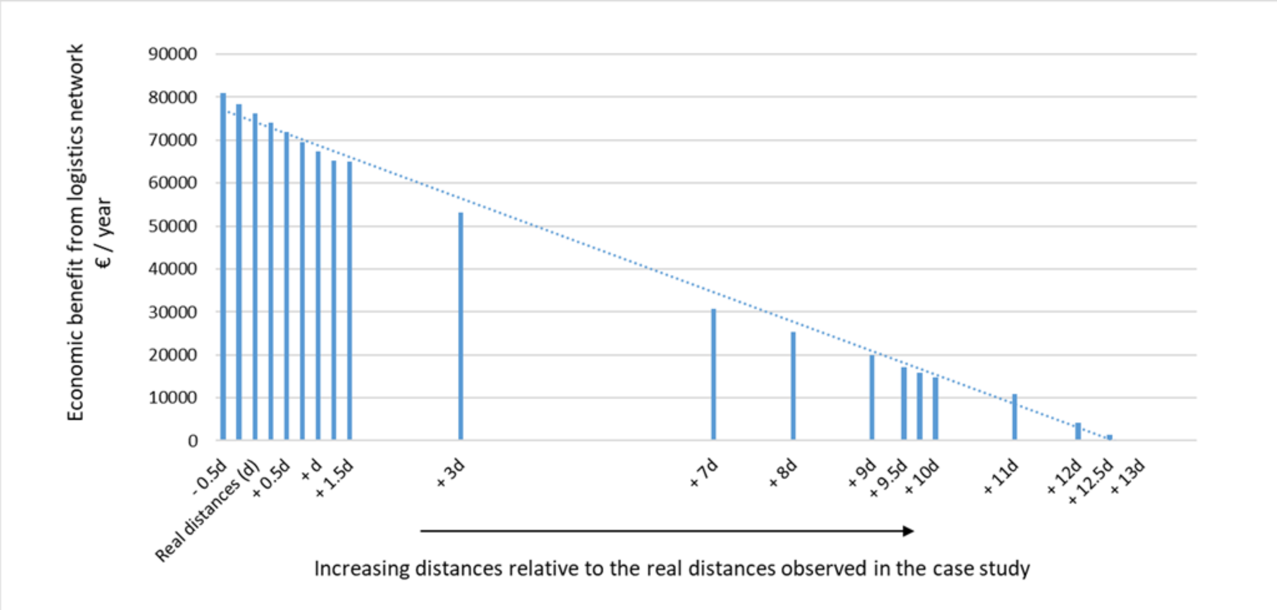


Figure 4.6 *– Logistics network model response to increasing distances.*

4.3.2.2.3 Variations in the value of maximum TRH to understand the selection of export plants

Another topic deserving further interpretation is the selection of export plants. Let us focus on alternative A2 under scenario Sc1. The visual representation of the network configuration for this alternative (Appendix 4B) shows that plants E14 and E16, which are very close to the destination plant D6, were not selected for the network; instead, plants located at a larger distance were preferred (E4, E6 and E8). These plants conduct dehydration, indicating that the model gave priority to the export plants that carry out said process.

In order to understand the selection of export plants better, the results of the alternative A2 under scenario Sc1 indicated in Table 4.4 were further analyzed. Here, it is shown that none of the destination plants selected (D2 and D6) was set to work at full planning capacity, since both destination plants were set to work at a TRH higher than the minimum value imposed (i.e., 20 days). Given this, we decided to test what would be the maximum TRH imposed, so that the destination plants selected could work at a higher planning capacity, and eventually at their full planning capacity (i.e., for a TRH = 20 days). In this case, one could expect the selection of export plants E14 and E16 to fulfill the higher demand. Accordingly, the value for the TRH maximum imposed, leading to a full capacity utilization, was 22 days, with destination plant D6 being the one selected in that case (see Table 4.6). However, the export plants selected were E1, E3, and E4, instead of plants E14 and E16, contrary to what was expected. Plants E1 and E4 conduct dehydration and plant E3 is the export plant without sludge dehydration with the highest value of saved costs. Hence, the amount of saved costs at export plants also plays a significant role in their selection. Furthermore, for the TRH maximum imposed of 22 days (which corresponds to a low flexibility in sludge demand), the optimization of the capacity utilization of the anaerobic digestion leads to a lower economic benefit - through the logistics network - and to a lower amount of biogas produced in the system[7] (see Table 4.6), namely when compared to the values obtained when considering the TRH maximum of 24 days. Therefore, one can also conclude that the flexibility in sludge demand (provided via the interval between the TRH min. and TRH max.) is a positive feature of the logistics model developed, since it enables a wider range of optimization opportunities.

[7] Keep in mind that the values of η and R are assumed invariable.

Table 4.6 – Effect on the organic loading, network configuration and other aspects, due to changes in the HRT_{max} imposed - considering alternative A2 under scenario Sc1.

HRT max imposed (days)	Economic benefit (€/year)	Export plants	Destination plants	Total amount of sludge to be transported (ton/day)	Number of digesters selected	Total amount of sludge at the entry of each digester (ton/day)	HRT (days)	Organic loading rate	Potential biogas produced in the system (m^3/day)
24	76 358	E1, E3, E6	D2	13.2	1	287.1	22	1.88	7 535
		E4, E6, E8, E15, E17	D6	29.3	1	45.8	24	1.75	
23	76 212	E1, E3, E6	D2	11.2	1	285.1	23	1.86	7 535
		E4, E6, E8, E15, E17	D6	31.3	1	47.8	23	1.83	
22	71 029	E1, E3, E4, E6, E8, E15, E17	D6	38.4	1	55	20	2.1	7 488

4.3.2.2.4 Variations in the amount of sludge available at export plants

As previously discussed, the model prioritizes the selection of export plants that conduct sludge dehydration. To test the model's sensitivity to variations in the amount of sludge available at export plants, the destination plant with the highest sludge production in our case study was considered an export plant with sludge dehydration. Also, one can easily verify that the model immediately selects this plant to be part of the logistics system (see Table 4.7). On the contrary, when the same plant is considered without sludge dehydration, the model does not select it. Thus, one can conclude that the model is susceptible to variations in the amount of sludge available at export plants, particularly in the cases where sludge dehydration is carried out. Therefore, if any export plant with a high sludge production is part of a given system of export and destination plants, it is very advisable to invest in sludge dehydration technology in case this technology is not yet adopted.

Table 4.7 – Effect on the economic benefit, network configuration and other aspects, after assuming the largest destination plant of the case study as an export plant with sludge dehydration - considering alternative A2 under scenario Sc1.

HRT max imposed (days)	Economic benefit (€/year)	Export plants	Destination plants	Total amount of sludge to be transported (ton/day)	Number of digesters selected	Total amount of sludge at the entry of each digester (ton/day)	HRT (days)	Organic loading rate	Potential biogas produced in the system (m^3/day)
24	312 367	D2	D3	232.3	2	147.7	20	2.07	7 593
		E1, E2, E3, E4, E6, E10, D2	D4	59.8	1	125	24	1.75	
		E8, E15, E17, D2	D6	29.3	1	45.8	24	1.75	

4.3.2.2.5 Comparison between decentralized and centralized systems

This model can also help decision-makers select whether a centralized system could work in the cases comprehending several destination plants. Accordingly, in an existing system (as the one in this case study), each destination plant is tested as the single destination plant in the system, while all the remaining destination plants are assumed as export plants with sludge dehydration. In addition, it is necessary to consider a set of potential digesters at the destination plant, large enough to receive all sludge produced in all the remaining destination plants (set as export plants). It is important to keep in mind that the model refers only to the operational aspects of the system. Thus, any investment associated with a potential centralized system (such as new digesters) should be subject to a cost-benefit analysis. In this case study, a centralized system for plant D2, and when considering plant D5, would not require any investment in new digesters, since each of these plants has sufficient capacity for all the sludge produced in the entire system. Table 4.8 displays the results concerning the potentially centralized systems of the case study, according to alternative A2 under scenario Sc1. Although the economic benefit is always higher than in the decentralized system (see Table 4.4), it is essential to point out that the value obtained is correlated to the amount of sludge transported and used to produce biogas. Also, and concerning the decentralized system, the majority of this sludge is producing biogas that is not being accounted for in the logistics model. Therefore, the economic benefit per se cannot be used to determine whether a centralized system is better than a decentralized one. The same happens with the potential biogas produced in the system, since it relates to the total amount of sludge available in the system, which does not vary by switching from a decentralized to a centralized management option.

Table 4.8 – Effect on the economic benefit, network configuration and other aspects, taking into account a centralized system - considering alternative A2 under scenario Sc1 and HRT maximum of 24 days.

Economic benefit (€/year)	Export plants	Centralized destination plant	Total amount of sludge to be transported (ton/day)	Number of digesters selected	Total amount of sludge at the entry of each digester (ton/day)	HRT (days)	Organic loading rate	Potential biogas produced in the system (m^3/day)
437 526	E1, E3, E4, E6, E8, E15, E17, D2, D3, D4, D5, D6	D1	604.4	5	128.1	20	2.08	7 532
288 679	E1, E3, E4, E6, E8, E15, E17, D1, D3, D4, D5, D6	D2	366.9	2	320.4	20	2.09	7 536
444 351	E1, E3, E4, E6, E8, E15, E17, D1, D2, D4, D5	D3	536.9	4	150	20	2.1	7 056
419 787	E1, E3, E4, E6, E8, E15, E17, D2, D3, D5, D6	D4	534.8	4	150	20	2.1	7 056
334 621	E1, E3, E4, E6, E8, E15, E17, D1, D2, D3, D4, D6	D5	497.6	3	213.6	21	1.99	7 536
412 560	E1, E3, E4, E6, E8, E15, E17, D2, D3, D4, D5	D6	587.8	11	54.9	20	2.1	7 102

Therefore, Table 4.9 displays a more comprehensive economic analysis respecting both the decentralized system and each centralized system.

Table 4.9 – Economic balance in centralized vs. decentralized systems considering A2 under Sc1.

	Economic benefit logistics network (€/year)	Economic benefit from remaining sludge (€/year)	Total economic benefit (€/year)
Decentralized System	76 358	815 110	891 468
Centralized system (D1)	437 526	49 456	486 982
Centralized system (D2)	288 679	373 154	661 833
Centralized system (D3)	444 351	85 966	530 317
Centralized system (D4)	419 787	88 827	508 614
Centralized system (D5)	334 621	195 089	529 710
Centralized system (D6)	412 560	22 615	435 175

According to Table 4.9, it is possible to discern that, in the absence of other operational aspects (e.g., higher costs associated with dehydrating a higher amount of sludge to be transported, in each new export plant; lower costs in each new export plant with the remaining process of sludge treatment; higher maintenance costs of the CHP system at the central destination plant; costs for transporting the digested sludge to a final destination, and other costs differences), the result is obvious, i.e., the decentralized system is preferable to the centralized one. The main reasons are the additional logistics costs required for any of the centralized systems tested, and the fact that these new export plants (that were destination plants in the decentralized system) do not present significant saved costs concerning aeration reduction time - since they are already operating with a reduced sludge age. Nonetheless, the same rationale could be used in a virtual system of WWTPs (i.e. a system not yet physically implemented, or only partially implemented), including large export plants without AD facilities and several destination plants. In this case, investment costs should also be accounted for, since they can influence the selection of a centralized system for sludge stabilization - through the process of AD - over a decentralized one. Table 4.9 provides some evidence supporting this hypothesis since when considering plant D2 as the centralized destination plant, the total economic benefit is close to that obtained in the decentralized system. Plant D2 is the one with the highest volumetric capacity in their digesters, and the plant with the highest sludge production.

4.4. Conclusion and future developments

This paper comprehends the development of a logistics network model for sewage sludge management. Said logistics model consists of a MILP model that is heavily dependent on the parameters associated with the structural conditions of the system. The model was implemented in a real-world case study of a Portuguese water utility. The fundamental drive was testing the economic feasibility and the magnitude of the impact regarding the differential in the biogas produced in the system. This differential was obtained by comparing the amount of biogas produced under the logistics network approach to the situation observed in the data reporting year and to the theoretical conditions of the system without the logistics network. In this regard, a set of six alternatives (A1 to A6) under two different scenarios (Sc1 and Sc2) was considered. It was proved that the model is economically feasible in all settings: the economic benefit obtained from sludge management under the proposed approach goes from 9.8% up to 23.5% of the current energy costs observed in the set of WWTPs selected to the correspondent networks (A3 under Sc1, and A6 under Sc2, respectively). In addition, the logistics network has a positive impact on the total biogas produced in the system in all settings considered. This positive impact corresponds to a minimum increase of 10.5% compared to the total amount of biogas that is currently produced in the system (in A1 under Sc1, which does not include the effect of co-digestion), and to a maximum increase of 456% (in A6 under Sc2, which includes the effect of co-digestion).

However, the existence of sludge dehydration observed at some export plants is a necessary condition to guarantee the feasibility of the model. This conclusion derives from the selection of all export plants undertaking sludge dehydration as part of the logistics network, regardless of their distance from the destination plant. Furthermore, the model is very sensitive to variations in the amount of sludge supplied exclusively by export plants with that particular feature. Thus, one can determine that any export plant with a high sludge production will represent an excellent contribution for the logistics system, particularly those conducting sludge dehydration, and if there is enough available capacity in the digesters at the destination plants.

Moreover, the significant reduction in the final residue obtained, which is inherent to the anaerobic digestion technology, is a positive environmental element that adds value to the model developed. Furthermore, the economic benefit that the utility obtains with the sludge management after the implementation of the logistics network could improve the quality of service provided to citizens. Besides, the logistics model contributes to improving the recovery of energy from a renewable source, and to a more efficient use of electricity - being these matters supported and encouraged by International Authorities, such as the European Commission, and others. Regarding the transportation component, the negative impact associated with GHG emissions is partially neutralized by the reduction of the electricity consumption at the export plants (which depends on the mix of energy sources of the country) and by the increase in biogas production. Furthermore, it can be minimized or even eliminated through the investment in biogas trucks.

The main managerial implication of the logistics approach proposed in this paper consists in its adoption as a tool for improving sludge treatment (or other organic waste) management in a system including several facilities. Furthermore, its practical implementation requires some initial investment to introduce slight operational changes in the facilities conducting anaerobic digestion that are part of the network, namely to enable the mixture of sludge before entering the digesters. The model can also be used as a tool for strategic asset management, since it supports the strategic choices by decision-makers that involve different types of facilities and their interconnection.

As future developments of the work presented here, co-digestion is one aspect that deserves further attention, particularly in terms of its direct integration in the mathematical model formulation (in this paper, the modeling of parameters only considered the effects of co-digestion). The development of the operational model for routing and scheduling is another topic deserving future attention. Furthermore, the decision regarding the final disposal of sludge, after the anaerobic digestion process, can be integrated into the mathematical model formulation. Additionally, developing a similar model under uncertainty could also be an option, in order to obtain a robust approach regarding the data used to model some of the parameters, particularly the amount of sludge produced in all WWTPs. However, this does not invalidate that the deterministic model designed could be run periodically to accommodate future technical and technological changes that may require adjustments in the strategies previously defined. Finally, the model for the logistics network developed in this paper can be perceived as a piece of a future puzzle representing a complete integrated system for sewage sludge treatment management, at a national level. This integrated system should also include co-digestion to maximize biogas production, thus improving the energy self-sustainability of WWTPs and the organic waste collection and treatment. Likewise, a process system engineering perspective should also be adopted to optimize the functioning of co-digestion within the planned integrated system.

Acknowledgments

The authors wish to acknowledge the financial support by Project "ETARs + Sustentáveis" (More Sustainable WWTPs). This project is funded by the Portuguese water utility Águas do Centro Litoral (AdCL). The authors are particularly grateful to all the support provided by Tiago Braga and Paulo Leitão from AdCL. The first author also wishes to acknowledge the financial support from FCT - Fundação para a Ciência e a Tecnologia (Portuguese national funding agency for science, research and technology), allocated through the Grant PD/BD/142815/2018 respecting the Doctoral Program on Sustainable Energy Systems by MIT-Portugal.

Appendix 4A

Computation of the amount of co-substrate to be blended in the sludge

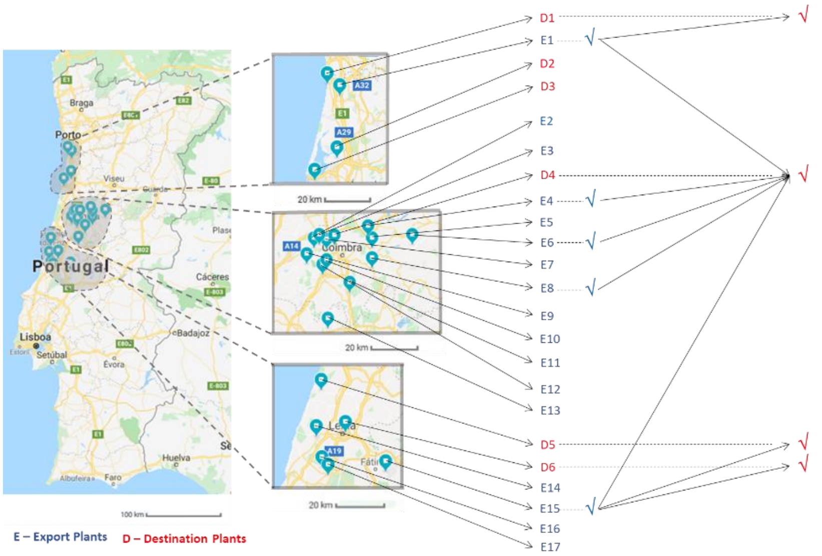
For the purpose of calculating the amount of co-substrate, x, to be blended in the sludge at each destination plant, and according to the objectives defined in section 4.3.2, we consider the example of a co-substrate with 25%S, from which 90% are VS. In order to increase 10% in the ratio VS over TS of the sludge (originally at 70% VS/TS) and to leverage the percentage of total solids in the sludge from, for example, 2% to 3%, it is necessary to add 4.98 ton of organic co-substrate to every 100 ton of sludge (see Figure 4A.1). The rationale used for the computation of x follows the modeling approach explained in section 4.2.2.2.6 to calculate the total amount of volatile solids that characterizes the sludge that is going through anaerobic degradation for biogas production. We recognize that this rationale may not lead to the optimum ratio between co-substrate and sludge according to some studies (c.f., Prabhu and Mutnuri, 2016), since we are ignoring other specificities of the mixture of co-substrate and sewage sludge. Nonetheless, we believe that it is possible to carefully select the type of co-substrate more suitable to be blended with the sewage sludge, in order to fulfill both the carbon to nitrogen (C:N) ratio, the increase in VS and the % of solids in the sludge, and to optimize the amount of biogas that can be produced through co-digestion - even if there is the need for an experimental stage.

$$\overbrace{(100+x)\times 3\times 80}^{\text{Intended mixture}} = \overbrace{100\times 2\times 70}^{\text{Existing sludge}} + \overbrace{x\times 25\times 90}^{\text{Co-substrate}} \Rightarrow \boldsymbol{x = 4.98}\ \text{ton } \textbf{of co-substrate}$$

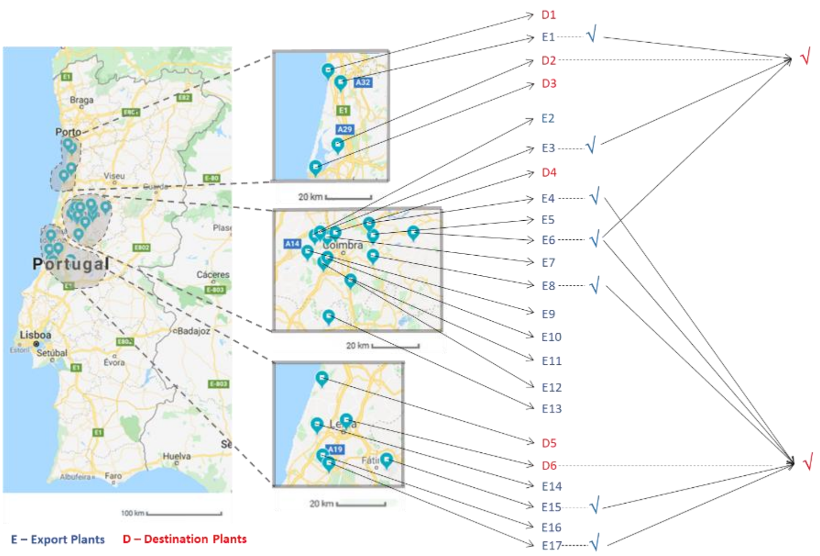
Figure 4A.1 *– Rationale for the blending quantities when considering co-digestion.*

Appendix 4B

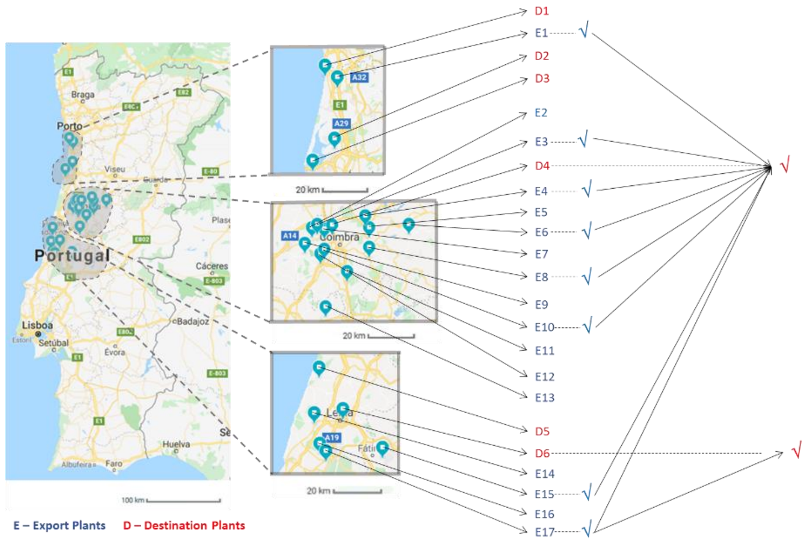
Schematic representation of the network design for alternatives A1-A6, scenario 1



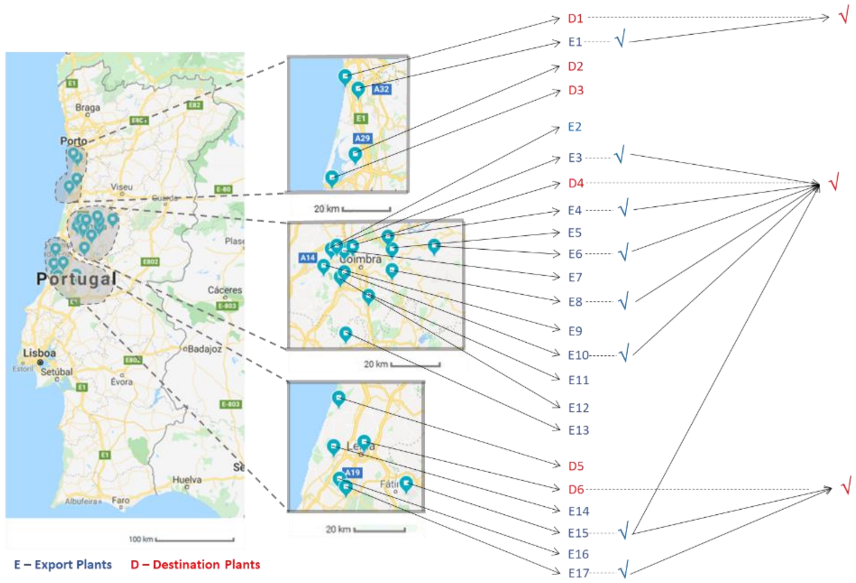
Alternative A1



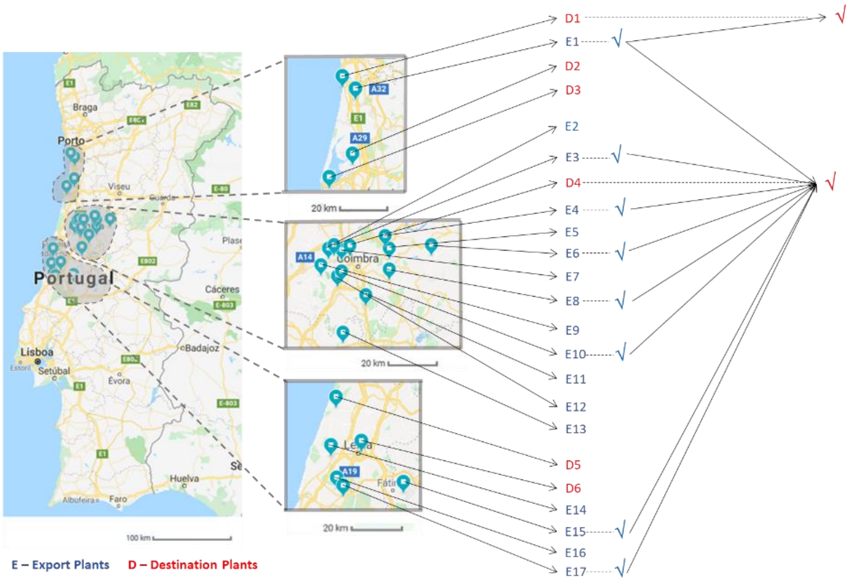
Alternative A2



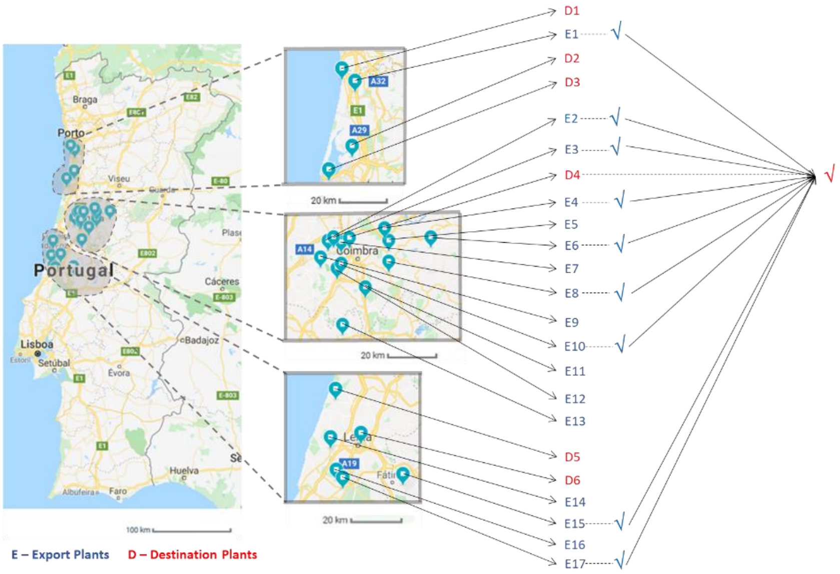
Alternative A3



Alternative A4



Alternative A5



Alternative A6

References

Ahmadi-Javid, A., Amiri, E. and Meskar, M. (2018). A profit-maximization location-routing-pricing problem: a branch-and-price algorithm. *European Journal of Operational Research*, 271, 866-881.

Andreoli, C. V., Sperling, M. and Fernandes, F. (Eds) (2007). Sludge Treatment and Disposal. IWA Publishing.

Balaman, S, Y. and Selim, H. (2014). A network design model for biomass to energy supply chains with anaerobic digestion systems. *Applied Energy*, 130, 289-304.

Barbosa-Póvoa, A. P., Silva, C. and Carvalho, A. (2018). Opportunities and challenges in sustainable supply chain: an operations research perspective. *European Journal of Operational Research*, 268, 399-431.

Beltramo, T., Klocke, M. and Hitzmann, B. (2019). Prediction of the biogas production using GA and ACO input features selection method for ANN model. *Information Processing in Agriculture*, Article in Press.

Chae, K. and Kang, J. (2013). Estimating the energy independence of a municipal wastewater treatment plant incorporating green energy resources. *Energy Conversion and Management*, 75, 664-672.

Entezaminia, A., Heydari, M. and Rahmani, D. (2016). A multi-objective model for multi-product multi-site aggregate production planning in a green supply chain: considering collection and recycling center. *Journal of Manufacturing Systems*, 40, 63-75.

Forsberg, J. (2014). Biogas grid in Mälardalen valley. SGC Rapport 2014:300.

Galvez, D., Rakotondranaivo, A. and Morel, L. (2015). Reverse logistics network design for a biogas plant: an approach based on MILP optimization and Analytical Hierarchical Process (AHP). *Journal of Manufacturing Systems*, 37, 616-623.

Ge, H., Batstone, D. J. and Keller, J. (2013). Operating aerobic wastewater treatment at very short sludge ages enables treatment and energy recovery through anaerobic sludge digestion. *Water Research*, 47, 6546-6557.

Gikas, P. (2017). Towards energy positive wastewater treatment plants. *Journal of Environmental Management*, 203, 621-629.

Görgüner, M., Aksoy, A. and Sanin, F. D. (2014). A transport cost-based optimization for recycling of municipal sludge through application on arable lands. *Resources, Conservation and Recycling*, 94, 146-156.

Gretzschel, O., Schmitt, T. G., Hansen, J., Siekmann, K. and Jakob, J. (2014). Sludge digestion instead of aerobic stabilization – a cost benefit analysis based on experiences in Germany. *Water Science & Technology*, 69(2), 430-437.

Gu, Y., Li, Y., Li, X., Luo, P., Wang, H., Robinson, Z. P., Wang, X., Wu, J. and Li, F. (2017). The feasibility and challenges of energy self-sufficient wastewater treatment plants. *Applied Energy*, 204, 1463-1475.

Hagos, K., Zong, J., Li, D., Liu, C. and Lu, X. (2017). Anaerobic co-digestion process for biogas production: progress, challenges and perspectives. *Renewable and Sustainable Energy Reviews*, 76, 1485-1496.

Hallaji, S. M., Kuroshkarim, M. and Moussavi, S. P. (2019). Enhancing methane production using anaerobic co-digestion of waste activated sludge with combined fruit waste and cheese whey. *BMC Biotechnology*, 19(19).

Hernández-Sancho, F., Molinos-Senante, M. and Sala-Garrido, R. (2011). Energy efficiency in Spanish wastewater treatment plants: a non-radial DEA approach. *Science of the Total Environment*, 409, 2693-2699.

Hobus, I., Kolisch, G., and Hansen, J. (2010). Increasing the energy efficiency of sludge stabilization by an interconnected operational approach. International Conference “Water and Energy”, Amsterdam.

Iakovou, E., Karagiannidis, A., Vlachos, D., Toka, A. and Malamakis, A. (2010). Waste biomass-to-energy supply chain management: a critical synthesis. *Waste Management*, 30, 1860-1870.

IEA (International Energy Agency) (2018). World Energy Outlook.

Kwasny, J. and Balcerzak, W. (2017). Production logistics and participation of biogas in obtaining primary energy in Poland. *Energy & Environment*, 28(4), 425-436.

Lashine, S. H., Fattouh, M. and Issa, A. (2006). Location/allocation and routing decisions in supply chain network design. *Journal of Modelling in Management*, 1(2), 173-183.

Liu, G. and Wang, J. (2015). Modeling effects of DO and SRT on activated sludge decay and production. *Water Research*, 80, 169-178.

Madsen, M., Holm-Nielsen, J. B. and Esbensen, K. H. (2011). Monitoring of anaerobic digestion processes: a review perspective. *Renewable and Sustainable Energy Reviews*, 15, 3141-3155.

Mao, C., Feng, Y., F. Wang, X. and Ren, G. (2015). Review on research achievements of biogas from anaerobic digestion. *Renewable and Sustainable Energy Reviews*, 45, 540-555.

Mitchell, G. and Beasley, J. E. (2011). Optimization of sludge treatment and transport. *Journal of the Operational Research Society*, 62(6), 939-948.

Molinos-Senante, M., Hernández-Sancho, F. and Sala-Garrido, R. (2010). Economic feasibility study for wastewater treatment: a cost-benefit analysis. *Science of the Total Environment*, 408, 4396-4402.

Mota, B., Gomes, M.I., Carvalho, A. and Barbosa-Povoa, A.P. (2015). Towards supply chain sustainability: economic, environmental and social design and planning. *Journal of Cleaner Production*, 105, 14-27.

Mota, B., Gomes, M. I., Carvalho, A. and Barbosa-Povoa, A. P. (2018). Sustainable supply chains: An integrated modeling approach under uncertainty. *Omega*, 77, 32-57.

Özceylan, E. and Paksoy, T. (2013). A mixed integer programming model for a closed-loop supply-chain network. *International Journal of Production Research*, 51(3), 718-734.

Pan, S., Du, M. A., Huang, I., Liu, I. and Chang, E. (2015). Strategies on implementation of waste-to-energy (WTE) supply chain for circular economy system: a review. *Journal of Cleaner Production*, 108, 409-421.

Pavan, P., Bolzonella, D., Battistoni, E. and Cecchi, F. (2007). Anaerobic co-digestion of sludge with other organic wastes in small wastewater treatment plants: an economic considerations evaluation. *Water Science & Technology*, 66(10), 45-53.

Pourjavad, E. and Mayorga, R. V. (2019). An optimization model for network design of a closed-loop supply chain: a study for a glass manufacturing industry. *International Journal of Management Science and Engineering Management*, 14(3), 169-179.

Prabhu, M.S. and Mutnuri, S. (2016). Anaerobic co-digestion of sewage sludge and food waste. *Waste Management & Research*, 34(4) 307-315.

Premier, G. C., Dinsdale, R., Guwy, A. J., Hawkes, F. R., Hawkes, D. L. and Wilcox, S. J. (1999). A comparison of the ability of black box and neural network models of ARX structure to represent a fluidized bed anaerobic digestion process. *Water Research*, 33(4), 1027-1037.

Rabbani, M., Hosseini-Mokhallesun, S. A. A., Ordibazar, H. A. and Farrokhi-Asl, H. (2018). A hybrid robust possibilistic approach for a sustainable supply chain location-allocation network design. *International Journal of Systems Science: Operations & Logistics*, DOI: 10.1080/23302674.2018.1506061

Ramudhin, A., Chaabane, A. and Paquet, M. (2010). Carbon market sensitive sustainable supply chain network design. *International Journal of Management Science and Engineering Management*, 5(1), 30-38.

Saffar, M. M., G. Shakouri, H. and Razmi, J. (2014). A new bi-objective integer linear programming for designing a supply chain considering CO2 emission. *Uncertain Supply Chain Management*, 2, 275-292.

Santibañez-Aguilar, J.E., González-Campos, J.B., Ponce-Ortega, J.M., Serna-González, M. and El-Halwagi, M.M. (2014). Optimal planning and site selection for distributed multiproduct biorefineries involving economic, environmental and social objectives. *Journal of Cleaner Production*, 65, 270-294.

Shah, M. S., Halder, P.K., Shamsuzzaman, A. S. M., Hossain, M. S., Pal, S. K. and Sarker, E. (2017). Perspectives of biogas conversion into Bio-CNG for automobile fuel in Bangladesh. *Journal of Renewable Energy*, 2017, 14 pp.

Wett, B., Buchauer, K. and Fimml, C. (2007). Energy self-sufficiency as a feasible concept for wastewater treatment systems. IWA Leading Edge Technology Conference: Singapore: Asian Water, pp. 22-25.

Zhu, X., Li, X., Yao, Q. and Chen, Y. (2011). Challenges and models in supporting logistics system design for dedicated-biomass-based bioenergy industry. *Bioresource Technology*, 102, 1344-1351.

Chapter 5 –

The Use of Order-*m* for a Benchmarking Analysis of a Wastewater Treatment Plants’ Problematic Sample

Chapter 5

The use of order-*m* for a benchmarking analysis of a wastewater treatment plants' problematic sample

This chapter refers to an adaptation of the paper that is currently (i.e., by the time the thesis is being concluded) under review by the journal *Annals of Operations Research*. The main differences between both versions concern the title, the abstract, and the introduction section.

Overarching research question to be explored with this paper:

RQ3: In what ways can a quantitative benchmarking assessment framework be explored from a water company perspective given the very particular characteristics of their assets' design and operation, with the aim of enhancements in resource consumption reduction and increased sustainability of the company?

Authors:

Alda A. Henriques, Milton Fontes, Ana S. Camanho, Giovanna D'Inverno, Pedro Amorim, J. Gabriel Silva

Thesis' author contribution:

Conceptualization (collaboration), data curation (collaboration), software usage (except the model implementation), formal analysis (collaboration), investigation, writing – original draft (except sections 5.2.2 and 5.2.3), writing – review and editing (collaboration), funding acquisition (collaboration).

The use of order-*m* for a benchmarking analysis of a wastewater treatment plants' problematic sample

Abstract

This paper explores an original use of the order-*m* method to support performance evaluation in real-world assessments, which includes best-practices identification and target setting. We use the example of a Portuguese sample of wastewater treatment plants to demonstrate that, in some specific cases, the order-*m* method can bring more adequate and useful insights as a performance benchmarking tool than a conventional data envelopment analysis application. Particularly, order-*m* delivers information to organizations' stakeholders about their units' relative performance so that practical improvements can be easily sought and implemented in the case their units' samples are small and with a high variability both in terms of the magnitude of the selected indicators and the units' contextual conditions. In this context, we also propose the term of problematic sample to use in cases as the one explored in this paper. We highlight that the application of the order-*m* method to problematic samples can be further useful when there are potential outliers whose removable from the sample is not desirable, and whose presence may be typical in such type of samples.

Keywords: Data Envelopment Analysis, Order-*m* method, Problematic Samples, Wastewater Treatment Plants

5.1. Introduction

In the context of nonparametric efficiency analysis, such as Data Envelopment Analysis (DEA) and Free Disposal Hull (FDH), a set of individual entities is comparatively assessed to search for potential improvements in their performance. Each individual entity, or decision-making unit (DMU), conducts a given productive activity that can be seen as the conversion of inputs into outputs. In order to conduct a fair efficiency assessment, the sample of DMUs used in the analysis must be homogeneous or, in other words, similar in a number of ways (Dyson et al., 2001). First, all DMUs are required to develop similar activities to obtain similar products or services (Pitfall 3.1, ibidem). However, although all DMUs may be converting the same type of inputs into the same type of outputs, in practice two additional issues can threaten the homogeneity assumption, namely the presence of atypical observations in the sample and different environmental conditions that the DMUs can face (Pitfall 3.2, ibidem). This is an issue that often affects real-world samples, where high variability in terms of input/output levels and exogenous conditions can be found, hence the name "problematic samples". The performance evaluation conducted on this type of samples might thus result in biased measures of performance. As an alternative to not conducting any performance analysis at all because some requirements are not perfectly met, or to implement a conventional framework with deficient results, the selection of an appropriate method may allow to get more and better insights upon the performance of the DMUs under assessment and to obtain more realistic results, which will allow extracting valuable information concerning those particular DMUs.

This paper conducts an exploratory analysis of a performance evaluation of DMUs in a problematic sample to show, from a novel perspective, how nonparametric efficiency evaluation tools can provide valuable information to assist companies or decision-makers in their management choices regarding a set of units in such type of samples. A problematic sample is defined as a sample with a small number of DMUs, and high variability in the indicators' values and/or environmental conditions. The meaning of "small" in the definition of problematic samples should be interpreted in relative terms, through a comparison between the number of DMUs and the inputs, outputs, and exogenous factors considered in the performance assessment. The curse of dimensionality in DEA is a well-known problem in the literature (e.g., see Charles et al., 2019). Even considering the works of Charles et al., 2019 and Lee and Cae (2020), for example, the curse of dimensionality still gives margin to challenges in real-world DEA benchmarking studies. Regarding only the case of small samples, that attribute should either not constitute an impediment to the performance evaluation and, depending on the purpose of the assessment, there are more classical solutions, such as the use of weight restrictions, that can minimize the inconvenient of obtaining an excessive number of DMUs identified as efficient (Cook et al., 2014). However, the variability in the magnitude of the indicators' values constitute an obstacle to a fair and successful performance assessment. Such variability can partially derive from the presence of atypical observations. In a more concrete way, a DMU may differ to a large extent from the rest of the sample mainly for two reasons: there may be errors in the data, or the observations may be potentially correct, but highly atypical. In a benchmarking setting, the DMUs located on the frontier should correspond to replicable DMU-level performance for the same set of circumstances. Therefore, DMUs with exceptionally high relative performance (i.e., outliers) should be removed from the sample for precautionary reasons. The efficiency literature proposes several outlier detection procedures to guide the removal of suspected units, attempting to ensure that the frontier is well populated and not too distant from the bulk of DMUs in the analysis (e.g., Anderson and Petersen, 1993; Wilson, 1995; Wilson, 2010; De Witte and Marques, 2010b). However, for analyses involving DMUs in a problematic sample, especially due to the limited number of DMUs available, removing potential outliers cannot be considered as a solution. The use of a partial frontier method instead of a full frontier method is more suited to mitigate the impact of atypical observations without removing them from the assessment (Cazals et al., 2002; Daraio and Simar, 2005 & 2007).

Another key challenge in the performance measurement of DMUs in a problematic sample is accounting for different operating contexts. To address the impact of contextual conditions on DMUs' efficiency levels, one-stage and two-stage approaches have been considered (see Avkiran and Rowlands, 2008 for a literature review). However, there is an unsettled debate in the literature regarding the most appropriate specification of these types of approaches. To overcome some of the limitations of the two-stage approaches, a conditional approach (Daraio and Simar, 2005 & 2007) has been proposed to circumvent the restrictive assumptions underlying efficiency assessments with heterogeneous environmental conditions (see De Witte and Marques, 2010a, for a critical discussion of different methodologies to incorporate heterogeneity).

In face of the previous discussion, we found that the nonparametric order-*m* method (Cazals et al., 2002; Daraio and Simar, 2005 & 2007) fits all the requirements to attain the objectives of this paper. Firstly, alternatively to a full-frontier method such as DEA, the partial-frontier order-*m* method is demonstrated to surmount the curse of dimensionality problem (Wheelock and Wilson, 2003), which is particularly useful when dealing with small samples and with the purpose of deriving as much insights regarding potential improvements as possible. Additionally, the order-*m* method also allows to mitigate the effect of potential outliers, either occurring due to errors in the data or due to real atypical observations, without the need to remove any DMU, which again is particularly useful when dealing with small samples. Additionally, this partial frontier method, considers a primary condition in its mathematical formulation that limits the comparison process of each DMU to those that are more suitable regarding the magnitude of their indicators. For this reason, the order-*m* method is also adequate to mitigate the variability in the magnitude of the indicators' values. Moreover, regarding the variability that occurs due to DMUs facing different contextual conditions, a conditional order-*m* method (Daraio and Simar, 2007) may provide additional insights into the impact caused by those contextual conditions in the performance of the DMUs. Nonetheless, more often than not, some structural contextual factors are, by nature, setting some of the boundaries for the magnitude observed in the DMUs' indicators. Therefore, a conventional (i.e., unconditional) order-*m* analysis, inherently, and indifferently, already accounts for some variability that is originated by those uncontrollable factors.

This paper identifies a suitable framework to be used for internal benchmarking exercises and thus to support real-world performance assessments. For this reason, it has a clear managerial implication. Accordingly, the ultimate objective is to provide information to guide practical improvements, particularly in the case where the DMUs to be assessed are in a problematic sample, or, in other words, for samples that at the first glance might discourage or invalidate the performance evaluation. The suggested framework focuses on a clear identification of potential reductions in resource consumption or reallocation of resources, rather than on a more general analysis of relative efficiency levels that would be suitable for policy design at more aggregated levels, and for which the evaluated units are commonly in samples with very low variability. By applying this toolbox, we illustrate how relevant is the role of the peers and targets for decision support as a complement to the standard results provided by the order-*m* method. We also exemplify how informative is the evidence provided by the conditional approach to guide more thoroughly comparative assessments. As a result, in real world case studies, intervening at a very detailed level can improve the effectiveness of the performance assessment exercise, making processes evolve towards more sustainable and productive systems at social, economic, and environmental levels (Moldan et al., 2012).

Although the proposed performance analysis framework can be applied to several areas, we explore a particular case study within the wastewater treatment industry. The efficient use of resources in the productive activity of cleansing the wastewater is a topic that deserves increased attention by the scientific community and practitioners. Since we are directing an exploratory study of performance evaluation of DMUs in a problematic sample, it is appropriate to highlight the concern expressed by Longo et al. (2018) regarding this type of industry, and more specifically the case of wastewater treatment plants (WWTPs). This constitutes an example of an application field where the homogeneity assumptions may be critical, although often disregarded in the literature.

The wastewater treatment industry plays a significant role in sustainable development worldwide. An improved functioning of the WWTPs would also support at least two goals of the Sustainable Development Agenda defined by the United Nations (UN General Assembly, 2015): Goal 6 ("By 2030, improve water quality 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", 6.3) and Goal 7 ("By 2030, double the global rate of improvement in energy efficiency", 7.3).

Our case study corresponds to a sample of 41 Portuguese WWTPs managed by the company Águas do Centro Litoral (AdCL). This sample presents remarkable variability in the input and output indicators and faces different operating characteristics. Therefore, it represents a good test bed to provide concrete examples of the advocated advantages of using the order-*m* method for performance measurement of DMUs in a problematic sample.

Two aspects are simultaneously considered to better assist decision-making: (1) The measurement of efficiency should be as fair as possible so as to provide the real picture of the relative performance behavior of the sample under evaluation. When this aspect is taken into account, a better prioritization of actions and eventual investments can be achieved. (2) The information on peers and targets for improvements should be provided. This allows plants to benefit from information sharing to guide performance improvements through adjustments of technical and operational aspects. The identification of attainable targets representing best practices that are aligned with management strategy to guide organizational improvements, is a practice commonly attached to the use of DEA, and it is also a topic that is gaining momentum in the literature in recent years (e.g., see Ruiz and Sirvent, 2019). DEA has also been widely used as the preferable methodology to assess the performance of wastewater treatment plants (WWTPs) (e.g., Hernández-Sancho and Sala-Garrido, 2009). Moreover, the majority of studies within the context of the wastewater treatment industry recognize the relevance of the environmental conditions faced by WWTPs on their performance (e.g., Fuentes et al., 2015). Having this in mind, and also for the reasons pointed out above, we adopt a robust DEA estimator as developed by Daraio and Simar (2007), which is an extension of the order-*m* method introduced by Cazals et al. (2002).

The proposed framework provides reliable and unbiased performance assessment results to DMUs in a problematic sample. In addition, as it allows a fair comparison of WWTPs along with the identification of their peers, the proposed framework helps in achieving deeper insights into the plants' activity. This directly allows the interested stakeholders, namely company's and plants' decision-makers, to receive tangible management solutions that can be practically implemented. Moreover, ensuring improvements in asset management is likely to have an impact beyond the direct benefits arising from organizing the WWTPs production process more efficiently and consequently saving resources at company level.

The contribution of this paper to the Operations Research and Management Science literature is twofold. First, it contributes to the nonparametric efficiency literature mostly oriented to internal managerial practices (Camanho and Dyson, 1999; Barros, 2006 & 2008; Camanho et al., 2009; Vaz et al., 2010; Rohacova, 2015) by suggesting the use of the order-*m* method with an operative point of view. A critical feature in internal benchmarking studies concerns the improvement of operational practices through learning from peers. In a recent paper by Lavigne et al. (2019), the authors emphasize the importance of identifying the peers underlying a robust conditional Benefit-of-the-doubt approach applied to waste management. Therefore, our paper extends this research line in two main directions. First, Lavigne et al. (2019) focused on the construction of a composite indicator to aggregate outputs, while we look at the transformation of resources into outputs in performance assessments involving problematic samples. Furthermore, rather than just focusing on estimating peers, we also estimate the targets for the inefficient DMUs to provide guidelines for the amount of resources that can be potentially saved. We also emphasize the advantages of using the order-*m* method rather than the DEA method when assessing the performance of DMUs in a problematic sample. This is done through comparatively exploring the results obtained from both methods to the DMUs in a particular example of problematic sample. We also emphasize the technical issues that must be considered when applying the order-*m* method, both in its conditional and unconditional forms, to the case of such type of samples.

The second contribution of this paper is related to its empirical application, specifically to the literature on WWTPs efficiency evaluation. Despite the increasing prominence of this topic in recent years (see D'Inverno et al., 2018), to the best of our knowledge only a few papers used a conditional approach to evaluate wastewater treatment plants efficiency (e.g., Fuentes et al., 2015; Guerrini et al., 2016). Furthermore, to the best of our knowledge, our study represents a development in relation to these approaches by proposing a novel WWTP output specification to fully capture the real output attained by each plant during the treatment process. This requires considering the rate of pollutant removal (which accounts for the concentration of the pollutant in the influent) and the total volume of wastewater treated. It is also the first study adopting an internal managerial perspective of WWTPs efficiency using Portuguese data.

The remainder of the paper is organized as follows. Section 5.2 describes the methodological toolbox proposed for the efficiency analysis with problematic samples. Section 5.3 describes the case study used to demonstrate the usefulness of the framework proposed. Section 5.4 presents the results obtained and discusses their managerial implications. Finally, section 5.5 concludes, highlighting the major findings and limitations of this study.

5.2. Methodology

In a first moment, this section introduces the nonparametric efficiency tools that are commonly used in performance evaluations. In a second moment, we will explain how these tools can be used in empirical applications in cases where DMUs under assessment are in problematic samples.

One of the main advantages of nonparametric efficiency methods is that the efficient frontier is derived directly from observations. Therefore, there is no need for prior specification of its functional form. Depending on the assumptions characterizing the production technology, different nonparametric efficiency models can be considered, such as Data Envelopment Analysis, Free Disposal Hull, Directional Distance Function, among others.

To keep a concise exposition, we focus the discussion on Data Envelopment Analysis (DEA) models, with an input orientation setting under Variable Returns to Scale (VRS). However, the approach can be straightforwardly extended to other nonparametric techniques, such as the ones mentioned above, as well as to different orientations or returns to scale assumptions.

5.2.1 DEA: A baseline nonparametric efficiency model

Data Envelopment Analysis (DEA) is a nonparametric method originated from the seminal work of Farrell (1957), who proposed the evaluation of relative efficiency as the radial distance to a production frontier estimated directly from empirical observations. Farrell (1957)'s concepts were operationalized for the first time using linear programming by Charnes et al. (1978).

DEA compares the efficiency of a relatively homogeneous set of DMUs in their use of multiple resources (inputs) to produce multiple outcomes (outputs). It derives a single summary measure of efficiency for each unit, which is based on a comparison with other units in the sample. DEA identifies two groups of units: radially efficient and inefficient. The DMUs located on the frontier are considered examples of best practices (i.e., the benchmarks) and obtain an efficiency score equal to one. The frontier is spanned by the efficient DMUs of the sample. For the inefficient DMUs, the magnitude of their inefficiency is evaluated as the distance to the frontier. For an input-oriented assessment, it represents the proportional reduction of inputs that is required to reach the frontier. One of the advantages of DEA is to allow each DMU to select its own weighting system for the performance evaluation, recurring to optimization. This allows emphasizing the strengths of each DMU.

The formulation of the input-oriented DEA model with Variable Returns to Scale (VRS) is shown in (5.1).

$$\min_{\theta_k, \lambda_j, s_i^-, s_r^+} \theta_k - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right) \quad \textbf{(5.1)}$$

Subject to

$$\theta_k\, x_{ik} - \sum_{j=1}^{n} \lambda_j x_{ij} - s_i^- = 0 \qquad i = 1, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{rk} \qquad r = 1, \ldots, s$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, s_i^-, s_r^+ \geq 0, \qquad \forall j, i, r$$

In model (5.1), x_{ij} corresponds to the value of input i $(i = 1, \ldots, m)$ and y_{rj} corresponds to the value of output r $(r = 1, \ldots, s)$ observed for DMU j $(j = 1, \ldots, n)$. The decision variables of the model (5.1) are θ_k , λ_j , s_i^- and s_r^+. θ_k corresponds to the radial efficiency score of the unit under assessment, represented by index k. λ_j is an intensity variable, which assumes a positive value whenever a unit j is used as peer for unit k under assessment. The VRS assumption is represented by the restriction imposing that the sum of λ_j, for $(j = 1, \ldots, n)$, is equal to one. s_i^- and s_r^+ are slack variables, representing potential non-radial adjustments to the input and output levels of the unit under assessment, beyond the proportional reduction to the input levels represented by the radial efficiency score. ε is an infinitesimal used to ensure that slacks are not ignored in the efficiency assessment.

When analyzing the solution to model (5.1), a DMU k under evaluation is fully efficient if and only if θ_k =1 and all slacks are equal to zero ($s_{ik}^- = s_{rk}^+ = 0, \forall i, r$). In other words, a unit is fully efficient if it is not possible to decrease the consumption of any of its inputs without also decreasing the amount of at least one of the outputs or increasing the consumption of another input. As by-products of the assessment, for inefficient DMUs it is also possible to obtain peers and targets. The input and output targets for unit k are obtained as shown in expressions (5.2) and (5.3). The values of the decision variables obtained at the optimal solution to model (5.1) are signaled with an asterisk.

$$x_{ik}^T = \sum_{j=1}^{n} \lambda_{jk}^* x_{ij} = \theta_k{}^* x_{ik} - s_{ik}^{*-} \qquad i = 1, \ldots, m \quad \textbf{(5.2)}$$

$$y_{rk}^T = \sum_{j=1}^{n} \lambda_{jk}^* y_{rj} = y_{rk} + s_{rk}^{*+} \qquad r = 1, \ldots, s \quad \textbf{(5.3)}$$

5.2.2 The order-*m* approach: mitigating the impact of atypical observations

The nonparametric nature of DEA is an important feature in situations where the production process is complex and the production function is unknown. However, its deterministic nature makes it very sensitive to atypical observations, as any unit can potentially contribute to shape the frontier, hence the name "full frontier". As a result, the presence of outliers and extreme values might lead to downwardly biased estimates (Fusco et al., 2019), as DMUs that differ to a large extent from the rest of the sample may lead to the shift of the frontier and have a significant impact on the evaluation of other DMUs. Several methods can be employed to detect outliers that can potentially affect the position of the frontier, which are candidates to be removed from the sample (e.g., super-efficiency of Anderson and Petersen, 1993 or the approach proposed by Wilson, 1995). From an applied perspective of internal benchmarking, the outlying units might be of great interest to the stakeholders to the extent that they could be potentially either the best or the worst performing units. Especially in a context of small samples, each piece of information is crucial for a sound managerial decision and the removal of observations should be avoided (De Witte and Marques, 2010b). To account for these issues and at the same time to mitigate the impact of atypical observations, the literature has proposed the use of "partial" or "robust" frontiers, which are the structural features of the order-*m* (Cazals et al., 2002; Daraio and Simar, 2005 & 2007) and order-α methods (Aragon et al., 2005; Daouia and Simar, 2007). In compliance with this latter stream of literature, this paper adopts the order-*m* method for the evaluation of performance as developed in Daraio and Simar, 2007. To keep the discussion focused on an operational perspective, we only present the basic idea and the main intuition of this robust approach. For an exhaustive exposition, we refer to Daraio and Simar (2007).

Computationally, in the full frontier estimation the linear programming model (5.1) is solved once for each unit under assessment. Instead, in the partial frontier estimation the linear programming model (5.1) is computed B times for each unit according to a Monte-Carlo procedure, where B is a large number (for more details, we refer the reader to Daraio and Simar, 2007). For an input-oriented assessment, in each of these B iterations, m units are drawn at random with replacement among those producing at least the same level of output as the unit under evaluation. By doing this, the impact of the atypical observations is remarkably mitigated and the unit under evaluation is compared with a less extreme benchmark. The order-*m* efficiency score $\hat{\theta}_k^m$ is then obtained as the average of the efficiency scores $\theta_k^{b,m}$ computed at each b-th iteration ($b = 1, \dots, B$), as shown in equation (5.4).

$$\hat{\theta}_k^m = \frac{1}{B}\sum_{b=1}^{B} \theta_k^{b,m} \tag{5.4}$$

Due to the subsampling, the evaluated unit might not be among the drawn m units. In this case, it does not constitute its own reference set and it may be located above the frontier. Accordingly, an efficiency score greater than one identifies a 'super-efficient unit' that is more efficient than the average m units in its reference set. m can have a twofold interpretation (Daraio and Simar, 2005; De Witte and Schiltz, 2018). First, it can represent the number of competitors producing greater or same output levels. Second, it can be seen as a trimming value to assert the robustness of the analysis. Among others, Cazals et al. (2002) and Rogge and De Jaeger (2013) note that the choice of m is not straightforward. m should not be set too low or too high, since "the order-*m* estimator will converge to the DEA frontier estimator if $m \to \infty$ but, for finite m, it will not envelop all the data points" (Cazals et al, 2002). Usually in practical applications $m < n$ and it should be set so to have a "sufficiently small decrease in the proportion of super-efficient observations" (Schiltz et al., 2019). A sensitivity analysis for different values of m helps to support the robustness of the findings.

The identification of peers and targets in evaluations with "partial" frontiers, particularly using the order-*m* method, is also possible. When the focus is on small samples representing real-world settings, these outcomes can be easily investigated and practically explored. For example, the number of times a unit is considered as a peer can be computed, while keeping in mind that now units are drawn from a subset of the sample. Furthermore, the intensity values

(λ_j^*) can be computed as an average value. Likewise, the target values can be obtained as the average target value across the B iterations, pointing at the average resource savings, once the units are evaluated against a "partial" frontier that envelops the data more tightly (representing a less demanding reference of best-practice). As a result, the order-*m* efficiency score will always be higher than the one obtained from the full frontier estimation, and consequently the order-*m* target values will be lower and more realistic to be achieved.

In this paper, we are going to explore, in a first moment, the unconditional order-*m* method, as developed in Daraio and Simar, 2007, in the context of the assessment of WWTPs that are in a problematic sample. The first challenge to be overcome is to understand which value of *m* should be selected given the particular features of this type of sample, particularly its small size. Secondly, we are going to identify, for each plant in the sample, the by-products from the order-*m* method application and, as mentioned, this is a practice not commonly undertaken in the related literature. We are here referring to the case of the peers' identification and targets computation.

5.2.3 The conditional order-*m* approach: the role of contextual variables

Besides softening the impact of outlying observations in a problematic sample, the role of the operating context needs to be accounted for when it comes to give operational guidelines to the units under assessment. This aspect enters both in the computation of the efficiency scores and in the detection of the influence of these external factors on the production process itself. The literature on efficiency has acknowledged the relevance of these aspects and has developed mainly two approaches to address them, namely the 'one-stage' approach and the 'two-stage' approach (Daraio and Simar, 2007). In the former, environmental variables are directly included in the model estimation, but they first need to be classified a priori for the analysis either as an input or as an output. In the latter, the efficiency scores are parametrically regressed in a second stage on nondiscretionary variables. However, a few issues might arise. First, the efficiency scores are serially correlated and the first stage efficiency scores are biased. Second, even if these problems would be addressed by bootstrap techniques, this approach would need a priori specification of the parametric regression model and it would imply the "separability condition". Estimating first the efficiency scores and then regressing them on the environmental variables would assume that these variables have no influence on the attainable set. In many practical applications, this assumption is difficult to defend (for a thorough explanation, we refer to Daraio and Simar (2007) and the references therein).

To circumvent the above-mentioned issues, conditional nonparametric frontier models have been introduced by Daraio and Simar (2005). Following insights from Daraio and Simar (2005, 2007) and De Witte and Kortelainen (2013), this paper integrates, in a second moment, the order-*m* method presented in the previous section with its conditional version.

Computationally, the conditional order-*m* frontier estimation requires a slight adjustment with respect to the (unconditional) order-*m* frontier estimation as described above. In an input-oriented setup as reference, the adjustment concerns the way m units are drawn among those producing at least the same output level as the evaluated unit (Rogge and De Jaeger, 2013). In the unconditional case, m units are drawn with replacement and with uniform probability, so that each unit is equally likely to be drawn. In the conditional case instead, m units are drawn with replacement but with a weight (probability) determined by an estimated kernel function. Accordingly, units that operate in a more similar environment will have a higher probability of being drawn and included in the reference set. Likewise, units that operate in a more dissimilar environment will have a lower probability of being drawn and considered in the reference set. In this sense, the efficiency score is 'conditional' upon the environmental variables z. The conditional order-*m* efficiency score $\hat{\theta}_k^{m,z}$ is obtained as the average of the conditional efficiency scores $\theta_k^{b,m,z}$ computed at each b-th iteration ($b = 1, \dots, B$) as indicated in equation (5.6).

$$\hat{\theta}_k^{m,z} = \frac{1}{B}\sum_{b=1}^{B} \theta_k^{b,m,z} \qquad \textbf{(5.6)}$$

In addition to using the conditional approach to estimate efficiency scores that account for different operating contexts, we are also going to provide a comparison between the conditional and the unconditional efficiency score so as to help detecting and to better understand the influence of the environmental variable on the efficiency. In practice, the ratio of the conditional and unconditional order-*m* efficiency estimates is computed as in (5.7) and nonparametrically regressed on the external variables z (Li and Racine, 2007).

$$Q_k = \frac{\hat{\theta}_k^{m,z}}{\hat{\theta}_k^{m}} \qquad \textbf{(5.7)}$$

The framework for the nonparametric regression smoothing is given by equation (5.8).

$$Q_k = g(z_k) + \epsilon_k, \qquad k = 1, \dots, n \qquad \textbf{(5.8)}$$

By making statistical inference, information about the sign and the statistical significance of the influence of environmental variables on the efficiency estimates can be retrieved (Rogge and De Jaeger, 2013). Moreover, it is possible to graphically visualize the ratio Q_k with a scatter plot when z is univariate or with a partial scatter plot when z is multivariate, so to help with the interpretation of the role of z on efficiency. In an input orientation, if Q_k is increasing, z plays a detrimental (unfavourable) influence on the efficiency. Vice versa, if Q_k is decreasing, z plays a conducive (favourable) influence on the efficiency (Daraio and Simar, 2005).

5.3. Case study

To demonstrate the practical usefulness of the methodological approach from a managerial and operational perspective, we use a case study of Portuguese wastewater treatment plants (WWTPs). It provides a practical example of the potential problems that may occur when attempting to conduct a fair efficiency assessment of DMUs in a problematic sample.

The Portuguese case study consists of a set of 41 WWTPs that are managed by Águas do Centro Litoral (AdCL), a public-owned company. All WWTPs are activated sludge systems, meaning that all WWTPs are converting inputs into outputs using the same secondary treatment technology (for more details on the functioning of the WWTPs, please refer to Appendix A).

The set of plants selected treated 80% of total volume of wastewater and were responsible for 91% of the electricity consumption of all WWTPs managed by AdCL in the year of 2015. All data used in this study was provided by AdCL and refers to the year 2015. The plants are quite heterogeneous in terms of dimension: 20 WWTPs serve less than 5000 inhabitants (measured in Population Equivalent, PE), 16 WWTPs serve between 5000 and 50000 inhabitants, and 5 WWTPs serve more than 50000 inhabitants. The smallest of these WWTPs was designed to serve 79 PE and the largest to serve 230020 PE. Only 8 WWTPs conduct a disinfection process. Only 5 WWTPs conduct anaerobic digestion with biogas production for energy recovery.

5.3.1 Modelling the WWTPs' activity: input and output specification

The literature reports several successful applications of DEA in the context of WWTPs performance assessment (e.g., Hernández-Sancho and Sala-Garrido, 2009). A critical phase of any DEA assessment concerns the selection of the most appropriate indicators to use as inputs and outputs of the productive process. Accordingly, input and output variables have been selected in compliance with the existing literature and the data availability. Moreover, the company managers and decision-makers of the plants under assessment have ultimately validated them to ensure a sound estimation of efficiency.

In terms of the input indicators, the majority of studies reviewed used costs (operation, maintenance and other costs) either expressed in €/m^3 (e.g., Fuentes et al., 2015; Sala-Garrido et al., 2011; Sala-Garrido et al., 2012a; Molinos-Senante et al., 2014a,b; Hernández-Sancho et al., 2011) or in €/year (e.g., Castellet & Molinos-Senante, 2016; D'Inverno et al., 2018; Guerrini et al., 2016; Hernández-Sancho et al., 2009; Molinos-Senante et al., 2016; Sala-Garrido et al., 2012b). The study by Lorenzo-Toja et al. (2015) considered the physical quantities of resources used per year (i.e., electricity and chemicals).

In this study, we selected as inputs the physical resources (energy and labour) used by each plant per year. We could not consider other resources in the efficiency evaluation, such as maintenance actions or materials consumed, due to data unavailability. Nonetheless, energy and labour represent the largest share of the operational expenditures for a typical WWTP (Silva and Rosa, 2015). The energy indicator corresponds to the energy balance at each WWTP, obtained as the difference between the sum of electricity and natural gas annually consumed by the plant and the total amount of electricity produced (measured in kWh/year). The labour indicator corresponds to the number of full-time equivalent workers assigned to each plant, obtained as the ratio between the number of weekly person-hours that are spent by workers at each plant and the number of weekly hours corresponding to a full-time worker. Note that our input specification also takes into account the scale of operation of the WWTP, as we do not normalize the resources used by the volume of wastewater treated (m^3). This ensures that no information inherent to the size of the plant is lost.

In terms of the output indicators used in WWTP assessments, the majority of studies reviewed use the amount of pollutants removed from the wastewater, either expressed in kg/year (e.g., Castellet & Molinos-Senante, 2016; Hernández-Sancho et al., 2009; Molinos-Senante et al., 2016; Sala-Garrido et al., 2011; Sala-Garrido et al., 2012b), or in g/m^3 (e.g., Hernández-Sancho et al., 2011; Molinos-Senante et al., 2014a,b; Sala-Garrido et al., 2012a; Fuentes et al., 2015), which corresponds to the difference between the pollutant in the influent and effluent. Dong et al. (2017) introduced a novel approach by using the rate of pollutants removal. This is interpreted as the ratio between the amount of pollutant removed and the amount of pollutant that is present in the raw wastewater (i.e., in the influent). Therefore, this indicator accounts for the concentration of each pollutant in the influent sewage. Indeed, considering two WWTPs (A and B) that treat the same volume of wastewater, it is not the same to remove 200 mg/l out of 600 mg/l on entry in plant A, or to remove 200 mg/l out of 1000 mg/l on entry in plant B. Although in terms of the absolute amount of pollutant removed both plants are identical, the rate of pollutant removal is higher in A than in B (e.g., A:200/600 > B:200/1000). Therefore, considering only the absolute amount of pollutant removed could imply an unfavourable evaluation of plant A since it demands a higher level of resource consumption than B. For this reason, the use of the rate of pollutant removal instead of the absolute amount of pollutant removed as the output specification allows a fairer efficiency comparison among plants.

However, since the scale is a factor that affects the performance of WWTPs (e.g., Hernández-Sancho and Sala-Garrido, 2009), the rate of pollutant removal still does not capture the real amount of pollutant removed from the influent wastewater, which varies in pollutant concentration that will have also an impact on the required effort in terms of the total amount of resource consumption. In this study, we contribute to the literature by proposing a novel WWTP output specification, namely the "amount of pollutant actually removed", that is defined as the rate of pollutant removal multiplied by the total volume of wastewater treated (m^3 per year). In this study, we used as outputs the "amount of pollutant actually removed", both in terms of suspended solids (SS) and chemical oxygen demand (COD), which is consistent with the input specification in terms of total amount of resources consumed (per year).

Table 5.1 shows the descriptive statistics of the inputs and outputs used in the DEA model. The variables display an important variability that emerges especially looking at the quartiles. This variability is a consequence of the characteristics of the set of plants that constitute the sample, both in terms of size and contextual factors affecting WWTP activity. As pointed out in the previous sections, the performance evaluation aims to provide a full picture of the WWTP system to enable the design of sound asset management policies by company managers, as well as improvements in operational procedures to be implemented by local managers of all plants.

Table 5.1 – Descriptive statistics for inputs and outputs of the DEA model (41 WWTPs, year of 2015).

		Mean	SD	Min	Max	Q1	Median	Q3
Inputs	Energetic balance (MWh/year)	376.90	781.48	21.50	4476.00	51.74	135.30	201.40
	Number of full-time equivalent workers	1.10	1.89	0.20	9.00	0.30	0.45	0.61
Outputs	$(COD_{removed}/COD_{entry})$×Vol. wastewater [thousand m^3 per year]	987.10	2468.53	2.01	12461	62.00	140.50	323.50
	$(SS_{removed}/SS_{entry})$×Vol. wastewater [thousand m^3 per year]	1022.00	2557	2.01	12875	57.29	141.90	331.00

The threat to the homogeneity of the units under analysis emerges not only from the observed variability on the magnitude of the inputs and the outputs, but also from the environmental conditions faced by the WWTPs. The majority of studies recognize the relevance of the impact of environmental conditions faced by WWTPs on their performance. In some of these studies, two-stage approaches have been used to explain the variability of the DEA efficiency scores (e.g., Hernández-Sancho et al., 2009; Dong et al., 2017; D'Inverno et al., 2018; Molinos-Senante et al., 2016; Hernández-Sancho et al., 2011; Molinos-Senante et al., 2014a,b; Sala-Garrido et al., 2012b). Alternatively, conditional frontier estimators have also been used within the context of WWTPs performance assessment, accounting for contextual conditions directly in the computation of the efficiency scores (e.g., Fuentes et al., 2015; Guerrini et al., 2016).

Table 5.2 provides a review of the contextual factors deemed to play a role in the WWTP activity, listing the most frequently investigated contextual variables, classified into structural factors, operational factors and other factors.

Table 5.2 – Overview of the contextual variables used in DEA related studies of WWTPs.

	Structural factors						Operational factors					Other factors	
Study	Plant capacity (or plant size)	Plant age	Secondary treatment technology	Technology used for sludge treatment	Nutrient removal and Tertiary treatment	Type of aeration in the bioreactor	Characteristics of influent wastewater	Capacity utilization	Energy consumption	Sludge generated	Compliance with regulatory effluent standards	Climate	Seasonality
D'Inverno et al. (2018)	x	x	x		x		x				x		x
Dong et al. (2017)	x		x				x	x				x	
Fuentes et al. (2015)	x	x				x	x						
Gómez et al. (2017)	x	x	x					x	x				
Hernández-Sancho et al. (2009)	x												
Hernández-Sancho et al. (2011)	x	x				x	x						
Guerrini et al. (2016)	x	x	x			x	x	x			x		
Lorenzo-Toja et al. (2015)	x				x		x	x				x	
Molinos-Senante et al. (2016)	x	x	x	x				x					
Molinos-Senante et al. (2014a)	x	x		x	x								
Molinos-Senante et al. (2014b)	x	x	x	x					x	x			
Sala-Garrido et al. (2012b)													x
No. times used	**11**	**8**	**6**	**3**	**2**	**3**	**6**	**6**	**2**	**1**	**1**	**2**	**1**

Among all variables considered in the literature, plant capacity (or plant size), plant age, secondary treatment technology, capacity utilization and characteristics of the influent wastewater, were the variables chosen more often to analyze their potential influence on the efficiency of the WWTPs.

As the contextual factors may strongly influence the performance of the WWTPs, we also discussed with AdCL's managers what would be the most relevant set of contextual factors to consider in the analysis of their plants.

As managers were very concerned with the topic of asset management, they demonstrated particular interest about the age of the plants. Also, they were concerned with five additional aspects that could be relevant to explain differences in relative performance. The first of these aspects is the plant dimension. There are very big and very small plants in the sample, and the operation is affected by their size. The second aspect is capacity utilization. This is a particularly critical aspect in at least one plant that was built to accommodate seasonal events (religious). The third aspect is the existence of disinfection treatment in some of the plants. This occurs due to legal requirements concerning the effluent quality. In the plants considered in our sample, the disinfection treatment is made through UV radiation, which implies additional energy consumption at these plants. The fourth aspect is the existence of sludge dehydration in some of the plants, which requires additional labour and electricity consumption. The fifth aspect is the existence of pumping facilities inside some of the plants. This is required by the geographical features of the place where the plant is located, and causes additional energy consumption.

In summary, we selected the following set of contextual factors: (i) age of the plant (expressed in terms of the number of years since the construction of the WWTP or since the last rehabilitation), (ii) installed capacity (expressed in terms of the volume of wastewater that can be treated at each plant on a daily basis, in m^3/day), (iii) capacity utilization (expressed as the ratio between the average volume of wastewater treated per day and the installed capacity), (iv) disinfection treatment (binary variable), (v) sludge dehydration (binary variable), and (vi) pumping facilities inside the WWTP (binary variable).

All the contextual factors listed above represent structural conditions at the plants, except the capacity utilization of the plant that corresponds to an uncontrollable operational aspect.

Table 5.3 shows the descriptive statistics of the contextual variables used in the conditional efficiency assessment.

Table 5.3 – Descriptive statistics for the contextual variables (41 WWTPs, year of 2015).

Contextual variable	Description	Mean	SD	Min	Max
Plant age	Years since construction or last intervention	11	5	4	23
Installed capacity	m^3/day	5400	12015.49	90	49000
Percentage of utilization of the installed capacity	Average volume of wastewater treated (m^3/day)/Installed capacity (m^3/day)	54.32%	32.33%	6.59%	155%
Disinfection treatment	Dummy = 1 if present	19.51%			
Dehydration of sludge	Dummy = 1 if present	26.83%			
Pumping facilities	Dummy = 1 if present	58.54%			

To conclude, Figure 5.1 depicts a visual summary of the input and output indicators, as well as the environmental variables considered in this study.

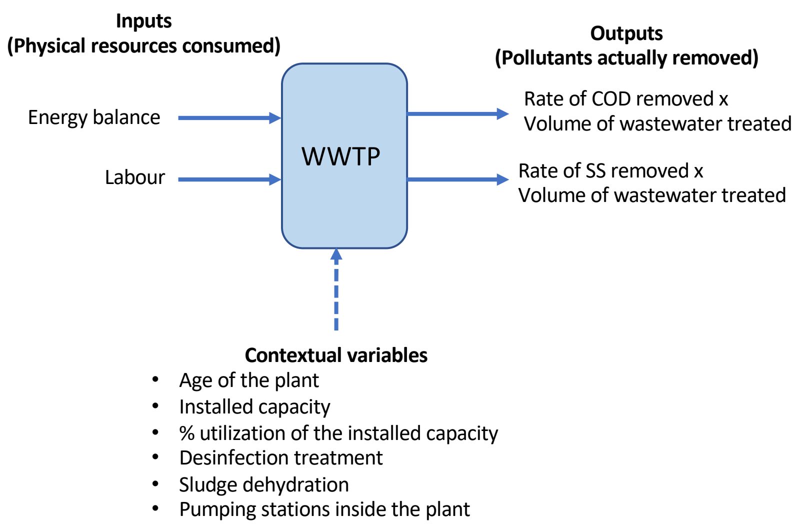


Figure 5.1 *– Schematic representation of the conditional DEA model used to assess WWTPs performance in this study.*

5.4. Results and discussion

5.4.1 Results of the order-*m* application compared with DEA results

This section presents the results of DEA in comparison with the order-*m* unconditional efficiency scores for the WWTPs in the problematic sample. We also explore the practical information to be delivered to WWTPs managers to guide performance improvements. Specifically, we used a VRS specification of the DEA model, in compliance with existing literature demonstrating that this kind of facilities exhibits VRS technology (e.g. Hernández-Sancho et al., 2009).

The first issue that must be resolved in assessments using the order-*m* method is the choice of the value of m (number of units that are drawn at random with replacement among those producing at least the same output as the unit under evaluation). As explained in section 5.2.2, the value of m can be selected by plotting the percentage of super-efficient units against the values of m, and then selecting the elbow value after which the percentage of super-efficient units stabilizes. Figure 5.2 shows the sensitivity analysis for different values of m, both in terms of % of super-efficient units and sample's average order-*m* efficiency scores. The detailed data at the DMU level underlying the construction of the graphs shown in Figure 5.2 is shown in Appendix 5B.

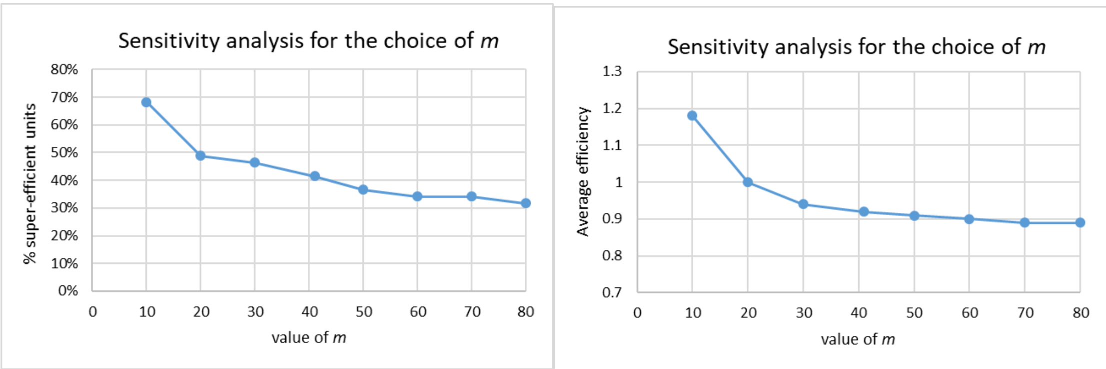


***Figure 5.2** – Sensitivity analysis for selection of the m-value.*

The analysis of these graphs led us to the choice of a value of m equal to 41, which corresponds to the number of DMUs in the sample. A value of m equal to the number of DMUs in the sample might be regarded as a recommended option for problematic samples due to the small dimension.

Table 5.4 shows the descriptive statistics of the order-*m* unconditional efficiency scores for different values of m, together with the DEA scores.

Table 5.4 – Summary information of efficiency results for different model specifications

Summary Statistics	DEA	Order-*m*							
		(m = 10)	(m = 20)	(m = 30)	(m = 41)	(m = 50)	(m = 60)	(m = 70)	(m = 80)
Mean	0.654	1.180	1.000	0.940	**0.920**	0.910	0.900	**0.890**	**0.890**
StDev	0.242	0.363	0.239	0.197	**0.180**	0.170	0.165	0.162	0.160
Min	0.245	0.647	0.573	0.547	**0.537**	0.532	0.529	0.528	0.527
Max	1	2.338	1.807	1.525	**1.367**	1.258	1.186	1.126	1.099
# of (super-)efficient units	8	29	23	23	**22**	21	21	21	21

The magnitude of the efficiency scores obtained for the DEA and the order-*m* approaches are quite different. The efficiency scores obtained under the DEA approach (mean value of 0.654) are much lower than in the order-*m* approaches, regardless of the value of m selected. This highlights the importance of using the order-*m* approach for the analysis of the problematic sample under study, as a full-frontier may lead to results with limited face-validity from a practical perspective, compromising the design of policies for continuous improvement.

Using the order-*m* method with m equal to 41, we found 19 plants performing inefficiently. Since the aim of this paper is to obtain deep insights upon the WWTPs' relative performance, Table 5.5 presents the information obtained regarding efficiency scores and targets for the detected inefficient DMUs.

Table 5.5 – Results of the efficiency analysis for the inefficient units, by using the order-*m* method.

Inefficient WWTPs	# units with ≥ output level than the evaluated unit	Order-*m* efficiency score	Target Energetic Balance (MWh/year)	Potential Energy savings (MWh/year)	Target Labor (# of FTE workers)	Potential Labour savings (# of FTE workers)
P5	4	0.924	1485.2	122.9	5.2	3.80
P9	7	0.857	489.6	421.0	1.2	0.20
P12	13	0.605	121.7	79.7	0.4	0.25
P14	18	0.572	126.7	94.5	0.3	0.26
P15	19	0.537	98.4	84.6	0.5	0.42
P16	13	0.670	115.2	56.3	0.4	0.21
P17	20	0.677	114.4	54.4	0.4	0.19
P19	11	0.870	153.1	22.8	0.2	0.20
P22	23	0.659	117.9	61.2	0.4	0.21
P23	22	0.769	103.9	31.4	0.5	0.14
P24	15	0.907	91.6	9.4	0.5	0.10
P26	40	0.696	24.9	11.0	0.3	0.15
P28	25	0.928	105.3	8.0	0.4	0.03
P29	29	0.678	105.3	50.4	0.3	0.15
P34	16	0.922	120.4	10.1	0.4	0.03
P35	21	0.674	145.1	72.2	0.3	0.13
P36	36	0.985	32.6	0.65	0.3	0.07
P38	34	0.820	34.4	7.6	0.3	0.13
P40	14	0.819	146.1	32.5	0.2	0.06
Mean		**0.766**	**196.4**	**64.8**	**0.7**	**0.35**
St.Dev		**0.136**	**317.7**	**93.0**	**1.1**	**0.84**
Min		**0.537**	**24.9**	**0.65**	**0.2**	**0.03**
Max		**0.985**	**1485.2**	**421.0**	**5.2**	**3.80**

From the analysis of Table 5.5 we can see that even despite the features underlying the order-*m* efficiency assessment (mitigation of the impact of outliers in the comparative evaluation process through resampling and construction of a partial frontier based only on units with the same or higher level of output than the unit under evaluation) there is still room for improvement. The average inefficiency is 23.4% (1-76.6%), which means that there is room for reducing the overall level of resource usage by 23.4%. In the case of the energy input, the total energy wasted due to inefficiency is 1230.8 MWh/year, corresponding to 8% of the total energy consumed by the 41 WWTPs. In the case of labor, the total number of full-time equivalent workers that can be potentially reduced is 6.7 (or equivalently 268 hours per week, assuming 1 FTE corresponds to 40 hours per week). This labour time could be applied to different functions within the company to improve the overall functioning of the WWTPs system.

Table 5.6 complements the insights gained by looking with more detail into the efficiency assessment exercise with the description of the peers identified. Table 5.6 presents the vector of intensities λ_j for each DMU, which indicates how relevant other (observed) WWTPs j are for constructing the benchmark against which the efficiency of plant k is assessed. The λ_j's presented in the intensity matrix are the average values of the λ_j's calculated for each plant in B iterations. Recall that the efficiency results are robustified by creating B (=2000) efficiency scores per plant given the subsample of m (=41) plants (the order-*m* approach) generated in each iteration.

Table 6 – List of peers and intensity values for the units considered inefficient under the order-*m* approach.

		Intensity values of WWTPs considered peers																					
		P1	P2	P3	P6	P10	P18	P19	P20	P21	P24	P25	P26	P28	P30	P32	P33	P34	P36	P37	P38	P40	P41
Inefficient WWTP	P5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	P12	0	0	0	0.46	0	0	0.03	0	0	0	0.51	0	0	0	0	0	0	0	0	0	0	0
	P14	0	0	0	0.5	0	0	0	0	0	0.03	0.38	0	0	0	0	0	0.07	0	0	0	0.02	0
	P15	0	0	0	0.08	0	0	0	0	0	0.08	0.8	0	0	0	0	0	0.03	0	0	0	0	0
	P16	0	0	0	0.37	0	0	0.02	0	0	0	0.62	0	0	0	0	0	0	0	0	0	0	0
	P17	0	0	0	0.3	0	0	0	0	0	0.06	0.54	0	0	0	0	0	0.09	0	0	0	0	0
	P19	0	0	0	0.98	0	0	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P22	0	0	0	0.32	0	0	0	0	0	0.07	0.46	0	0	0	0	0	0.14	0	0	0	0.01	0
	P23	0	0	0	0.15	0	0	0	0	0	0.09	0.69	0	0	0	0	0	0.06	0	0	0	0	0
	P24	0	0	0	0	0	0	0	0	0	0.05	0.94	0	0	0	0	0	0	0	0	0	0	0
	P26	0	0.01	0	0	0	0	0	0	0	0	0	0.01	0	0.08	0.23	0.63	0	0.04	0	0	0	0
	P28	0	0	0	0.14	0	0	0	0	0	0	0.12	0	0.02	0	0	0	0.02	0	0.67	0	0.02	0
	P29	0	0	0	0.1	0	0.1	0	0.03	0	0	0	0	0	0	0	0	0.01	0	0	0	0.03	0.72
	P34	0	0	0	0.42	0	0	0	0	0	0.03	0.47	0	0	0	0	0	0.07	0	0	0	0	0
	P35	0	0	0	0.69	0	0	0.02	0	0	0.02	0.14	0	0	0	0	0	0.02	0	0	0	0.11	0
	P36	0	0	0.02	0	0	0	0	0	0	0	0	0	0	0.69	0	0	0	0.21	0	0.08	0	0
	P38	0	0	0.06	0	0.02	0	0	0	0	0	0	0	0	0.72	0	0	0	0	0	0.21	0	0
	P40	0	0	0	0.82	0	0	0	0	0	0	0.13	0	0	0	0	0	0	0	0	0	0.05	0

From an individual perspective, we can take the example of WWTP P15 that was assigned the lowest value of efficiency to give an in-depth understanding of the practicality inherent to the order-*m* approach. This plant was considered comparable to 19 other WWTPs of AdCL, meaning that these other 19 plants produce the same or more output that WWTP P15 (see Table 5C.1 in the Appendix 5C). P15 uses a very high percentage of the installed capacity (in fact, on average, it was working with overload in the year analyzed). It is a relatively new facility (with 9 years in 2015). Table 5.6 shows that plant P25 was considered the most relevant peer for plant P15, with an intensity value of 0.8. This suggests that P15 should look for the best practices implemented in plant P25 to foster performance improvements. In fact, the similarities between P15 and P25 are noticeable. Both were overloaded in the period analyzed, and P25 is only 6 years older than P15, indicating that both are in useful age. Note that using the DEA approach the efficiency score of WWTP P15 was only 0.334 (a difference of 20.3% in relation to the order-*m* approach).

Another example is WWTP P13, with the biggest difference between the results obtained in the DEA and order-*m* approaches (64.7%). In the order-*m* approach this plant is considered comparable to 7 other plants (see Table 5C.1 in the Appendix 5C), and was considered slightly super-efficient (1.0004). This suggests that plant P13 was being unfairly evaluated when using DEA (efficiency score equal to 0.353) due to comparisons with very different WWTPs.

Finally, we consider the case of WWTP P9, which was set as the most inefficient plant in the DEA approach (efficiency score equal to 0.245). This plant was assigned an efficiency score of 0.857 in the order-*m* approach. Looking at the characteristics of this particular WWTP, we see that it had a very low average percentage of utilization of the installed capacity in the year of

2015 (37.9%). This plant is very affected by seasonality due to religious events that attract a lot of people to the city of Fátima for a few days per year, in the area served by this WWTP. In addition, this WWTP also undertakes a disinfection treatment. Although it may be difficult to improve the efficiency of this plant due to the adverse contextual conditions, the order-*m* approach allows comparisons with 7 other plants in the sample and found WWTP P21 as its peer. This plant also had a very low average percentage of utilization of the installed capacity (18.1%) in the year of 2015, and can be used as a benchmark to share best practices and guide the implementation of enhanced management and operational procedures at plant P9.

As expected, all units that were considered efficient in the DEA approach continue to be considered efficient in the order-*m* approach. However, the number of efficient units increases significantly in the order-*m* approach by mitigating the effect of outliers in the assessment and by imposing a more restricted assumption regarding the comparability among DMUs. Accordingly, a total of 22 WWTPs were identified as (super-)efficient in the order-*m* approach, against only 8 efficient WWTPs under the DEA approach. Consequently, the results of the order-*m* efficiency assessment match more closely stakeholders' perceptions concerning WWTPs performance than the results of the DEA model. The targets obtained using DEA would be very demanding, and perceived as impractical to accomplish with the technology available at the WWTPs.

Other important conclusions can be drawn by comparing the number of times each WWTP is identified as peer and the average intensity value (λ_j) obtained in the DEA versus order-*m* DEA approaches for each WWTP. In the case of the order-*m* approach, the number of times a plant is identified as peer is obtained by counting the number of times that the average results of the intensity values obtained in the B iterations were greater than zero. This means that even a WWTP considered inefficient (on average) may have been considered efficient at least once in the B iterations and therefore be counted as a peer of itself. Table 5.7 summarizes this information concisely.

Table 5.7 – List of WWTPs, number of times each WWTP is identified as peer and average intensity values.

	WWTP	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
DEA	# of times used as peer				29		29	4							
	Average intensity				0.058		0.267	0.025							
Order-*m*	# of times used as peer	2	5	9	1		27	2	2		9	5		1	
	Average intensity	0.049	0.019	0.035	0.024		0.165	0.024	0.024		0.033	0.002		0.024	

	WWTP	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
DEA	# of times used as peer													35	
	Average intensity													0.649	
Order-*m*	# of times used as peer		1		7	14	9	4		3	13	21	2	1	5
	Average intensity		0.000		0.028	0.002	0.028	0.049		0.000	0.012	0.171	0.000	0.015	0.000

	WWTP	P29	P30	P31	P32	P33	P34	P35	P36	P37	P38	P39	P40	P41
DEA	# of times used as peer													
	Average intensity													
Order-*m*	# of times used as peer		12	4	4	3	13	2	6	8	6	9	19	6
	Average intensity		0.084	0.024	0.029	0.036	0.012	0.000	0.008	0.041	0.007	0.006	0.007	0.038

There are three main aspects to retain from Table 5.7. First, WWTP P6 is considered the preferable peer both in the DEA and the order-*m* approaches. This signals outstanding relative performance at this plant in the year 2015.

Secondly, WWTP P4 was considered 29 times as peer for the underperforming units in the DEA benchmarking exercise (although with an average intensity value of only 0.06), whereas in the order-*m* approach P4 is only peer to itself. In fact, WWTP P4 is the largest of the sample in terms of output levels, so although all other plants could have used this DMU as peer, they have selected other plants as benchmarks in the order-*m* analysis (it was only used once as peer).

On the other side of the dimension spectrum, the smallest plant of the sample is P27. It has been considered 35 times as peer in the DEA approach, with the highest intensity value (0.65). Conversely, in the order-*m* approach, this WWTP could not be included in the comparator set of any other plant (see Table 5C.1 in Appendix 5C). Therefore, given the assumptions underlying the order-*m* approach, this plant could not be a peer to any other plant except itself (and indeed it is a self-comparator, despite having a small lambda value in its own efficiency evaluation).

Thirdly, WWTP P25 is considered as a preferable peer in the order-*m* approach (from Table 7 it can be observed that 22 plants have used it as a benchmark in the efficiency evaluation) whereas it is considered inefficient in the DEA approach (efficiency score of 0.739). In fact, P25 is comparable only to 12 plants in the overall sample (see the line corresponding to P25 in Table 5C.1 of Appendix 5C). This shows that by mitigating the effect of outliers, some plants that are considered inefficient in the DEA approach are able to stand out as examples of best practices in the order-*m* assessment.

The information retrieved from the analysis of Tables 5.6 and 5.7 leads to summarize the two essential aspects that, combined, constitute the novelty of using the order-*m* approach for efficiency evaluation of DMUs in a problematic sample. The first aspect is that valuable managerial information can be retrieved by exploring the by-products of a robust non-parametric model. Said by-products consist in the additional specification of peers and computation of targets. This is an issue that is not commonly covered in the related literature but which nevertheless is very relevant for internal benchmarking studies. More precisely, and in the first place, it is the identification of peers that paves the way for information sharing among managers, operators, and technical staff. Thus, it constitutes a starting point for planning the action-taking that leads to performance improvements. In the second place, the computation of targets gives both managers and operational staff a direct indication of necessary improvements and a need for further insights gathering from both the current practices and the best practices observed in the identified peers.

The second aspect concerns the comparison of the results corresponding to robust and deterministic approaches. Best-practice sharing can be best promoted using a robust approach, namely the order-*m* method, as the peers identified are more similar to the unit under assessment. This is actually the essential ingredient for fair benchmarking evaluations that instill confidence in companies' managers and decision-makers in the results obtained.

5.4.2 Results of the conditional order-*m* application

As demonstrated in the previous section, the order-*m* approach allows a trustworthy comparative evaluation of units, which serves as a baseline reference for the estimation of the efficiency levels of the WWTPs. To complement this analysis, we also explored the relationship of contextual factors with plants' performance. This exploratory study aims to grasp additional insights concerning the evaluation of small samples with variability in the values of the input and output indicators and facing heterogeneous environmental conditions.

The statistical significance of the models is an issue that must be dealt with caution. Also, in the presence of small samples, it is not possible to include all potentially relevant contextual factors in the same model, which prevents exploring interactions among variables. Nevertheless, conditional models can still guide decision makers concerning the significance and direction of influence of each contextual variable on the performance of the units under assessment.

Table 5.9 shows the summary results obtained for the conditional models tested. This table also indicates the direction of influence of the conditional factor on WWTP efficiency (favorable or unfavorable). Appendix 5D provides the complete information on the efficiency scores obtained for each WWTP in the models considered.

Table 5.9 – Results of the conditional order-*m* efficiency analysis and statistical inference.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Variables	Age	Installed pacity	% of utilization installed capacity	Disinfection atment	Sludge hydration	Pumping ilities
Mean	0.945	0.904	0.940	0.916	0.922	0.910
St.Dev.	0.111	0.168	0.129	0.168	0.146	0.170
Min.	0.639	0.528	0.526	0.535	0.535	0.536
Max.	1.058	1.195	1.166	1.323	1.155	1.249

Direction of fluence	Unfavorable	Unfavorable	Unfavorable	Unfavorable	Unfavorable	Unfavorable
p-value	0.067	0.049*	0.407	0.001*	0.044*	0.018*

*Statistically significant at a 5% confidence level

We can conclude that two of the factors considered (age and percentage of utilization of installed capacity) do not have a statistically significant relationship with efficiency. Regarding the variable age, similar results were found, for example, by Hernández-Sancho et al. (2011) and D'Inverno et al. (2018).

Regarding the percentage of utilization of the installed capacity, Guerrini et al. (2016) found that a high use of the installed capacity has a positive influence on the performance levels. Although the analysis of our sample did not confirm this result at a statistically significant level, the analysis of the scatter plot for this conditional model (see Figure 5.3) shows a decreasing slope for a % of the utilization of the installed capacity between 50% and 110%, signaling that in this range of values the effect of increasing the use of the installed capacity may have a favourable influence on efficiency. This result is in accordance with what is described by Lorenzo-Toja et al. (2015) and by Dong et al. (2017).

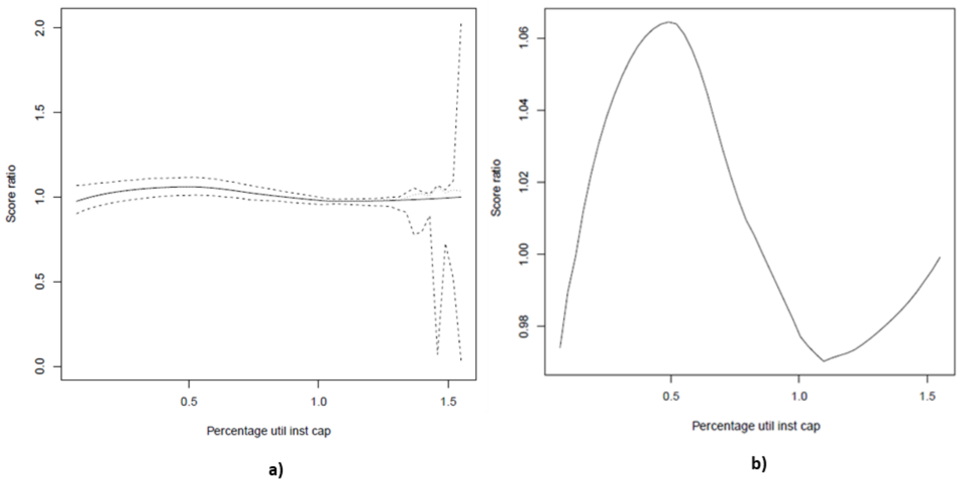


Figure 5.3 *– Univariate scatter plot for the influence of the percentage of utilization of the installed capacity a) with confidence interval; b) without confidence interval.*

According to our sample, only severe over-sizing (below 50% of utilization) and under-sizing (more than 120% of utilization) are detrimental to good performance. In our sample, 18 WWTPs are oversized and 4 WWTPs are undersized (totalizing more than half of the plants in the sample), which justifies the observed unfavorable trend of this contextual factor on performance. Furthermore, as the average percentage of utilization of the installed capacity is around 50%, it means that there is considerable room for increasing the capacity utilization in the long run. In the future, increasing the volume of wastewater treated at the 18 currently oversized plants might enhance efficiency. Accordingly, to face a potential increase in wastewater treatment needs, it could be wise to analyze the feasibility of draining the additional volumes to these existing WWTPs, as an alternative, if possible, to the construction of new facilities.

One of the contextual factors that is commonly studied in the WWTPs performance assessment literature is the plant dimension (see Table 5.2). This factor is proxied in our study by the installed capacity. Table 5.9 shows that the installed capacity has an unfavourable effect on WWTP efficiency for the AdCL sample, meaning that higher installed capacity is associated with lower levels of efficiency. Although the study by Fuentes et al. (2015), also using a conditional order-*m* approach, reached a similar conclusion regarding the unfavourable effect of dimension on efficiency, this is a result that is contrary to what is commonly described in literature (e.g., Guerrini et al., 2016). Although for our sample this effect was found to be statistically significant at a 5% level, the *p*-value is quite high (0.049), so the results should be viewed with caution. Note also that in our sample the majority of plants (36 out of 41) have a small installed capacity (i.e., lower than 6500 m^3/day), while only 5 plants have a big installed capacity (i.e., greater than 20000 m^3/day).

Concerning the disinfection treatment, sludge dehydration and the existence of pumping facilities, all these factors have a statistically significant association with the efficiency of WWTPs, and the direction of their influence is unfavorable. Thus, some of the inefficiency identified in the unconditional order-*m* model can be attributed to the influence of these contextual conditions. This result is as expected, as all these factors imply the consumption of electricity and require additional workforce.

It has been argued in the literature (e.g., Lorenzo-Toja et al., 2015) that the plants with a tertiary treatment, which includes disinfection, should not be comparatively assessed with plants that don't conduct this kind of treatment. However, in the presence of small samples, it is not possible to separate the plants in groups according to the combination of contextual conditions that they face. This is precisely what motivated this paper. We provide an alternative approach to evaluate small samples, consisting of using an unconditional order-*m* model to estimate a base-line efficiency level for all plants under evaluation, followed by a conditional order-*m* approach, such that the decision maker becomes aware of the particular issues (contextual conditions) that may be correlated with the performance of the plants. This way, the design of policy measures to improve efficiency can be done in the light of the additional information provided by the conditional analysis.

It should also be noted that although the conditional analysis revealed that some of the factors have a statistically significant relationship with the efficiency levels, the correlation between the efficiency scores obtained with conditional and unconditional order-*m* models is very high (see Table 5.10).

Table 5.10 - Spearman correlation matrix for different models' specifications

	Age	Installed capacity	% utilization of installed capacity	Disinfection treatment	Sludge dehydration	Pumping facilities
Unconditional order-*m*	0.7673	0.9606	0.8112	0.9770	0.9737	0.9922

In summary, our results demonstrate that four contextual variables have a statistically significant association with the efficiency levels. Therefore, the results revealed by the conditional models should be taken in due account, especially when considering the identification of peers to guide performance improvements at the plants.

5.5. Conclusion

In order to guide the design of policies for performance improvement, efficiency studies at macro level are often directed to a sector or area to understand trends and reveal opportunities and threats. In these cases, the samples to analyze are carefully selected, such that it is possible to reach unbiased conclusions. At a micro level, companies may also be interested in conducting efficiency studies to better tackle internal challenges for improvement of their own utilities. In these internal benchmarking situations, the samples available may be problematic, undermining the reliability of the evaluation process. Problematic samples, as defined in this paper, are small samples whose units present high variability in terms of input-output indicators and/or environmental conditions.

Considering the importance of performance assessment studies in the search for improvements within organizations, the variability within the samples should be treated without compromising the quality of the information obtained. This must be done with respect to two major concerns that typically are brought by managers and practitioners of companies or entities that manage a small number of units. The first one is the ability for undertaking fair comparisons as this is the primary block for trusted results. The second one is the ability to include all units regardless of their particular features.

In this paper, we showed the potential of an efficiency analysis framework based on the combination of unconditional and conditional order-*m* models. This procedure was explored using a real-world problematic sample of a Portuguese wastewater company. It was demonstrated that an order-*m* efficiency approach can better tackle the problem of identifying best practices and uncovering the real potential for improvement within a company than a conventional DEA method. Particularly, the order-*m* method allows the identification of a more realistic set of targets than the DEA formulation.

For example, in the case study consisting of 41 WWTP from the AdCL company, the baseline unconditional order-*m* value for reduction of the total energy consumed was 8%. This is more conservative and realistic than the value proposed by the conventional DEA approach (28%). Furthermore, by using a conditional order-*m* approach it was possible to derive additional information about the effect of the environmental factors that were not previously considered directly in the unconditional order-*m* efficiency assessment. More precisely, it was concluded that the installed capacity of the plants, the existence of disinfection treatment, sludge dehydration and pumping facilities inside the plants have a statistically significant association with the plants' efficiency levels. Therefore, these factors must be taken into account in the identification of the peers where best-practices should be observed to foster performance improvements at underperforming WWTPs. However, the use of the conditional order-*m* method applied to the case of problematic samples has a limitation. Due to the small size of the sample, it is not possible to build a model with all relevant contextual variables since it would lack statistical significance. Therefore, when conducting a performance assessment of DMUs in a small sample using a conditional order-*m* method, it is not possible to capture the full effect of the contextual conditions faced by the plants in their efficiency scores, best practices selection, or potential targets for improvement. Nevertheless, using one relevant contextual variable at a time, from the point of view of practitioners, is still of added value in comparison with only conducting the unconditional analysis since it allows shedding some light on the likely effect each variable may have in the real operational context. This limitation can be seen as a limitation that is inherent to the conditional order-*m* application to the case of problematic samples.

The information gathered can assist the decision-making process as it helps identify priorities for investments and determine the type of technical and operational measures to implement at inefficient units based on information sharing and target setting.

The practical application of the efficiency assessment framework proposed in this paper also constitutes an example of a formative evaluation procedure that may be adopted by other water companies in the pursuit of improved sustainability of their assets. Although the performance assessment framework proposed in this paper was applied to the specific case of the wastewater industry, it can also be adopted in other industry contexts, especially in internal benchmarking exercises where the samples available are small and the operating characteristics and environmental conditions differ significantly among DMUs.

Our study only covers information from one year, so future research should explore the evolution of performance over time. This could mitigate the variability in input and output levels that may occur over the years, and identify in a systematic way the plants that consistently maintain high or low levels of performance over the years. It would also allow monitoring the effect of the environmental conditions on DMUs performance, and investigate the changes in productivity levels over time.

Acknowledgements

The authors wish to acknowledge the financial support of Project "More Sustainable WWTPs”. This project is financed by the Portuguese water company Águas do Centro Litoral (AdCL). The authors are especially grateful to all the support provided by Tiago Braga and Paulo Leitão from AdCL. The authors also wish to gratefully acknowledge Kristof De Witte and Laura Carosi for their valuable comments and suggestions on a previous version of the paper. The first author also wishes to acknowledge the financial support from FCT - Fundação para a Ciência e a Tecnologia (Portuguese national funding agency for science, research and technology), given through the Grant PD/BD/142815/2018 respecting the Doctoral Program on Sustainable Energy Systems by MIT-Portugal. The fourth author also gratefully acknowledges financial support from Research Foundation-Flanders, FWO (Postdoctoral Fellowship 12U0219N).

The wastewater treatment plants activity

Wastewater treatment plants (WWTPs) are the facilities that allow to prevent pollution of surface water, by removing the contaminants that are present in the wastewater, and then delivering it back into a natural water course. Clean water is therefore linked to sanitation. These are two important aspects that constitute the basis of health and food security. For this reason, Goal 6 of the Sustainable Development Agenda for 2030 defined by the United Nations (UN General Assembly, 2015) is to ensure availability and sustainable management of water and sanitation for all. In Europe, the directive concerning urban wastewater (Council Directive, 1991) determined the collection and treatment of wastewater in all agglomerations of more than 2000 population equivalents. This requirement triggered the enlargement of the sewage system in Europe. Currently, in Europe and other world regions where sanitation is already largely implemented, one of the major challenges concerns the overall sustainability of the wastewater treatment systems, from which WWTPs constitute an important part.

The treatment processes undertaken in a WWTP are physical, mainly gravitationally driven, biological and occasionally chemical. Briefly, the wastewater treatment process comprises two different operational pathways, corresponding to water and sludge that is generated during the treatment of the wastewater. Concerning the water line, the biological treatment is called secondary treatment and allows the removal of dissolved and colloidal organic matter. There are alternative technologies available for secondary treatment, whose choice affects the configuration of the entire system of the WWTP. The most common technology used is activated sludge (Hreiz et al., 2015). In this type of WWTPs, the secondary treatment is responsible for about 50% to 60% of the total energy demand (Gu et al., 2017) due to the aeration requirements. Regarding the sludge line, some large WWTPs undertake anaerobic digestion (a type of biological treatment that occurs in the absence of oxygen) to stabilize the sludge and recovering its energy content through the biogas that is produced in such biological treatment. The biogas can then be used to produce heat and electricity in a combined heat and power (CHP) system. In other WWTPs, the sludge is stabilized aerobically, which demands more electricity consumption. In some cases, it is used an alternative chemical stabilization of the sludge. In some WWTPs, the sludge can also be dehydrated. This process also consumes resources, namely energy, reagents and workforce. For a description of the overall process of wastewater treatment, Sala-Garrido et al. (2011) emphasizes the specificities of different technologies, and Guerrini et al. (2016) focuses on the operations and processes that allow the removal of different types of pollutants.

Appendix 5B

Supplementary information on the choice of the *m*-value.

Table 5B. 1 – Order-*m* efficiency scores of WWTPs according to each *m*-value tested.

WWTP	DEA	Order-*m* (*m* = 10)	Order-*m* (*m* = 20)	Order-*m* (*m* = 30)	**Order-*m* (*m* = 40)**	Order-*m* (*m* = 50)	Order-*m* (*m* = 60)	Order-*m* (*m* = 70)	Order-*m* (*m* = 80)	Order-*m* (*m* = 90)
P1	0.815	1	1	1	**1**	1	1	1	1	1
P2	1	1.325	1.114	1.048	**1.024**	1.013	1.006	1.004	1.003	1.002
P3	1	1.467	1.205	1.1	**1.057**	1.032	1.02	1.013	1.01	1.005
P4	1	1	1	1	**1**	1	1	1	1	1
P5	0.433	0.928	0.924	0.924	**0.924**	0.924	0.924	0.924	0.924	0.924
P6	1	2.139	1.346	1.122	**1.04**	1.014	1	1	1.001	1
P7	1	1.017	1	1	**1**	1	1	1	1	1
P8	0.409	1.134	1.015	1	**1**	1	1	1	1	1
P9	0.245	0.952	0.866	0.859	**0.857**	0.857	0.857	0.857	0.857	0.857
P10	1	1.658	1.359	1.21	**1.14**	1.085	1.062	1.04	1.026	1.018
P11	1	1.247	1.077	1.03	**1.007**	1.003	1.001	1	1	1
P12	0.439	0.763	0.654	0.619	**0.605**	0.599	0.595	0.594	0.593	0.593
P13	0.353	1.132	1.012	1.002	**1**	1	1	1	1	1
P14	0.396	0.664	0.603	0.583	**0.572**	0.568	0.566	0.564	0.563	0.563
P15	0.334	0.647	0.573	0.547	**0.537**	0.532	0.529	0.528	0.527	0.527
P16	0.487	0.842	0.723	0.687	**0.671**	0.663	0.66	0.658	0.658	0.657
P17	0.400	0.787	0.713	0.688	**0.677**	0.672	0.669	0.666	0.665	0.664
P18	0.543	1.212	1.081	1.043	**1.025**	1.015	1.009	1.007	1.005	1.003
P19	0.619	1.068	0.905	0.879	**0.871**	0.869	0.868	0.868	0.868	0.867
P20	0.536	1.237	1.097	1.058	**1.036**	1.025	1.016	1.01	1.006	1.004
P21	0.647	1.386	1.065	1.009	**1**	1	1	1	1	1
P22	0.333	0.768	0.696	0.672	**0.658**	0.652	0.648	0.645	0.643	0.642
P23	0.381	0.902	0.817	0.787	**0.77**	0.762	0.756	0.754	0.751	0.751
P24	0.572	1.094	0.956	0.922	**0.906**	0.904	0.901	0.9	0.9	0.899
P25	0.739	1.373	1.123	1.051	**1.023**	1.011	1.003	1.002	1.001	1
P26	0.611	0.954	0.769	0.72	**0.695**	0.682	0.674	0.671	0.667	0.665
P27	1	1.625	1.295	1.2	**1.139**	1.107	1.086	1.063	1.05	1.035
P28	0.445	1.042	0.965	0.94	**0.928**	0.923	0.92	0.917	0.916	0.915
P29	0.444	0.797	0.72	0.692	**0.677**	0.67	0.667	0.663	0.662	0.661
P30	0.982	2.338	1.807	1.525	**1.367**	1.258	1.186	1.126	1.099	1.059
P31	0.681	1.588	1.182	1.051	**1.018**	1.002	1.002	1.002	1	1
P32	0.968	1.652	1.285	1.18	**1.12**	1.087	1.067	1.047	1.038	1.024
P33	0.969	1.48	1.194	1.092	**1.049**	1.033	1.017	1.012	1.007	1.003
P34	0.637	1.04	0.957	0.934	**0.922**	0.916	0.914	0.912	0.911	0.911
P35	0.5	0.848	0.737	0.693	**0.672**	0.663	0.656	0.653	0.65	0.649
P36	0.751	1.371	1.094	1.017	**0.982**	0.967	0.955	0.951	0.949	0.943
P37	0.486	1.176	1.061	1.026	**1.01**	1.006	1.004	1.002	1.001	1.001
P38	0.635	1.18	0.938	0.861	**0.824**	0.802	0.786	0.774	0.768	0.76
P39	0.696	1.371	1.139	1.054	**1.009**	0.986	0.973	0.966	0.958	0.957

P40	0.667	1.031	0.878	0.836	**0.818**	0.813	0.81	0.808	0.807	0.807
P41	0.667	1.216	1.096	1.052	**1.026**	1.019	1.01	1.006	1.004	1.003
# of (super-)efficient units	8	29	23	23	**22**	21	21	21	21	21
Mean	0.654	1.182	1.001	0.944	**0.918**	0.906	0.898	0.893	0.890	0.887
StDev	0.242	0.363	0.239	0.197	**0.180**	0.170	0.165	0.162	0.160	0.158
Min	0.245	0.647	0.573	0.547	**0.537**	0.532	0.529	0.528	0.527	0.527
Max	1	2.338	1.807	1.525	**1.367**	1.258	1.186	1.126	1.099	1.059

Appendix 5C

Supplementary information upon the peers in the order-*m* approach.

In the case study of WWTPs from AdCL, the peers used for the evaluation of DMU k are those shown in the columns of Table 5C.1.

Table 5C.1 – Matrix describing the fulfilment of the comparability criterium for the WWTPs under evaluation

	WWTPs that produce at least the same quantity																																								
WWTPs under evaluation	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28	P29	P30	P31	P32	P33	P34	P35	P36	P37	P38	P39	P40	P41
P1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1
P3	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	0	1	0	1	1	1
P4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P5	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P6	1	0	0	1	1	1	1	1	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
P7	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P8	1	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P9	1	0	0	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P10	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	0	1	0	0	1	1
P11	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	0	1	1	1	1	1
P12	1	0	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
P13	1	0	0	1	1	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P14	1	0	0	1	1	1	1	1	1	0	0	1	1	1	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0
P15	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	0	0	1	0	1	0	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0
P16	1	0	0	1	1	1	1	1	1	0	0	0	1	0	0	1	0	0	1	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
P17	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	1	0	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0
P18	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	0	0	1	1	0	1	0	0	1	0
P19	1	0	0	1	1	1	1	1	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
P20	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	1	0	0	1	0	0	1	1	0	1	0	0	1	0
P21	1	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P22	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1	0
P23	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1	0
P24	1	0	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0
P25	1	0	0	1	1	1	1	1	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
P26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P28	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	0	1	0	0	1	0	0	1	1	0	1	0	0	1	0
P29	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	0	0	1	1	0	1	0	0	1	1
P30	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	0	1	0	0	1	1
P31	1	0	0	1	1	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
P32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1
P33	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P34	1	0	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0
P35	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	1	0	0	1	1	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1	0
P36	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1
P37	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	1	0
P38	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	0	1	1	1	1	1
P39	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	0	1	0	1	1	1
P40	1	0	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0
P41	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	1	0	0	1	0	0	1	1	0	1	0	0	1	1

Appendix 5D

Detailed information resulting from DEA, unconditional and conditional order-*m*.

Table 5D.1 – Complete table of efficiency scores

WWTP	DEA	Unconditional order-*m*	Conditional order-*m* (*Age)*	Conditional order-*m* (*Installed capacity)*	Conditional order-m (Percentage utilization of installed capacity)	Conditional order-*m* (D*isinfection treatment)*	Conditional order-*m* (Sludge *dehydration)*	Conditional order-*m* (Pumping facilities)
P1	0.8149	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
P2	1.0000	1.0219	1.0000	1.0091	1.0347	1.0136	1.0083	1.0263
P3	1.0000	1.0514	1.0000	1.0252	1.0000	1.0322	1.0223	1.0332
P4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
P5	0.4333	0.9236	0.9998	0.9832	0.9236	0.9236	0.9236	0.9236
P6	1.0000	1.0275	1.0000	1.0000	1.0000	1.0124	1.0221	1.0171
P7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
P8	0.4089	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
P9	0.2454	0.8574	1.0000	1.0000	0.8579	0.9174	0.9461	0.8573
P10	1.0000	1.1291	1.0000	1.0715	1.0370	1.0950	1.0685	1.0902
P11	1.0000	1.0099	1.0000	1.0042	1.0000	1.0025	1.0023	1.0129
P12	0.4393	0.6052	1.0000	0.5929	0.9942	0.5985	0.6678	0.6003
P13	0.3532	1.0004	1.0000	1.0000	1.0000	1.0001	1.0001	1.0003
P14	0.3961	0.5719	0.6420	0.5652	0.6367	0.5707	0.6286	0.5704
P15	0.3343	0.5365	0.7510	0.5279	0.5263	0.5343	0.5345	0.5358
P16	0.4873	0.6701	0.9283	0.6580	0.9998	0.6635	0.7159	0.6661
P17	0.3995	0.6766	0.7684	0.6667	0.7388	0.6739	0.7012	0.6799
P18	0.5427	1.0244	1.0038	1.0091	1.0176	1.0001	1.0084	1.0187
P19	0.6191	0.8703	0.8886	0.8674	0.8694	0.8688	0.9254	0.8690
P20	0.5357	1.0353	1.0005	1.0149	1.0002	1.0000	1.0107	1.0284
P21	0.6467	1.0022	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
P22	0.3333	0.6592	0.7650	0.6468	0.7119	0.7609	0.7358	0.6553
P23	0.3809	0.7686	0.9484	0.7544	0.8579	0.7649	0.9088	0.7729
P24	0.5717	0.9070	1.0000	0.8998	1.0004	0.9038	0.8995	0.9043
P25	0.7393	1.0187	1.0000	1.0003	1.0000	1.0068	1.0000	1.0127
P26	0.6108	0.6963	0.7006	0.6784	0.6650	0.6825	0.6788	0.6909
P27	1.0000	1.1417	1.0267	1.0965	1.0012	1.1166	1.0923	1.1152
P28	0.4448	0.9277	1.0000	0.9192	0.9795	0.9250	0.9453	0.9308
P29	0.4444	0.6775	0.6984	0.6665	0.7279	0.7893	0.6780	0.6735
P30	0.9824	1.3476	1.0486	1.1909	1.1458	1.3637	1.1711	1.2490
P31	0.6811	1.0188	1.0000	1.0000	1.0020	1.0038	1.0000	1.0047
P32	0.9684	1.1183	1.0000	1.0813	1.0124	1.1000	1.0757	1.0955
P33	0.9691	1.0474	1.0115	1.0225	1.0003	1.0295	1.0209	1.0550
P34	0.6369	0.9216	0.9818	0.9115	1.0016	0.9187	0.9747	0.9184

P35	0.5000	0.6739	0.6753	0.6563	0.8414	0.6659	0.7643	0.6783
P36	0.7508	0.9853	0.9424	0.9641	0.9962	0.9928	0.9605	0.9975
P37	0.4861	1.0120	1.0000	1.0033	1.0023	1.0077	1.0027	1.0084
P38	0.6351	0.8198	1.0000	0.7913	0.9386	0.8013	0.7849	0.8079
P39	0.6961	1.0050	0.9557	0.9747	1.0045	0.9912	1.0000	0.9858
P40	0.6667	0.8187	1.0000	0.8071	1.0021	0.8147	0.9311	0.8177
P41	0.6667	1.0269	1.0000	1.0111	1.0055	1.0301	1.0148	1.0229

References

Anderson, P. and Peterson, N.C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science*, 39(10), 1261-1264.

Aragon, Y., A. Daouia, and C. Thomas-Agnan. (2005). Nonparametric frontier estimation: a conditional quantile based approach. *Econometric Theory,* 21, 358–389.

Avkiran, N. K. and Rowlands, T. (2008). How to better identify the true managerial performance: state of the art using DEA. *Omega*, 36, 317-324.

Barros, C.P. (2008). Efficiency analysis of hydroelectric generating plants: A case study for Portugal. *Energy Economics*, 30, 59-75.

Barros, C.P. (2006). Productivity Growth in the Lisbon Police Force. *Public organization Review*, 6, 21-35.

Camanho, A.S. and Dyson, R.G. (1999). Efficiency, size, benchmarks and targets for bank branches: an application of Data Envelopment Analysis. *Journal of the Operational Research Society*, 50(9), 903-915.

Camanho, A.S., Portela, M.C., and Vaz, C.B. (2009). Efficiency analysis accounting for internal and external non-discretionary factors. *Computers & Operations Research*, 36(5), 1591-1601.

Castellet, L. and Molinos-Senante, M. (2016). Efficiency assessment of wastewater treatment plants: A data envelopment analysis approach integrating technical, economic, and environmental issues. *Journal of Environmental Management*, 167, 160-166.

Cazals, C., Florens, J. P., and Simar, L. (2002). Nonparametric frontier estimation: A robust approach. *Journal of Econometrics*, 106 (1), 1-25.

Charles, V., Aparicio, J., Zhu, J. (2019). The curse of dimensionality of decision-making units: a simple approach to increase the discriminatory power of data envelopment analysis. *European Journal of Operational Research*, 279, 929-940.

Charnes, A., Cooper, W.W., and Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2(6), 429–444.

Cook, W., Tone, K., Zhu, J. (2014). Data envelopment analysis: prior to choosing a model. (Technical note), *Omega*, 44, 1-4.

Council Directive (1991). Council Directive of 21 May 1991 concerning urban waste water treatment concerning urban waste water treatment. *Official Journal of the European Union*, 135, 40-52.

D'Inverno, G., Carosi, L., Romano, G., and Guerrini, A. (2018). Water pollution in wastewater treatment plants: An efficiency analysis with undesirable output. *European Journal of Operational Research*, 269(1), 24-34.

Daouia, A. and Simar, L. (2007). Nonparametric efficiency analysis: a multivariate conditional quantile approach. *Journal of Econometrics*, *140*(2), 375-400.

Daraio, C. and Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: A probabilistic approach. *Journal of Productivity Analysis, 24*(1), 93–121.

Daraio, C. and Simar, L. (2007). Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. *Journal of Productivity Analysis*, 28, 13-32.

De Witte, K. and Kortelainen, M. (2013). What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. *Applied Economics, 45*(17), 2401–2412.

De Witte, K. and Marques, R.C. (2010a) Incorporating heterogeneity in non-parametric models: a methodological comparison, *International Journal of Operational Research*, 9(2), 188–204.

De Witte, K. and Marques, R.C. (2010b). Influential observations in frontier models, a robust non-oriented approach to the water sector. *Annals of Operations Research*, *181*(1), 377-392.

De Witte, K. and Schiltz, F. (2018). Measuring and explaining organizational effectiveness of school districts: Evidence from a robust and conditional Benefit-of-the-Doubt approach. *European Journal of Operational Research*, *267*(3), 1172-1181.

Dong, X., Zhang, X., and Zeng, S. (2017). Measuring and explaining eco-efficiencies of wastewater treatment plants in China: An uncertainty analysis perspective. *Water Research*, 112, 195-207.

Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S., and Shale, E.A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, 132, 245-259.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*. Series A (General) 120(3), 253-290.

Fuentes, R., Torregrosa, T., and Ballenilla, E. (2015). Conditional order-*m* efficiency of wastewater treatment plants: the role of environmental factors. *Water*, 7, 5503-5524.

Fusco, E., Vidoli, F., and Rogge, N. (2019). Spatial directional robust Benefit of the Doubt approach in presence of undesirable output: An application to Italian waste sector. *Omega*.

Gómez, T., Gémar, G., Molinos-Senante, M., Sala-Garrido, R., and Caballero, R. (2017). Assessing the efficiency of wastewater treatment plants: A double-bootstrap approach. *Journal of Cleaner Production*, 164, 315-324.

Gu, Y., Li, Y., Li, X., Luo, P., Wang, H., Robinson, Z. P., Wang, X., Wu, J., and Li, F. (2017). The feasibility and challenges of energy self-sufficient wastewater treatment plants. *Applied Energy*, 204, 1463-1475.

Guerrini, A., Romano, G., Mancuso, F., and Carosi, L. (2016). Identifying the performance drivers of wastewater treatment plants through conditional order-*m* efficiency analysis. *Utilities Policy*, 42, 20-31.

Hernández-Sancho, F. and Sala-Garrido, R. (2009). Technical efficiency and cost analysis in wastewater treatment processes: A DEA approach. *Desalination*, 249, 230-234.

Hernández-Sancho, F., Molinos-Senante, M., and Sala-Garrido, R. (2011). Energy efficiency in Spanish wastewater treatment plants: A non-radial DEA approach. *Science of the Total Environment*, 409(14), 2693–2699.

Hreiz, R., Latifi, M. A., and Roche, N. (2015). Optimal design and operation of activated sludge processes: State-of-the-art. *Chemical Engineering Journal*, 281, 900–920.

Lavigne, C., Jaeger, S.D., and Rogge, N. (2019). Identifying the most relevant peers for benchmarking waste management performance: a conditional directional distance Benefit-of-the-Doubt approach. *Waste Management*, 89, 418-429.

Lee, C. Y., and Cai, J. Y. (2020). LASSO variable selection in data envelopment analysis with small datasets. Omega, 91, 102019.

Li, Q. and Racine, J. (2007). *Nonparametric econometrics: Theory and practice*. Princeton: Princeton University Press.

Longo, S., Hospido, A., Lema, J. M., and Mauricio-Iglesias, M. (2018). A systematic methodology for the robust quantification of energy efficiency at wastewater treatment plants featuring Data Envelopment Analysis. *Water Research*, 141, 317-328.

Lorenzo-Toja, Y., Vázquez-Rowe, I., Chenel, S., Marín-Navarro, D., Moreira, M. T., and Feijoo, G. (2015). Eco-efficiency analysis of Spanish WWTPs using the LCA + DEA method. *Water Research*, 68, 651-666.

Moldan, B., Janoušková, S., and Hák, T. (2012). How to understand and measure environmental sustainability: Indicators and targets. *Ecological Indicators*, *17*, 4-13.

Molinos-Senante, M., Gémar, G., Gómez, T., Caballero, R., and Sala-Garrido, R. (2016). Eco-efficiency assessment of wastewater treatment plants using weighted Russel directional distance model. *Journal of Cleaner Production*, 137, 1056-1075.

Molinos-Senante, M., Hernández-Sancho, F., Mocholí-Arce, M., and 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., Hernández-Sancho, F., and Sala-Garrido, R. (2014b). Benchmarking in wastewater treatment plants: a tool to save operational costs. *Clean Technologies and Environmental Policy*, 16, 149-161.

Rohacova, V. (2015). A DEA based approach for optimization of urban public transport system. *Central European Journal of Operations Research*, 23, 215-233.

Rogge, N. and De Jaeger, S. (2013). Measuring and explaining the cost efficiency of municipal solid waste collection and processing services. *Omega*, *41*(4), 653-664.

Ruiz, J. L., and Sirvent, I. (2019). Performance evaluation through DEA benchmarking adjusted to goals. Omega, 87, 150-157.

Sala-Garrido, R., Hernández-Sancho, F., and 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., and Hernández-Sancho, F. (2011). Comparing the efficiency of wastewater treatment technologies through a DEA metafrontier model. *Chemical Engineering Journal*, 173(3), 766–772.

Sala-Garrido, R., Molinos-Senante, M., and Hernández-Sancho, F. (2012b). How does seasonality affect water reuse possibilities? An efficiency and cost analysis. *Resources, Conservation and Recycling*, 58, 125-131.

Schiltz, F., De Witte, K., and Mazrekaj, D. (2019). Managerial efficiency and efficiency differentials in adult education: a conditional and bias-corrected efficiency analysis. *Annals of Operations Research*, 1-18.

Silva, C. and Rosa, M. J. (2015). Energy performance indicators of wastewater treatment: a field study with 17 Portuguese plants. *Water Science & Technology*, 72(4), 510-519.

UN General Assembly (2015). Transforming our world: the 2030 agenda for sustainable development. New York: United Nations.

Vaz, C.B., Camanho, A.S., and Guimarães, R.C. (2010). The assessment of retailing efficiency using network Data Envelopment Analysis. *Annals of Operations Research*, 173(1), 5-24.

Wheelock, D. C., & Wilson, P. W. (2003). Robust nonparametric estimation of efficiency and technical change in U.S. commercial banking. *Unpublished working paper. The John E. Walker Department of Economics, 222 Sirrine Hall, Clemson University, Clemson, South Carolina 29634, USA*.

Wilson, P. (1995). Detecting influential observations in data envelopment analysis. *Journal of Productivity Analysis,* 6(1), 27-45.

Wilson, P.W. (2010). Detecting outliers in deterministic nonparametric frontier models with multiple outputs: Correction. Unpublished working paper, Department of Economics, Clemson University, Clemson, South Carolina 29634 (available on-line at: http://www.clemson.edu/economics/faculty/wilson/Papers/ap-corrected.pdf)

Chapter 6 –

Measuring Productivity Change in Portuguese Wastewater Treatment Plants using an Order-*m* Global Malmquist Index

Chapter 6

Measuring productivity change in Portuguese wastewater treatment plants using an order-*m* global Malmquist index

Overarching research question to be explored with this paper:

RQ4: What additional insights and guidelines can be drawn from a productivity analysis over time respecting a set of WWTPs' activity in a problematic sample?

Measuring productivity change in Portuguese wastewater treatment plants using an order-*m* global Malmquist index

Abstract

Performance measurement is a key component of a good management system. Considering the case of the wastewater sector, wastewater treatment plants (WWTPs) are structural assets deserving increased management attention. In this context, this paper describes the application of a global order-*m* Malmquist Index (GMI_m) to a set of 32 WWTPs managed by a Portuguese water company. The aim is to obtain insights upon productivity change over a four-year period so to gain increased knowledge upon plants' performance as well as understanding the potential causes of productivity variations. The use of the GMI_m further allows to derive information upon overall productivity change, i.e., from the first to the last year, instead of only between two adjacent years, due to the comparison of performance in one given year with the metafrontier that is composed by the performance observations respecting to all years. This paper contributes to the literature of performance measurement of WWTPs through the innovative application of the GMI_m to a real-world case study of a problematic sample of plants that are managed by a Portuguese water utility.

Keywords: Global Malmquist index; Order-m efficiency; Productivity; Data Envelopment Analysis; Wastewater treatment plants; Performance assessment

6.1. Introduction

The wastewater treatment industry is nowadays becoming a prominent focus of research (e.g., Cantor et al., 2021; Chae and Kang, 2013; Eggiman et al., 2017; Gandiglio et al., 2017; Garrido-Baserba et al., 2020; Grigg, 2016; Iribarnegaray et al., 2012; Nowak et al., 2015; Qadir et al., 2020; Sarvari et al., 2021) due to the external drivers of climate change, population growth, and the need to effectively pursuit sustainability worldwide (IEA, 2018; UN General Assembly, 2015). We would like to emphasize here the particular attention that is being given to the area of performance measurement of this industry since it is a recommended practice by the OECD international authority toward increased and better water governance (Akhmouch, 2016; OECD, 2011), which ultimately leads to increased sustainability.

A brief search in the literature related to performance measurement studies within the scope of the wastewater treatment industrial activity delivers a considerable amount of studies that encompass a panoply of specific purposes and methodological approaches (e.g., Abbott and Cohen, 2009; Guerrini et al., 2016; Hernández-Chover et al., 2018; Newhart et al., 2019).

In this paper, we are concerned about analyzing the variations in the productivity observed on a set of wastewater treatment plants (WWTPs) over time. We highlight two specific aspects that motivated our study. First, we had previously conducted an in-depth performance analysis of a very similar set of plants (see Chapter 5 of this thesis) from a management perspective. The insights derived concern the relative efficiency of the plants in a problematic sample of 41 WWTPs managed by a Portuguese water company, and the benchmarking results to guide improvements through best-practices' information sharing. We also derived some insights regarding the contextual conditions that might be influencing the differences observed in the efficiency results. However, such study was based on data from only one year. Therefore, in this paper, we aim at gathering more information about the plants' performance by considering data from four consecutive years, namely to understand whether there is any productivity change over time. Additionally, if any changes in productivity are observed, we would also like to explore their eventual causes. Second, since we are dealing with approximately the same sample of plants, which is a problematic sample, as in the previous study, we are interested in exploring the same methodological approach as for the efficiency measurement (i.e., the order-*m* method), which is also a fundamental issue to obtain robust results in the assessment of productivity change over time. From the information gathered in the literature (Hernández-Sancho et al., 2011; Molinos-Senante et al., 2015; Molinos-Senante et al., 2016a,b), no study has yet been developed that applies an order-*m*-based method for productivity measurement to the case of WWTPs. Therefore, we will explore such methodological approach as an application to the specific case of our sample, following the work developed by Wheelock and Wilson (2003).

This paper contributes to the literature through the application of a global order-*m* Malmquist Index (GMI_m) to a sample of wastewater treatment plants. The global index will allow to directly compare the productivity among any two-yearly period, including the first and the last years observed, since the comparison is being made against the metafrontier that is obtained through considering the data of the entire four-year period observed, so that it is possible to draw a clear vision upon the evolution of the plants over time.

The remainder of the paper is organized as follows: in section 6.2 we will present the methodological steps used for the construction of the global order-*m* Malmquist Index for measuring productivity changes over time. In section 6.3 we present the sample and variables used in the analysis and then the results and discussion are presented in section 6.4. Finally, we present the main conclusions in section 6.5.

6.2. Methodology

In this paper we will use the order-*m* global Malmquist index. The order-*m* method is an original construction developed by Cazals et al. (2002), which lies in the nonparametric area of performance measurement that is typically based on Free Disposal Hull (FDH) and Data Envelopment Analysis (DEA) type technologies, with the additional statistical inference potential that stems from an inherent probabilistic nature. This approach overcomes some drawbacks associated with the deterministic nature of the FDH and DEA estimators.

In order to introduce the order-*m* global Malmquist index formulation as utilized in this paper, we will first refer to the basic concepts behind it, starting with the deterministic DEA estimator of efficiency, followed by a brief introduction to the probabilistic approach that can be assigned to it and which is structural for the order-*m* approach.

6.2.1 The DEA estimator of efficiency

In any given production activity, which transforms inputs (resources) into outputs (products or services), there is a technology that enables the activity to occur. This technology also limits the amount of outputs that is possible to produce from a certain amount of inputs (considering the perspective of an output-oriented case) or limits the minimum amount of inputs required to produce a certain amount of outputs (in an input-oriented case).

In this context, let $x \in \mathbb{R}_+^p$ denote the set of p inputs and $y \in \mathbb{R}_+^q$ represent the set of q outputs of the production activity. The technology is the frontier of the set of all possible combinations of inputs and outputs, which is also known as the production possibility set.

Although the complete (i.e., true) set that represents the technology is not usually known (Cook et al., 2014), it can be symbolically translated by Ψ and formally defined as:

$$\mathrm{Y} = \{(x, y) \in \mathbb{R}_{+}^{p+q} |\ x\ can\ produce\ y\} \tag{6.1}$$

In order to determine the Farrell's efficiency (i.e., radial efficiency), one needs to estimate the frontier of all observations. Considering an input-oriented setting (as it will be explored in this paper), the technology (i.e., the input efficient frontier) allows to derive the input-oriented efficiency score, θ, for a unit operating at the level (x, y), as follows:

$$\theta(x, y) = inf\{\theta|(\theta x|y)\epsilon \mathrm{Y}\} \tag{6.2}$$

Using the DEA envelopment estimator developed by Charnes et al. (1978), in the form of the variable returns to scale (Banker et al., 1984), we are jointly assuming free disposability regarding the feasible combinations of inputs and outputs, and convexity regarding the shape of the technology. In this context, one can nonparametrically estimate the unobserved frontier of the production possibility set using a finite sample of n units, as follows (kneip et al., 1998):

$$\widehat{\mathrm{Y}}_{DEA} = \left\{(x, y) \in \mathbb{R}_{+}^{p+q} |\ x \geq \sum_{\mathrm{j}=1}^{n} l_j x_j\ ;\ y \leq \sum_{\mathrm{j}=1}^{n} l_j y_j\ , for\ some\ (l_1, \ldots, l_n),\right.$$
$$\left. such\ that\!: \sum_{\mathrm{j}=1}^{n} l_j = 1;\ l_j \geq 0, \mathrm{j} = 1, \ldots, n\right\} \tag{6.3}$$

The DEA estimator of efficiency scores for each unit under assessment, $\hat{\theta}_{DEA}(x, y)$, is then obtained by solving the following linear programming model:

$$\hat{\theta}(x, y) = min\left\{\theta > 0|\ \theta x \geq \sum_{j=1}^{n} l_j x_j\ ;\ y \leq \sum_{j=1}^{n} l_j y_j\ \ for\ some\ (l_1, \ldots, l_n),\right.$$
$$\left. such\ that \sum_{j=1}^{n} l_j = 1; l_j \geq 0, \qquad j = 1, \ldots, n\ \right\} \tag{6.4}$$

which is equivalent to substituting Y in equation (6.2) by $\widehat{\mathrm{Y}}_{DEA}$ (Daraio and Simar, 2007).

6.2.2 Probabilistic approach to the definition of the technology and of the efficiency

The equations (6.1) to (6.4) presented in the previous sub-section provide a complete set for the deterministic version of the DEA estimator of efficiency, which is typically used in empirical studies of performance assessment with descriptive purposes. However, within a broader context of performance assessment, for example in the case of a company aiming to analyze the management policy related to its assets, the deterministic method does not hold an inference power. Only a probabilistic approach to the DEA method is able to provide enough reliability in regards to the efficiency results obtained. In fact, the statistical properties of the DEA required for its treatment as of a probabilistic nature were demonstrated by Banker (1993). Furthermore, a clear information upon the data generating process (DGP) associated to a DEA model can be found in the work of Simar (1996). In fact, the core difference between the deterministic and the probabilistic approaches to a frontier framework such as DEA lies on the data generating process (DGP). Accordingly, in the deterministic case, the underlying DGP assumes that all data points belong to the production possibility set with 100% of probability, which means that the reason for any data point to lie below the frontier is pure inefficiency. In other words, the deterministic approach assumes that the assessment conducted is of interest only for the observations considered in the analysis, exactly as they are expressed by the values of the indicators given. Therefore, it holds a descriptive power over the sample of observations at hand. In the probabilistic (or stochastic) approach, the previously mentioned assumption does not hold, which means that there is some probability that some data points may lie outside of the frontier of the production possibility set due to observational errors and random noise (Cazals et al., 2002).

In a deterministic setting, there is a sample of n observations $\left(x_j, y_j\right) \in \mathbb{R}_+^{p+q}, j = 1, \dots, n$ that are comparatively assessed. In a probabilistic setting, the sample of n observations, is able to be transformed, by means of two-dimensional random variables, generally described as (X, Y), into different samples of observations, $[X(x), Y(y)]$, where the functional nature of the two-dimensional random variable is usually omitted (Meyer, 1983). In other words, we can also say that the sample at hand can be resampled (see assumption A1 of the DEA statistical model defined in Kneip et al. (1998). For abbreviation, we will skip the further details associated with the statistical model definition that underlies a probabilistic approach to the DEA framework, and focus only on the aspects that are more relevant to introduce the order-*m* approach. We will also base our explanation on the rationale presented in Daraio and Simar (2007).

As such, in the probabilistic case, the technology, Υ, is defined in terms of the joint probability measure on $(X, Y) \in \mathbb{R}_+^{p+q}$, as opposed to the deterministic case, where it is only defined in terms of the observed $\left(x_j, y_j\right) \in \mathbb{R}_+^{p+q}, j = 1, \dots, n$ (see expression (6.3)). The joint probability measure associated with the two-dimensional random variable (X, Y) is characterized by the joint probability function $H_{XY}(\cdot,\cdot)$. Under the free disposability assumption, this probability function is defined according to the expression (6.5) below.

$$H_{XY}(x, y) = Prob(X \leq x, Y \geq y) \tag{6.5}$$

In other words, $H_{XY}(x, y)$ corresponds to the probability of the observation (x, y) to be dominated.

Here it is important to mention that each two-dimensional random variable (X, Y) can be rather seen as an association of two one-dimensional random variables, i.e., X and Y, which is useful when it is necessary to work with conditional settings (Meyer, 1983). In an input-oriented performance assessment the goal is to maximize the input efficiency for a given level of output. For this reason, the probability function described in expression (6.5) is decomposed into its components, considering the association of (X, Y) with X and Y, as follows:

$$\begin{aligned} H_{XY}(x, y) &= Prob(X \leq x | Y \geq y) Prob(Y \geq y) \\ &= F_{X|Y}(x|y) S_Y(y) \end{aligned} \tag{6.6}$$

where $F_{X|Y}(x|y)$ is the conditional distribution of X given Y, and $S_Y(y)$ is the marginal survivor function of Y.

The input efficiency scores, $\theta(x, y)$, can then be defined in terms of the support of the probabilities shown in equation (6.6), i.e., for $(x, y) \in$ Υ. Therefore, for all y with $S_Y(y) > 0$, we have:

$$\begin{aligned} \theta(x, y) &= inf\left\{\theta | F_{X|Y}(\theta x | y) > 0\right\} \\ &= inf\left\{\theta | H_{X|Y}(\theta x, y) > 0\right\} \end{aligned} \tag{6.7}$$

6.2.3 The order-*m* method for estimating efficiency

Regarding the order-*m* method, introduced by Cazals et al. (2002), it is the probabilistic approach briefly presented in section 6.3.2, which gives the foundation for the corresponding mathematical formulation.

However, the first relevant notion to keep in mind is that this approach has a different perspective into what is considered the benchmark obtained as reference for each DMU in the original sample under assessment. More precisely, contrary to the existence of a full frontier that envelops all the data points (i.e., all the DMUs) as in the FDH or DEA frameworks, in the order-*m* approach there is the concept of order-*m* (partial) frontiers that will be obtained for each DMU and that, in each case, will not envelop all data points, leading to a less extreme measure of performance (Cazals et al., 2002).

In an input-oriented setup, for a given level of output y in the interior of the support of Y, (i.e., for a given observation $(x, y) \in \Upsilon$) we consider now m independent and identically distributed (i.i.d.) random variables X_j, which means following a process of randomly selecting, and with replacement, values of x from the original setting to which correspond a value of output greater or equal than y, with $j = 1, \dots, m$. More precisely, the generation of X_j follow the conditional p-variate distribution function $F_X(x|y) = Prob(X \leq x, Y \geq y)$. In regards to the selection of the more suitable value of m, please refer to Cazals et al. (2002). It is then possible to define the random technology set estimator, as follows:

$$\widetilde{\Upsilon}_m(y) = \left\{(x, y') \in \mathbb{R}_+^{p+q} | \; x \geq X_j, y' \geq y, \mathrm{j} = 1, \dots, m\right\} \tag{6.8}$$

In the same context, the order-*m* input efficiency score is defined, for any $x \in \mathbb{R}_+^p$, as follows:

$$\theta_m(x, y) = E_{X|Y}\left(\tilde{\theta}_m(x, y)|Y \geq y\right) \tag{6.9}$$

where $\tilde{\theta}_m(x, y)$ is the random minimal input efficiency obtained for each DMU when compared to the m DMUs randomly drawn with replacement, according to the condition of having at least the same level of outputs, as follows:

$$\tilde{\theta}_m(x, y) = inf\left\{\theta|(\theta x, y)\epsilon\widetilde{\Upsilon}_m(y)\right\} \tag{6.10}$$

and $E_{X|Y}$ is the expectation relative to the distribution $F_X(x|y)$. In other words, the order-*m* input efficiency for any particular DMU in the sample is given by the expectation of the minimal achievable amount of input.

As pointed out in Cazals et al. (2002), the order-*m* input efficiency given by equation (6.9) can assume values greater than one if the DMU (x, y) under assessment is compared to a set of peers, from the m randomly selected DMUs, that, on average, are less efficient.

Following Daraio and Simar (2007), the order-*m* DEA input efficiency score is obtained based on the technology resulting from the convexification of the original order-*m* technology presented in equation (8), as generically represented by equation (6.11).

$$\widetilde{\Upsilon}_{m,DEA}(y) = CH\left(\widetilde{\Upsilon}_m(y)\right) \quad \textbf{(6.11)}$$

$$= \left\{(x, y') \in \mathbb{R}_+^{p+q} | x \geq \sum_{\mathrm{j}=1}^{m} l_j X_j \,, for\ (l_1, \dots, l_m)\ such\ that \sum_{\mathrm{j}=1}^{m} l_j = 1;\ \ l_j \geq 0, y' \geq y, \quad \mathrm{j} = 1, \dots, m\right\} \quad \textbf{(6.12)}$$

where the random variables X_j are generated by $F_X(x|y)$.

The random order-*m* DEA minimal input efficiency is then given by:

$$\tilde{\theta}_{m,DEA} = inf\{\theta | (\theta x, y) \epsilon \widetilde{\Upsilon}_{m,DEA}(y)\} \quad \textbf{(6.13)}$$

So that the order-*m* DEA input efficiency score for each DMU may be given by:

$$\theta_{m,DEA} = E_{X|Y}\left(\tilde{\theta}_{m,DEA}(x, y) | Y \geq y\right) \quad \textbf{(6.14)}$$

A nonparametric estimator $\hat{\theta}_{m,DEA}(x, y)$, is given by:

$$\hat{\theta}_{m,DEA} = \hat{E}_{X|Y}\left(\tilde{\theta}_{m,DEA}(x, y) | Y \geq y\right) \quad \textbf{(6.15)}$$

which can be approximated using the Monte-Carlo procedure described in Daraio and Simar (2007).

6.2.4 The order-*m* global Malmquist index

In order to measure productivity change over time that accounts for both efficiency causes and technology causes, and without assuming any functional form for the technology, Färe et al. (1992) developed the nonparametric Malmquist productivity index, using a deterministic setting. Later on, Wheelock and Wilson (2003) introduced the order-*m* Malmquist index, which maintains all the previous characteristics, i.e., a non-parametric index to measure productivity changes over time that can be decomposed into the efficiency and technology components, but with the robustness gathered from the statistical support enabled by the probabilistic nature. In this paper, we aim to develop a productivity change analysis within the wastewater treatment industry using a small sample of DMUs. Therefore, we follow the order-*m* based approach of Wheelock and Wilson (2003) and use a global version of it (Pastor and Lovell, 2005) to enable drawing conclusions in terms of the evolution of the productivity from the first period to the last of several periods thanks to the inherent property of circularity.

Let us consider the sets of inputs and outputs vectors, $x^{t_{\mathrm{u}-1}} \in \mathbb{R}_+^p$ and $y^{t_{\mathrm{u}-1}} \in \mathbb{R}_+^q$, respectively, that characterizes a sample of n DMUs in a given time period, $t_{\mathrm{u}-1}$, and the sets $x^{t_\mathrm{u}} \in \mathbb{R}_+^p$, and $y^{t_\mathrm{u}} \in \mathbb{R}_+^q$, that characterizes the same sample of n DMUs in a time period t_u, with $t_\mathrm{u} > t_1$ and $u = 1,2, \dots, z$. The order-*m* global Malmquist index can be written as indicated in equation (6.16).

$$GMI_m\left(x^{t_{u-1}}, y^{t_{u-1}}, x^{t_u}, y^{t_u} | \mathrm{Y}_{m,DEA}^{t_{u-1}}, \mathrm{Y}_{m,DEA}^{t_u}, \mathrm{Y}_{m,DEA}^{M}\right) = \frac{\theta_m(x^{t_u}, y^{t_u})|\mathrm{Y}_{m,DEA}^{M}}{\theta_m(x^{t_{u-1}}, y^{t_{u-1}})|\mathrm{Y}_{m,DEA}^{M}} \quad \textbf{(6.16)}$$

In equation (6.16), the term $\theta_m(x^{t_\mathrm{u}}, y^{t_\mathrm{u}})|\mathrm{Y}_{m,DEA}^{M}$ is the input efficiency score of a DMU, which is characterized by a set of inputs and outputs corresponding to the time period t_u relatively to the metafrontier, i.e., the set of all DMUs in all time periods analyzed, or, in other words, relatively to the reunion of $\mathrm{Y}_{m,DEA}^{t_{\mathrm{u}-1}}$ and $\mathrm{Y}_{m,DEA}^{t_\mathrm{u}}$. The same rationale is applied to the term $\theta_m(x^{t_{\mathrm{u}-1}}, y^{t_{\mathrm{u}-1}})|\mathrm{Y}_{m,DEA}^{M}$.

The index $GMI_m(x^{t_{\mathrm{u}-1}}, y^{t_{\mathrm{u}-1}}, x^{t_\mathrm{u}}, y^{t_\mathrm{u}})$ indicated in equation (6.16) can then be estimated using the order-*m* Malmquist index estimator, $\widehat{M}_{m,n_{t_{\mathrm{u}-1}},n_{t_\mathrm{u}}}(x^{t_{\mathrm{u}-1}}, y^{t_{\mathrm{u}-1}}, x^{t_\mathrm{u}}, y^{t_\mathrm{u}})$, which is obtained by plugging in equation (6.16) the order-*m* input efficiency scores estimator defined in equation (6.15). Notice that this estimator should be built upon the robust order-*m* technology set obtained as the reunion of all technologies in the sample under study that refer to the n DMUs in each of the periods under analysis, $t_{\mathrm{u}-1}$ and t_u.

Similar to what is indicated in Färe et al. (1992), the order-*m* global Malmquist productivity index (GMI_m) can be decomposed into the efficiency change (EC) and an equivalent form of the technology change, which is named best-practice change (BPC). A value equal to one respecting the GMI_m corresponds to a stagnation in terms of productivity, a value greater than one indicates an improve in productivity, and a value inferior to one reveals a decline in productivity. Regarding the EC index, a value greater than one indicates that, on average, the DMUs are closer to the frontier in t_u compared to the period $t_{\mathrm{u}-1}$, a value equal to one indicates that, on average, the DMUs keep their distance to the frontier in both periods, and a value inferior to one indicates a longer distance to the frontier in the period t_u compared to the period $t_{\mathrm{u}-1}$. Finally, the index BPC, as it refers to movements at the frontier level, give indication of an increase in the productivity of the benchmarks identified in t_u in comparison to those identified in $t_{\mathrm{u}-1}$, if it has a value greater than one. The opposite is true if BPC has a value inferior to one. If BPC has a value equal to one it indicates that the frontier remains in the same position in both periods $t_{\mathrm{u}-1}$ and t_u.

The GMI_m decomposition is shown in equation (6.17), following the demonstration presented in Pastor and Lovell (2005).

$$\begin{aligned}
GMI_m^{t_{u-1},t_u} &= \frac{\theta_m(x^{t_u}, y^{t_u})|\mathrm{Y}_{m,DEA}^{t_u}}{\theta_m(x^{t_{u-1}}, y^{t_{u-1}})|\mathrm{Y}_{m,DEA}^{t_{u-1}}} \\
&\quad \times \left\{\frac{\theta_m(x^{t_u}, y^{t_u})|\mathrm{Y}_{m,DEA}^{M}}{\theta_m(x^{t_u}, y^{t_u})|\mathrm{Y}_{m,DEA}^{t_u}} \times \frac{\theta_m(x^{t_{u-1}}, y^{t_{u-1}})|\mathrm{Y}_{m,DEA}^{t_{u-1}}}{\theta_m(x^{t_{u-1}}, y^{t_{u-1}})|\mathrm{Y}_{m,DEA}^{M}}\right\} \\
&= \frac{\theta_m(x^{t_u}, y^{t_u})|\mathrm{Y}_{m,DEA}^{t_u}}{\theta_m(x^{t_{u-1}}, y^{t_{u-1}})|\mathrm{Y}_{m,DEA}^{t_{u-1}}} \times \left\{\frac{\frac{\theta_m(x^{t_u}, y^{t_u})|\mathrm{Y}_{m,DEA}^{M}}{\theta_m(x^{t_u}, y^{t_u})|\mathrm{Y}_{m,DEA}^{t_u}}}{\frac{\theta_m(x^{t_{u-1}}, y^{t_{u-1}})|\mathrm{Y}_{m,DEA}^{M}}{\theta_m(x^{t_{u-1}}, y^{t_{u-1}})|\mathrm{Y}_{m,DEA}^{t_{u-1}}}}\right\} \\
&= EC \times \left\{\frac{BPG^{M,t_u}(x^{t_u}, y^{t_u})}{BPG^{M,t_{u-1}}(x^{t_{u-1}}, y^{t_{u-1}})}\right\} \\
&= EC \cdot \mathrm{BPC} \qquad \textbf{(6.17)}
\end{aligned}$$

The efficiency change component (EC) in equation (6.17) relates the efficiency score obtained by a given DMU in each of the time periods analyzed, $t_{\mathrm{u}-1}$ and t_u, when it is evaluated against the technology of the same time period. The best practice change component (BPC) in equation

(6.17) refers to how each DMU's efficiency evolves over time relative to the metafrontier given this component is calculated as the ratio of the best practice gap (BPG) observed for the DMU under analysis in each time period. In other words, the BPG of a given time period corresponds to the difference in the radial distance that goes from the DMU's position in the input-output space towards the best-practice frontier of that given time period and towards the metafrontier.

In order to compute the $GMI_m^{t_1,t_z}$, being $z > 2$, and considering a general formulation for the time period as t_u, with $u = 1,2,\dots,z$, then equation (6.18), which represents the circular property of the global Malmquist Index, shall be used. Nonetheless, equation (6.17) also apply.

$$GMI_m^{t_1,t_z} = GMI_m^{t_1,t_2} \times GMI_m^{t_2,t_3} \times \dots \times GMI_m^{t_{z-1},t_z} \quad \textbf{(6.18)}$$

6.3. Sample and variables

The framework proposed was applied to a case study of a set of 32 WWTPs managed by the Portuguese water company Águas do Centro Litoral (AdCL). The data refer to a set of four indicators (two inputs and two outputs), according to the indicator definition provided in Chapter 5 of this thesis. The observations refer to a four-year period, from 2015 to 2018 (see Table 6.1).

Table 6.1 – Summary statistics for inputs and outputs (32 WWTPs).

Variable		Mean	SD	Min	Max	Q1	Median	Q3
Input 1 – Energetic balance = Energy used – Energy produced (MWh/year)	2015	230.1	388.4	23.6	1787.6	49.3	122.0	178.7
	2016	272.5	455.0	10.4	2055.8	64.9	141.6	261.8
	2017	322.0	672.2	9.7	3135.8	65.9	140.4	228.3
	2018	315.5	639.3	23.4	2966.5	66.2	145.0	222.8
Input 2 – Number of full-time equivalent workers	2015	0.84	1.42	0.20	7.00	0.38	0.45	0.60
	2016	0.84	1.42	0.20	7.00	0.38	0.45	0.60
	2017	0.84	1.42	0.20	7.00	0.38	0.45	0.60
	2018	0.84	1.42	0.20	7.00	0.38	0.45	0.60
Output 1 – COD actually removed = [(COD concentration in the raw wastewater – COD concentration in the effluent)/COD concentration in the raw wastewater] x volume of wastewater treated (m^3/year)	2015	820739	2507608	5104	12460949	63932	117696	280983
	2016	1057647	3144707	12565	15658023	81304	179197	372141
	2017	844421	2591045	9224	12731234	47361	136974	249234
	2018	1071357	3121001	14676	15604927	74842	162870	366300
Output 2 – SS actually removed* (m^3/year)	2015	838565	2569955	4860	12875260	65073	119517	266852
	2016	1097115	3260349	12108	16351689	81730	182108	377453
	2017	872690	2677860	8709	13169622	49875	139725	261249
	2018	1118229	3261508	17792	16392406	76067	166305	373904

*Indicator built on the basis of the same rationale as Output 1.

6.4. Results and discussion

This section is composed of three subsections. First, we will present and discuss the results of order-m efficiency scores that constitute the basis for the subsequent analysis of the performance behavior of the sample of WWTPs over the four yearly periods of time considered, i.e., from 2015 to 2018. Second, we will present and discuss the results of the order-m global Malmquist index to deliver information upon the productivity changes, including decomposing those productivity changes into their efficiency and technology components, to identify clear trends of performance over time and see what relevant information can be further extracted from the analysis so to be considered in future management and policy developments. All calculations related to these two subsections were conducted in the R software. Third, we will analyze the hypothesis that relates the yearly precipitation to the productivity changes observed over the years.

6.4.1 Order-*m* efficiency results

As explained in section 6.2.3, the order-m efficiency scores of each WWTP in the 32-units sample of AdCL were estimated according to equation (6.15), using the Monte-Carlo procedure described in Daraio and Simar (2007). It was considered a B value of 2000 and a m value equal to the size of the sample, following the same rationale used in Chapter 5 of this thesis since the sample analyzed in this paper is also considered problematic. The results obtained are displayed in Table 6.2.

Table 6.2 – Order-m efficiency scores.

Period	Mean	SD	Min	Max	# Efficient WWTPs ($\theta_m \geq 1$)[8]
2015	0.936	0.173	0.541	1.376	19
2016	0.939	0.218	0.516	1.588	20
2017	0.961	0.229	0.530	1.629	19
2018	0.926	0.177	0.514	1.271	16

When analyzing Table 6.2, we see that the mean values of the order-m efficiency scores in each of the four years assessed denote a quite homogenous performance among WWTPs. Two comments can immediately be drawn from this observation. The first observation stems from the probabilistic nature of the order-m method. Also, the order-m method allows for mitigating the impact of outliers by integrating the condition relative to the output level. This condition excludes the comparison of the DMU under assessment with those DMUs that could exhibit a biased extraordinary efficient behavior due to their relatively small size (in an input-oriented perspective). The type of results obtained can also be linked to the small size of the sample.

Secondly, after mitigating the effects of the aspects that could mostly distort the comparative analysis among the set of WWTPs by using the order-m method, it is plausible to assert that the plants in the sample demonstrate similar performance behavior among them, since they are

[8] We refer the reader to the paper of Daraio and Simar (2005) for the explanation related to the fact that the order-m efficiency score is not bounded by one.

managed by the same entity and because in terms of the type of secondary treatment technology the plants are also quite similar.

Nonetheless, it is observed that in the year of 2017 the plants exhibit quite a more homogeneous behavior compared to the other years, as indicated by the highest level of average order-*m* efficiency score, even though the number of benchmarks is quite similar to the previous years. On the contrary, the results for the year 2018 show that the plants exhibit a slightly worse performance than in the remaining years, which is also supported by the smaller number of benchmarks observed. We will come back to these results ahead in this discussion section to identify potential reasons for those observations.

6.4.2 Order-*m* global Malmquist index results

Regarding the productivity change observed in the same period of time, i.e., from 2015 to 2018, the order-*m* global Malmquist index was estimated according to the description provided in section 6.2.4. The corresponding results are displayed in Table 6.3, including the values of the efficiency and technology components of the GMI_m. We further identify, on Table 6.4, the number of WWTPs that demonstrated whether an increase, a decrease or a stagnation of the MI_m and its components to better understand the overall relative behavior of the plants.

An important note regarding the calculation of the GMI_m needs to be highlighted here. As the GMI_m is calculated based on the four different components identified in equation (6.16), special attention is required when calculating the efficiency of the DMUs in relation to the metafrontier due to the *m*-value selection. In this case, the *m*-value cannot be the same as when calculating the within-period efficiency due to the $n \times z$ dimension of the sample, where z is the number of periods. Figure 6.1 shows the analysis conducted for the selection of the most adequate *m*-value, according to the procedure proposed by Cazals et al. (2002). Based on this analysis, it was selected a *m*-value equal to 80 since it is a good representation of the elbow point in all four curves represented.

Table 6.3 – Order-*m* global Malmquist index, efficiency change, and best practice change.

Period	**Order-*m* global Malmquist index (GMI_m)**				**Efficiency change (EC)**				**Best practice change (BPC)**			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
2015-2016	1.106	0.270	0.627	1.994	1.006	0.159	0.663	1.406	1.091	0.130	0.946	1.425
2016-2017	0.898	0.147	0.524	1.140	1.030	0.108	0.673	1.269	0.877	0.142	0.478	1.047
2017-2018	1.102	0.321	0.487	2.503	0.985	0.150	0.542	1.453	1.126	0.305	0.789	2.290
Overall (2015-2018)	1.045	0.247	0.555	1.711	1.011	0.219	0.558	1.602	1.032	0.131	0.713	1.427

Table 6.4 – Number of WWTPs that maintained, decreased, and increased productivity over time.

Period	**# Maintained**	**# Decreased**	**# Increased**
2015-2016	3	10	19
2016-2017	2	22	8
2017-2018	2	11	19
Overall (2015-2018)	3	13	16

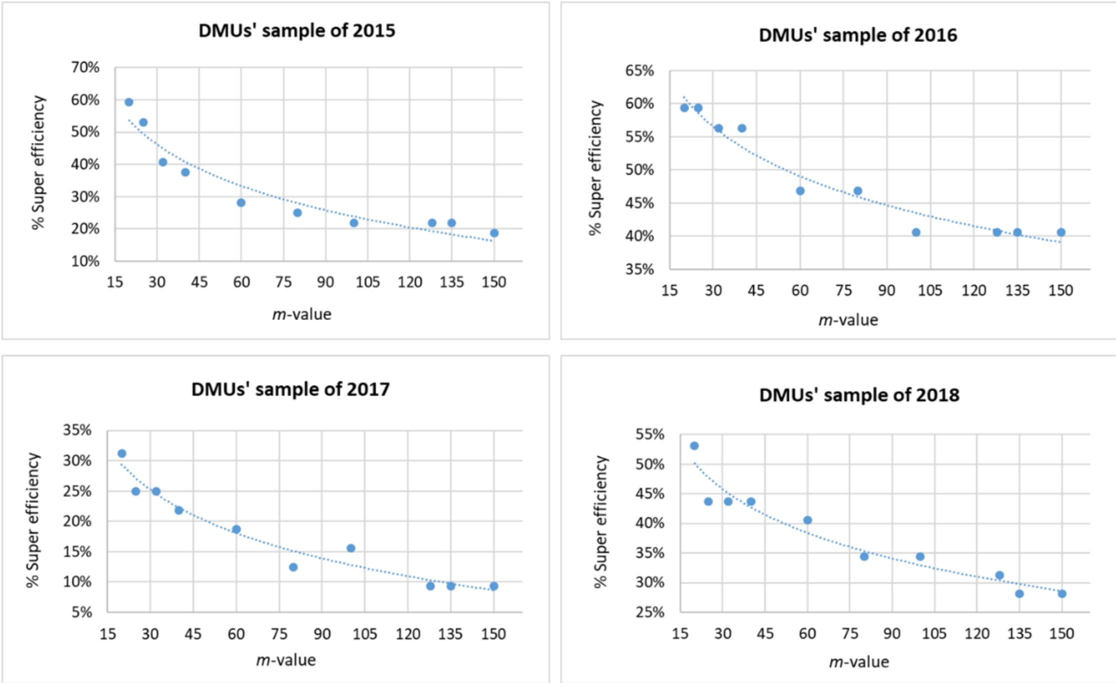


Figure 6.1 *– Evolution of the percentage of superefficient DMUs with m.*

Through the analysis of Table 6.3, one can observe that only between the period of 2016 and 2017 there was a decrease in the sample's average productivity since the value of the GMI_m is lower than one. In the remaining time periods the sample's average productivity increased. However, what is interesting to observe is that the reason for the productivity decline was a technology decline, instead of the more natural efficiency decline.

The year of 2017 was particularly atypical in regards to the global precipitation (i.e., much lower than in common years as demonstrated in Table 6.5), which may explain why the majority of WWTPs showed similar negative trends in the productivity variations between 2016 and 2017. The normal functioning conditions of the plants were reestablished immediately in the following period, when the global precipitation became usual, as it is confirmed by the value of the GMI_m being again greater than one between 2017 and 2018 due to an increase in the technology component.

A productivity decline due to a technology decline from a rainy year to a less rainy year means that the benchmarks defining the efficient frontier are not able to perform in less rainy years as efficiently as in more rainy years compared to the metafrontier. The total number of plants that decreased their productivity is also the highest in 2017, as indicated in Table 6.4.

Through the analysis of the input-output indicators setting described in Table 6.1, one possible reason for the decline in the benchmarks' efficiency is a lower amount of output attained in the productive activity of wastewater cleansing during less rainy years than during more rainy years due to changes in the wastewater volume. Note that the wastewater volume is a component of the output indicators used in the DEA model used to assess the efficiency of the plants. This leads to a perception of technology decline when going from a high rainy year to a low rainy year derived from the need to cope with a lower amount of output, while keeping approximately unchangeable the level of resources due to technological reasons.

The reasoning discussed above leads us to hypothesize that, in fact, the variations in precipitation from one year to the other have an impact on the productivity changes eventually observed. Another interesting thing reaffirming this hypothesis is that in the year of 2017 the efficiency component is the highest of the four yearly periods, giving the indication that with less rain in the sewage system (as happened in 2017), the plants are able to perform more homogeneously. This result is displayed on Table 6.2 of relative efficiencies. The hypothesis here described regarding the potential impact of precipitation in the WWTP's performance will be tested in the following section.

Before going forward into the analysis of the hypothesis identified, we highlight the fact that the overall productivity (i.e., from 2015 to 2018) of the sample of WWTPs under study have increased (GMI_m = 1.045), and this increase was both due to technology (BPC = 1.032) and efficiency (EC = 1.011) improvements. This is a positive result for the company since it demonstrates that, despite any exception on performance due to external conditions, in the overall the benchmarks are making progress in their performance over the years and the remaining plants are also demonstrating to be able of equally improving their performance in a similar or even better way along time.

6.4.3 Testing the hypothesis of the impact of precipitation on WWTPs' performance variation

This section is organized in two stages. In the first stage, we will conduct a verification process of a set of five complementary hypotheses, which underly the core hypothesis that precipitation has a determinant impact on WWTPs' activity observed over the years under analysis. In a second moment, we will test the core hypothesis of the impact of precipitation on WWTP's performance variation.

The five complementary hypotheses are the following:

1) The precipitation data used, which is aggregated at the region level (within Portugal), is adequate for the analysis of the impact of precipitation in the WWTPs' performance variation over time.
2) The precipitation varies significantly over the years.
3) The volume of the wastewater treated, which is one component of the DEA model's output indicators, varies significantly over the years.
4) The concentration of pollutants (i.e., COD and SS) in the raw wastewater vary significantly over the years.
5) The energy balance is significantly different over the years.

In regards to the first hypothesis above indicated, it was used the data freely available at PORDATA[9] respecting the Portuguese total precipitation. Considering the Mainland Portugal, the data were only available for seven regions, corresponding to districts in Portugal, including the North, Center, and South of the Country (Viana do Castelo, Bragança, Porto, Castelo Branco, Lisboa, Beja, and Faro).

Therefore, it was necessary to verify if the differences observed in precipitation from one year to the subsequent year, in each of those regions, were identical for all regions, so to be possible to assume that differences in the AdCL plants' locations, which correspond to districts that are not included in those above mentioned, were equally identical in the same time interval. This is the core information needed to support the average precipitation data associated with each year (see the last column of Table 6.5) can be used as reference for the main precipitation hypothesis verification.

[9] www.pordata.pt

It was considered the data that were simultaneously available for all regions, which corresponded from the year 1970 to 2011 and from 2015 to 2019, and then run a one-way ANOVA test. The *p*-value found was 0.9844, meaning that there are no differences among Portuguese regions in terms of differences of precipitation felt from one year to the subsequent year. Therefore, it is safe to consider the indication of the average yearly precipitation of all regions above mentioned as a guideline on the yearly precipitation felt in AdCL plants' locations. Table 6.5 displays the data used as guideline for the analysis presented on this paper. Figure 6.2 is a visual representation that supports the validity of the hypothesis considered, using ten adjacent yearly intervals.

Table 6.5 – Total precipitation (mm) in Portugal Mainland by region.

Year	Viana do Castelo	Bragança	Porto	Castelo Branco	Lisboa	Beja	Faro	**Mean**
2015	1139.5	593.8	994	471	469.6	415	343	**632.3**
2016	1416.8	971	1529.1	840.5	894.8	728	475.7	**979.4**
2017	837.8	555	816.7	419.5	505.1	434.4	312.6	**554.4**
2018	1411.2	922.6	1014	795	767.1	589.8	492	**856.0**
1970-2011 (average)	1420.4	758.2	1136.1	762.2	759.5	560.8	508.8	**843.7**

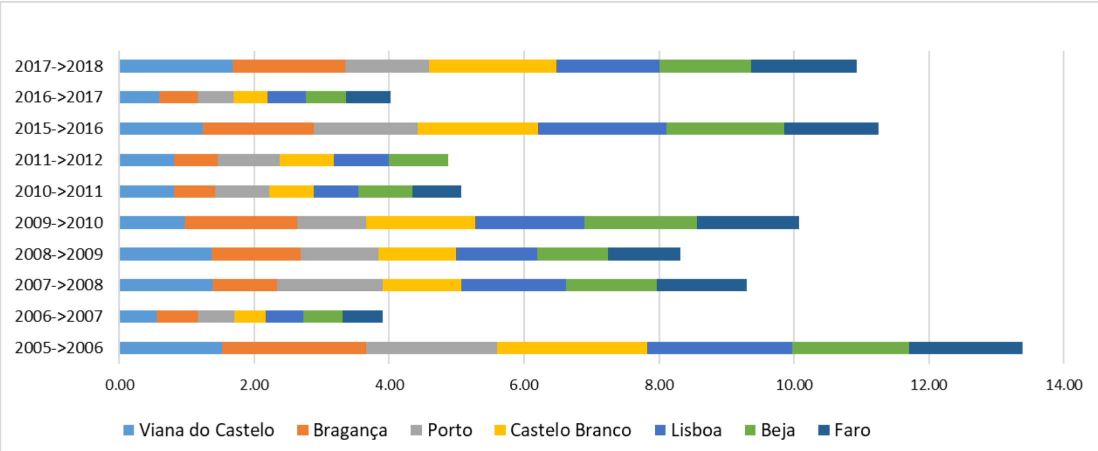


Figure 6.2 *– Relative variation of annual precipitation.*

In regards to the second complementary hypothesis listed above, it was verified for the years 2016, 2017, and 2018, as demonstrated in Table 6.6 (note that the null hypothesis was that there were no differences in precipitation among years). Particularly, the year of 2017 was significantly less rainy than 2016 and also than 2018, although there are no significant differences in the precipitation occurred in 2015 and in 2016 or in the remaining years. An interesting aspect that matches these observations is depicted in Figure 6.3, where the average efficiency scores are plotted against the precipitation over the analyzed years. Both variables demonstrate an opposite trend over time whenever the precipitation is significantly different (i.e., between the years of 2016 and 2017), reflecting the fact that lower precipitation leads to more homogenous performance, which supports the observation already pointed out in section 6.4.2.

Table 6 – Results of the hypothesis test analysis to verify significant differences among the annual precipitation in Portugal.

Group	Average total precipitation in Portugal[10] (mm)	Test p_value
2015	632.3	
2016	979.4	0.054
2017	554.4	
2018	856.0	
2015	632.3	0.083
2016	979.4	
2016	979.4	0.021
2017	554.4	
2017	554.4	0.049
2018	856.0	

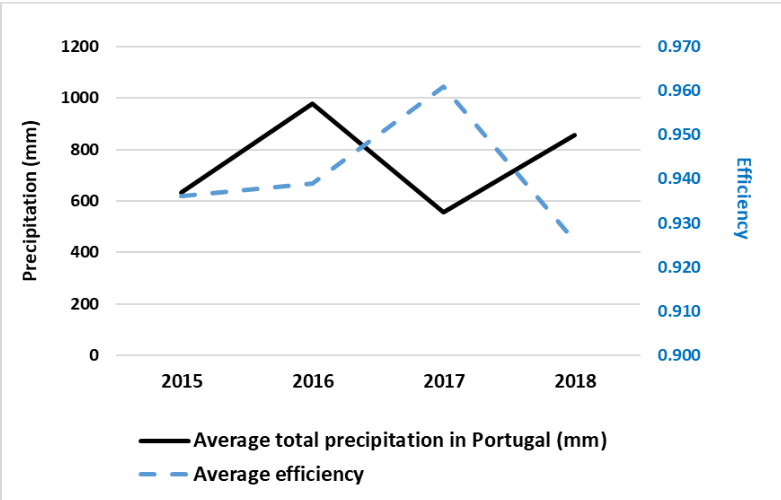


Figure 6.3 *– Relation between precipitation and efficiency.*

The third complementary hypothesis was proved incorrect. As such, we must admit that the volume of the wastewater does not change significantly over the years. Nonetheless, Figure 6.4 demonstrates a visually similar variation pattern of this DEA output variable component with the total precipitation occurred over the years.

[10] www.pordata.pt

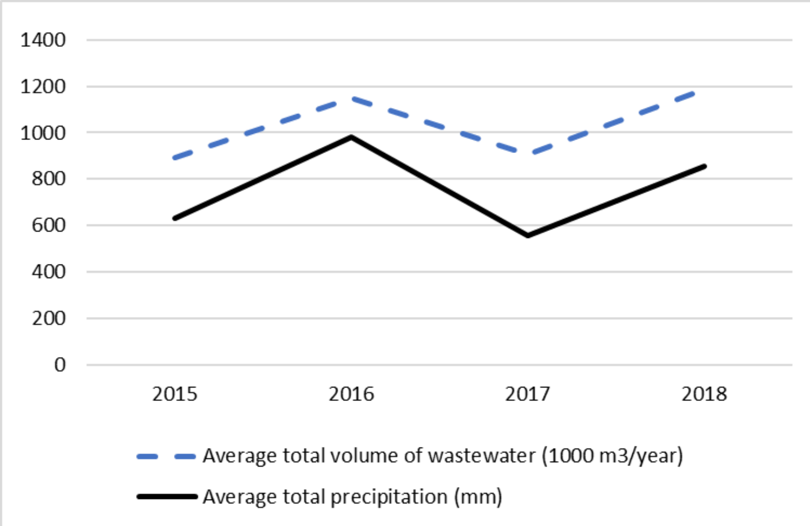


***Figure 6.4** – Relation between the volume of wastewater and precipitation.*

The fourth complementary hypothesis was proved correct, as shown in Table 6.7 (note that the null hypothesis considered was that there were no differences among groups). Figure 6.5 further denotes a coherent and opposite trend between the evolution of the concentration of pollutants in the raw wastewater and the average total precipitation over the years.

Table 6.7 – Results of the hypothesis test analysis to verify significant differences among the concentration of pollutants in the raw wastewater over the years.

Variable	Group	Average concentration of pollutants	Average precipitation (mm)	Test *p*_value
COD (gO_2/m^3)	2015	731.8	632.3	0.016
	2016	560.7	979.4	
	2017	814.6	554.4	
	2018	558.2	856.0	
SS (g/m^3)	2015	386.9	632.3	0.005
	2016	240.6	979.4	
	2017	367.9	554.4	
	2018	239.0	856.0	

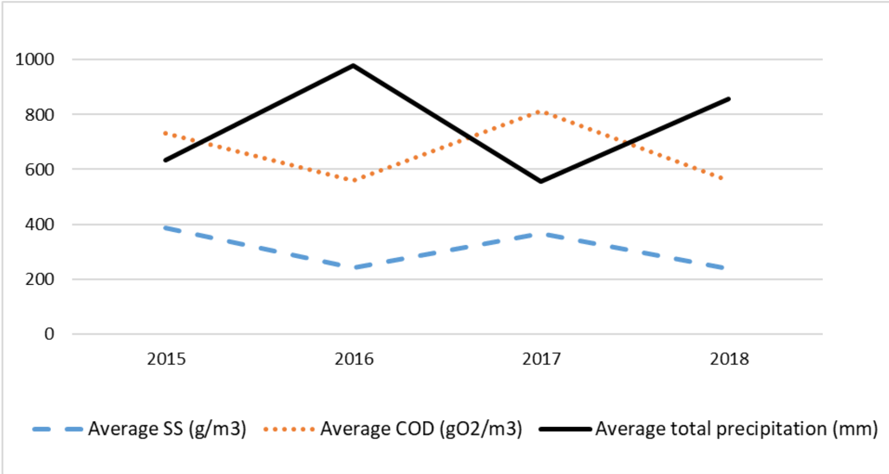


Figure 6.5 *– Relation between the concentration of pollutants in the raw wastewater and precipitation.*

Furthermore, the correlation coefficient between the volume of wastewater treated and the precipitation occurred in the same period is 0.914, indicating a clear association between both variables. The same happens in terms of the association between the output indicators of the DEA model and the precipitation over the 4 years analyzed, with a correlation factor of 0.927 (in the case of COD actually removed) and 0.916 (in the case of SS actually removed).

Figure 6.6 shows, in the same graphic representation, two components of the output indicator (i.e., concentration of pollutants in the raw wastewater and volume of wastewater), the output indicators and the precipitation occurred over the years, emphasizing all that had been said in terms of the observed trends in variations over time. Note that the output indicators of the DEA model demonstrate the higher similarity with the precipitation pattern.

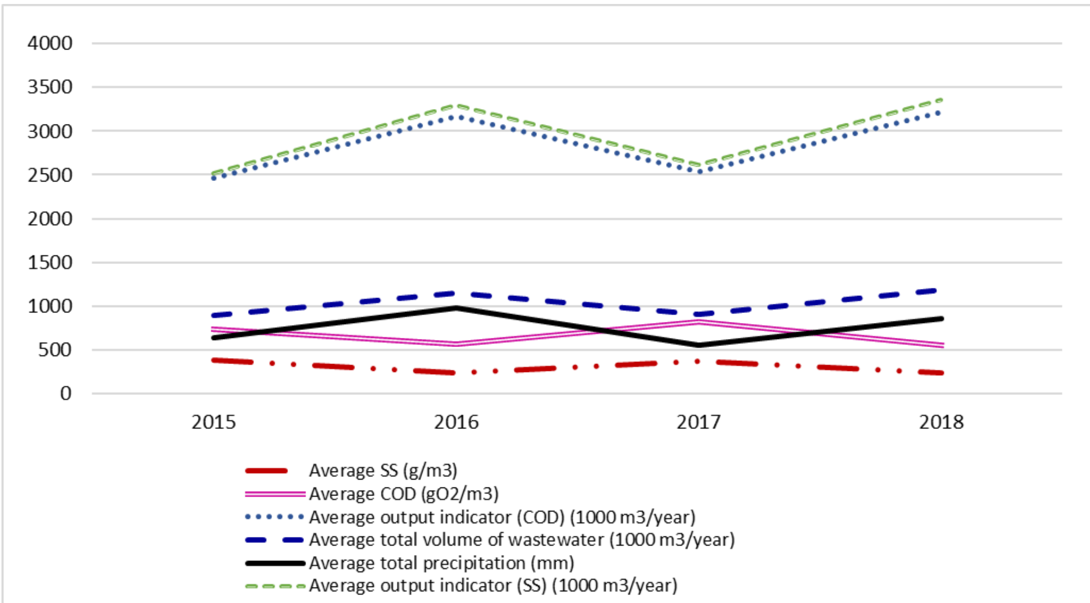


Figure 6.6 *– Relation between the output indicators, volume of wastewater treated, concentration of pollutants in the raw wastewater and precipitation.*

The fifth complementary hypothesis is that the energy balance (see Table 6.1) is significantly different over the years. However, this hypothesis was proved incorrect since the *p*-value of the hypothesis test was 0.902. Figure 6.7 further highlights that there are no apparent differences in the energy balance of the plants over the years and also that, on average, the energy balance of the plants is not affected by variations in precipitation.

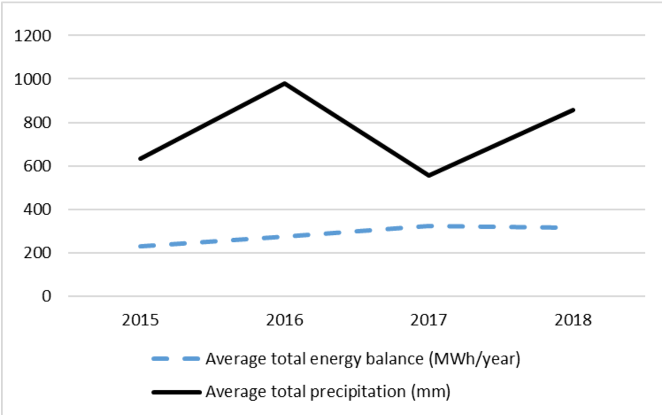


Figure 6.7 *– Relation between the energy balance of WWTPs and precipitation.*

From the above discussion, we find that precipitation has a significant influence only in the concentration of pollutants in the raw wastewater, whilst also demonstrating an association with the volume of wastewater treated. This leads us to conclude that there is a relation between precipitation and the output indicators of the empirical model.

The reason to not consider directly an analysis between the DEA output indicators and precipitation was to allow a discrimination of the impact of precipitation in each of the two type of components that make up the output indicators. In other words, assuming that the absolute amount of pollutants is constant over the years, the precipitation could have a significant impact only on the concentration of pollutants in the raw wastewater (as in fact was verified) and not in the total amount of the wastewater volume (which was also verified), being the wastewater volume the component that gives the order of magnitude to the output indicators. Without this separate analysis of the DEA model's output components important information could be diluted in the output indicator and therefore overlooked.

Any variations found in the energy balance of the plants (input indicator of the empirical model) have no clear relationship with precipitation. Therefore, we can assert that the differences observed in the GMI_m are due to the impact that precipitation has on the output indicators. Figure 6.8 provides a visual representation of the observed trends over time to support this affirmation.

In the overall, we have observed that, from one side, efficiency reaches the highest value in the less rainy year, which means a more homogeneous behavior of the plants when there is lower amount of precipitation. From the other side, productivity declines in the same less rainy year. As productivity can be decomposed into an efficiency change and a technology change component, it is in fact the technology change component that is driving the productivity decline. This information is also shown in Table 6.3. Therefore, we can conclude that, although on average WWTPs are more efficient in the less rainy year, the benchmarks are not able to perform as efficiently, compared to the metafrontier, in the less rainy year. This means that the reduction observed in the total amount of output attained in the less rainy year (as observed in Figure 6.6), even if not significantly lower than in other years, is not accompanied by an equally reduction in the amount of resources, with particular emphasis on energy, in the same period. This is an important information that decision-makers can retrieve to improve the management of their plants. Particularly, as the fraction of the wastewater volume that is due to rain cannot be directly managed by the water company in the short term, the core matter would be to find ways of adjusting the energy consumption to the real concentration of pollutants in the wastewater, which was demonstrated to have a clear correlation with precipitation. On the other hand, this information can also be used at a policy level to ensure that the rain water is not captured by the sewage system in a longer run.

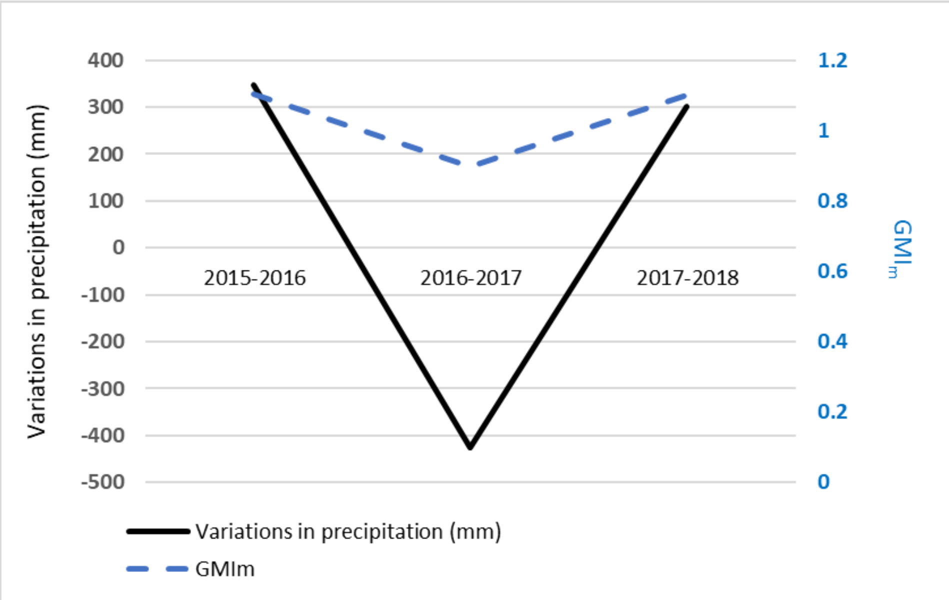


Figure 8 *– Relation between the evolution of GMI and variations in precipitation.*

6.5. Conclusions

Performance measurement is an essential component of a successful management system within an organization. In this paper, we demonstrate that the application of a performance measurement exercise can give important insights into the relative performance of the units within an organization so to contribute to the enhanced understanding of such system of units and its behavior over time. In this particular case, we applied a global order-*m* Malmquist Index to a set of wastewater treatment plants (WWTPs) to measure productivity changes over a period of four years. The results show that these WWTPs' productivity is, on average, affected by the total amount of precipitation felt in each year, demonstrating a decline on the productivity in particularly low rainy years compared to normal rainy years. The observed decline on productivity was due to a technological decline. We explored the hypothesis of the impact of the yearly amount of rain on productivity, and we found that it is correlated with the output indicators of the empirical model used to characterize the activity of the plants, namely because the wastewater volume is impacted by the rain, which contributes for a dilution in the concentration of the pollutants to be removed. Therefore, it was concluded that the technology decline observed was due to the fact that the benchmarks identified in a rainier year are not able to adjust their energy input (for example, due to technical reasons) to a correspondent lower amount, so as to cope with the conditions faced in less rainy years. This indicates that, in the short run, technical and operational adjustments could be sought to avoid identical declines in the productivity in face of identical weather conditions. Furthermore, this result gives the suggestion that, in general, the plants are designed to cope with higher volumes of wastewater, which in fact should not be necessary if the rain water was not being mixed with the sewage. Therefore, in the long run, policy adjustments could be advisable in regards to the rain urban water management, following the information retrieved from this practical example.

Nonetheless, it was also found that in the overall, i.e., from the first analyzed year of 2015 to the last of the four-year period, 2018, there was a productivity increase, both due to technological and efficiency improvements, which demonstrates the positive performance of the majority of the plants in the analyzed sample. This could also serve as an incentive to keep on going with the current practices but also to grow in the ambition of pursuing even higher performances, by investing in innovation and to further engage the work-force toward continuous improvement and sustainability gains.

A similar analysis to the one conducted in this paper could be applied to a sample of more plants within the same company to derive even more in-depth insights regarding the overall plants' performance. Moreover, the assessment could be conducted at a cross-company level so to increase the benefits that can be accrued with a benchmarking exercise and contribute to transparency and stakeholder engagement within the sector.

References

Abbott, M., Cohen, B. (2009). Productivity and efficiency in the water industry. *Utilities Policy*, 17, 233-244.

Akhmouch, A. (2016). The 12 OECD principles on water governance – When science meets policy. *Utilities Policy*, 43, 14-20.

Banker, R. D. (1993). Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation. *Management Science*, 39(10), 1265-1273.

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.

Cantor, A., Sherman, L., Milman, A., Kiparsky, M. (2021). Regulators and utility managers agree about barriers and opportunities for innovation in the municipal wastewater sector. *Environmental Research Communications*, 3, 031001.

Cazals, C., Florens, J.-P., & Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106, 1-25.

Chae, K. J., Kang, J. (2013). Estimating the energy independence of a municipal wastewater treatment plant incorporating green energy resources. *Energy Conversion and Management*, 75, 664-672.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429-444.

Cook, W., Tone, K., Zhu, J. (2014). Data envelopment analysis: prior to choosing a model. (Technical note), *Omega*, 44, 1-4.

Daraio, C., & Simar, L. (2007). Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. *Journal of Productivity Analysis*, 28, 13-32.

Eggimann, S., Mutzner, L., Wani, O., Schneider, Y., Spuhler, D., de Vitry, M. M., Beutler, P., Maurer, M. (2017). The potential of knowing more: a review of data-driven urban water management. *Environmental Science & Technology*, 51, 2538-2553.

Färe, R., Grosskopf, S., Lindgren, B., & Roos, P. (1992). Productivity changes in Swedish Pharmacies 1980-1989: A non-parametric Malmquist approach. *The Journal of Productivity Analysis*, 3, 85-101.

Gandiglio, M., Lanzini, A., Soto, A., Leone, P., Santarelli, M. (2017). Enhancing the energy efficiency of wastewater treatment plants through co-digestion and fuel cell systems. *Frontiers in Environmental Science*, 5(70).

Garrido-Baserba, M., Corominas, L., Cortés, U., Rosso, D., Poch, M. (2020). The fourth revolution in the water sector encounters the digital revolution. *Environmental Science & Technology*, 54, 4698-4705.

Grigg, N. S. (2016). The business of water in a changing world: organizations, connectors and support sector. *International Journal of Water Resources Development*, 32(5), 708-720.

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.

Hernández-Chover, V. Bellver-Domingo, A., Hernández-Sancho, F. (2018). Efficiency of wastewater treatment facilities: the influence of scale economies. *Journal of Environmental Management*, 228, 77-84.

Hernández-Sancho, F., Molinos-Senante, M., Sala-Garrido, R. (2011). Techno-economical efficiency and productivity change of wastewater treatment plants: the role of internal and external factors. *Journal of Environmental Monitoring*, 13, 3448-3459.

IEA (International Energy Agency) (2018). World Energy Outlook.

Iribarnegaray, M. A., Copa, F. R., D'Andrea, M. L. G., Arredondo, M. F., Cabral, J. D., Correa, J. J., Liberal, V. I., Seghezzo, L. (2012). A comprehensive index to assess the sustainability of water and sanitation management systems. *Journal of Water, Sanitation and Hygiene Development*, 2, 205-222.

Kneip, A., Park, B. U., & Simar, L. (1998). A note on the convergence of nonparametric DEA estimators for production efficiency scores. *Econometric Theory*, 14, 783-793.

Meyer, P. L. (1983). Probabilidade - Aplicações à Estatística. Rio de Janeiro: *LTC - Livros Técnicos e Científicos Editora S.A*.

Molinos-Senante, M., Hernández-Sancho, F., Mocholí-Arce, M., Sala-Garrido. (2016a). Productivity growth of wastewater treatment plants – accounting for environmental impacts: a Malmquist-Luenberger index approach. *Urban Water Management*, 13(5), 476-485.

Molinos-Senante, M., Hernández-Sancho, F., Sala-Garrido, R. (2015). Comparing the dynamic performance of wastewater treatment systems: a metafrontier Malmquist productivity index approach. *Journal of Environmental Management*, 161, 309-316.

Molinos-Senante, M., Sala-Garrido, R., Hernández-Sancho, F. (2016b). Development and application of the Hicks-Moorsteen productivity índex for the total factor productivity assessment of wastewater treatment plants. *Journal of Cleaner Production*, 112, 3116-3123.

Newhart, K.B., Holloway, R.W., Hering, A.S., Cath, T.Y. (2019). Data-driven performance analyses of wastewater treatment plants: a review. Water Research, 157, 498-513.

Nowak, O., Enderle, P., Varbanov, P. (2015). Ways to optimize the energy balance of municipal wastewater systems: lessons learned from Austrian applications. Journal of Cleaner Production, 88, 125-131.

OECD (2011). Water in OECD countries: a multi-level approach – highlights.

Pastor, J. T., Lovell, C. A. K. (2005). A global Malmquist productivity index. *Economics Letters*, 88, 266-271.

Qadir, M., Drechsel, P., Cisneros, B. J., Kim, Y., Pramanik, A., Mehta, P., Olaniyan, O. (2020). Global and regional potential of wastewater as a water, nutrient and energy source. Natural Resources Forum, 44, 40-51.

Sarvari. H., Chan, D.W.M., Banaitiene, N., Noor, N.M., Beer, M. (2021). Barriers to development of private sector investment in water and sewage industry. Built Environment Project and Asset Management, 11(1), 52-70.

Simar, L. (1996). Aspects of statistical analysis in DEA-type frontier models. *The Journal of Productivity Analysis*, 7, 177-185.

UN General Assembly (2015). Transforming our world: the 2030 Agenda for Sustainable Development. A/RES/70/1, available at: https://www.refworld.org/docid/57b6e3e44.html [accessed 12 July 2021].

Wheelock, D. C., & Wilson, P. W. (2003). Robust nonparametric estimation of efficiency and technical change in U.S. commercial banking. Unpublished working paper. *The John E. Walker Department of Economics, 222 Sirrine Hall, Clemson University, Clemson, South Carolina 29634, USA.*

Chapter 7 –

A System-Level Optimization Framework for Efficiency and Effectiveness Improvement of Wastewater Treatment Plants

Chapter 7

A system-level optimization framework for efficiency and effectiveness improvement of wastewater treatment plants

This chapter corresponds to a paper that is currently (i.e., by the time the thesis is being concluded) under review by the journal *International Transactions in Operational Research*.

Overarching research question to be explored with this paper:

RQ5: How to conduct a performance measurement of a set of WWTPs aiming to integrate improvements both in resource's use efficiency and treatment effectiveness? What insights and guidelines can be derived from such exercise?

Authors:

Ana S. Camanho, Flávia Barbosa, Alda Henriques

Thesis' author contribution:

Conceptualization (collaboration), data curation, investigation (collaboration), writing – original draft (section 1, section 2, section 4 only partially, and section 6), writing – review and editing (collaboration), funding acquisition (collaboration).

A system-level optimization framework for efficiency and effectiveness improvement of wastewater treatment plants

Abstract

Wastewater treatment plants constitute an essential part of the sewage system. They have the role of removing pollutants from wastewater to enable the safe disposal of the treated effluent in the natural environment. This research seeks to evaluate plants' efficiency and effectiveness, which involves minimizing energy consumption while obtaining a quality level of the treated water aligned with legislation requirements. We explore two policy scenarios regarding the measurement of effluent quality. The first assumes that pollutants' emission quotas are fixed at each plant. The second assumes that quotas are set for the receiving waters (e.g., river or watercourse in the natural environment) so that trade-offs in emission quotas among plants sharing the same discharge site are possible. This latter policy scenario requires a system-wide analysis to identify optimal targets for pollutants removal at each plant that allow fulfilling the expected average quality levels of the effluent discharged. This paper develops a methodology to fully realize the potential for energy savings based on an innovative Mixed-Integer Linear Programming model. This model follows the Data Envelopment Analysis axioms to estimate the frontier of the production possibility set. The approach proposed is tested in a real-world context using the plants of a Portuguese water company. The results show that the two scenarios combining efficiency and effectiveness perspectives have advantages in terms of energy savings compared to the conventional situation focused only on efficiency gains. The saving potential is slightly higher in the scenario allowing reallocation of emission quotas among plants.

Keywords: Data Envelopment Analysis, Efficiency, Effectiveness, Wastewater treatment plants, System level optimization

7.1. Introduction

This paper focuses on the activity of wastewater treatment plants (WWTPs), which are the infrastructures that carry out the process of treating wastewater to remove pollutants. Among other resources (inputs of the process), wastewater treatment requires energy to remove different types of pollutants from the wastewater. The objective is to obtain an effluent that matches the quality criteria defined by regulators. The maximum level of pollutants allowed in treated water should respect the environmental requirements of the discharging area (Directive, 1991), corresponding to the natural watercourse that is associated to each WWTP (Sala-Garrido et al., 2012a). A detailed description of the functioning of wastewater treatment plants, including the process of wastewater treatment, the alternative technologies, as well as the input-output framework that is typically used in efficiency studies is available in Gandiglio et al. (2017); Guerrini et al. (2016); Sala-Garrido et al. (2011).

In the performance measurement literature, the efficiency of wastewater treatment activity usually adopts an input-oriented perspective. The objective of the studies is to explore the potential for resource-savings at each WWTP, given the amount of pollutants removed from the wastewater (Castellet and Molinos-Senante, 2016; Hernández-Sancho et al., 2011a; Molinos-Senante et al., 2014b).

The total amount of pollutants removed from the wastewater depends on the volume of water treated, the level of pollutants in the influent wastewater, and the legal requirements for the maximum level of pollutants allowed in the effluent, as noted by Dong et al. (2017). However, the dimension associated with the fulfillment of the quality criteria for the effluent water has been disregarded in most studies in the wastewater sector. Consequently, previous efficiency assessments of WWTPs provide an incomplete view of the plants' activity. Guerrini et al. (2016) highlighted the importance of considering the quality criteria in the performance evaluation of WWTPs, but this study only tackled this issue by considering emission quotas as a contextual factor.

The quality criteria that should be accomplished in the treatment process (or maximum limits of pollutants concentration on the effluent on exit of the WWTP) varies according to the features of the discharging area or the type of watercourse associated with each plant (Directive, 1991). If the plants do not meet the legislation requirements, the outputs representing the pollutants removed need to be adjusted, which will have an impact on energy consumption. Although the plants usually achieve the limits imposed to pollutants' concentration in wastewater, operational, technological, or contextual/external conditions may compromise the fulfilment of the levels envisaged.

Furthermore, it might also be the case that the WWTPs are removing more pollutants than it is required. In that case, the WWTPs are using more resources than expected or necessary, so there is a potential for resource savings. The framework developed in this paper aims to explore input savings that could be achieved if the water quality criteria were exactly matched during the treatment process.

Different scenarios regarding the pollutants' emission quotas are considered, which involve quotas specified for each plant, or a more flexible regulatory context that allows trade- offs among WWTP, such that the quotas are fulfilled at the system level. This latter scenario involves an assessment of the average concentration of pollutants for effluent waters from various WWTPs converging to the same watercourse.

In this context, it is crucial to develop enhanced frontier methods for performance assessment that can look beyond input-oriented efficiency estimates, and consider simultaneous adjustments to inputs and outputs, given the regulatory framework that imposes target values for output levels. The output adjustments represent an effectiveness dimension that complements the efficiency evaluation.

Therefore, the objective of this paper is to propose an innovative approach that allows a comprehensive evaluation of a set of DMUs, involving the optimization of the resources used to fulfill the objectives set at a system-wide level.

From a methodological perspective, there are studies that also consider trade-offs for the optimization of effectiveness after the projection of DMUs on the efficient frontier. For example, in the study of Golany (1988), the trade-off analysis is conducted interactively by the decision-maker in a five-step interactive multi-objective linear programming procedure, in an effectiveness goal setting context. Golany and Tamir (1995) developed a DEA-based framework for resource allocation that enables the comprehensive integration of three criteria, namely efficiency, effectiveness, and equality, within the context of public sector activities. Here, the authors allow trade-offs to enable a fair compromise between effectiveness and equality criteria. However, these are accomplished through allowable small deviations from the efficient frontier, compromising the efficiency criterium.

The main contribution of this paper is the development of an innovative three-step approach to achieve both efficiency and effectiveness at the system-level. All models used in the different stages are defined with an input-oriented focus, based on the assumption that the efficient frontier is always the same. However, different input minimizing targets are pro- posed based on adjustments to output levels that enable reaching the desired outcomes, at the individual DMU level or at the overall system level. In the latter case, the optimization involved requires a DEA-inspired Mixed-Integer Linear Programming formulation.

This paper also contributes to the performance assessment literature in the water sector by proposing a performance measurement framework that considers both the efficiency of the treatment process and effectiveness regarding the fulfillment of the effluent quality criteria. This is a problem that has not yet been tackled in previous studies in the water sector. The simultaneous consideration of efficiency and effectiveness in the way described in this paper enhances the accuracy of the performance evaluation of WWTPs since the effluent quality requirements is what dictates the cleansing level of the wastewater that WWTPs should undertake. Given the input-oriented focus of the wastewater treatment activity, allowing for adjustments to the cleaning effort required at the individual WWTP level enables additional gains in resource consumption at the overall system level. These gains are likely to be significant if some plants in the system are removing more pollutants than what is required by legislation.

The remaining of the paper is organized as follows. Section 7.2 presents a brief review of the literature on efficiency and effectiveness assessments using frontier techniques. Section 7.3 describes the methodology developed in this research. Section 7.4 presents an application of the methodology to a real-world case study, in the context of wastewater treatment plants from a Portuguese company. The results and their managerial implications are discussed in Section 7.5. Section 7.6 concludes, reviewing the lessons learned and proposing directions for future research.

7.2. Literature Review

The first studies dedicated to the assessment of performance in the water sector go back to the year of 1978, according to the seminal work of Abbott and Cohen (2009). This study makes a systematic review of the literature dedicated to the measurement of efficiency and productivity in the water sector, including water and wastewater services, using statistics and optimization techniques. The authors observed that the studies reviewed fall into one of four categories, namely, economies of scale, economies of scope, public versus private ownership, and effects of regulation.

Focusing on the more recent years, the studies applying quantitative techniques to assess performance in the water sector were mostly conducted at the company or utility level (e.g. Thanassoulis, 2002; De Witte and Marques, 2009; Guerrini et al., 2013; Castellet and Molinos-Senante, 2016; Güngör-Demirci et al., 2018; Zhang and Xu, 2019; D'Inverno et al., 2021). There are also a few studies devoted specifically to the assessment of efficiency of wastewater treatment plants (WWTP). Guerrini et al. (2016) conducted a review of the studies focusing on WWTPs' performance published between 2003 to 2015. This study systematized, among other characteristics, the selected inputs and outputs and the techniques applied and concluded that DEA is by far the most used technique in this context.

The DEA literature on performance measurement of WWTPs has tackled several problems using different modeling approaches. The studies have focused on the assessment of efficiency (Gómez et al., 2017; Hernández-Sancho et al., 2011a; Hernández-Sancho and Sala- Garrido, 2009; Molinos-Senante et al., 2014b; Sala-Garrido et al., 2012a, 2011, 2012b), the evolution of productivity over time (Hernández-Sancho et al., 2011b; Molinos-Senante et al., 2015, 2016b), ecological improvements of WWTP activity (Castellet and Molinos-Senante, 2016; D'Inverno et al., 2018; Dong et al., 2017; Lorenzo-Toja et al., 2015; Molinos-Senante et al., 2016a,b, 2014a).

Despite the different objectives and modelling approaches underlying each study, there is consensus in the literature regarding the specification of the inputs and outputs that reflect WWTP activity, and the returns to scale assumption that should underlie efficiency evaluations based on DEA. Most studies adopted an input-oriented focus to enable the reduction of the resources consumed, considering a variable returns to scale technology.

However, despite the potential to reduce further the inputs via adjustments to the output levels (effort to remove pollutants from wastewater), previous studies did not explore the fulfillment of the quality criteria in the effluent discharged by WWTPs in the receiving waters. Therefore, this issue deserves further attention, as it opens the possibility of enhancing performance via achieving efficient production alongside effective levels of pollutants removal, in accordance with legislation requirements.

The concept of effectiveness is associated with the achievement of the desired goals (Asmild et al., 2007), or doing the right things (Drucker, 1977; Griffin, 1987; Anthony et al., 1989), i.e. the things that are supposed to be attained. In a performance assessment context, it involves measuring how the DMUs perform in regards to the intended goals (Golany, 1988).

By contrast, efficiency is associated with conducting activities as well as possible or "doing things right".

In this broad context of effectiveness analysis, Golany (1988) states that the DEA methodology can be extended from its original objective of efficiency measurement to a concurrent efficiency and effectiveness analysis. This can be achieved following different procedures. One possibility concerns the assessment of economic efficiency (pursuing the economic objectives of cost minimization, revenue or profit maximization), that involves the use of functions relating input-output quantities with the corresponding prices. Alternatively, it may be necessary to restrict the weights associated with inputs or outputs according to their relative importance in terms of the desired goals.

Another alternative for the measurement of effectiveness consists of using goal programming, which requires the specification of the desired level of inputs or outputs to be attained, and minimizing deviations from these goals. A typical application of the effectiveness analysis is the pollutants' emission permits allocation problem. Different modeling developments were proposed to identify the best allocation scheme according to the specificity of the problem. Many of these studies share a particular feature related to the fact that the inputs and outputs of the DMUs under evaluation are not independent, as stated by Wang et al. (2013). These authors developed a multi-input zero-sum gains DEA model to disaggregate the total CO_2 emissions target for 2020 assumed by China in the Copenhagen Summit on Climate Change. This allowed the allocation of sub-targets to each province of China, so that the national targets can be more easily attained. The study also provides a detailed literature review on total emission control and emission allowance allocation approaches.

In the same context of the pollutants' emission allocation problem, a panoply of other studies has been conducted more recently. For example, Sun et al. (2014) developed allocation emission permit models based on DEA. These are used as mechanisms to control the total emissions of a group of manufacturing companies. In one of those mechanisms, a common governing body is considered under a centralized approach. The objective of the model is to maximize overall efficiency of the group of DMUs. The second mechanism, a dominating DMU in the group is considered (under an individual approach). The objective is to maximize the dominant DMU efficiency at the expense of all other DMUs. In both cases, the efficiency levels estimated correspond to the emission permits reduction that should be assigned to each DMU, in accordance with the overall reductions that are imposed to the group of DMUs.

Xie et al. (2020) proposed a fixed-cost allocation DEA-based model for the allocation of tradable pollution permits in the context of the industrial enterprise wastewater discharge permits, which also satisfies the fairness principle. The model was applied to 30 Chinese provinces and demonstrated that all provinces are Pareto-efficient (from an environmental perspective) after the allocation process.

Rogers and Louis (2007) proposed a method to determine the optimal allocation of limited financial resources among community water systems. The framework proposed used DEA to estimate efficiency levels that were subsequently incorporated into a decision support system that solved a linear programming model to determine the optimal allocation of resources among the water systems.

In the context of wastewater treatment plants literature and allocation of pollutant emission permits, there aren't any previous studies focusing on the effectiveness of the water treatment process in terms of the fulfillment of the quality level expected in the effluent discharged. The approach developed in this paper aims at tackling this literature gap. Due to the specificities of the production process, the definition of effectiveness underlying the framework proposed relates to the broader definition of doing the right things, or, more precisely, of doing specifically the treatment activity required by regulation. This means removing the pollutants from the wastewater precisely at the level that is legally required. In doing so, a three-stage DEA-based approach is proposed, which includes some of the developments proposed by Golany and Thore (1997), particularly regarding the use of an envelopment constraints to limit the projection region on each facet of the efficient frontier, and the adoption of a system vision for the problem (Golany et al., 1993; Wang et al., 2013; Sun et al., 2014; Xie et al., 2020).

The novelty of this approach also stems from allowing trade-offs among WWTPs in the third modeling stage of the framework proposed. The trade-offs consist of allowing deviations in terms of the pollutant emission value at the WWTP level, given the assumption that the quality level imposed by regulation must be fulfilled at the discharging watercourse shared by a group of WWTPs.

The objective function aims at minimizing resource consumption for the system comprising all WWTPs. As all WWTPs must operate on the efficient frontier, it is expected that the model derives additional resource savings from the flexibility allowed by trade-offs in the pollutants' emission quotas. Ultimately, the framework developed enables a comprehensive evaluation of WWTPs by providing a measure of performance that combines both efficiency and effectiveness of the treatment process.

7.3. Methodology

In this study, the performance evaluation is conducted with a procedure structured in three steps to optimize the plants' operation to comply with the EQ at the receiving waters while maximizing the energy savings. This approach was motivated by the possibility of the regulatory context evolving to a more flexible approach, which allows the effluent quality monitoring to change from the plant level to the receiving waters level. This requires a system-wide analysis of performance, exploring the opportunity to look at the average quality of the effluent converging to the receiving waters instead of controlling the effluent discharged by each WWTP.

The first step involves the assessment of efficiency, and minimizes the inputs levels given the observed levels of outputs. The second step looks beyond efficiency and seeks the attainment of effectiveness, without compromising the efficiency level accomplished in the previous step. More precisely, effectiveness is assessed by adjusting the outputs to meet regulatory goals.

In the third step, we develop a mixed-integer linear programming model to determine the impact on input savings of the reallocation of pollutants emission quotas among plants. The objective is to minimize resource consumption at the system level rather than focusing on an optimization at the DMU level, whilst maintaining production at the frontier and meeting the regulatory targets regarding water quality. Although optimizing resource consumption at a DMU level or system level would be equivalent in a conventional DEA setting where output levels are fixed (Beasley, 2003), the model proposed in this paper is prepared to accommodate a reallocation of pollutant emission quotas, providing the possibility of trade-offs among DMUs that can be beneficial from a system-level analysis.

These trade-offs imply deviations in the quality level imposed to the wastewater treated in different plants, provided that the average quality levels of effluents disposed in the natural environment respect the limits imposed by legislation. The trade-offs allowed in the model formulation constitutes the major novelty of the methodology proposed.

Figure 7.1 schematically represents the main aspects of the different steps proposed in the methodology.

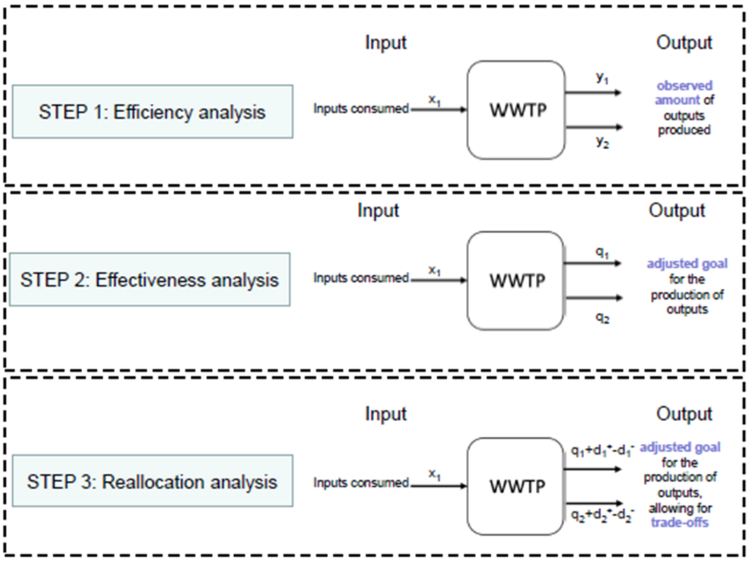


Figure 7.6 *– Sensitivity analysis for different values of M*

7.3.1. Step 1: efficiency analysis

In the first step, the estimation of efficiency scores is done using Data Envelopment Analysis (DEA) (Charnes et al., 1978). The linear programming technique compares the efficiency of a set of decision-making units (DMUs) in their use of resources (inputs) to produce outcomes (outputs). DEA derives a single summary measure of efficiency for each unit based on comparisons with other units in the sample. The benchmarks constitute the peers underlying the construction of the efficient frontier. The frontier is constructed from piece-wise linear segments connecting the efficient DMUs of the sample under analysis. The nonparametric nature of the DEA technique avoids the *a-priori* specification of the functional form for the frontier, which is an essential property when the production process is complex and the production function is unknown.

We consider conventional Data Envelopment Analysis models enhanced with a set of peer restrictions. These models constitute a deterministic alternative to the robust approach developed by Cazals et al. (2002). Given the input-oriented setting adopted in this paper, the model forces any given DMU to be compared only with others with a comparable amount of output levels, which requires bounding the comparison by a threshold concerning the minimum and maximum output levels required for the peers used in the benchmarking exercise. This type of constraints was introduced by Golany and Thore (1997) in the envelopment formulation of a DEA model to reflect external conditions affecting the production process. However, the use of this type of restriction is not commonly explored in the DEA literature. It allows deriving an efficiency score restricting the comparison to peers with a comparable scale of operation. The specification of a maximum output bound also prevents the occurrence of bias due to the disturbance caused by potential outliers in the estimation of the efficient frontier.

In a real-world setting, from the practitioners' perspective, the peer restrictions are particularly relevant, as they allow a fair benchmarking process, especially when the sample is small and with high variability in the magnitude of input and output indicators. In the area of wastewater treatment management, restricting peers to comparable plants is essential to ensure the validity of the model and give managers the confidence that the comparison is fair.

We now proceed with the specification of the DEA model formulation, and definition of the models' parameters and decision variables. Consider j DMUs ($j = 1, \dots, n$), each consuming x_{ij} inputs i $(1, \dots, m)$, and producing y_{rj} outputs r ($r = 1, \dots, s$). The DEA model proposed for Step 1 estimates efficiency taking into account the comparison of peers with a similar scale of operations. Therefore, the constant returns to scale (CRS) model should be used, as homogeneity in terms of scale size between the DMU under assessment and its peers (DMUs on the frontier) is ensured via the peers' restrictions. In the context of a small sample assessment, we consider this approach more adequate than the imposition of VRS, as VRS would only restrict the comparison of the DMU under assessment with a virtual peer on the frontier constructed as a linear combination of observed DMUs from the sample (with no extrapolation allowed). This implies that in a VRS model the scale size of peers may be quite different from the DMU under assessment, provided that their linear combination has a comparable size to the unit under assessment.

The envelopment formulation of the input-oriented DEA model with CRS is shown in Model (7.1).

$$\begin{aligned} & Min\ \theta_k \\ \text{s.t.}\ & \textstyle\sum_{j \in P_k} l_j x_{ij} \leq \theta_k x_{ik} \quad i = 1, \dots, m \\ & \textstyle\sum_{j \in P_k} l_j y_{rj} \geq y_{rk} \quad r = 1, \dots, s \\ & l_j \geq 0 \quad \forall j \in P_k \end{aligned} \tag{7.1}$$

The objective function θ_k corresponds to the efficiency score of the unit k under assessment. θ_k is a decision variable of the model. l_j are the intensity decision variables, which assume a positive value whenever a unit j is used as a peer for the unit under assessment. The set P_k defines the peers with a comparable scale of operation to the unit under assessment k:

$$P_k = \cap_{r=1,\dots,s} \{ j : L_{rk} \leq y_{rj} \leq U_{rk} \} \tag{7.2}$$

where L_{rk} is the minimal and U_{rk} is the maximal threshold levels for output r and DMU k.

In the empirical application presented on section 7.5, we used expression (7.3) to restrict the best practice comparison sets.

$$y_{rk} \leq y_{rj} \leq (1+\alpha) \times y_{rk} \qquad \textbf{(7.3)}$$

This means that DMU j can be peer of DMU k (under assessment) if and only if the observed values for output r of DMU j are:

- greater or equal the observed values for output r of DMU k;
- not more than a certain percentage (α) higher than the observed values for output r of DMU k (we used α = 0.25, meaning that the output levels of peers cannot exceed the observed values of the plant under assessment by more than 25%).

7.3.2. Step 2: efficiency and effectiveness analysis

The second step adds an effectiveness component to the efficiency assessment conducted in the first step. Effectiveness is associated with achieving the desired goals. Therefore, in the context of this paper, effectiveness arises from minimizing input consumption when plants have the outputs adjusted to meet regulatory goals. In addition to the parameters previously defined for model (7.1), let q_{rk} be the adjusted goal for output r of unit k under assessment. The objective function of Model (7.4) (β_k) corresponds to a performance score that reflects both the efficiency and effectiveness of the DMU k. Analogous to model (7.1), the decision variables l_j indicate whether a unit j is used as a peer for the unit under assessment. Note that as the left-hand side of the restrictions of Model (7.4) are identical to the left-hand side of restrictions of Model (7.1), the frontier constructed from the assessment of the DMUs in step 2 remains unchanged in relation to the frontier used in step 1.

$$\begin{aligned} & Min\ \beta_k && \textbf{(7.4)} \\ \text{s.t.}\ & \textstyle\sum_{j \in P_k} l_j x_{ij} \leq \beta_k x_{ik} \quad i = 1, \dots, m \\ & \textstyle\sum_{j \in P_k} l_j y_{rj} \geq q_{rk} \quad r = 1, \dots, s \\ & l_j \geq 0 \quad \forall j \in P_k \end{aligned}$$

It is also important to note that the regulatory goal value q_{rk} for output r of unit k can be equal, lower, or higher than the observed value, and the corresponding implications on the model's objective function and input savings are:

- If, for all r, $q_{rk} = y_{rk} \Rightarrow \beta_k = \theta_k$
  i.e., if for unit k the adjusted goal q_{rk} of all outputs r coincides with the observed y_{rk}, then the input savings with efficiency are equal to the input saving with efficiency and effectiveness. Therefore, the savings associated with effectiveness are equal to zero;
- If, for all r, $q_{rk} > y_{rk} \Rightarrow \beta_k < \theta_k$
  i.e., if for unit k the adjusted goal q_{rk} of all outputs r is greater than the original output value y_{rk}, then the input savings with efficiency are greater than the input

saving with efficiency and effectiveness. Therefore, the savings associated with effectiveness are lower than zero;

- If, for all r, $q_{rk} < y_{rk} \Rightarrow \beta_k > \theta_k$
  i.e., if for unit k the adjusted goal q_{rk} of all outputs r is smaller than the original output y_{rk}, then the input savings with efficiency are lower than the input saving with efficiency and effectiveness. Therefore, the savings associated with effectiveness are greater than zero.

7.3.3. Step 3: trade-offs among goals

The third step also assumes that the input-output process will be done at the frontier of the production possibility set.

A traditional DEA assessment requires solving one linear programming model for each DMU k under assessment. Consequently, the estimation of efficiency for a sample with n DMUs requires solving n linear programming models. However, it is also possible to run a single linear programming model for all DMUs in the sample and obtain the DEA scores or each individual DMU, as shown by Beasley (2003). In standard DEA assessments, each model k has a different vector $l_j = (l_1, l_2, \dots, l_n)$ representing the decision variables that specify the linear combination of peer DMUs used to measure the distance of DMU k to the frontier of the production possibility set. If we add the index k to the decision variables, and define the matrix l_{jk}, ($j = 1, \dots, n; k = 1, \dots, n$) it is possible to run a single linear programming model to evaluate all DMUs at the same time.

Following this procedure, running Model (7.4) n times (once for each DMU k under assessment) is equivalent to running a single linear programming model as shown in (7.5):

$$
\begin{aligned}
& Min\ \textstyle\sum_{k=1}^{n} \beta_k && \textbf{(7.5)}\\
& \text{s.t. } \textstyle\sum_{j \in P_k} l_{jk} x_{ij} \leq \beta_k x_{ik} \quad i = 1, \dots, m; k = 1, \dots, n \\
& \qquad \textstyle\sum_{j \in P_k} l_{jk} y_{rj} \geq q_{rk} \quad r = 1, \dots, s; k = 1, \dots, n \\
& \qquad l_{jk} \geq 0 \ \ \forall j, k
\end{aligned}
$$

In order to allow trade-offs among the goals specified for the DMUs in the sample under assessment, we propose an innovative DEA approach that involves running a single mixed-integer linear programming problem to evaluate all DMUs in the system simultaneously.

These variations around the goals specified for each DMU individually are expressed in Model (7.6) by d_{rk}^{+} for a positive deviation, and d_{rk}^{-} for a negative deviation.

$$
\begin{aligned}
& Min \sum_{k=1}^{n} f_k && \textbf{(7.6)} \\
& \text{s.t. } \sum_{j \in P_k} l_{jk} x_{ij} \leq f_k x_{ik} \quad i = 1, \dots, m; k = 1, \dots, n \\
& \quad \sum_{j \in P_k} l_{jk} y_{rj} \geq q_{rk} + d_{rk}^{+} - d_{rk}^{-} \quad r = 1, \dots, s; k = 1, \dots, n \\
& \quad d_{rk}^{+} \leq M \times q_{rk} \times z_{rk} \quad r = 1, \dots, s; k = 1, \dots, n \\
& \quad d_{rk}^{-} \leq M \times q_{rk} \times (1 - z_{rk}) \quad r = 1, \dots, s; k = 1, \dots, n \\
& \quad \sum_{k=1}^{n} V_k d_{rk}^{+} = \sum_{k=1}^{n} V_k d_{rk}^{-} \quad r = 1, \dots, s \\
& \quad d_{rk}^{+}, d_{rk}^{-} \geq 0 \; \forall r, k \\
& \quad z_{rk} \in \{0,1\} \; \forall r, k \\
& \quad l_{jk} \geq 0 \;\; \forall j, k
\end{aligned}
$$

The decision variables f_k define the proportional input adjustments that are possible for each DMU k. The value of f_k can also be interpreted as the factor by which the input levels of the DMU under analysis can be decreased radially (equiproportionally), considering the output levels set for each DMU. Therefore, the objective function of Model (7.6) minimizes this factor for all DMUs in the system, to detect the overall potential for resource savings.

The first set of constraints in Model (7.5) remains unchanged in Model (7.6). The right-hand side of the second set of constraints is modified to allow deviations to the output levels based on the trade-offs among DMUs. Therefore, the goal q_{rk} is replaced by $q_{rk} + d_{rk}^{+} - d_{rk}^{-}$. The decision variables d_{rk}^{+} and d_{rk}^{-} correspond to positive and negative deviations to the goals q_{rk}. In the context of the wastewater treatment plants assessment, these deviations are trade-offs allowed among the quantities of pollutants removed at each plant.

The set of the first two constraints ensures that production is done at the frontier of the production possibility set, constructed from the original observations of DMUs' inputs and outputs. This is mathematically defined by the left-hand side of these constraints, corresponding to a linear combination of x_{ij} and y_{rj} values observed at the DMUs, using the values of l_{jk} as the multipliers.

The decision variables d_{rk}^{+}, d_{rk}^{-}, and q_{rk} are continuous and positive.

The third and fourth set of constraints ensure that the units are allowed at most a certain variation (M) around the original goals specified at plant level (M is expressed as a percentage). These restrictions also guarantee that either a positive or a negative variation to the goal is allowed via the specification of the binary decision variables z_{rk}. If $z_{rk} = 1$, then there is a positive variation in the goal r of DMU k, and if $z_{rk} = 0$ then there is a negative variation in the goal r of DMU k.

The fifth constraint imposes that weighted positive deviations are offset by weighted negative deviations at the system level. The deviations are weighted by the parameters V_k that represent the relative importance of the DMU in the overall system, guaranteeing that at the system level, the goals are fulfilled. For example, in the context of wastewater treatment plants, the output could correspond to the amount of pollutants removed (kg per m^3). Thus, a positive deviation of 10 kg/m^3 in a WWTP treating 20% of the total volume of water in the system must be offset by a negative deviation of 2.5 kg/m^3in a WWTP treating the remaining 80% of the total volume of water.

Note that, if $d_{rk}^{+} = d_{rk}^{-} = 0$, Model (7.6) is reduced to Model (7.5) and $f_k = \beta_k \; \forall k$.

For the special case of a single input or multiple inputs measured in monetary terms, if the objective is to minimize the total input savings of the system, an alternative specification of the objective function that takes into account the relative scale size of the DMUs can be adopted, as shown in (7.7). This was the objective function used in the empirical application reported in this paper.

$$Min \sum_{k=1}^{n} f_k\left(\sum_{i=1}^{m} x_{ik}\right) \quad \textbf{(7.7)}$$

7.3.4. A small illustrative example

This section illustrates the methodology proposed in this paper using a small contrived example. Suppose we have 8 DMUs, consuming one input and producing one output. Each unit has a goal corresponding to a target output level. Table 7.1 reports the data.

Table 7.1 – Illustrative data.

DMU	Input (x)	Output (y)	Goal (q)	Size (V)
A	41	51.04	51.04	1112
B	25	43.19	40.55	273
C	15	13.5	13.50	160
D	45	71.89	75.00	900
E	18	36.76	31.36	638
F	52	78.67	87.88	943
G	32	65.34	47.47	414
H	55	98.9	74.38	410
Total	283	459.29	421.18	4850

At the system level, the DMUs are producing 9% more outputs than the goal $\left(\frac{459.29-421.8}{421.18} = 0.09\right)$. DMUs D and F are the only ones that do not reach the goal since the amount of output produced is lower than the goal imposed. DMUs A and C are the only ones that meet exactly the production goal, which signals effectiveness. The total input consumption is 283.

Due to the reduced number of DMUs considered in this sample, we do not impose restrictions to peers to force comparisons only with DMUs with a similar scale size. This implies that in this illustrative example P_k specified in expression (7.2) corresponds to the entire sample.

Step 1: efficiency analysis

Table 7.2 reports the results of the traditional efficiency evaluation (Model (7.1)).

Table 7.2 – Results for the illustrative efficiency analysis.

DMU	Input (x)	Efficiency score (θ)	Input target (θ_x)	Input savings $(1-\theta)x$
A	41	0.61	25.01	15.99
B	25	0.85	21.25	3.75
C	15	0.44	6.6	8.4
D	45	0.78	35.1	9.9
E	18	1	18	0
F	52	0.74	38.48	13.52
G	32	1	32	0
H	55	0.88	48.4	6.6
Total	283		224.84	58.16

The efficient DMUs are E and G. Therefore, adjusting the input levels of the other DMUs to the efficient frontier will enable a reduction to input consumption. At the system level, the input savings corresponding to the attainment of efficiency correspond to 58.16, representing a saving potential of 21% $\left(\frac{283-224.84}{283}=0.21\right)$.

Step 2: efficiency and effectiveness analysis

Table 7.3 reports the results of the performance assessment using Model (7.4), which minimizes input consumption while meeting the goal defined for the output level. This involves the simultaneous achievement of efficiency (via the reduction of inputs until reaching the frontier), and effectiveness (via the adjustments of observed outputs to meet a pre-defined goal).

Table 7.3 – Results for the illustrative efficiency and effectiveness analysis.

DMU	Input (x)	Efficiency & effectiveness (β)	Input target (β_x)	Total savings $(1-\beta)x$	Efficiency savings $(1-\theta)x$	Effectiveness savings $(\theta-\beta)x$
A	41	0.61	25.01	15.99	15.99	0
B	25	0.79	19.75	5.25	3.75	1.5
C	15	0.44	6.6	8.4	8.4	0
D	45	0.82	36.9	8.1	9.9	-1.8
E	18	0.85	15.3	2.7	0	2.7
F	52	0.83	43.16	8.84	13.52	-4.68
G	32	0.73	23.36	8.64	0	8.64
H	55	0.66	36.3	18.7	6.6	12.1
Total	283		206.38	76.62	58.16	18.46

Only DMUs A and C achieve effectiveness. This happens because the output levels coincide with the goals. Therefore, adjusting of output levels to reach effectiveness for the remaining DMUs enables savings in input consumption. At system level, 76.62 of input savings can be achieved (58.16 associated with efficiency and 18.46 with effectiveness), representing a total saving potential at the system level of 27% $\left(\frac{283-206.38}{283}=0.27\right)$.

In Figure 7.2 we illustrate how the model leads to input savings. Consider DMU G. In the first step, DMU G had an efficiency value equal to 1. It means that this unit is on the frontier and there is no margin for decreasing the input consumption. However, when we adjust the output value to meet the goal (corresponding to a lower output level), this unit moves to a position inside the production possibility set. This effectiveness adjustment allows decreasing input consumption, as it opens the possibility of projecting it on the frontier. Although unit G has no saving potential concerning efficiency (see Table 7.3), the effectiveness assessment brings a saving potential of 8.64.

Step 3: Trade-offs analysis

Table 7.4 reports the results of the evaluation from Model (7.6), considering M = 20%.

At system level, an input reduction of 88.53 is achieved with efficiency and effectiveness, allowing for deviations in the goals, representing a saving potential of 31% in input consumption.

To illustrate the process underlying the optimization of step 3, consider again DMU G in Figure 7.3. In step 2, it was possible to save 8.64 of inputs by adjusting the output to meet precisely the goal (G^g). However, if we allow the goal to have some minor variations to decrease the input consumption, the optimum solution to Model (7.6) leads to a negative deviation ($d^{-*} = 0.27$) for DMU G, corresponding to an adjustment of the original goal (47.47) to 47.20. This decrease allowed reducing the consumption of inputs for DMU G by 8.89.

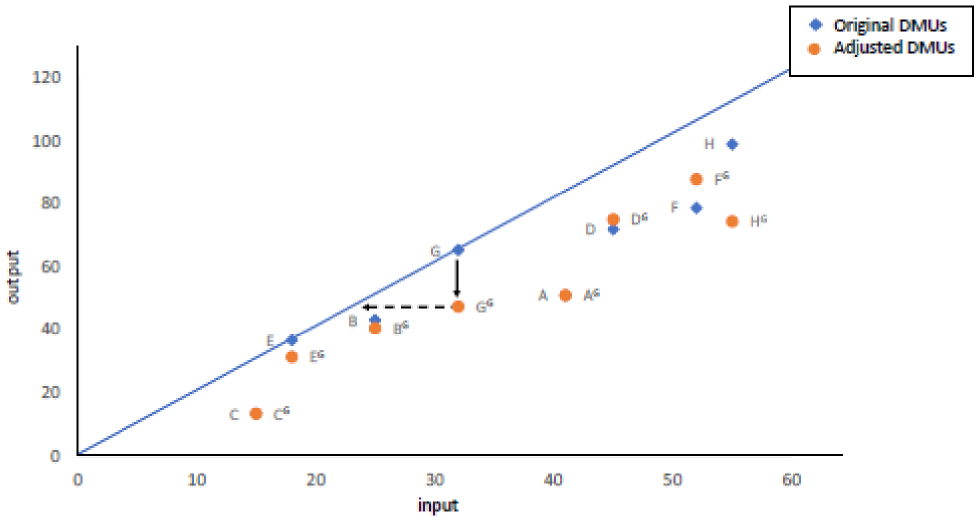


Figure 7.2 *– Frontier representation and efficiency and effectiveness analysis.*

Moreover, the balance constraint between positive and negative deviations (see the 5th constraint in model (7.6)) is met when considering the size (V) presented in Table 7.1:

$$\frac{1112 \cdot 0 + 273 \cdot 9.89 + 160 \cdot 13.5 + 900 \cdot 0 + 638 \cdot 0 + 943 \cdot 0 + 414 \cdot 0.27 + 410 \cdot 6.52}{4850}$$
$$= \frac{1112 \cdot 1.59 + 273 \cdot 0 + 160 \cdot 0 + 900 \cdot 3 + 638 \cdot 1.53 + 943 \cdot 2.33 + 414 \cdot 0 + 410 \cdot 0}{4850}$$

The parameter V_k represents the relative importance of the DMU in the overall system, which in this case is the proportion of the total water treated at plant k.

Finally, we highlight that in our case study, variable V will be expressed as the proportion of water treated at each WWTP (volume of water treated at the WWTP divided by the total water treated at all WWTPs).

Table 7.4 – Results for the illustrative trade-off analysis.

DMU	Input (x)	Output (y)	Goal (q)	Model (7.6) Score f	Input target fx	Deviations d^-	d^+	Adj. goal $(q+d^+-d^-)$	Total saving $(1-f)x$	Step 2 savings $(1-\beta)x$	Trade-off savings $(\beta-f)x$
A	41	51.04	51.04	0.63	25.777	0	1.59	52.63	15.23	15.99	-0.76
B	25	43.19	40.55	0.60	15.01	9.89	0	30.66	9.99	5.25	4.74
C	15	13.5	13.50	0.00	0	13.5	0	0.00	15.00	8.40	6.60
D	45	71.89	75.00	0.85	38.195	0	3	78.00	6.81	8.10	-1.29
E	18	36.76	31.36	0.89	16.10	0	1.53	32.89	1.90	2.70	-0.80
F	52	78.67	76.34	0.83	43.04	0	2.33	78.67	8.96	8.84	0.12
G	32	65.34	47.47	0.72	23.11	0.27	0	47.20	8.89	8.64	0.25
H	55	98.9	74.38	0.60	33.23	6.52	0	67.85	21.77	18.70	3.07
Total	283				194.47				88.53	76.62	11.91

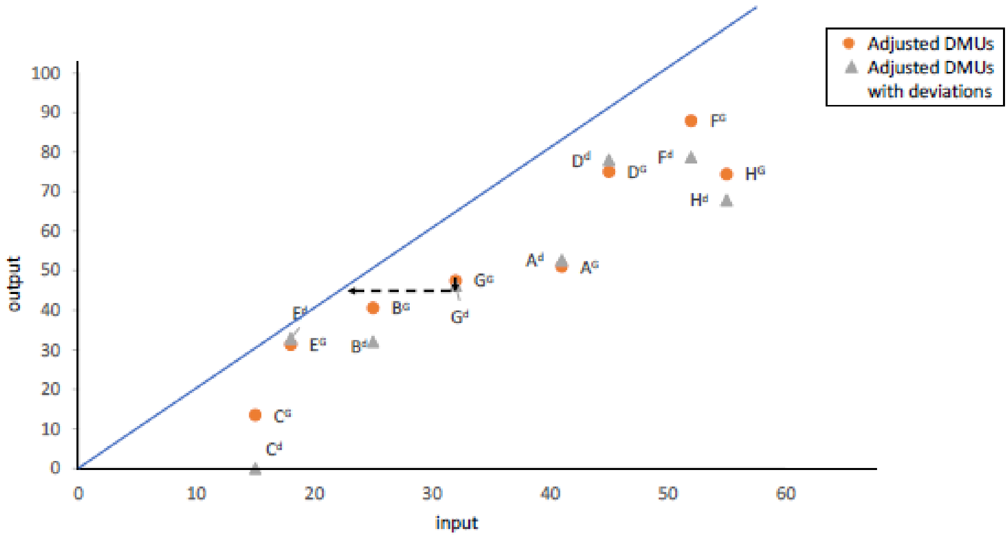


Figure 7.3 – Representation of the performance assessment allowing for trade-offs.

7.4. Case study

The methodological developments proposed in this paper were motivated by a real-world situation faced by a Portuguese water company.

Following the requirements for discharges from urban wastewater treatment plants determined by the Council of the European Communities (Directive, 1991), and translated to the Portuguese framework law through the Law 152/97 of June 19, each WWTP must comply with the Legal Value for Emission (Emission Quota (EQ)) specified for the effluent discharged. The EQ is specified in accordance with the level of sensitivity of the receiving waters. This information is specifically highlighted by the Portuguese Agency for the Environment (APA) and the Portuguese regulator for water and waste services (ERSAR) in Simões et al. (2008). Among the accountability for the environmental impact associated with all phases of WWTP design, implementation, operation and dismantling, this document establishes the framework for monitoring the environmental impact associated with WWTPs during the operation phase.

The effluent quality at each WWTP is periodically monitored and reported to ERSAR to check the degree of compliance with the EQ imposed by legislation. This information is published in the annual report of ERSAR so that it can be publicly accessed.

By the time the data used in this study was collected, the majority of the WWTPs of the Portuguese water company analyzed in this research were operating with an effluent quality higher than the values imposed by regulation (i.e., with a pollutant emission limit lower than the EQ). Only a few plants revealed difficulties in obtaining acceptable effluent quality levels.

Given the legislative framework that allows pollutant levels to be assessed for the receiving environment, a hypothetical future scenario was considered by the company, where the effluent quality monitoring could change from the plant level to the receiving waters level. This scenario is rooted in the fact that several WWTPs of the company discharge their effluents to the same receiving watercourse, representing an opportunity to look at the average quality of the effluent converging to the receiving waters instead of controlling the effluent discharged by each WWTP.

The research question to be explored in this study is how to optimize the plants' operation in order to comply with the EQ at the discharge points (receiving waters) whilst maximizing the energy savings allowed with this scenario. To conduct this study, we used data from a sample of 41 WWTPs that are operated by the Portuguese water company. These plants are all activated sludge systems, and the data collected refers to the year 2015.

The formulation of the DEA model underlying the assessment of the potential saving in energy consumption adopted the input-output specification proposed by Henriques et al. (2021). The DEA model seeks to minimize the energy balance given the pollutants removed (Chemical Oxygen Demand (COD) and Suspended Solids (SS)) at the WWTP, as shown in Figure 7.4.

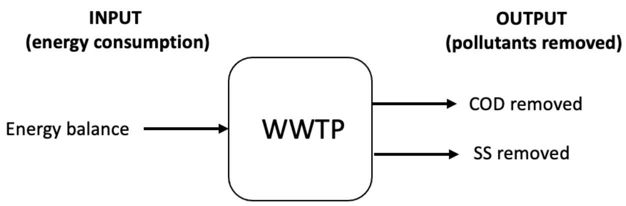


Figure 7.4 *– Schematic representation of the DEA model used.*

The input indicator representing the energy balance is computed as follows:

$$\text{energy balance} = (\text{electricity consumed} + \text{natural gas consumed}) - (\text{electricity produced})$$

Concerning the input indicator, most DEA studies that focus on the WWTP efficiency consider operation costs as inputs, either disaggregated by resource type or aggregated in a single cost item. Alternatively, some studies use the amount of different physical resources as input indicators.

In terms of overall operation costs at WWTPs, energy represents the second largest share (18%), with the first one corresponding to personnel costs (Molinos-Senante et al., 2010). In activated sludge plants, electricity consumption in the aeration phase of the secondary treatment represents between 50% and 60% of total energy consumption (Hernández-Sancho et al., 2011a). Despite the importance of other resources to ensure the fulfillment of the treatment process, in this study we only considered energy consumption, as for other resources (e.g., reagents, maintenance costs) data was not available at the plant level. As the company was specifically focused on energy savings, this was not considered a major limitation. Furthermore, the adjustments to WWTP operation resulting from the flexibility allowed by the reallocation of emission quotas are expected to be made during the aeration phase of the secondary treatment, where energy is the most relevant input.

Concerning the outputs, they allow a comparative analysis of the total amount of pollutants removed during the wastewater treatment process. The output variables reflect two critical aspects of WWTP operation: the proportion of pollutants removed from the raw influent and the volume of wastewater treated.

In (7.8) we define the output y_{rj} for Model (7.1) corresponding to the removal of pollutant r (COD or SS) at plant j:

$$y_{rj} = \frac{p_{rj}^{in} - p_{rj}^{out}}{p_{rj}^{in}} \times vol_j \quad \textbf{(7.8)}$$

As shown in expression (7.8), the removal of pollutant r is measured as the ratio between the quantity of pollutant removed (obtained by the difference between the pollutant measured in the raw influent p_{rj}^{in} and treated effluent p_{rj}^{out}) and the quantity of pollutant observed at the plant's entry, multiplied by the total volume of wastewater treated annually (vol_j). Both values (p_{rj}^{in} and p_{rj}^{out}) correspond to the mean of all values measured periodically during the year under assessment.

Table 7.5 shows the descriptive statistics of the original data provided by the company.

Such data was used to compute the inputs and outputs considered in the empirical application.

Table 7.5 – Descriptive statistics for inputs and outputs of the DEA model (41 WWTPs, year of 2015)

	Mean	StDev	Min	Max	Q1	Median	Q3
Energy balance (MWh/year)	376.85	781.48	21.50	4476.24	51.74	135.30	201.38
Volume of water treated (1000m^3)	1083.74	2701.74	2.16	13540.77	68.68	143.73	373.77
Pollutants observed in raw influent							
COD ((gO_2/m^3)/year)	702.58	435.75	107.50	2742.50	473.33	633.77	803.33
SS (g/m^3)	351.31	328.58	63.45	2137.50	196.86	285.83	377.50
Pollutants observed in treated influent							
COD ((gO_2/m^3)/year)	49.53	36.13	7.50	174.50	23.38	36.67	63.13
SS (g/m^3)	21.15	22.99	3.17	107.50	8.00	10.50	26.59
Outputs of DEA model:							
COD removal (1000m^3/year)	987.08	2468.53	2.01	12460.95	62.00	140.52	323.47
SS removal (1000m^3/year)	1022.19	2557.33	2.01	12875.26	57.29	141.85	331.01

The values for the emission quotas depend on the river that WWTP discharges the pollutants into. For plants P1 to P33, the emission quota for COD is 125 gO_2/m^3 and for SS is 35 g/m^3. For plants P34 and P35, the emission quotas for COD is 125 gO_2/m^3 and for SS is 60 g/m^3. For plants P36 to P41, the emission quota for COD is 150 gO_2/m^3 and for SS is 60 g/m^3.

For conducting the effectiveness analysis, the outputs should be adjusted to meet the legislation regarding the emission quotas. Let EQ_{rj} be the emission quotas considering the discharge levels of pollutants r (COD or SS) for each plant j. Then, in (7.9) we compute the goal (q_{rj}) in terms of pollutants removal to be used in Model (7.4).

$$q_{rj} = \frac{p_{rj}^{in} - EQ_{rj}}{p_{rj}^{in}} \times vol_j \tag{7.9}$$

Finally, the third DEA model allows variations in the emission quotas at each plant provided that the emission quotas for the river are met. Positive deviations in some plants must be compensated by negative deviations in other plants. This implies adding a set of constraints to enforce that for all plants k discharging in the same river R_t, the average quality of the water discharged must meet the quota EQ_{R_t} defined by legislation. We have $s = 2$ outputs, one input, and $p = 3$ rivers where the effluent is discharged to the natural environment.

$$Min \sum_{k=1}^{n} f_k x_{1k} \tag{7.10}$$

$$\text{s.t. } \sum_{j \in P_k} l_{jk} x_{ij} \leq f_k x_{1k} \quad k = 1, \dots, n$$

$$\sum_{j \in P_k} l_{jk} y_{rj} \geq \frac{p_{rk}^{in} - [EQ_{rk} + d_{rk}^{+} - d_{rk}^{-}]}{p_{rk}^{in}} \times vol_k \quad r = 1, \dots, s; k = 1, \dots, n$$

$$d_{rk}^{+} \leq M \times EQ_{rk} \times z_{rk} \quad r = 1, \dots, s; k = 1, \dots, n$$

$$d_{rk}^{-} \leq M \times EQ_{rk} \times (1 - z_{rk}) \quad r = 1, \dots, s; k = 1, \dots, n$$

$$EQ_{R_t} = \frac{\sum_{k \in R_t} [EQ_{rk} + d_{rk}^{+} - d_{rk}^{-}] \times vol_k}{\sum_{k \in R_t} vol_k} \quad r = 1, \dots, s; t = 1, \dots, p$$

$$d_{rk}^{+}, d_{rk}^{-} \geq 0 \; \forall r, k$$

$$z_{rk} \in \{0,1\} \; \forall r, k$$

$$l_{jk} \geq 0 \;\; \forall j, k$$

7.5. Results and discussion

This section compares the results in terms of energy savings obtained in the three steps of the methodology previously described.

7.5.1. Results for efficiency

The first step aims at improving the potential to reduce the use of resources by adopting best practices observed in other WWTPs, given the pollutants currently removed.

First, we compare the results obtained from a traditional (deterministic) DEA with the ones obtained using peers' constraints, using a value of $\alpha = 0.25$ in expression (7.3). Table 7.6 shows the descriptive statistics of the efficiency scores.

According to Golany and Thore (1997), the DEA with peers restricted can be seen as a dynamic clustering. The efficiency scores obtained under the conventional approach are much lower than the mean value of DEA with peers restricted. This can be explained by the conditional assumption, where each WWTP can only be compared with those that produce comparable amounts of outputs.

Table 7.6 – Summary information of efficiency results.

	conventional DEA	peers restricted DEA
Mean	0.23	0.81
StDev	0.20	0.23
Min	0.01	0.31
Max	1	1

Figure 7.5 highlights the importance of conducting the benchmarking evaluation only with comparable WWTPs, as a full-frontier may compromise the design of policies due to the heterogeneity of plants' features.

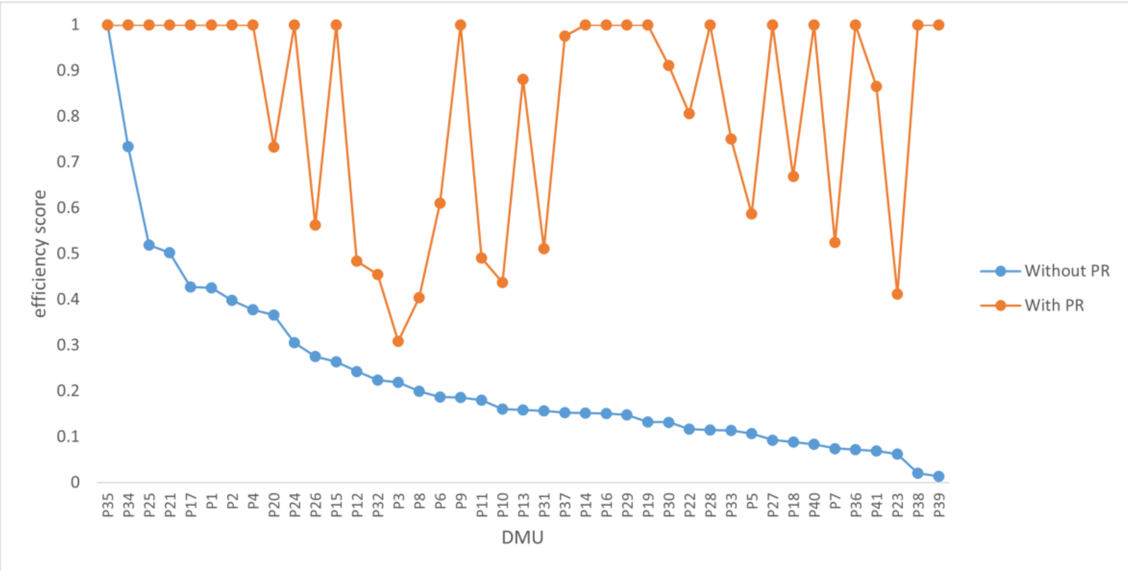


***Figure 7.5** – Comparison of Conventional DEA with peer restricted DEA.*

Using the DEA Model (7.1) with peers' constraints we found 28 plants performing inefficiently. Table 7.7 presents the information regarding efficiency scores and targets for the inefficient DMUs. The description of the peers identified is also reported.

At the system level, 4385.92 MWh/year of energy savings can be achieved through efficiency improvements, corresponding to 28.39% of the total energy consumed by the 41 WWTPs (15450.91 MWh/year).

7.5.2. Results for efficiency and effectiveness

The second step aims at reducing further the energy consumption by adjusting the effort to remove pollutants to meet exactly the values of emission quotas imposed by legislation. Table 7.8 reports the results of the effectiveness analysis obtained using Model (7.4).

On average, the WWTPs are removing 62% more COD than necessary and 49% more SS than necessary. Therefore, when meeting the exact emission quota for each WWTP, the total energy savings are 5291.52 MWh/year (4385.92 MWh/year corresponding to savings from efficiency and 905.60 MWh/year from effectiveness). At the system level, this represents a total saving potential of 34.25% in energy consumption.

Only one WWTP did not meet the COD EQ (plant P38), and five WWTPs did not meet the SS EQs (plants P6, P12, P22, P33, and P38). These WWTPs did not comply with the legislation, as they did not reach the levels required for quality of the water discharged, i.e., the effluents have pollutant amounts higher than the maximum allowed. This may have been caused by exogenous conditions or atypical events that the plants could not cope with as would be desirable.

Table 7.7 – Inefficient WWTPs from efficiency analysis.

WWTP	energy used (x)	peers	Efficiency (θ)	energy target (θx)	energy savings $(1-\theta)x$
P3	4476.24	P34	0.31	1378.63	3097.61
P5	910.58	P09	0.59	534.73	375.85
P6	47.63	P24	0.61	29.09	18.54
P7	71.98	P37	0.52	37.72	34.25
P8	201.38	P21	0.4	81.27	120.11
P10	221.19	P20	0.44	96.66	124.53
P11	183.02	P20	0.49	89.69	93.33
P12	171.46	P21	0.48	82.91	88.55
P13	168.76	P11	0.88	148.61	20.15
P18	179.07	P19	0.67	119.69	59.38
P20	101.06	P21	0.73	74.1	26.95
P22	113.44	P29	0.81	91.37	22.07
P23	155.76	P14	0.41	64.17	91.59
P26	130.57	P21	0.56	73.5	57.07
P30	42.01	P37	0.91	38.27	3.74
P31	51.74	P24	0.51	26.43	25.31
P32	178.61	P21	0.45	81.16	97.45
P33	100.88	P14	0.75	75.65	25.22
P37	42.15	P6	0.98	41.12	1.04
P41	23.56	P40	0.87	20.38	3.18
Total					4385.92

Table 7.8 – Results step 2.

WWTP	Energy used (x)	EQ COD	EQ SS	Performance score (β)	Energy target (βx)	Total savings $(1-\beta)x$	Efficiency savings $(1-\theta)x$	Effectiv. Savings $(\theta-\beta)x$
P1	1608.07	125	35	0.93	1500.36	107.71	0	107.71
P2	152.62	125	35	0.98	149.02	3.61	0	3.61
P3	4476.24	125	35	0.26	1181.71	3294.54	3097.61	196.92
P4	598.14	125	35	0.86	515.92	82.21	0	82.21
P5	910.58	125	35	0.5	456.9	453.68	375.85	77.83
P6	47.63	125	35	0.58	27.57	20.06	18.54	1.52
P7	71.98	125	35	0.45	32.44	39.54	34.25	5.28
P8	201.38	125	35	0.34	68.18	133.2	120.11	13.09
P9	532.73	125	35	0.96	509.24	23.49	0	23.49
P10	221.19	125	35	0.39	86.56	134.63	124.53	10.1
P11	183.02	125	35	0.47	86.2	96.82	93.33	3.49
P12	171.46	125	35	0.48	81.69	89.77	88.55	1.22
P13	168.76	125	35	0.87	146.16	22.6	20.15	2.45
P14	74.81	125	35	0.84	62.97	11.84	0	11.84
P15	175.95	125	35	0.7	123.96	51.98	0	51.98
P16	80.85	125	35	0.9	72.99	7.86	0	7.86
P17	488.93	125	35	0.92	448.1	40.83	0	40.83
P18	179.07	125	35	0.62	111.8	67.27	59.38	7.89
P19	135.3	125	35	0.93	126.11	9.19	0	9.19
P20	101.06	125	35	0.56	56.34	44.72	26.95	17.76
P21	90.89	125	35	0.52	47.65	43.24	0	43.24
P22	113.44	125	35	0.84	95.19	18.24	22.07	-3.82
P23	155.76	125	35	0.38	59.64	96.12	91.59	4.52
P24	31.18	125	35	0.88	27.57	3.62	0	3.62
P25	169.71	125	35	0.93	157.14	12.56	0	12.56
P26	130.57	125	35	0.49	64.02	66.55	57.07	9.48
P27	217.35	125	35	0.92	199.8	17.55	0	17.55
P28	33.29	125	35	0.9	29.8	3.49	0	3.49
P29	94.65	125	35	0.84	79.72	14.93	0	14.93
P30	42.01	125	35	0.86	36.02	5.99	3.74	2.26
P31	51.74	125	35	0.42	21.47	30.26	25.31	4.95
P32	178.61	125	35	0.4	71.99	106.62	97.45	9.17
P33	100.88	125	35	0.74	74.54	26.33	25.22	1.11
P34	1485.19	125	60	0.95	1412.59	72.6	0	72.6
P35	1787.58	125	60	0.99	1767.97	19.6	0	19.6
P36	41.25	150	60	0.96	39.47	1.78	0	1.78
P37	42.15	150	60	0.75	31.55	10.61	1.04	9.57
P38	35.91	150	60	1.15	41.41	-5.5	0	-5.5
P39	21.5	150	60	0.86	18.56	2.94	0	2.94
P40	23.95	150	60	0.89	21.29	2.66	0	2.66
P41	23.56	150	60	0.75	17.77	5.79	3.18	2.61
Total	15450.91				10159.39	5291.52	4385.92	905.60

7.5.3. Results for efficiency and effectiveness with EQs trade-offs

The third step aims at enhancing resource savings through a reallocation of pollutant emission quotas among WWTs (i.e., allowing for trade-offs among WWTPs).

We performed a sensitivity analysis regarding the percentage variation (M) allowed around the emission quotas to be used in Model (7.10). For $M = 0$, the optimum solution of Model (7.5) and Model (7.10) coincide. Figure 7.6 represents the energy savings (in percentage) for each value of M. Recall that the observed energy consumption was 15450.91 MWh/year.

For $M = 0$, at the system level, 5291.52 MWh/year of energy savings can be achieved, representing a total saving potential of 34.25% in energy consumption (as estimated in the previous section). For $M = 100\%$, 5862.10 MWh/year of energy savings can be achieved at the system level, representing a total saving potential of 37.98% in energy consumption.

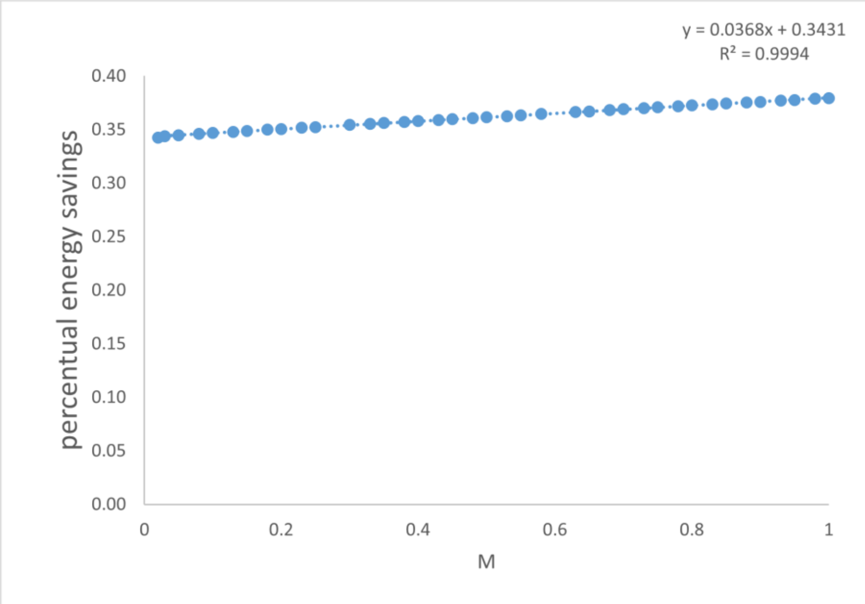


Figure 7.6 *– Sensitivity analysis for different values of M.*

From the rate $\frac{dy}{dx} = 0.0004$ in the trendline shown in Figure 7.6, we see that the incremental gain in the saving potential is minimal when varying M (where y represents the saving potential in percentage, and x represents the variations M around the goals). With this sensitivity analysis, we see that the model is robust and has minor sensitivity to the value of M. The company studied considered that a margin of deviation around 25% is a realistic value that would be feasible to implement in practice at the plants. Therefore, Table 7.9 reports the results of the performance evaluation for $M = 25\%$ using Model (7.10).

Table 7.9 – Results for the reallocation analysis with $M = 25\%$.

WWTP	Input (x)	Output		Goal EQ		Model (7.10) score (f)	Input target fx	Adj. Goal		Savings		
		y_1	y_2	COD	SS			COD	SS	Total $(1-f)x$	Step 2 $(1-\beta)x$	Reallocation $(\beta-f)x$
P1	1608.07	4765959	4919922	125	35	0.89	1433.76	125	43.75	174.31	107.71	66.60
P2	152.62	411072	437342	125	35	1.00	152.77	125	26.25	-0.14	3.61	-3.75
P3	4476.24	6436703	7043061	125	35	0.28	1241.19	125.32	26.25	3235.05	3294.54	-59.49
P4	598.14	1532103	1628018	125	35	0.82	490.05	127.15	43.75	108.09	82.21	25.88
P5	910.58	679276	673304	125	35	0.47	429.76	156.25	34.89	480.82	453.68	27.14
P6	47.63	62002	48593	125	35	0.55	26.09	156.25	35.05	21.54	20.06	1.48
P7	71.98	37240	37533	125	35	0.43	30.69	156.25	43.75	41.29	39.54	1.75
P8	201.38	280435	285853	125	35	0.32	63.66	125	43.75	137.72	133.2	4.52
P9	532.73	676737	713117	125	35	0.93	497.71	125	43.75	35.02	23.49	11.53
P10	221.19	246715	246370	125	35	0.38	83.48	125	43.75	137.71	134.63	3.08
P11	183.02	228924	231092	125	35	0.48	87.52	125	26.25	95.50	96.82	-1.32
P12	171.46	290066	249971	125	35	0.48	83.10	125	26.25	88.36	89.77	-1.41
P13	168.76	185878	186877	125	35	0.87	146.77	125	26.25	21.99	22.6	-0.61
P14	74.81	79054	80448	125	35	0.79	59.41	125	43.75	15.40	11.84	3.56
P15	175.95	323467	331014	125	35	0.60	105.86	125	43.75	70.09	51.98	18.11
P16	80.85	84672	86562	125	35	0.87	70.45	93.75	43.75	10.40	7.86	2.54
P17	488.93	1373632	1503468	125	35	0.95	463.37	125	26.25	25.56	40.83	-15.27
P18	179.07	110492	111644	125	35	0.61	109.47	93.75	43.75	69.60	67.27	2.33
P19	135.3	124901	127390	125	35	0.91	123.45	101.14	43.75	11.85	9.19	2.66
P20	101.06	257932	260642	125	35	0.51	51.34	93.75	43.75	49.72	44.72	5.00
P21	90.89	317979	319683	125	35	0.36	33.00	93.75	43.75	57.89	43.24	14.65
P22	113.44	92521	92909	125	35	0.82	92.76	93.75	43.75	20.68	18.24	2.44
P23	155.76	67810	682254	125	35	0.37	58.30	107.71	43.75	97.46	96.12	1.34
P24	31.18	66466	67666	125	35	0.85	26.54	156.25	43.75	4.64	3.62	1.02
P25	169.71	613218	627741	125	35	0.95	161.63	125	26.25	8.08	12.56	-4.48
P26	130.57	250361	258515	125	35	0.50	65.15	125	31.19	65.42	66.55	-1.13
P27	217.35	140523	141853	125	35	0.90	194.70	125	43.75	22.65	17.55	5.10
P28	33.29	26667	26962	125	35	0.87	28.80	125	43.75	4.49	3.49	1.00
P29	94.65	97641	96243	125	35	0.78	73.58	111.58	43.75	21.07	14.93	6.14
P30	42.01	38557	38080	125	35	0.83	34.67	93.75	43.75	7.34	5.99	1.35
P31	51.74	56330	57294	125	35	0.39	19.94	125	43.75	31.80	30.26	1.54
P32	178.61	277956	285483	125	35	0.38	68.57	125	43.75	110.04	106.62	3.42
P33	100.88	79948	68898	125	35	0.72	72.58	156.25	36.41	28.29	26.33	1.96
P34	1485.19	7598923	7587447	125	60	0.92	1365.25	155.71	75	119.94	72.6	47.34
P35	1787.58	12460949	12875260	125	60	1.00	1793.88	106.44	50.93	-6.31	19.6	-25.91
P36	41.25	20736	20299	150	60	1.00	41.35	150	45	-0.10	1.78	-1.88
P37	42.15	44803	41943	150	60	0.66	27.71	157.39	72.84	14.45	10.61	3.84
P38	35.91	5104	4860	150	60	1.10	39.68	112.5	75	-3.76	-5.5	1.74
P39	21.5	2012	2007	150	60	0.81	17.41	187.5	69.59	4.08	2.94	1.14
P40	23.95	13330	14378	150	60	0.94	22.62	132.31	45	1.33	2.66	-1.33
P41	23.56	11344	11729	150	60	0.75	17.61	156.13	45	5.95	5.79	0.16
Total	15450.91									5445.27	5291.53	153.74

At the system level, a total input reduction of 5445.27 MWh/year is possible, with 5291.53 MWh/year corresponding to gains from efficiency and effectiveness, and 153.74 MWh/year corresponding to additional savings allowing for trade-offs of emission quotas among plants. The total saving potential estimated in step 3 corresponds to 35.24% of the original input consumption. Although the additional 1% gains from reallocation of emission quotas ($35.24\% - 34.25\%$) may seem a small gain (in percentage), the amount of energy saved can be a significant contribution towards a more sustainable operation, with important financial benefits.

Consider, for example, plants P34 and P35 which discharge the pollutants in the same river. Their emission quota for COD is 125 gO_2/m^3 and for SS is 35 g/m^3. From the reallocation analysis, some variations to these values translate, in practice, in lower values of energy consumption, while preserving the average amount of pollutants discharged in the river.

In terms of COD, P34 had the emission quota adjusted to 149.41 gO_2/m^3, while P35 had the emission quota adjusted to 110.25 gO_2/m^3. The average amount of pollutants discharged in the river weighted by the volume of water treated is:

$$\frac{149.41 \times 8184075 + 110.25 \times 13540767}{8184075 + 13540767} = 125 gO_2/m^3$$

For SS, P34 had the emission quota adjusted to 72 g/m^3, while P35 had the emission quota adjusted to 52.75 g/m^3. The average amount of pollutants discharged in the river weighted by the volume of water treated is:

$$\frac{72 \times 8184075 + 52.75 \times 13540767}{8184075 + 13540767} = 60g/m^3$$

With those adjustments, P34 decreases the energy consumption by 110.47 MWh/year while P35 increases the energy consumption by 1.00 MWh/year. With the reallocation, the decrease in the energy consumption of these plants is 109.47 MWh/year.

7.6. Conclusion

This paper proposes a combined evaluation of the efficiency and effectiveness of wastewater treatment plants (WWTPs) using a DEA-inspired approach. The methodology developed is new both in terms of the models used to measure performance and the challenges explored in the water sector. This research brings a novel perspective of effectiveness to the assessment of performance corresponding to the compliance with quotas imposed by regulation.

In the wastewater context, this involves fulfilling water quality levels set by legislation for the effluent discharged.

Two policy scenarios were considered in this study, requiring different modeling solutions. The first scenario is currently in force at the company, which dictates that the water quality should be measured at each WWTP. In this context, an input-oriented DEA model with peers restricted was proposed. The second scenario assumes that the water quality should be measured at the discharging areas. As several WWTPs are assigned the same natural watercourse for effluent discharge, the quality levels should be monitored considering the average amount of pollutants for DMUs sharing the same discharging area. This seems appropriate, as the main objective of ensuring high-quality effluents is to preserve the natural watercourses from potential negative impacts of the effluents discharged.

In this context, a DEA-based mixed-integer linear programming (MILP) model was developed. This model uses the same efficient frontier as a traditional DEA efficiency evaluation. However, it enables additional resource savings due to the trade-offs allowed in the emission quotas among plants discharging the effluent in the same watercourse.

With the classical efficiency analysis, we found that adjusting the energy consumption to efficient levels enables a reduction of 28% in the total energy consumed by the 41 WWTPs. Next, when meeting the exact emission quota at each WWTP, the efficiency and effectiveness analysis resulted in identifying a saving potential in energy consumption of 34% (28% corresponding to savings from efficiency and 6% from effectiveness). Finally, when the reallocation of emission quotas among the WWTPs is enabled, a total input reduction of 35% is achieved, meaning that allowing for deviations in the emission quotas saves a further 1% in energy consumption.

Overall, the methodology developed in this paper enables a comprehensive evaluation of performance considering efficiency alongside effectiveness (i.e., the attainment of goals aligned with regulatory requirements). Furthermore, to obtain a fair performance evaluation regarding comparability with peers, we propose using peer restrictions to estimate the efficient frontier. This approach is aligned with the ideas initially presented by Golany and Thore (1997), which is beneficial for small samples with high variability in the values of input and output indicators. These contributions may be appreciated by researchers in operations research and managers/authorities in the water sector.

The work developed has a limitation that can be seen as a doorway for future research. As discussed in Section 7.4, other resources (alongside energy) are used in the wastewater treatment process. It would be interesting to specify a DEA model with a comprehensive representation of all resources used in the wastewater treatment process. At the moment this is not possible due to data unavailability at the WWTP level, so systematic procedures of data collection need to be implemented to allow future studies on this topic.

Acknowledgements

Alda Henriques wishes to acknowledge the financial support from FCT – Fundação para a Ciência e a Tecnologia (Portuguese national funding agency for science, research and technology), allocated through the Grant PD/BD/142815/2018 respecting the Doctoral Program on Sustainable Energy Systems by MIT-Portugal.

References

Abbott, M. and Cohen, B. (2009). Productivity and efficiency in the water industry. *Utilities Policy*, 17(3):233-244.

Anthony, R. N., Dearden, J., and Bedford, N. M. (1989). Management control systems, sixth ed.

Asmild, M., Paradi, J. C., Reese, D. N., and Tam, F. (2007). Measuring overall efficiency and effectiveness using DEA. *European Journal of Operational Research*, 178(1):305-321.

Beasley, J. (2003). Allocating fixed costs and resources via data envelopment analysis. *European Journal of Operational Research*, 147(1):198-216.

Castellet, L. and Molinos-Senante, M. (2016). Efficiency assessment of wastewater treatment plants: A data envelopment analysis approach integrating technical, economic, and environmental issues. *Journal of Environmental Management*, 167:160-166.

Cazals, C., Florens, J.-P., and Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106(1):1-25.

Charnes, A., Cooper, W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6):429-444.

De Witte, K. and Marques, R. C. (2009). Capturing the environment, a metafrontier approach to the drinking water sector. *International Transactions in Operational Research*, 16(2):257-271.

D'Inverno, G., Carosi, L., Romano, G., and Guerrini, A. (2018). Water pollution in wastewater treatment plants: An efficiency analysis with undesirable output. *European Journal of Operational Research*, 269(1):24-34.

Directive, E. (1991). Council directive of 21. may 1991 concerning urban waste water treatment (91/271/eec). *J. Eur. Commun*, 34:40.

Dong, X., Zhang, X., and Zeng, S. (2017). Measuring and explaining eco-e_ciencies of wastewater treatment plants in china: An uncertainty analysis perspective. *Water Research*, 112:195-207.

Drucker, P. (1977). An Introductory View of Management. Harper College Press, New York.

D'Inverno, G., Carosi, L., and Romano, G. (2021). Environmental sustainability and service quality beyond economic and financial indicators: A performance evaluation of Italian water utilities. *Socio-Economic Planning Sciences*, 75:100852.

Gandiglio, M., Lanzini, A., Soto, A., Leone, P., and Santarelli, M. (2017). Enhancing the energy efficiency of wastewater treatment plants through co-digestion and fuel cell systems. *Frontiers in Environmental Science*, 5:70.

Golany, B. (1988). An interactive MOLP procedure for the extension of DEA to effectiveness analysis. *The Journal of the Operational Research Society*, 39(8):725-734.

Golany, B., Phillips, F. Y., and Rousseau, J. J. (1993). Models for improved effectiveness based on DEA efficiency results. *IIE Transactions*, 25(6):2-10.

Golany, B. and Tamir, E. (1995). Evaluating efficiency-effectiveness-equality trade-offs: A data envelopment analysis approach. *Management Science*, 41(7):1172-1184.

Golany, B. and Thore, S. (1997). Restricted best practice selection in DEA: An overview with a case study evaluating the socio-economic performance of nations. *Annals of Operations Research*, 73(0):117.

Griffin, R. W. (1987). Management, second ed. Houghton Mifflin Co., Boston.

Guerrini, A., Romano, G., and Campedelli, B. (2013). Economies of scale, scope, and density in the Italian water sector: a two-stage data envelopment analysis approach. *Water resources management*, 27(13):4559-4578.

Guerrini, A., Romano, G., Mancuso, F., and Carosi, L. (2016). Identifying the performance drivers of wastewater treatment plants through conditional order-m efficiency analysis. *Utilities Policy*, 42:20-31.

Güngör-Demirci, G., Lee, J., and Keck, J. (2018). Assessing the performance of a California water utility using two-stage data envelopment analysis. *Journal of Water Resources Planning and Management*, 144(4):05018004.

Gómez, T., Gémar, G., Molinos-Senante, M., Sala-Garrido, R., and Caballero, R. (2017). Assessing the efficiency of wastewater treatment plants: A double-bootstrap approach. *Journal of Cleaner Production*, 164:315-324.

Henriques, A. A., Fontes, M., Camanho, A. S., Amorim, P., and Silva, J. G. (2021). Opening the black box for performance evaluation of problematic samples: a robust nonparametric approach for wastewater treatment plants. *FEUP Working Paper*.

Hernández-Sancho, F., Molinos-Senante, M., and Sala-Garrido, R. (2011a). Energy efficiency in Spanish wastewater treatment plants: a non-radial DEA approach. *The Science of the total environment*, 409(14):2693-2699.

Hernández-Sancho, F., Molinos-Senante, M., and Sala-Garrido, R. (2011b). Techno-economical efficiency and productivity change of wastewater treatment plants: the role of internal and external factors. *J. Environ. Monit*., 13:3448-3459.

Hernández-Sancho, F. and Sala-Garrido, R. (2009). Technical efficiency and cost analysis in wastewater treatment processes: A DEA approach. *Desalination*, 249(1):230-234.

Lorenzo-Toja, Y., Vázquez-Rowe, I., Chenel, S., Marín-Navarro, D., Moreira, M. T., and Feijoo, G. (2015). Eco-efficiency analysis of Spanish WWTPs using the LCA + DEA method. *Water Research*, 68:651-666.

Molinos-Senante, M., Gémar, G., Gómez, T., Caballero, R., and Sala-Garrido, R. (2016a). Eco-efficiency assessment of wastewater treatment plants using a weighted Russell directional distance model. *Journal of Cleaner Production*, 137:1066-1075.

Molinos-Senante, M., Hernández-Sancho, F., Mocholí-Arce, M., and 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., Hernández-Sancho, F., and Sala-Garrido, R. (2010). Economic feasibility study for wastewater treatment: A cost-benefit analysis. *Science of the Total Environment*, 408(20):4396-4402.

Molinos-Senante, M., Hernández-Sancho, F., and Sala-Garrido, R. (2014b). Benchmarking in wastewater treatment plants: a tool to save operational costs. *Clean Technologies and Environmental Policy*, 16(1):149-161.

Molinos-Senante, M., Hern_andez-Sancho, F., and Sala-Garrido, R. (2015). Comparing the dynamic performance of wastewater treatment systems: A metafrontier Malmquist productivity index approach. *Journal of Environmental Management*, 161:309-316.

Molinos-Senante, M., Sala-Garrido, R., and Hernández-Sancho, F. (2016b). Development and application of the Hicks-Moorsteen productivity index for the total factor productivity assessment of wastewater treatment plants. *Journal of Cleaner Production*, 112:3116-3123.

Rogers, J. W. and Louis, G. E. (2007). A financial resource allocation model for regional water systems. *International Transactions in Operational Research*, 14(1):25-37.

Sala-Garrido, R., Hernández-Sancho, F., and 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., and Hernández-Sancho, F. (2011). Comparing the efficiency of wastewater treatment technologies through a DEA metafrontier model. *Chemical Engineering Journal*, 173(3):766-772.

Sala-Garrido, R., Molinos-Senante, M., and Hernández-Sancho, F. (2012b). How does seasonality affect water reuse possibilities? an efficiency and cost analysis. *Resources, conservation and recycling*, 58:125-131.

Sun, J., Wu, J., Liang, L., Zhong, R. Y., and Huang, G. Q. (2014). Allocation of emission permits using DEA: centralised and individual points of view. *International Journal of Production Research*, 52(2):419-435.

Thanassoulis, E. (2002). Comparative performance measurement in regulation: The case of English and welsh sewerage services. *The Journal of the Operational Research Society*, 53(3):292-302.

Wang, K., Zhang, X., Wei, Y.-M., and Yu, S. (2013). Regional allocation of CO2 emissions allowance over provinces in china by 2020. *Energy Policy*, 54:214-229.

Xie, Q., Xu, Q., Zhu, D., Rao, K., and Dai, Q. (2020). Fair allocation of wastewater discharge permits based on satisfaction criteria using data envelopment analysis. *Utilities Policy*, 66:101078.

Zhang, Y. and Xu, Z. (2019). Efficiency evaluation of sustainable water management using the HF-TODIM method. *International Transactions in Operational Research*, 26(2):747-764.

PART III

CONCLUSION

Conclusion

In order to draw some final conclusions, let us recall the overall problem to be tackled with this thesis work. This was to identify innovative ways able to expand the current research state-of-the-art toward the successful transformation of the wastewater sector into a fully sustainable part of the society's existence, organization, and activity.

The first aspect to highlight is that taking a system-wide approach for units' analysis and further optimization is a practice that is worth exploring more often in the research areas concerning the wastewater sector. This was a gap identified in the motivational literature review in Chapter 1. As such, it was demonstrated that taking a system-wide approach to previously individually-considered entities is beneficial in the pursuit of resource (including energy) efficiency improvement opportunities at the utilities and assets level (see Chapters 5-7). Similarly, the system-wide approach also constitutes a potential solution for wastewater's embedded-energy recovery improvements whilst contributing to other sustainability enhancements at the utility level (see Chapter 4). Furthermore, it was also demonstrated its usefulness at the regulator and utility level, leading to increased transparency, information objectivity and complementary to other less aggregated approaches (see Chapter 3).

Another aspect to point out stems from the motivational literature review conducted and corresponds to the current transition phase that is being experienced by the wastewater sector worldwide, although felt at different paces and triggered by different needs in developing and developed countries. Therefore, it was considered that the transition phase constitutes a great context for research in areas directly related to the wastewater sector. In this thesis, the attention was placed in structures that are typically present in developed countries, such as an organized regulator, utilities that are well established in their practices with a consistent and reliable service provision, as well as the assets that they manage and operate. Therefore, the research opportunity considered in this thesis was to tackle optimization problems in the pursuit of enhanced sustainability, and having the energy topic as a central focus.

Further conclusions were already drawn at the end of each chapter, particularly the ones composing the Part II of the thesis.

In the overall, this thesis contribution for the research field of the wastewater sector is a combination of multiple optimization modelling approaches that are mainly oriented to the management and operation of the service provision mechanisms at multiple levels of organizational structures, so that potential enhancements in sustainability, particularly related to the energy currency, are identified and thus able to be followed through.

In regards to future developments, some suggestions were identified in each chapter respecting the research papers. Nonetheless, departing from a more integrated vision upon the ideas presented throughout the thesis, an interesting future research topic could be to develop a combination of a DEA-based composite indicator and the balanced scorecard applied either to a set of wastewater treatment plants or to a set of utilities, to name some examples. The purpose would be to achieve an increased engagement and collaboration of stakeholders toward enhanced transparency, information sharing, and improved overall management so to derive better solutions for sustainability improvements.

Another proposal is the development of a network DEA model to better cope with the complexity of the wastewater treatment production activity developed in wastewater treatment plants. Having in mind the overall objective of maximizing the total energy balance of a system of plants (considering the energy balance as the total energy production minus the total energy consumption), my suggestion is to consider the sludge production as an output of the first DEA model, and as an input of the second DEA model. Furthermore, the energy consumption would be regarded as input of each DEA model and the energy production would be regarded as the output of the second DEA model, in the case where anaerobic digestion is the only technology for energy recovery.

ANNEX I

Literature review on sustainability challenges in the water sector: water-energy nexus, performance assessment, digitalization/ big data

Table I.1 provides a systematized overview of some of the studies focusing on the challenges and the discussion over the many ways sustainability can be more consistently attained within the wastewater sector worldwide under the call for transition that is triggered by the great drivers of climate change and population growth. We divided the studies reviewed into three main areas of knowledge to facilitate the understanding of the different pathways conducive to sustainability, which corresponds to an aggregated and simplified way of exposing information related to the five topics previously identified in section 1.1 since those five topics are very often intertwined among each other. The first area of knowledge comprises resource recovery, the water-energy nexus, and the more direct sustainability discussion. The second area refers to governance, management, and performance measurement. The third area of knowledge concerns the current and futuristic topic of big data analytics and computer science trends along with the corresponding influences over the wastewater sector development.

Table I.1 – Reviewed studies regarding the pursuit of a fully sustainable wastewater sector.

Resource recovery, water-energy nexus, and sustainability	Chae and Kang, 2013	• It is proposed a set of methodologies to estimate the energy independence of a municipal WWTP with a capacity of 30 000 m^3/d. • The methodologies allow for the proper selection of power capacity and technologies in order to capture the energy available of the so-called green energy resources through photovoltaics, small hydropower and effluent heat recovery system to balance and surpass all plant energy needs. • It is highlighted that the implementation of those type of systems, although necessary to improve the energy independence of WWTPs, is only economically viable under external financial support such as the case described of Korea. • It is recommended to further explore the heat embedded in the wastewater for district heating and cooling since its amount is superior to all treatment energy demand.
	Foladori et al., 2015	• It develops a methodology for conducting an energy audit that considers each stage, process, or unit of a WWTP, which includes a set of energy indicators. • It validates the proposed methodology through data collected continuously over two years in five WWTPs treating up to 10 000 PE in the North of Italy. • It identifies the causes of excessive energy consumption thus providing real opportunity for energy consumption reduction.
	Gandiglio et al., 2017	• It is conducted an overview of the methods for reducing energy consumption at WWTPs as well as the maximization of energy production towards energy positive WWTPs. • It is presented a case study respecting to the application of co-digestion of sewage sludge with the organic fraction of municipal solid waste (OFMSW), and the utilization of a high-efficiency solid oxide fuel cell (SOFC) as the combined heat and power (CHP) system in an Italian WWTP serving around 2.5 million of person equivalent (PE). • It is reported an increase of the electrical production of 228% compared to the situation of the WWTP before co-digestion and the use of SOFC.
	Gikas, 2017	• It is proposed an alternative treatment process to the conventional activated sludge that combines enhanced primary treatment for total solids removal and partial biochemical oxygen demand (BOD) removal, with additional secondary treatment for the remaining BOD and nitrogen removal with a set of adequate technologies. • The biosolids with the adequate dewatering and thermal treatments are subject to gasification or, alternatively, to anaerobic digestion, both processes having associated energy recovery in a co-generation system.

		• The proposed sequence of methods enables a pilot facility with a capacity of 380 m^3/d to gain in efficiency and energy recovery enough to surpass all energy needs.
	Gu et al., 2017	• It conducts an analysis of energy consumption and recovery in WWTPs and a characterization of the factors that influence the energy use such as treatment techniques, treatment capacities, and regional differences. • It describes current technological trends used in energy self-sufficient WWTPs along with the identification of major challenges in their implementation. • It presents some examples of energy self-sufficient WWTPs. • It highlights the use of benchmarking as a strategic means to enhance the energy efficiency of WWTPs.
	Gurung et al., 2018	• It describes an energy benchmarking analysis of 22 Finnish municipal WWTPs on the basis of a questionnaire and a subsequent statistical analysis, which demonstrated potential for significant unit energy consumption reduction. • It proposes a retrofitting strategy based on four emerging technologies to effectively achieve a state of energy neutrality, which was implemented in the case-study of a Finnish WWTP (Mikkeli WWTP) that treats about 5 million m^3 of wastewater per year.
	Hamiche et al., 2016	• It highlights the existing links between water and electricity across the three major functions of both industries, namely production, transportation, and consumption. • It acknowledges the multidimensional nature of the water-energy nexus, identifying five dimensions: environmental, technological, economic, political, and social. • It reviews papers focusing on the study of the water-energy nexus, by examining their objectives, scope, methodologies and main findings. • It asserts the need for widening the perspectives from which the nature of the water-energy nexus is analyzed, and, therefore, proposes a general research framework for further investigation upon that topic, which is based on Integral Theory.
	Iribarnegaray et al., 2012	• It proposes a framework for the construction of a water and sanitation sustainability index (WASSI) built upon a definition of sustainability that ensures the combination of territorial, temporal, and personal aspects. The objective is to contribute with an inclusive sustainability assessment method to be applied at local level. • It discusses the results of the application of the framework developed to the case study of the city of Salta, Argentina.
	Maktabifard et al., 2018	• It highlights the fact that the WWTPs are crucial factors in the water-energy nexus and presents a review of the existing approaches for energy neutrality achievement by WWTPs, including methods for energy consumption reduction, methods for increasing energy recovery, and external sources of renewable energy. • It highlights the relevance of energy performance evaluation in the identification of potential energy consumption reduction and energy conservation measures.

	• It identifies the decision-support methods for energy-related aspects of WWTPs.
Marques et al., 2015	• It reviews sustainability actions undertaken in different urban water contexts in several countries. • It discusses the sustainability concept and argues that additional dimensions (assets and governance) should be considered along with the triple bottom line (economic, social, and environmental dimensions) when assessing the sustainability of urban water services. • It proposes a set of criteria to evaluate different aspects of the urban water services so as to include all dimensions of sustainability. • It uses a multicriteria model to aggregate all proposed criteria that enables to calculate a sustainability score for each water service potentially assessed. • It describes the model structuring process using the case of the Portuguese water context.
Neugebauer et al., 2015	• It highlights the role of WWTPs as an economically competitive technology for recovery of energy from wastewater and to potentially provide energy surpluses to the vicinity of plants, according to demand, type of resources in excess, and in coordination with relevant stakeholders. It points out that thermal energy is one interesting possibility to provide to WWTP's external points in the urban or industrial neighborhood environment, as demonstrated in some northern European countries. • It presents a methodology to estimate thermal energy resource potential of WWTPs that considers the wastewater infrastructure in integration with regional energy supply concepts. • It demonstrates the proposed methodology using the Austrian case.
Nowak et al., 2015	• It discusses major ways of optimizing the energy balance of municipal WWTPs, including the use of thermal energy, anaerobic digestion and co-digestion, as well as the optimization of all mechanical equipment along with optimal aeration control. • It gives the example of two Austrian WWTPs that are energy self-sufficient and with surplus energy production.
Qadir et al., 2020	• It provides estimates on the global potential for resource recovery from wastewater, including water, nutrient, and energy, by making use of existing data basis available worldwide that cover 95% of the global population. • It highlights the crucial role of wastewater-embedded resources in the sustainable development goals (SDG) context along with the equally crucial value of wastewater-related data information collected at local-, regional- or country-level for policymakers, researchers, practitioners and public institutions. • It points out the factors that can both hinder or promote the progress on wastewater resource recovery and reuse, depending on how they are considered, explored, and/or implemented, advocating an integrated resource recovery approach across sectors within urban systems.

	Scott et al., 2011	• It expands the water-energy nexus discussion beyond the input-output relationship (resource linkages) into the realm of resource governance, policy, and climate change adaptation. • It highlights the need to consider both energy and water resources at multiple scales as well as multitiered institutional solutions, such as regulatory cooperation, based on the assertion of institutional and policy dimensions of water and energy coupling. • It argues that finding mechanisms for internalization of impacts along with identification of trade-offs between water and energy resources across regional areas are some of the central points for increased articulation between water and energy policies and institutions.
	Shen et al., 2015	• It analyzes the utilization of anaerobic digester technology in US WWTPs and highlights the favorable economic context for biogas production under the expanded renewable fuel standard (RFS2) to foster the increase of the biogas utilization at US WWTPs, which is currently very low. • It reviews the barriers, gaps and challenges in deploying biogas production technology. • It highlights the benefits of co-digestion to boost biogas production and overcoming some of the identified barriers to the deployment of biogas production technology.
	Solon et al., 2019	• It develops conventional activated sludge (CAS)-based unit WWTPs process models to be integrated in new plant layouts along with the results of simulations run on the basis of the same influent characteristics for all tested process schemes. • It demonstrates that there is potential for achieving future energy-positive water resource recovery facilities (WRRFs) only through using novel integrated combinations of existing CAS-based WWTP technologies.
	Spinosa et al., 2011	• It focuses on the development of innovative and integrated systems for sustainable management of sludge produced in WWTPs, which includes recovery of nutrients and other materials. • It proposes an integrated system for sludge treatment and management that includes anaerobic digestion, dewatering/drying, and pyrolysis/gasification processes with the inherent objective of recovering materials for reuse in agriculture or construction industry and also for energy purposes.
	Voulvoulis, 2018	• It argues that a transition to a circular economy paradigm could boost a wider adoption of water reuse as part of an integrated water management strategy so as to facilitate an alternate water supply solution for the increasing water stress related problems. • It examines the opportunities and risks associated with such economy approach. • It provides examples of water reuse programs around the world.

	Wett et al., 2007	• It generically describes the utilization of energy manuals for WWTPs by Central European countries and the benefits in terms of short- medium- and long-term optimization measures for energy savings that were thus possible to identify, plan and implement. • It highlights the case of Austria, which instead of utilizing energy manuals has adopted a national benchmarking program. A five-year period of the benchmarking practice allowed a decrease in 30% of energy costs, showing the enormous potential for energy efficiency improvements worldwide in systems that had not yet been subject to any optimization. • It describes the optimization of the two-stage biological treatment in the municipal WWTP Strass (Austria), which treats the load of 90 000 to more than 200 000 PE weekly averages, depending on tourist seasons. The optimization undertaken in this plant is solely based in strategic modifications in the overall treatment process and some technology improvements. It was reached a positive energy balance without the addition of any co-substrates.
Governance, management, and performance measurement	Abbott and Cohen, 2009	• It conducts a review of studies focusing on efficiency and productivity measures undertaken within the water and wastewater sectors in several countries over a period of more than twenty years. • The focus is to identify trends in optimal structures of operation within the water and wastewater industries, particularly in regards to economies of scale, scope, public versus private ownership, as well as the impact of regulation. • It identifies several future research trends that may have a relevant impact on the efficiency of the water and wastewater sector's operation, such as environmental management activities, the role of wastewater as a potential source of reusable water, and the relationship between water supply and urban planning.
	Admiraal and Helden, 2003	• It describes a benchmarking project conducted by the Dutch water boards to the wastewater treatment. The benchmarking project was based on the application of the Balance Scorecard methodology. • It compares the Dutch benchmarking project with the Best Value project in the UK.
	Burkhard et al., 2000	• It presents a review of management areas concerning the urban wastewater, namely the rainwater management, the domestic wastewater management, and the management of water re-use (both domestic and rain-water), by identifying the respective techniques, the efficiency of each technique in terms of purpose achievement, as well as the correspondent economic and social aspects. • It emphasizes the social dimension of the sustainability in water management. • It highlights the need to consider an integrated planning approach of all areas of the urban water and wastewater management, for which an adequate tool should be developed.
	Cantor et al., 2020	• It highlights that innovation in the municipal wastewater sector in the United States have been slow despite the existence of the drivers for change, such as urban growth, climate change, and aging infrastructure.

	• It states that wastewater utilities managers as well as wastewater regulators are the primary decision-makers leading to innovation in the wastewater sector. • It describes the application and results of a nation-wide survey to wastewater utilities managers and regulators aiming at understanding the key barriers to innovation in the wastewater sector. The objective is to understand how to guide improvements in regulation that promote innovation along with environmental protection.
Grigg, 2016	• It describes the water attributes that makes it both a sector and a connector to other sectors, so that a framework is created to model integrative water management and identify business opportunities. • It describes the drivers of change that are calling for innovative responses in regards to the water as a resource, as a sector and as a connector to other sectors. • It points out the need to focus of an integrated way on both the common public water sector activities (access to safe water and security against flooding) and the business opportunities for the private sector to intervene with innovation and structural support to the public services provision, namely infrastructure, equipment and supportive services to the water-handling sector and subsectors.
Marques et al., 2011	• It highlights the importance of benchmarking in the context of water and wastewater services, which are usually provided in monopolistic regimes that characterizes the public sector, to bring artificial competition toward increased motivation for efficiency improvements and innovation. • It surveys the use of benchmarking within more than 50 water regulatory contexts that cover the five world continents and describes the different types of benchmarking approaches identified as well as the main objectives to it (quality service regulation and tariff setting). • It exemplifies the different types of benchmarking used in case studies of several countries.
Murray et al., 2011	• It describes the evolving context of wastewater-related matters in India, China, and Ghana, and demonstrates that there are increasing expansion trends in both matters of access to sanitation as well as greater treatment and safe management of human excreta, as a result of a set of significant drivers of change. • It highlights the fact that governments are showing increasing interest in the engagement of the private sector, but points out the need for the prior development of adequate models of engagement. • It argues that business models focused on reuse-oriented sanitation might be the solution for the desirable increased engagement of the private sector since it provides an opportunity for reliable revenue that not depends on users' fees.
Nguyen et al., 2019	• It presents a critical overview of the aspects associated with conventional water management, and highlights the water-related problems that can no longer be tackled only through the old management way, with particular emphasis on developing countries. Major water-related problems are severe draughts and floods, which are linked to climate changes, and rapid urbanization, which demands for water-related adequate responses.

	• It then describes the evolution of the concept of Sponge City and presents the arguments in favor of this new water management comprehensive strategy, which is inherently linked with a sustainable urban planning, as a promising alternative to overcome major water-related problems along with ecological benefits. • It presents the principles of the Sponge City concept, and discusses the challenges to its successful implementation.
Picazo-Tadeo et al., 2011	• It reports a technical efficiency assessment of the different stages of the urban water cycle (i.e., water treatment and delivery; sewage collection; sewage treatment) in Andalusia (Southern European region), using directional distance functions and data envelopment analysis techniques. • It uses a sample data of 35 water and sewage utilities in the analysis. • It demonstrates that there is potential for delivering increased volumes of water, as well as treating increased volumes of wastewater without decreasing other outputs or increasing inputs of the processes, which is a key aspect within the context of water scarcity that characterizes the region of Andalusia.
Sarvari et al., 2021	• It identifies and analyzes the barriers to the private sector investment on the water and sewage industry (WSI) through the primary means of a literature review and interviews of experts on the related matters. • It highlights directions to overcome the obstacles previously identified, such as more transparent economic policies, increased interactions among public and private sector, and the adoption of a slow and steady approach to privatization, so as to facilitate the engagement of the private sector, which is recognized as a means of improving the performance of economic activities. • It highlights that the barriers to private sector investment are significantly neglected in developing countries such as Iran. • It points out the need for further research on context-specific factors that affect the development of private sector investment across different countries.
Seppälä, 2015	• It reflects on the concepts of metric and process benchmarking, as well as their distinction. • It is made a comparison among the Nordic benchmarking systems. • It highlights that a large amount of operational and process data is produced by water utilities but its information potential is under-explored for management and decision-making. Therefore, it is argued that utilities should prepare themselves for more sophisticated information and communication technology systems to a facilitated information transfer among systems and to enable holistic smart water management systems.
Silva et al., 2015	• It provides an overview of the factors that characterizes a general wastewater treatment system structure and operation to better understand the need for energy management systems. • It summarizes relevant information of the ISO 50001 energy management system framework and justifies its application to the specific case of the wastewater systems. • It establishes a connection between energy management and infrastructure asset management.

		• It presents a framework of energy performance indicators and indices for WWTPs.
	Tecco, 2008	• It conducts an analysis on the private actors' participation on the water sector along with the description of the economic features of the water sector functioning that undermines a more desirable intervention of the private sector. • It proposes a set of necessary conditions to promote the sustainable financial investment in the water sector in developing countries. • It highlights the importance of public governance and the commitment of all actors involved, including local people, to the creation of the appropriate environment for the private sector's sustainable financial support and the overall water services provision.
	Tupper and Resende, 2004	• It provides an efficiency assessment of water and sewage (WS) companies in Brazil using data from the period 1996-2000. The methodology used is data envelopment analysis (DEA) in a first stage to obtain the efficiency scores and then a Tobit regression model as an econometric treatment to control for regional heterogeneities. • It highlights the relevance of efficiency measurements as a form of yardstick competition in the case where companies, such as the WS, are under the state and municipal public firms' jurisdiction combined with a fuzzy regulatory structure. However, it also asserts that DEA-based measures cannot provide direct jurisdiction but with additional information in regards to the returns to scale of each specific company are able to identify relevant information to guide appropriate jurisdiction changes.
	Ujang and Buckley, 2002	• It summarizes the main topics addressed at the Kuala Lumpur IWA Conference, which took place on 29-31 October, 2001, to address critical issues of water and wastewater management in developing countries. • It points out the tendency of fragmentized and short-term-problem-solving-focused solutions that play out in developing countries to face water challenges derived from rapid urbanization, industrialization, population growth, and depleting water resources. • It proposes an integrated urban water management (IUWM) strategy in parallel with cleaner industrial production, waste minimization and financial arrangements to sustainably cope with those challenges. • It proposes the establishment of a Directorate for Water and Wastewater Management for developing countries.
	Walker et al., 2020	• It highlights the role of benchmarking to foster efficiency improvements of utilities. • It uses a double-bootstrap data envelopment analysis method to obtain bias-corrected energy and economic efficiency scores of water companies. • It undertakes an exogenous factors analysis using a truncated regression method to understand their impact on the efficiency measures. • It uses a sample of 17 water only and water and sewerage companies from England and Wales.

	Walsh et al., 2015	• It highlights the need for water to be managed as a valuable resource by industries and organizations for the reasons of increased cost efficiency and corporate social responsibility. • It highlights the water-energy nexus as a factor that adds relevance to the appropriate management of water. • It argues the widespread application of water standards similarly to the case of energy as a way to strategically cope with water-related challenges. • It describes the successful application of the standard ISO 50001 adapted to the case of water management at the University College Cork (UCC).
	Zayed et al., 2015	• It identifies the need for rehabilitation decision tools, given the context of U.S.A. and Canada WWTPs facing infrastructure deterioration due to aging and inadequate maintenance. • It develops a condition-rating index (CRI) model for prioritization of rehabilitation activities.
Big data analytics and computer science trends	Eggimann et al., 2017	• It focuses on reviewing and critically discussing data-driven urban water management (UWM) approaches as potential drivers and supporters of fundamental changes in the UWM toward more sustainable services. • It identifies the challenges associated with data-driven approaches. • It schematically represents the relation between data availability and correspondent utility in UWM.
	Garrido-Baserba et al., 2020	• It provides detailed information on the current topic of big data and its role as a lever for advanced enhancements in the urban water management complex problems such as urban water asset management, operation of urban water infrastructure, and recovering information on health and lifestyle of individuals and communities. • It points out the challenges related to big data management. • It establishes the link between big data and artificial intelligence in the context of the urban water sector. • It identifies socio-economic changes, cross-sector boundaries, and new research needs, as a result of the widespread influence of big data analytics.
	Moy de Vitry et al., 2019	• It extensively reflects on the risks associated with smart urban water cities and propose guiding principles to avoid potential pitfalls associated with the increased digitalization of the water sector.