To be cited as: Zeferino, J., Antunes, A.P., Cunha, M.C., "Regional wastewater systems design under population dynamics uncertainty", Journal of Water Resources Planning and Management (ASCE), 140 (3), pp. 322-331, 2014.
doi: 10.1061/(ASCE)WR.1943-5452.0000334

REGIONAL WASTEWATER SYSTEM PLANNING UNDER POPULATION DYNAMICS UNCERTAINTY

João A. Zeferino[1], António P. Antunes[2] and Maria C. Cunha[3]

Abstract

Regional wastewater systems are used for the collection and treatment of the wastewater generated in a region, and aimed at guaranteeing surface water quality. The volumes of wastewater to process depend on future population, and thus are affected by the uncertainty inherent to population dynamics. In this article, we present a robust approach to the planning of regional wastewater systems under population dynamics uncertainty. The approach searches for the optimal configuration of the sewer networks and for the best location, type, and size of the possible pump stations and treatment plants to include in the system. It assumes uncertainty to be described by a given number of discrete scenarios of known probabilities, and relies on discrete nonlinear optimization models whose objective is to minimize the expected regret of solutions with respect to total costs. As demonstrated through a case study developed for the North Baixo Mondego area, in central Portugal, the results obtained through the proposed approach provide clear insights into the wastewater system planning decisions to make and do not require excessive computational effort.

Keywords

Wastewater systems, population projections, robust optimization, simulated annealing

[1] Department of Civil Engineering, University of Coimbra, Polo II, 3030-788, Coimbra, Portugal

[2] Department of Civil Engineering, University of Coimbra, Polo II, 3030-788, Coimbra, Portugal

[3] Department of Civil Engineering, University of Coimbra, Polo II, 3030-788, Coimbra, Portugal

INTRODUCTION

Population growth and urbanization continue to take place at fast speed in many parts of the world, and this affects essential resources – in particular, water resources. As a consequence, water bodies are under heavy stress and contaminated with large volumes of pollutants, attributable to a great extent to domestic household sewage. Wastewater systems play an important role in guaranteeing surface water quality, which is vital for sustainable development. The investments needed to build, operate, and maintain such systems are often very large, but they can be fully recouped through the social benefits they are expected to generate (WBCSD 2008). For this to happen, it is essential that investment plans are made bearing in mind all the costs involved, preferably at regional level, to cash in on possible economic and environmental scale advantages (Cunha et al. 2009).

One of the main difficulties faced by wastewater system planners in their studies relates with the uncertainty affecting the volumes of wastewater to plan for, which in general closely depend on the population to serve in a distant future (the horizon for such studies is at least 20 years). Population projections are, therefore, an essential ingredient of wastewater system planning (as well as of any other infrastructure system planning activities). Typically, these projections lead to the definition of reference population values that do not properly reflect the uncertainty inherent to demographic dynamics. However, neglecting uncertainty in such studies can result in either over-conservative or over-optimistic solutions. To avoid this, a robust planning approach ensuring, by design, solutions less sensitive to uncertainty needs to be adopted (Mulvey et al. 1995). This kind of approach takes into account all (or most) possible realizations of the uncertain variables, enabling to find solutions that are close to optimal and feasible for all (or most) scenarios considered, but are not necessarily optimal in any of them.

The goal of this article is to present a robust approach to the planning of a regional wastewater system (i.e., the set of facilities operated to collect and treat the wastewater generated in a region) under population dynamics uncertainty. The system comprises the following types of facilities: wastewater treatment plants (WWTP), to process the wastewater before it is discharged into a river; sewer networks connecting the population centers with the WWTP; and pump stations, to lift wastewater if it is unfeasible or uneconomic to drain it by gravity. The approach we have adopted assumes uncertainty to be described by a given number of discrete scenarios of known probabilities, and relies on discrete nonlinear optimization models whose objective is to minimize the expected regret of the solution with respect to total costs while ensuring that the volumes of treated wastewater discharged from the WWTP into the river do not compromise water quality standards. Regret is the deviation between the payoff (total costs in this case) of a solution selected with limited information and the best payoff that could be obtained if all information was available at the time the solution was selected (Loomes and Sugden 1982). Three optimization models are proposed, each one handling uncertainty in a different way. The models are solved through a heuristic method combining a simulated annealing algorithm with a local improvement procedure. The results that can be obtained through the approach are illustrated for a case study involving the North Baixo Mondego area, in central Portugal.

LITERATURE OVERVIEW

Regional wastewater system planning models and population projection methods are the two main bodies of literature implicated in this article. Below, in separate subsections, we provide overviews for both.

Regional wastewater system planning models

The history of the application of optimization models to regional wastewater system planning is already quite long, dating back to the 1970s. The best-known early contributions have dealt with the problem of finding a minimum-cost solution for the sewer network, or more generally, the wastewater system to be installed in a region (e.g., Converse 1972, Deininger and Su 1973, and Joeres et al. 1974), in some cases taking into account quality issues in the receiving water bodies (e.g., McNamara 1976, Smeers and Tyteca 1982, and Vieira 1989). These contributions, which are thoroughly reviewed in Melo and Câmara (1994) and Whitlatch (1997), either used optimization models that overly simplify the planning problems under consideration or employed greedy methods prone to miss global optima (or both). This started to change with the works of Sousa et al. (2002) and Wang and Jamieson (2002), which are, to the best of our knowledge, the first ones where modern heuristic methods (respectively, simulated annealing and genetic algorithms) were applied to regional wastewater system planning models. The majority of recent articles on this and similar subjects apply the same types of methods. This includes Zechman and Ranjithan (2007), Álvarez-Vasquez et al. (2008), and Brand and Ostfeld (2011) on the genetic (or evolutionary) algorithm side, and Cunha et al. (2009) and Yeh et al. (2011) on the simulated annealing side. Genetic algorithms have also been used for solving regional waste load allocation models (e.g., by Cho et al. 2004, Yandamuri et al. 2006, and Aras et al. 2007), a kind of models closely related with the ones used in regional wastewater systems planning.

A feature common to all the works quoted above (and many others) is that they are deterministic – i.e., they do not take uncertainty into account. To the best of our knowledge, the only article concerned with regional wastewater system planning where this occurs is Zeferino et al. (2012). There, three robust optimization models are proposed, corresponding to three different ways of capturing uncertainty in a particular variable – the flow of the river where the treated wastewater is discharged. These models are inspired by the robust optimization framework introduced in Mulvey et al. (1995). The uncertainty is handled through scenario planning, that is, is represented with a set of possible states of the world called scenarios whose probability is assumed to be known (Rockafellar and Wets 1991). The same kind of models has been recently applied to other water-related problems, namely by Maeda et al. (2010), to handle a waste load allocation multi-objective problem considering uncertainties in the inflow rates, by Rosenberg and Lund (2009), to address water supply system enhancements and conservation actions taking into account water shortage uncertainties, and by Cunha and Sousa (2010) to design robust water supply systems capable of responding to uncertain extreme events.

Population projection methods

Population projections are an essential ingredient of planning studies (in particular those involving infrastructure systems), being the subject of an extremely vast literature that includes, among many others, renowned textbooks by Newell (1988), Smith et al. (2001), and Rowland (2003), as well as a new wide-ranging book edited by Stillwell and Clarke (2011). A recent review by Booth (2006) sets out three types of methods:

(trend) extrapolation methods, which predict the future based on historical patterns; expectation methods, which resort to subjective prospects; and explanation methods, which rely on structural models. The latter include component methods that combine projections of births, deaths, and migrations to update a population. The cohort-component method is a type of method developed in terms of gender and age groups, which is useful in planning situations where a detailed knowledge of the population characteristics is needed. However, it is worth noting here that the more complex methods – namely, the component methods – do not necessarily lead to more accurate forecasts of total population than those achieved with simpler methods (Smith 1997). The reason is mainly because there is some irreducible level of uncertainty about the future that no method can cope with, no matter sophisticated it is. Keilman (2008) came to a similar conclusion that demographic forecasts are intrinsically uncertain after realizing that the population projections made by several statistical agencies are no more accurate today than they were twenty-five years ago.

Among the factors that influence the accuracy of population projections are the time frame and the spatial aggregation. The degree of uncertainty increases and projections are more inaccurate for smaller regions (errors tend to cancel each other out over large space scales) and long term horizons, leading to values affected by considerable uncertainty (Smith et al. 2001). The growth rate of a region's population depends on what occurs in the country as a whole. However, the internal variability of a demographic trend at regional scale is larger and more complex than at national scale, and fewer works dealing with small-area projections have been published (Wilson and Bell 2007). There are several causes that might generate significant internal migration within the urban centers of a region, even if this does not affect the region's total population. After evaluating the accuracy of small-area population projections, Murdock et al. (1991) suggested that growth patterns are inclined to be accentuated or muted by the demographic characteristics of an area, and thus presented relevant groups of characteristics.

A possible way of improving population projections is to take advantage of ex post forecasting errors, assuming that future errors can be drawn from the same statistical distribution as past errors. Keyfitz (1981) and Stoto (1983) pioneered the study of this topic. For instance, the distribution of past errors can be used to construct empirical confidence intervals for population forecasts (Smith and Sincich 1988, and De Beer 2000). Small-area forecasts can also take advantage of ex post forecasting errors as proposed by Tayman et al. (1998) and Rayer et al. (2009) to obtain confidence intervals for county and sub-county areas.

OPTIMIZATION MODELS

In this section, we present the three optimization models upon which the proposed robust approach to regional wastewater system planning under population dynamics uncertainty is based. We also provide essential information on the method used to solve them.

The first model (designated as expected-regret model) extends the deterministic model described in Cunha et al (2009) to a stochastic formulation, making use of scenario planning to find solutions that are expected to perform well (i.e., are close to optimum and/or to feasibility) under the set of possible future states of the world. The objective consists in minimizing the expected regret of the solution with respect to total costs (that is, the expected deviation between the costs of a solution selected with limited information and the minimum costs that could be obtained if all information was available at the time the solution was selected). The second model (alpha-reliable model) pursues the same expected-regret minimization objective but taking only into account a fraction of scenarios with a given global probability of occurrence, α. The third model (beta-reliable model) also pursues the expected-regret minimization objective but facilities (specifically, sewers) are required to work in proper conditions only in a fraction of scenarios with a given global probability of occurrence, β. In all the three models, water quality concerns are handled through constraints on the maximum volume of wastewater discharged from each WWTP.

Expected-regret model

This optimization model aims to find the solution whose expected regret with respect to total costs is as low as possible, while working properly in all scenarios considered.

Using the notation introduced in Table 1, the model can be formulated as follows:

$$\text{Minimize } W \tag{1}$$

subject to:

$$QR_{is} + \sum_{j \in N_S \cup N_I} Q_{jis} = \sum_{j \in N} Q_{ijs}, \quad i \in N_S, s \in \boldsymbol{S} \tag{2}$$

$$\sum_{j \in N_S \cup N_I} Q_{jls} = \sum_{j \in N} Q_{ljs}, \quad l \in N_I, s \in \boldsymbol{S} \tag{3}$$

$$\sum_{j \in N_S \cup N_I} Q_{jks} = QT_{ks}, \quad k \in N_T, s \in \boldsymbol{S} \tag{4}$$

$$\sum_{i \in N_S} QR_{is} = \sum_{k \in N_T} QT_{ks}, \quad s \in \boldsymbol{S} \tag{5}$$

$$\sum_{p \in T} V_{kp} \leq 1, \quad k \in \boldsymbol{N_T} \tag{6}$$

$$Q\min_{ij} X_{ij} \leq Q_{ijs} \leq Q\max_{ij} X_{ij}, \quad i \in \boldsymbol{N_S} \cup \boldsymbol{N_I}, j \in \boldsymbol{N}, s \in \boldsymbol{S} \tag{7}$$

$$QT_{ks} \leq \sum_{p \in T} QT\max_{kp} V_{kp}, \quad k \in \boldsymbol{N_T}, s \in \boldsymbol{S} \tag{8}$$

$$R_s = C_s - C_s^{\,*}, \quad s \in \boldsymbol{S} \tag{9}$$

$$C_s = \sum_{i \in N_S \cup N_I} \sum_{j \in N} C_{ijs}\left(Q_{ijs}, L_{ij}, E_{ijs}, X_{ij}, Y_{ij}\right) + \sum_{k \in N_T} \sum_{p \in T} C_{kps}\left(QT_{ks}, V_{kp}\right), \quad s \in \boldsymbol{S} \tag{10}$$

$$W = \sum_{s \in S} P_s R_s \tag{11}$$

$$Q_{ijs}, E_{ijs} \geq 0, \quad i \in \boldsymbol{N_S} \cup \boldsymbol{N_I}, j \in \boldsymbol{N}, s \in \boldsymbol{S} \tag{12}$$

$$QT_{ks} \geq 0, \quad k \in \boldsymbol{N_T}, s \in \boldsymbol{S} \tag{13}$$

$$V_{kp} \in \{0,1\}, \quad k \in \boldsymbol{N_T}, p \in \boldsymbol{T} \tag{14}$$

$$X_{ij}, Y_{ij} \in \{0,1\}, \quad i \in \boldsymbol{N_S} \cup \boldsymbol{N_I}, j \in \boldsymbol{N} \tag{15}$$

$$R_s, C_s \geq 0, \quad s \in \boldsymbol{S} \tag{16}$$

The objective function (1) of this model minimizes the expected regret of the solution (represented with *W*) with respect to total costs. Constraints (2), (3), and (4) are the continuity equations for three types of network nodes: wastewater sources, possible intermediate nodes (representing slope changes and possible sewer intersections), and possible WWTP locations (Figure 1). Constraints (5) ensure that all the wastewater generated in the region will be treated at some WWTP. Constraints (6) guarantee that there will be at most one WWTP, of a specific type, in each possible WWTP location. Constraints (7) ensure that the flow carried by sewers will be within given minimum and maximum desirable values. These values depend on the diameter and slope of sewers, and on flow velocity requirements. The calculations needed to determine the diameter and slope of sewers are performed using a hydraulic simulation model. Constraints (8) ensure that the wastewater sent to any WWTP will not exceed given maximum values. These values depend on the water quality standards defined for the river where the wastewater is discharged and vary with the type of WWTP. Constraint (9) defines the regret associated with each scenario in terms of the (total discounted) costs of the solution, as discussed previously (note that C_s^* are model parameters, and their values should have been previously calculated). Constraints (10) specify the costs of solutions, separating between sewer network and pump station costs (C_{ijs}) from WWTP costs (C_{kps}). The former depend, for each sewer and scenario, on wastewater flow, sewer length, and hydraulic head difference. The latter depend, for each type of WWTP and scenario, on wastewater flow. Both types of costs include capital (setup) and operation costs, properly discounted. Constraint (11) defines the expected regret of the solution to implement, summing over the scenario set the probability of each scenario multiplied by the respective regret. Expressions (12) to (16) specify the domain of the decision variables.

Alpha-reliable model

This optimization model is aimed at finding a solution that also minimizes the expected regret with respect to total costs, but disregarding the most unfavorable scenarios in terms of expected regret according to the α-reliable concept. This concept was introduced by Daskin et al. (1997) in relation with maximum regret minimization problems (minimax regret). The set of scenarios to take into account (called reliability set) is endogenously determined with a global probability of occurrence $\alpha < 1$, defined by the decision-maker. The more averse to risk the decision-maker is, the higher should be the value of α.

In the formulation of this model, constraints (7), (8), and (11) of the expected-regret model are replaced with the following ones:

$$Q\mathrm{min}_{ij} X_{ij} Z_{\alpha s} - M_{Q\mathrm{min}} X_{ij}\left(Z_{\alpha s}-1\right) \le Q_{ijs} \le Q\mathrm{max}_{ij} X_{ij} Z_{\alpha s} - M_{Q\mathrm{max}} X_{ij}\left(Z_{\alpha s}-1\right), \quad i \in N_S \cup N_I, j \in N, s \in S \tag{17}$$

$$QT_{ks} \le \sum_{p \in T} QT\mathrm{max}_{kp} V_{kp} Z_{\alpha s} - M_{QT}(Z_{\alpha s}-1), \quad k \in N_T, s \in S \tag{18}$$

$$W = \sum_{s \in S} P_s R_s Z_{\alpha s} \tag{19}$$

$$\sum_{s \in S} P_s Z_{\alpha s} \ge \alpha \tag{20}$$

$$Z_{\alpha s} \in \{0,1\}, \quad s \in S \tag{21}$$

Constraints (17) ensure that, for the scenarios included in the reliability set (the subset of scenarios over which the regret is computed), the flow carried by sewers will be within given minimum and maximum desirable values. M_{Qmin} and M_{Qmax} are constants that must be set small and large enough, respectively, so that the size of sewers will not be dependent on scenarios excluded from the reliability set. Constraints (18) ensure that, for the scenarios included in the reliability set, the wastewater sent to any WWTP will not exceed given maximum values. M_{QT} is a constant that must be set large enough so that the maximum capacity of a WWTP will only be applied to scenarios included in the reliability set. Constraint (19) defines the expected regret of the solution to implement, considering only the scenarios included in the reliability set. Constraint (20) states that the global probability of occurrence of these scenarios is at least α. Expressions (21) specify the domain of the additional decision variables ($Z_{\alpha s}$).

Beta-reliable model

This optimization model is aimed at finding a solution that minimizes the expected regret with regard to costs for all scenarios (similarly to the expected-regret model), but requiring each sewer to work in proper conditions only in a fraction of scenarios endogenously determined with a global probability of occurrence $\beta < 1$, defined by the decision-maker. The more concerned with the sewers' level of service the decision-maker is, the higher should be the value of β. By proper conditions we mean, in this context, that the maximum desirable flow in a sewer, $Q_{max_{ij}}$, which corresponds to a depth of flow no larger than 0.50 of the sewer's diameter, is not exceeded (such depth of flow still ensures good ventilation and prevents septicity problems in the sewers). In the remaining fraction of scenarios (with a global probability of occurrence of 1-β), the depth of flow is allowed to be up to 0.94 times the sewer's diameter, which corresponds to the maximum tolerable flow in a sewer, $Q_{MAX_{ij}}$.

In the formulation of this model, constraints (7) of the expected-regret model are replaced with the following ones:

$$Q_{min_{ij}} X_{ij} \le Q_{ijs} \le Q_{max_{ij}} X_{ij} Z_{\beta s} - Q_{MAX_{ij}} X_{ij} \left(Z_{\beta s} - 1\right), \quad i \in N_S \cup N_I, j \in N, s \in S \tag{22}$$

$$\sum_{s \in S} P_s Z_{\beta s} \ge \beta \tag{23}$$

$$Z_{\beta s} \in \{0,1\}, \quad s \in S \tag{24}$$

Constraints (22) ensure that the flow carried by sewers will be above a given minimum desirable value in any scenario, below a maximum desirable value in the scenarios included in the reliability set, and below a maximum tolerable value in the scenarios excluded from the reliability set. Constraint (23) states that the global probability of occurrence of the scenarios included in the reliability set is at least β. Expressions (24) specify the domain of the additional decision variables ($Z_{\beta s}$).

Model solving

The optimization models presented in the previous section are nonlinear and include discrete decision variables. Even for small-scale instances, such models can be extremely difficult to solve to exact optimality, and need to be handled through heuristic methods. In line with the work we have been developing with respect to regional wastewater system planning (Cunha et al. 2009, Zeferino et al. 2009, and Zeferino et al. 2012), we used a heuristic method combining a simulated annealing (SA) algorithm and a local improvement (LI) procedure to solve the model. The SA algorithm was proposed by Kirkpatrick et al. (1983) and has been applied to a wide range of problems. Comprehensive information about the SA algorithm and its evolution through time can be found in Eglese (1990) and Suman and Kumar (2006).

The heuristic method starts with the SA algorithm. Accordingly, candidate model solutions chosen at random in the neighborhood of some incumbent solution are sequentially generated. For each candidate solution, the hydraulic simulation model is applied to design sewers (using the Manning equation), possible pump stations, and WWTP complying with all relevant regulations, and then the total costs of the regional wastewater system are calculated. Neighborhood moves to solutions better (less costly) than the incumbent solution are always accepted. The SA algorithm attempts to avoid being trapped in a local optimum by sometimes accepting candidate solutions worse than the incumbent solution. The transition between solutions is regulated by a parameter called temperature, according to a cooling schedule. Initially, even very negative transitions will be accepted, but, as temperature falls, the acceptance of such transitions will become increasingly less frequent. The SA algorithm proceeds until the incumbent solution ceases to improve, and then the LI procedure starts. This procedure searches all the solutions in the neighborhood of the incumbent solution and moves into the best (least cost) of these solutions if it is better than the incumbent solution. By doing this in successive iterations until no further total cost reductions are achieved, the LI procedure is expected to improve on the solution obtained by the SA algorithm.

For detailed information about the implementation of this type of heuristic algorithm for solving regional wastewater system planning models, the reader is referred to Zeferino et al. (2009).

CASE STUDY

The results that may be obtained by applying the optimization models presented in the previous section are illustrated below with a case study involving the North Baixo Mondego area, in central Portugal (Figure 2). This area occupies a territory of 1,222 km^2 on the right banks of River Mondego, the longest all-Portuguese river. It is divided in 6 municipalities and subdivided in 56 communities (“freguesias”, the smallest administrative unit in Portugal). In 2001, date of the latest census (when the study was made), the total population of this area was 229,625. The terrain is quite flat downstream of Coimbra, the largest city located in the area, and particularly along the Mondego banks, which is characterized with an intense farming activity. The only exception is the Boa Viagem Hills area, near the Atlantic Ocean, north of the second-largest city, Figueira da Foz. Upstream of Coimbra, terrain is much rougher, reaching a maximum altitude of more than 500 m, and forestry is the predominant activity.

The case study consists in determining a planning solution for the wastewater system of the North Baixo Mondego area taking into account the uncertainty that affects the evolution of its population. The horizon year considered in the study was 2021, which signifies a 20-year time span with respect to the date of the latest census. At present, the region is served by 8 separate systems, each one built around one WWTP. Since a large part of the existing infrastructure is expected to be in poor condition in 2021, we have decided to plan for a whole new system (note, however, that considering some parts of the existing infrastructure as possible components of the future system would not pose any significant challenge). Detailed information about the data and results of this case study, as well as about computational issues, is provided below in consecutive subsections.

Study Data

The data used in the study can be organized into three categories: network; population; and costs.

Network Data

The network consists of three types of features: wastewater sources; possible sewer networks; and possible WWTP locations.

The wastewater sources were assumed to be the 56 communities of the North Baixo Mondego area, represented with their respective geometric centers (Figure 3). The daily volume of wastewater generated in each community was taken to be 200 liters per inhabitant. The base sewer network (i.e. the superset of possible sewer networks) was defined to allow for direct connections between each community and neighboring communities or intermediate nodes representing slope changes and possible sewer intersections. In total, this network comprises 77 nodes (including 21 intermediate nodes) and 482 possible sewers. The possible WWTP locations were considered to be the locations of the existing plants, which, according with what was said before, were assumed to cease operations. The maximum capacity for a WWTP was taken to be 30,000 cubic meters per day (150,000 inhabitants), that is, the capacity of the largest plant currently operating in the area.

Population Data

The population data employed in the case study were obtained in two stages: first, using (linear) regression analysis, we modeled the population growth rates (PGR) of the 105 communities of the Baixo Mondego region as a function of their characteristics (we used the whole Baixo Mondego region for this purpose – and not only its northern part – to have a larger data sample and increase the accuracy of population projections); second, using the regression model and taking into account the properties of the respective error term, we generated 20 scenarios with the same 5% probability for the population of the communities of the North Baixo Mondego area in 2021.

The regression model for the PGR of the Baixo Mondego communities was estimated with data from 1981, 1991, and 2001 considering the following community characteristics as independent variables: PGR in the previous census period; distance to the region main town (Coimbra); distance to the municipality main town; total population; population density; average population age; literacy rate; economically active population rate; and unemployment rate (thus following the recommendations contained in Murdock et al., 1991). Since some of these characteristics are correlated, we applied linear regression in a (backward) stepwise way, discarding in each step the less significant variable until all variables were statistically significant (Draper and Smith, 1998).

The result was as follows:

$$\begin{aligned} PGR_T &= 62.8024 + 0.1856 \times PGR_{T-1} - 0.7548 \times DST \\ &\quad - 2.9052 \times DNS_t - 1.4337 \times AGE_t + \varepsilon_T \end{aligned} \qquad \left(R^2_{adj} = 0.31\right) \tag{25}$$

where PGR_T is the population growth rate in period T (starting in year t); DST is the (road) distance to the municipality main town; DNS_t is the population density in year t, and AGE_t is the average population age in year t; and ε_T is the error term.

This regression model shows that, *ceteris paribus*, the PGR of the communities of the Baixo Mondego region in a given period is larger the larger the PGR in the previous period, the shorter the distance to the municipality main town, and the lower the population density and average age of the community – which is generally in line with what could be expected. It also shows that, in this region, population forecasts are subject to considerable uncertainty, thus reinforcing the need of using a robust approach in the planning of its infrastructure systems. Indeed, despite the large number of community characteristics considered in the analysis, the adjusted correlation coefficient, R^2_{adj}, was just 0.31, meaning that those characteristics explain only 31 percent of the variation observed in the data. The large dispersion of PGR values around the expected values (given by the regression model with $\varepsilon_T = 0$) is illustrated in Figures 4 and 5, where we show the observed values of PGR plotted against the modeled values and a histogram of the PGR regression residuals (realizations of the error term), respectively. This histogram clearly suggests that the error term follows a normal distribution.

The generation of the population scenarios was also accomplished in two stages: first, the regression model (25) was used to design 20 PGR scenarios for the communities of the North Baixo Mondego area in the periods 2001-2011 and 2011-2021, with realizations of the error term drawn at random from a normal distribution with a mean of -1.4885 and a standard deviation of 8.9358 (note that the mean would be zero if it referred to all the communities of the Baixo Mondego region used in the regression analysis); second, for each scenario, the populations of the communities in 2021 were obtained by multiplying the respective PGR in the periods 2001-2011 and 2011-2021 with the respective population in 2001 and projected population in 2011.

Cost Data

The cost data needed to run the case study consists of the total costs for the 20 optimal wastewater systems corresponding to the population scenarios considered. These values were obtained by solving 20 times, one for each scenario, the deterministic wastewater system planning model presented in Cunha et al. (2009), considering the maximum capacity for a WWTP to be 30,000 cubic meters per day. An SA algorithm enhanced with an LI procedure similar to the one described previously was used for solving the model. The cost values obtained for the various scenarios are displayed in Table 2. The total costs of the optimal systems range between 39.73 and 43.79 M€ (for a discount rate of 4% per year), and the capital costs between 28.36 and 30.88 M€, the average costs are 41.37 and 29.25 M€, and the mean deviations with respect to the average are 0.82 and 0.55 M€ (1.97 and 1.88%), respectively. An example of optimal wastewater system for one of the scenarios – specifically Scenario 19 (the highest-cost scenario) – is shown in Figure 6.

Study Results

The application of the three optimization models to the case study led to the results presented and compared below.

Expected-regret model

The results obtained through the expected-regret model are displayed in Table 3 and Figure 7, together with the results obtained through the other two models. The capital costs for the optimal wastewater system are 31.96 M€, and the annual operation costs vary between 0.84 M€ and 0.95 M€, depending on which scenario actually occurs. This corresponds to total costs ranging between 44.25 M€ to 45.16 M€. Such costs are 3.1 to 11.4% higher than the minimum costs obtained for the different scenarios (Table 2), and 8.1% (3.33 M€) higher on average. The additional costs are justified with the fact that the optimal system obtained through the expected-regret model is fully reliable (works in proper conditions for all scenarios), whereas the optimal systems for the different scenarios are not – even the most expensive one, corresponding to Scenario 19, would not work properly if any one of the other scenarios arise. The optimal system configuration for the expected-regret model is displayed in Figure 7 (top). This system comprises 7 WWTP, 2 of them – the ones located near Coimbra and Figueira da Foz – with vast catchment areas, and 19 pump stations. The total length of sewer networks is 179.93 km.

Alpha-reliable model

The α-reliable model was applied to the case study considering an α value of 0.90. This means that the optimal wastewater system will work in proper conditions for a set of scenarios with a global probability of occurrence of at least 90%, and that the remaining scenarios are disregarded. The total costs of the system range between 43.19 to 43.97 M€, of which 31.21 M€ correspond to capital costs (Table 3). It is not surprising that the total costs of this system are lower (1.07 M€ on average) than the costs of the optimal system obtained through the expected-regret model – the latter correspond to a system that will work in proper conditions for all scenarios. Two scenarios are excluded from the reliability set – Scenarios 12 and 19, coincidentally the ones corresponding to the two most expensive solutions (note that this had not to occur as the objective is to minimize expected regret). The optimal system configuration for the α-reliable model is shown in Figure 7 (middle). It comprises 4 WWTP (again, 2 of them, in the same locations as before, with large catchment areas), 21 pump stations, and sewer networks with a total length of 187.73 km. The main difference with respect to the solution obtained through the expected-regret model is the reduction of the number of WWTP from 7 to 4. This happens because, in this solution, the largest WWTP, located close to Coimbra, is near the maximum capacity. When the two scenarios are excluded from the reliability set, this WWTP becomes able to receive the wastewater generated in more communities, and it is possible to avoid the construction of 3 (small) WWTP. Such change implies the installation of more pump stations and a longer sewer network, but this is worthwhile because of the savings on the WWTP.

Beta-reliable model

The β-reliable model was applied to the case study considering a value of β of 0.75. Therefore, sewers are required to work in proper conditions only in scenarios with a global probability of occurrence of 75%, and can work in inadequate conditions in the remaining 25% – but still need to work (flow depth should not exceed 0.94 of the diameter of the sewer). The total costs of the wastewater system range between 44.03 to

44.93 M€, of which 31.72 M€ correspond to capital costs (Table 3). The decrease in total costs with respect to the optimal system obtained through the expected-regret model is 0.22 M€ on average. The optimal system configuration for the β-reliable model is displayed in Figure 7 (bottom). It comprises 7 WWTP, 19 pump stations, and sewer networks with a total length of 178.05 km (the darker sewers are those that will work in inadequate conditions with respect to ventilation and septicity if some scenarios occur). This configuration is very similar to the one obtained through the expected-regret model – the difference is that some sewer networks are shorter and some sewers have a smaller diameter.

Results summary

As one could expect, there exists a tradeoff between the cost and the reliability of the wastewater system solutions provided by the three models. The expected-regret model has led to the most expensive solution because it provides a wastewater system that is fully reliable, working properly in all scenarios, even the most unfavorable ones. With the beta-reliable model, solutions are less costly but also less reliable than with the expect-regret model, because the system, in particular the sewer network, is allowed to work in inadequate conditions for a fraction of scenarios. The alpha-reliable model is the one that has led to the less expensive solution, but the wastewater system will fail if the most unfavorable scenarios occur. The expect-regret model is therefore the model that the more risk-averse decision-makers should use, whereas the more risk-prone decision-makers will tend to prefer the alpha-reliable model.

Computational Issues

The models used in the case study were solved in a 2.50 GHz Intel Core2 Quad computer with 2 GB of RAM running *Microsoft Windows 7* through the *OptWastewater* software package (developed at the University of Coimbra for research purposes). The computation time taken to solve the deterministic model was around one minute for each scenario. The expected-regret model took around 20 minutes, and the β-reliable model took about 5 minutes more. These are very reasonable computational efforts given the large size of the models. In contrast, the α-reliable model has taken around 4.5 hours to solve. The reason for this to happen is because, in the case of this model, and unlike with the β-reliable model, it is necessary to analyze which scenarios are excluded from the reliability set to calculate expected regret, and there are numerous combinations of such scenarios.

CONCLUSION

Until some time ago, infrastructure system planning solutions were generally designed through deterministic approaches, often complemented with sensitivity analyses. Recently, there is a clear trend to incorporate uncertainty issues explicitly into planning processes, because uncertainty is no longer accepted as an excuse for infrastructure solutions that fail to perform correctly.

The robust approach to regional wastewater system planning proposed in this article tallies with this trend. Among the many types of models upon which such approach could rely (see Snyder 2006), we opted for models aiming at minimizing the expected regret of solutions with respect to total costs. The solutions provided through these types of models are, by design, close to the yet unknown optimal solutions (they depend on which future scenario will prevail), while being feasible for all (or most) scenarios under consideration. The models therefore contemplate two of the main concerns of decision-makers: costs and reliability. As demonstrated through the case study developed for the North Baixo Mondego area, in central Portugal, the results obtained through the models provide clear insights into the planning decisions to make and do not require excessive computational effort (particularly when all scenarios are to be taken into account).

The source of uncertainty considered in the optimization models upon which the robust approach is based is population dynamics. This is an important uncertainty source that planners need to face, including because of the impact climate change may have on the spatial distribution of population, but it is far from being the only one. Other major sources of uncertainty are costs and river flows. In order to increase the practical usefulness of the proposed robust approach, the various sources of uncertainty need to be dealt with at the same time. On the basis of what we have learned up to now, expected-regret models and derivatives, such as the α- and the β-reliable models, should be valuable tools for this endeavor. But, of course, the underlying models will be more complex and the number of scenarios to take into account will necessarily increase, thus the models will be more difficult to solve. Our future efforts along the line of research where this article fits into will surely include some work on these more complex models.

ACKNOWLEDGMENTS

The authors would like to thank the *Fundação para a Ciência e a Tecnologia* for its financial support through grant SFRH/BD/31080/2006.

REFERENCES

Álvarez-Vázquez, L. J., Balsa-Canto, E., and Martínez, A. (2008). "Optimal design and operation of a wastewater purification system." *Mathematics and Computers in Simulation*, 79(3), 668-682.

Aras, E., Togan, V., and Berkun, M. (2007). “River water quality management model using genetic algorithm.” *Environmental Fluids Mechanics*, 7(5), 439-450.

Booth, H. (2006). “Demographic forecasting: 1980 to 2005 in review.” *International Journal of Forecasting*, 22(3), 547-581.

Brand, N., and Ostfeld, A. (2011). “Optimal design of regional wastewater pipelines and treatment plant systems.” *Water Environment Research*, 83(1), 53-64.

Cho, J. H., Sung, K. S., and Ha, S. R. (2004). “A river water quality management model for optimising regional wastewater treatment using a genetic algorithm.” *Journal of Environmental Management*, 73(3), 229-242.

Converse, A. O. (1972). “Optimum number and location of treatment plants.” *Journal of the Water Pollution Control Federation*, 44(8), 1629-1636.

Cunha, M. C., and Sousa, J. (2010). “Robust design of water distribution networks for a proactive risk management.” *Journal of Water Resources Planning and Management*, 136(2), 227-236.

Cunha, M. C., Pinheiro, L., Zeferino, J., Antunes, A., and Afonso, P. (2009). “Optimization model for integrated regional wastewater systems planning.” *Journal of Water Resources Planning and Management*, 135(1), 23-33.

Daskin, M. S., Hesse, S. M., and ReVelle, C. S. (1997). “α-reliable p-minimax regret: a new model for strategic facility location modeling.” *Location Science*, 5(4), 227-246.

De Beer, J. (2000). *Dealing with Uncertainty in Population Forecasting.* Department of Population Statistics Netherlands, Central Bureau of Statistics (CBS), Voorburg, Netherlands.

Deininger, R. A., and Su, S. Y. (1973). “Modelling regional waste water treatment systems.” *Water Research*, 7(4), 633-646.

Draper, N., and Smith, H. (1998). *Applied Regression Analysis*. John Wiley & Sons, New York, USA.

Eglese, R. W. (1990). “Simulated annealing: A tool for operational research.” *European Journal of Operational Research*, 46(3), 271-281.

Joeres, E. F., Dressler, J., Cho, C. C., and Falkner, C. H. (1974). “Planning methodology for the design of regional waste water treatment systems.” *Water Resources Research*, 10(4), 643-649.

Keilman, N. (2008). “European demographic forecasts have not become more accurate over the past 25 years.” *Population and Development Review*, 34(1), 137-153.

Keyfitz, N. (1981). “The limits of population forecasting.” *Population and Development Review*, 7(4), 579-593.

Kirkpatrick, S., Gellatt, C., and Vecchi, M. (1983). “Optimization by simulated annealing.” *Science*, 220(4598), 671-680.

Loomes, G., and Sugden, R. (1982). “Regret theory: An alternative theory of rational choice under uncertainty.” *Economic Journal*, 92(368), 805-824.

Maeda, S., Kawachi, T., Unami, K., Takeuchi, J., and Ichion, E. (2010) “Controlling wasteloads from point and nonpoint sources to river system by GIS-aided Epsilon Robust Optimization model.” *Journal of Hydro-environment Research*, 4(1), 27-36.

McNamara, J.R. (1976). “An optimization model for regional water quality management.” *Water Resources Research*, 12(2), 125-134.

Melo, J. J., and Câmara, A. S. (1994). “Models for the optimization of regional wastewater treatment systems.” *European Journal of Operational Research*, 73(1), 1-16.

Mulvey, J. M., Vanderbei, R. J., and Zenios, S. A. (1995). “Robust optimization of large-scale systems.” *Operations Research*, 43(2), 264-281.

Murdock S. H., Hamm R. R., Voss, P. R., Fannin, D., and Pecote, B., (1991). “Evaluating small-area population projections.” *Journal of the American Planning Association*, 57(4), 432-443.

Newell, C. (1988). *Methods and Models in Demography*. The Guilford Press, New York, USA.

Rayer, S., Smith, S. K., and Tayman, J. (2009). “Empirical prediction intervals for county population forecasts.” *Population Research and Policy Review*, 28(6), 773-793.

Rockafellar, R. T., and Wets, J. B. (1991). “Scenarios and policy aggregation in optimization under uncertainty.” *Mathematics of Operations Research*, 16(1), 119-147.

Rosenberg, D. E., and Lund, J. R. (2009). “Modeling integrated decisions for a municipal water system with recourse and uncertainties: Amman, Jordan.” *Water Resources Management*, 23(1), 85-115.

Rowland, D. T. (2003). *Demographic Methods and Concepts*. Oxford University Press, New York, USA.

Smeers, Y., and Tyteca, D. (1982). “Optimal location and design of wastewater treatment plants under river quality constraints.” In: S. Rinaldi, ed., *Environmental Systems Analysis and Management*. North- Holland, Amsterdam, Netherlands, 289-310.

Smith S. K. (1997). “Further thoughts on simplicity and complexity in population projection models.” *International Journal of Forecasting*, 13(4), 557-565.

Smith S. K., Tayman J., and Swanson D. A. (2001). *State and Local Population Projections: Methodology and Analysis*. Kluwer Academic/Plenum Publishers, New York, USA.

Smith, S. K., and Sincich, T. (1988). “Stability over time in the distribution of population forecast errors.” *Demography*, 25(3), 461-474.

Snyder, L.V. (2006). “Facility location under uncertainty: A review”. *IIE Transactions* 38(7), 537-554.

Sousa, J., Ribeiro, A., Cunha, M. C., and Antunes, A. (2002). "An optimization approach to wastewater systems planning at regional level." *Journal of Hydroinformatics*, 4(2), 115-123.

Stillwell, J., and Clarke, M., eds. (2011). *Population Dynamics and Projection Methods*. Springer, London, UK.

Stoto, M. A. (1983). "The accuracy of population projections." *Journal of the American Statistical Association*, 78(381), 13-20.

Suman, B., and Kumar P. (2006). "A survey of simulated annealing as a tool for single and multiobjective optimization." *Journal of the Operational Research Society*, 57(10), 1143–1160.

Tayman J., Schafer E., and Carter L. (1998). "The role of population size in the determination and prediction of population forecast errors: An evaluation using confidence intervals for sub-county areas." *Population Research and Policy Review*, 17(1), 1-20.

Vieira, J.M.P., and Lijklema, L. (1989) "Development and application of a model for regional water quality management." *Water Research*, 23(6), 767-777.

Wang, C. G., and Jamieson, D. G. (2002). "An objective approach to regional wastewater treatment planning." *Water Resources Research*, 38(3), 4/1-4/8.

WBCSD (2008). *2008 – UN International Year of Sanitation: It Is Time for Business to Act*. World Business Council for Sustainable Development, Geneva, Switzerland.

Whitlatch, E. (1997). "Siting regional environmental facilities." In: C. ReVelle, and A. E. McGarity, eds., *Design and Operation of Civil and Environmental Engineering Systems*. Wiley, New York, NY, USA, 615-656.

Wilson, T., and Bell, M. (2007) "Probabilistic regional population forecasts: The example of Queensland, Australia." *Geographical Analysis*, 39(1), 1-25.

Yandamuri S. R. M., Srinivasan K., and Bhallamudi S. M. (2006). "Multiobjective optimal waste load allocation models for rivers using Nondominated Sorting

Genetic Algorithm-II." *Journal of Water Resources Planning and Management*, 132(3), 133-143.

Yeh, S.-F., Chu, C.-W., Chang, Y.-J., and Lin, M.-D. (2011). "Applying tabu search and simulated annealing to the optimal design of sewer networks." *Engineering Optimization*, 43(2), 159-174.

Zechman, E.M., and Ranjithan, R.S. (2007). "Generating alternatives using evolutionary algorithms for water resources and environmental management problems." *Journal of Water Resource Planning and Management*, 133(2), 156-165.

Zeferino, J. A., Antunes, A. P., and Cunha, M. C. (2009). "An efficient simulated annealing algorithm for regional wastewater system planning." *Computer-Aided Civil and Infrastructure Engineering*, 24(5), 359-370.

Zeferino, J. A., Cunha, M. C., and Antunes, A. P. (2012). "Robust optimization approach to regional wastewater system planning." *Journal of Environmental Management*, 109,113-122.

Table 1 - Notation for the robust optimization models

Symbol	Descripton
Sets	
$\boldsymbol{N_S}$	set of wastewater sources
$\boldsymbol{N_I}$	set of possible intermediate nodes
$\boldsymbol{N_T}$	set of possible WWTP and related river reaches
$\boldsymbol{N}$	set of nodes (wastewater sources plus possible intermediate nodes plus possible WWTP)
$\boldsymbol{T}$	set of WWTP types
$\boldsymbol{S}$	set of scenarios
Decision variables	
Q_{ijs}	wastewater flow carried from node i to node j under scenario s
QT_{ks}	wastewater flow sent to a WWTP located at node k under scenario s
V_{kp}	binary variable that takes the value 1 if there is a WWTP of type p at node k, and is equal to 0 otherwise
X_{ij}	binary variable that takes the value 1 if there is a sewer to carry wastewater from node i to node j, and is equal to 0 otherwise
R_s	regret associated with scenario s
C_s	cost of the solution to be implemented under scenario s
E_{ijs}	difference of hydraulic heads between node i and node j under scenario s
Y_{ij}	binary variable that takes the value 1 if there is a pump station for taking wastewater from node i to node j, and is equal to 0 otherwise
$Z_{\alpha s}$ and $Z_{\beta s}$	binary variables that take the value 1 if scenario s is included in the reliability set, and are 0 otherwise
Parameters	
QR_{is}	wastewater flow generated at node i under scenario s
$Qmin_{ij}$ and $Qmax_{ij}$	minimum and maximum desirable flows allowed in the sewer linking node i to node j
$QTmax_{kp}$	maximum wastewater flow that may be treated at node k in a WWTP of type p
C_s^*	minimum cost of the solution for scenario s
L_{ij}	length of the sewer linking node i to node j
P_s	probability of scenario s
M_{Qmin}	very small constant
M_{Qmax} and M_{QT}	very large constants
α and β	reliability parameters
Q_{MAXij}	maximum tolerable flow allowed in the sewer linking node i to node j

Table 2 - Costs of the optimal wastewater system for the different scenarios

Scenario	Capital costs (M€)	Operating costs (M€/year)	Total costs (M€)
1	28.64	0.87	40.49
2	28.62	0.87	40.44
3	29.48	0.89	41.60
4	29.16	0.89	41.23
5	28.36	0.84	39.73
6	29.49	0.91	41.80
7	29.75	0.91	42.12
8	28.51	0.86	40.25
9	28.49	0.88	40.44
10	28.51	0.87	40.37
11	29.08	0.89	41.23
12	30.32	0.94	43.06
13	29.46	0.89	41.62
14	29.75	0.91	42.12
15	30.28	0.93	42.92
16	28.62	0.86	40.34
17	29.29	0.89	41.43
18	29.25	0.89	41.34
19	30.88	0.95	43.79
20	28.96	0.89	41.01

Table 3 - Costs of the robust optimal wastewater systems

Scenario	Expected-regret model		α-reliable model		β-reliable model	
	Capital costs	Total costs	Capital costs	Total costs	Capital costs	Total costs
	(M€)					
1		44.52		43.45		44.30
2		44.50		43.46		44.27
3		44.72		43.65		44.50
4		44.66		43.59		44.43
5		44.25		43.19		44.03
6		44.81		43.74		44.58
7		44.88		43.83		44.69
8		44.44		43.38		44.23
9		44.57		43.50		44.33
10	31.96	44.52	31.21	43.46	31.72	44.30
11		44.66		43.59		44.44
12		45.07		-		44.84
13		44.79		43.70		44.54
14		44.93		43.85		44.69
15		45.06		43.97		44.83
16		44.41		43.33		44.18
17		44.71		43.64		44.47
18		44.72		43.65		44.50
19		45.16		-		44.93
20		44.61		43.57		44.41

Figure 1
Click here to download Figure: Figure 1.pdf

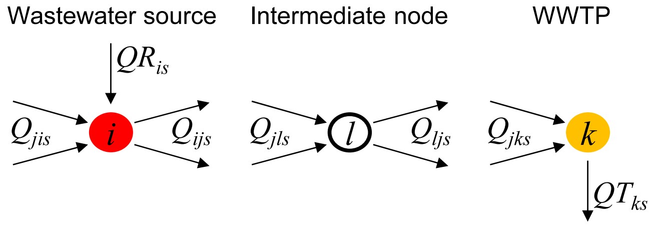

Figure 1 - Notation for continuity equations

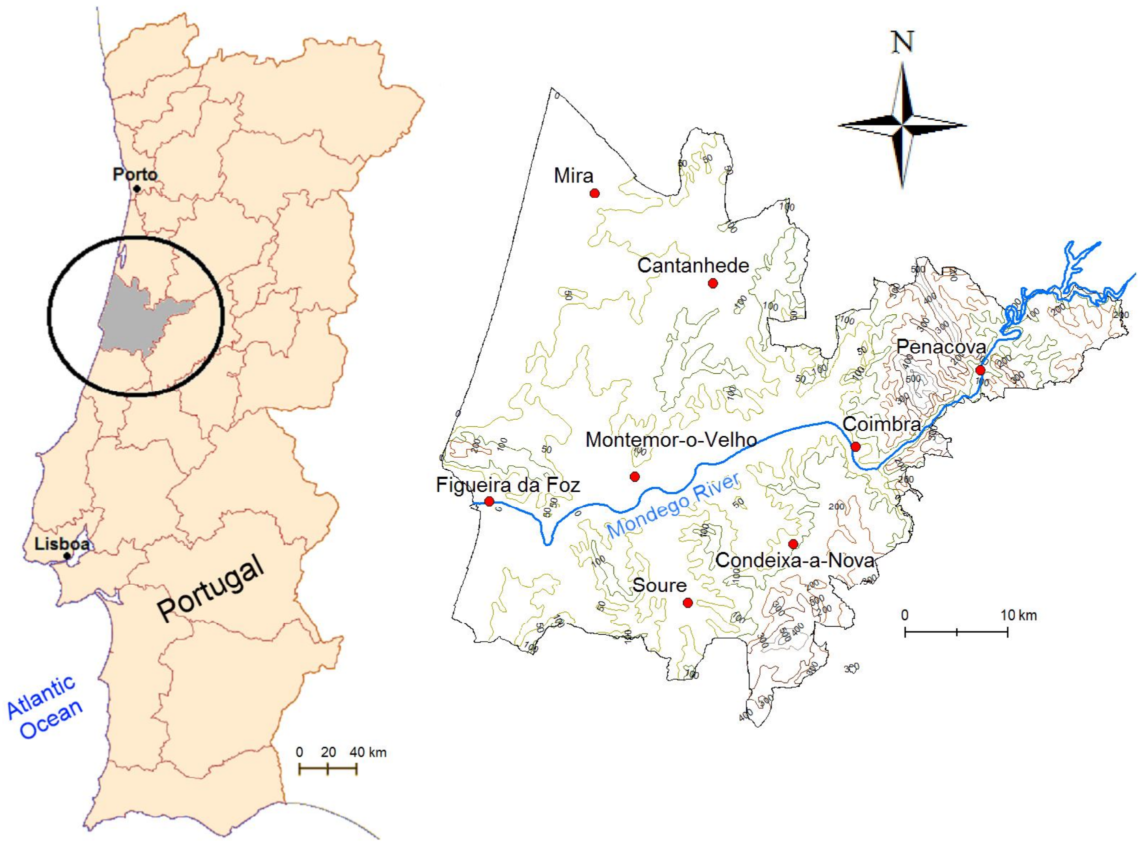

Figure 2 - Baixo Mondego region

Figure 3
Click here to download Figure: Figure 3.pdf

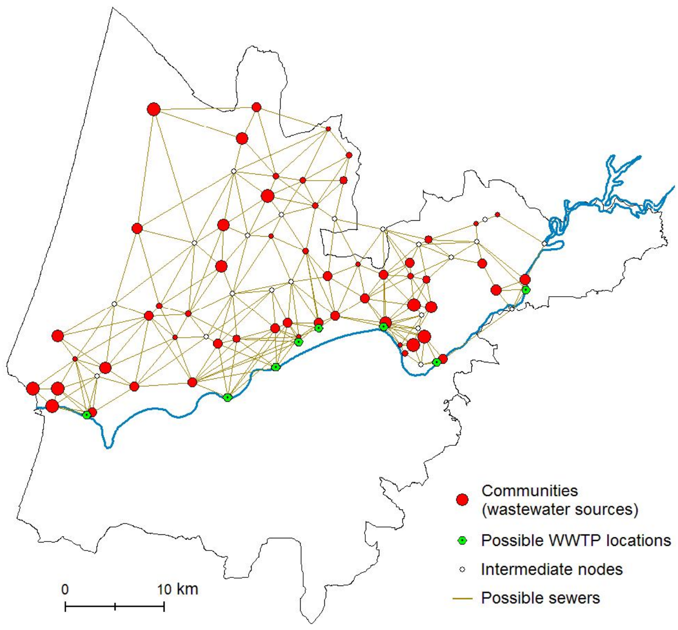

Figure 3 - North Baixo Mondego network data

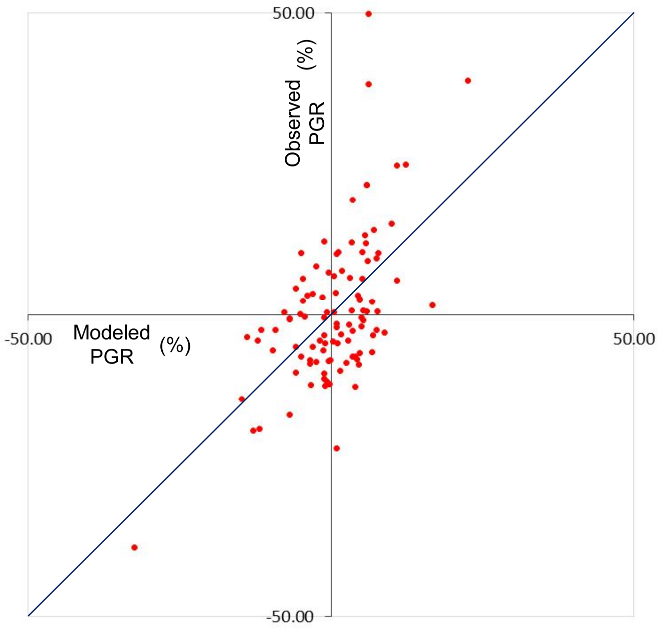

Figure 4 - Observed vs. modeled PGR

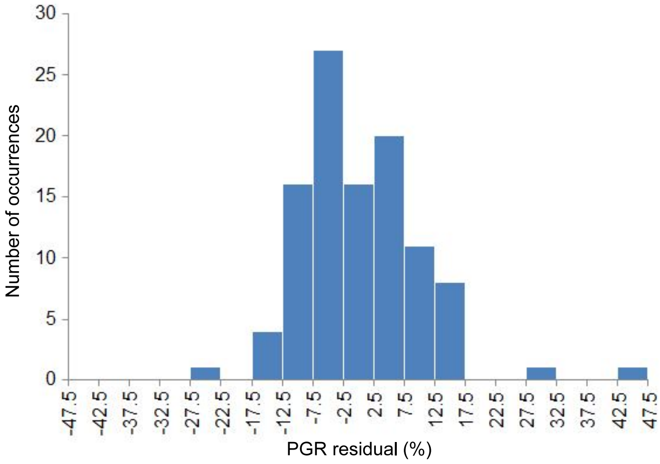

Figure 5 - Histogram of PGR residuals

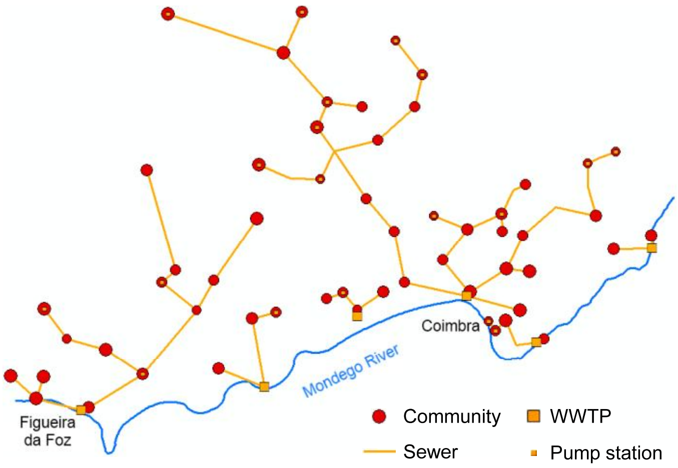

Figure 6 - Optimal wastewater system configuration for Scenario 19

Figure 7
Click here to download Figure: Figure 7.pdf

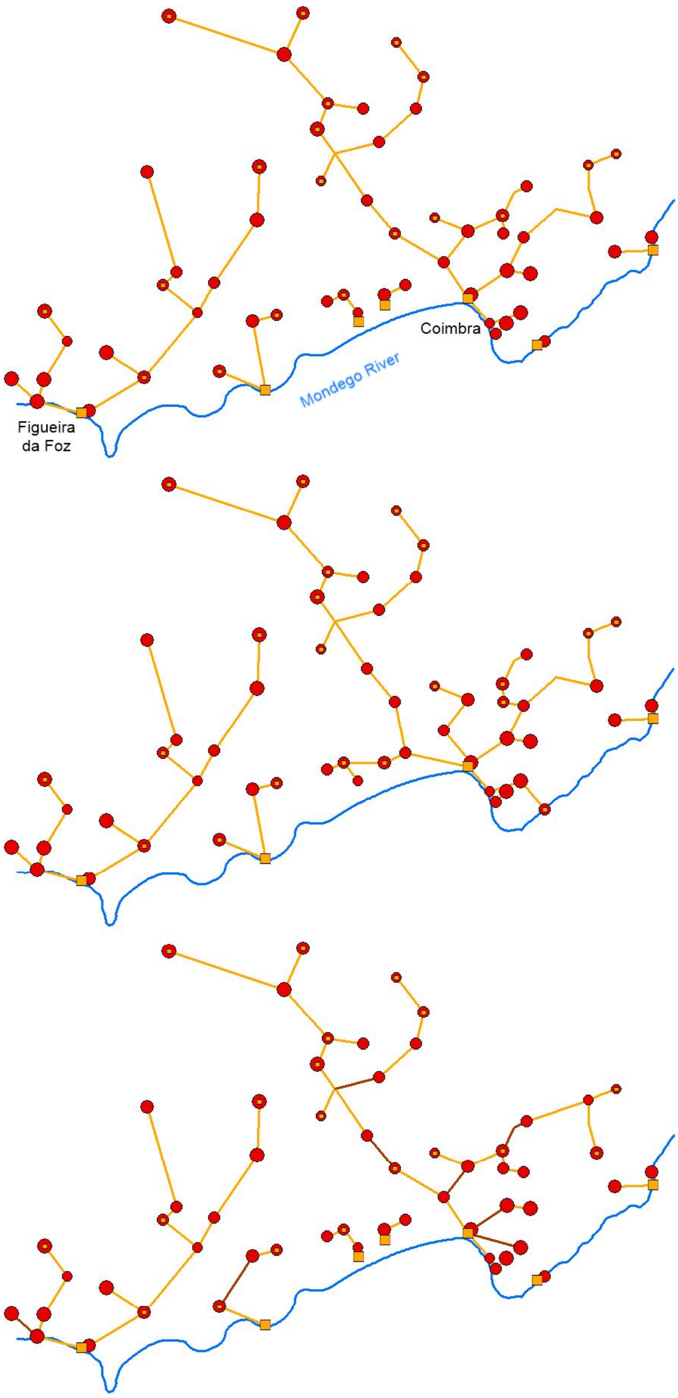

Figure 7 - Optimal wastewater system configurations: expected-regret model (top), α-reliable model (middle), and β-reliable model

Figure captions

Figure 1 - Notation for continuity equations

Figure 2 - Baixo Mondego region

Figure 3 - North Baixo Mondego network data

Figure 4 - Observed vs. modeled PGR

Figure 5 - Histogram of PGR residuals

Figure 6 - Optimal wastewater system configuration for Scenario 19

Figure 7 - Optimal wastewater system configurations: expected-regret model (top), α-reliable model (middle), and β-reliable model

Table headings

Table 1 - Notation for the robust optimization models

Table 2 - Costs of the optimal wastewater system for the different scenarios

Table 3 - Costs of the robust optimal wastewater systems