Ann Oper Res (2014) 219:141–167
DOI 10.1007/s10479-011-1036-7

An improved decomposition-based heuristic to design a water distribution network for an irrigation system

Graça Marques Gonçalves · Luís Gouveia · Margarida Vaz Pato

Published online: 20 December 2011

Abstract In this paper the authors address a pressurized water distribution network design problem for irrigation purposes. Two mixed binary nonlinear programming models are proposed for this NP-hard problem. Furthermore, a heuristic algorithm is presented for the problem, which considers a decomposition sequential scheme, based on linearization of the second model, coupled with constructive and local search procedures designed to achieve improved feasible solutions. To evaluate the robustness of the method we tested it on several instances generated from a real application. The best solutions obtained are finally compared with solutions provided by standard software. These computational experiments enable the authors to conclude that the decomposition sequential heuristic is a good approach to this difficult real problem.

Keywords Pressurized water distribution network design problem · Mixed binary nonlinear programming problem · Linearization · Heuristics

G.M. Gonçalves (✉)
Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Monte da Caparica, Portugal
e-mail: gmsg@fct.unl.pt

G.M. Gonçalves · L. Gouveia · M.V. Pato
Centro de Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal

L. Gouveia
Departamento de Estatística e Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal
e-mail: legouveia@fc.ul.pt

M.V. Pato
Departamento de Matemática, Instituto Superior de Economia e Gestão, Universidade Técnica de Lisboa, 1200-781 Lisboa, Portugal
e-mail: mpato@iseg.utl.pt

1 Introduction

This paper addresses a pressurized water distribution network for the purpose of irrigation. In such a network the water is conveyed from the water supply sources through a set of interconnected hydraulic components, such as pipes and pumps, and is delivered to the consumers at the right rates and quantities, beside the adequate pressure to operate the pressurized on-farm irrigation systems. As the network is designed for agricultural purposes, the reliability factor is not relevant, which explains why this is a non-looped type network.

The pressurized water distribution network design problem (WDNP) is both challenging and complex. It involves making numerous inter-related decisions, which are difficult to consider as a whole. These decisions particularly concern the network layout, the flow of water to be conveyed on each arc of the network, the diameter and thickness of the pipes, plus the location of the pumps on the network nodes and their dimensions. The choice of hydraulic equipment sizes depends, to a large extent, on the water flows through each arc of the network which, in turn, are determined by the network layout.

Conceptually, all three types of decisions should be considered simultaneously, which accounts for the difficulty in computationally solving the overall problem. For this reason, similar problems studied in the literature were broken down into three independent sub-problems or, alternatively, the sub-problems were grouped into pairs and solved sequentially at all times.

Most studies on water distribution networks address the problem of the municipal network design and seek to minimize the cost of the distribution network on the basis of a predetermined layout. Although they all envisage looped networks they share the hydraulic requirements with WDNP.

The paper by Bragalli et al. (2006) is devoted to optimization of a municipal water distribution network. The author assumes a pre-determined geographical layout and presents a mixed integer nonlinear programming formulation for the problem. The nonlinearities of the model are mainly due to the condition of energy balance. To reduce these nonlinearities, the authors propose a reformulation in which they define the cross sectional area of the pipelines as variables rather than the diameters. Computational tests on the reformulated model produce good results.

In the paper by Zhang and Zhu (1996), the authors studied the problem of dimensioning a gas distribution looped network. Here the network layout involved is also pre-determined. A mixed integer nonlinear programming model is defined, the flow in each arc defined as continuous variables, as well as the pressure at each node, and the diameter of the pipe defined as discrete variables. To solve the problem, the authors transform the initial model into a nonlinear model that considers pipes segmented into different diameters. The new model continues to be a difficult nonlinear problem, which is solved as a bi-level programming problem.

Sherali and Smith (1997) studied a municipal water distribution looped network, assuming a predetermined layout and pipes segmented into different diameters, as reported in the previous work. They also assumed that the network may already contain pipe sections of a given diameter and length and that one may have to replace them. The authors presented a branch-and-bound algorithm that applies the reformulation-linearization technique and solves a linear problem to calculate lower bounds for the optimal value of the problem. Upper bounds were obtained from a heuristic procedure. Some problems in the literature for which, to date, only local optimal solutions had been known, were optimally solved using this global optimization algorithm.

Most studies published before those described above also address the problem of designing municipal water distribution looped networks on the assumption of a pre-determined layout, while only solving the problem of sizing the hydraulic components. Such is the case of the works of Karmeli et al. (1968), Knowles et al. (1976), Hansen et al. (1991), Alperovits and Shamir (1977) and Kessler and Shamir (1991). As for the study by Ionescu et al. (1981) they present a branched network, though they also assume a pre-defined layout. Most of these authors adopted a solving method which breaks the original problem down into sub-problems. These are solved sequentially, by using a solution of one sub-problem to solve the following one.

In the study by Loganathan et al. (1990), the layout is not pre-determined. The authors assume a certain degree of redundancy in the network, which is why the final network may include cycles. They devise a two-phase method to solve the problem. In the first phase, the tree of shortest paths is determined, then, by solving linear sub-problems this same tree is modified to reduce network costs. In the second phase, some arcs are added to the tree that was obtained in the previous phase, to ensure there are at least two paths from the source point to each consumer.

It should be noted that a particular version of the WDNP had previously been studied by Gonçalves and Pato (2000), where a decomposition-based heuristic was proposed. In this heuristic, three sub-problems are sequentially solved: the tree network layout, which is the first sub-problem, followed by the calculation of the flow in each arc of the previously obtained tree. Finally, the pipeline sizes are computed, the locations and sizes of the pumps are decided for the tree and flows that were determined from the first and second sub-problems. A solution for the first sub-problem is obtained through a constructive heuristic. This is based on the shortest path insertion heuristic for the minimal length Steiner tree, due to Takahashi and Matsuyama (1980). Then, the arc flows on the second sub-problem are calculated by using a probabilistic hydraulic formula which estimates the flow for each arc. The pipes' diameters and thickness, besides the location of the pumps and respective pumping heights are defined on the basis of the solution of a mixed integer linear problem.

Most of these studies deal with looped networks that are inapplicable to water distribution networks for irrigation purposes in view of the unnecessary additional costs involved in this domain. The few studies involving branched networks do not optimize the layout and sizes of the hydraulic components simultaneously. They consider a pre-defined layout and pipes which are segmented into different diameters. Most studies consider the installation of a hydraulic pump only at the point of water withdrawal.

The problem studied in this paper differs from the one mentioned in Gonçalves and Pato (2000), in terms of the problem, its formulation and also the method used to solve the problem. As for the problem itself, the previous study does not consider the network operating costs. The formulation presented here is also new. As for the method, the previous paper (Gonçalves and Pato 2000) studies a simple Steiner constructive heuristic with one version only and no improvement phase. For this reason the network design heuristic is less elaborate than the one presented in this paper; the flow formula assumes that not all the consumers irrigate at the same time. Paper (Gonçalves and Pato 2000) only studies one real instance whereas here we present a computational study with a wide set of simulated instances. Finally, the solutions obtained from the heuristic of that paper are not accessed with lower bounds for the optimum whereas the current paper involves a comparison with lower bounds and known optimum values. Hence, the development and evaluation of improved heuristics for WDNP continue to be a topical field of research. Moreover, an efficient heuristic for the WDNP can be a useful tool for adaptation to other difficult nonlinear discrete optimization problems.

The paper is organized as follows. In Sect. 2 we introduce the pressurized water distribution network design problem under investigation. In Sect. 3 we present two mathematical formulations for the problem and in Sect. 4 we detail the decomposition-based heuristic in its several versions. The computational experiments performed in order to evaluate the proposed heuristic versions are reported in Sect. 5. The paper ends with some conclusions drawn from the work undertaken and some directions for future research.

2 Problem description

Given a set of consumer points, a set of source points and the set of all the potential pipe routes between these points, the WDNP calls for the selection of the choice of pipe routes that will connect each consumer to a source point. Moreover, it also demands specification of the dimension of the network's hydraulic components. Besides determination of the power of the hydraulic pumps needed in the network, as well as the decision as to their location in the network, the WDNP requires specification of the diameter and thickness of each pipe on each pipe route.

The aim of the network is to connect all consumers to supply sources by employing the most economical network, in terms of investment and operating costs, while satisfying a set of hydraulic constraints. Among others, they should include consumer demand for a minimum flow rate and a minimum pressure.

Before formally stating the problem in Sect. 3, we discuss its key features. The geographical locations of the source and consumer points are assumed to be known, and these points will be denoted in this paper as *water withdrawals* and *hydrants*, respectively. The water and pressure requirements at each hydrant are also known, as well as all the potential pipe routes between the withdrawals and the hydrants (see Fig. 1 which represents a real network with one water withdrawal only). The connection between these points may be direct, from the withdrawal point to the hydrant, or indirect, via other hydrants or *junction* points. These are additional optional points. Though they are neither demand nor source points, here they are considered for a purpose: to obtain a shorter network or install a hydraulic pump there, should electricity be available. The geographical location of the junction points is also known in advance. Other data are the ground elevation of any of the three types of points defined above, as well as the distance between them, besides the land area to be irrigated by each hydrant.

We should note that choices to be made as regards pipe sizing specifically refer to diameter and thickness here represented by the pressure class of the pipe. Where pump sizing is concerned, the choice is related to the pumping height of the pump. Moreover, it should also be remembered that this network is designed for agricultural purposes. For this reason, rather than resorting to a more costly looped network, the authors envisage a non-looped network. As the reliability factor is not that relevant, the network is built up of independent acyclic sub-networks, each one rooted on a withdrawal. In short, to solve this problem, the following decisions must be made:

(i) selection of the pipe routes, definition of a path from a water withdrawal to each hydrant;
(ii) sizing of each pipe in each route, i.e., computation of the diameter and thickness of each pipe;
(iii) selection of the nodes where one should install a hydraulic pump and computation of its appropriate height;
(iv) computation of the water flow to be conveyed on each pipe route.

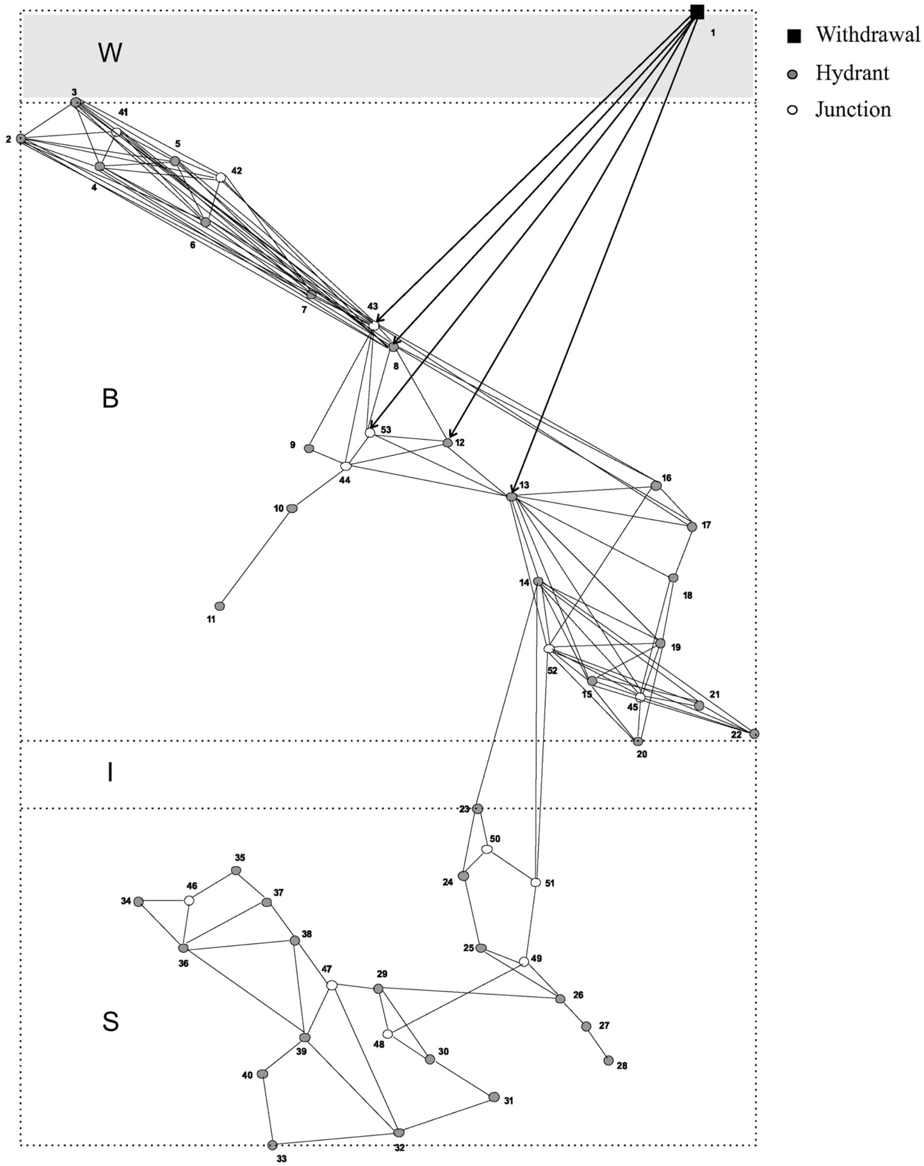
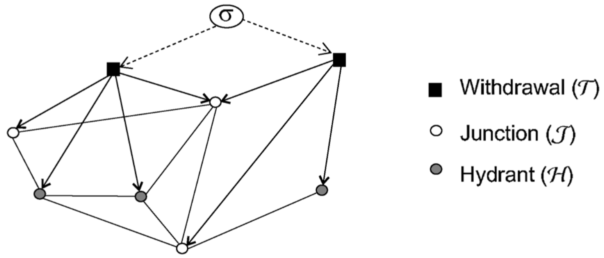


Fig. 1 Nodes and potential pipe routes of a real network

Finally, we observe that the WDNP is a Steiner tree problem with additional complicating constraints. These are related to the dimensioning of the hydraulic components of the system (e.g., pumps' location and pumping heights, pressure class and diameters of pipes). Thus, the problem is NP-hard since it is a generalization of a well known NP-hard problem (see Karp 1972).

Fig. 2 Graph $\mathcal{G}_\sigma = (\mathcal{N}_\sigma, \mathcal{A}_\sigma)$



3 Mathematical formulation

In Sect. 3.1 firstly we introduce the notation that will be used throughout the paper. Since the heuristic algorithm presented in Sect. 4 of this paper is based on a mixed binary nonlinear programming (MBNLP) formulation for the WDNP, we also present this new model in Sect. 3.2. Finally, in order to obtain a model with fewer nonlinear terms, in Sect. 3.3, we propose a reformulation of the previous model.

3.1 Notation

As is normal practice when formulating such a problem, we define a connected directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ to represent the pipe routes, where $\mathcal{N}$ is the set of nodes and $\mathcal{A}$ is the set of arcs. The arcs correspond to pipes and the node set $\mathcal{N}$ is partitioned into water withdrawal nodes ($\mathcal{T}$), hydrant nodes ($\mathcal{H}$) and junction nodes ($\mathcal{J}$).

Assume that there are no arcs into $\mathcal{T}$ nodes in the graph $\mathcal{G}$ and define another directed graph $\mathcal{G}_\sigma$ by adding a dummy node σ to $\mathcal{G}$, as well as the arcs (σ, i), $i \in \mathcal{T}$, with an associated length $L_{\sigma i} = 0$, while, for the other arcs $(i, j) \in \mathcal{A}$, the associated length L_{ij} is equal to the Euclidean distance between the end points of the arc. The resulting graph is $\mathcal{G}_\sigma = (\mathcal{N}_\sigma, \mathcal{A}_\sigma)$, where $\mathcal{A}_\sigma = \mathcal{A} \cup \{(\sigma, i) \mid i \in \mathcal{T}\}$ and $\mathcal{N}_\sigma = \mathcal{N} \cup \{\sigma\}$ (as an illustration see Fig. 2, which represents a small instance with two withdrawals). For each $j \in \mathcal{N}_\sigma$, δ_j^+ denotes the set of nodes $i \in \mathcal{N}_\sigma$, such that $(j, i) \in \mathcal{A}_\sigma$ and δ_j^- is the set of nodes $i \in \mathcal{N}_\sigma$ such that $(i, j) \in \mathcal{A}_\sigma$.

The topological support of the model presented in Sect. 3.2 is the graph $\mathcal{G}_\sigma$ and the model's sets and indexes are presented in Table 1.

The elements of the set $\mathcal{D}_{ij}$ of allowable diameters for the arc (i, j), illustrated in Table 1 are the diameters of set $\mathcal{D}$, for which the speed of the water flow to be conveyed by arc (i, j) satisfies the minimum and maximum speed limits established for each diameter.

Note that if the flow q (in m^3/h) to be conveyed through the arc (i, j) is known in advance one can compute which commercially available diameters d (in mm) satisfy the following water speed bounding inequalities:

$$V_d^L \leq \frac{\frac{q}{3600}}{\pi(\frac{(\frac{d}{1000})}{2})^2} = \frac{10^4}{9\pi}\frac{q}{d^2}\left[\frac{\text{m}^3/\text{s}}{\text{m}^2}\right] \leq V_d^U, \tag{1}$$

where V_d^L and V_d^U (in m/s) stand for the minimum and maximum permissible speed for diameter d, respectively.

The flow to be conveyed through the arc (i, j) is not known beforehand. However, if $j \in \mathcal{H}$ we know that the minimum flow is Q_j, that is the water flow required by the hydrant.

Table 1 Sets and indexes of the model

Symbol	Description
$\mathcal{A}, \mathcal{N}$	set of original network arcs and nodes, respectively
$\mathcal{H}$	hydrants (demand nodes)
$\mathcal{J}$	junctions (optional nodes)
$\mathcal{T}$	water withdrawals (source nodes)
σ	dummy node
i, j	general network nodes
$\mathcal{A}_\sigma$	$\mathcal{A} \cup \{(\sigma, i) \mid i \in \mathcal{T}\}$
$\mathcal{N}_\sigma$	$\mathcal{N} \cup \{\sigma\}$
(i, j)	general network arc, directed
δ_j^-	$\{i \in \mathcal{N}_\sigma \mid (i, j) \in \mathcal{A}_\sigma\}$
δ_j^+	$\{i \in \mathcal{N}_\sigma \mid (j, i) \in \mathcal{A}_\sigma\}$
$\mathcal{D}$	discrete set of pipe diameters commercially available
$\mathcal{P}$	discrete set of pipe pressure classes commercially available
$\mathcal{D}_{ij}$	set of allowable pipe diameters for the arc (i, j)
d	general pipe diameter (in millimeters—mm)
p	general pipe pressure class (in meters—m)
$\mathcal{DP}_{ij}$	set of allowable pairs $(d, p) \in \mathcal{D}_{ij} \times \mathcal{P}$ for the arc (i, j)

If $j \in \mathcal{J}$ and (i, j) is used, then the minimum flow through the arc is $\min\{Q_j \mid j \in \mathcal{H}\}$. Furthermore, we know that the maximum flow required for an arc is $\sum_{j \in \mathcal{H}} Q_j$. The minimum and maximum allowable diameters for each arc (i, j) were calculated on the basis of these minimum and maximum flow values respectively, thus defining the set $\mathcal{D}_{ij}$.

The set $\mathcal{DP}_{ij}$ is thus the set of pairs of commercially available pipes with a diameter $d \in \mathcal{D}_{ij}$ and applicable for all pressure class $p \in \mathcal{P}$.

Other hydraulic input data are presented in Table 2 below.

It should be added that the coefficients H_j^U, M_{ijd}^1, M_{ijd}^2, M_{ij}^3, PD_j^U, PS_j^U, S_{ij}^L, S_{ij}^U, Y_{ij}^U and W_j^U in Table 2 were specified by using some logical preprocessing, which will be detailed later.

The decision variables of the model, as well as the variables of the reformulated model presented in Sect. 3.3, are introduced in Table 3 below.

3.2 A new mathematical model

In this study the WDNP is formulated as a mixed binary nonlinear program which will be presented in this subsection. To provide a clearer explanation of the model some additional notation is introduced.

Objective function

$$\sum_{(i,j)\in\mathcal{A}} L_{ij} \sum_{(d,p)\in\mathcal{DP}_{ij}} CC_{dp}\, \alpha_{ijd}\, \varepsilon_{ijp} + \sum_{(i,j)\in\mathcal{A}_\sigma} CP\, y_{ij}\, (pd_j - ps_j)$$
$$+ \sum_{(i,j)\in\mathcal{A}_\sigma} CE\, s_{ij}\, (pd_j - ps_j). \tag{2}$$

Table 2 Hydraulic data

Symbol	Description
CC_{dp}	constant for the cost of a pipe with diameter d and pressure class p
CE	constant for the energy cost
CP	constant for the cost of the pumps and the power cost
E_{ijd}	constant for the formula for pressure loss in (i, j) relative to diameter d
H_j^L	practical lower bound (in m) imposed on the head pressure at node j
H_j^U	specified upper bound (in m) on the head pressure at node j
L_{ij}	length (in m) of the arc (i, j)
M_{ijd}^1	constant specified by using some logical preprocessing
M_{ijd}^2	constant specified by using some logical preprocessing
M_{ij}^3	constant specified by using some logical preprocessing
P_i^{min}	minimum pressure (in m) required at node $i \in \mathcal{H}$
PD_j^U	specified upper bound (in m) on the leaving pressure at node j
PS_j^U	specified upper bound (in m) on the entering pressure at node j
Q_j	water flow rate (in m^3/h) needed at node $j \in \mathcal{H}$. $Q_j = 0, \forall j \in \mathcal{T} \cup \mathcal{J}$
S_j	land area (in ha) to irrigate by the node $j \in \mathcal{H}$. $S_j = 0, \forall j \in \mathcal{T} \cup \mathcal{J}$
S_{ij}^L	specified lower bound (in ha) on the land area to irrigate by (i, j)
S_{ij}^U	specified upper bound (in ha) on the land area to irrigate by (i, j)
V_d^L	minimum permissible water speed (in m^3/s) for diameter d
V_d^U	maximum permissible water speed (in m^3/s) for diameter d
Z_j	ground elevation (in m) of node j
Z_{ij}	$Z_i - Z_j$
Y_d^L	$\frac{9\pi}{10^4} V_d^L d^2$
Y_d^U	$\frac{9\pi}{10^4} V_d^U d^2$
Y_{ij}^U	specified upper bound (in m^3/h) on the water flow in arc (i, j)
W_j^U	specified upper bound on node j

Table 3 Variables of the model and variables of the reformulated model

Symbol	Description
x_{ij}	indicates whether the arc (i, j) is in the solution $(= 1)$ or not $(= 0)$
α_{ijd}	indicates if a pipe diameter d is assigned to the arc (i, j) $(= 1)$ or not $(= 0)$
ε_{ijp}	indicates if a pipe pressure class p is assigned to the arc (i, j) $(= 1)$ or not $(= 0)$
u_j	indicates if a hydraulic pump should be installed at node j $(= 1)$ or not $(= 0)$
y_{ij}	water flow (in m^3/h) in arc (i, j)
s_{ij}	land area (in ha) to irrigate downstream the arc (i, j)
ps_j, pd_j	water pressure (in m) entering and leaving node j, respectively
θ_{ijdp}	indicates if a pipe of diameter d and pressure class p is assigned to the arc (i, j) $(= 1)$ or not $(= 0)$
h_j	pumping height (in m) of a pump installed at node j
w_j	auxiliary variable for each node j

This function to be minimized includes the investment costs of pipes, represented by the first summation in (2), the investment costs of pumps represented in the second summation, which also represents the power costs and, finally, the function comprises the energy costs represented by the third summation.

The coefficient CC_{dp} on the pipe capital cost formula is defined as $\frac{C_{dp}}{Am}$, where C_{dp} (in monetary units/m) is the unit investment cost per length of a pipe with a diameter d and pressure class p, and Am is the number of years required to amortize the investment.

The cost coefficient CP is obtained from $CP = \frac{C_1}{Am}\frac{9.8}{3600}\frac{1}{\eta} + \frac{C_2}{Am}\frac{9.8}{3600}\frac{1}{\eta}$, where C_1 (in monetary units/kW) is the investment cost of a pump per unit of power and C_2 (in monetary units/kW) is the cost of the power rate per unit of power. The investment cost of a pump, besides the power rate cost, is proportional to the pump power (in kW), defined for the specified units of flow and pressure as $\frac{9.8}{3600}\frac{y_{ij}(pd_j - ps_j)}{\eta}$, where η is the pump's efficiency and $pd_j - ps_j$ represents the pump height.

Finally, the energy cost coefficient CE is defined as $\frac{C_3}{Am}\,\nu\frac{9.8}{3600}\frac{1}{\eta}$, where C_3 (in monetary units/kWh) is the energy cost per kWh consumed, and ν (in m^3/ha/year) is an estimate of the volume of water required per hectare per year.

This model has different sets of constraints, including assignment constraints, satisfaction of water and pressure demand at the hydrants, the flow conservation conditions for the water flow and the area of land to irrigate, the water speed bounds on pipes, the energy balance equations, the maximum permissible pressure on pipes, bounds for the height of the pumps and some forcing constraints. These constraints are given below.

Assignment constraints for each node

$$\sum_{i\in\delta_j^-} x_{ij} = 1 \quad \forall j \in \mathcal{H}, \tag{3}$$

$$\sum_{i\in\delta_j^-} x_{ij} \le 1 \quad \forall j \in \mathcal{T} \cup \mathcal{J}. \tag{4}$$

The first set of these constraints guarantees that each hydrant is visited by exactly one path, while the second set ensures that each junction or withdrawal is visited at the most by one path.

Selection of a pipe diameter and a pressure class at each arc

$$\sum_{d\in\mathcal{D}_{ij}} \alpha_{ijd} = x_{ij} \quad \forall (i,j) \in \mathcal{A}_\sigma, \tag{5}$$

$$\sum_{p\in\mathcal{P}} \varepsilon_{ijp} = x_{ij} \quad \forall (i,j) \in \mathcal{A}. \tag{6}$$

These equations state that a single diameter in the case of (5), and one class of pressure, where (6), are selected for an arc (i, j) such that $x_{ij} = 1$ and none is selected if $x_{ij} = 0$.

Flow conservation at each node

$$\sum_{i\in\delta_j^-} y_{ij} - \sum_{i\in\delta_j^+} y_{ji} = Q_j \quad \forall j \in \mathcal{N}, \tag{7}$$

$$\sum_{i\in\delta_j^-} s_{ij} - \sum_{i\in\delta_j^+} s_{ji} = S_j \quad \forall j \in \mathcal{N}. \tag{8}$$

The constraints (7) are the usual flow conservation conditions for the water flow variables. They also guarantee that the demand for water at the hydrants is satisfied, while the balance equations (8) define the area of land for irrigation downstream of the arc (i, j), which is required to define the energy costs.

Water speed bounds at each arc

$$\sum_{d\in\mathcal{D}_{ij}} Y_d^L \alpha_{ijd} \le y_{ij} \le \sum_{d\in\mathcal{D}_{ij}} Y_d^U \alpha_{ijd} \quad \forall (i,j) \in \mathcal{A}_\sigma. \tag{9}$$

These inequalities ensure that the water speed in each pipe does not exceed the minimum and maximum values corresponding to the selected diameter d, for which $\alpha_{ijd} = 1$. Note that, for the selected diameter, the above constraints can be expressed as $Y_d^L \le y_{ij} \le Y_d^U$, or as $V_d^L \le \frac{10^4}{9\pi}\frac{y_{ij}}{d^2} \le V_d^U$ if the value of the constants Y_d^L and Y_d^U, according to their definition in Table 2, is replaced. The latter inequalities define the bounds of the water flow speed imposed on y_{ij}, in accordance with conditions (1).

Energy balance at each arc

$$pd_i - ps_j + Z_{ij} - E_{ijd}\, y_{ij}^{1.79} \le M_{ijd}^1 (1 - \alpha_{ijd}) \quad \forall (i,j) \in \mathcal{A},\ d \in \mathcal{D}_{ij}, \tag{10a}$$

$$-pd_i + ps_j - Z_{ij} + E_{ijd}\, y_{ij}^{1.79} \le M_{ijd}^2 (1 - \alpha_{ijd}) \quad \forall (i,j) \in \mathcal{A},\ d \in \mathcal{D}_{ij}. \tag{10b}$$

The energy balance equation states that the total head loss from node i to node j must equal the friction losses plus the minor head losses. Frictional head losses were calculated through the *Scimeni* formula. Thus, the unit head loss for arc (i, j) with pipe diameter equal to d is $K\frac{y_{ij}^{1.79}}{d^{4.79}}$, where K is a constant dependent on the roughness of the pipes and units used. As for the minor head losses, on the whole due to bends, we consider them to represent 10% of frictional head losses.

The constant E_{ijd} in these constraints is defined as $1.1\,K\,L_{ij}\,d^{-4.79}$, for each arc (i, j) with length L_{ij} and for each diameter d, and $E_{ijd}\, y_{ij}^{1.79}$ represents the total head loss from node i to node j in a pipe of diameter d.

Note that for the selected diameter d, i.e., for an arc such that $\alpha_{ijd} = 1$, the above conditions can be written as $pd_i - ps_j + Z_{ij} = E_{ijd}\, y_{ij}^{1.79}$, which is the energy balance equation for the arc (i, j). Also remember that the constants M_{ijd}^1 and M_{ijd}^2 in (10a)–(10b) are defined such that these conditions are redundant if $\alpha_{ijd} = 0$. We have considered $M_{ijd}^1 = PD_i^U + Z_{ij}$ and $M_{ijd}^2 = PS_j^U + E_{ijd}(Y_{ij}^U)^{1.79} - Z_{ij}$, where $Y_{ij}^U = \sum_{t\in\mathcal{H}} Q_t$.

Maximum permissible pressure value at each arc

$$1.1\,pd_i - 0.1\,ps_j + 1.1\,Z_{ij} \le \sum_{p\in\mathcal{P}} \frac{p}{2}\,\varepsilon_{ijp} + M_{ij}^3 (1 - x_{ij}) \quad \forall (i,j) \in \mathcal{A}. \tag{11}$$

An upper bound on the pressure in every arc is imposed by these inequalities. In other words, the pressure at the origin of an arc, plus an overpressure (as a rule set at 10% of the total head losses), plus the end points' ground level difference, should not exceed a maximum permissible pressure. In this paper we consider this maximum value as equal to 50% of the pipe pressure class.

Note that if $x_{ij} = 1$ then, for d and p such that $\alpha_{ijd} = 1$ and $\varepsilon_{ijp} = 1$, respectively, we obtain $ps_j = -E_{ijd}\, y_{ij}^{1.79} + pd_i + Z_{ij}$ from (10a)–(10b) and, consequently, replacement of ps_j in (11) results in $pd_i + 0.1\, E_{ijd}\, y_{ij}^{1.79} + Z_{ij} \leq p/2$, which is the pressure bounding condition mentioned above. As the constant M_{ij}^3 must take a value such that (11) is redundant if $x_{ij} = 0$, we have considered $M_{ij}^3 = 1.1\, PD_i^U + 1.1\, Z_{ij}$, where $PD_i^U = \max_{t \in \mathcal{N}} Z_t - Z_i + \frac{1}{2} \max_{p \in \mathcal{P}}\ p$ was defined from the pressure bounding constraints. From these values and the energy balance equations above we have defined $PS_i^U = \max_{t \in \mathcal{N}} PD_t^U + \max_{t \in \mathcal{N}} Z_t - Z_i$.

Minimum pressure value required at each hydrant

$$pd_j \geq P_j^{min} \quad \forall j \in \mathcal{H}. \tag{12}$$

The minimum pressure value required at each hydrant is guaranteed by these constraints.

Minimum and maximum height for a hydraulic pump at each node

$$ps_j = 0 \quad \forall j \in \mathcal{T}, \tag{13}$$

$$H_j^L u_j \leq pd_j - ps_j \leq H_j^U u_j \quad \forall j \in \mathcal{N}. \tag{14}$$

The first set of constraints defines the entering pressure value equal to zero in each source node, while the second imposes bounds for the pumping height of the pumps actually in place. The lower limit imposed on the lifting height of the pumps is due to the high cost of pumping water, while the upper limit is not the result of a technical requirement. Thus we have considered $H_j^U = PD_j^U$.

Forcing constraints

$$ps_j \leq PS_j^U \sum_{i \in \delta_j^-} x_{ij} \quad \forall j \in \mathcal{J}, \tag{15}$$

$$u_j \leq \sum_{i \in \delta_j^-} x_{ij} \quad \forall j \in \mathcal{T} \cup \mathcal{J}, \tag{16}$$

$$S_{ij}^L x_{ij} \leq s_{ij} \leq S_{ij}^U x_{ij} \quad \forall (i,j) \in \mathcal{A}_\sigma. \tag{17}$$

The sets of constraints defined above establish connections between different variables. The bounds on s_{ij} (constraints (17)) were defined as $S_{ij}^L = S_j$ and $S_{ij}^U = \sum_{t \in \mathcal{H}} S_t$.

Binary and non-negativity conditions

$$x_{ij} \in \{0, 1\} \quad \forall (i,j) \in \mathcal{A}_\sigma, \tag{18}$$

$$\alpha_{ijd} \in \{0, 1\} \quad \forall (i,j) \in \mathcal{A}_\sigma,\ d \in \mathcal{D}_{ij}, \tag{19}$$

$$\varepsilon_{ijp} \in \{0, 1\} \quad \forall (i,j) \in \mathcal{A},\ p \in \mathcal{P}, \tag{20}$$

$$u_j \in \{0, 1\} \quad \forall j \in \mathcal{N}, \tag{21}$$

$$y_{ij} \in [0, Y_{ij}^U], \qquad s_{ij} \in [0, S_{ij}^U] \quad \forall (i,j) \in \mathcal{A}_\sigma, \tag{22}$$

$$pd_j \in [0, PD_j^U], \qquad ps_j \in [0, PS_j^U] \quad \forall j \in \mathcal{N}. \tag{23}$$

The above conditions define binary and non-negative variables of the model.

Note that definition of the two sets $\mathcal{D}_{ij}$ and $\mathcal{DP}_{ij}$ in each arc (i, j), based on limits to the speed of the water, led to a reduction in the total number of binary variables to consider in the model. According to Labye et al. (1988) selection of the diameters that lead to permissible water speed values is a common practice in this kind of problem.

The model defined above for the WDNP is the following mixed binary nonlinear programming problem:

$$\begin{aligned}&\text{Minimize } (2)\\&\text{subject to} \quad (3)\text{–}(23).\end{aligned}$$

In fact, this problem includes in the objective function the bilinear terms $\alpha_{ijd}\,\varepsilon_{ijp}$ depending on the binary variables, as well as the bilinear terms $(y_{ij}+s_{ij})\,(pd_j-ps_j)$ depending on the non-negative variables. Also constraints (10a) and (10b) include nonlinear terms on non-negative variables. Proof that this model is valid for the WDNP can be found in Gonçalves (2008).

With a view to reducing the number of bilinear terms of the objective function in the formulation above, we propose a reformulation of the model in the next subsection.

3.3 Reformulation

Define the additional binary variables $\theta_{ijdp}\in\{0,1\}$, for all $(i,j)\in\mathcal{A}$, $(d,p)\in\mathcal{DP}_{ij}$, such that:

$$\sum_{\substack{p\in\mathcal{P}:\\(d,p)\in\mathcal{DP}_{ij}}}\theta_{ijdp}=\alpha_{ijd}\quad\forall(i,j)\in\mathcal{A},\ d\in\mathcal{D}_{ij},\tag{24a}$$

$$\sum_{\substack{d\in\mathcal{D}_{ij}:\\(d,p)\in\mathcal{DP}_{ij}}}\theta_{ijdp}=\varepsilon_{ijp}\quad\forall(i,j)\in\mathcal{A},\ p\in\mathcal{P},\tag{24b}$$

and also consider the additional non-negative variables $h_j, w_j \geq 0$, $\forall j\in\mathcal{N}$, such that:

$$h_j=pd_j-ps_j\quad\forall j\in\mathcal{N},\tag{25}$$

$$w_j=\sum_{i\in\delta_j^-}(CP\,y_{ij}+CE\,s_{ij})\quad\forall j\in\mathcal{N}.\tag{26}$$

By replacing $\alpha_{ijd}\,\varepsilon_{ijp}=\theta_{ijdp}$ in the objective function (2) and using (25)–(26), we obtain the following reformulated model:

$$\text{(P1)}\quad\text{Minimize}\ \sum_{(i,j)\in\mathcal{A}}L_{ij}\sum_{(d,p)\in\mathcal{DP}_{ij}}CC_{dp}\,\theta_{ijdp}+\sum_{j\in\mathcal{N}}w_j\,h_j\tag{27}$$

$$\text{subject to}\quad(3)\text{–}(26)$$

$$\theta_{ijdp}\in\{0,1\}\quad\forall(i,j)\in\mathcal{A},(d,p)\in\mathcal{DP}_{ij}\tag{28}$$

$$h_j\in[0,H_j^U],\qquad w_j\in[0,W_j^U]\quad\forall j\in\mathcal{N}.\tag{29}$$

The constant W_j^U is an upper bound for the w_j variable, for each $j \in \mathcal{N}$, which was specified as $W_j^U = \sum_{i\in\delta_j^-}(CPY_{ij}^U + CES_{ij}^U)$.

The objective function of this new model does not include bilinear terms that depend on the binary variables. Moreover, it has fewer bilinear terms that depend on the non-negative variables.

The problem described is very complicated and is difficult to solve, even for small sized instances. However, if we fix the value of some variables (corresponding to taking decisions on the network layout and on the water flows and area to irrigate) we obtain a mixed binary linear program (MBLP) which, though still NP-hard (it contains the minimum Steiner tree as a particular case), can be solved by any available MILP solver, such as CPLEX. This has led us to devise a method in which some decisions are heuristically taken in order to obtain the more restricted version of the problem.

4 Decomposition-based heuristic

We propose a simple but efficient heuristic algorithm by devising a decomposition sequential scheme, coupled with a local search procedure. The decomposition scheme is based on three sub-problems, namely, the network layout sub-problem, computation of the water flows as well as the irrigated areas. Finally, we have the sub-problem of dimensioning the network pipes and pumps and locating the pumps, defined for the tree, flows and irrigated areas resulting from the previous sub-problems. The sub-problems are sequentially tackled, beginning with the first, which is solved through a heuristic process based on a Steiner tree constructive heuristic, followed by an improvement heuristic; then the flows and the irrigated area for each arc of the previously obtained network are trivially calculated. Finally, the solution to the last sub-problem is obtained by solving an MBLP problem.

Algorithm 1 below presents the decomposition heuristic. Section 4.1 then indicates the different versions of the constructive procedure developed to build the network layout. Section 4.2 describes the local search to improve the constructive solutions and, finally, Sect. 4.3 explains the MBLP model, which corresponds to the third sub-problem.

4.1 Constructive procedure

As mentioned in Sect. 2, the network studied in this paper has a non-looped topology, based on a Steiner tree rooted at a dummy node σ (Sect. 3.1). The constructive procedure does not consider the dummy node and respective arcs. Therefore the solution $\hat{\mathcal{G}} = (\hat{\mathcal{N}}, \hat{\mathcal{A}})$ of the first sub-problem of the decomposition heuristic is a Steiner forest of trees rooted at withdrawal nodes, defined in the graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ introduced in Sect. 3.1, such that $\hat{\mathcal{N}} \subseteq \mathcal{N}$, $\hat{\mathcal{A}} \subseteq \mathcal{A}$ and $\hat{\mathcal{N}} \supset \mathcal{H}$. Each basic node, or hydrant, is included in $\hat{\mathcal{N}}$, regardless of the fact that Steiner nodes—junctions—are included in $\hat{\mathcal{N}}$ or not.

As noted above, the pure Steiner problem is still NP-hard. Therefore we consider a constructive heuristic for the network layout, which is a straightforward adaptation of the *shortest path insertion heuristic* due to Takahashi and Matsuyama (1980) for the minimal length Steiner tree in an undirected graph. We should, however, emphasize that in some versions of this heuristic we consider arc lengths, as in the previous work by Gonçalves and Pato (2000), besides the slopes of the arcs. This amounts to an attempt to obtain the network layout by considering operating costs beyond the lengths of the arcs. In another version of the constructive heuristic the costs criteria are based on the optimal solution to a relaxation of model (P1).

Algorithm 1 Decomposition heuristic

Input: $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ and the input data of Tables 1 and 2
Output: $\hat{\mathcal{G}} = (\hat{\mathcal{N}}, \hat{\mathcal{A}})$; $\hat{y}_{ij}, \hat{s}_{ij}, \forall (i,j) \in \hat{\mathcal{A}}$; $\hat{pd}_j, \hat{ps}_j, \hat{h}_j, \hat{u}_j, \forall j \in \hat{\mathcal{N}}$; $\hat{\theta}_{ijdp}, \forall (i,j) \in \hat{\mathcal{A}}, (d,p) \in \hat{\mathcal{DP}}_{ij}$; or no feasible solution

1: Compute the network layout by a constructive procedure (Sect. 4.1) and denote the solution by $\bar{\mathcal{G}} = (\bar{\mathcal{N}}, \bar{\mathcal{A}})$;

2: Compute a new network layout by applying an improvement procedure (Sect. 4.2) to the solution $\bar{\mathcal{G}}$ previously obtained and denote the improved solution by $\hat{\mathcal{G}} = (\hat{\mathcal{N}}, \hat{\mathcal{A}})$;

3: Compute the water flows $y_{ij}, \forall (i,j) \in \hat{\mathcal{A}}$, by solving the following system of equations:

$$\sum_{(i,j)\in\hat{\mathcal{A}}} y_{ij} - \sum_{(j,i)\in\hat{\mathcal{A}}} y_{ji} = Q_j, \quad \forall j \in \hat{\mathcal{N}} \setminus \mathcal{T}$$

and denote its solution by $\hat{y}_{ij}, \forall (i,j) \in \hat{\mathcal{A}}$;

4: Compute the land area $s_{ij}, \forall (i,j) \in \hat{\mathcal{A}}$, by solving the following system of equations:

$$\sum_{(i,j)\in\hat{\mathcal{A}}} s_{ij} - \sum_{(j,i)\in\hat{\mathcal{A}}} s_{ji} = S_j, \quad \forall j \in \hat{\mathcal{N}} \setminus \mathcal{T}$$

and denote its solution by $\hat{s}_{ij}, \forall (i,j) \in \hat{\mathcal{A}}$;

5: **if** $\hat{\mathcal{DP}}_{ij} \neq \emptyset, \forall (i,j) \in \hat{\mathcal{A}}$ **then**

6: Solve the MBLP problem (P2) (Sect. 4.3) on the variables $h_j, pd_j, ps_j \geq 0, \forall j \in \hat{\mathcal{N}}$; $u_j \in \{0,1\}, \forall j \in \hat{\mathcal{N}}$; $\theta_{ijdp} \in \{0,1\}, \forall (i,j) \in \hat{\mathcal{A}}, (d,p) \in \hat{\mathcal{DP}}_{ij}$, resulting from replacing the solution of the previous sub-problems into the model (P1)

7: **end if**

As formulated in pseudocode below, Algorithm 2 begins by initially including all withdrawal nodes. Then, in each iteration, the least cost path for a given cost function, between the nodes of the current solution $\mathcal{G}_\ell = (\mathcal{N}_\ell, \mathcal{A}_\ell)$ and the hydrants absent from the current solution, is calculated and added to $\mathcal{G}_\ell$. The algorithm stops when all hydrants belong to the solution.

To obtain feasible solutions for the network layout sub-problem by means of the constructive heuristic, we consider the different cost functions $cost1(i,j)$ and $cost2(i,j)$ displayed in Table 4. As mentioned above, in previous work (Gonçalves and Pato 2000) we considered only the first option for these costs.

The least cost path required in each iteration of Algorithm 2 is obtained by computing, for each node $i_1 \in \mathcal{N}_\ell$, a tree $\mathcal{T}_{i_1} = (\mathcal{N}_{i_1}, \mathcal{A}_{i_1})$ in the graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, rooted at the node i_1, such that $\mathcal{N}_{i_1} \cap \mathcal{N}_\ell = \{i_1\}$ and spanning all hydrants outside the current solution, that minimizes a given cost function $\sum_{(i,j)\in\mathcal{A}_{i_1}} cost1(i,j)$. Finally, the path sought is selected from the set of all trees calculated, and corresponds to path $(i_1, i_2, \dots, i_k)$ between $i_1 \in \mathcal{N}_\ell$ and $i_k \in \mathcal{H} \setminus \mathcal{N}_\ell$, that minimizes another given cost function $\sum_{j=1}^{k-1} cost2(i_j, i_{j+1})$.

The solution of Algorithm 2 is a forest of trees rooted at withdrawal nodes and possibly one or more isolated withdrawals. These isolated points will be eliminated from the solution because they define potential withdrawal locations that will not be effective in the network.

Algorithm 2 Steiner-based heuristic for the network layout

Input: $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, L_{ij}, $cost1(i,j)$, $cost2(i,j)$, $\forall (i,j) \in \mathcal{A}$

Output: $\bar{\mathcal{G}} = (\bar{\mathcal{N}}, \bar{\mathcal{A}})$, $Length\_of\_\bar{\mathcal{G}}$

1: Let $\mathcal{G}_0 = (\mathcal{N}_0, \mathcal{A}_0)$ be the initial solution
$\mathcal{N}_0 := \mathcal{T}$; $\mathcal{A}_0 := \emptyset$; $Length\_of\_\mathcal{G}_0 := 0$;
2: $\ell := 0$;
3: **repeat**
4: $cost^* := +\infty$;
5: **for** each node $i_1 \in \mathcal{N}_\ell$ **do**
6: Compute the tree $\mathcal{T}_{i_1} = (\mathcal{N}_{i_1}, \mathcal{A}_{i_1})$ in the graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, rooted at the node i_1, such that $\mathcal{N}_{i_1} \supset \mathcal{H} \setminus \mathcal{N}_\ell$, $\mathcal{N}_{i_1} \cap \mathcal{N}_\ell = \{i_1\}$, with least cost $\sum_{(i,j)\in\mathcal{A}_{i_1}} cost1(i,j)$;
7: Find the path $path_{i_1} = (i_1, i_2, \ldots, i_{k-1}, i_k)$ in the tree $\mathcal{T}_{i_1}$, with least cost $cost\_of\_path_{i_1} = \sum_{j=1}^{k-1} cost2(i_j, i_{j+1})$;
8: **if** $path_{i_1}$ does not exist **then**
9: $cost\_of\_path_{i_1} := +\infty$
10: **end if**
11: **if** $cost\_of\_path_{i_1} < cost^*$ **then**
12: $cost^* := cost\_of\_path_{i_1}$;
13: $nodes^* := \{i_1, i_2, \ldots, i_{k-1}, i_k\}$;
14: $arcs^* := \{(i_1,i_2), (i_2,i_3), \ldots, (i_{k-1}, i_k)\}$
15: **end if**
16: **end for**
17: $\mathcal{N}_{\ell+1} := \mathcal{N}_\ell \cup nodes^*$;
18: $\mathcal{A}_{\ell+1} := \mathcal{A}_\ell \cup arcs^*$;
19: $Length\_of\_\mathcal{G}_{\ell+1} := Length\_of\_\mathcal{G}_\ell + \sum_{(i,j)\in arcs^*} L_{ij}$;
20: $\ell := \ell + 1$
21: **until** $\mathcal{H} \subset \mathcal{N}_\ell$
22: $\bar{\mathcal{N}} := \mathcal{N}_\ell$; $\bar{\mathcal{A}} := \mathcal{A}_\ell$; $Length\_of\_\bar{\mathcal{G}} := Length\_of\_\mathcal{G}_\ell$

Table 4 Different options for $cost1(i,j)$ and $cost2(i,j)$ in Algorithm 2

Constructive heuristic	Algorithm 2	
	Line 6 $cost1(i,j)$	Line 7 $cost2(i,j)$
1	L_{ij}	L_{ij}
2	$Z_j - Z_i$	$Z_j - Z_i$
3	$\lvert Z_j - Z_i \rvert$	$\lvert Z_j - Z_i \rvert$
4	L_{ij}	$\lvert Z_j - Z_i \rvert$
5	L_{ij}	$\frac{1}{4}[3\lvert Z_j - Z_i \rvert + L_{ij}]$
6	$1 - \overline{x}_{ij}$	$1 - \overline{x}_{ij}$

In Version 1 of Algorithm 2, optimization is only performed on the lengths of arcs, whereas in Versions 2–5 we consider the slopes or a combination of lengths and slopes. The purpose is to obtain a more balanced network, in an attempt to reduce the number of pumps required in the network in view of the very high cost of pumping water.

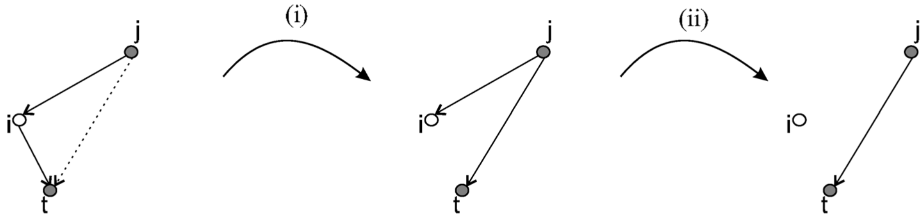


Fig. 3 Neighbors generation scheme

Version 6 of the heuristic developed to build the network design uses the values $\overline{x}_{ij}$ of the binary variables x_{ij} given by the optimal solution of the model (P1) continuous relaxation in Sect. 3.3. The basic goal of this version is to choose the arcs associated with the highest $\overline{x}_{ij}$ values.

4.2 Improvement procedure

In this subsection we present the main characteristics of our local search procedure (line 2 of Algorithm 1). This was implemented in order to improve the feasible solutions from the different versions of the constructive heuristic, as previously described.

Given a particular solution of any version of the constructive heuristic, denoted by $\bar{\mathcal{G}} = (\bar{\mathcal{N}}, \bar{\mathcal{A}})$, the neighborhood of such a solution consists of the set of Steiner forests that can be obtained by removing and adding precisely one arc from/to the current solution $\bar{\mathcal{G}}$.

Each neighbor solution is obtained by performing the following two operations in the order to be found in Fig. 3:

(i) if (j,i), $(i,t) \in \bar{\mathcal{A}}$ and if $(j,t) \in \mathcal{A}$, where $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ is the original graph, then arc (j,t) is added to graph $\bar{\mathcal{G}}$ and (i,t) is removed from $\bar{\mathcal{G}}$;
(ii) if i is a Steiner node of degree one on graph $\bar{\mathcal{G}}$, then node i and arc (j,i) are removed from $\bar{\mathcal{G}}$.

Note that the size of the neighborhood is bounded by the number of arcs of the initial solution $\bar{\mathcal{G}}$. Therefore, as the neighborhood is not so large, we adopted the best-improvement strategy to select the next neighbor in the search procedure. The final solution denoted by $\hat{\mathcal{G}} = (\hat{\mathcal{N}}, \hat{\mathcal{A}})$ is a local optimum in terms of total length.

4.3 Mixed binary linear programming problem

As mentioned in line 6 of Algorithm 1, the MBLP problem corresponding to the last sub-problem is obtained by replacing the solution of the first two sub-problems in the model (P1) defined in Sect. 3.3. To define the MBLP problem, let us consider, for each arc $(i,j) \in \hat{\mathcal{A}}$, selected for the solution of the first sub-problem, the set $\hat{\mathcal{DP}}_{ij}$ of diameters in $\mathcal{DP}_{ij}$ such that the water speed corresponding to the flow $\hat{y}_{ij}$ resulting from the second sub-problem, is feasible. Note that, in theory, the set $\hat{\mathcal{DP}}_{ij}$ could be empty, in which case the algorithm will not produce a feasible solution.

By replacing in the model (P1) the values $x_{ij} = 1$, $y_{ij} = \hat{y}_{ij}$ and $s_{ij} = \hat{s}_{ij}$ for each arc $(i,j) \in \hat{\mathcal{A}}$, selected for the solution of the first sub-problem, $x_{ij} = 0$, $y_{ij} = 0$ and $s_{ij} = 0$ for each arc $(i,j) \in \mathcal{A}_\sigma \setminus \hat{\mathcal{A}}$, $\theta_{ijdp} = 0$ for each pair $(d,p) \notin \hat{\mathcal{DP}}_{ij}$ and for each arc $(i,j) \in \hat{\mathcal{A}}$, $\theta_{ijdp} = 0$ for each arc $(i,j) \in \mathcal{A} \setminus \hat{\mathcal{A}}$ and for each pair $(d,p) \in \mathcal{DP}_{ij}$, $\alpha_{ijd} = 0$ for each arc $(i,j) \in \hat{\mathcal{A}}$ and for each diameter d such that $(d,p) \notin \hat{\mathcal{DP}}_{ij}$, for each $p \in \mathcal{P}$, $\alpha_{ijd} = 0$ for each arc $(i, j) \in \mathcal{A}_\sigma \setminus \hat{\mathcal{A}}$ and for each diameter $d \in \mathcal{D}_{ij}$, $\varepsilon_{ijp} = 0$ for each arc $(i, j) \in \mathcal{A} \setminus \hat{\mathcal{A}}$ and for each pressure class $p \in \mathcal{P}$, $w_j = \sum_{i\in\hat{\mathcal{N}}:(i,j)\in\hat{\mathcal{A}}}(CP\,\hat{y}_{ij} + CE\,\hat{s}_{ij}) = \hat{w}_j$ for each node $j \in \hat{\mathcal{N}}$ and, finally, $w_j = u_j = pd_j = ps_j = h_j = 0$ for each node $j \in \mathcal{N} \setminus \hat{\mathcal{N}}$, we reach the following MBLP problem:

$$\text{(P2)}\quad \text{Minimize} \sum_{(i,j)\in\hat{\mathcal{A}}} L_{ij} \sum_{(d,p)\in\hat{\mathcal{DP}}_{ij}} CC_{dp}\,\theta_{ijdp} + \sum_{j\in\hat{\mathcal{N}}} \hat{w}_j\, h_j \tag{30}$$

$$\text{subject to} \sum_{(d,p)\in\hat{\mathcal{DP}}_{ij}} \theta_{ijdp} = 1 \quad \forall (i, j) \in \hat{\mathcal{A}} \tag{31}$$

$$pd_i - ps_j + Z_{ij} = \hat{y}_{ij}^{1.79} \sum_{(d,p)\in\hat{\mathcal{DP}}_{ij}} E_{ijd}\theta_{ijdp} \quad \forall (i, j) \in \hat{\mathcal{A}} \tag{32}$$

$$1.1pd_i - 0.1ps_j + 1.1Z_{ij} \le \sum_{(d,p)\in\hat{\mathcal{DP}}_{ij}} \frac{p}{2}\theta_{ijdp} \quad \forall (i, j) \in \hat{\mathcal{A}} \tag{33}$$

$$pd_j \ge P_j^{min} \quad \forall j \in \mathcal{H} \tag{34}$$

$$ps_j = 0 \quad \forall j \in \mathcal{T} \cap \hat{\mathcal{N}} \tag{35}$$

$$h_j = pd_j - ps_j \quad \forall j \in \hat{\mathcal{N}} \tag{36}$$

$$H_j^L u_j \le h_j \le H_j^U u_j \quad \forall j \in \hat{\mathcal{N}} \tag{37}$$

$$\theta_{ijdp} \in \{0, 1\} \quad \forall (i, j) \in \hat{\mathcal{A}},\ (d, p) \in \hat{\mathcal{DP}}_{ij} \tag{38}$$

$$u_j \in \{0, 1\} \quad \forall j \in \hat{\mathcal{N}} \tag{39}$$

$$h_j \in [0, H_j^U], \qquad pd_j \in [0, PD_j^U], \qquad ps_j \in [0, PS_j^U] \quad \forall j \in \hat{\mathcal{N}}. \tag{40}$$

The objective function (30) represents the total cost of the network; whereas constraints (31) state that only one pair (d, p) is selected for each arc; equations (32) are energy balance conditions, while constraints (33) impose an upper bound to the pressure in each arc, and inequalities (34) establish the minimum pressure required at each hydrant. Conditions (35)–(37) define the minimum and maximum height for a pump at each node and, finally, conditions (38)–(40) represent the domain of the variables.

5 Computational experiments

This section reports on computational tests performed to evaluate Algorithm 1. As benchmark instances are not available for this problem, we used randomly generated test instances built from real data to evaluate the quality of the solutions provided by this algorithm.

We begin by describing the test instances in Sect. 5.1. We then report in Sect. 5.2 a summary analysis, evaluating the impact of the local search (line 2 of Algorithm 1) on the quality of the decomposition heuristic's solutions. Finally, still in Sect. 5.2, we compare the cost of the algorithm's solution for different versions of the constructive heuristic (line 1 of Algorithm 1) with the optimum value of WDNP when it is known, or with a lower bound for the optimum value of WDNP. The optimum was obtained through global optimization software BARON (GAMS 2004), whereas the lower bound was chosen between the one given by BARON and another lower bound obtained from a linearization-based relaxation of model (P1), presented in Gonçalves (2008). This last relaxation is based on linearizations of model (P1), where the bilinear terms $h_j w_j$ in the objective function (27) are replaced by a relaxation of its convex and concave envelopes proposed by McCormick (1976) and also by Al-Khayyal and Falk (1983). The concave terms $-y_{ij}^{1.79}$ in constraints (10a) are replaced by a piecewise-linear approximation defined in the interval $[0, Y_{ij}^U]$, whereas the convex terms $y_{ij}^{1.79}$ in constraints (10b) are also replaced by an outer-linear approximation in the interval $[0, Y_{ij}^U]$. The result is an MBLP, which was solved through the CPLEX 11.1 (CPLEX Optimization 2007) software.

5.1 Test instances

The instances of the WDNP used in our computational experiments were randomly generated on the basis of a real instance described in Gonçalves and Pato (2000), corresponding to the network illustrated in Fig. 1 of Sect. 2.

The real instance is characterized by three different zones, namely the large area indicated in Fig. 1 by the letter B, whose area for irrigation ranges between 19 and 84.2 hectares. The small area indicated by the letter S has an area for irrigation of between 3.1 and 8.6 hectares and, finally, the intermediate area denoted by the letter I, forms the connection between the two areas. This network has one withdrawal and 39 hydrants, 18 of which are located in the S area and 21 in the B area. The number of junction nodes is 13, 7 of which are located in the B area and the remaining 6 in the S area. The total number of arcs of this network is 279, 70 of which are in the S area, 198 in the B area, 6 in the intermediate area I and the remaining 5 arcs connect the withdrawal to the nodes in the large area.

The main features of this real case are summarized in Table 5. On the basis of these data and on the characteristics of the network described above, we generated 12 small test problems consisting of five different types, each possessing 10 nodes and a number of arcs, ranging between 20 and 40. The main difference between these five types of instances is the dimension of the area to irrigate by the network, which can amount to between 19 and 85 hectares, corresponding to the large area (B), or between 3 and 9 hectares equating to the small area (S), or it can even include both areas. All generated instances have one water withdrawal, whose location area, denoted by W, is different for each instance type, as indicated in Fig. 4. In this figure, BA represents the instance type for which the area to be irrigated is type B, while for the SA instance type the area is type S. In the remaining types of instances (denoted by BW, SW and BSW) the two areas coexist.

As mentioned above, Table 5 provides the parameters of the real instance, as well as the corresponding values adopted to generate the test instances. Note that we considered two arcs at the most, between the withdrawal and the remaining nodes, as well as two arcs at the most in area I, in the case of BW, SW and BSW instance types. We should also add that the arc lengths were defined as the Euclidean distance between the end points of the arc. As for the coefficients in the cost function of model (P2), we considered $\eta = 0.75$ for pump efficiency, $Am = 15$ for the number of years' amortization, $C_1 = 17.644$ (monetary units/kW), $C_2 = 0.6554$ (monetary units/kW), $C_3 = 0.011$ (monetary units/kWh) and $\nu = 7000$ (m^3/ha/year). For the investment cost on a pipe C_{dp} we adopted the values provided by Table 8 of Appendix. The bounds V_d^L and V_d^U (m/s) for water speed in pipes, used to model constraints (9), are given in Table 9 in the same Appendix.

5.2 Computational results

Computational tests were performed with the six versions of the constructive Steiner-based heuristic for the network's layout (Algorithm 2 in Sect. 4.1), as well as with the local search

Table 5 Parameters adopted to generate the test instances

Zone	Symbol	Real case	Instances
	$\|\mathcal{N}\|$	53	10
	$\|\mathcal{T}\|$	1 ($\simeq$ 2% of 53)	1
	$\|\mathcal{H}\|$	39 ($\simeq$ 73.5% of 53)	$\simeq$ 73.5% of $\|\mathcal{N}\|$
S	$\|\mathcal{H}\|$	18 ($\simeq$ 46% of 39)	$\simeq$ 46% of $\|\mathcal{H}\|$
B	$\|\mathcal{H}\|$	21 ($\simeq$ 54% of 39)	$\simeq$ 54% of $\|\mathcal{H}\|$
	$\|\mathcal{J}\|$	13 ($\simeq$ 24.5% of 53)	0, 1 or 2
	$\|\mathcal{A}\|$	279	between 20 and 40
S	$\|\mathcal{A}\|$	70 ($\simeq$ 25% of 279)	$\simeq$ 25% of $\|\mathcal{A}\|$
B	$\|\mathcal{A}\|$	198 ($\simeq$ 71% of 279)	$\simeq$ 71% of $\|\mathcal{A}\|$
I	$\|\mathcal{A}\|$	6 ($\simeq$ 2% of 279)	1 or 2
W	$\|\mathcal{A}\|$	5 ($\simeq$ 2% of 279)	1 or 2
S	S_j (ha)	[3.1, 8.6]	[3, 9]
B	S_j (ha)	[19, 84.2]	[19, 85]
S	Q_j (m^3/h)	{15, 20, 25, 35, 40}	{15, 20, 25, 35, 40}
B	Q_j (m^3/h)	{100, 120, 145, 175, 200, 250}	{100, 120, 145, 175, 200, 250}
	Ox	[30 655, 37 322.5]	according to Fig. 4
S	Oy	[−135 030, −134 005]	according to Fig. 4
B	Oy	[−133 935, −128 618]	according to Fig. 4
I	Oy	[−134 005, −127 750]	according to Fig. 4
W	Oy	[−128 618, −127 750]	according to Fig. 4
	Z_j (m)	[189, 222]	[180, 230]
	P_j^{min} (m)	45	[30,50]
	H_j^L (m)	25	25

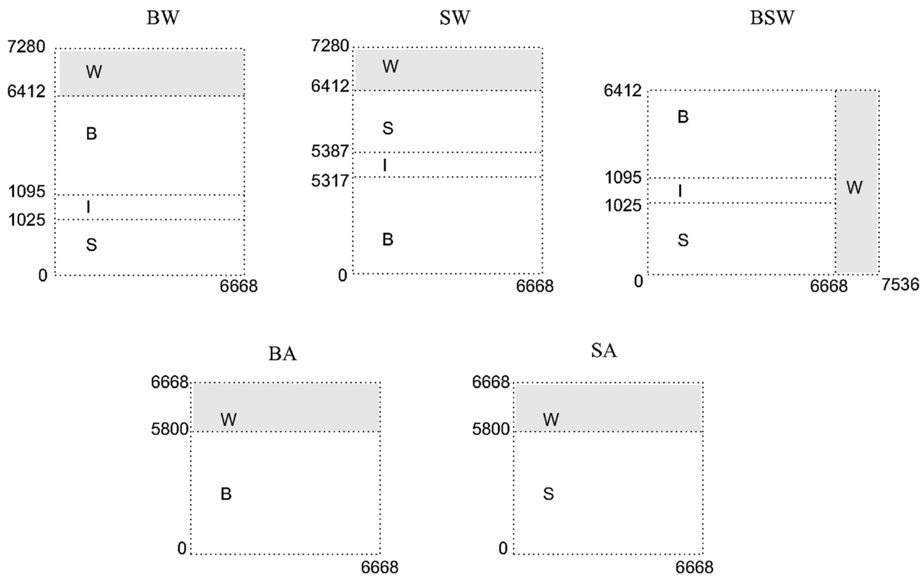


Fig. 4 Test instances

procedure (algorithm in Sect. 4.2). Local search starts from the solutions provided by these six constructive heuristic versions in order to improve the length of the network, while maintaining the balance as regards the slopes, in the case of Versions 2–5.

The different versions of the constructive heuristic algorithm and the local search algorithm were implemented in *Pascal*. The model (P2)—which arises within the three-phase decomposition heuristic—was solved using software CPLEX 11.1 with standard parameters. Version 6 of the constructive heuristic starts from the solution of the continuous relaxation of the problem (P1) obtained with GAMS-BARON using $\epsilon_r = 1E$–09. All experiments were performed on a Pentium®, with a 3.6 GHz processor and 3.25 GB RAM.

Firstly, we intended to evaluate the different versions of the constructive heuristic, besides the impact of the local search procedure on the quality of the decomposition heuristic solution. The cost used to evaluate the solutions is the total cost described in (2), including investment, power and energy costs as defined for the WDNP.

In Table 6 we compare the results of the decomposition heuristic when the network layout is calculated by each Version 1–5 of the simple constructive heuristic, denoted by CH_1 to CH_5, plus local search, with the results when the layout is built by the relaxation-based heuristic, Version 6, denoted by CH_6, also plus local search.

The first columns of this table show the main characteristics of the 60 tested instances: the first column identifies the instance; the second one indicates its type in accordance with the generation process in Sect. 5.1; and the third displays the number of network arcs.

Columns 4 to 9 of Table 6 show the percentage improvement $[UB(CH_i) - UB(LS_i)]/UB(CH_i) \times 100\%$ made by the local search ($UB(LS_i)$) over the solution value ($UB(CH_i)$) of the constructive heuristic, version CH_i. Bold figures indicate that the respective constructive plus local search attained the best upper bound value. The last three rows of the table show the average improvements over the entire 60 instances, along with the number of times each heuristic version attained the best value among the six possibilities and, finally, the average CPU times of the decomposition heuristic for the different initial network layouts.

The results found in this table reveal that local search did not act on any of the solutions from CH_1 (see fourth column). This conclusion was to be expected insofar as the improvement criterion is arc length, which is precisely the minimization criterion of Version 1 of the constructive process (in Table 4). Local search over CH_2 and CH_5 led to few improvements (see fifth and eighth columns). It should also be noted that local search produced a greater effect when the initial solution is given by CH_3 and CH_4 (sixth and seventh columns). For two cases—instances 40 and 59—the best solution among the six was precisely obtained from local search.

The constructive heuristic based on continuous relaxation produced solutions that were seldom improved upon by local search, as confirmed in the last column.

Hence, it is clear that in some cases there was a substantial improvement in the cost of the solution as a result of local search application, even if in other cases no improvement was achieved. From all these figures no relation as to heuristic solution quality and characteristics of the instances is evident. Solution quality is irregular, and clearly no single heuristic always dominates the others. However, if we start with CH_1, CH_5 or CH_6 solutions, Algorithm 1 produced better results, even if for some cases this algorithm came up with the best solutions starting from network structures built by CH_2, CH_3 and CH_4.

The CPU time of the relaxation-based heuristic is much higher than the other CPU times (see last row of Table 6). In fact, constructive heuristic CH_6 calls BARON algorithm to solve the continuous relaxation, which is a nonlinear programming problem.

Finally, it may be concluded that the use of local search incorporated in the decomposition heuristic, starting from solutions of CH_1–CH_5, is justified insofar as, for some test

Table 6 Percentage improvement by local search

Instance			Constructive plus local search heuristic					
Id.	Type	$\lvert\mathcal{A}\rvert$	1	2	3	4	5	6
1	BW	23	0	0	**0**	**0**	0	**0**
2	BW	21	0	0	7.26	7.26	0	**0**
3	BW	23	**0**	0	4.78	4.78	**0**	**0.07**
4	BW	26	**0**	0.90	0.98	0.98	**0**	**0**
5	BW	24	**0**	0	0	0	**0**	**0**
6	BW	25	**0**	13.51	13.51	13.51	**0**	**0**
7	BW	34	**0**	0	0	0	**0**	**0**
8	BW	34	**0**	11.51	4.66	4.66	**0**	**0**
9	BW	33	0	9.32	**10.05**	**10.05**	0	**0**
10	BW	34	**0**	10.85	1.90	1.90	**0**	0
11	BW	34	**0**	**12.75**	0	0	**0**	**0**
12	BW	34	**0**	0	17.09	17.09	**0**	0
13	SW	22	**0**	0	0	0	**0**	**0**
14	SW	23	**0**	0	0	0	**0**	**0**
15	SW	22	0	0	0	0	0	**0**
16	SW	23	**0**	0	0	0	**3.61**	**0**
17	SW	22	**0**	0	0	0	**0**	**0**
18	SW	22	**0**	0	**0**	**0**	**0**	**0**
19	SW	33	0	0	0	0	0	**0**
20	SW	33	0	0	0	0	0	**0**
21	SW	34	0	21.71	29.19	22.84	0	**0**
22	SW	33	0	0	2.45	2.45	**0**	**0**
23	SW	34	**0**	0	0	0	**0**	**0.17**
24	SW	33	0	0	11.11	11.11	0	**0**
25	BSW	22	**0**	5.45	6.95	6.95	**0**	**0**
26	BSW	22	**0**	**0**	0	0	**0**	0
27	BSW	24	**0**	0	0	0	**0**	0
28	BSW	24	**0**	0	1.51	1.51	**0**	0
29	BSW	24	0	**0**	0	0	0	0
30	BSW	24	0	0	0	0	0	**0**
31	BSW	34	0	14.52	3.56	3.56	0	**0**
32	BSW	32	**0**	0	8.43	8.43	**5.31**	0
33	BSW	34	**0**	0	0	0	**0**	**0**
34	BSW	34	0	0	3.46	3.46	0	**0**
35	BSW	34	0	0	**0**	**0**	0	0
36	BSW	34	**0**	0	2.53	2.53	**0**	**0**
37	SA	24	0	0	0	0	0	**0**
38	SA	25	**0**	7.71	0	0	**0**	0
39	SA	26	**0**	11.80	7.28	7.28	**0**	0
40	SA	24	0	0	**1.41**	**1.41**	0	0
41	SA	26	**0**	0	4.01	4.01	**0**	0
42	SA	25	**0**	0	0	0	**0**	1.24

Table 6 (*Continued*)

Instance			Constructive plus local search heuristic					
Id.	Type	$\|\mathcal{A}\|$	1	2	3	4	5	6
43	SA	36	**0**	0	0	1.62	**0**	**0**
44	SA	36	0	0	0	0	0	**0**
45	SA	36	**0**	0	0	0	**0**	**0**
46	SA	36	0	0	11.14	11.14	0	**0**
47	SA	37	0	0	**0**	**0**	0	0
48	SA	35	0	0	2.50	2.50	0	**0**
49	BA	22	**0**	5.79	5.79	5.79	**0**	**0**
50	BA	24	**0**	0	0	0	**0**	0
51	BA	23	**0**	0	0	0	**0**	**0**
52	BA	22	0	0	0	0	0	**0**
53	BA	24	**0**	0	3.27	3.27	**0**	0
54	BA	24	**0**	0	**0**	**0**	**0**	0
55	BA	36	0	0	2.48	2.48	0	**0**
56	BA	37	**0**	14.86	30.94	30.94	**0**	0
57	BA	36	**0**	1.99	2.07	2.07	**0**	**0**
58	BA	36	0	0	7.93	7.93	0	**0**
59	BA	37	0	0	**1.43**	**1.43**	0	0
60	BA	37	0	**0**	5.51	5.51	0	0
Average improvement			0.00	2.38	3.59	3.51	0.15	0.02
Number of best solutions			35	4	8	8	36	40
Average CPU time (sec)			0.14	0.14	0.13	0.14	0.10	172.96

problems, we obtained a substantially lower cost solution. Moreover, in 44 cases, the best result was achieved by starting from CH_1–CH_5 heuristics solutions, while in 40 cases the best result was obtained by starting from a CH_6 solution. When starting from a CH_1–CH_5 solution, the computing time expenses of Algorithm 1 were also very low.

In Table 7 we evaluate the quality of upper bounds for the optimum value of the WDNP generated by the decomposition heuristic (Algorithm 1). Lower bounds or optimum values are used to evaluate the optimality gap associated with the upper bounds.

Table 7 identifies the instances in the first column and, in the second and third columns, respectively, presents the linearization-based lower bounds and the lower bounds or the optimum values (labeled with an asterisk symbol) from BARON using $\epsilon_r = 1E$–09. The lower bound from BARON was achieved after 18000 CPU seconds (i.e., 5 CPU hours). The values reported in the fourth and fifth columns represent the deviation $\frac{UB(LS)-LB}{LB} \times 100\%$, where $UB(LS)$ is the upper bound given by the decomposition heuristic and LB is the best known lower bound, from the second and third columns. Naturally, these figures represent assured optimality gaps.

In the fourth column we consider the best upper bound from the five upper bounds obtained by the decomposition heuristic, starting from network structures built by each heuristic CH_1–CH_5. The fifth column, however, is devoted to the decomposition heuristic, starting from the network structure built by CH_6. In comparing the fourth and fifth columns, the best result is identified in bold characters for the instances with a known optimum. The sixth

Table 7 Assured optimality gaps

Inst. id.	Lower bound		Gap		Time	
	Lin.-based rel. (m.u.)	BARON (m.u.)	$1+\cdots+5$ (%)	6 (%)	$1+\cdots+5$ (sec)	BARON (sec)
1	12154.588	11039.417	9.32	9.32	0.49	18000.00
2	7614.000	*8243.377	3.99	**2.53**	0.30	9827.80
3	9404.475	8991.461	8.23	8.23	0.43	18000.00
4	10401.432	5887.511	20.57	20.57	0.41	18000.00
5	14991.810	11840.533	8.27	8.27	0.47	18000.00
6	13353.947	10360.679	5.43	5.43	1.28	18000.00
7	7948.580	5325.593	19.25	19.25	1.41	18000.00
8	5558.161	4533.898	5.91	5.91	0.31	18000.00
9	5154.216	3821.561	14.07	14.07	0.41	18000.00
10	12329.848	4201.156	19.37	19.98	0.43	18000.00
11	12389.895	6566.216	9.73	9.73	0.32	18000.00
12	8903.354	3621.683	10.54	10.58	0.45	18000.00
13	9038.321	9621.197	10.26	10.26	1.35	18000.00
14	11359.083	6495.145	3.73	3.73	0.52	18000.00
15	8026.769	6270.361	12.23	7.22	0.53	18000.00
16	17341.175	13487.387	9.36	9.36	0.53	18000.00
17	11427.269	6306.305	23.82	23.82	0.35	18000.00
18	8770.907	5634.541	21.97	21.97	0.30	18000.00
19	8087.303	4055.794	9.89	6.81	0.36	18000.00
20	5720.685	4191.672	26.12	16.42	0.31	18000.00
21	4983.075	3721.461	25.60	24.90	0.51	18000.00
22	22002.585	8388.513	11.17	11.17	0.44	18000.00
23	10833.487	4402.369	11.02	11.02	0.41	18000.00
24	11577.698	7391.853	19.43	13.84	1.34	18000.00
25	9725.748	9440.253	9.84	9.84	1.41	18000.00
26	8145.722	7483.543	20.73	20.82	1.34	18000.00
27	7362.953	*7620.413	**11.37**	29.09	0.48	14935.25
28	13876.183	12053.854	14.39	23.35	0.36	18000.00
29	17014.597	*18500.404	**1.88**	23.02	0.45	17067.93
30	13021.986	13000.025	23.76	14.88	1.27	18000.00
31	8700.249	5767.625	13.56	12.05	0.44	18000.00
32	5819.757	3597.274	2.85	11.57	1.33	18000.00
33	5199.326	3269.159	14.13	14.13	0.63	18000.00
34	16385.975	8254.827	24.81	16.61	1.54	18000.00
35	15010.697	8035.919	8.62	11.14	0.40	18000.00
36	10739.134	8503.496	7.03	7.03	0.44	18000.00
37	5389.369	*5892.141	2.77	**0.29**	0.33	636.72
38	5563.483	*5930.646	**0.42**	0.49	0.34	8895.11
39	3841.975	*4030.160	**0.29**	6.28	1.29	516.30
40	6383.666	*7081.819	**1.75**	1.85	0.43	13703.87
41	6216.492	*7073.200	**0.43**	6.77	0.49	4247.53
42	4840.352	*5239.238	**0.96**	6.74	0.44	2376.14

Table 7 (*Continued*)

Inst. id.	Lower bound		Gap		Time	
	Lin.-based rel. (m.u.)	BARON (m.u.)	$1+\cdots+5$ (%)	6 (%)	$1+\cdots+5$ (sec)	BARON (sec)
43	3724.284	*3884.988	0.51	0.51	0.38	7068.13
44	3310.648	*3619.361	5.71	**0.55**	0.46	6370.24
45	3666.495	*3852.484	0.27	0.27	0.23	9187.46
46	4555.850	3431.321	26.41	12.64	0.44	18000.00
47	5455.316	3787.855	9.41	11.16	0.40	18000.00
48	5989.150	4937.307	7.33	4.39	0.34	18000.00
49	13323.775	13137.036	21.10	21.10	0.42	18000.00
50	12100.258	*12693.238	**0.97**	8.77	0.53	2572.07
51	12681.066	10532.731	13.38	13.38	0.43	18000.00
52	22541.144	15992.933	14.37	14.03	1.46	18000.00
53	23044.391	*24976.608	**0.83**	1.01	0.33	15751.34
54	20224.609	18047.137	12.87	13.26	1.42	18000.00
55	7752.085	3497.131	25.21	20.44	1.30	18000.00
56	7595.410	4397.164	19.72	19.81	0.52	18000.00
57	10221.327	5127.696	8.28	8.28	0.30	18000.00
58	20415.053	7099.431	31.10	17.71	1.53	18000.00
59	24895.765	9107.817	8.98	9.78	1.30	18000.00
60	13618.601	3358.203	7.70	16.17	0.42	18000.00
Average values			11.55	11.73	0.65	15685.93
Average best values			7.70	11.58		
Average gaps for * cases			2.30	6.30		

column provides the sum of the CPU times of the five decomposition heuristic runs that produced the results in the fourth column. The last column of this table contains the BARON CPU time.

From Table 7 we may conclude that the five runs of the decomposition heuristic starting from each CH_1–CH_5 heuristic's solution produced slightly better upper bounds (results in the fourth column) as compared to the upper bounds from the decomposition heuristic starting from the CH_6 network (results in the fifth column). In fact, on average the assured gaps are 11.55% and 11.73%, respectively, as seen in the antepenultimate row of the table. Moreover, the five decomposition heuristic runs attained 20 times the best upper bound for the optimum of WDNP, whereas the other version attained 16 times the best upper bound (as seen in Table 6). If one takes the average gaps within the subsets where the heuristics attained the best results, we find 7.7% and 11.58% (penultimate row of Table 7), again a most favorable behavior for the combination of five heuristics evaluated in the fourth column.

We also calculated the averages of the relative gaps when the optimum values are known. The figures in question are 2.30% and 6.30%, which again show a better performance for the decomposition heuristic found in the fourth column.

As for computation times, results clearly reveal that the average of the sum of the CPU time required by all the five decomposition heuristic runs, given in the sixth column of Table 7, is very low as compared with the CPU time of the other decomposition heuristic version (given as an average in the last row of the ninth column in Table 6). For this set of small instances the difference in the computational time is not relevant for the strategic problem studied. However, if we want to solve medium or large problem instances, usually from real contexts, a low CPU time will be relevant in choosing a version of the heuristic. As to the results produced by BARON, the computing time taken was excessively high: on average it took 8082.56 seconds to attain the 14 optimal values and, for each of the other cases, it terminated at the computational limit, without yielding an ϵ_r-optimal solution.

6 Conclusions

In this paper we have addressed the design of a pressurized water distribution network for irrigation purposes in a realistic context which considers investment and operating costs. We have proposed two new mixed binary nonlinear formulations which were specially devised for this problem: an initial model, as well as a reformulation of the model, designed to reduce its nonlinearities.

From the computational perspective, the WDNP is complex as it is NP-hard. In addition, the WDNP involves discrete options, as well as nonlinearities—both within the optimization objective and in the constraints. A simple but efficient heuristic was therefore devised to determine feasible solutions of WDNP and the respective upper bounds for the optimum. The approach is a sequential three-phase procedure, based on decomposition of the problem. It starts from constructive heuristics—either greedy or relaxation based—and a local search heuristic, both of which are typical for Steiner trees, thus providing the design of the pipes' network. In an intermediate phase, the water flows and the irrigated areas are obtained from a system of linear equations. Finally, the hydraulic components of the irrigation system (pipes and pumps) are located and dimensioned, using a standard software, by resolving an MBLP restriction of the reformulated nonlinear model.

The computational experiments were undertaken with sets of randomly- built instances to simulate different real case situations. The five independent runs of the decomposition heuristic algorithm, starting from the network layout built by the simple constructive heuristic in Versions 1–5, ran in very short computational times. Moreover, the quality of the feasible solutions obtained is promising for such a difficult combinatorial and nonlinear problem. In fact, the relative optimality gaps calculated for all cases with known optimum are, on average, 2.30%.

Future work will focus on incorporating these upper bounds within a branch and bound algorithm that exactly solves the WDNP. Further local search strategies for the first phase sub-problem can be suggested to improve the upper bounds, namely by using arc costs other than arc lengths. New heuristics to determine feasible solutions for the WDNP and respective upper bounds will also be developed, by dealing in a different way with the intrinsic nonlinearities and discrete decisions. For instance, exclusion of the pressure technical conditions would result in a MBLP problem depending only on flows and network layout. A feasible solution for the WDNP could then be obtained by enforcing the conditions hitherto unsatisfied. Other feasible solutions for the WDNP could be obtained through heuristics based on different relaxations of model (P1).

Acknowledgements This work is supported by Portuguese National Funding from FCT—Fundação para a Ciência e a Tecnologia, under the project: PEst-OE/MAT/UI0152.

The authors are grateful to the referees for the extensive suggestions that have considerably improved the paper.

Appendix: Input data

Table 8 Costs of the pipes commercially available (Cimianto, Sociedade Técnica de Hidráulica, S.A.)

	Pressure class—p (bar)				
	6	12	18	24	30
Diameter d (mm)	Cost C_{dp} (m.u./m)				
50	–	–	1.40	1.50	1.57
60	–	–	1.50	1.58	1.70
70	–	–	1.57	1.65	1.86
80	–	1.62	1.64	1.81	2.08
100	–	1.76	1.88	2.22	2.53
125	–	1.93	2.20	2.78	3.09
150	2.26	2.31	2.70	3.29	3.83
200	2.77	2.99	3.59	4.77	5.88
250	3.26	3.89	4.65	6.28	7.39
300	3.68	4.95	6.13	7.46	8.73
350	4.79	6.25	7.94	8.73	10.67
400	6.07	7.69	9.72	10.87	12.85
450	7.10	9.17	12.38	14.03	15.28
500	8.59	10.87	14.53	16.41	17.80
600	11.41	14.37	17.72	20.00	22.35
700	14.29	18.55	22.70	26.12	–
800	16.79	22.21	27.72	33.40	–
900	20.87	26.68	34.33	–	–
1000	24.59	31.99	41.08	–	–

1 bar ≈ 10.33 m,
1 m.u. ≈ 4.99 €

Table 9 Bounds for the water speed on pipes (Labye et al. 1988)

Diameter d (mm)	Minimum speed V_d^L (m/s)	Maximum speed V_d^U (m/s)
50, 60, 70, 80, 100	0.20	1.80
125	0.25	1.85
150	0.25	1.95
200	0.35	2.05
250	0.40	2.15
300	0.40	2.25
350	0.50	2.30
400	0.50	2.50
450, 500	0.50	2.85
600, 700, 800, 900, 1000	0.50	3.10

References

Al-Khayyal, F. A., & Falk, J. E. (1983). Jointly constrained biconvex programming. *Mathematics of Operations Research*, *8*(2), 273–286.

Alperovits, E., & Shamir, U. (1977). Design of optimal water distribution systems. *Water Resources Research*, *13*(6), 885–900.

Bragalli, C., D'Ambrosio, C., Lee, J., Lodi, A., & Toth, P. (2006). *An MINLP solution method for a water network problem* (IBM Research Report, RC23893 (W0602-210)). February 28, pp. 1–17.

CPLEX Optimization (2007). ©ILOG USA.

GAMS: The Solver Manuals (2004). ©GAMS Development Corporation.

Gonçalves, G. M., & Pato, M. V. (2000). A three-phase procedure for designing an irrigation system's water distribution network. *Annals of Operations Research*, *94*, 163–179.

Gonçalves, G. M. (2008). *Modelos de Optimização para o Desenho de uma Rede de Distribuição de Água sob Pressão em Sistemas de Rega*. Ph.D. Dissertation, Lisboa.

Hansen, C. T., Madsen, K., & Nielsen, H. B. (1991). Optimization of pipe networks. *Mathematical Programming*, *52*, 45–58.

Ionescu, V., Pantu, D., Berar, U., & Hutanu, V. (1981). Optimizing the dimensioning of a ramified network of pipe lines with flow variable in time. *Economic Computation & Economic Cybernetics Studies & Research*, *15*(3), 41–49.

Karmeli, D., Gadish, Y., & Meyers, S. (1968). Design of optimal water distribution networks. *Journal of the Pipeline Division*, *94*, 1–10.

Karp, R. M. (1972). In R. Miller & J. Thatcher (Eds.), *Complexity of computer computations*. New York: Plenum.

Kessler, A., & Shamir, U. (1991). Decomposition technique for optimal design of water supply networks. *Engineering Optimization*, *17*, 1–19.

Knowles, T., Gupta, I., & Hassan, M. (1976). Decomposition of water distribution networks. *AIIE Transactions*, *8*(4), 443–448.

Labye, Y., Olson, M., Galand, A., & Tsiourtis, N. (1988). *FAO irrigation and drenage paper 44*. *Design and optimization of irrigation distribution networks*. Rome: Food and Agriculture Organization of the United Nations.

Loganathan, G. V., Sherali, H. D., & Shah, M. P. (1990). A two-phase network design heuristic for minimum cost water distribution systems under a reliability constraint. *Engineering Optimization*, *15*, 311–336.

McCormick, G. P. (1976). Computability of global solutions to factorable nonconvex programs: Part I—Convex underestimating problems. *Mathematical Programming*, *10*, 147–175.

Sherali, H. D., & Smith, E. P. (1997). A global optimization approach to a water distribution network design problem. *Journal of Global Optimization*, *11*, 107–132.

Takahashi, H., & Matsuyama, A. (1980). An approximate solution for the Steiner problem in graphs. *Mathematica Japonica*, *24*(6), 573–577.

Zhang, J., & Zhu, D. (1996). A bilevel programming method for pipe network optimization. *SIAM Journal on Optimization*, *6*(3), 838–857.