2nd International Joint Conference on Water Distribution
Systems Analysis & Computing and Control in the Water Industry
Valencia (Spain), 18-22 July 2022
doi: https://doi.org/10.4995/WDSA-CCWI2022.2022.14823

REDUCTION OF THE SEARCH SPACE FOR THE OPTIMIZATION PROBLEM OF THE DESIGN OF THE PUMPING STATION THROUGH THE AUTOMATIC IDENTIFICATION OF INFEASIBLE FLOW DISTRIBUTIONS

Jimmy H. Gutiérrez-Bahamondes[1], Bastián Valdivia-Muñoz[2], Daniel Mora-Melia[3] and Pedro Iglesias- Rey[4]

[1,2]Doctorado en Sistemas de Ingeniería. Facultad de Ingeniería, Universidad de Talca, Curicó(Chile)
[3]Dep. de Ingeniería y Gestión de la Construcción. Facultad de Ingeniería, Universidad de Talca, Curicó (Chile).
[4]Dep. Ingeniería Hidráulica y Medio Ambiente. Universitat Politècnica de València, Valencia (Spain).

[1]jgutierrezb@utalca.cl, [2] bvaldivia18@alumnos.utalca.cl, [3]damora@utalca.cl, [4] piglesia@upv.es

Abstract

The pumping station design is a critical process in water distribution networks. This set of decisions will have an immediate impact on construction costs and determine energy consumption over the entire lifetime of the system. However, the minimization of investment and operational costs at the same time is a complex problem that has been approached from different perspectives.

To achieve this goal, in recent years, it has been shown that it is possible to optimize the selection of pumps, accessories, and control systems while optimizing the flow distribution provided by the pumping stations by using the setpoint curve. However, the huge number of possible combinations and the non-linearity of the equations rule out the use of exact methods to solve the proposed mathematical model. Despite this, some metaheuristic techniques, specifically population-based evolutionary algorithms, have shown good performance against case studies in networks with a high level of simplification. Each objective function evaluation involves at least one hydraulic simulation during the analysis periods. Therefore, the computational effort grows considerably as the size of the network increases, affecting the efficiency of these algorithms and limiting their use to smaller networks. Thus, optimization of the design of pumping stations in real-size networks is a problem that has not yet been fully resolved.

To reduce the number of evaluations of the objective function during the optimization process, this work presents a new method for the reduction of the search space based on the automatic identification of infeasible flow ranges as part of the network preprocessing. The method considers the maximum capacity of the available pumps, the minimum pressure required, and the demand patterns of the network. In this way, each pumping station has different restrictions for the decision variables of the mathematical model related to the flow contributions.

From this point on, the algorithm does not waste any computational effort evaluating solutions that represent flow distributions previously classified as infeasible. Therefore, it is possible to accelerate the convergence of the algorithms while preserving the quality of the solutions obtained. This new method can be applied to any direct injection network. The amount of solution space reduction will depend on the characteristics of each network. To clarify, this work includes the analysis of one case study and a genetic algorithm was implemented to resolve the model. Finally, the results show a reduction of the solutions space of 80% for the largest network presented.

Keywords
Pumping station design, setpoint curve, metaheuristic, search space.

1 INTRODUCTION

The design of pumping stations (PS) in water distribution networks is a complex problem [1], [2]. The work of these systems directly affects the quality of life of consumers. However, its construction implies a high cost, and its operation is carried out through a large energy consumption throughout the year. In this way, the design of PSs has short- and long-term consequences for the budget of any city [3]. On the one hand, in the immediate term, it will determine the investment costs necessary for the construction of the facilities. While the operating costs will also be affected throughout the life of the project [4]. This means that efficient design could significantly reduce the costs involved in the water supply. Especially in networks where there is not enough elevation to install tanks, or the network is fed directly from groundwater [5], [6].

During the design process, pump models, accessories, and control systems must be selected [7]. Later, through network operations, the control system determines the on/off status of each pump depending on demand [4]. Therefore, the decisions involved in the design must consider the operational conditions and, consequently, cannot be optimized independently [8].

Different approaches have been proposed to optimize the efficiency of the PS design [1], [6], [8]. The authors use different criteria, like location and minimizing the leakage [9], maximizing energy production [10], and minimizing maintenance and energy costs at the same time [4] among other criteria. However, in recent times it has been shown that the determination of the optimal flow distribution can be used as an effective tool in the design of PSs [11], [12]. In particular, the research carried out by the authors de [8] implements a novel methodology based on the minimization of operational and investment costs at the same time using the optimization of flow distribution. They propose a mathematical optimization model where the decision variables are the fractions of flow provided by each PS, the model of the pumps, and the number of fixed and variable speed pump included. And to solve the model they use a pseudo genetic algorithm proposed in [13]. However, this methodology has a problem; each evaluation of a solution involves going through all the nodes to verify compliance with the minimum pressure constraints. Therefore, the algorithm loses search capacity as the network size grows. Additionally, in the case of large networks, the calculation time grows significantly due to the scarcity of feasible solutions.

This work proposes a new method to speed up the heuristic search processes of evolutionary algorithms applied to the problem posed by [8] using automatic pre-processing of the analyzed network. This method allows an empirical, systematic, and exhaustive mapping of the relationships between the possible flow contributions of the pumping stations that are being designed heads and flows in all the nodes of the network. Thus, it is possible to significantly limit the flow ranges provided by each pumping station. Specifically, it is possible to find flow rate combinations that will always give rise to an infeasible solution. The application of the methodology presented here is aimed at optimizing large networks, however, it can also be used in small networks without the need for changes.

The remainder of the paper is organized as follows: Section 2 describes the proposed methodology. Then, the developed methodology is applied to a case study, and an optimization method is implemented. Next, Section 3 provides the results, and a discussion is detailed in Section 4. Finally, the conclusions of the research can be found in Section 5.

2 MATERIALS AND METHODS

This work proposes a new method to speed up the process of searching for solutions to the problem posed in [8]. The implementation of this procedure does not imply changes in the mathematical model proposed in that investigation. For a better understanding, this model is briefly presented in the following sections.

2.1 Mathematical model

Decision variables are the key aspect of this study. They conform the solution to the problem. First, x_{ij} defines the percentage of the flow supplied from PS i (PS_i) at each time step j. Second, m_i indicates the number of fixed speed pumps and third, b_i corresponds to the identifier of the pump model to be installed in PS_i The parameters N_t and N_{ps} represent the total number of time steps and the total number of PSs, respectively; Once these values are known, it is possible to calculate the maximum flow for each PS, the number of total pumps ($N_{B,i}$), the number of VSPs.

The optimization model minimizes the sum of capital (CAPEX) and operational (OPEX) costs at the same time. Equation 1 presents the total annualized cost of the project. Where Fa is the amortization factor applying an interest rate r during Np periods.

$$\mathrm{F} = F_a \cdot CAPEX + OPEX \tag{1}$$

$$F_a = \frac{r \cdot (1+r)^{Np}}{(1+r)^{Np-1}} \tag{2}$$

The CAPEX is calculated according to Equation 3.

$$C_{CAPEX} = \mathrm{N_B} \cdot \mathrm{C_{pump}} + \mathrm{n} \cdot \mathrm{C_{inv}} + \mathrm{C_{facility}} + \mathrm{C_{control}} \tag{3}$$

Where NB is the number of pumps, C_{pump} is the purchase cost of a pump, n is the number of frequency inverters, $C_{facility}$ represents the cost of accessories including pipes, and C control is the sum of a pressure transducer, flowmeter, and programmable logic controller.

$$\mathrm{OPEX} = \sum_{j=1}^{N_t} \left\{ \sum_{i=1}^{N_{ps}} \left[\left(\sum_{k=1}^{m_{i,j}} \frac{\gamma \cdot (H_{o,i} - A_i \cdot Q_{i,jk}^2)}{(E_i - F_i \cdot Q_{i,j,k})} + \sum_{k=1}^{n_{j,i}} \frac{\gamma \cdot (H_{o,i} \cdot \alpha_{i,j,k} - A_i \cdot Q_{i,jk}^2)}{\left(\frac{E_i}{\alpha_{i,j,k}} - \frac{F_i}{\alpha_{i,j,k}^2} \cdot Q_{i,j,k} \right)} \right) \cdot p_{i,j} \right] \Delta t_j \right\} \tag{4}$$

where for each PS_i, the parameters $H_{0,i}$, Ai, E_i and F_i are the characteristic coefficients of the pump head and the performance curve and are extracted from an existing database depending on the pump model; $Q_{i,j,k}$ represents the discharge of pump k during time step j in PS i; $p_{i,j}$ is the energy cost; ϒ is the specific gravity of water; Δt_j is the discretization interval of the optimization period; and the numbers of FSPs and VSPs running at time step j are represented by $m_{i,j}$ and $n_{i,j}$, respectively. These values depend on the selected pump model and the system selected to control the operation point.

The optimization model is restricted by continuity and momentum equations and by minimum head requirements in the demand nodes. Equations 5 y 6 guarantees that the total flow supplied by the PS is equal to the flow demand.

$$x_{i,j} \geq 0 \ \forall i,j \tag{5}$$

$$\sum_{i=1}^{N_{ps}} x_{i,j} = 1\ \forall j \tag{6}$$

2.2 The infeasibility problem

The decision variable x_{ij} determines the fraction of flow that PS_i contributes during period j. This variable can take ranges from 0 to 100 (expressed as a percentage). Where 0 means that the PS does not supply water in that period. So, on the contrary, a value of 100 indicates that it will be the only SP that operates in the period. So, there is a huge number of possible combinations, and many of them are infeasible solutions.

The main causes of infeasibility are listed below:

- The distribution of flow generates sectors of the network where it is not possible to reach the minimum required pressures.
- Some of the PSs must provide a pressure greater than the maximum head of the largest pump that exists in the available catalog.
- The sum of the flows supplied is greater than the demand.

Each evaluation of the objective function supposes an increase in the computational effort made by the algorithm. It is for this reason that this work proposes to analyze the network previously to establish minimum and maximum limits for the variable $x_{i,j}$. Unfortunately, the non-linearity of the relationships between the hydraulic variables does not allow these values to be fixed, but depends on the piezometric head of the main PS. Which supplies all the water that is not provided by the rest of the PSs. To simplify the analysis, in this paper it is assumed that there is at least one PS called PS1. Then, PS1 is always the main station. And therefore, the lower and upper limits are expressed as curves as a function of PS1.

2.3 Calculation of curve of minimum and maximum flows

The steps to find the minimum and maximum flow curves for PSs other than PS1 are described.

1) Determine the sum of the product of the base demands of the nodes according to the consumption pattern (Q_B), corresponding to the lowest consumption in all the periods of the analysis.
2) Determine the limit of the piezometric head of PS1. The minimum value is the initial piezometric head of the network (H_0). The maximum value will be the initial head plus the maximum pressure given by the pump catalog (H_{bMax}). Assign the head to the corresponding PS1 (H_{design}).
3) Select one of the pumping stations, and Q_{min} is assigned to it.
4) Set a distribution flow rate (Q_r), which corresponds to the difference between Q_B and Q_{min}.
5) Set a list of flow distribution Q_r with all possible combinations of the remaining pumping stations. These can range from 0% to 100%.
6) Initialize the hydraulic analysis, assigning and testing the possible combinations of Q_r at the remaining stations.

7) If the difference between P_{min} and $P_{Reached}$ is less than or equal to the allowed threshold (U) or the pressure of PS_i is greater than H_{Max}, then, the minimum flow rate that the station can provide is Q_{min}.

8) Select another PS in the network and go to step 3. If $Q_{mín}$ is greater than Q_B the analysis is finished. Else, update Q_{min}= Q_{min}+1 and go to step 4.

9) If there are no more pumping stations in the network, increase the height of PS1. The value of the increase will determine the number of points on the curve. A small increment will mean more computational effort while very large values can affect the performance of the algorithm.

10) If H_{design} is greater than H_{max} the analysis must finish, else go to step 2.

Subsequently, the analysis is carried out for the maximum flows in a similar way. First, replace in step 1 the period of lower demand for the greatest network demand. Second, in step 7, the pressure condition must be deleted.

All the points that are outside the calculated curves give rise to infeasible solutions, however, it is not possible to ensure that all the points between the curves are feasible. That is, there may still be combinations of flow distributions that can be classified according to section 2.2. These solutions must be discarded by the algorithm during the respective evolutionary process.

2.4 Decoding solution

Once the curves have been calculated they are used by the evolutionary algorithm in each evaluation of the objective function. In this new method, the value of x_{ij} represents a fraction of the difference between the highest value of the maximum flow curve and the lowest value of the minimum flow curve. Figure 1 shows how it is possible to determine the pressure of PS1 from the intersection of the flow fraction and the minimum curve. At this point, PS1 provides enough head to reach the minimum pressure in the entire network. On the left, the curves are bounded by the PS1 level and on the right, the curves extend until they reach the maximum height that a pump can provide in the catalog. Outside this range, it would be impossible to achieve a technically feasible solution.

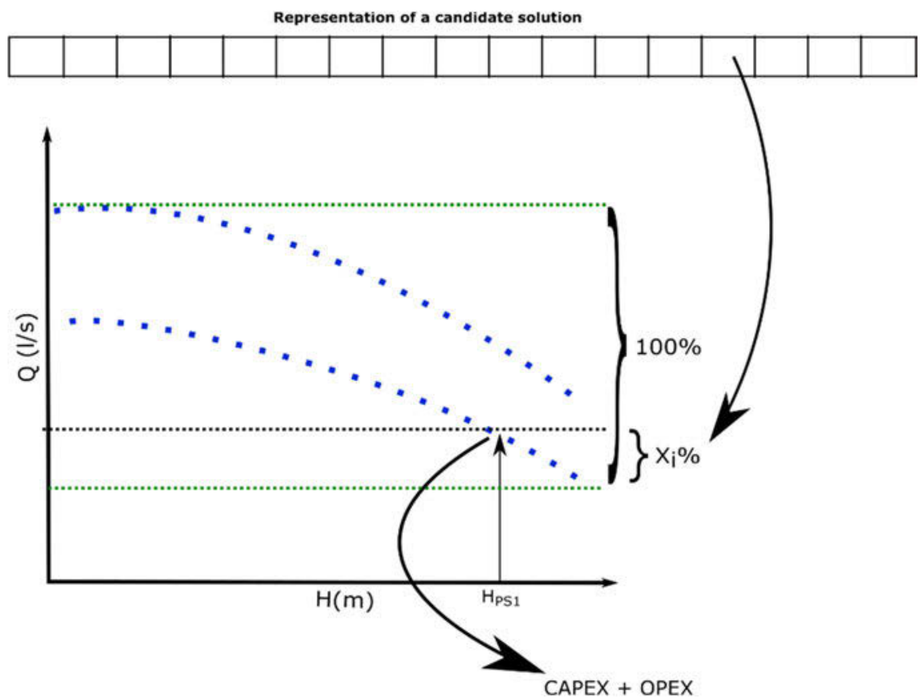

Figure 1: Decoding a solution

2.5 Case Study

To apply the methodology described above, one case study was conducted. Figure 2 shows a new WDN called Curicó Network.

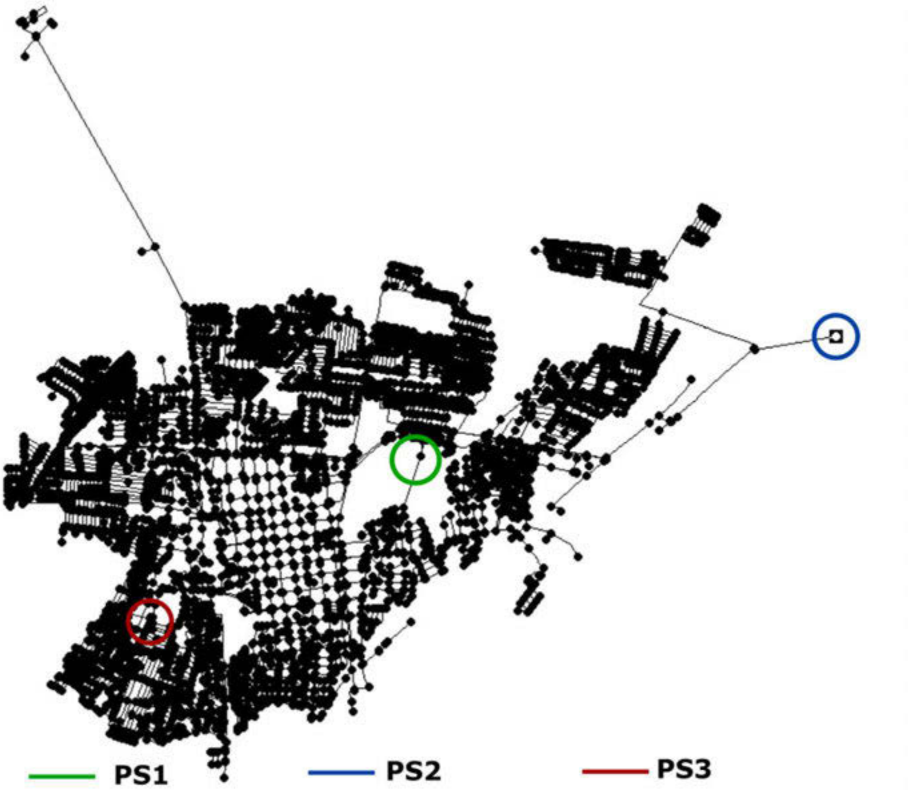

Figure 2: Red de Curicó

This network has 3 PSs (PS1, PS2 and PS3), 7630 consumption nodes, and 8359 pipes. A hydraulic analysis was carried out for one day, and a time step of one hour was considered. The demanded average flow rate is 172 L/s, the minimum pressure at the node is 15m and the roughness coefficient is 0.1. The pattern used to characterize time variation in demand is shown in Figure 3.

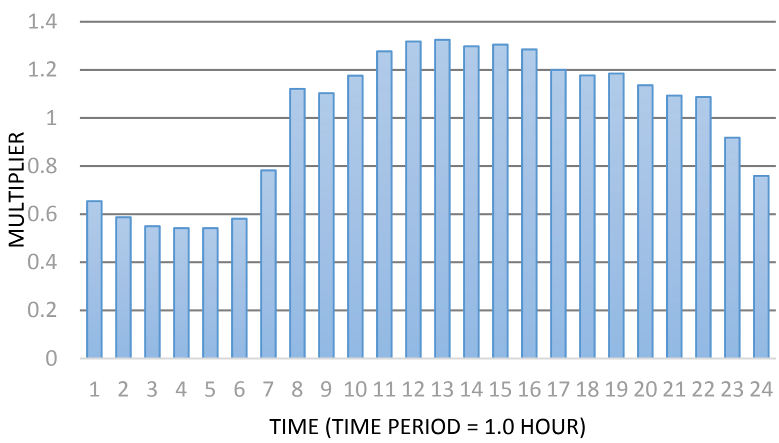

Figure 3: Demand pattern Curicó network

To calculate OPEX, Table 1 shows the hourly electricity rate used for each PS in the network. Por otro lado, todos los coeficientes necesarios para estimar CAPEX fueron obtenidos desde [8] . The maximum flow rate of the pumps in the database varied between 9 L/s and 50.7 L/s.

Table 1: Electricity for the case study (€/kWh)

Time(h)	PS1	PS2	PS3
1–8	0.094	0.092	0.09
9–18	0.133	0.131	0.129
19–22	0.166	0.164	0.162
23–24	0.133	0.131	0.129

To perform the optimization process, a database with 67 pump models was used. The annualized costs of these models were calculated using an interest rate of 5% per year and a projection time of 20 years. This led to an amortization factor Fa= 7.92%.

2.6 Optimization Method

The objective of this work is to show the advantages of the application of the incorporation of the calculation of the maximum and minimum curves as a support tool for evolutionary algorithms. For this comparison to be fair, the resolution of the proposed model was carried out using the same algorithm used in [8]. Specifically, it is the pseudo genetic algorithm (PGA) developed by the authors of [13]to solve problems of an integer nature. In this way, it is possible to compare directly with the methodology proposed in avoiding unnecessary biases. In addition, the same parameters of population size (P), crossover frequency (Pc), and mutation frequency (Pm) recommended by the authors of the previous investigation were considered.

The implementation was developed according to the instructions exposed in [14] This system can conduct massive simulations and is integrated with the hydraulic network solver EPANET using the programmer's toolkit [15].

3 RESULTS

After applying the methodology proposed in section 2.3, it was possible to calculate the minimum and maximum flow curves for the PSs of the Curicó network. Figures 4 and 5 show in red the area of infeasible solutions during the period of highest demand. And in blue, is the area where there is the possibility of finding feasible solutions. In both cases about 80% of the total solutions are infeasible. It is necessary to consider that PS1 does not appear in this analysis because it is considered by the methodology as the main station. That is, it provides all the flow that has not been supplied by the rest. In fact, PS1 is not a direct part of the representation of the solutions but corresponds to the complement of the work done by PS2 and PS3.

In the case of the PS3 analysis, if PS1 has a height greater than 312m, the minimum, and maximum curves are equal to zero. That is, from that limit PS3 cannot operate. Otherwise, the supply of water to the network will exceed the demand. Unlike PS3, the PS2 curves have a regular shape. Given the topographic conditions of the network, it is possible to appreciate that PS3 is very sensitive to the contributions of the rest of the stations.

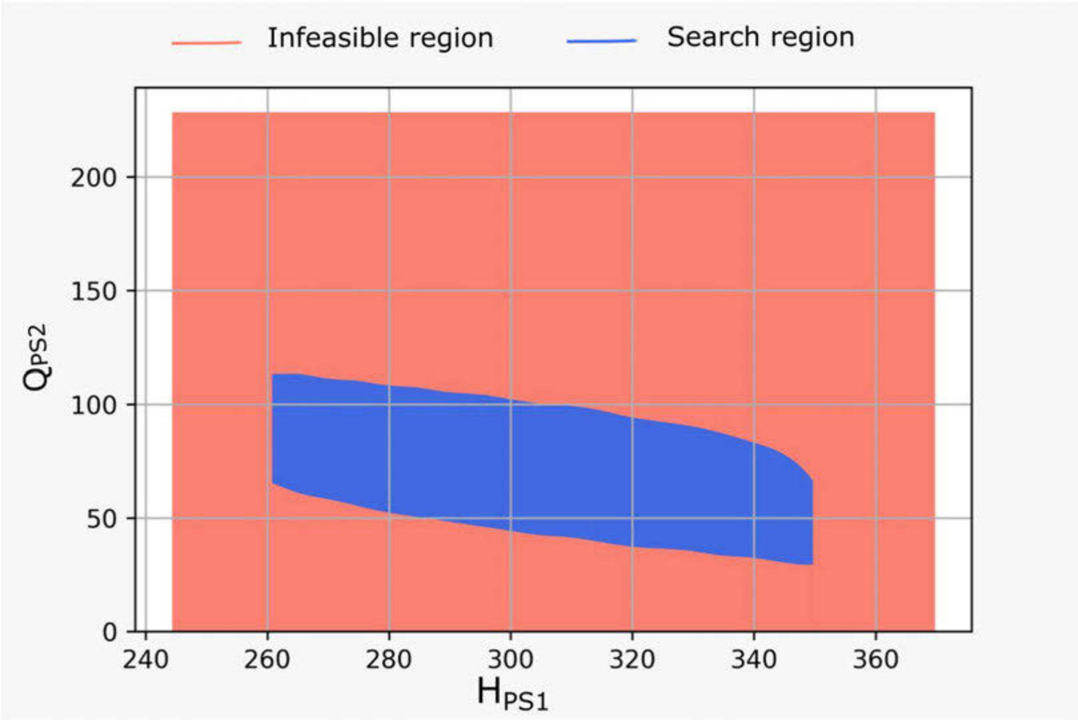

Figure 4a: Region infeasible vs search region for PS2

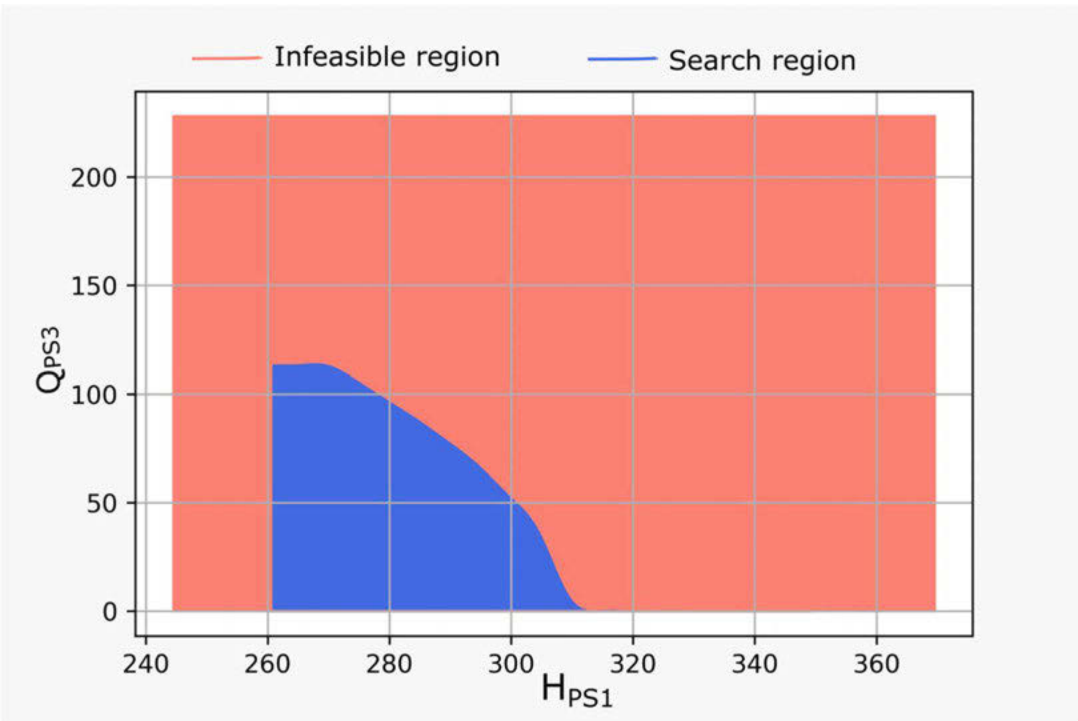

Figure 4b: Region infeasible vs search region for P3

Figure 5 shows the superposition of both curves, this analysis is important because it allows determining the influences that the behavior of one station has on the other. During the evaluation of the objective function, the PS are randomly ordered for the decoding of the solutions. In this way, it is possible to rule out infeasible solutions caused by the intersection of the two curves.

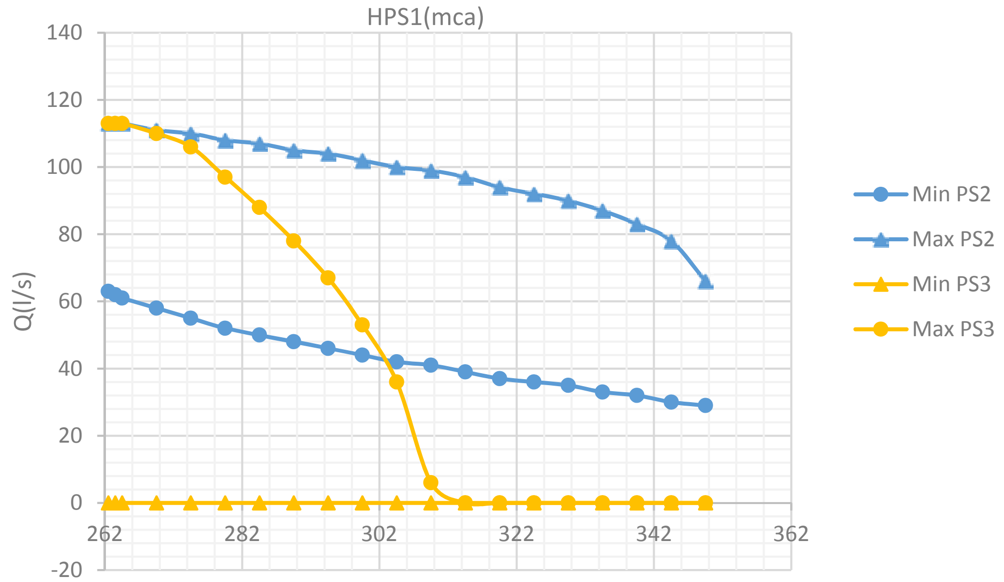

Figure 5: PS2 and PS3 curve overlay

Once the curves are calculated, the comparison of the PGA implemented in the methodology proposed in [8] vs the incorporation of the calculation to delimit the ranges of x_{ij}. Figure 6 shows the evolution of the values of the objective functions during 25,000 evaluations for the Curicó network problem.

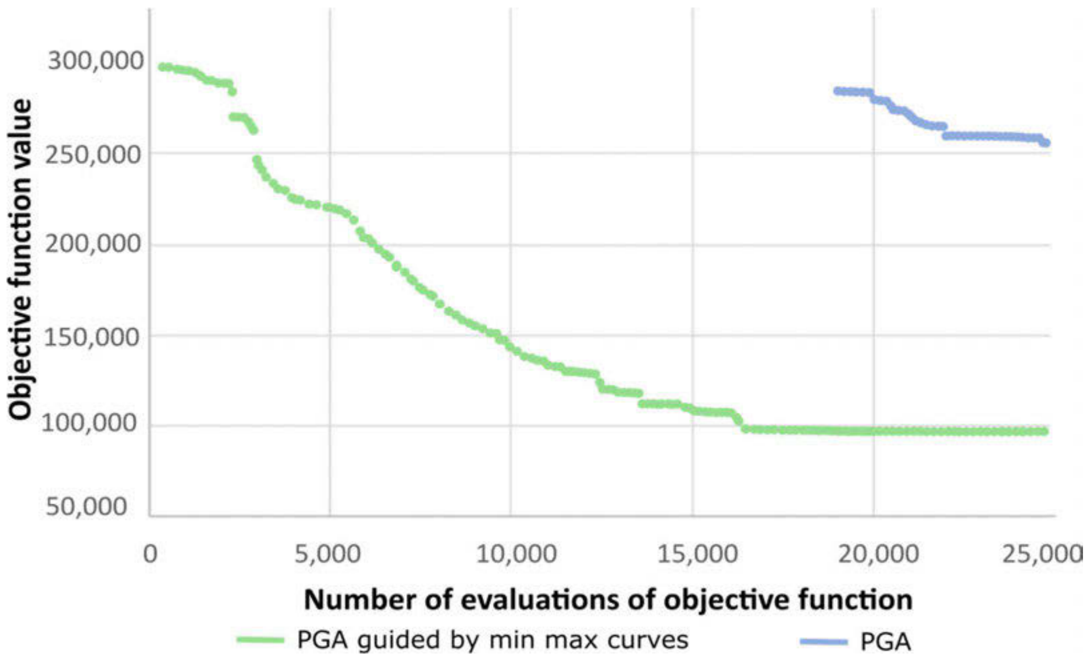

Figure 6: PGA Performance Comparison

4 DISCUSSION

The results of Figure 6 show a significant improvement in the performance of the implemented algorithm. The curve-guided algorithm finds feasible solutions very quickly and the nearly 16,000 evaluations of the objective function manage to converge. While the algorithm without the support of the minimum and maximum curves takes almost 18,000 evaluations alone to find a solution. Also, the convergence process is slower because it evaluates many infeasible solutions until a better solution is found.

On the other hand, in the work presented by [8] the calculation of the setpoint curve is used as a tool to determine the minimum energy point that must be provided by the pumps. This is an iterative procedure where each iteration of the loop must perform a computational simulation. On the other hand, in this work the setpoint curve is not used, instead, the curve of minimums calculated allows for approximate in a single iteration of the value of the height of the main PS. Therefore, this procedure is particularly advantageous in large networks.

In this case study, the results are significantly better. However, the new method presented supposes a computational effort before the implementation of an evolutionary algorithm. Therefore, it is necessary to add all the evaluations that were carried out to obtain the curves. Specifically, in the case study, these evaluations are approximately 10% of all the evaluations carried out. It is an iterative procedure that allows mapping infeasibility zones that are only performed once and can be used by any evolutionary algorithm to speed up the search for solutions. It is probable that in case studies where the number of nodes and pipes does not represent a problem, the implementation of this previous calculation will not be necessary. It could mean an increase in the number of computational simulations carried out.

5 CONCLUSIONS

The design of pumping stations in water distribution networks is a complex problem that can be optimized from different approaches. Decisions during design will affect investment costs and operational costs throughout the life of the project.

That is why the authors of [8] proposed a methodology to minimize investment and operational costs to it. And they showed that it is possible to completely design all the PSs in the network using the created model. However, this methodology is not very efficient compared to real-size networks where it is necessary to know the information of all the nodes in each evaluation of the objective function. To avoid this problem, a new methodology was proposed in this paper to speed up the search process for solutions of evolutionary algorithms compared to the mathematical model proposed above.

A systematic procedure was designed to map the infeasibility zones before running the algorithm. Then, the solutions are decoded using the result of the previous step, avoiding the evaluation of previously determined infeasible solutions. It is important to keep in mind that this procedure allows ruling out infeasibility zones, however, it does not ensure that all the solutions found within them are feasible.

The results show that the application of this new methodology can considerably reduce the number of evaluations of the objective function needed to find feasible solutions in a real size network. However, as discussed in the paper, the presented work has some limitations. In the case of small networks, the computational effort to find the infeasibility zones can generate an excess of evaluations of the objective function to achieve the same result.

Finally, experimentation with different types of networks, algorithms and parameters will be part of the future work that will allow determining the scope of the new proposal.

6 REFERENCES

[1] H. Mala-Jetmarova, N. Sultanova, and D. Savic, "Lost in optimisation of water distribution systems? a literature review of system design," Water (Basel), vol. 10, no. 3, Mar. 2018, doi: 10.3390/w10030307.

[2] T. M. Müller, P. Leise, I.-S. Lorenz, L. C. Altherr, and P. F. Pelz, "Optimization and validation of pumping system design and operation for water supply in high-rise buildings," Optimization and Engineering, vol. 22, no. 2, pp. 643–686, 2021, doi: 10.1007/s11081-020-09553-4.

[3] L. J. Blinco, A. R. Simpson, M. F. Lambert, and A. Marchi, "Comparison of pumping regimes for water distribution systems to minimize cost and greenhouse gases," Journal of Water Resources Planning And Management, vol. 142, no. 6, Jun. 2016, doi: 10.1061/(ASCE)WR.1943-5452.0000633.

[4] Y. Makaremi, A. Haghighi, and H. R. Ghafouri, "Optimization of pump scheduling program in water supply systems using a self-adaptive NSGA-II; a review of theory to real application," Water Resources Management, vol. 31, no. 4, pp. 1283–1304, Mar. 2017, doi: 10.1007/s11269-017-1577-x.

[5] H. Mala-Jetmarova, N. Sultanova, and D. Savic, "Lost in optimisation of water distribution systems? a literature review of system operation," Environmental Modelling & Software, vol. 93, pp. 209–254, Jul. 2017, doi: 10.1016/j.envsoft.2017.02.009.

[6] C. F. León-Celi, P. L. Iglesias-Rey, F. J. Martínez-Solano, and D. Savic, "Operation of multiple pumped-water sources with no storage," Journal of Water Resources Planning and Management, vol. 144, no. 9, p. 04018050, 2018, doi: 10.1061/(asce)wr.1943-5452.0000971.

[7] I. Pulido-Calvo and J. C. Gutiérrez-Estrada, "Selection and operation of pumping stations of water distribution systems," Environmental Research Journal, vol. 5, no. 3, pp. 1–20, 2011.

[8] J. H. Gutiérrez-Bahamondes, D. Mora-Meliá, P. L. Iglesias-Rey, F. J. Martínez-Solano, and Y. Salgueiro, "Pumping station design in water distribution networks considering the optimal flow distribution between sources and capital and operating costs," Water , vol. 13, no. 21, Nov. 2021, doi: 10.3390/w13213098.

[9] O. Fecarotta and A. McNabola, "Optimal location of pump as turbines (pats) in water distribution networks to recover energy and reduce leakage," Water Resources Management, vol. 31, no. 15, pp. 5043–5059, 2017, doi: 10.1007/s11269-017-1795-2.

[10] N. Nagkoulis and K. L. Katsifarakis, "Minimization of Total Pumping Cost from an Aquifer to a Water Tank, Via a Pipe Network," Water Resources Management, vol. 34, no. 13, pp. 4147–4162, Oct. 2020, doi: 10.1007/s11269-020-02661-x.

[11] C. Leon-Celi, P. L. Iglesias-Rey, and F. Javier Martinez-Solano, "Energy optimization of supplied flows from multiple pumping stations in water distributions networks," Procedia Engineering, vol. 186, pp. 93–100, 2017.

[12] C. Leon-Celi, P. L. Iglesias-Rey, F. Javier Martinez-Solano, and D. Mora-Melia, "A methodology for the optimization of flow rate injection to looped water distribution networks through multiple pumping stations," Water (Basel), vol. 8, no. 12, 2016, doi: 10.3390/w8120575.

[13] D. Mora-Melia, P. L. Iglesias-Rey, F. J. Martinez-Solano, and V. S. Fuertes-Miquel, "Design of water distribution networks using a pseudo-genetic algorithm and sensitivity of genetic operators," Water Resources Management, vol. 27, no. 12, pp. 4149–4162, Sep. 2013, doi: 10.1007/s11269-013-0400-6.

[14] J. H. Gutierrez-Bahamondes, Y. Salgueiro, S. A. Silva-Rubio, M. A. Alsina, D. Mora-Meliá, and V. S. Fuertes-Miquel, "jHawanet : An Open-Source Project for the," no. September, 2019.

[15] L. a. Rossman, "EPANET 2.0 User's manual (EPA/600/R-00/057)," 2000.