Multiple Linear Regression Analysis of Pumps Performance in Water Pumping Plants

Fouad LAAJINE\textsuperscript{1,2}, Mohammed MACHKOR\textsuperscript{3}, Driss MAZOUZI\textsuperscript{1}

\textsuperscript{1}R.N.E Laboratory, Multidisciplinary Faculty of Taza, University Sidi Mohamed Ben Abdellah, Fez, Morocco
\textsuperscript{2}National Office of Electricity and Drinking Water Morocco

Correspondence: Driss MAZOUZI (driss.mazouzi@usmba.ac.ma)

Abstract. Energy use in water supply systems represents a consequent part of global energy consumption across all sectors. This consumption is expected to rise, due to the increasing demand and the recourse to unconventional water resources. Regarding water utilities, most of their operating costs are related to energy consumption, especially pumping systems consumption. In this context, the main objective of this study was to model accurately the energy consumption of pumping systems in order to optimize the whole water supply system, thus improving its efficiency, especially in the case of a limited renovation. For this purpose, Multiple Linear Regression was fitted to model the produced kWh/m\textsuperscript{3} ratio according to the following parameters, active and reactive energies, the daily produced water volume, the power factor (Cos\(\phi\)), and the operating time of each pump. The final model describes accurately the consumption per cubic meter produced with R-square statistic reaching 0.91 and value standard error is close to 5\% were found. Therefore, this model could be considered a good estimator for the calculated ratio, which was close to the experimental one. In addition, this approach considers the system's behavior while most of the comparable studies focus on the pump scheduling problem estimator for the calculated ratio which was close to the experimental one. In addition, this approach considers the real-time-data behavior while most of the comparable studies focus on the pump scheduling problem.

Key Words: Energy efficiency, multiple linear regression, pumping systems, water supply system.
1. Introduction

In the upcoming decades, water demand is forecasted to significantly increase up to 35% in 2050 (UN-Water 2020). Therefore, energy needs that are related to the water pumping and distribution processes could contribute to the expected increase in energy demand (EIA 2019). Indeed, several studies have concluded that the energy consumption related to these processes accounts for 7% of the total energy used across the globe (Coelho et al., 2014). His share is expected to get bigger due to the increasing distances between resources and populations, especially in water-scarce countries, and the growing consumption per capita due to the industrialization and improvement of living standards.

To reach the sixth Sustainable Development Goal (SDG6) that aims to generalize the access to drinking water supply, the water production cost mustn't impact its price which should stay affordable to the population. Therefore, production costs should be reduced via the control of energy consumption which is typically the largest marginal cost for the production (Helena et al., 2017).

Pumping processes consume the largest fraction of total energy (Plappally and Lienhard, 2012). The pumps consumption often presents 80% to 90% of the total energy consumption (Sarbu, 2016). However, this consumption may depend on many factors such as surface water or ground water, transport differences, flat or mountain regions, etc (Rothausen and Conway, 2011) (Plappally and Lienhard, 2012). Thus, by achieving energy efficiency improvements measures, we may reduce this consumption by 25% (Moreira, et al., 2013). In this context, few studies were interested in modeling pumping systems and evaluated the influence of parameters such as the aging of the components, which can reduce the performance of the pump by up to 12% (Durmus, et al., 2008).

The improvement of the efficiency of the pumping has always been an important part. Since many specific factors impact the overall energy efficiency of a water pumping system such as surface water or ground water, transport differences, flat or mountain region... One important factor is the operating point which should be as close as possible to the maximum energy efficiency region (Carravetta, et al.,
Another factor is the age of the system, especially the pumps which can lose 1.5% of their efficiency every year (Zhou et al., 2021). Many attempts have been made to address the subject of pumping optimization using different approaches to optimize either energy consumption or the overall cost of energy. To optimize the energy consumption, a model based on ant colony was utilized (Ostfeld and Tabultzev, 2008) to provide an effective model that minimizes the design and the operating costs. This model acts on parameters such as: the diameters of the pipes, the pumping station’s maximum power, and the tank's storage. Another work targeting the minimization of energy consumption (Giustolisi O et al., 2012) used a multiobjective strategy to minimize losses by controlling pressure in the water distribution system and thus minimizing the energy consumption.

On the other hand, various studies have focused on optimizing the operation of the existing water systems by using deterministic methods such as Dynamic Programming (Fallside and Perry, 1975) which worked on minimizing the energy cost by load shifting using a hierarchical optimization to reduce the problem to 10 control parameters. Given the difficulty of the deterministic approach, different metaheuristic methods were developed, the widely used among them is Genetic Algorithms (Wu et al., 2014) which used evolutionary computation to maximize the efficiency of a system consisting of 2 pumps. Although these efforts employ evolutionary optimization techniques. However, in literature, particular attention has been paid to the applicability of linear regression (Puleo et al., 2014) and multiple linear regression (Adamowski et al., 2012) to the pumping scheduling optimization problem.

This work introduces a method of using multiple regression method to analyze the key factors affecting the efficiency of the pumping for drinking water production. Hence, this paper presents the results targets the pumping system of the Bab Louta drinking water production located in the province of Taza, Morocco. In this perspective, a multiple linear regression (MLR) was fitted to model the produced kWh/m³ ratio according to the input parameters by using real-Time-data.
such as the active energy and reactive energy consumed by the pumping station, the
daily produced volume, the power factor and the pump operating time.

2. Materials and methods

Conventional water supply systems (WSS) consist of sets of structures and facilities
providing products with a suitable quantity and quality for domestic and industrial
use (Luna et al., 2019). To develop energy and hydraulic model, WSS systems
should be evaluated in terms of mass and energy (Figure 1).

There are various methods to enhance energy efficiency in WSS. These methods
are ranging from leakages control through simple monitoring operations to massive
investments by reviewing the design of the infrastructures, equipment upgrade,
pump system optimization, and real-time control. These Methods were classified
into 3 major sub-categories is shown in Figure 2 (Shankar, et al., 2016). These
factors affecting system efficiency, and formulate targeted improvement plans.
Figure 2: Scheme of the main of factors affecting efficiency enhancement opportunities in WSS.

2.1 Study area and description of the drinking water production system

The province of Taza is located in the center North of Morocco and it is one of the 9 provinces of the region of Fez-Meknes with a population of roughly 530 000 inhabitants (Morocco 2014). The water treatment plant of Tahla provides a large population of the province, mainly the urban areas. It is situated 60 km from the city of Taza. The plant treats the raw water of the Bab Louta reservoir.

The production system consists of a pumping station SP0 for raw water, a water treatment plant then another pumping station of SP3 to reach the city of Taza, which is the chief town of the province.

Figure 3: Location of the Bab Louta reservoir.
2.2 Description of the pumping station

The pumping station SP0 is located about 3km from the reservoir. It is responsible for overcoming the difference in altitude between the raw water intake and the water treatment plant. The water is taken by gravitation from the reservoir to the pumping station SP0 through a pipe of a nominal diameter (ND) of 700 mm which converge to a diameter of 500 mm just before reaching the SP0. Then, the water is pumped through a pipe of a nominal diameter of 600 mm to a tank destined to provide the water treatment plant with raw water, it is called RMC0 and having a capacity of 1000 m³, which provides the water treatment plant with raw water through a pipe of a nominal diameter of 500 mm (ONEE 2019).

Figure 4: A simplified example of a water treatment plant with water source (Bab Louta reservoir), pumping station (SP0) and one water tank (RMC0).

The pumping station, the object of the current study, is situated at an altitude of 498.20 m and tank RMC0 is situated at an altitude of 737.80 m. SP0 consists of four pumps (Figure 5) with a flow of 457 m³/h each and a manometric pressure of 209 m. They are manufactured by the company HIDROTECAR powered by an electric motor which is manufactured by LEROY SOMER with a nominal power of 455 kW (ONEE, 2019).
**Figure 5:** Hydromechanical scheme of the pumping station SP0.

### 2.3 Experimental procedure

To obtain a reliable model of our system, measurements of the key parameters influencing this system were conducted. These measurements were taken in the period between January 2015 and December 2018 daily for the parameters. Based on the above analysis, a total of eight factors affecting the efficiency of the pumping for drinking water production are selected, namely:

- **P**: the active energy consumed by the pumping station (measured by a wattmeter),
- **Q**: the reactive energy consumed by the pumping station (measured by a wattmeter),
- **V**: the daily produced volume (measured by an electromagnetic flowmeter),
- **Cosφ**: the power factor,
- **HMG1**: the pump operating time “1”,
- **HMG2**: the pump operating time “2”,
- **HMG3**: the pump operating time “3”,
- **HMG4**: the pump operating time “4”.

This study investigates the impact of these several independents variables on the energy use of a drinking water production. The starting from these parameters, the principal component analysis is performed, and the comprehensive evaluation function equation and the regression equation can be obtained, respective.

2.4 Modelling

This study aimed to use Multiple Linear Regression (MLR) analysis to predict an output from a range of inputs. MLR model with multiple input variables can be expressed as follows (Longo et al., 2016):

\[
Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon
\]  

Eq 1

Where, \( Y \) is the output variable, \( \beta_i \) are the regression parameters, \( X_i \) are the input variables \((i = 1, \ldots , n)\) and \( \varepsilon \) is the random error.

In order to assess the influence of the included parameters on the cubic meters ratio produced by the pumping station, the following parameters were considered: the active energy \( (E_P) \), the reactive energy \( (E_Q) \), the daily produced volume \( (V) \), the power factor \( (\text{Cos} \phi) \), and the operating time of each pump \( (\text{HMG}_i) \). Therefore, the principal parameter analysis is used to establish the evaluation model to achieve more objective and accurate analysis.

The effect of eight variables on the produced Kwh/m³ ratio was evaluated. Of note, 1388 experiments were conducted during 4 years. The set of analysis data is shown in Table 1.

**Table 1:** Problem characteristics

<table>
<thead>
<tr>
<th>Objective of the study</th>
<th>The effect of the variables on the ratio KWh/m³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Variables</td>
<td>8</td>
</tr>
<tr>
<td>Number of experiments</td>
<td>1388</td>
</tr>
<tr>
<td>Number of the coefficients</td>
<td>8</td>
</tr>
<tr>
<td>Number of responses</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2: Measurements summary of the mean and standard deviation during period 2015 and 2018

<table>
<thead>
<tr>
<th></th>
<th>$E_P$ (kWh)</th>
<th>$E_Q$ (kVar)</th>
<th>Cos$\phi$</th>
<th>Production (m$^3$)</th>
<th>HMG1 (m)</th>
<th>HMG2 (m)</th>
<th>HMG3 (m)</th>
<th>HMG4 (m)</th>
<th>Ratio (kWh/m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>14 837.47</td>
<td>8 339.58</td>
<td>0.87</td>
<td>18 181.40</td>
<td>7.95</td>
<td>7.63</td>
<td>10.43</td>
<td>9.01</td>
<td>0.82</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>9 158.63</td>
<td>8 529.85</td>
<td>0.05</td>
<td>5 081.13</td>
<td>8.90</td>
<td>5.88</td>
<td>11.51</td>
<td>8.85</td>
<td>0.43</td>
</tr>
</tbody>
</table>

The descriptive statistics of the pumping station parameters under study are given in Table 2. It provides a summary of the mean and standard deviation values of eight measured parameters during 4 years. In this case, generally for each year, the mean direct production about 18181 m$^3$ related active energy of 14837 kWh with the operating time of each pump (HMG$_i$) to range between 8 and 10 m. Figure 6 displays the real water consumption versus the month and year that is measured over four years.

Figure 6: Representation the dataset collected of the variation of the consumptions of water versus the months during period 2015 and 2018, (A) consumption variation through the months (*: P<0.05, consumption variation through the years; #: P<0.05, consumption variation through the months), (B) consumption variation through the years (P<0.05 is considered as statistically significant. Each point represents the consumption average of a given month during the year).
The dataset collected of the water production across the year for each year from 2015 to 2018 presented in figure 6-A, and it the trend shows that clearly demonstrates a higher production which reflects a higher consumption of water during the summer months in province of Taza. On the other hand, figure 6-B shows the evolution of the production through the years 2015 to 2018 and it has also the same increasing trend due to the continuous commissioning of new networks leading to a growing number of consumers.

It can be seen from Table 1, 2 and Figure 6 that the original data has large differences and many influencing factors, and it is difficult to conduct comprehensive and systematic analysis by conventional methods. Therefore, the MLR analysis method is used to establish the evaluation model to achieve more objective and accurate analysis.

3. Results and discussion

3.1 Statistical interpretation

The correlation coefficients shown in a matrix (Table 3) are the results of statistical analyses for possible relationships between different parameters monitored. It was found that the studied variables are strongly correlated:

- The active energy consumed by the pumping station was dependent on the production, reactive energy, pumps operating hours (4;1), power factor (Cosφ), HMG2, and HMG3 respectively.
- The reactive energy consumed is highly was dependents on the production, active energy, Cosφ, pumps operating hours (4;1;2;3) respectively.
- Cosφ was dependent on the reactive energy, the production, active energy, pumps operating hours (4;1;3;2) respectively.
- Production was dependent on the active energy, the reactive energy pumps operating hours (4;1;2;3).
Table 3: Correlation matrix of water pumping station parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Means</th>
<th>Std. Dev.</th>
<th>P</th>
<th>Q</th>
<th>Cosφ</th>
<th>Prod</th>
<th>HMG1</th>
<th>HMG2</th>
<th>HMG3</th>
<th>HMG4</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>-0.022423</td>
<td>0.487229</td>
<td>1.000000</td>
<td>0.883071</td>
<td>-0.190305</td>
<td>0.939746</td>
<td>0.333340</td>
<td>0.202424</td>
<td>0.169873</td>
<td>0.625872</td>
<td>0.463030</td>
</tr>
<tr>
<td>Q</td>
<td>-0.027636</td>
<td>0.345041</td>
<td>0.883071</td>
<td>1.000000</td>
<td>-0.585775</td>
<td>0.883506</td>
<td>0.332525</td>
<td>0.170169</td>
<td>0.170505</td>
<td>0.612964</td>
<td>0.298339</td>
</tr>
<tr>
<td>Cosφ</td>
<td>0.026763</td>
<td>0.761882</td>
<td>-0.190305</td>
<td>-0.585775</td>
<td>1.000000</td>
<td>0.316528</td>
<td>-0.197066</td>
<td>0.024943</td>
<td>-0.076562</td>
<td>-0.271961</td>
<td>0.193819</td>
</tr>
<tr>
<td>prod</td>
<td>0.000741</td>
<td>0.973866</td>
<td>0.939746</td>
<td>0.883506</td>
<td>-0.316528</td>
<td>1.000000</td>
<td>0.340309</td>
<td>0.222533</td>
<td>0.175107</td>
<td>0.613597</td>
<td>0.150594</td>
</tr>
<tr>
<td>HMG1</td>
<td>-0.001400</td>
<td>1.002171</td>
<td>0.333340</td>
<td>0.332525</td>
<td>-0.197066</td>
<td>0.340309</td>
<td>1.000000</td>
<td>-0.164919</td>
<td>-0.016355</td>
<td>0.236341</td>
<td>0.111733</td>
</tr>
<tr>
<td>HMG2</td>
<td>-0.001865</td>
<td>0.998758</td>
<td>0.202424</td>
<td>0.170169</td>
<td>0.024943</td>
<td>0.222533</td>
<td>-0.164919</td>
<td>1.000000</td>
<td>-0.086627</td>
<td>-0.141591</td>
<td>0.006575</td>
</tr>
<tr>
<td>HMG3</td>
<td>0.001362</td>
<td>1.003257</td>
<td>0.169873</td>
<td>0.170505</td>
<td>-0.076562</td>
<td>0.175107</td>
<td>-0.016355</td>
<td>-0.086627</td>
<td>1.000000</td>
<td>0.074751</td>
<td>0.036371</td>
</tr>
<tr>
<td>HMG4</td>
<td>-0.017745</td>
<td>0.708400</td>
<td>0.625872</td>
<td>0.612964</td>
<td>-0.271961</td>
<td>0.613597</td>
<td>0.236341</td>
<td>-0.141591</td>
<td>0.074751</td>
<td>1.000000</td>
<td>0.0257315</td>
</tr>
<tr>
<td>Ratio</td>
<td>-0.051987</td>
<td>0.197427</td>
<td>0.463030</td>
<td>0.298339</td>
<td>0.193819</td>
<td>0.150594</td>
<td>0.111733</td>
<td>0.006575</td>
<td>0.036371</td>
<td>0.0257315</td>
<td>1.000000</td>
</tr>
</tbody>
</table>
Therefore, it is necessary to comprehensively consider the efficiency of the pumping unit, summarized in Table 4. Table 4 lists the standardized regression coefficients and regression coefficients of various factors on system efficiency. The standardized regression coefficients can be used to construct the multiple Regression equation. ‘T-test’ shows that the probability of rejection of truth for all factors except the constant term is less than 0.05, indicating that the above regression factors have significant effects on the dependent variables. Moreover, it was concluded that the factors influencing the kwh/m³ ratio in descending order are:

- Ratio is positively correlated with the active energy consumed by the pumps;
- Ratio is negatively correlated with the production;
- Ratio is positively correlated with Cosφ;
- Ratio is negatively correlated with the reactive energy consumed by the pumps;
- Ratio is positively correlated with the operating hours of pumps 1 and 4.

### Table 4: Regression summary for dependent variable

<table>
<thead>
<tr>
<th></th>
<th>Standardized regression coefficient (b*)</th>
<th>Standard error of b*</th>
<th>Regression coefficient (b)</th>
<th>Standard error of b</th>
<th>T-test (1379)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.026073</td>
<td>0.001608</td>
<td>-0.026073</td>
<td>0.001608</td>
<td>-16.2112</td>
</tr>
<tr>
<td>Prod</td>
<td>-2.52837</td>
<td>0.026164</td>
<td>-0.512563</td>
<td>0.005304</td>
<td>-96.6366</td>
</tr>
<tr>
<td>HMG₁</td>
<td>0.02116</td>
<td>0.008985</td>
<td>0.004168</td>
<td>0.001770</td>
<td>2.3546</td>
</tr>
<tr>
<td>HMG₂</td>
<td>0.01693</td>
<td>0.009327</td>
<td>0.003346</td>
<td>0.001844</td>
<td>1.8150</td>
</tr>
<tr>
<td>HMG₃</td>
<td>-0.00080</td>
<td>0.008299</td>
<td>-0.000157</td>
<td>0.001633</td>
<td>-0.0960</td>
</tr>
<tr>
<td>HMG₄</td>
<td>0.04287</td>
<td>0.011275</td>
<td>0.011948</td>
<td>0.003142</td>
<td>3.8023</td>
</tr>
<tr>
<td>P</td>
<td>2.85701</td>
<td>0.043662</td>
<td>1.157669</td>
<td>0.017692</td>
<td>65.4347</td>
</tr>
<tr>
<td>Q</td>
<td>-0.08315</td>
<td>0.040932</td>
<td>-0.047577</td>
<td>0.023421</td>
<td>-2.0314</td>
</tr>
<tr>
<td>Cosφ</td>
<td>0.09614</td>
<td>0.020445</td>
<td>-0.024913</td>
<td>0.005298</td>
<td>-4.7023</td>
</tr>
</tbody>
</table>
The variance analysis of data in Table 5 has shown that eight variables are significant which are Prod, HMG1, HMG2, HMG3, HMG4, P, Q and Cosφ. It can be observed that the regression mean squares was superior to the residual mean squares, the Fischer value is very high and the p-value is strictly tended towards zero. Hence, the variance analysis result so all meet the regression requirements.

Besides as shown in Table 6 that, the R-square and Adjusted R-square statistic of this model reaching 0.91 were found, which can reflect the good degree of influence of various factors on model. From the error analysis, it can be seen that the value standard error estimate is close to 0.05 (5%), thus error value is within a reasonable interval, and the regression equation is statistically significant. In view of these results, we can say that the parameters studied are highly variable in the research area, thus, they also confirm the performance of the developed model.

### Table 5: Analysis of Variance for MLR Model

<table>
<thead>
<tr>
<th>Effect</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Mean Squares</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>49.31283</td>
<td>8</td>
<td>6.164104</td>
<td>1790.044</td>
<td>0.00</td>
</tr>
<tr>
<td>Residual</td>
<td>4.74865</td>
<td>1379</td>
<td>0.003444</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>54.06149</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 6: Statistical evaluation of model

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>0.95507172</td>
</tr>
<tr>
<td>R²</td>
<td>0.91216198</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.9116524</td>
</tr>
<tr>
<td>Fischer (8.1379)</td>
<td>1790</td>
</tr>
<tr>
<td>p-value</td>
<td>0.00</td>
</tr>
<tr>
<td>Standard error of estimate</td>
<td>0.0586817579</td>
</tr>
</tbody>
</table>

Based on this presented Statistical data analysis, the following set of the parameters in the standardized regression equation are the data after standardized processing is proposed in the final model expressed by the equation below:

\[
R_{ratio} = 2.85701 P - 2.52837 \text{Prod} - 0.08315 Q + 0.09614 \text{Cosφ} + 0.02116 \text{HMG1} + 0.01693 \text{HMG2} - 0.00080 \text{HMG3} + 0.04287 \text{HMG4}. \quad \text{Eq 2}
\]
Where, \( P \)-active energy consumed by the pumping station; Prod-production volume;
\( Q \)-reactive energy consumed by the pumping station; \( \cos \varphi \)-power factor; HMG-
pump operating time (\( i=1;2;3,4 \)).

The regression equation obtained indicating that the above regression factors have
significant effects on the dependent variables.

3.2 Technical Interpretations

From this analysis, this study concluded that in order to improve the ratio per
produced cubic meter, a reduction of the active energy consumed by the pumps
should be considered. As the function point of the pumps is already set by the
system characteristics, only active energy could be reduced by the improvement of
the pumps efficiency.

The ratio is negatively correlated with production means; i.e., there is an
economy of scale. It means the most the production increases the least the ratio is.

The operating hours of the pumps 1 and 4 are positively correlated, which means
that more these pumps are used higher more the ratio is higher. Therefore, it is
recommended to use the pump 3, and if there is an operation of renovation of the
pumping station, it is recommended to start with upgrading the pumps 1 and 4. In
the event of a new investment, the company can increase the capacity of the RMC0
storage tank which, according to the model, will decrease the significantly the ration
and also allows a load shifting to the off-peak hours.

A regression analysis was performed, which identified the quantitative
relationships between input parameters and energy consumption ratio per cubic
meter water, involved in a water pumping station. Using the variance analysis, the
degree of confidence for the achieved results by the regression equation was
determined, and the suitability of this equation at every point of the experimental
field.

The model which was elaborated in this study, was successfully validated in the
statistical analysis. It shows that the R-square statistic reaching 0.91, and a standard
error of estimate of 0.05. Thus, due to the lack of previous studies using multiple linear regression, we compared the results with a previous study involving five data-mining approaches (Kusiak et al., 2013).

This study had for objective to model the energy consumption in a comparable case of a wastewater pumping station that has 3 pumps that transfer the wastewater to a treatment plant. Although there are differences between the flow capacities and the pressure with drinking water supply facilities but the approach remains the same. The five data mining approaches are the multi-layer, perceptron, neural network (MLP), the boosted-tree (regression) algorithm (BT), the random-forest algorithm (RF), the support-vector machine (SVM), and the k-nearest neighbor algorithm. These approaches had all provided more than 90% of accuracy which is the case in the model of this study. The benefit of our method goes beyond the control methods used in most of the optimization approaches which only provide a method to operate the system in an efficient way but don’t account for other factors such as the aging of the pumps, factors that are crucial when upgrading the system.

4. Conclusions

A Linear multiple regression was conducted to assess and study the influence of multiple parameters on the energy consumption ratio per cubic meter water, involved in a water pumping station.

This unique approach has allowed evaluating the real response of the system relying on data that is measured over a 4 years period. Modelling the ratio will be a tool to take decisions on which pump should the work be done first. This method combined with a cash flow analysis, can help to take decisions on establishing priorities in case of renovations, to change the pumps 1 and 4 with more efficient pumps. To validate this model, we performed the performance test by determining the correlation R to show the link between the produced kWh/m$^3$ ratio and the following parameters such as active and reactive energies, the daily produced water volume, the power factor (Cos$\phi$), and the operating time of each pump. the
regression coefficients, thus validating the models. The final model describes accurately the consumption per cubic meter produced with a R-square statistic reaching 0.91.

After this study, we retain that the developed model can predict the energy consumption ratio per cubic meter water, involved in a water pumping station. Thus, the model would be useful when the next renovation will be undertaken by the office which will conduct a replacement of the pumps in the year 2024, can more accurately and reasonably evaluate the efficiency pumping, according to the pumping unit model, motor power…

Besides that, the above findings demonstrate the potential of method for solving real-time pump scheduling problems in large water distribution systems with many pumps. However, this requires further work with other metaheuristic methods such as Genetic Algorithms before relevant conclusion can be made.

Data availability. The source data used for the illustrations of the cases are available upon request.

Conflict of interest. All authors declare no conflicts of interest in this paper.
References


