

Article

Optimization Methodology for Estimating Pump Curves Using SCADA Data

Elad Salomons ¹, Uri Shamir ² and Mashor Housh ^{1,*}¹ Department of Natural Resources and Environmental Management, University of Haifa, Haifa 3498838, Israel; selad@optiwater.com² Civil and Environmental Engineering, Technion, Haifa 32000003, Israel; shamir@technion.ac.il

* Correspondence: mhoush@univ.haifa.ac.il

Abstract: Water distribution systems (WDSs) deliver water from sources to consumers. These systems are made of hydraulic elements such as reservoirs, tanks, pipes, valves, and pumps. A pump is characterized by curves which define the relationship of the pump's head gain and efficiency with its flow. For a new pump, the curves are provided by the manufacturer. However, due to its operating history, the performance of a pump deteriorates, and its curves decline at an estimated rate of about 1% per year. Pump curves are key elements for planning and management of WDSs and for monitoring system efficiency, to determine when a pump should be rehabilitated or replaced. In practice, determining pump curves is done by field tests, which are conducted every few years. This leaves the pump's performance unmonitored for long time periods. Moreover, these tests often cover only a small range of the curves. This study demonstrates that in the era of IoT and big data, the data collected by Supervisory Control And Data Acquisition (SCADA) systems can be used to continuously monitor pumps' performance and derive updated pump characteristic curves. We present and demonstrate a practical methodology to estimate fixed and variable speed pump curves in pumping stations. The proposed method can estimate individual pump curves even when the measurements are given only for the pumping station as a whole (i.e., total flow, pumping station head gain). The methodology is demonstrated in a real-world case study of a pumping station in southern Israel.

Keywords: pump curves; water distribution systems modeling; pump monitoring by SCADA



Citation: Salomons, E.; Shamir, U.; Housh, M. Optimization Methodology for Estimating Pump Curves Using SCADA Data. *Water* **2021**, *13*, 586. <https://doi.org/10.3390/w13050586>

Academic Editor: Joaquim Sousa

Received: 25 January 2021

Accepted: 19 February 2021

Published: 24 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The flow of water requires energy. In nature, rivers flow from the mountains to the sea due to gravitational energy. Similarly, in some water distribution systems (WDSs), the flow of water in pipes is also governed by gravity. However, to supply water to higher areas and to overcome the energy losses in the water delivery system, additional energy must be added. This is the role of pumps, which convert electrical power to mechanical power, and then into hydraulic energy (head). Energy consumption for pumping water in distribution systems constitutes an important element of the overall energy budget. For example, the energy consumed for treating and pumping of water and wastewater in the United States in 2010 was estimated to be 4% of the total energy consumption [1,2]. In a survey conducted by Lam et al. [3], it was found that the energy used for water provision in 30 cities over 15 years, is up to 1 kWh per m³ which amounts to an average annual per-capita energy use of 100 kWh. Pump performance and efficiency deteriorate over time [4]. It has been estimated that the energy use of a potable water pump will increase by about one percent per year [5]. The reasons and rates of performance degradation over time have been analyzed by Eaton et al. [6]. They conclude that the head at constant flow declines in a non-linear fashion, reaching 1% after two years of service, and growing to 10% after 9 years of service. They recommend that pump performance should be tracked over time but also point to the practical difficulties of doing this under field conditions.

Thus, monitoring the performance and condition of pumps is important. The operation of a pump is described by its characteristic curves that show the pump's head gain, power, and efficiency over a range of flow rates. For new pumps, these curves are supplied, together with the pump, by the pump's manufacturer. For pumps already installed and in operation, it is customary to perform field tests to evaluate their current performance curves [7]. For example, Israeli regulations require that a certified pump test should be performed once every 30 months or after 7500 h of operation (whichever comes later) [8]. According to these regulations, it is not permitted to operate a pump with an efficiency of less than 65%, or a well pump with an efficiency of less than 55%. The pump curves are used for monitoring its performance as well as a central component in hydraulic models of WDSs. These models are used for design, operation, and monitoring tasks in WDSs, including: pipes' sizing and expansion [9,10], network calibration [11], detecting cyber-attacks [12], optimizing pumps operation [13,14], water quality modelling [15,16], sensor placement [17], minimizing greenhouse gas emissions [18], and analysis of water hammer effects [19].

Obtaining the pump curve by a pump test is a simple procedure for a fixed speed pump (FSP) as the pump can be tested at a few operational points (flows and head), with different suction and/or discharge pressures. By utilizing flow, pressure and power measurements, the curves can be constructed. On the other hand, variable speed pumps (VSP) are used to maintain a desired flow or pressure. Therefore, it is not feasible to test the pump across its range of speeds, so they are usually tested at a single speed, which is its nominal (i.e., maximum) speed. VSPs are common because they provide a number of advantages [20] including: (a) the pump flow can change gradually and give the upstream process (e.g., treatment plant) time to adjust; (b) no water storage is required on the demand side, as the pump can adjust to changing demands while maintaining the required pressure in the demand zone; (c) the flow can be changed gradually to reduce water hammer, and (d) motor life can be extended since fewer starts and stops are needed [21].

On the other hand, Gottliebson et al. [20] argue that VSPs also have disadvantages compared to FSPs: (a) they are more expensive in both installation and maintenance; (b) they may be less efficient; (c) controlling a VSP is more complex, and (d) a VSP may not be suitable for flat H-Q system curves as high efficiency is difficult to maintain over the entire flow range. In spite of these disadvantages, VSPs are most popular in systems in which no water storage is available, and where there is a need to regulate the flow using a demand-following mechanism.

With the VSPs gaining popularity in practice, they have been modelled in most simulation software, such as EPANET [22], and their modelling and simulation continue to be an active research topic [23–26]. Many studies of VSPs address the operation of WDSs and optimization of pumps scheduling [27–31]. In a recent review, Wu et al. [32] report improved system efficiency due to the use of VSPs, as well as other benefits of increased levels of flexibility in controlling WDSs in real time. Lima et al. [33] suggested the use of VSPs as a tool to recover energy and reduce leakage in WDSs. Wu et al. [34] incorporated VSPs in the design stage of water networks and transmission lines. Huo et al. [35] explored the option of using VSPs in deep injection well systems. All these studies assumed a fully known pump curves for their VSPs. Thus, as the effort in VSPs modelling and simulation grows, the need for accurate representation of the VPS curves increases.

Although not explicitly mentioned in most published research, the performance of pumps, including VSPs, may be evaluated from the abundance of historical records which is often available in WDSs that have Supervisory Control And Data Acquisition (SCADA) systems [5,36]. To produce the FSP's curves, one needs the flow, suction and discharge pressures (in fact only the difference is needed), and power readings for the pump. For a VSP the speed readings (or the motor frequency) are also required. Due to budget restrictions, many water utilities do not install all the required sensors in their pumping stations. In some cases, flow meters are not installed for each individual pump and only a single flow measurement is available for the entire pumping station. In other cases, the

individual pump's speed is not recorded. Power measurements are missing in many cases for the individual pumps and, if at all, are available only for the entire station. Facing these problems in real systems motivates the development of a methodology for deriving pump curves under limited data availability.

The methodology presented in this study is designed to produce the pump's curves when some of the data are missing. The implementation is demonstrated on a pressure zone in the water supply system of Mey-Sheva, which is the largest water utility in southern Israel. The entire system contains 6 pumping stations, 11 water tanks, and serves a population of 143,000 in Be'er-Sheva and the adjacent town Ofakim. The system has 670 km of water pipelines, of which about 100 km are part of the pressure zone.

2. Methodology

The methodology for pump curve calculation is designed to estimate them using SCADA data. Following the problem statement, the procedures are developed for single and combination of FSPs, and for single and combination of VSPs.

2.1. Problem Statement

The objective is to determine the curves for individual pumps in a pumping station that contains pumps in parallel, using SCADA data. The pumps may operate alone or in various combinations with other pumps. Ideally, a pumping station that consists of n pumps, $p \in Pumps$, where $Pumps$ is the set of pumps in the station, will have real-time measurements of each pump's flow, power, and on/off state, as well as the suction and discharge pressures of the station. However, due to budget limitations, individual pump flow and power are often unavailable and only the total station's flow is measured. In this study, we consider that the following data are available at multiple times: each pump's state, I_p (a binary variable where a value of 0 denotes that the pump is off and 1 when it is running), the total station's flow, Q_{obs} , and the suction and discharge pressures. The difference between the discharge and suction pressures is the station's head gain, H_{obs} . For each variable speed pump, its speed, n_p , is also recorded. All these values change over time and are available for each time step, $t \in T$, where T is a set of given historical time steps. We aim to estimate the pumps' curves in a quadratic form, which is a common form used to approximate the concave nonlinear association between head gain and pump flow [22,37].

$$H_p = a_p - b_p \cdot Q_p^2 \quad p \in Pumps \quad (1)$$

where, H_p and Q_p are the pump's head gain and flow, respectively. a_p (the shutoff head) and b_p are the function parameters.

2.2. Fixed Speed Pumps

We start with the simple case of a single fixed speed pump (Figure 1) operating alone. Using Equation (1) and the SCADA data, we calculate, in Equation (3), the head errors, e_H , between the estimated pump head, H_{est} (i.e., the head resulted from the H-Q curve, Equation (2)), and the observed head, H_{obs} . Since we are dealing with a single pump, the index p is omitted.

$$H_{est,t} = a - b \cdot (Q_{obs,t})^2 \quad \forall t \in T \quad (2)$$

$$e_{H,t} = H_{est,t} - H_{obs,t} \quad \forall t \in T \quad (3)$$

To estimate the parameters of Equation (1) the sum of the squares of e_H could be minimized, namely using least square regression. However, this method is sensitive to outliers in the SCADA data, and large deviations may bias unduly the fitting procedure. To avoid this deficiency of a least square regression, we use the least absolute errors (LAE) method in which the sum of absolute errors, E_H , as defined in Equation (4), is minimized.

$$E_H = \sum_{t \in T} |e_{H,t}| \quad (4)$$

In the ordinary case of curve fitting, it is customary to assign the errors to vertical distances (i.e., the head in our context) between the observations and the assumed curve. However, in general, one can assign the errors to horizontal distances (i.e., the flow in our context) between the observations and the curve (Figure 1). Thus, the curve fitting process can also be defined on the flows, by minimizing E_Q as defined in Equations (5)–(7).

$$Q_{est,t} = \sqrt{\frac{a - H_{obs,t}}{b}} \quad \forall t \in T \quad (5)$$

$$e_{Q,t} = Q_{est,t} - Q_{obs,t} \quad \forall t \in T \quad (6)$$

$$E_Q = \sum_{t \in T} |e_{Q,t}| \quad (7)$$

where, e_Q is the flow error, Q_{est} is the estimated flow, Q_{obs} is the observed flow, and E_Q is the sum of absolute flow errors. Figure 1 shows the vertical and horizontal distances of a measured point from the estimated pump curve. Generally, the estimated curve will not be the same for H and Q error minimization. While for a single fixed speed pump there is no clear advantage of one method over the other, and both vertical and horizontal errors can be easily calculated, the advantages of the horizontal calculations will become evident for more complicated cases that will be discussed next.

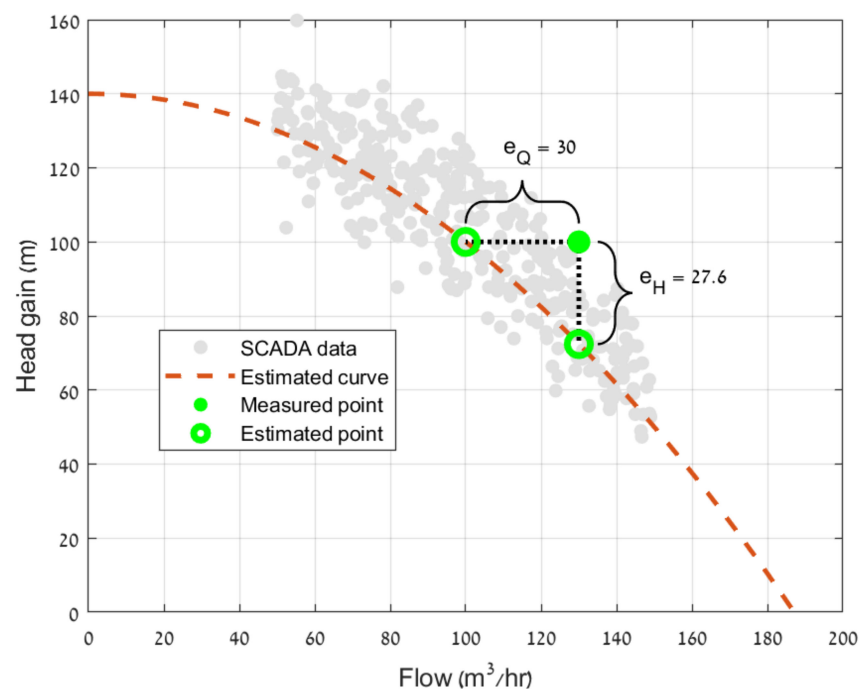


Figure 1. Vertical and horizontal errors between observations and assumed curve.

Pumping stations may have more than one pump operating in different combinations at different times. Consider, for example, a station with two identical pumps, Figure 2a shows the measured points of each pump alone and the two pumps working together. The case in which the two pumps operate together must be included in the procedure for deriving the curves for the individual pumps, since if only data for the pumps operating individually are considered, the curve's parameters may be biased because the entire operating region of the pump is not covered. To see this, consider the system curve in Figure 2b, when a pump operates alone it will be on the “right side” of its curve with high flow rates, as can be seen in Point A in Figure 2b. For two pumps operating together in parallel, the combined curve is obtained by summing the flows of the two pumps for each value of the head. The operating point in this case is Point B which is at the intersection of

the combined pump curve with the system curve. This operating point corresponds to a higher head compared to Point A, and thus each pump provides less flow at a higher head (Point C) than when it operates alone (Point A).

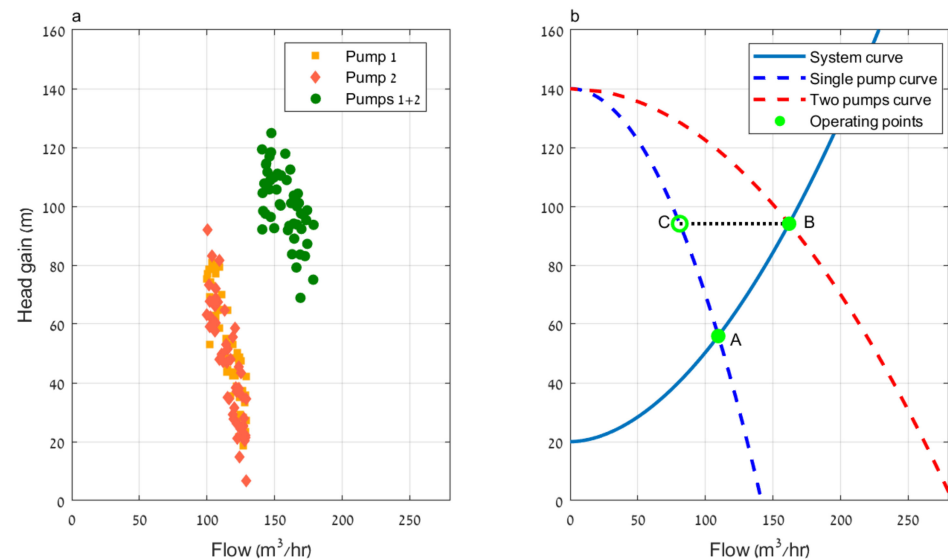


Figure 2. Two pumps operating in parallel (a) SCADA data (b) pumps and system curve.

This example demonstrates how individual pumps can work in different regions on the curve depending on the active pump combination. This behavior will be even more pronounced when there are several non-identical pumps in the pumping station with many combinations of operations. If we only use the data of a single pump operating alone (e.g., data near Point A) the estimated pump curve will tend to fit points in that region without considering the entire range of possible pump flows. To estimate the parameters of the individual curves based on the entire set of SCADA points (i.e., all combinations) there are two options:

Method 1

The combined curve for any specific combination can be derived as a function of the individual pumps' parameters. For example, given two pumps with curves, Equations (8) and (9), the combined function must be written explicitly, as Equation (10), and curve fitting technique with vertical error, e_H , can be used to estimate the parameters for each of the two pumps curves.

$$H_1 = a_1 - b_1 Q_1^2 \quad (8)$$

$$H_2 = a_2 - b_2 Q_2^2 \quad (9)$$

$$H_{est} = f(a_1, b_1, a_2, b_2, Q_{obs}) \quad (10)$$

While this procedure is valid in theory, deriving an explicit analytical function, even for only two pumps and certainly for more pumps operating together, is not practical due to the nonlinearity of the curves. To derive the function $f(\cdot)$, Equations (8) and (9) are written for Q as given in Equations (11) and (12).

$$Q_1 = \sqrt{\frac{a_1 - H_1}{b_1}} \quad (11)$$

$$Q_2 = \sqrt{\frac{a_2 - H_2}{b_2}} \quad (12)$$

It should be noted that for any number of pumps operating in parallel, their head gain is equal, thus $H_1 = H_2 = H_{est}$. Summing the two flows yields the total flow of the pumps station as given in Equation (13)

$$Q_{obs} = \underbrace{\sqrt{\frac{a_1 - H_{est}}{b_1}} + \sqrt{\frac{a_2 - H_{est}}{b_2}}}_{g(a_1, b_1, a_2, b_2, H_{est})} \quad (13)$$

The function $f(\cdot)$ is the inverse of the function $g(\cdot)$ defined in Equation (13). Thus, even for the simple case of two pumps, one cannot derive an explicit form for the function $f(\cdot)$. At best, the procedure is tractable for two pumps, by a numerical solution of Equation (10) or (13) for given values of a_1, b_1, a_2, b_2 . For a larger number of pumps this is not practical, since one needs to prepare in advance the combined functions for all possible combinations of pumps and then numerically solve all combination for each time step. Method 2 addresses this challenge by considering the horizontal error.

Method 2

Figure 3 demonstrates the calculation of the horizontal error for two pumps operating together. The observed flow, Q_{obs} , of the station is the combined flow of the two pumps, and H_{obs} is the measured head gain of the pumping station, which is equal for all pumps running in parallel. To calculate the horizontal error, e_Q in Equation (6), the estimated total flow, Q_{est} , is calculated based on the individual pumps' curves as given in Equation (14). That is, as shown in Figure 3, the observed head is used to estimate the flows from each pump, $Q_{est,1}$ and $Q_{est,2}$, which are then summed as an estimate for the total flow of the station. Given this total flow estimate the horizontal error, e_Q , can be calculated.

$$Q_{est} = \sqrt{\frac{a_1 - H_{obs}}{b_1}} + \sqrt{\frac{a_2 - H_{obs}}{b_2}} \quad (14)$$

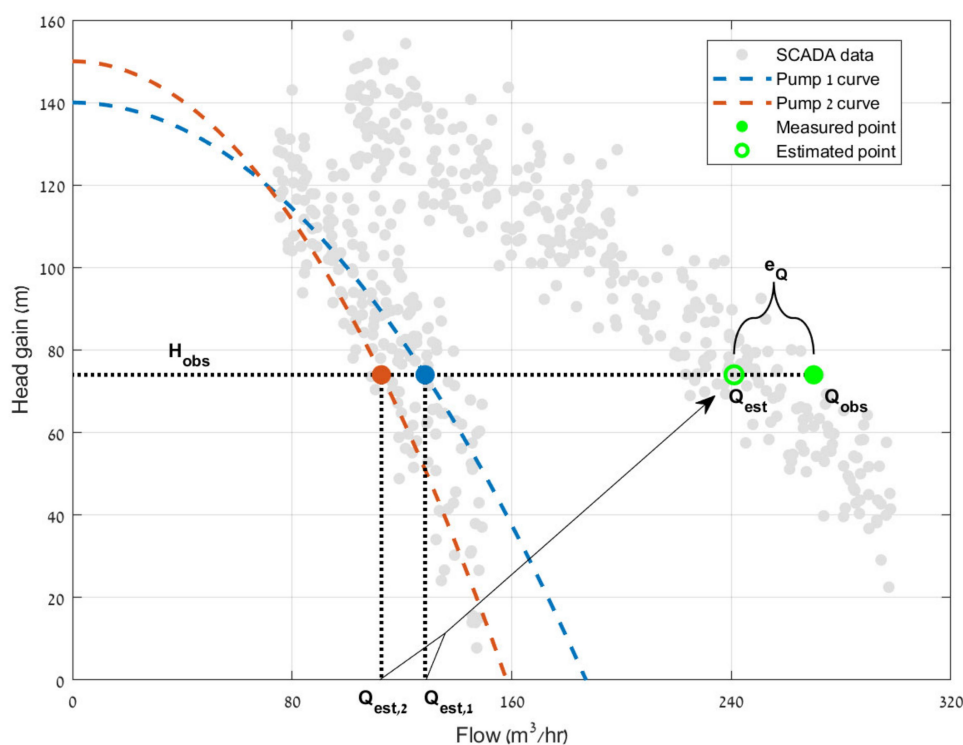


Figure 3. The horizontal error for two fixed speed pumps.

The advantage of this method is that it does not require an explicit derivation of the combined curve based on the individual curves' parameters (i.e., it does not require knowing the explicit function $f(\cdot)$). In fact, unlike Method 1, which requires preprocessing the curves of all combinations, using Method 2 leads to a generic optimization problem, which can consider any arbitrary number of pumps, Equations (15)–(20).

$$\min_{a_p, b_p} \sum_{t \in T} |e_{Q,t}| \quad (15)$$

Subject to:

$$Q_{est,p,t} = \sqrt{\left| \frac{a_p - H_{obs,t}}{b_p} \right|} \quad \forall p \in Pumps, \forall t \in T \quad (16)$$

$$Q_{est,t} = \sum_{p \in Pumps} I_{p,t} \cdot Q_{est,p,t} \quad \forall t \in T \quad (17)$$

$$e_{Q,t} = Q_{est,t} - Q_{obs,t} \quad \forall t \in T \quad (18)$$

$$a_p \geq 0 \quad \forall p \in Pumps \quad (19)$$

$$b_p \geq 0 \quad \forall p \in Pumps \quad (20)$$

Equation (15) is the objective function which minimizes the summed absolute horizontal errors by deciding on the pumps' curves parameters a_p and b_p . In Equation (16) the estimated individual pumps flows, $Q_{est,p,t}$, are calculated, and the total estimated flow for the pumping station is defined in Equation (17). It should be noted that the estimated flow is summed only for the pumps which are operating at each time by multiplying the individual pump's flow by its given binary state, $I_{p,t} \in \{0, 1\}$. In Equation (18) the horizontal error is calculated, and Equations (19) and (20) maintain the non-negativity of the parameters. While the absolute value in the objective function might be converted to set of linear constraints, the nonlinearity in the constraint in Equation (16) cannot be eliminated. Therefore, the obtained optimization problem is nonlinear. Readily available solvers such as `fmincon` shipped within Matlab [38] and the IPOPT [39] open source solver can handle this type of nonlinearity in the constraints. Moreover, since the number of decision variables is not expected to be high (e.g., a pumping station with 10 pumps will only result in 20 decision variables) and since this problem will be solved offline, one can utilize more computationally demanding global solvers such as Baron [40] to solve the optimization problem.

2.3. Variable Speed Pumps

A change in the rotational speed of a VSP changes its curves and operating point, according to the affinity laws:

$$\frac{Q_{p,t}}{\bar{Q}_{p,t}} = \frac{n_{p,t}}{\bar{n}_p} \quad \forall p \in Pumps, \forall t \in T \quad (21)$$

$$\frac{H_{p,t}}{\bar{H}_{p,t}} = \left(\frac{n_{p,t}}{\bar{n}_p} \right)^2 \quad \forall p \in Pumps, \forall t \in T \quad (22)$$

$$\frac{P_{p,t}}{\bar{P}_{p,t}} = \left(\frac{n_{p,t}}{\bar{n}_p} \right)^3 \quad \forall p \in Pumps, \forall t \in T \quad (23)$$

where, $\bar{n}_p$ is the nominal speed of pump p , and $n_{p,t}$, $\bar{Q}_{p,t}$, $\bar{H}_{p,t}$ and $\bar{P}_{p,t}$ are the speed, nominal flow, nominal head gain, and nominal power for time t respectively. We examine first the case in which a single pump is operating. Its flow and head are that of the pumping station, $Q_{p,t} = Q_{obs,t}$ and $H_{p,t} = H_{obs,t}$. Since $\bar{n}_p$ is a known constant value (usually equivalent to the pump's speed at 50 Hz), $\bar{Q}_{p,t}$ and $\bar{H}_{p,t}$ can be calculated from Equations (21) and (22) for each time t according to the speed of the pump, $n_{p,t}$. Once these

values are calculated for different points in time, a curve fitting method can be employed to derive the pump's curve (i.e., $\bar{H}_p = a - b \cdot \bar{Q}_p^2$) in the same way as for a fixed speed single pump (Section 2.2).

Thus, for times of a single operating pump we can easily map observed points to the nominal pump curve. However, as discussed previously, because the time instances of single pump operation are only a subset (sometimes only a small subset) of the entire operation, we may end up with a limited number of points on the curve, which causes a bias in the pump curve estimation process. This is especially true, when the points for the single pump operation fall on a narrow range of the pump curve. To overcome this bias, the pump's operation with other pumps must be included.

Next, we derive the equations for fitting all the pumps' curves simultaneously without restriction to specific time instances in which an individual pump is operating alone. That is, we also consider the time instances with joint operation of different pump combinations. From Equations (21) and (22) we obtain Equations (24) and (25).

$$\bar{Q}_{p,t} = Q_{p,t} \left(\frac{\bar{n}_p}{n_{p,t}} \right) \quad \forall p \in Pumps, \forall t \in T \quad (24)$$

$$\bar{H}_{p,t} = H_{p,t} \left(\frac{\bar{n}_p}{n_{p,t}} \right)^2 \quad \forall p \in Pumps, \forall t \in T \quad (25)$$

For VSPs we aim at fitting the nominal pumps' curves as shown in Equation (26):

$$\bar{H}_{p,t} = a_p - b_p \cdot \bar{Q}_{p,t}^2 \quad \forall t \in T, p \in Pumps \quad (26)$$

Substituting Equations (24) and (25) in Equation (26) and extracting $Q_{p,t}$ will yield a relationship between the individual pump heads and flows at any speed.

$$Q_{p,t} = \sqrt{\frac{\left(\frac{n_{p,t}}{\bar{n}_p} \right)^2 a_p - H_{p,t}}{b_p}} \quad \forall t \in T, p \in Pumps \quad (27)$$

Since the measured head is the same for all pumps, $H_{p,t} = H_{obs,t} \quad \forall p \in Pumps \quad \forall t \in T$, by substituting $H_{obs,t}$ in Equation (27) we can obtain an estimate for the flows of the individual pumps.

Owing to the advantages of the horizontal error formulation, the optimization problem is formulated in Equations (28)–(33). The objective function, Equation (28) is minimized by deciding on the pump's curve parameters a_p and b_p . The individual pumps' flows are estimated in Equation (29). To obtain the total estimated flow of the pumping station, the individual pumps' flows are summed in Equation (30) and the horizontal error can be calculated in Equation (31), while maintaining the non-negativity of the parameters in Equations (32) and (33).

$$\min_{a_p, b_p} \sum_{t \in T} |e_{Q,t}| \quad (28)$$

Subject to:

$$Q_{est,p,t} = \sqrt{\frac{\left(\frac{n_{p,t}}{\bar{n}_p} \right)^2 a_p - H_{obs,t}}{b_p}} \quad \forall t \in T, p \in Pumps \quad (29)$$

$$Q_{est,t} = \sum_{p \in Pumps} I_{p,t} \cdot Q_{est,p,t} \quad \forall t \in T \quad (30)$$

$$e_{Q,t} = Q_{est,t} - Q_{obs,t} \quad \forall t \in T \quad (31)$$

$$a_p \geq 0 \quad \forall p \in Pumps \quad (32)$$

$$b_p \geq 0 \quad \forall p \in \text{Pumps} \quad (33)$$

The process for calculating the horizontal error is shown in Figure 4. First, the normalized heads for the two pumps are calculated (at points A and B) using their operating speed by Equation (25). Then, using the normalized heads the normalized flows can be estimated by inverting Equation (26). Next, using each pump's speed, their estimated individual flows ($Q_{est,1}$, $Q_{est,2}$) are estimated (at Points C and D) by inverting Equation (24). In the optimization model, these operations are lumped together in Equation (29). Finally, the estimated flows are summed in Equation (30) to obtain the total estimated flow Q_{est} , which is then used in calculating the horizontal error in Equation (31).

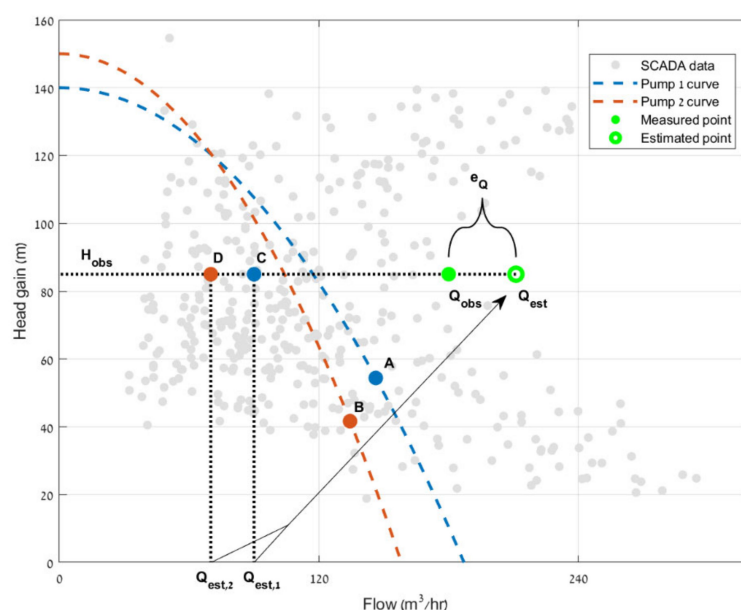


Figure 4. Calculation of the horizontal error for two variable speed pumps.

One should note the similarity between the formulation of the fixed speed pumps in Equations (15)–(20) and the variable speed pumps formulation in Equations (28)–(33). In fact, the former is a particular case of the latter when $n_{p,t} = \bar{n}_p \forall t$. Furthermore, in this way, the hybrid case of a pumping station featuring both fixed and variable speed pumps could also be addressed. As such, to use this methodology it is sufficient to implement the formulation in Equations (28)–(33) for an arbitrary combination of fixed and variable speed pumps.

3. Test Case and Results

As a test case, we consider a large pressure zone (PZ) in the Southern Israeli city of Be'er-Sheva which is operated by the Mey-Sheva (<https://mey7.co.il/en> accessed on 25 January 2021) water cooperation.

The PZ is supplied by a single pumping station and has no storage tanks. The station has four variable speed pumps, and each can operate over a range of frequency settings which are correlated with their speeds. SCADA measurements (31 December 2019–03 May 2020) are available at 30 s intervals: suction and discharge pressures, total station flow, and individual pump frequencies (SCADA data is available in Data Set S1 in the Supplementary Materials). The frequencies are recorded as percentage [0, 100] for a range of 35–50 Hz. When the value is 0% it indicates that the pump is turned off or operating at its minimum frequency. There are no individual pump flow data and no power data. Power data are not available even for the entire station.

The pumps are operated to maintain the discharge pressure of the station within pre-specified range. As demand in the PZ increases, the pressure in the network decreases, and

the speed of the operating pump is raised to meet the required pressure. The speed is raised to the maximum speed and then another pump is switched on at its lowest speed, which is increased if the pressure continues to drop. The controlled pressure is set to be around 47 m during daytime (06:30–23:00) and to about 42 m during night hours (23:00–06:30) as seen in Figure 5.

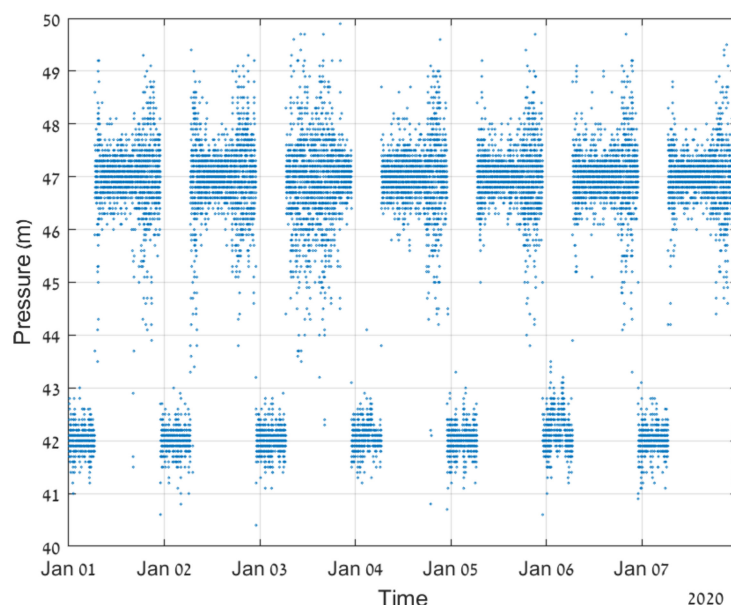


Figure 5. Pumping station discharge pressure.

Observing the station's overall Flow-Head Gain plot in Figure 6, it can be seen that the station operates to produce a relatively constant discharge head for a wide range of flows, owing to the control strategy setting stated above. The two regions ("data clouds") in Figure 6 correspond to the day and nighttime discharge-head points.

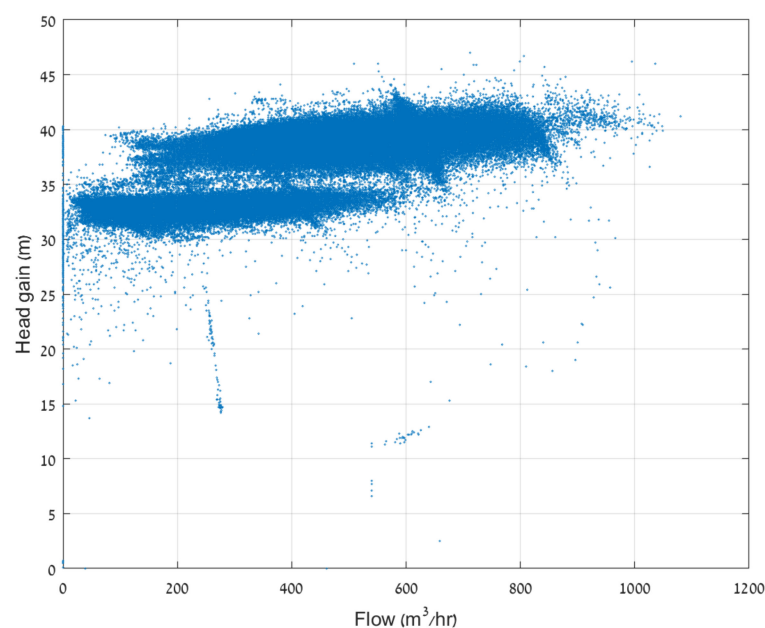


Figure 6. Pumping station Flow-Head Gain points.

We begin by scanning the SCADA data for points that correspond to a single pump in operation, to determine the curves for each pump alone. This was done by searching for

times where only one pump has a frequency greater than 35 Hz ($>0\%$ in the data) while all other pumps are at a frequency of 35 Hz ($=0\%$). Figure 7 shows the pump curves at the nominal speed (as per Equations (21) and (22)) for the individual pumps as recorded by the SCADA system. Figure 7 also shows the most recent available pump tests, which were performed in September 2017 by the Israeli Water Works Association (IWWA).

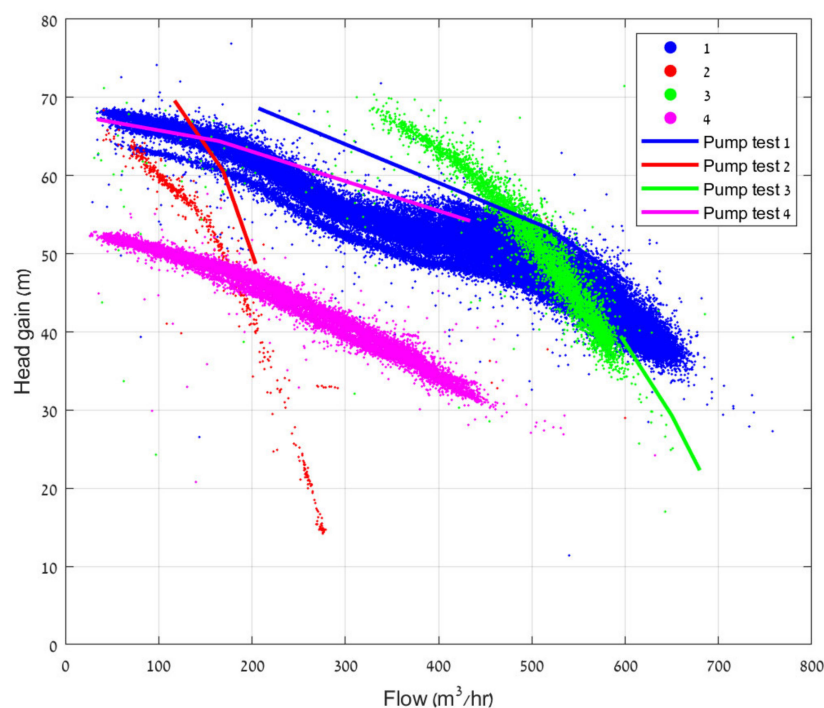


Figure 7. Individual pump curves for the test case pumping station.

The following observations can be made from Figure 7: (a) the pump test curves are parallel to the SCADA data; (b) the SCADA data are below the pump tests, as expected for deterioration of pump performance over time; (c) the SCADA data for pump 4 is about 20 m below the test curve, which seems excessive and requires further investigation; and (d) for pumps 2 and 3 the test covers only a small part of the actual flows experienced in operation. In fact, for pump 3, the test only covers a small region which was rarely used in the system.

Since the pumps operate alone only part of the time, it is necessary to evaluate their curves while they operate in different combinations. Utilizing the optimization methodology outlined in Section 2.3, (Equations (28)–(33)) we calculated the updated curves (Figure 8). The curve parameters appear in Table 1. For the case studied herein, the updated curve coefficients are not very different from the ones obtained when the pumps operate individually, except for pump 1 where the difference is more pronounced. Still, our generic methodology accounts for all available data for deriving the pump curves, thus it can deal with situations where the pumps operate in different regions under different conditions as was discussed in Figure 2.

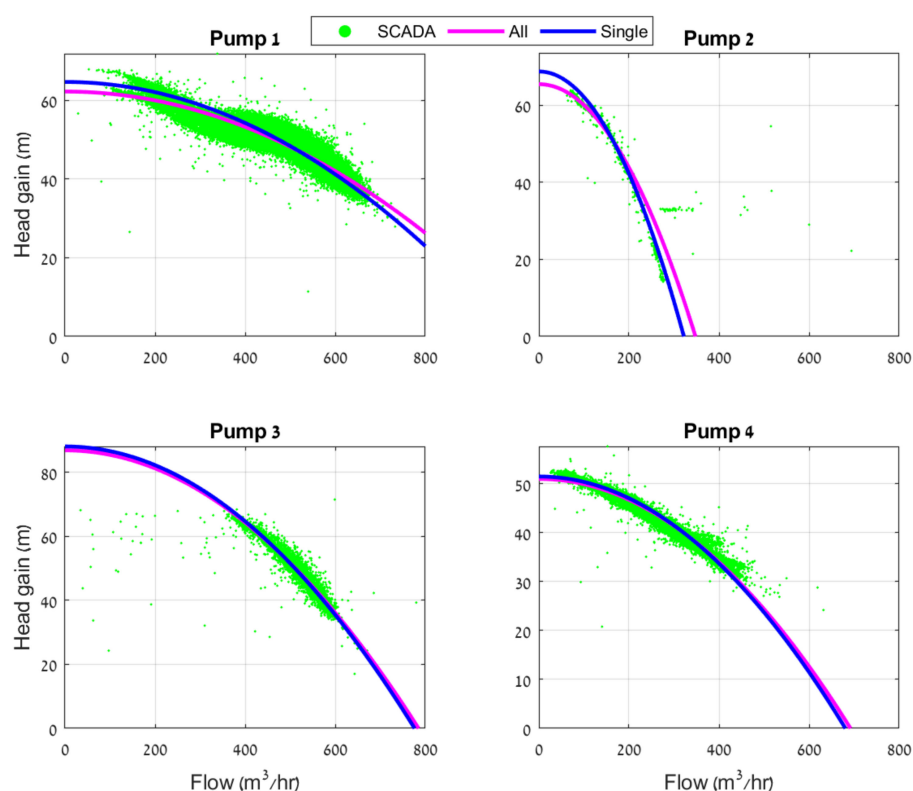


Figure 8. Estimated pumps curves for the test case pumping station.

Table 1. Pumps' curve coefficients.

Pump	Single		All	
	a	b ($\times 10^{-4}$)	a	b ($\times 10^{-4}$)
1	64.00	0.621	66.29	0.701
2	66.63	6.163	65.78	5.826
3	85.62	1.364	83.93	1.309
4	51.05	1.072	51.07	1.073

4. Conclusions

In this study we present a practical pump curve constructing methodology for obtaining individual pumps' curves from partially available SCADA data. Using it, each pump's performance can be monitored continuously between physical on-premises inspections. We considered the case in which the analyst aims at estimating the individual pump curves whilst the measurements are only performed on the pump station level (i.e., only total flow is available). We show that, unlike ordinary curve fitting techniques that minimize the vertical error between observations and estimating curve, it is advantageous to use the notion of horizontal errors (i.e., the deviation in terms of flow) since it does not require deriving explicit functions for all pump combinations in the pumping station.

The proposed methodology is formulated as a non-linear optimization problem with a small number of decision variables (two for each pump) which can be solved with open source or commercial global solvers. The proposed optimization formulation is generic for any number and type of pumps. Thus, it can be utilized for a single or multiple pumps, and for fixed or variable speed pumps (and a combination of these types). The method can estimate the pumps' curves for any given amount of historical SCADA dataset, without increasing the size of the optimization problem.

The methodology developed herein constitutes a contribution to an increasing trend to "back-figure" system information from SCADA data. Relying on SCADA data for estimat-

ing pump curves has significant practical implications. Field pump tests are expensive and may even be impractical. Indeed, our results on a real test case show that the individual pumps' SCADA data and their corresponding field test results can be significantly different. Furthermore, the field tests, for some of the pumps, cover only part of operating range. This emphasizes the importance of the proposed methodology that derive the pump's curve over the entire operating range under different operating conditions. That is, the method utilizes measurements not only from times when the pump operates alone, but also when it operates together with other pumps. It is worth noting that these situations, where different pump combinations operate at different time slots, are more common when pumps are connected in parallel. Our methodology is tailored for pumps connected in parallel, but still, a dedicated methodology for pumps connected in series is warranted and will be considered for future work.

Unfortunately, in our test case the station power consumption data were not available, and, therefore, the power and individual efficiency curves could not be calculated. Still, if the station power consumption data were available, the methodology could be used to estimate the individual power consumption curves.

Supplementary Materials: The following are available online at <https://www.mdpi.com/2073-4441/13/5/586/s1>, Dataset S1: The SCADA data used in this study.

Author Contributions: Conceptualization, E.S., M.H. and U.S.; methodology, E.S. and M.H.; software, E.S.; validation, E.S. and M.H.; writing—original draft preparation, E.S.; writing—review and editing, M.H. and U.S.; supervision, M.H. and U.S.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are available in the Supplementary Materials.

Acknowledgments: This study was supported by the Israeli Water Authority. The help and support of Mey-Sheva management and staff is highly appreciated.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Copeland, C.; Carter, N.T. Energy—Water Nexus: The Water Sector's Energy Use R43200. 2017. Available online: <https://crsreports.congress.gov/product/pdf/R/R43200> (accessed on 25 January 2021).
2. Sanders, K.T.; Webber, M.E. Evaluating the energy consumed for water use in the United States. *Environ. Res. Lett.* **2012**, *7*, 034034. [CrossRef]
3. Lam, K.L.; Kenway, S.J.; Lant, P.A. Energy use for water provision in cities. *J. Clean. Prod.* **2017**, *143*, 699–709. [CrossRef]
4. ETSU; CETIM; Reeves, D.T.; NESA; Technical University Darmstadt. Study on Improving the Energy Efficiency of Pumps. European Commission: UK, 2001. Available online: <http://www.jakob-albertsen.dk/komposit/Darmstadtrapport.pdf> (accessed on 25 January 2021).
5. Bunn, S.M.; Reynolds, L. The energy-efficiency benefits of pumpscheduling optimization for potable water supplies. *IBM J. Res. Dev.* **2009**, *53*, 5:1–5:13. [CrossRef]
6. Eaton, A.; D'Alessandro, F.; Ahmed, W.; Hassan, M. On the performance degradation of centrifugal pumps. In Proceedings of the International Conference on Fluid Flow, Heat and Mass Transfer, Niagara Falls, Canada, 7–9 June 2018; Avestia Publishing: Orléans, ON, Canada, 2018; p. 158.
7. Sanks, R.L.; Tchobanoglous, G.; Bosserman Ii, B.E.; Jones, G.M. *Pumping Station Design*, 2nd ed.; Butterworth-Heinemann: Oxford, UK, 1998.
8. Israeli Regulations: Energy Efficiency Test in Pumping Facilities. Available online: https://www.nevo.co.il/law_html/law01/99_9_315.htm (accessed on 10 December 2020).
9. Costa, A.L.H.; De Medeiros, J.L.; Pessoa, F.L.P. Optimization of pipe networks including pumps by simulated annealing. *Braz. J. Chem. Eng.* **2000**, *17*, 887–895. [CrossRef]
10. Stokes, C.S.; Simpson, A.R.; Maier, H.R. A computational software tool for the minimization of costs and greenhouse gas emissions associated with water distribution systems. *Environ. Model. Softw.* **2015**, *69*, 452–467. [CrossRef]

11. Ostfeld, A.; Salomons, E.; Ormsbee, L.; Uber, J.G.J.G.; Bros, C.M.C.M.; Kalungi, P.; Burd, R.; Zazula-Coetzee, B.; Belrain, T.; Kang, D.; et al. Battle of the Water Calibration Networks. *J. Water Resour. Plan. Manag.* **2012**, *138*, 523–532. [\[CrossRef\]](#)
12. Housh, M.; Ohar, Z. Model-based approach for cyber-physical attack detection in water distribution systems. *Water Res.* **2018**, *139*, 132–143. [\[CrossRef\]](#)
13. Marchi, A.; Simpson, A.R.; Lambert, M.F. Optimization of Pump Operation Using Rule-Based Controls in EPANET2: New ETTAR Toolkit and Correction of Energy Computation. *J. Water Resour. Plan. Manag.* **2016**, *142*, 04016012. [\[CrossRef\]](#)
14. Khatavkar, P.; Mays, L.W. Model for real-time operations of water distribution systems under limited electrical power availability with consideration of water quality. *J. Water Resour. Plan. Manag.* **2018**, *144*, 04018071. [\[CrossRef\]](#)
15. Shang, F.; Uber, J.G.; Rossman, L.A. Modeling reaction and transport of multiple species in water distribution systems. *Environ. Sci. Technol.* **2008**, *42*, 808–814. [\[CrossRef\]](#)
16. Rossman, L. Modeling of chlorine residual in drinking water distribution system. *J. Environ. Eng. ASCE* **1995**, *120*, 803–820, ISSN 0733-9372/94/0004-0803. [\[CrossRef\]](#)
17. Ostfeld, A.; Salomons, E. Optimal layout of early warning detection stations for water distribution systems security. *J. Water Resour. Plan. Manag.* **2004**, *130*, 377–385. [\[CrossRef\]](#)
18. Blinco, L.J.; Simpson, A.R.; Lambert, M.F.; Marchi, A. Comparison of Pumping Regimes for Water Distribution Systems to Minimize Cost and Greenhouse Gases. *J. Water Resour. Plan. Manag.* **2016**, *142*, 04016010. [\[CrossRef\]](#)
19. Wood, D.J. Waterhammer Analysis—Essential and Easy (and Efficient). *ASCE* **2005**, *131*, 1123–1131. [\[CrossRef\]](#)
20. Gottliebson, M.; Sanks, R.L.; Vause, A.; Abelin, S.M.; Benfell, R.S.; Bicknell, H.C.; Densmore, B.; Duncan, R.W.; Hong, S.S.; Jones, C.; et al. Variable-Speed Pumping. In *Pumping Station Design*; Elsevier Ltd.: Amsterdam, The Netherlands, 2008; ISBN 9781856175135.
21. Lansey, K.; Awumah, K. Optimal Pump Operations Considering Pump Switches. *J. Water Resour. Plan. Manag.* **1994**, *120*, 17–35. [\[CrossRef\]](#)
22. Rossman, L.A. *EPANET2 User's Manual*; U.S. Environmental Protection Agency: Cincinnati, OH, USA, 2000.
23. Marchi, A.; Simpson, A.R. Correction of the EPANET Inaccuracy in Computing the Efficiency of Variable Speed Pumps. *J. Water Resour. Plan. Manag.* **2013**, *139*, 456–459. [\[CrossRef\]](#)
24. Ulanicki, B.; Kahler, J.; Coulbeck, B. Modeling the efficiency and power characteristics of a pump group. *J. Water Resour. Plan. Manag.* **2008**, *134*, 88–93. [\[CrossRef\]](#)
25. Marchi, A.; Simpson, A.R.; Ertugrul, N. Assessing variable speed pump efficiency in water distribution systems. *Drink. Water Eng. Sci.* **2012**, *5*, 15–21. [\[CrossRef\]](#)
26. Todini, E.; Tryby, M.E.; Wu, Z.Y.; Walski, T.M. Direct computation of Variable Speed Pumps for water distribution system analysis. In *Proceedings of the Combined International Conference of Computing and Control for the Water Industry, CCWI2007 and Sustainable Urban Water Management, SUWM2007*, Leicester, UK, 3–5 September 2007; pp. 411–418.
27. AbdelMeguid, H.; Ulanicki, B. Feedback Rules for Operation of Pumps in a Water Supply System Considering Electricity Tariffs. *Water Distrib. Syst. Anal.* **2010**. [\[CrossRef\]](#)
28. Coulbeck, B.; Orr, C.H. Real-Time Optimized Control of a Water Distribution System. *IFAC Proc. Vol.* **1989**, *22*, 441–446. [\[CrossRef\]](#)
29. Abdallah, M.; Kapelan, Z. Fast Pump Scheduling Method for Optimum Energy Cost and Water Quality in Water Distribution Networks with Fixed and Variable Speed Pumps. *J. Water Resour. Plan. Manag.* **2019**, *145*, 04019055. [\[CrossRef\]](#)
30. Hashemi, S.S.; Tabesh, M.; Ataekia, B. Ant-colony optimization of pumping schedule to minimize the energy cost using variable-speed pumps in water distribution networks. *Urban Water J.* **2014**, *11*, 335–347. [\[CrossRef\]](#)
31. Bene, J.G.; Hos, C.J. Finding least-cost pump schedules for reservoir filling with a variable speed pump. *J. Water Resour. Plan. Manag.* **2012**, *138*, 682–686. [\[CrossRef\]](#)
32. Wu, W.; Maier, H.R.; Dandy, G.C.; Arora, M.; Castelletti, A. The changing nature of the water–energy nexus in urban water supply systems: A critical review of changes and responses. *J. Water Clim. Chang.* **2020**, *11*, 1095–1122. [\[CrossRef\]](#)
33. Lima, G.M.; Luvizotto, E.; Brentan, B.M.; Ramos, H.M. Leakage control and energy recovery using variable speed pumps as turbines. *J. Water Resour. Plan. Manag.* **2018**, *144*, 04017077. [\[CrossRef\]](#)
34. Wu, W.; Simpson, A.R.; Maier, H.R.; Marchi, A. Incorporation of variable-speed pumping in multiobjective genetic algorithm optimization of the design of water transmission systems. *J. Water Resour. Plan. Manag.* **2012**, *138*, 543–552. [\[CrossRef\]](#)
35. Huo, J.; Eckmann, D.H.; Hoskins, K.P. Hydraulic simulation and variable speed pump selection of a dual-function pumping station: Pumping treated water to a reclaimed water system or deep injection well system. In *World Environmental and Water Resources Congress*; American Society of Civil Engineers: Reston, VA, USA, 2007; pp. 1–7.
36. Sunela, M.I.; Puust, R. A visual tool to calculate optimal control strategy for non-identical pumps working in parallel, taking motor and VSD efficiencies into account. *Water Sci. Technol. Water Supply* **2015**, *15*, 1115–1122. [\[CrossRef\]](#)
37. Gostoli, C.; Spadoni, G. Linearization of the head-capacity curve in the analysis of pipe networks including pumps. *Comput. Chem. Eng.* **1985**, *9*, 89–92. [\[CrossRef\]](#)
38. *MATLAB 9.5.0.1033004 (R2018b)*; The MathWorks Inc.: Natick, MA, USA, 2018.
39. Wächter, A.; Biegler, L.T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* **2006**, *106*, 25–57. [\[CrossRef\]](#)
40. Tawarmalani, M.; Sahinidis, N.V. A polyhedral branch-and-cut approach to global optimization. In *Proceedings of the Mathematical Programming*; Springer: Berlin/Heidelberg, Germany, 2005; Volume 103, pp. 225–249.