processes

Article

Predicting Centrifugal Pumps' Complete Characteristics Using Machine Learning

Jiangping Yu [1], Emmanuel Akoto [1,*], Derek Kweku Degbedzui [2] and Liren Hu [1]

[1] College of Petrochemical Engineering, Lanzhou University of Technology, Lanzhou 730050, China
[2] Department of Biomedical Engineering, Erciyes University, Kayseri 38039, Turkey
* Correspondence: eakoto@st.ug.edu.gh

Abstract: The complete characteristics of centrifugal pumps are crucial for the modeling of hydraulic transient phenomena occurring in pipe systems. However, due to the effort required to obtain these curves, pump manufacturers typically only provide basic information, particularly when the pump operates under normal conditions. To acquire the full characteristic curves based on the manufacturer's normal performance curve, a machine learning (ML) model is proposed to predict full, complete Suter curves using a pump's specific speed with the known parts of the Suter curve. The training data for the model are sourced from the available Suter curves from laboratory experiments. Subsequently, the proposed ML model combines several types of regression models in an attempt to find the most accurate prediction in terms of the root mean square error (RMSE). The result proved highly efficient, as the experiments attained a maximum RMSE value of 0.032 across the three categories of centrifugal pumps based on their specific speeds, hence demonstrating the potential of machine learning in the study of pump characteristic curves.

Keywords: Suter curves; pump hydraulic transient; pump specific speed; dimensionless torque curve; dimensionless head curve; performance curve

Citation: Yu, J.; Akoto, E.; Degbedzui, D.K.; Hu, L. Predicting Centrifugal Pumps' Complete Characteristics Using Machine Learning. *Processes* **2023**, *11*, 524. https://doi.org/10.3390/pr11020524

Academic Editor: Jiaqiang E

Received: 30 December 2022
Revised: 2 February 2023
Accepted: 7 February 2023
Published: 9 February 2023

1. Introduction

A lot of research has been conducted in relation to the transient flow mechanisms in pumps and valves. Among these, pump startups and failures are found to be a major source of hydraulic transients [1]. The methods of simulating and analyzing these kinds of hydraulic transients have so far included the analytical method, the graphical method, and the use of the method of characteristics. However, the latter has precedence over all methods and is widely adopted for the computation of transient flow for pump failures and startups [2,3]. During the application of the method of characteristics analysis, the convenience of using non-dimensional forms of pump performance characteristics to represent pump behaviors through graphical curves presents another hurdle, as performance characteristic values approach infinity for certain flow conditions, which are of interest in transient analyses. In 1965, a researcher named Suter found a solution to this problem by introducing a new parameter [4–6].

It is fair to say that Suter is the pioneer of modern studies of the characteristics of pumps, as his experimental findings in 1966 are still of significant relevance and are in application in today's transient studies of pumps. Nevertheless, prior to the publication of Suter's findings, Stepanoff's [7] and Donsky's [8] complete pump characteristic curves of three specific speeds (Ns), Ns = 25 (one radial pump), Ns = 147 (one semi-axial pump), and Ns = 261 (one axial pump), served as benchmarks for the transient calculations of pumps [9]. Engineers back then chose the curves of specific speeds closest to their machines and then, with or without interpolation, approximated fluid transient calculations [10]. These curves are essential in determining the full operational zones of a pump, that is, the normal operation, the turbine operation, the energy dissipation, the reverse pumping, and the braking zones during steady and transient states [10]. All these zones together form a complete characteristic curve and are delineated into four quadrants. Pump performance curve studies comprise a flow ratio to head ratio coordinate system with four quadrants, namely, Quadrants I, II, III, and IV. Quadrant I, known as the pump drive, operates as a motor; in Quadrant II, the pump operates as a generator as the runner rotates in the direction of the pump, and the water flows into the turbine, representing the braking of the aggregate; Quadrant III is considered a turbine drive because the electric machine operates as a generator; and in Quadrant IV, the runner rotates in the direction of the turbine, the water flows into the pump, and the electrical machine operates as a motor [11]. In each quadrant, the pump operational zones can be plotted as a set of speed ratio and torque ratio [7,8,10] curves. However, since all of the pump operational parameters (rotation, flow rate, head, and torque) can attain a value of zero, the use of ratios results in singularities or asymptotes when the denominator of ratios is zero. Before the emergence of Suter's curves, pump characteristics were divided into separate parts and then switched between one another during analyses [6–8]. However, this approach still had issues, as a single pump operation could fall into all four quadrants without a specific order [10,12]. With Suter finding a way to eliminate the possibility of singularities in performance curves, a game changer emerged.

Suter [5] represented the four-quadrant curves with three dimensionless variables, namely, the dimensionless head ratio (Wh), representing the head; the dimensionless torque ratio (Wb), representing the torque; and theta (θ), which was the new independent variable that he introduced. With these variables, it was possible to completely represent the full pump characteristics with just the dimensionless head and torque curves plotted over the domain of θ, which is between 0 and 2π without any asymptotes. Suter's curve revolutionized the studies of pump–turbine performance characteristics in light of digital computation. The data from the curves are very much suited to digital computation, as the values of Wh and Wb corresponding to a particular θ can be easily computed by interpolation. The torque values can be used to calculate the running speed of the pump at a particular time period, and the head across the pump is then used in the boundary condition calculation [10]. However, it is very unlikely that anything other than the normal head versus flow rate and the pressure versus flow rate curves of the pump turbines in the normal pumping mode or the normal turbine mode will be available, so it is necessary to devise a reasonable approximation to the Suter curve for the other modes of pump-turbine operation [2,13]. The assumptions needed in order to generate Suter's curves of unknown pump-specific speeds through interpolation raise an eyebrow. For instance, the four-quadrant data are made dimensionless relative to the best efficiency point (BEP) or pump-rated parameters [14]. This brings about a concern of correlation, as it is rare to find pumps that operate at their BEP. In light of this, Brown and Rogers in 1980 questioned the validity of using pump-specific speed as a correlating factor for four-quadrant pump data, as their field work showed a much weaker correlation between four-quadrant characteristics and the specific speed for radial pumps but a better correlation for mixed and axial flow pumps [14]. In spite of these concerns, Suter's curve is still used to date, as there has not been any model that better suits the demands of modern numerical computations.

In addition, some researchers explored the use of computational fluid dynamics (CFD) in finding complete characteristic curves. Despite the complexities of setting up a grid mesh, flow and pressure field equations, and boundary conditions, some researchers chalked some success. Gros et al. [15] using CFD obtained very satisfactory results for the four quadrants of a centrifugal pump based on two transient equations of turbulence models. Similar work was carried out by Muttalli [16], who used ANSYS-CFX and went further ahead to discuss the impact of cavitation on numerical simulation. Höller et al. [17] investigated the steady and transient state conditions surrounding a mixed flow pump also using CFD and obtained fairly accurate results. Frosina et al. [18] studied three centrifugal pumps of different specific speeds using CFD and used the obtained results to evaluate the inverse characteristics and the BEP of those pumps. Wang et al. [19] pulled ahead by constructing a three-dimensional CFD model and proposing a new method for determining the complete Wh and Wb curves for double-suction centrifugal pumps. These CFD methods required extensive computational power and time to be able to obtain fairly accurate results and, hence, were expensive.

In a bid to avoid costly experiments, many curve-fitting methods for complete characteristic curves have come into play [2]. Shao [20] formulated the curves by applying surface fitting and the least square method. In addition, Zhang et al. [21] experimented with the use of B-spine interpolation to develop complete characteristic curves. Other methods, such as neural network models, have also been used to compute pump characteristic curves. Moreover, Han et al. were able to accurately predict a centrifugal pump performance curve in the normal operation mode by using the specific speed, flow rate, impeller inlet and outlet diameters, hub diameter, blade outlet width, and blade number as inputs [22]. Wuyi and Wenrui derived the inversion method to compute complete pump performance characteristics from a normal performance curve [2]. Although this inversion method was derived without any approximate assumptions, it could only compute a complete third-quadrant curve. More recent researchers, such as Li et al., have also tackled this problem. Li et al. proposed a new method by combining the mathematical modeling of a centrifugal pump with the correlation relationships of some known complete characteristic curves [23]. However, according to their demonstration in the paper, this method could only accurately predict within the domain of $0 > \theta < 5\pi/4$. This paper proposes a machine learning model that incorporates several regression learners, including a linear regression learner, a support vector machine, a Gaussian process regression, and a three-layered neural network, to find a more generalized complete characteristic curve prediction model over the entire spectrum of the various specific speeds of a pump. The proposed model attempts to accurately predict complete characteristics over the entire domain of $0 > \theta < 2\pi$. This model receives the specific speeds and the known parts of Suter's curve as inputs and then generates the complete curve characteristics with the combined lowest minimum RMSE. This paper first starts with a brief introduction to the modeling of Suter's curve and pump operational zones, and then it continues with an in-depth description of the machine learning model. It finally ends with an evaluation and discussion of the obtained predicted complete characteristic curve results.

2. Pump Operational Zones, and Suter Curve Modeling

Pump operational studies basically revolve around four variables, namely, the pump rotational speed (N), flow rate (Q), head (H), and torque (T). The relationship between the changes in these four variables is what is responsible for the various operational zones of a pump. In all, there are eight operational zones conventionally labeled from A to H, and one would need to fully comprehend the dynamics of all these zones to be able to accurately predict a pump's hydraulic transients. Donsky et al., Knapp et al., and Giljen et al. dissect the makings of these zones brilliantly, and, hence, a summary of the definitions of these is presented in Table 1. Therefore, for convenience in modeling, these variables are standardized into their dimensionless forms using their respective rated pump values.

The ratios of the dimensionless forms of rotation speed (α), flow rate (v), head (h), and torque (β) can be plotted to represent the pump head and torque characteristic curve. For instance, in obtaining the pump head curve, the ratio of the dimensionless flow rate (v) to the dimensionless rotation speed (α) is plotted against the ratio of the dimensionless head (h) to the dimensionless rotation speed (α). Based on Table 1, it is evident that transitioning between some operational zones is bound to introduce singularities into the plotting dataset as some variables approach zero. This accentuates Suter's transform curves, as it finds a way to produce both dimensionless pump head and torque characteristic curves without discontinuities. The set of Equations (1)–(3) constitutes Suter's transform [5]:

$$W_h(\theta) = \frac{h}{\alpha^2 + v^2} \tag{1}$$

$$W_b(\theta) = \frac{\beta}{\alpha^2 + v^2} \tag{2}$$

$$\theta = \arctan\frac{\alpha}{v} \tag{3}$$

Table 1. Summary of Definitions of Pump Operation Zones [10].

Zone	Mode	Rotation	Flow	Head	Torque
A	Normal Pumping	N > 0	Q > 0	H > 0	T > 0
B	Energy Dissipation	N > 0	Q > 0	H < 0	T > 0
C	Reverse Turbining	N > 0	Q > 0	H < 0	T < 0
D	Energy Dissipation	N < 0	Q > 0	H > 0	T < 0
E	Reverse Pumping	N < 0	Q > 0	H > 0	T < 0
F	Braking	N < 0	Q < 0	H > 0	T < 0
G	Normal Turbining	N < 0	Q < 0	H > 0	T > 0
H	Energy Dissipation	N > 0	Q < 0	H > 0	T > 0

Through this transformation, the complete pump head and torque characteristic curve can be plotted over the domain of θ spanning from 0 to 2 π. This domain can further be divided into four: (0 > θ < π/2) the first quadrant, (π/2 > θ < π) the second quadrant, (π > θ < 3π/2) the third quadrant, and lastly (3π/2 > θ < 2π) the fourth quadrant. The four-quadrant characteristic model is a tool used to classify and evaluate the operational characteristics of different devices and systems. In the context of centrifugal pumps, Quadrant I is referred to as the Turbine Zone, Quadrant II is the Dissipation Zone, Quadrant III is the Normal Zone, and Quadrant IV is the Reversed Speed Dissipation Zone [2]. The dimensionless head and torque, as determined using Equations (1) and (2), are plotted against theta in Equation (3), forming the Suter curve. Figure 1 shows a sample of complete characteristic or Suter curves plotted in the four quadrants with the operating zones (A-H) indicated.

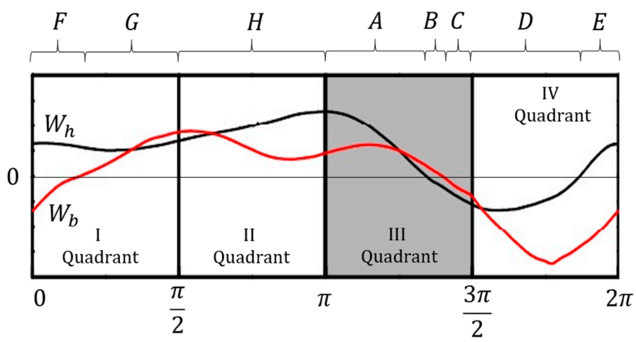


Figure 1. Rectangular coordinate showing pump operation zones and quadrants.

3. Machine Learning (ML) Model

An overview of the entire ML model is shown in Figure 2. The datasets retrieved from existing complete experimental curves are prepared and fed into an ML regression model. This model comprises five regression learner models: a linear regression model, a support vector machine (SVM), a Gaussian process regression (GPR), a regression tree, and an artificial neural network (ANN). The prepared data are separately trained by these algorithms for prediction, and the algorithm with the lowest RMSE is selected for the final prediction in a forward stepwise prediction model.

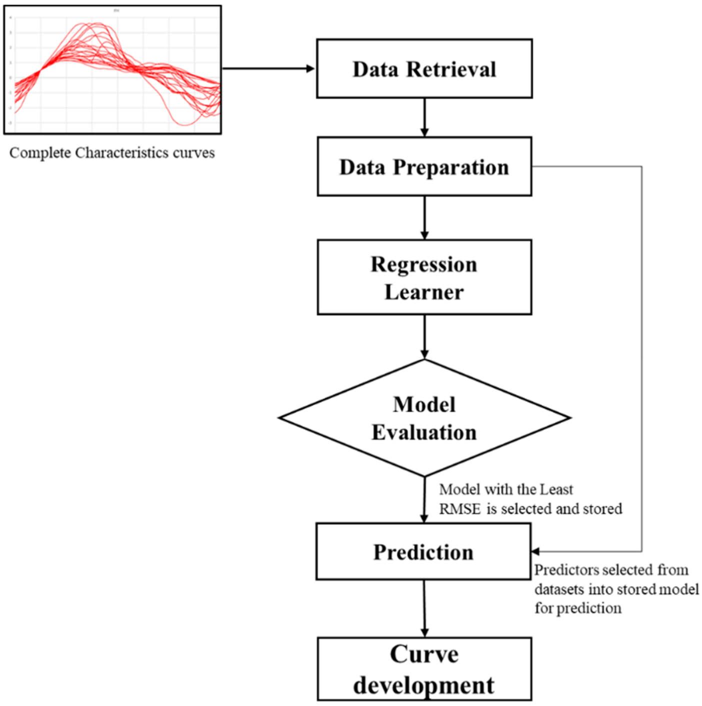


Figure 2. Overview of ML model.

3.1. Data Sourcing and Preparation

The datasets are sourced from complete experimental characteristic curves compiled and published by Trey et al. [24]. These experiments are designed and performed in a laboratory setting using a test rig and various other apparatuses, such as torque, flow, pressure, and power meters, to measure the rate of change in the pump head and torque of real model pumps. The results of the experiments are plotted on graphs as the pump dimensionless head and torque against theta [9,25–29]. The total number of complete curves selected for the ML experiment is 28, with the specific speeds ranging from 15.7 to 261.6. The data points of a stepwise increment of 5 units of θ are chosen from each respective curve. Given each curve, a total of 28 observations with 74 variables are recorded, as shown in Figure 3. The dataset is then first normalized using the column-wise normalization of the specific speeds by dividing them with the respective absolute maximum to attain a range of 0 to 1 for all data in the specific speed column. The same is carried out for the curve data points (Theta0–Theta360), but they are normalized row-wise. This routine of data preparation is carried out for both the Wh and Wb curves.

Variables

Observations

No.	Pump Specific Speed	Theta0	Theta5	Theta10	· · ·	Theta360
1	0.060015	−0.2953	−0.2282	−0.1544	· · ·	−0.2953
2	0.076453	−0.3295	−0.2659	−0.1629	· · ·	−0.3295
3	0.078364	−0.3438	−0.2890	−0.1641	· · ·	−0.3438
⋮	⋮	⋮	⋮	⋮	⋮	⋮
28	1	−0.2685	−0.2662	−0.2556	· · ·	−0.6789

Figure 3. Representation of prepared dataset for training.

3.2. Regression Learners

These learners find the relationship existing between dependent and independent variables under a certain error function. The obtained relationship is used to make predictions on a new set of data. The ML model in this paper trains the dataset using the 16 different ML algorithms categorized below.

3.2.1. Linear Regression

During training, the linear and interaction linear regression models were found to be more accurate than the other types of linear regression models. Interaction linear regression models differ from linear regression by establishing interaction terms between the predictors, in addition to the constant and linear terms derived in a linear regression model.

3.2.2. Support Vector Machine (SVM)

The SVM algorithm was first developed by Cortes and Vapnik [30], and through calculus, Lagrange multipliers and vector geometry have evolved into some of today's most popular ML tools. SVM regression relies on kernel functions and is hence regarded as a nonparametric approach. The kernel functions considered for this ML experiment were linear, Gaussian, and quadratic, as shown in Equations (4)–(6), respectively [31], where x_j and x_k are predictors.

$$G(x_j, x_k) = x_j' x_k \tag{4}$$

$$G(x_j, x_k) = \exp\left(- \| x_j - x_k \|^2\right) \tag{5}$$

$$G(x_j, x_k) = (1 + x_j' x_k)^2 \tag{6}$$

3.2.3. Gaussian Process Regression (GPR)

This is a stochastic probabilistic model based on Gaussian distribution. It is also a nonparametric mainly relying on covariance functions to define the behavior of the model. The covariance functions selected are the squared exponential, Matern, exponential, and rational quadratic, as seen in Equations (7)–(10), respectively [32]:

$$K_{SE}(r) = \exp\left(-\frac{r^2}{2\ell^2}\right) \tag{7}$$

$$K_{Matern}(r) = \frac{2^{1-v}}{\Gamma(v)}\left(\frac{\sqrt{2vr}}{\ell}\right)^{v} K_v\left(\frac{\sqrt{2vr}}{\ell}\right) \tag{8}$$

$$K(r) = \exp\left(-\frac{r}{\ell}\right) \tag{9}$$

$$K_{RQ}(r) = \exp\left(1 + \frac{r^2}{2\alpha\ell^2}\right)^{-\alpha} \tag{10}$$

where $r = x_k - x_j'$, ℓ measures the characteristic length scale between the data points, K_v is the modified Bessel function of order v evaluated in gamma function $\Gamma(v)$, and $\alpha \geq 0$.

3.2.4. Artificial Neural Network (ANN)

Four ANNs, namely, feedforward three-layered, bilayer, narrow, and wide ANN models, were added to the regression learners. ANNs are great algorithms for data mining due to their architecture of layers with interconnecting nodes used for prediction.

3.2.5. Regression Tree

A fine regression tree with a minimum leaf size of 4 without a surrogate decision split is used in an iterative operation to split the dataset into tree branches as the operand moves up the tree via binary recursive partitioning. This binary recursive split is decided based on the RMSE, and then the algorithm selects the branch that minimizes the RMSE.

3.3. Curve Prediction and Development

The ML code is developed using MATLAB on a laptop with a Windows 10 operating system, an Intel i5 processor, 8 GBs read-access memory, and Nvidia GeForce GTX 1650 Max-Q. The experiment is conducted using a 5-fold cross-validation and performance test, validated in terms of accuracy using RMSE. During the training and testing of the datasets, the parameters suitable for predicting the lowest RMSE for each model are established. After which, the best model with regard to the RMSE for each column is selected and stored in a struct. The stored models are then called in a forward stepwise prediction model for the complete curve. The prediction model takes a sequence of curve data points in a block as input to predict the subsequent data point after the block. This block consists of the datasets from the known part of the characteristic curves. This continues until the last data point is reached. Figure 4 demonstrates this prediction model. It is worth noting that the number of predictors selected from the block for each theta model is not fixed and depends on the sequence of datasets within the block with the lowest RMSE.

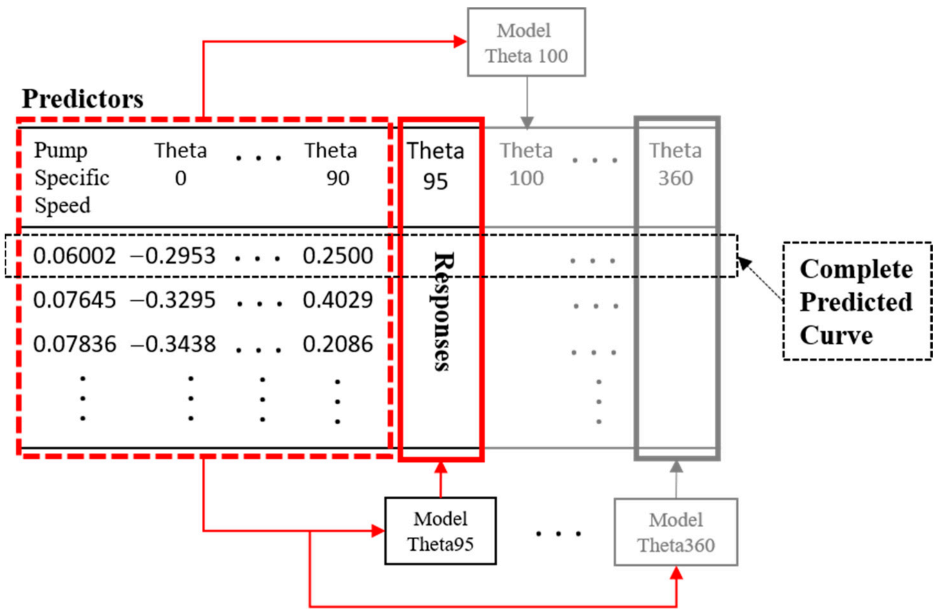
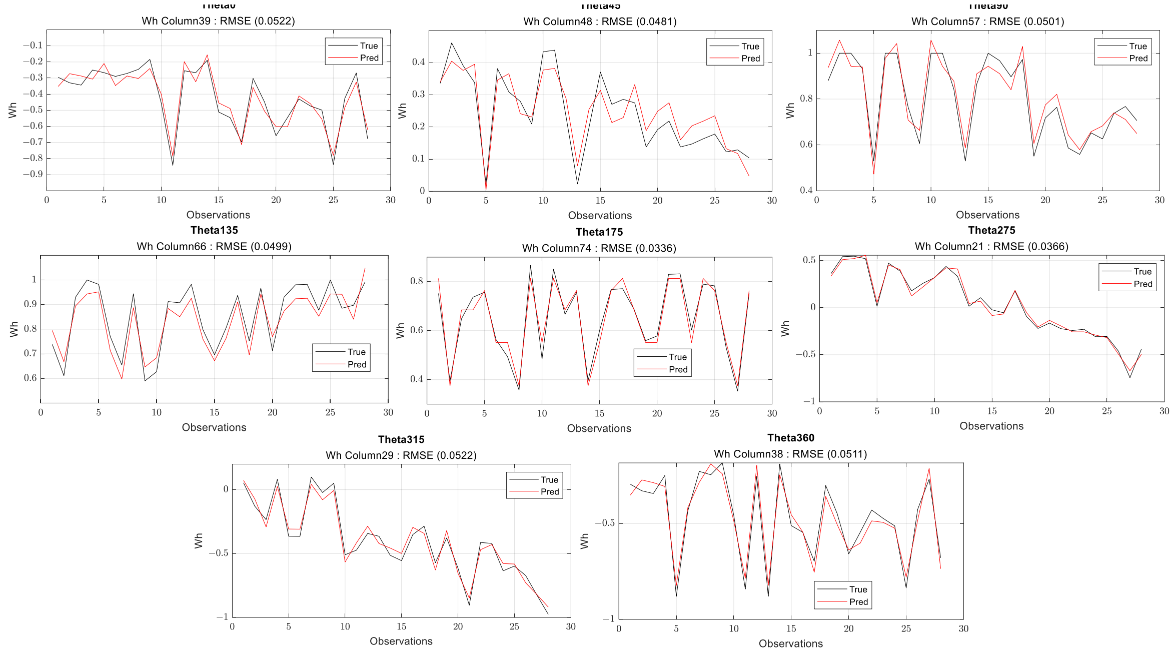


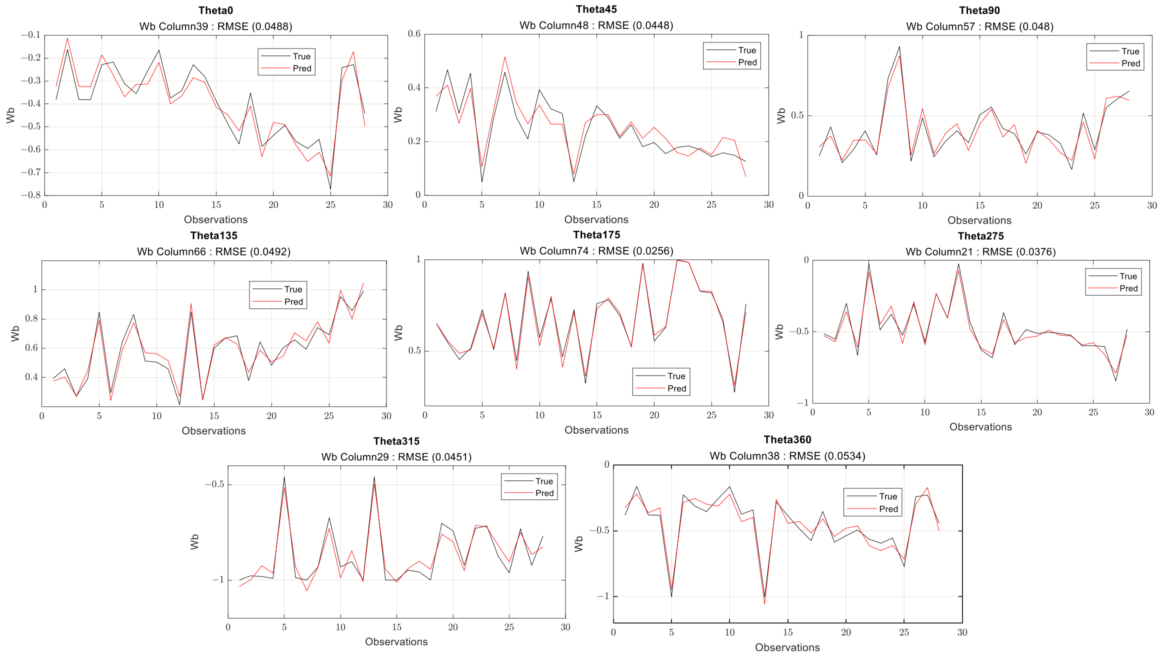
Figure 4. ML prediction model.

4. Results and Discussion

As already mentioned in the Introduction, the pump data provided by most manufacturers are only good enough to determine the third quadrant, which comprises zones A to C, as tabulated in Table 1. Moreover, since the ML prediction model was developed to move in the forward direction, there was a need to rearrange the dataset in the order of pump operational zones A–H, as illustrated in Figure 1. This meant that the third quadrant was first, followed by the fourth, the first, and the second. Figure 5 shows the training test validation results across the entire observations respective to theta (0, 45, 90, 135, 175, 275, 315, 360) for both the dimensionless head and torque. The results of the training test validation averaged an RMSE of 0.0484 with a standard deviation of 0.0047 for the dimensionless head and an RMSE of 0.0472 with a standard deviation of 0.0042 for the dimensionless torque. The size of the standard deviation for both the dimensionless head and torque demonstrates the robustness of the ML model across all three quadrants. The performance of the implemented ML model was evaluated across a spectrum of centrifugal pumps categorized by specific speeds, namely, a radial pump (<40), a mixed-flow pump (between 40 and 175), and an axial pump (>175). The third quadrant dataset of a complete curve selected from each of the mentioned pump categories was fed into the exported model for test validation. Tables 2 and 3 show the selected specific speeds with their respective third-quadrant input datasets.



(**a**)



(**b**)

Figure 5. (**a**) Normalized dimensionless head test validation during training for all observations for column theta (0, 45, 90, 135, 175, 275, 315, 360). (**b**) Normalized dimensionless torque test validation during training for all observations under theta (0, 45, 90, 135, 175, 275, 315, 360).

Pump-specific speeds are a key factor in describing a pump's performance, but they are not the only factor since different types of pumps can have the same specific speed. Therefore, an instance is added to the input dataset to test the ML model's response in these scenarios.

Now, these input data points become the block or pool of datasets that the ML model fetches input data from to generate the datasets of Quadrants I, II, and IV. The normalization of the predicted datasets is reversed, and the dataset is rearranged from theta 0 to

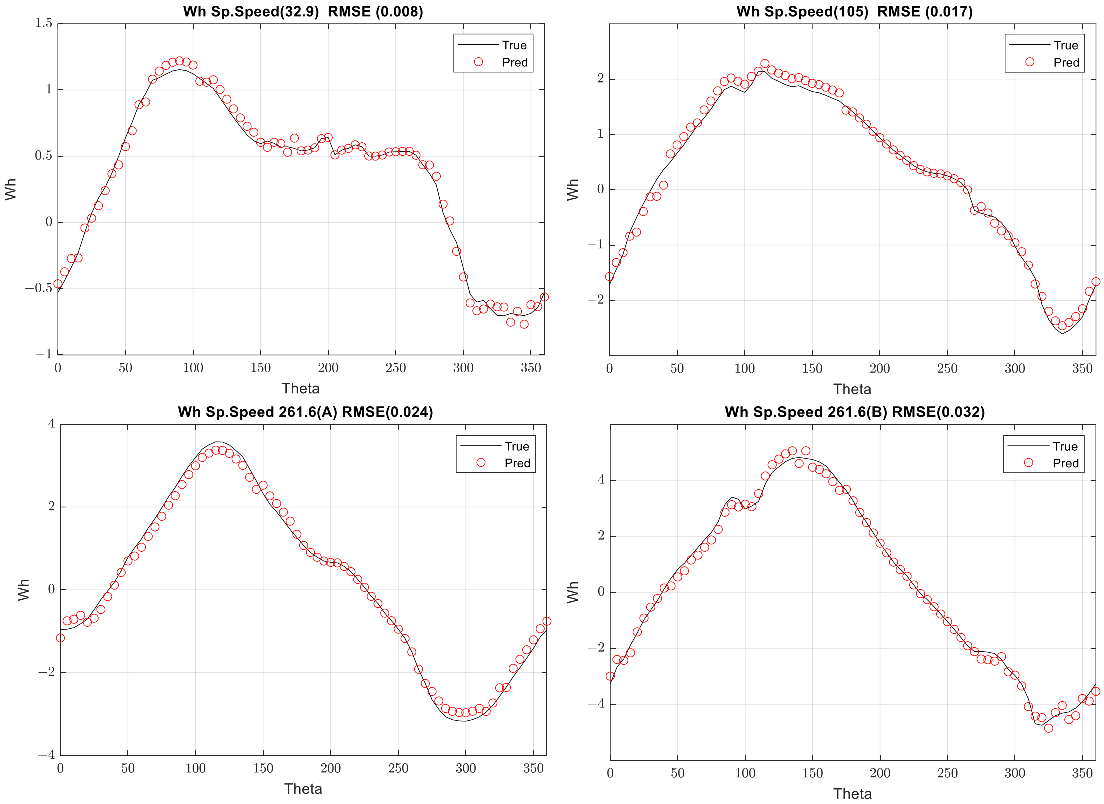
360 and then plotted against the actual curve, as seen in Figure 6. The graph results show that the ML model is able to predict the change in the dimensionless head and torque over theta with great accuracy, recording a maximum RMSE of 0.032. It is also able to significantly differentiate between pumps of the same specific speed. The input data have a fixed maximum variable count of 20, made up of the specific speed and Quadrant III datasets, as seen in Tables 2 and 3. However, during the training and column response model generation, the number of predictors or variables required for each model may vary but would not exceed 20. The ML model for each column selects a sequence of predictors starting from the specific speed variable to theta270, the 20th variable that computes the lowest RMSE. For instance, the prediction model for the dimensionless torque column theta355 model requires 16 predictors (that is, a sequence of data points from the specific speed to column theta250), whilst the theta360 prediction model requires an input of 10 predictors (that is, a sequence of data points from the specific speed to column theta220) to attain the lowest RMSE. From this, it is clear that each column theta model according to the training results of the datasets might have different requirements for the number of predictors and the type of regression model to be adopted. In the 54-column models generated from the training of the dimensionless torque datasets, the average number of predictors per column is 17, with the SVM regression model selected for 96.3% of the 54 columns. The SVM regression model is similarly favored in the dimensionless head ML model.

Table 2. Normalized dimensionless head third-quadrant input datasets for test validation.

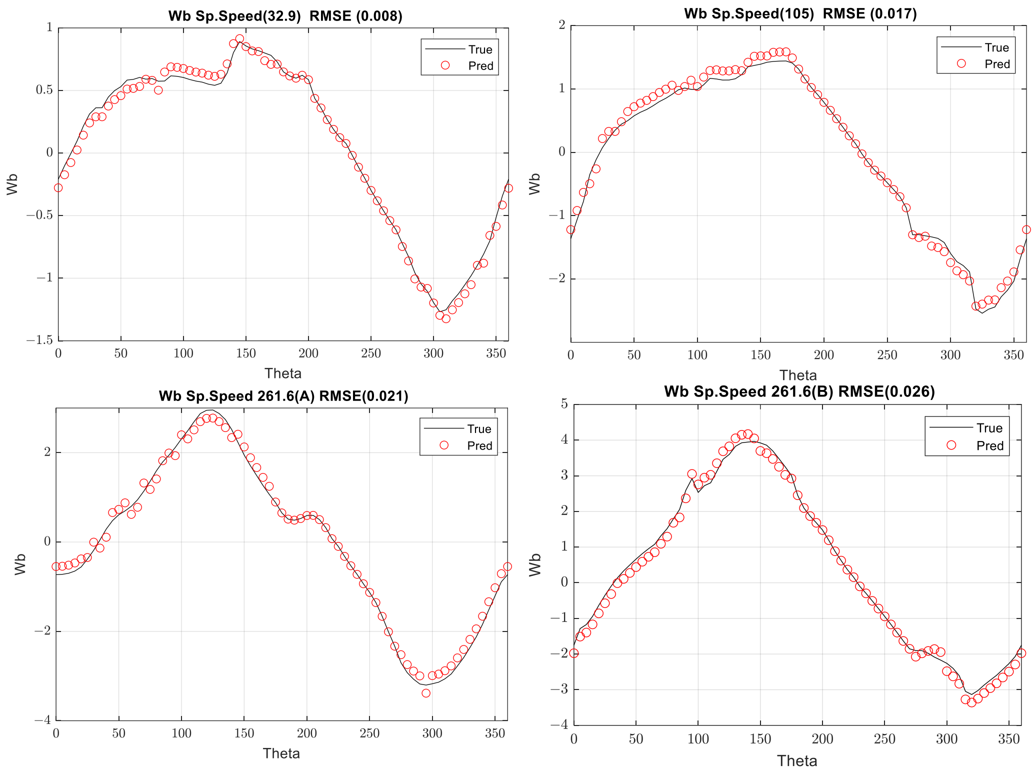
Specific Speed	Theta 180	Theta 185	Theta 190	Theta 195	Theta 200	Theta 205	Theta 210	Theta 215	Theta 220
32.9	0.4683	0.4735	0.4883	0.5482	0.5557	0.4420	0.4734	0.4847	0.5077
105	0.5390	0.4978	0.4546	0.4051	0.3608	0.3170	0.2762	0.2385	0.2039
261.6	0.2998	0.2527	0.2202	0.1944	0.1854	0.1820	0.1579	0.1226	0.0704
261.6	0.6785	0.5911	0.5170	0.4397	0.3641	0.2923	0.2237	0.1684	0.1187
Theta 225	**Theta 230**	**Theta 235**	**Theta 240**	**Theta 245**	**Theta 250**	**Theta 255**	**Theta 260**	**Theta 265**	**Theta 270**
0.4952	0.4345	0.4339	0.4411	0.4606	0.4625	0.4647	0.4655	0.4396	0.3777
0.1665	0.1410	0.1237	0.1140	0.1096	0.0959	0.0764	0.0504	−0.0015	−0.1428
0.0171	−0.0441	−0.0929	−0.1574	−0.2090	−0.2657	−0.3291	−0.4194	−0.5367	−0.6331
0.0540	−0.0102	−0.0545	−0.1072	−0.1616	−0.2177	−0.2757	−0.3354	−0.3970	−0.4404

Table 3. Normalized dimensionless torque third-quadrant input datasets for test validation.

Specific Speed	Theta 180	Theta 185	Theta 190	Theta 195	Theta 200	Theta 205	Theta 210	Theta 215	Theta 220
32.9	0.5116	0.4860	0.4719	0.4904	0.4624	0.3447	0.2843	0.2104	0.1494
105	0.5174	0.4560	0.4025	0.3583	0.3107	0.2610	0.2087	0.1559	0.1028
261.6	0.2037	0.1602	0.1528	0.1642	0.1854	0.1865	0.1568	0.1007	0.0235
261.6	0.6191	0.5289	0.4718	0.4241	0.3730	0.3012	0.2232	0.1567	0.0929
Theta 225	**Theta 230**	**Theta 235**	**Theta 240**	**Theta 245**	**Theta 250**	**Theta 255**	**Theta 260**	**Theta 265**	**Theta 270**
0.0973	0.0603	−0.0144	−0.0888	−0.1589	−0.2354	−0.3008	−0.3638	−0.4259	−0.4840
0.0531	−0.0099	−0.0641	−0.1109	−0.1473	−0.1888	−0.2315	−0.2760	−0.3449	−0.5123
−0.0303	−0.1001	−0.1659	−0.2259	−0.2917	−0.3524	−0.4233	−0.5194	−0.6276	−0.7282
0.0393	−0.0258	−0.0751	−0.1283	−0.1826	−0.2381	−0.2948	−0.3527	−0.4120	−0.4675



(**a**)



(**b**)

Figure 6. (**a**) Predicted dimensionless head complete pump characteristic curve plotted against actual experimental curve with respect to Table 2. (**b**) Predicted dimensionless torque complete pump characteristic curve plotted against actual experimental curve with respect to Table 3.

5. Conclusions

In this paper, a supervised ML model is proposed to measure and predict the relationship between Quadrant III data points and those of Quadrants I, II, and IV of a centrifugal performance characteristic curve. The centrifugal pump details provided by manufacturers are only enough to compute the pump characteristic curve in Quadrant III. A four-quadrant characteristic curve is vital to the investigation of the hydraulic transient phenomenon in pumps and any other installation surrounding it, and, hence Quadrants I, II, and IV can never be overlooked if accurate transient results are to be obtained. Moreover, current methods, such as CFD and other prediction methods, are either laborious, expensive, or incomplete. The ML model proposed in this study computes complete characteristic curves and the dimensionless head and torque curves from the dataset of Quadrant III with great accuracy, as demonstrated in Figure 6. Hence, the observations from this work open up a new way of looking at four-quadrant data, as the high accuracy of the prediction emphasizes the capability of ML tools to measure the relationship between the data points of four-quadrant curves. Further research should focus on the impact of operating pumps below the best efficiency point on the accuracy of pump characteristic prediction.

Author Contributions: Supervision J.Y.; conceptualization, E.A.; methodology, E.A., D.K.D. and L.H.; data curation, E.A. and L.H.; writing—original draft preparation, E.A.; writing—review and editing, J.Y., E.A., D.K.D. and L.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The dataset used in this experiment is available and published in [24].

Acknowledgments: The corresponding author of this paper would respectfully like to thank the Chinese Government Scholarship Council (CSC) for fully sponsoring his graduate studies and thanks to the Major Science and Technology Project in Gansu Province (Project No. 19ZD2GA004), the Key Research and Development program in Gansu Province (Project No. 20YF8GA013), and the Natural Science Foundation of Gansu Province (Project No. 1610RJZA029) for their support.

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

References

1. Wan, W.; Huang, W. Investigation on complete characteristics and hydraulic transient centrifugal pump. *J. Mech. Sci. Technol.* **2011**, *25*, 2583–2590. [CrossRef]
2. Wylie, E.B.; Streeter, V.L. *Fluid Transients*; McGraw-Hill International Book Company: New York, NY, USA, 1978.
3. Thanapandi, P.; Prasad, R. Centrifugal pump transient characteristics and analysis using the method of characteristics. *Int. J. Mech. Sci.* **1995**, *37*, 77–89. [CrossRef]
4. Marchal, M.; Flesh, G.; Suter, P. The calculation of water hammer problems by means of the digital computer. In Proceedings of the International Symposium on Water hammer Pumped Storage Projects, ASME, Chicago, IL, USA, 7–11 November 1965.
5. Suter, P. Representation of pump characteristics for calculation of water hammer. *Sulzer Tech. Rev.* **1966**, *66*, 45–48.
6. Dörfler, P.K. 'Neo-Suterian' pump-turbine characteristics, and their benefits. *Inst. Phys. Conf. Ser. Earth Environ. Sci.* **2010**, *12*, 012027.
7. Stepanoff, A.J. *Centrifugal and Axial Flow Pumps*; Springer-Verlag Berlin Heidelberg GmbH: New York, NY, USA, 1959.
8. Donsky, B. Complete pump characteristics and the effects of specific speeds on hydraulic transients. *J. Basic Eng. Trans ASME* **1961**, *83*, 685–699. [CrossRef]
9. Knapp, R.T. Complete characteristics of centrifugal pumps and their use in the prediction of transient behavior. *Trans ASME* **1937**, *59*, 683–689.
10. Giljen, Z.; Nedeljkovic, M.; Cheng, Y.G. Pump-turbine characteristics for analysis of unsteady flows. In Proceedings of the 28th IAHR Symposium on Hydraulic Machinery and Systems, Grenoble, France, 4–8 July 2016; Volume 49.
11. Pejovic, S.; Bryan, K. Guidelines for transients are in need of revision. In Proceedings of the 27th IAHR, Symposium on Hydraulic Machinery and Systems, Montreal, Canada, 22–26 September 2014; Volume 22, pp. 1155–1165.
12. Martin, C.S. *Pumped Storage, 12.14 Section of Pump Handbook*; McGraw-Hill Companies: New York, NY, USA, 2008.
13. Fox, J.A. *Hydraulic Analysis of Unsteady Flow in Pipe Networks*; The MacMillan Press Ltd.: London, UK, 1977.
14. Walters, T.W.; Lang, S.A.; Miller, D.O. Unappreciated challenges in applying four quadrant pump data to water hammer simulation part 1: Fundamentals, and part 2: Application examples. In Proceedings of the 13th International Conference on Pressure Surges, Bordeaux, France, 14–16 November 2018; BHR Group: Bordeaux, France, 2018; Volume 1, pp. 741–769.

15. Gros, L.; Couzinet, A.; Pierrat, D.; Landry, L. Complete pump characteristics and 4-quadrant representation investigated by experimental and numerical approaches. In Proceedings of the ASME-JSME-KSME 2011 Joint Fluids Engineering Conference, Hamamatsu, Japan, 24-29 July 2011; Volume 1, pp. 359–368.
16. Muttalli, R.S.; Agrawal, S.; Warudkar, H. CFD simulation of centrifugal pump impeller using ANSYS-CFX. *Int. J. Innov. Res. Sci. Eng. Technol.* **2014**, *3*, 15553–15561.
17. Höller, S.; Benigni, H.; Jaberg, H. Investigation of the 4-Quadrant behaviour of a mixed flow diffuser pump with CFD-methods and test rig evaluation. *Inst. Phys. Conf. Ser. Earth Environ. Sci.* **2016**, *49*, 032018. [CrossRef]
18. Frosina, E.; Buono, D.; Senatore, A. A performance prediction method for pumps as turbines (PAT) using a computational fluid dynamics (CFD) modeling approach. *Energies* **2017**, *10*, 103.
19. Wang, L.; Li, M.; Wang, F.J.; Wang, J.B.; Yao, C.G.; Yu, Y.S. Study on three-dimensional internal characteristics method of Suter curves for double-suction centrifugal pump. *J. Hydraul. Eng.* **2017**, *48*, 113–122.
20. Shao, W.Y.; Zhang, X. A new simulation method of complete characteristic curves of reversible pump turbine moving least square approximation. *J. Hydroelectr. Eng.* **2004**, *23*, 102–106.
21. Zhang, L.; Xu, H.; Yu, Y.H. Fitting method for pump complex characteristic curve based on B-spline. *Drain. Irrig. Mach.* **2007**, *25*, 50–53.
22. Han, W.; Nan, L.; Su, M.; Chen, Y.; Li, R.; Zhang, X. Research on the prediction method of centrifugal pump performance based on a double hidden layer BP neural network. *Energies* **2019**, *12*, 2709. [CrossRef]
23. Li, H.; Lin, H.; Huang, W.; Li, J.; Zeng, M.; Ma, J.; Hu, X. A New Prediction Method for the Complete Characteristic Curves of Centrifugal Pumps. *Energies* **2021**, *14*, 8580. [CrossRef]
24. Walters, T.W.; Dahl, T.; Rogers, D. Pump Specific Speed and Four Quadrant Data in Waterhammer Simulation—Taking Another Look" auxiliary data files. In Proceedings of the ASME 2020, Pressure Vessels and Piping Conference, Online, 3 August 2020; Volume 4. PVP2020-21069. Available online: https://www.aft.com/technical-papers/pump-specific-speed-and-four-quadrant-data-in-waterhammer-simulation-taking-another-look (accessed on 23 October 2022).
25. Brown, R.J.; Rogers, D.C. Development of Pump Characteristics from Field Tests. *ASME J. Mech. Des.* **1980**, *102*, 807–817.
26. Ayder, E.; Ilikan, A.N.; Şen, M.; Özgür, C.; Kavurmacıoğlu, L.; Kirkkopru, K. Experimental Investigation of the Complete Characteristics of Rotodynamic Pumps. In Proceedings of the ASME Fluids Engineering Division Summer Meeting, Vail, CO, USA, 2–6 August 2009; pp. 35–40, FEDSM2009-78052.
27. Kittredge, C.P. Hydraulic Transients in Centrifugal Pump Systems. *ASME J. Basic Eng.* **1956**, *78*, 1307–1322. [CrossRef]
28. Thorley, A.R.D.; Chaudry, A. "Pump Characteristics for Transient Flow, Pressure Surges and Fluid Transients. In Proceedings of the 7th International Conference on Pressure Surges, Harrogate, UK, 16–18 April 1996; BHR Group: Harrogate, UK, 1996.
29. Swanson, W.M. Complete Characteristic Circle Diagrams for Turbo-machinery. *Trans. ASME* **1953**, *75*, 819–826. [CrossRef]
30. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995.
31. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*, 2nd ed.; Springer: New York, NY, USA, 2008.
32. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; The MIT Press: London, UK, 2006; ISBN 026218253X.