Cole-Puli
International Conference Publishing Group

Science, Engineering and Technology Proceedings

Article

2025 International Conference on Science Technology, Architecture, Power and Intelligent Information Technology (APIIT 2025)

A Fault Diagnosis Method for the Charging and Discharging Water System of a Certain Ship Lift Based on Historical Data Analysis

Binbin Zou [1,*], Yansong Lei [1], Gengbang Huang [1] and Lingzhi Yan [1]

[1] Three Gorges Navigation Authority, Yichang, Hubei, China
* Correspondence: Binbin Zou, Three Gorges Navigation Authority, Yichang, Hubei, China

Abstract: This paper takes the hydraulic action of the discharge pipeline in the filling and discharging water system of a Certain ship lift as a case study. Based on analysis of operational data from different historical stages, a mathematical model is established to determine whether the equipment in the filling and discharging water system shows trends toward abnormal conditions. The proposed model will provide early warnings for equipment maintenance and reduce the frequency of timeout fault alarms for pipeline water extraction.

Keywords: timeout of discharge pipeline water; data analysis; regression linear equation

Received: 26 April 2025
Revised: 01 May 2025
Accepted: 18 May 2025
Published: 11 June 2025

1. Introduction

The water filling and discharging system is a system used by the ship lift to adjust the water depth of the cabin, and its main functions are compartment drainage, compartment water replenishment and pumping pipeline water, wherein pumping pipeline water action is an inevitable step in the operation of each compartment, taking a ship lift as an example, the pumping pipeline water running time is about 10-13 minutes, accounting for about 50% of the total running time. At present, the failure of the water filling and discharging system is mainly concentrated in the pumping pipeline water link, when the pumping pipeline water rate decreases, the pumping pipeline water duration increases, which will affect the connection of the corresponding process, and even alarm shutdown.

During the operation of the ship lift, many factors influence pipeline water extraction time, such as the water volume in the interstitial pipeline, the valve opening size, and the power output of the pump. Any abnormality in these factors may lead to a timeout failure in the water extraction process. However, the water volume in the interstitial pipeline is a factor that cannot be controlled by maintenance personnel. Its volume is affected by the interstitial water depth behind the primary gate valve and the ship chamber gate, and it has an upper limit. Long-term observation shows that even under the upper limit, pipeline water extraction can still be completed within the specified time when the equipment is in normal condition. Therefore, the operational condition of pipeline water extraction equipment itself has a fundamental impact on whether a failure occurs. Since the valve opening size and the pump power output do not have feedback values transmitted to the upper-level control system and the data server, the operational status of such equipment has become a blind spot for maintenance personnel. Thus, it is necessary to use indirect methods, specifically historical data analysis, to determine whether there is a trend of deterioration in the operational status of this type of equipment.

2. Sorting Out the Operation Process of Pumping Pipeline Water

First, sort out the interlocking mechanism of each device during the process of extracting pipeline water. After the water in the gap between the primary gate valve and the ship chamber gate leaks into the interstitial pipeline, the water is stored in the interstitial pipeline. Once the docking seal frame retracts, the pipeline water extraction and the drive mechanism start synchronized operations. Second, the corresponding butterfly valve of the filling and discharging system opens. After the valve is fully opened, the electric control device controls the pump to start extracting water. Since the interstitial pipeline, the filling and discharging system and the ship chamber are arranged from bottom to top, a check valve is installed in the filling and discharging pipeline to prevent backflow during the valve opening process, ensuring that water from the ship chamber does not flow into the interstitial pipeline. This also protects the pump impeller from reversing.

When the water level in the interstitial pipeline reaches the specified level, the PLC (Programmable Logic Controller) controls the pump to stop, the corresponding butterfly valve to close, and the check valve to return to its initial state. Once all actions are completed, the PLC outputs a "pipeline water extraction completed" signal.

Then, analyze the direct factors of each device affecting the duration of pipeline water extraction. The analysis results are shown in Table 1:

Table 1. Factors Affecting Pipeline Water Extraction Duration.

Device	Interstitial Pipeline	Gate Valve	Pump	Check Valve
Factors	Water in the interstitial pipeline	Valve opening and closing time	Pump output power	Check valve opening degree

Among the factors affecting the duration of pipeline water extraction mentioned above, the pump output power and the check valve opening do not provide feedback values due to equipment limitations. Consequently, there is no corresponding database, making it difficult for operators to monitor their working states. Therefore, we can analyze the operational state of such equipment inversely through the results. Here, the pump power and the check valve opening degree both affect the flow rate per unit of time and can be treated as a single entity.

The water in the interstitial pipeline, one of the influencing factors, cannot be directly measured but can be converted based on the water level in the interstitial pipeline. However, after interstitial water is discharged, the water in the interstitial pipeline fluctuates significantly under the impact force, and the fluctuations are relatively complex. As a result, the readings of the interstitial pipeline water level are prone to considerable errors. Therefore, an alternative approach must be identified.

By reviewing the operation process of the Three Gorges Ship Lift, two alternative solutions were analyzed:

1) Use the gap water depth value recorded after the primary gate valve and ship chamber gate are fully closed but before the gap discharge valve is opened.
2) Use the difference in the ship chamber water depth values before and after pipeline water extraction.

However, the second method is affected by factors such as strong winds, which can cause fluctuations in the ship chamber water level. Additionally, the large area of the ship chamber amplifies any slight deviations in readings, leading to significant errors in the calculated results.

Therefore, the first alternative is adopted: using the gap water depth value recorded after the primary gate valve and ship chamber gate are fully closed but before the gap discharge valve is opened.

In summary, the key data to be collected are the gap water depth value, the completion time of the opening of the shut-off valve, and the time when the shut-off valve starts to close the valve.

3. Pipeline Water Extraction Data Collection

During the pipeline water extraction data collection process, incomplete data may occasionally be encountered. This can be caused by data loss, an incomplete operation of a Certain Ship Lift, or a process interruption due to a malfunction. The collection of such incomplete data will affect the output results. Therefore, after data collection, it is necessary to identify and filter these data [1].

The filtering method is as follows:

First, organize the action time points for collecting this information, as shown in Figure 1:

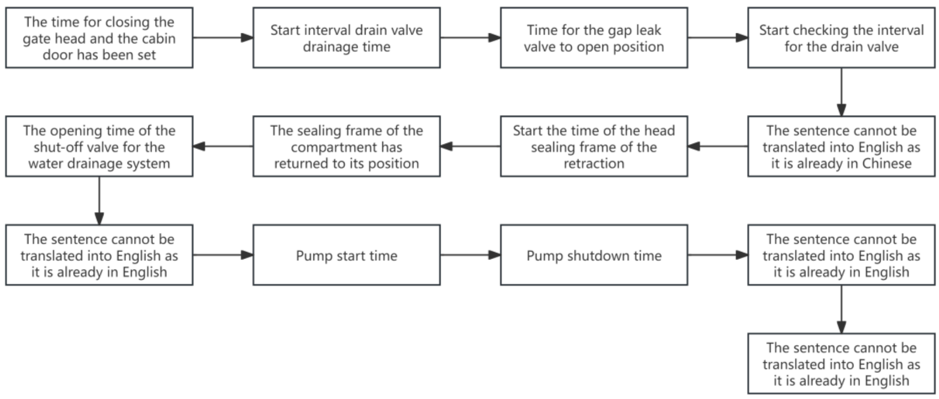

Figure 1. Flowchart of Mechanism Action Time Points.

Through analysis of the operational time data of each mechanism, the minimum and maximum values within the normal range are determined. Adding the minimum values and the maximum values, we conclude that the time from the butterfly valve of the filling and discharging system being fully closed to the primary gate valve and the ship chamber gate being fully closed should be no less than 780 seconds and no more than 960 seconds. This range is used as the criterion for determining the completeness of the entire data set.

If the above conditions are not met, the data is considered invalid and excluded from the data analysis. Although a very small number of valid data points may be misclassified as invalid during this process, it does not affect the overall data analysis results.

4. Pipeline Water Extraction Data Conversion and Analysis

4.1. Data Conversion

After collecting a large amount of complete data and considering the trigger conditions for pipeline water extraction timeout failures, the collected data needs to be converted accordingly as follows [2]:

Since the water volume in the interstitial pipeline equals the volume of the discharged interstitial water, and the discharged water volume is related to the interstitial water height (the higher the height, the larger the water volume), the interstitial water level is generally above 9m. The increase in water volume above 9m rises in quadratic growth with the height difference, but the height increment does not exceed 0.7 m. Therefore, the relative increase in the total interstitial water volume is minimal. This model can approximately be considered ideal, meaning the water volume in the gap water pipe is directly proportional to the gap water level height. Hence, in the formula, the volume of the interstitial water pipe can be replaced by the interstitial water level height.

The water extraction time is calculated as the difference between the time when the gate valve starts to close and the time when the gate valve finishes opening.

4.2. Data Analysis

Using actual collected data as an example, this paper continuously collected pipeline water extraction data from August and December 2020. After data screening and conversion, the processed data is shown in Tables 2 and 3:

Table 2. Pipeline Water Extraction Data in August 2020.

Interstitial Water Depth (m)	**9.02**	**9.37**	**9.171**	**9.212**	**9.218**	**9.295**	**9.296**	**9.335**	**9.338**	**9.467**
Extraction Time (s)	617	623	697	695	731	737	722	737	659	776
Interstitial Water Depth (m)	9.542	9.545	9.547	9.579	9.6	9.641	9.692	9.695	9.73	
Extraction Time (s)	683	693	651	727	751	703	759	727	747	

Table 3. Pipeline Water Extraction Data in December 2020.

Interstitial Water Depth (m)	**9.109**	**9.157**	**9.186**	**9.212**	**9.302**	**9.36**	**9.383**	**9.483**
Extraction Time (s)	743	739	744	746	752	741	754	748
Interstitial Water Depth (m)	9.574	9.582	9.595	9.611	9.619	9.636		
Extraction Time (s)	741	772	773	754	763	736		

The two sets of data above are presented on the same scatter plot, as shown in Figure 2:

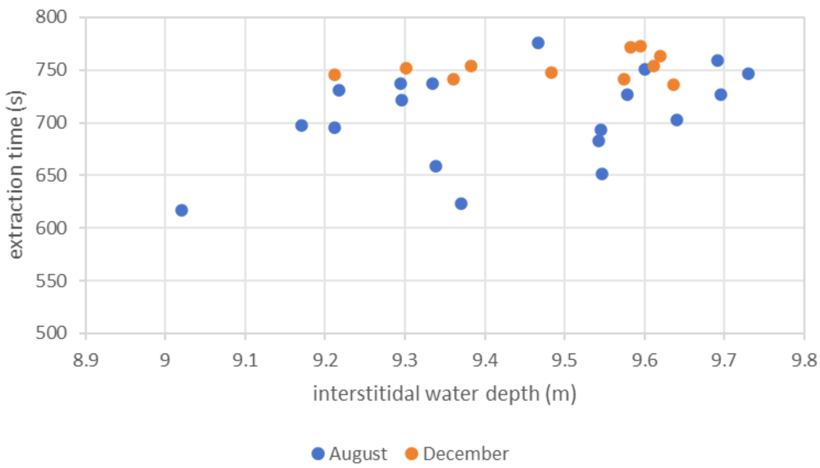

Figure 2. Scatter Plot of Interstitial Water Depth vs. Extraction Time.

From the scatter plot, it can be observed that the data points for August are significantly lower than those for December. This indicates that for the same interstitial water depth, the pipeline water extraction time in August is noticeably shorter than in December. Conversely, for the same extraction time, the interstitial water depth extracted in August is significantly higher than in December. This suggests that the extraction rate in August is higher than that in December.

In practical applications, these variables are collected continuously. We need an algorithm to convert the observed phenomena into one or two numerical outputs, followed by setting thresholds [3]. If the output exceeds the threshold, it indicates a deteriorating trend in equipment performance, prompting maintenance personnel to investigate and resolve the issue promptly.

From the scatter plot, it can be roughly inferred that there may be a linear relationship between extraction time and interstitial water depth [4]. Assuming a linear relationship exists, linear regression equations are used for analysis.

Let the gap water depth samples be X, with each sample value denoted as x_i, and the extraction time samples be Y, with the corresponding values denoted as y_i. The formula for the least squares method is as follows:

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2 \quad (1)$$

$$S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{1}{n}(\sum_{i=1}^{n} y_i)^2 \quad (2)$$

$$S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{1}{n}(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i) \quad (3)$$

$$\hat{b} = \frac{S_{xy}}{S_{xx}} \quad (4)$$

$$\hat{a} = \frac{1}{n}\sum_{i=1}^{n} y_i - (\frac{1}{n}\sum_{i=1}^{n} x_i)\hat{b} \quad (5)$$

Using the least squares method, the data for August and December 2020 were calculated separately, with the results shown in Table 4:

Table 4. Least Squares Method Calculation Data.

Data Type	S_{xx}	S_{yy}	S_{xy}	$\hat{b}$	$\hat{a}$
August	0.8826	36029.7894	96.7239	109.58	324.89
December	0.4897	1799.4285	13.8964	28.375	483.28

Therefore, the regression line equation for August 2020 is:

$$\hat{y} = 324.89 + 109.58x \quad (6)$$

The regression line equation for December 2020 is:

$$\hat{y} = 483.28 + 28.375x \quad (7)$$

Estimation of σ^2 for the regression line equation of August 2020:
Sum of squared residuals

$$Q_e = S_{yy} - \hat{b}S_{xy} = 36029.7894 - 109.58 \times 96.7239 = 25430.7844 \quad (8)$$

$$\hat{\sigma}^2 = Q_e/(n-2) = 25430.7844/(19-2) = 1495.9284 \quad (9)$$

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{1495.9284} = 38.6772 \quad (10)$$

Significance test of the linear hypothesis for the regression line equation of August 2020:

$$|t| = \frac{|\hat{b}|}{\hat{\sigma}}\sqrt{S_{xx}} = \frac{109.58}{38.6772}\sqrt{0.8826} = 2.6617 \geq t_{\alpha/2}(n-2) = t_{0.025}17 = 2.1098 \quad (11)$$

That is, the regression line equation is significant.
The confidence interval for the coefficient \ (b \) is:

$$\left(\hat{b} \pm t_{\alpha/2}(n-2) \times \frac{\hat{\sigma}}{\sqrt{S_{xx}}}\right) = (22.724,196.436) \quad (12)$$

As can be seen from Table 4, in December, $\hat{b}$ is still within the confidence interval of coefficient b, but the coefficient value is significantly lower.

Based on actual operations, the data collected in August 2020 represents normal operating conditions, whereas in December, the a Certain ship lift frequently reported pipeline water extraction timeout failures. An inspection of the filling and discharging system in December revealed that the valve opening of the check valve was significantly reduced compared to normal conditions. The reduced opening decreased the water flow through the check valve per unit time, thereby extending the pipeline water extraction time. This corresponds to the behavior of $\hat{b}$ in the regression line equation [5]. Therefore, the calculated $\hat{b}$ value can be used to determine whether there are issues such as blockages, abnormal valve openings, or motor power anomalies in the filling and discharging system during the pipeline water extraction process.

4.3. Establishment of the Mathematical Model

From the above calculation and analysis, it is concluded that $\hat{b}$ in the regression linear equation has a positive relationship with the pipeline water discharge rate. Therefore, the range of $\hat{b}$ can be used to represent the condition of the filling and discharging system equipment. By integrating the aforementioned formulas (4), (8), (9), (10), (11), and (12), the formula for determining the trend of the pipeline water discharge equipment status is derived as shown in formula (13) below.

$$\left(\hat{b} \pm t_{\varepsilon/2}(n-2) \times \frac{\sqrt{(S_{yy}-\frac{S_{xy}^2}{S_{xx}})/(n-2)}}{\sqrt{S_{xx}}}\right) \tag{13}$$

By integrating formulas (1), (2), and (3), the mathematical model is derived as:

$$\begin{cases} S_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}x_i^2-\frac{1}{n}(\sum_{i=1}^{n}x_i)^2 \\ S_{yy}=\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}y_i^2-\frac{1}{n}(\sum_{i=1}^{n}y_i)^2 \\ S_{xy}=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) \\ =\sum_{i=1}^{n}x_iy_i-\frac{1}{n}(\sum_{i=1}^{n}x_i)(\sum_{i=1}^{n}y_i) \\ \hat{b}=\frac{S_{xy}}{S_{xx}} \\ \hat{b}\pm t_{\varepsilon/2}(n-2)\times\frac{\sqrt{(S_{yy}-\frac{S_{xy}^2}{S_{xx}})\big/(n-2)}}{\sqrt{S_{xx}}} \end{cases} \tag{14}$$

In the practical application process, the $\hat{b}$ values calculated from continuously generated new data sets can be monitored. When the $\hat{b}$ values gradually approach the boundary of the confidence interval, it can be determined that the state of the pipeline water extraction equipment is about to become abnormal, thereby alerting operation and maintenance personnel to perform relevant maintenance in advance [6].

5. Conclusion

In this paper, the linear regression analysis method is used to reverse verify the data changes under the normal operating conditions and abnormal operating conditions of the pumping pipeline water, and through calculation, the value of $\hat{b}$ under the normal operating condition is higher than the value of $\hat{b}$ under the abnormal operating condition, that is, the slope of the corresponding linear regression equation is larger, and the large slope indicates that the pumping rate of the pumping pipeline water is faster, and it also shows that the pumping pipeline water equipment such as the pumping pump equipment is operating normally, and the opening of the valve is normal, which confirms the feasibility of the linear regression analysis method in the operation state analysis of the pumping pipeline water equipment from the side. This paper provides an idea for analyzing and judging the state change of the pumping pipe and water pouring equipment of a Certain ship lift.

Through the mathematical model established in this paper, the online monitoring and alarm of the water pumping equipment of a Certain Ship Lift can be realized, which provides a basis for the operators to grasp the status of the equipment at all times.

References

1. Y. Xu, Y. Sun, J. Wan, X. Liu, and Z. Song, "Industrial Big Data for Fault Diagnosis: Taxonomy, Review, and Applications," *IEEE Access*, vol. 5, pp. 17368–17380, 2017, doi: 10.1109/ACCESS.2017.2731945.
2. S. Mujeeb, T. A. Alghamdi, S. Ullah, A. Fatima, N. Javaid, and T. Saba, "Exploiting Deep Learning for Wind Power Forecasting Based on Big Data Analytics," *Appl. Sci.*, vol. 9, no. 20, p. 4417, 2019, doi: 10.3390/app9204417.
3. J. Helsen, G. De Sitter, and P. J. Jordaens, "Long-Term Monitoring of Wind Farms Using Big Data Approach," in *Proc. 2016 IEEE 2nd Int. Conf. Big Data Comput. Serv. Appl. (BigDataService)*, Oxford, UK, 2016, pp. 265–268, doi: 10.1109/BigDataService.2016.49.
4. W. Yang, R. Court, and J. Jiang, "Wind Turbine Condition Monitoring by the Approach of SCADA Data Analysis," *Renew. Energy*, vol. 53, pp. 365–376, 2013, doi: 10.1016/j.renene.2012.11.030.

5. A. Sasinthiran, S. Gnanasekaran, and R. Ragala, "A Review of Artificial Intelligence Applications in Wind Turbine Health Monitoring," *Int. J. Sustain. Energy*, vol. 43, no. 1, 2024, doi: 10.1080/14786451.2024.2326296.
6. X. Zhang, Y. Meng, T. Yan, et al., "A New Method for Total Organic Carbon Prediction of Marine-Continental Transitional Shale Based on Multivariate Nonlinear Regression," *Front. Earth Sci.*, 2025, doi: 10.1007/s11707-025-1149-y.