Steam Trap Maintenance-Prioritizing Model Based on Big Data
Jiwon Roh, Subean Jang, Suyeun Kim, Hyungtae Cho,* and Junghwan Kim*

ABSTRACT: Steam traps in large facilities need continuous maintenance to prevent corrosion and other damage that could pose a considerable threat to a facility and its workers. However, a significant amount of human resources is required for the maintenance of steam traps. An automatic method to inform stakeholders regarding maintenance cycles will be beneficial for the maintenance process. Therefore, an optimal maintenance priority decision model is developed in this study to establish an efficient steam trap management system. First, the frequency of failures, installation locations, and specifications of steam traps were determined as parameters causing a failure. A relative score and conversion score are calculated for each parameter. The final conversion score is the sum of the conversion score multiplied by the corresponding steam trap data weight factor. Steam traps within the range requiring inspection are classified as high priority cases. Experimental results confirmed that the failure accuracy rate is approximately 95%, and the average failure error rate is within 3%. Additionally, the number of steam traps to be checked was reduced by 3616. The proposed model significantly reduces maintenance in commercial industries.

1. INTRODUCTION
A steam trap is a valve designed to discharge a condensate using the latent heat from the steam equipment and piping. Unlike regular valves, it is an automatic valve that opens and closes in response to changes in the flow rate of the condensate. It prevents steam leakage by selectively discharging the condensate and saving the live steam during the process, thus saving energy and adding stability to the steam facility.1

Typically, steam traps are divided into mechanical, temperature-controlled, and thermodynamic types, depending on the operation method. Mechanical traps use the difference in density between the condensate and steam. Because condensate is denser than steam, a floating barrel or sphere is used to open and close the trap. Temperature-controlled traps, also known as thermostatic traps, exploit the difference in the coefficients of thermal expansion of two metals. The mechanical device that opens and closes the valve is, essentially, a bimetallic element driven by the temperature difference between the steam and condensate. Thermodynamic traps operate on the Bernoulli’s principle and comprise a single moving part, a disk. At the beginning of the operation, the condensate is freely discharged because the valve is open. As steam enters the system, the flow velocity increases, thus, reducing the pressure and closing the valve. When the re-evaporated steam condenses, the disk opens the valve.2

When the condensate fails to discharge, the temperature of the trap decreases owing to ambient air cooling. When the temperature of the trap is out of the normal range, it is said that the trap is in the COLD state. Additionally, water hammering, a sudden change in fluid flow due to the remaining condensed water, causes a pressure surge and the propagation of a pressure wave. In the case of a LEAK failure, which indicates a steam leak along with condensate discharge due to the failure of the steam trap valve, the latent heat of the steam is not consumed efficiently, resulting in a loss of energy.3

If these problems are not resolved, corrosion or similar damage can occur; consequently, shortening the life of the facility and increasing the risk of injury for workers. Therefore,
steam traps require regular maintenance. For example, in the process of diagnosing several steam traps scattered throughout a factory during steam trap maintenance, there is a substantial consumption of human resources and time; therefore, a model that could predict maintenance cycles and automatically inform the stakeholders would be instrumental in improving the maintenance management process.

When the remaining useful life (RUL) of installed equipment is known, it is meaningful to use mathematical optimization models to determine the optimal production and minimal maintenance costs; such models have been widely used in the past. Having understood the commonly identified constraints, such as the required production volume and required maintenance cycle, we can determine the variables that optimize the operating costs and employ an objective function for rational decision-making. However, such models still require an accurate mathematical formulation for the constraints. In the case of steam traps, the constraints related to the predicted lifetime of the equipment are unclear; consequently, establishing a mathematical model is a challenging task.

Methods for predicting the RUL of a device include models based on experience and data and models based on physical and mathematical calculations. In the case of steam traps, because no physical and mathematical models can predict RUL, they were excluded from consideration. Methods for automating decision-making and evaluating equipment status using data-driven algorithms include neural networks, support vector machines, and fault tree analysis. The failure pattern and predicted equipment lifetime can be found by setting a functional significance parameter (FSP) and by monitoring and analyzing the data of the relevant FSPs. Gibbs suggested attaching a wireless sensor to the steam trap to collect and manage the steam traps’ operational data effectively in real time, but attaching and managing sensors on more than 30,000 steam traps installed at a refinery is not economically feasible. In a typical and realistic configuration of a steam trap without a sensor, it is difficult to provide real-time data; hence, conventional equipment maintenance using FSP is challenging.

To create a maintenance model based on the big data of the maintenance history of the steam trap, the failure mode and effects analysis (FMEA) method has been used to prioritize and determine the maintenance. The FMEA method is a structured method that analyzes the influence of the failure modes and occurrences that affect the life of a system to ensure the safety and reliability of the system. FMEA prioritizes risk as the product of the risk factors of the system, the occurrence and severity of the failure, and the ability to detect a pending failure. Because it can be challenging to create an accurate mathematical model for maintenance in various fields, a study was conducted to calculate the priority by setting the criteria based on experience. Additionally, a method of assigning different weights to each risk factor to avoid errors in the standard FMEA that treats all risk factors equally, called failure mode effects and criticality analysis (FMECA), was proposed. Carpitella et al. used both FMECA and multi-criteria decision methods to optimize the maintenance of complex systems. All possible failure modes were identified through FMECA, and the failure modes were ranked through fuzzy techniques for order preference by similarity to ideal solution. The analytic hierarchy process method was applied to give weights to use qualitative judgments made by an expert group and to determine the evaluation criteria. Yang et al. improved the productivity of the process and performed dynamic maintenance scheduling by presenting a quantitative method to determine the priority of maintenance work orders. To overcome the shortcomings of the existing methods of determining the risk priority number, Zammori and Gabbirelli proposed prioritizing the risk using a new multi-criteria decision-making approach in which decisions can be made considering qualitative and political aspects, even when attributes and alternatives are mutually dependent; the proposed method is classified as a ranking model.

We now discuss studies using the prioritizing decision maintenance model used in this study. Trojan and Morais have developed a prioritizing decision maintenance model to reduce unnecessary water use and wastage in the water supply sector. Priority was calculated by determining the ranking of the final score based on Copeland’s method. Taghipour et al. developed a prioritizing maintenance management program by creating multi-criteria and applying a score to mitigate functional failures of medical equipment. The criteria for a critical failure occurring in the device were determined using an analytical hierarchy process. The model calculated total risk to estimate the different failure modes and for assessing frequencies realistically. Sharma et al. used an analytical hierarchy process to prioritize decisions. The study was conducted in two phases, i.e., qualitative followed by quantitative.

In this study, to develop a model that predicts and informs maintenance cycles based on qualitative and quantitative evaluations, a method of calculating the conversion score and ranking the priority was used. The big data we used in calculating the steam trap failure rate and for deriving parameters were the maintenance history data from 2010 to 2019 from Yajung Ltd., a steam trap maintenance company.

2. FAILURE PROBABILITY OF THE STEAM TRAP

2.1. Main Parameters of Failure. It is challenging to acquire accurate quantitative operation and maintenance data because in most cases, there is no sensor attached to a steam trap. For this reason, it is necessary to identify the qualitative tendencies indicated by parameters that are likely to cause failure. This study identified qualitative trends using the maintenance history data of 34,226 steam traps. Herein, we examine an instance of a failure and whether the measured values of size, pressure, type, and location show a causal relationship with each failure. The trap specifications consisted of three categories: size, pressure, and type, classified into six sizes, five types, and three pressure ranges. As shown in Table 1, the regions are divided into 104 locations according to the type of process used to install the steam trap.

Figure 1 shows the frequency and probability of occurrence of the COLD and LEAK failure types over 5 years. Table 1 shows the number of steam traps per location. When comparing Figure 1 and Table 1, it can be seen that the failure tendency is slightly different for each corresponding location regardless of the total number of steam traps. The installed location was also selected as a parameter to be considered by confirming that the number of traps installed in a specific location and the frequency of failure is not proportional.

The size refers to the diameter of the steam trap outlet pipe. The reason that the size can be a factor of failure is because many refineries often re-select a size that has been used in the
past; however, the size of the trap outlet side should be set based on the condensate flow rate and pressure loss of the steam installation. Depending on the outlet size, the flow rate of the condensate that can be processed can vary by more than 4800 kg/h; therefore, considerable failure can occur if a proper size is not selected.

Each type of trap has its own advantages and disadvantages. Their methods of functioning are different; therefore, a trap suitable for each specific application should be selected. In contrast, if the type of the trap installed in the steam main is unsuitable, condensate stagnation can occur, which can increase the probability of a water hammering event. Pressure has not been directly investigated in terms of its qualitative features or quantitative values that affect the failure. However, when calculating the probability of failure for each level of pressure, the probability of failure increases rapidly as the pressure increases. It was hence necessary to select the pressure as a probable cause of failure.

Except for device removal or movement, the specification and type of location for an installed steam trap are fixed. In contrast, the frequency of a steam trap failure exhibits apparent time-variational behavior as visible in the observed period. Owing to different time characteristics, the degrees to which the specifications, locations, and observed frequency of failures contribute to future failures were calculated differently. The specifications and locations were quantified and standardized by calculating the probability of failure. Furthermore, the frequency of failure was analyzed by comparing the correlation between failure history and the current state.

2.2. Failure Probability by Main Parameters. The probability of each failure was calculated from the frequencies of COLD and LEAK occurrences, according to the trap specifications and locations. $P_{\text{Cold}}$ and $P_{\text{Leak}}$ which indicate the probability of COLD and LEAK failures, respectively, were calculated by dividing the number of COLD and LEAK failures by the number of steam traps corresponding to each parameter. Table 2 shows the failure probability calculation process. As shown column (1) in Table 2, the number of steam traps for each level of pressure is known. Column (2) displays the number of COLD failures that occurred during the last five years.

| location | number of steam traps | number of steam traps |
|----------|-----------------------|-----------------------|
| A        | 125                   | AF                    |
| B        | 242                   | AG                    |
| C        | 307                   | AH                    |
| D        | 279                   | AI                    |
| E        | 50                    | AJ                    |
| F        | 187                   | AK                    |
| G        | 411                   | AL                    |
| H        | 215                   | AM                    |
| I        | 54                    | AN                    |
| J        | 433                   | AO                    |
| K        | 123                   | AP                    |
| L        | 376                   | AQ                    |
| M        | 82                    | AR                    |
| N        | 338                   | AS                    |
| O        | 162                   | AT                    |
| P        | 1011                  | AU                    |
| Q        | 491                   | AV                    |
| R        | 670                   | AW                    |
| S        | 165                   | AX                    |
| T        | 192                   | AY                    |
| U        | 171                   | AZ                    |
| V        | 154                   | BA                    |
| W        | 77                    | BB                    |
| X        | 68                    | BC                    |
| Y        | 497                   | BD                    |
| Z        | 99                    | BE                    |
| AA       | 101                   | BF                    |
| AB       | 300                   | BG                    |
| AC       | 344                   | BH                    |
| AD       | 239                   | BI                    |
| AE       | 931                   | BJ                    |

Figure 1. Number of failures and failure rate at each location for 5 years.
halves; thus, the count is divided by 5. In column (3), the average number of COLD failures per semi-annual period is displayed. At this time, (3) is divided by (1) to obtain the probability of failure (4). Hence,

\[ P_{\text{Cold}} \% = \frac{(2) \text{ in Table 3}}{5 \times (1) \text{ in Table 3}} \times 100 \]

\[ = \frac{(3) \text{ in Table 3}}{(1) \text{ in Table 3}} \times 100 \]

2.3. Correlation Analysis Method. It is assumed that the most significant influence on the current state of a steam trap in the current quarter is the state of the steam trap in the last three quarters. Correlation analysis among trap phases per quarter was conducted to check this assumption. Table 3 shows the quarterly correlations from 2010 to 2015; a correlation coefficient indicating the degree of impact of the previous three quarters on the current state of a steam trap is marked in blue for each quarter. For example, the impacts of 2010.2, 2010.3, and 2010.4, in the first quarter of 2011, are represented by 0.4565, 0.5223, and 0.6206, respectively. Table 4 shows the correlation coefficients among the three previous consecutive quarters and the current quarter. Finally, the correlation coefficients were obtained by averaging the correlation coefficients for the third, second, and first quarters. As a result of listing the degrees of influence in chronological order, the average correlation coefficients were found to be 0.5652, 0.6086, and 0.7375 for the third, second, and first quarters, respectively, ranging from the oldest to the newest. Therefore, it was confirmed that the quarter that immediately preceded each quarter shows the highest correlation coefficient.

3. DEVELOPMENT OF A MODEL FOR OPTIMAL MAINTENANCE DECISIONS

3.1. Optimal Decision Flow. A flow chart for selecting steam trap maintenance priorities is shown in Figure 2. The final conversion score \( (C_{\text{Final}}) \) for selecting the steam trap maintenance priority is calculated using eq 2.

\[ C_{\text{Final}} = \alpha C_F + \beta C_L + \gamma C_S \]

\( C_{\text{Final}} \) is based on a 100 point scale, and if \( C_{\text{Final}} \) fits within the diagnostic range, it is recommended as a maintenance priority and calculation is terminated. Parameters \( \alpha \), \( \beta \), and \( \gamma \) in

| Table 2. Calculating \( P_{\text{Cold}} \) |
|------------------------------------------|
| pressure | number of steam traps by type | cold for 5 semi-annuals | cold for 1 semi-annual | \( P_{\text{Cold}} \) |
|----------|-------------------------------|--------------------------|------------------------|----------------|
| 150      | 28,763                        | 23,249                   | 4649                   | 16%  |
| 300      | 4986                          | 2680                     | 536                    | 11%  |
| 600      | 477                           | 245                      | 49                     | 10%  |

| Table 3. Quarterly Correlation Coefficient |
|--------------------------------------------|
| Quarter | 10.2 | 10.3 | 10.4 | 11.1 | 11.2 | 11.3 | 11.4 | 11.5 | 12.1 | 12.2 | 12.3 | 12.4 | 12.5 | 12.6 | 12.7 | 12.8 | 12.9 | 13.1 | 13.2 | 13.3 | 13.4 | 14.1 | 14.2 | 14.3 | 15.1 | 15.2 | 15.3 |
|        | 0.39 | 0.46 | 0.62 | 0.63 | 0.57 | 0.55 | 0.57 | 0.56 | 0.49 | 0.47 | 0.49 | 0.47 | 0.56 | 0.56 | 0.64 | 0.62 | 0.77 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76 |
each term are weight factors multiplied by the conversion scores \( C_C, C_L, \) and \( C_S \) for each parameter, respectively, and thus become the weight factors for correcting the degree of influence of the three parameters on failure. The weighting process was perfected through a case study, which will be presented in Section 4.1. Parameters \( C_C, C_L, \) and \( C_S \) refer to the frequency, location, and specification conversion scores, respectively, which are determined based on the scores per each frequency \( R_F \), location \( R_L \), and relative specification scores \( R_S \). Scores \( R_F, R_L, \) and \( R_S \) are calculated based on the calculated failure probability, correlation coefficient, and failure history, obtained as described in Sections 2.2 and 2.3. The method of obtaining the converted and relative scores will be further discussed in detail.

3.2. Optimal Decision Conversion Score. To calculate \( C_C, C_L, \) and \( C_S \), the rating criteria for each parameter were determined from the distribution of \( R_F, R_L, \) and \( R_S \), respectively. As shown in Figure 3, when the relative scores were arranged in descending order from the highest point, the first grade was set to the score range corresponding to 10% of the total number of traps. The second grade was set to the score range corresponding to the cumulative 30% of the total number of traps starting from the lowest relative first-grade score. The third grade was set to the score range of the remaining traps, corresponding to 70% of the total starting from the lowest relative second-grade score. It is classified according to the grade standard, and a converted score is assigned to each grade, where the converted score is a number between 0 and 1. The distribution of relative scores that spreads over a wide range for each parameter is inferred as three-scaled distributions.

Table 4. Correlation Coefficient of the Previous Three Quarters

| quarter | three quarters ago | two quarters ago | one quarter ago |
|---------|-------------------|-----------------|----------------|
| 11.1    | 0.4565            | 0.5223          | 0.6206         |
| 11.2    | 0.5222            | 0.5333          | 0.6712         |
| 11.3    | 0.4750            | 0.5932          | 0.7533         |
| 11.4    | 0.6411            | 0.7198          | 0.7399         |
| 12.1    | 0.5889            | 0.5617          | 0.6395         |
| 12.2    | 0.5147            | 0.5283          | 0.7269         |
| 12.3    | 0.4928            | 0.6518          | 0.8243         |
| 12.4    | 0.6537            | 0.6062          | 0.7458         |
| 13.1    | 0.5276            | 0.5972          | 0.7996         |
| 13.2    | 0.5853            | 0.6453          | 0.7549         |
| 13.3    | 0.5799            | 0.6226          | 0.7983         |
| 13.4    | 0.6162            | 0.6280          | 0.7165         |
| 14.1    | 0.4895            | 0.5060          | 0.6912         |
| 14.2    | 0.3492            | 0.5707          | 0.7728         |
| 14.3    | 0.3027            | 0.6435          | 0.7810         |
| 15.1    | 0.6745            | 0.6483          | 0.6982         |
| 15.2    | 0.6646            | 0.7030          | 0.7739         |
| 15.3    | 0.6392            | 0.6774          | 0.7688         |
| average | 0.5652            | 0.6086          | 0.7375         |

Figure 2. Maintenance prioritizing model.

Figure 3. Relative score criteria for grading.

Table 5 shows the matching of parameter grades to the appropriate ranges in the manner described above. For each parameter, the range of relative scores corresponding to the three grades is given, and the converted score is calculated as the grade of the range that includes the calculated relative scores.

3.3. Relative Score Calculation Method for Optimal Decisions. In this section, the relative scores for parameters are calculated to define the criteria for separating the three grades, as shown in Table 5. Relative scores were calculated using the failure probability, as obtained in Section 2.2, and the average correlation coefficient, as obtained in Section 2.3.

3.3.1. Calculation of \( R_S \) and \( R_L \). Parameters \( R_S \) and \( R_L \) were calculated by the method shown in Table 6 using \( P_{\text{Leak}} \) and \( P_{\text{Cold}} \) which correspond to the specification and location, respectively. Each parameter is calculated by multiplying \( P_{\text{Leak}} \) and \( P_{\text{Cold}} \) as shown in Section 2.2, with their appropriate weight factors. Because the specification consists of size, type, and pressure, \( R_S \) is calculated as the sum of the relative scores of size, type, and pressure (\( R_{\text{Size}}, R_{\text{Type}}, \) and \( R_{\text{Pressure}} \)).
Because LEAK and COLD present different types of accident damage, weight factors are multiplied in proportion to the severity of the damage. The numbers 2 and 8 are based on the big data and experience of Yajung Co., Ltd., a steam trap maintenance company. When cold and leaks occur, the amount of damage and financial losses that occur to the process are estimated. The value was set assuming that it incurs a loss of about four times.

3.3.2. \( R_f \) Calculation Method. Before obtaining the relative score, \( R_p \), corresponding to the failure rate, the steam trap state was converted into a number corresponding to each state. The steam trap state can be OK (normal), COLD (clogged), or LEAK (leak). The steam trap state was quantified as 0 for OK, 1 for COLD, and 2 for LEAK during the last three quarters based on the time when maintenance and inspection were required. \( R_p \) is calculated using the conversion coefficient values obtained in Section 2.3 and the converted state value according to the corresponding state.

\[
R_p = 0.5652 \times X_{-3} + 0.6086 \times X_{-2} + 0.7375 \times X_{-1}
\]

where \( X_{-n} \) is the quantified state value, such as 0, 1, and 2, and \( n \) indicates the number of previous quarters.

4. RESULTS AND DISCUSSION

4.1. Case Study of Parameter Weights. Because each parameter has a different effect on failure, the accuracy of the failure prediction model varies according to the weight factor of each parameter. To develop a high-accuracy failure prediction model, weighting factors (\( \alpha \), \( \beta \), and \( \gamma \)) were applied to the conversion score for each parameter when calculating the final conversion scores for a maintenance priority diagnosis, as shown in eq 2. A case study was conducted to determine the weight factors and importance of all three parameters. Table 7 shows the parameter weight factors selected for each case. For a clearer understanding of Table 7, case A is used as an example. For case A, the weight factor of frequency is 60, location is 10, and specification is 30. Case A deals with how the model will prioritize maintenance when frequency has the most significant effect on failure, and location has a minor effect on failure. Table 8 shows the scores calculated by multiplying the weight factors selected for each case by the corresponding converted scores for frequency, location, and specification. The failure prediction rate and the failure occurrence rate were determined as the predicted performance indicators for the case selection criteria. The failure prediction rate (eq 4) is the ratio of the number of steam traps predicted to fail (those to be checked, the number of risk states, \( N_R \)) to the number of actual steam trap failures.

\[
\text{failure prediction rate (\%)} = \frac{\text{number of actual failed steam traps in } N_R}{\text{number of actual failed steam traps}} \times 100
\]

The prediction target was the steam trap data from the first half of 2016, with 34,226 steam trap data points. The actual number of faulty traps in the semi-annual period was 4072. Table 9 shows the failure prediction rate and prediction error rate for each case according to changes in weight factors as a result of the case studies. Case B shows the highest failure rate and the lowest prediction error rate, which can be interpreted as the case with the best performance. In Figure 4, it is possible to confirm the changing trend of the \( N_R \) and failure prediction rate for each case. As shown in Figure 4, when the results of case A and case B are compared, the failure prediction rate in case B was found to be improved.

When comparing the trend of the parameters and weight factors for each case shown in Table 8 and the prediction accuracy for each case shown in Table 9, \( \beta \) increased and \( \gamma \) decreased in order from case A to case B. As a result, the magnitude relation in case A, where \( \beta < \gamma \), was reversed in B to \( \beta > \gamma \) therefore, the failure rate increased due to this change. Based on this finding, while actual operating conditions, such

### Table 5. Relative Score Standard for the Conversion Score by Each Parameter

| grade | range of \( R_f \) | \( C_f \) |
|-------|-------------------|-------|
| 1     | 1.35 < \( R_f \)  | \( C_F, 3 \) |
| 2     | 0.55 < \( R_f < 1.35 \) | \( C_F, 2 \) |
| 3     | \( R_f < 0.55 \)   | \( C_F, 1 \) |

### Table 6. Calculation of \( R_{\text{spec}} \) and \( R_{\text{location}} \)

| \( R_{\text{size}} \) | \( R_{\text{spec}} \) | \( R_{\text{type}} \) | \( R_{\text{location}} \) |
|-----------------|-----------------|-----------------|-----------------|
| \( \times \) \( (P_{\text{cold}}, \text{size}) \) + 8 \( (P_{\text{cold}}, \text{type}) \) | \( \times \) \( (P_{\text{cold}}, \text{pressure}) \) + 8 \( (P_{\text{cold}}, \text{location}) \) | \( \times \) \( (P_{\text{cold}}, \text{type}) \) + 8 \( (P_{\text{cold}}, \text{pressure}) \) | \( \times \) \( (P_{\text{cold}}, \text{location}) \) |

### Table 7. Weight Factor by Case

| case | \( \alpha \) | \( \beta \) | \( \gamma \) |
|------|----------|-------|-------|
| case A | 60       | 10    | 30    |
| case B | 50       | 30    | 20    |
| case C | 50       | 30    | 20    |
| case D | 40       | 35    | 25    |

The prediction target was the steam trap data from the first half of 2016, with 34,226 steam trap data points. The actual number of faulty traps in the semi-annual period was 4072. Table 9 shows the failure prediction rate and prediction error rate for each case according to changes in weight factors as a result of the case studies. Case B shows the highest failure rate and the lowest prediction error rate, which can be interpreted as the case with the best performance. In Figure 4, it is possible to confirm the changing trend of the \( N_R \) and failure prediction rate for each case. As shown in Figure 4, when the results of case A and case B are compared, the failure prediction rate in case B was found to be improved.
as pressure and temperature, vary depending on the process, the installation location has a greater effect on the steam trap failure frequency than the specifications of the device. Cases B and C showed the same failure prediction rate and \( N_R \). The CL was changed to vary the distribution of the final conversion score. As a result of this change, there was a difference in grade 2 of case B and case C, but there was no difference in the final result. When the results of case B and case D are compared, the probability of failure in case B is significantly higher, and the \( N_R \) is approximately 3600. In Table 8, \( a \) decreases, and \( \beta \) and \( \gamma \) increase from case B to case D; therefore, the correlation between the result and the weight factor is substantial in the following order: frequency > location > specification. Based on case B, which shows the highest failure prediction rate in the case study results, the number of steam traps that need to be checked was reduced by 67%, from 34,226 to 11,465, and the failure prediction rate was 90.05%.

4.2. Comparison of Predictive Results. The steam trap data of 2011 Q1 (quarter 1) were predicted by applying the weight factor of each parameter for case B. The total number of steam trap data points for this period is 18,756, and the total number of failures is different for each quarter. Table 10 shows the prediction performance. For all four quarters, there is a high failure rate of approximately 95%, and the \( N_R \) indicates that the number of steam traps to be checked was reduced by

![Figure 4. Comparison of \( N_R \) and prediction accuracy by case.](https://dx.doi.org/10.1021/acsomega.0c05784)
at least 3616. In addition, it was confirmed that the failure occurrence rate is 2.775% on average.

5. CONCLUSIONS

Regular diagnosis of the conditions of steam traps is essential to avoid energy loss and equipment damage. However, a significant amount of resources, both in human effort and time, is required to diagnose the numerous steam traps in an industrial facility. This study developed a maintenance prioritizing model for steam traps, using big data to reduce resource consumption.

Based on big data, parameters that affect the failures were selected. Relative scores were calculated with the failure probability and correlation coefficient for each parameter, and a rating criterion was prepared in consideration of the distribution of the relative scores. Additionally, the conversion score corresponding to each rating was given. Case studies were conducted on the weight factor for each parameter to improve the accuracy of the prediction model. As a result of the case study, it was confirmed that the accuracy of the predictive model in case B, whose weight factors were $\alpha = 50$, $\beta = 30$, and $\gamma = 20$, was the highest. Maintenance priority was determined according to the final conversion score, and the steam traps were classified as those that need current maintenance and those that can be postponed.

In a case study based on steam trap maintenance data from 2018, i.e., case B, the failure prediction rate was 90.05%, the average failure rate was 2.5%, and the number of steam traps that needed to be inspected decreased from 34,226 to 11,465. It showed excellent prediction accuracy. To confirm the prediction accuracy, the steam trap data for four quarters, using the same model as case B, was predicted, and the weight factor confirmed a failure prediction rate of approximately 95% and an average failure rate of within 3%. Additionally, the number of steam traps to be checked was reduced by at least 3616, demonstrating an excellent improvement in efficiency.

The steam trap maintenance prioritizing model can be applied to commercial industries. Because each weight factor has a considerable difference in accuracy, in the future, the weight factors must be fine-tuned and adjusted to improve the failure prediction rate and the failure rate of the maintenance prioritization model.

■ AUTHOR INFORMATION

Corresponding Authors

Hyungtae Cho — Korea Institute of Industrial Technology, Green Materials & Processes R&D Group, Ulsan 44413, Republic of Korea; Email: htcho@kitech.re.kr
Junghwan Kim — Korea Institute of Industrial Technology, Green Materials & Processes R&D Group, Ulsan 44413, Republic of Korea; o orcid.org/0000-0002-2311-4567; Email: kjh31@kitech.re.kr

Authors

Jiwon Roh — Korea Institute of Industrial Technology, Green Materials & Processes R&D Group, Ulsan 44413, Republic of Korea; Department of Chemical and Biomolecular Engineering, Yonsei University, Seoul 03722, Republic of Korea
Subean Jang — Department of Chemical and Biochemical Engineering, Dongguk University, Seoul 04620, Republic of Korea

Suyeun Kim — Yajung Co., Ltd, Ulsan 44988, Republic of Korea

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.0c05784

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study has been conducted with the support of the Korea Institute of Industrial Technology as “Development of Global Optimization System for Energy Process (KITECH EE-20-0019)” and “Development of AI Platform Technology for Smart Chemical Process (KITECH JH-21-0005)”.

■ ABBREVIATIONS

COLD one of the major problems with steam traps is that the temperature of the steam trap falls because of the condensate being blocked
CFinal final conversion score for prioritizing the maintenance of the steam trap
CF conversion score by frequency required for calculating the final conversion score
CL conversion score for each location required to calculate the final conversion score
CS conversion score for each specification required to calculate the final conversion score
FMEA failure mode and effects analysis
FMECA failure mode, effects, and criticality analysis
FSP functional significance parameters
LEAK one of the major problems that occurs in steam traps is that steam leaks when condensate is discharged, causing energy loss
NRisk number of risk-grade steam traps requiring inspection
NRg number of normal-grade steam traps that do not require inspection
PCold probability of a cold failure occurring
PLeak probability of a leak occurring
RF relative score by frequency of steam trap, same as RFrequency
RL relative score for each steam-trap location, same as RLocation
RP pressure relative score by pressure of a steam trap
RS relative score by specification of a steam trap; same asRSpecification
RSsize relative score by size of a steam trap
RTtype relative score by type of a steam trap
RUL remaining useful life
Xn quantified state value, such as 0, 1, and 2, and $n$ indicates the number of previous quarters
$\alpha$ weight factor of the converted score for each frequency used in the final conversion score calculation (%)
$\beta$ weight factor of the converted score for each location used for calculating the final converted score (%)
$\gamma$ weight factor of the conversion score for each specification used in calculating the final conversion score (%)
REFERENCES

(1) Risko, J. R. Understanding Steam Traps. Chem. Eng. Prog. 2011, 107, 21–26.
(2) Risko, J. R. Steam Trap Management: Do Something; Anything. Please! Chem. Eng. Prog. 2017, 113, 64–72.
(3) Sarco, S. The Steam and Condensate Loop: An Engineer’s Best Practice Guide for Saving Energy; Spirax Sarco Limited, 2008.
(4) Kwon, H.; Oh, K. C.; Choi, Y.; Chung, Y. G.; Kim, J. Development and Application of Machine Learning-based Prediction Model for Distillation Column. Int. J. Smart Sens. Intell. Syst. 2021, DOI: 10.1002/ijss.22368.
(5) Kim, J.; Tak, K.; Moon, I. Optimization of Production and Production Planning Model in Refinery Processes Considering Corrosion Effect. Ind. Eng. Chem. Res. 2012, 51, 10191–10200.
(6) Tak, K.; Kim, J. Corrosion Effect on Inspection and Replacement Planning for a Refinery Plant. Comput. Chem. Eng. 2018, 117, 97–104. https://doi.org/10.1016/j.compchemeng.2018.05.027.
(7) Ma, Z.; Ren, Y.; Xiang, X.; Turk, Z. Data-Driven Decision-Making for Equipment Maintenance. Autom. Constr. 2020, 112, 103103.
(8) Si, X.-S.; Wang, W.; Hu, C.-H.; Zhou, D.-H. Remaining Useful Life Estimation — A Review on the Statistical Data Driven Approaches. Eur. J. Oper. Res. 2011, 213, 1–14.
(9) Duer, S.; Zajkowski, K.; Duer, R.; Paš, J. Designing of an Effective Structure of System for the Maintenance of a Technical Object with the Using Information from an Artificial Neural Network. Neural. Comput. Appl. 2013, 23, 913–925.
(10) Zhao, B.; Chen, S.; Wang, Y.; Li, J. Maintenance Decision Methodology of Petrochemical Plant Based on Fuzzy Curvelet Neural Network. Appl. Soft Comput. 2018, 69, 203–212.
(11) Chen, X.; Shen, Z.; He, Z.; Sun, C.; Liu, Z. Remaining Life Prognostics of Rolling Bearing Based on Relative Features and Multivariable Support Vector Machine. Proc. Inst. Mech. Eng., Part C 2013, 227, 2849–2860.
(12) Widodo, A.; Yang, B.-S. Support Vector Machine in Machine Condition Monitoring and Fault Diagnosis. Mech. Syst. Signal Process. 2007, 21, 2560–2574. https://doi.org/10.1016/j.ymssp.2006.12.007.
(13) Lipol, L. S.; Haq, J. Risk Analysis Method: FMEA/FMECA in the Organizations. Int. J. Basic Appl. Sci. 2011, 11, 74–82.
(14) Gibbs, C. Wireless Diagnostics for Steam Traps. Chem. Eng. 2007, 114, 52–54.
(15) Huang, J.; Li, Z.; Liu, H.-C. New Approach for Failure Mode and Effect Analysis Using Linguistic Distribution Assessments and TODIM Method. Reliab. Eng. Syst. Saf. 2017, 167, 302–309.
(16) Committee, I. E. C. 60812 T. Analysis Techniques for System Reliability Procedure for Failure Mode and Effects Analysis (FMEA); Acoustical Society of America: IEC 60812; 2006.
(17) Kim, K. O.; Zuo, M. J. General Model for the Risk Priority Number in Failure Mode and Effects Analysis. Reliab. Eng. Syst. Saf. 2018, 169, 321–329.
(18) Anes, V.; Henriques, E.; Freitas, M.; Reis, L. A New Risk Prioritization Model for Failure Mode and Effects Analysis. Qual. Reliab. Eng. Int. 2018, 34, 516–528.
(19) Chin, K.-S.; Wang, Y.-M.; Poon, G. K. K.; Yang, J.-B. Failure Mode and Effects Analysis by Data Envelopment Analysis. Decis. Support Syst. 2009, 48, 246–256.
(20) Xiao, N.; Huang, H.-Z.; Li, Y.; He, L.; Jin, T. Multiple Failure Modes Analysis and Weighted Risk Priority Number Evaluation in FMEA. Eng. Fail. Anal. 2011, 18, 1162–1170.
(21) Carmignani, G. An Integrated Structural Framework to Cost-Based FMECA: The Priority-Cost FMECA. Reliab. Eng. Syst. Saf. 2009, 94, 861–871.
(22) Carpitella, S.; Certa, A.; Izquierdo, J.; La Fata, C. M. A Combined Multi-Criteria Approach to Support FMECA Analyses: A Real-World Case. Reliab. Eng. Syst. Saf. 2018, 169, 394–402.
(23) Yang, Z.; Chang, Q.; Djurdjanovic, D.; Ni, J.; Lee, J. Maintenance Priority Assignment Utilizing On-Line Production Information. J. Manuf. Sci. Eng. 2006, 129, 435–446.

(24) Zammori, F.; Gabbielli, R. ANP/RPN: A Multi Criteria Evaluation of the Risk Priority Number. Qual. Reliab. Eng. Int. 2012, 28, 85–104.
(25) Trojan, F.; Morais, D. C. Prioritising Alternatives for Maintenance of Water Distribution Networks: A Group Decision Approach. Water Sa 2012, 38, 555–564.
(26) Taghipour, S.; Banjevic, D.; Jardine, A. K. S. Prioritization of Medical Equipment for Maintenance Decisions. J. Oper. Res. Soc. 2011, 62, 1666–1687.
(27) Sharma, M.; Gupta, R.; Acharya, P. Prioritizing the Critical Factors of Cloud Computing Adoption Using Multi-Criteria Decision-Making Techniques. Glob. Bus. Rev. 2020, 21, 142–161.