Assessment of fouling in plate heat exchangers using classification machine learning algorithms




SEYIT AHMET KUZUCANLI
Department of Hydraulic Module R&D, Bosch Thermotechnology, Manisa, Turkey
ahmet.kuzucanli@tr.bosch.com

CEREN VATANSEVER
Department of Hydraulic Module R&D, Bosch Thermotechnology, Manisa, Turkey

ALP EMRE YASAR
Department of Hydraulic Module R&D, Bosch Thermotechnology, Manisa, Turkey

ZIYA HAKTAN KARADENIZ
Department of Energy Systems Engineering, İzmir Institute of Technology, İzmir, Turkey

Keywords: Fouling, plate heat exchanger, ML, classification, Naïve Bayes, k-nearest neighbours, decision tree.

Introduction

Fouling is the continuous deposition and accumulation of undesirable particles on heat transfer surface areas [1]. The primary heat exchanger (HC) provides heat transfer by combusting fuel to central heating (CH) water which is used to transfer heat to the domestic water via a plate heat exchanger (PHE). PHE consists of two channels that guide two immiscible fluids, CH and domestic water. Both sides of the PHE are under the influence of dirt accumulation from the system components. In addition, the domestic water side of the PHE is also affected by calcification due to existing calcium compounds in sanitary water. Crystallization and calcification of calcium compounds are investigated successfully by Lee et al. [2] and Pääkkönen et al. [3,4]

The CH side of the PHE is mostly influenced by the particles coming from the system components such as aluminium-made HC (HC). Due to the corrosion of HC and particles that come from the pipeline, particulate type fouling is encountered on this side of the plates. The main particle that is seen is Al_2O_3. Experimental and numerical investigations are published in the literature about particulate and composite fouling. [5–8]

Fouling modelling and prediction algorithms have mostly been based on statistical methods and ML techniques. Various ML algorithms are studied such as fouling prediction and detection algorithms based on support vector machines (SVM) [9,10], autoassociative kernel regression (AAKR) [11], autoregressive integrated moving average (ARIMA) [12] and artificial neural networks (ANN) [13–17]. Also, model-based fouling prediction research has been investigated by Kalman filter usage [18]. A predictive maintenance approach has been employed to predict fouling behaviour. Predictive maintenance system usually consists of data acquisition, pre-processing, fault diagnostics, and failure prognostics. Fault diagnostics generally focus on statistical approaches that provide classification and clustering. Most failure mechanisms can be associated with degradation processes [19]. The data acquisition process can be maintained by health system monitoring [20]. The ML algorithms used for classification are commonly Naïve Bayes, k-nearest neighbours (kNN), decision trees, and random forests. These algorithms are successfully studied to classify the faults of boilers using simulated data [21]. As a result, the decision tree algorithm gave the best result with an accuracy of 97.8%. As can be seen from references, the ML techniques are commonly used in the HVAC sector, especially in heat exchangers but when the open resources are considered the classification ML algorithms on PHEs have been investigated rarely ever.

In this paper, a multi-classification study to determine fouling levels with an aim of to generate a warning on combi-boiler appliances is carried out for PHEs. The Naïve Bayes, kNN and decision tree ML algorithms are studied with cross-validation methods for experimentally acquired data. A comparison of multi-classification ML algorithms is provided as an introductory study.

Research methods

2.1 Experimental conditions

The algorithm development process starts with data acquisition regarding healthy and faulty conditions. To get a dataset for both conditions, PHEs are tested as stand-alone, i.e., separate from combi-boiler devices. Two PHEs one having 30 and the other 32 plates are considered to examine the clogging status. Each PHE is designed for a specific combi-boiler heating capacity and dynamic system behaviour. Specifications indicate that the given inlet and outlet temperatures and flow rates of the PHE should be in the desired range. Investigated PHEs are already meeting these specifications therefore they are considered as a reference.

The tests are conducted in the flow rates shown in **Table 1**. The domestic water inlet temperatures, i.e., domestic cold water (DCW), are kept constant at 10°C. The CH inlet temperatures are applied in a range of 72 ± 1°C.

The applied conditions of both PHEs represent the technical specifications of 30 and 32 plates (**Table 1**). The technical specifications of 30 and 32 plates PHE are applied to PHEs have different plate numbers, 28, 26, 24, 22, 20, 18 and 16. It is assumed that the effect of fouling on the performance of PHE is the same as the effect that would occur if the PHE with fewer plates was used. 50% clogged PHE as the maximum clogging percentage is considered by using technical specifications of 32 plate PHE that are implied to 16 plate PHE.

2.2 Experimental set-up

The experimental setup contains two lines representing CH and DHW circuits (**Figure 1**). CH line is a closed circuit. The static pressure of the circuit is 2 ± 0.1 bar which is measured inside the tank. A pump provides circulation through the closed circuit. A flow control valve is located on the CH line, together with the pump that can be controlled manually to adjust the required flow rates. On the CH circuit, there is a bypass line that is used for heating the water without preheating the PHE.

Table 1. *Experimental conditions applied to demonstrate the clogging behaviour of the PHEs.*

32 Plates		30 Plates	
CH Flow Rate (l/min)	DHW Flow Rate (l/min)	CH Flow Rate (l/min)	DHW Flow Rate (l/min)
29	18	26	10.3

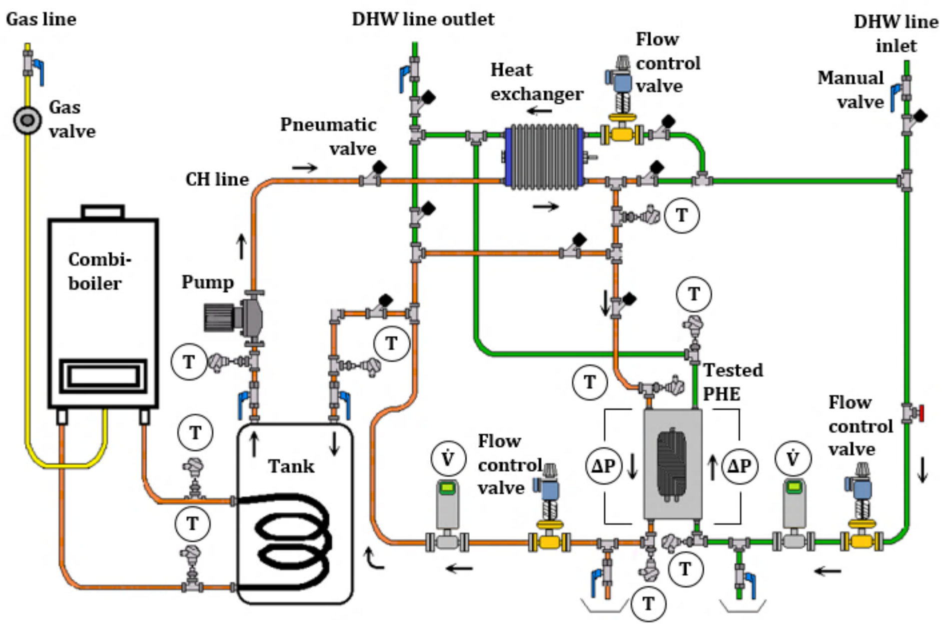

Figure 1. Schematic diagram of experiment set-up and its components.

CH line is interacted with a cold-water line to achieve the ability of necessary cooling with an additional plate heat exchanger. This cooling process control is carried out manually by a flow control valve located on the cold-water line. DHW line is supplied from a main chiller unit. The required flow rates are provided by a flow control valve manually. There are temperature sensors to measure water temperature at the inlets and outlets of the tested PHE besides important points such as the tank inlet and outlet. The sensor locations can also be seen in **Figure 1**. Flow sensors are located on both lines to measure the volume flow rate. Two differential pressure meters are placed to measure the pressure drop over the inlets and outlets of the tested PHE.

2.3 Data processing

The main effect of fouling on the PHEs is functional performance decreasing. The accumulated fouling particles create a film layer on the plates, which pretends like an insulation layer that results in degression of heat transfer. This film layer can be represented as fouling resistance regarding the thermal resistance concept. The logarithmic mean temperature difference (LMTD) method is used to calculate U (Equation. (1)). Heat transfer rates of DHW and CH sides are calculated by Equation (2) and (3). The material properties are taken at the average temperatures of the inlets and outlets for both fluids. The total heat transfer rate is determined by taking the average of the heat transfer rates of the CH and DHW sides. This method is also implied successfully by Zhang et al. [6]

$$\dot{Q}_{total} = U \cdot A \cdot LMTD \quad (1)$$

$$\dot{Q}_i = \dot{m}_i \, c_{p,i} \, \Delta T_i \quad (2)$$

$$\dot{Q}_j = \dot{m}_j \, c_{p,j} \, \Delta T_j \quad (3)$$

$$U = R_i + R_{wall} + R_j + R_f \quad (4)$$

Here, $\dot{Q}$ denotes the heat transfer rate. DHW and CH are indicated as i and j, respectively. Specific heat is indicated as c_p, mass flow rate is $\dot{m}$, temperature difference is ΔT. A denotes the heat transfer area. Fouling resistance (R_f) is obtained from equation (4), where R_i and R_j denote the convective thermal resistances of the DHW and CH sides, respectively. The above results are obtained from the CFD simulations generated by using the test conditions as boundary conditions. R_{wall} denotes conduction heat resistance which is neglected in this study.

2.4 Classification method

The main algorithm development method used is fault diagnosis to obtain the current fouling status of the PHE. Classification implied in this study is a type of supervised ML in which an algorithm learns to classify new data. The training data, from the experimental results, is used in an algorithm to teach the zones to be predicted. The zones that are the fault labels (1 to 8) of the algorithm represent the test conditions. Experimental conditions and measured parameters are predictors, while the zones are categorized responses in the classifier algorithm. While the deviation from zero-hour performance, initial status, is increasing, the deterioration of PHE will be increased also as expected. Zones represent the comfort loss and cost increase levels till the required maintenance time comes and finally when the PHE is needed to be changed. (**Figure 2**). As there are more than one classes to be predicted, multi classification algorithms are used.

Naïve Bayes, kNN (k-nearest neighbours) and decision tree algorithms are chosen due to their applicability to multi classification cases. The algorithms are applied to data by using Classification Learner App in MATLAB program. Before training of the algorithms,

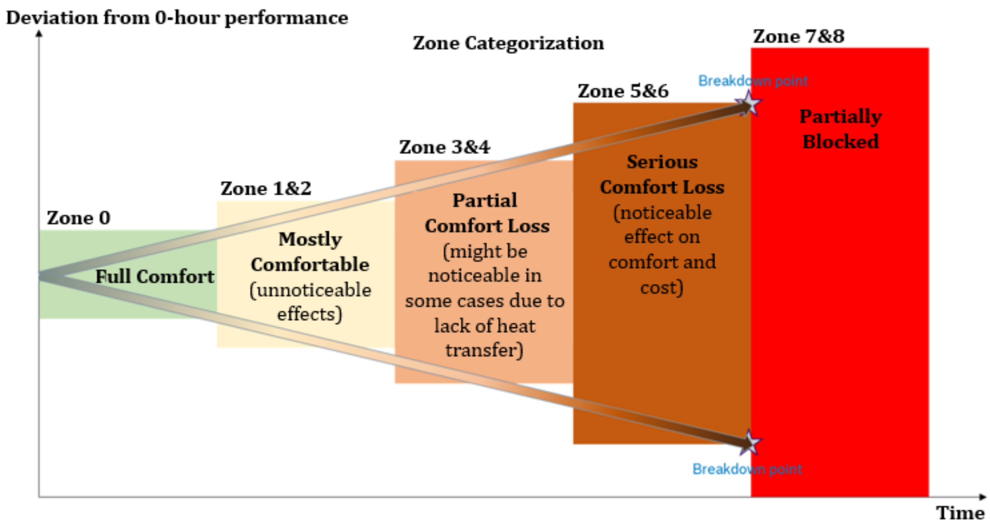

Figure 2. Zone categorization.

cross-validation is used in the process of creating the testing and training data. Cross-validation is a model assessment technique used to evaluate the algorithm's performance. Basically, this offers several techniques that split the data differently to be protected against overfitting. The k-fold cross-validation technique, which is used, partitions data into k randomly chosen subsets (or folds) of roughly equal in size as described in **Figure 3**. One subset is used to validate the trained model using the remaining subsets [22]. The average error across all k partitions is chosen to determine the overall accuracy percentage. The k value is chosen as 5 in this study for all used algorithms.

Naive Bayes is a classification algorithm that applies density estimation to the data. The algorithm uses Bayes theorem and assumes that the predictors are conditionally independent, given the class. Naive Bayes classifiers assign observations to the most probable class [23]. Kernel distribution is used as a numerical predictor where the kernel width is automatically determined using an underlying fitting function via MATLAB [24].

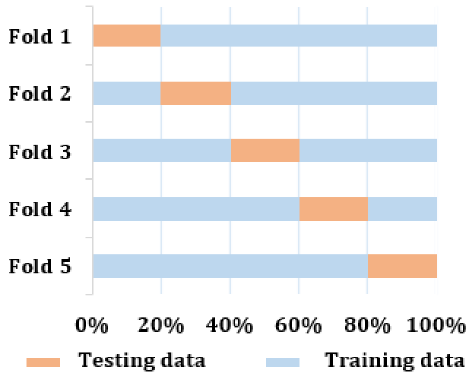

Figure 3. k-fold cross validation designation.

Given a data size of n number points and a distance function, the k-nearest neighbours (kNN) algorithm provides finding the k closest points in data [25]. Number of nearest neighbours (k) to find for classifying each point when predicting, specified as 10 in this study while Euclidean distance is implied as default in MATLAB Classification learner. Decision trees create a hierarchical partitioning of the data, which relates the different partitions at the leaf level to the different classes [26]. The hierarchical partitioning at each level is created using a split criterion. The Gini's diversity index is chosen as a split criterion in this study while the maximum number of splits is implied as 100.

Results and discussion

The experimental results are processed as grouped by classes to be predicted. Outlet temperatures of PHE and differential pressure between inlets and outlets of the PHE of both CH and DHW lines are presented in scatter plots for 30 and 32 plates (**Figure 4**).

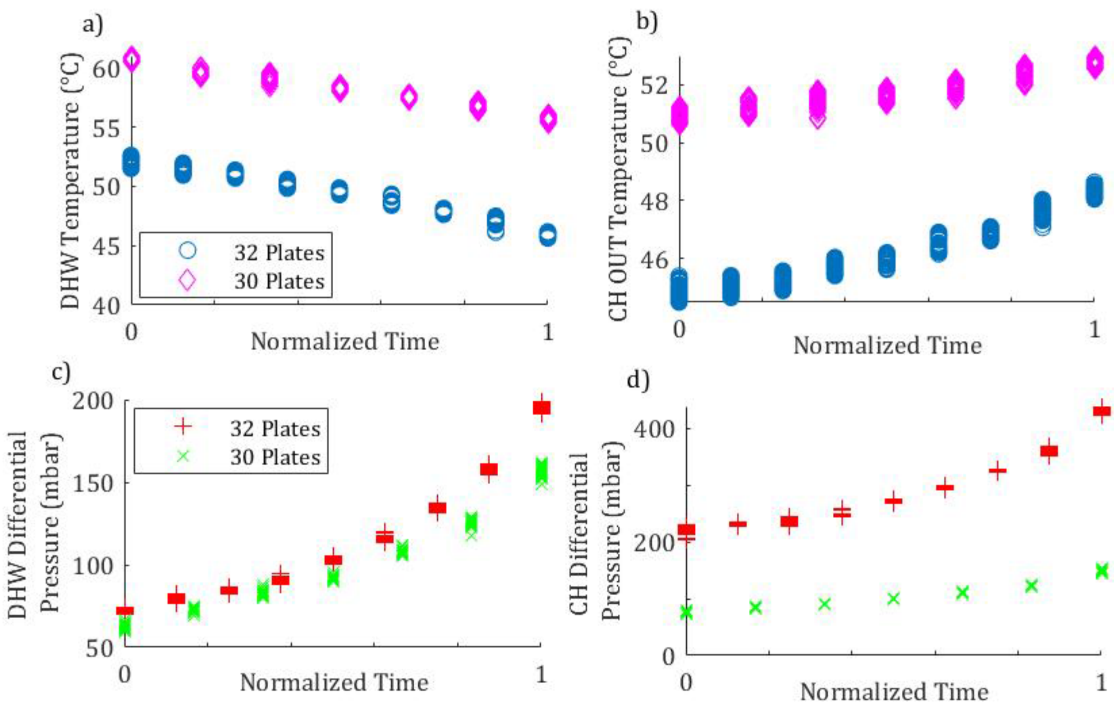

Figure 4. The performance trends for 32 and 30 plates a) DHW temperature behaviour b) CH outlet temperature behaviour c) Pressure drop behaviour of DHW d) Pressure drop behaviour of CH.

The results are presented in normalized time scale, here the normalized time axis represents the zones since they stand for the degradation of PHE from zone 0 to 8, respectively.

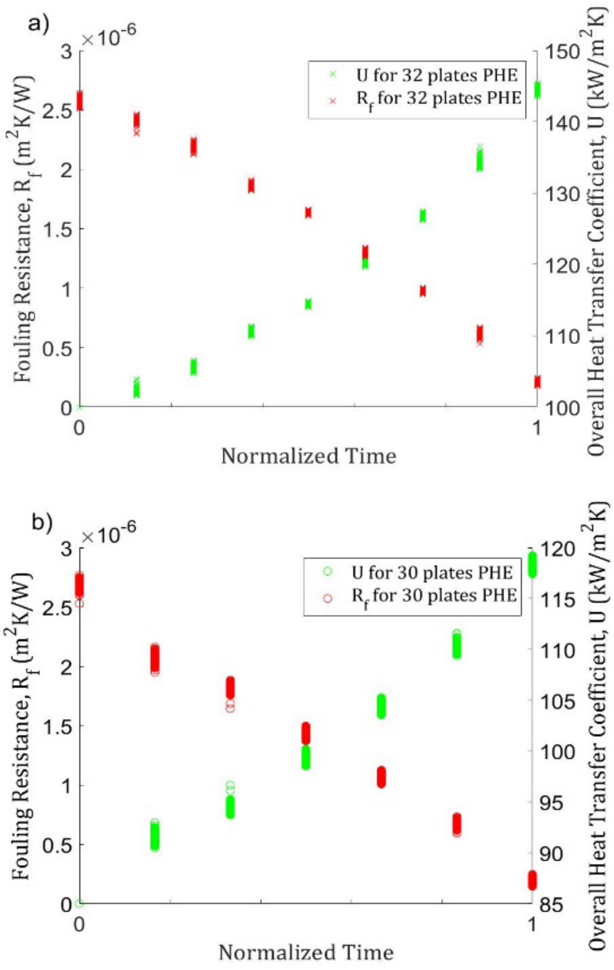

Figure 5. Fouling resistance and *U* behaviours a) for 32 plates, b) for 30 plates.

Since clogging of plates results in a performance decrease in PHEs, trends of DHW outlet temperature for both 30 and 32 plates decrease additionally, trends of CH outlet temperature and pressure drops of both CH and DHW lines for 30 and 32 plates increase as shown in **Figure 4**. The calculated overall heat transfer coefficients for both 30 and 32 plates are shown in **Figure 5**. The trends of U are decreasing as expected.

The fouling resistance values also follow the expected trends in contrast to the *U* as shown in **Figure 5**. In addition, fouling resistance graphs show similar trends to the study by Zhang et al. [6]

Additionally, the Reynolds number range is similar to the range in this study. [6] As a result, trends of the calculated fouling resistance can be considered as a realistic representation of the fouling behaviour with the help of the experimental method implied in this study.

The confusion matrices of the decision tree and kNN algorithms as the results of the predicted performances of the trained models are shown in **Figure 6**. Naïve Bayes predicted the response classes perfectly with 100% accuracy. As the decision tree algorithm can predict fouling with an accuracy of 99.2%, it is followed by the kNN algorithm with 96.7% accuracy. The true positive rates (TPR) and false negative rates (FNR) are designated in confusion matrices with the prediction accuracies portioned by classes. The standard deviation of the imported data can be seen in the parallel coordinates plot for Naïve Bayes trained model in **Figure 7**. In the figure, the classes show more distinguishable

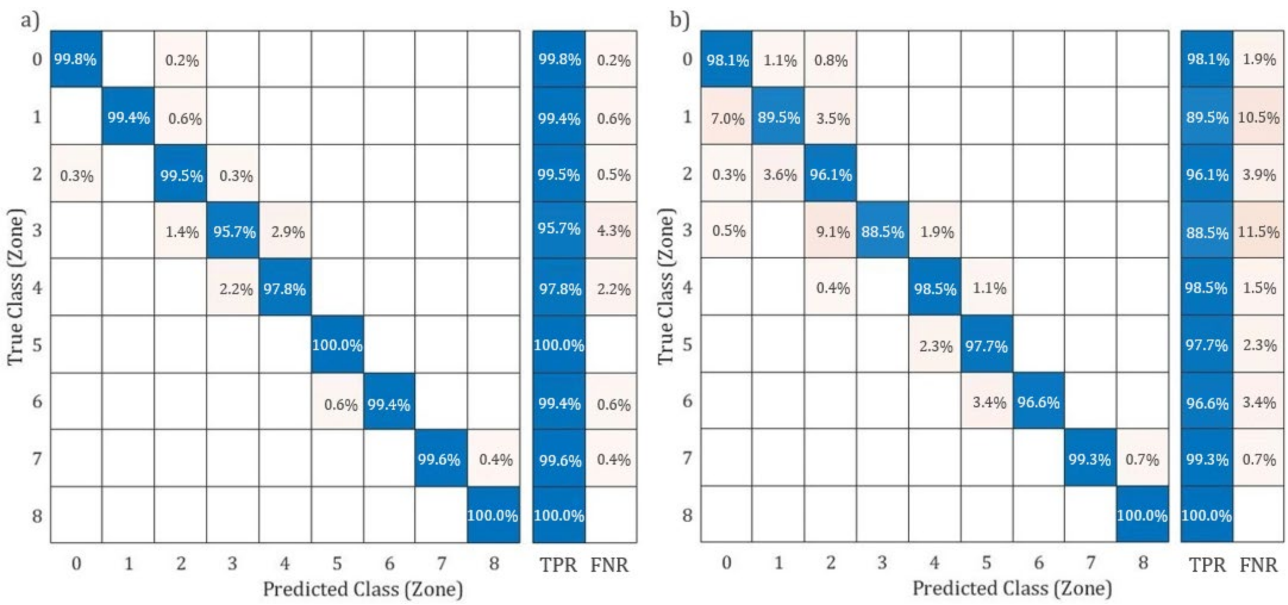

Figure 6. Confusion matrices of a) decision tree model b) kNN model.

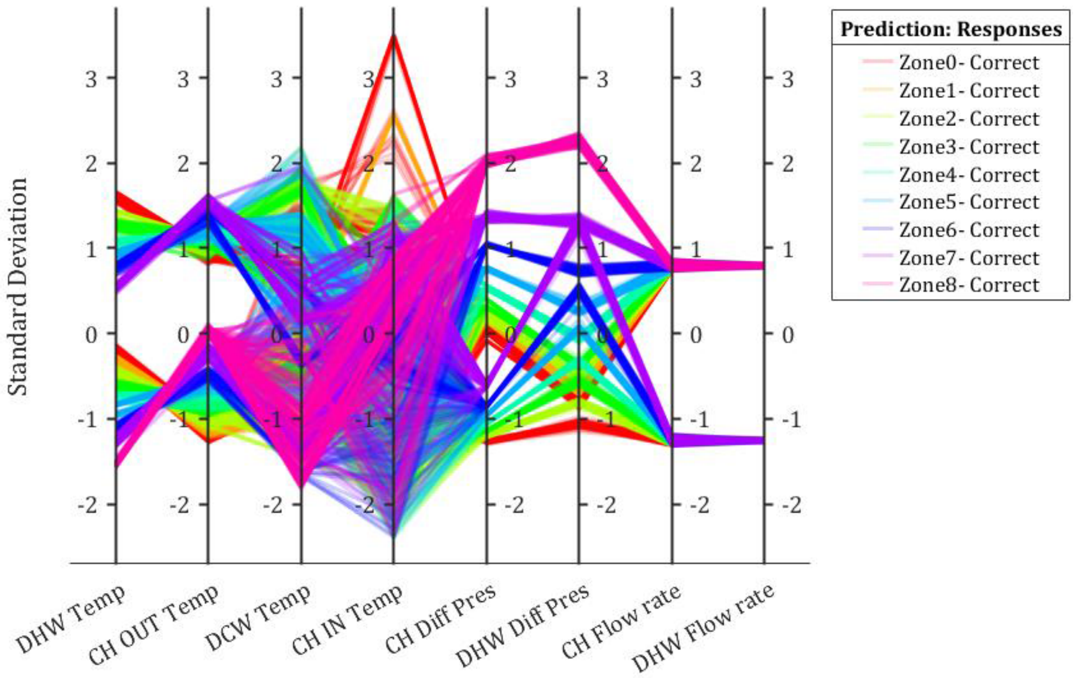

Figure 7. Parallel coordinates plot of experiment data shown in standard deviation scales for the Naïve Bayes model.

distribution on the CH and DHW pressure difference data rather than CH inlet and DCW temperatures. This results in the pressure difference is a more convenient parameter to predict classes correctly rather than other parameters. DHW and CH outlet temperature data are also helpful in distinguishing the classes according to the distribution shown in the parallel coordinate plot (**Figure** 7). With this representation of high dimensional experiment data, the relation of standard deviation between the predictors can be seen.

Conclusion

In this study, an algorithm is developed to imply on combi-boiler appliances with to generate a warning that indicates the fouling level of PHEs. Naïve Bayes, kNN and decision tree are used as the multi-classification ML algorithms.

The data is acquired from an experiment set-up for PHEs having 32 and 30 plates are tested. The experimental conditions are selected as the technical specifications of the PHEs. The experimental data is grouped by zones representing the fouling levels of PHE. The attempted models of ML algorithms result in that Naïve Bayes has better accuracy compared to other models and it is followed by decision tree algorithm with an accuracy of 99.2% and kNN algorithm with 96.7% prediction accuracy. The results of trained models with tested data are shown in confusion matrices. The standard deviation of the data can be represented in parallel coordinates plot which results in the pressure drop values being seen to have the best distinguishing feature among the predictors.

Overall, this study demonstrates the possibility of generating a warning for the current fouling level classification of PHEs in combi-boiler appliances by implying ML algorithms with high accuracy. Generation of fouling level warnings results in the possibility of releasing a feature that can be the major effect of cost saving by retrenching on maintenance.

The framework of this study can be refined by taking a time-dependent dataset into account to assess optimum time schedule of maintenance in future studies. ■

Acknowledgement

Sincere gratitude is expressed deeply to acknowledge the help provided by the technical and support staff in Bosch Thermotechnology.

References

Please find the full list of references in the original article at: https://proceedings.open.tudelft.nl/clima2022/article/view/127