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

Seyit Ahmet Kuzucanlı [a], *Ceren* Vatansever [b], *Alp Emre* Yaşar [c], *Ziya Haktan* Karadeniz [d]

[a] Department of Hydraulic Module R&D, Bosch Thermotechnology, 45030, Manisa, Turkey, ahmet.kuzucanli@tr.bosch.com

[b] Department of Hydraulic Module R&D, Bosch Thermotechnology, 45030, Manisa, Turkey, ceren.vatansever@tr.bosch.com

[c] Department of Hydraulic Module R&D, Bosch Thermotechnology, 45030, Manisa, Turkey, alpemre.yasar@tr.bosch.com

[d] Department of Energy Systems Engineering, İzmir Institute of Technology, İzmir, Turkey, haktankaradeniz@iyte.edu.tr

Abstract. Plate heat exchangers (PHE) used in combi-boilers are continuously affected by small particles, both from the heating circuit and other components of the heating system. These particles can accumulate in the heat exchanger and create clogging that affect the performance of the heat exchanger over time by generating a insulating layer. To avoid unexpected blockage and other kinds of mechanical failure caused by unintended particles that originate from the pipeline and other components, it is crucial to design an effective predictive maintenance system for PHE used in the combi-boiler. In this study, the early stage of blockage in a PHE is investigated experimentally to minimize the field failure rate. The data is acquired from an experimental set-up in which just the PHEs are tested. The PHEs with the same plate pattern and different plate numbers are tested using varied flow rate and inlet temperatures as parameters. The overall heat transfer coefficient and fouling resistance are calculated to associate with the functionality of PHE. A comparison study of multi-classification algorithms has been investigated to present an algorithm which gives the most accurate model trained by experimental data. K-folds cross validation are studied using Naïve Bayes, k-nearest neighbours (kNN) and decision tree machine learning algorithms. As a result, the behaviour of overall heat transfer coefficient and fouling resistance in normalized time scale show the expected trends. The attempted models of machine learning algorithms result in Naïve Bayes predicting the classes of test data perfectly and it is followed by decision tree algorithm with an accuracy of 99.3% and kNN algorithm with 96.3%.

Keywords. Fouling, plate heat exchanger, machine learning, classification, Naïve Bayes, k-nearest neighbours, decision tree.
DOI: https://doi.org/10.34641/clima.2022.127

1. Introduction

Fouling is the process of continuous accumulation of deposited undesirable particles on heat transfer surface areas (1). In heating appliances, such as combi-boilers, a plate heat exchanger (PHE) is essential in heating up domestic water to the desired temperature. The primary heat exchanger provides heat transfer by combusting fuel (e.g natural gas) to central heating (CH) water which is used to transfer heat to the domestic water via PHE, which consists of two channels that guide two immiscible fluids, CH and domestic water. Both sides of the PHE are affected by dirt accumulation from the system components. In addition, domestic water channel of the PHE is also affected by calcification due to calcium compounds in sanitary water. Crystallization and calcification are investigated by Lee et al. (2) and Pääkkönen et al. (3,4) The CH side of the PHE is mostly affected by the particles coming from system components such as the primary heat exchanger. Due to continuous interaction with water, corrosion occurs on the surface of the aluminium primary heat exchanger. The corrosion of primary heat exchanger and particles from the pipe systems cause particulate type fouling on this side of the plates. The main particle that is seen, although in small amounts, is AI2O3, along with other calcium compounds. There are several published experimental and numerical investigations into particulate and composite fouling. (5–8)

In recent years, one of the trending research topics centered around fouling prevention is modelling and prediction algorithms. These algorithms have mostly been based on statistical methods and machine learning techniques and include fouling prediction and detection algorithms based on support vector machines (SVM) (9,10), auto-associative kernel regression (AAKR) (11), autoregressive integrated moving average (ARIMA) (12) and artificial neural networks (ANN) (13–17). Model-based fouling prediction research have also been investigated by Kalman filter usage (18).

The predictive maintenance approach has also been looked at with an algorithm to predict fouling behaviour. The predictive maintenance techniques are designed based on mainly fault diagnostics using data analysis. The purpose of this algorithm determines the presence of anomalies and fault diagnostics generally focus on statistical approaches that provide classification and clustering. Most failure mechanisms can be associated with the degradation processes (19). The data acquisition process can be maintained by health system monitoring as in Ref. (20). The machine learning algorithms used for classification are Naïve Bayes, k-nearest neighbours (kNN), decision trees, random forests. In the Ref. (21) these algorithms are successfully studied to classify the faults of boilers by using simulated data. As a result, decision tree algorithm gave the best result with an accuracy of 97.8%. As can be seen from references, the machine learning techniques are generally used in HVAC industry especially in heat exchangers but when the open resources are considered the classification machine learning algorithms on PHEs have been investigated rarely ever.

In this study, a multi-classification study to determine fouling levels with the aim of generating a warning on combi-boiler appliances is carried out for compact brazed plate heat exchangers. The Naïve Bayes, kNN and decision tree machine learning algorithms are studied with cross-validation method for experimentally acquired data. A comparison of multi-classification machine learning algorithms, which are applied for detecting the fouling level of PHEs in the combi-boiler appliances, is provided as an introductory study.

2. Research methods

2.1 Experimental conditions

Algorithm development process starts with data acquiring regarding to get healthy and faulty conditions. To get a dataset for both conditions, just PHEs are tested. 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. The technical specifications that are generated from the design parameters, includes the working range information of CH and domestic water channel of PHEs. Specifications indicate that for the given inlet temperatures and flow rates, the 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 the reference and their performance is considered as zero-hour performance.

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 as 10°C. The CH inlet temperatures are applied in a range of 72±1°C. The first conditions of both PHEs represent the technical specifications of 30 and 32 plates. Therefore, the results with respect to first test conditions are considered as the initial status (at t_0) of fouling process. Other test conditions are the technical specifications of PHEs with 28, 26, 24, 22, 20, 18 and 16 plates, respectively. By performing this test procedure, 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 instead of the designed one. Thus, a representation of the clogging behavior has been demonstrated. 50% clogged PHE as maximum clogging percentage is considered in test condition 8 (Tab. 1), technical specifications of 16 plate PHE that are implied to 32 plate PHE.

Tab. 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)
Test 1	29	18	26	10.3
Test 2	26	10.3	25.1	10.1
Test 3	25.1	10.1	21.5	8.5
Test 4	21.5	8.5	21.5	8.7
Test 5	21.5	8.7	17	6.9
Test 6	17	6.9	17.2	6.9
Test 7	17.2	6.9	17.6	6.9
Test 8	17.6	6.9		

2.2 Experimental set-up

The experimental setup contains two lines that represent CH (orange line) and DHW (green line) circuits (Fig.1). CH line is a closed circuit, and the demand of hot water is met with a combi-boiler. CH water is supplied from a tank, which is heated by the combi-boiler's system circuit with a heater coil. The static pressure of the closed circuit is 2 ± 0.1 bar which is measured inside the tank. A pump circulates water 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. CH line is interacted with a cold- water

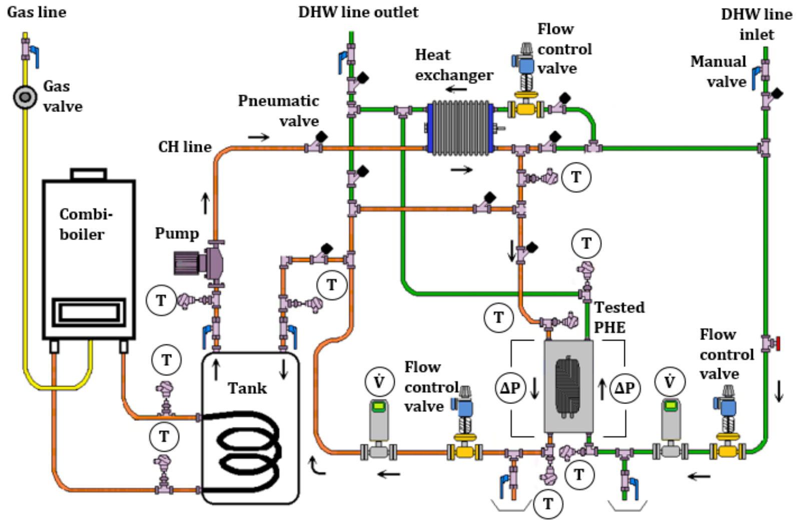

Fig. 1 – Schematic diagram of experiment set-up and its components

line to achieve the ability of necessary cooling with an additional plate heat exchanger. This cooling process control is carried out manually with the help of a flow control valve that is 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 the temperature of water at the inlets and outlets of the tested PHE beside the other important points such as tank inlet and outlet. The sensor locations can also be seen in Fig. 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 in the PHEs is functional performance decreasing. The accumulated fouling particles creates 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 to the thermal resistance concept. The logarithmic mean temperature difference (LMTD) method is used to calculate overall heat transfer coefficient (U) (Equation. (1)). Heat transfer rates of DHW and CH side 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 th average of the heat transfer rates of CH and DHW sides. This overall heat transfer coefficient calculation 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 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

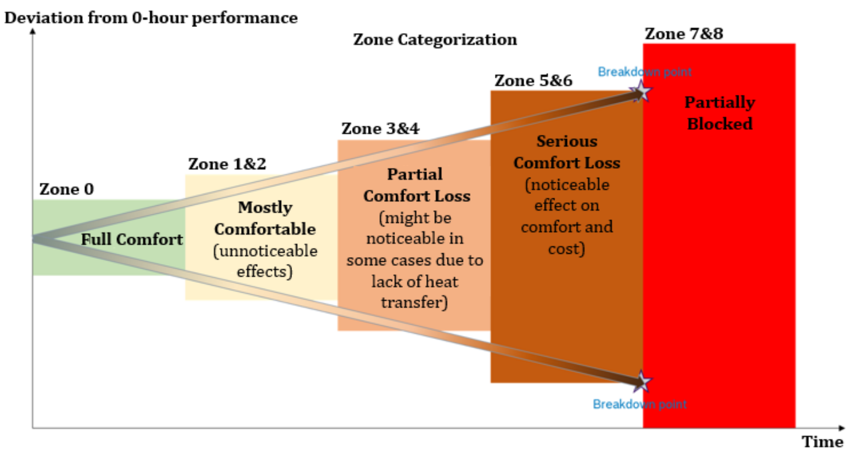

Fig. 2 – Zone categorization.

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. They 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 used main algorithm development method is fault diagnosis to obtain the current fouling status of the PHE. There are several machine learning techniques that can be applied to diagnose faults and current situations. Classification, which is implied in this study, is a type of supervised machine learning in which an algorithm learns to classify the 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 0-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 required to be changed. (Fig. 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

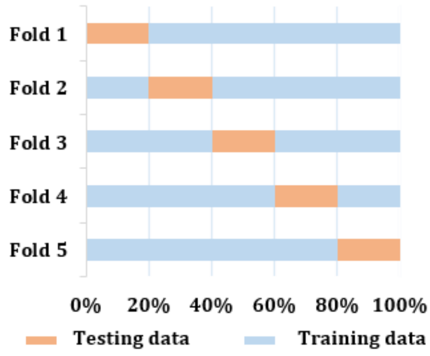

Fig. 3 – k-fold cross validation designation.

multi classification cases. The algorithms are applied to data using Classification Learner App in MATLAB. Before training of the algorithms, cross-validation is used in the process of creation 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 Fig. 3. One subset is used to validate the model that is trained using the remaining subsets (22). The average error across all k partitions is chosen to determine overall accuracy percentage. The k value is chosen as 5 in this study for all used algorithms.

Naïve Bayes is a classification algorithm that applies

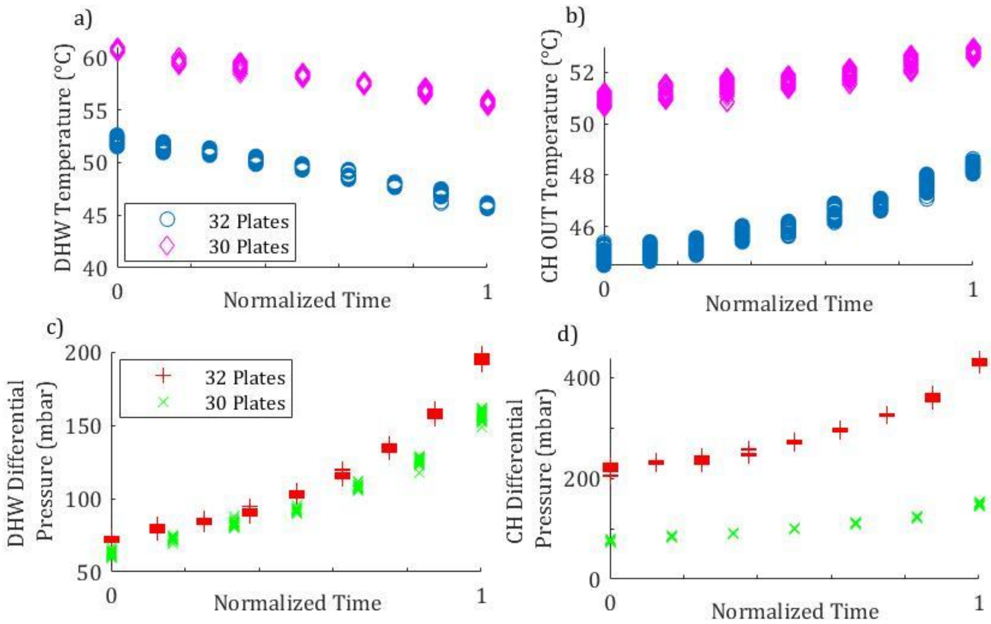

Fig. 4 - The performance trends for 32 and 30 plates a) DHW temperature behaviour b) CH outlet temperature behaviour c) Pressure drop behaviour of DHW d) Pressure drop behaviour of CH.

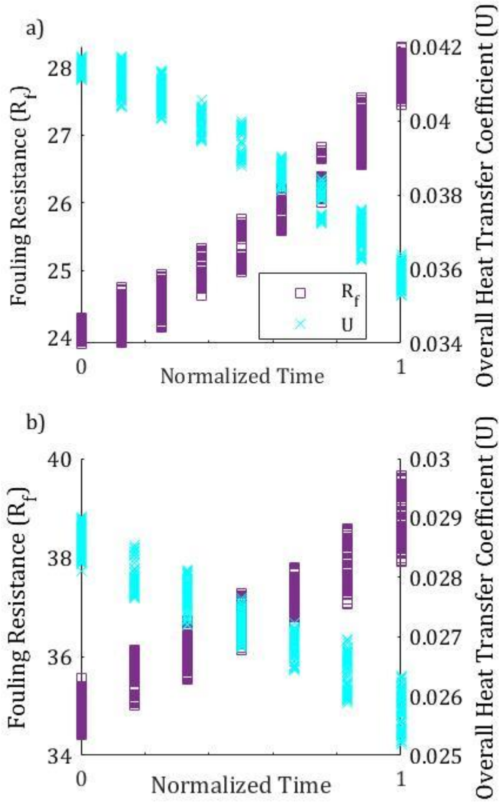

Fig. 5 – Fouling resistance and overall heat transfer coefficient behaviours a) for 32 plates b) for 30 plates.

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).

Given data for n number of points and a distance function, k-nearest neighbours (kNN) algorithm finds 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 a default in MATLAB Classification learner application. 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 split criterion in this study while the maximum number of splits is implied as 100.

3. Results and discussion

The experimental results are processed as grouped to be predicted. The outlet temperatures of PHE and the differential pressure between inlets and outlets of the PHE of both CH and DHW lines are presented in scatter plot for 30 and 32 plates (Fig. 4). 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.

Since clogging of plates results in performance decrease in PHEs, trends of the DHW outlet temperature for both 30 and 32 plates decrease as shown in Fig. 4. In contrast, trends of the CH outlet temperature and the pressure drops of both CH and DHW lines for 30 and 32 plates increase as shown in Fig. 4. The calculated overall heat transfer coefficients for both 30 and 32 plates are shown in Fig. 5. The trends of overall heat transfer coefficient are decreasing as an expected statement of the decreasing performance.

The fouling resistance values also follow the expected trends in contrast to the overall heat transfer coefficient as shown in Fig. 5. In addition to that, the fouling resistance graphs show similar trends for particulate and composite fouling behaviours of four PHEs having different geometries studied by Zhang et al. (6).

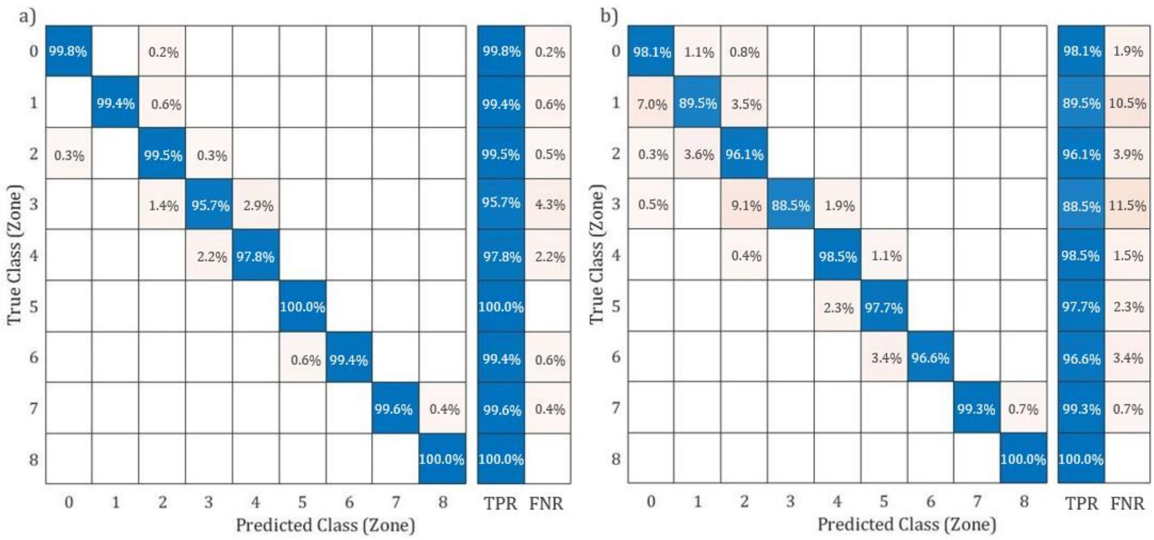

Fig. 6 – Confusion matrices of a) decision tree model b) kNN model.

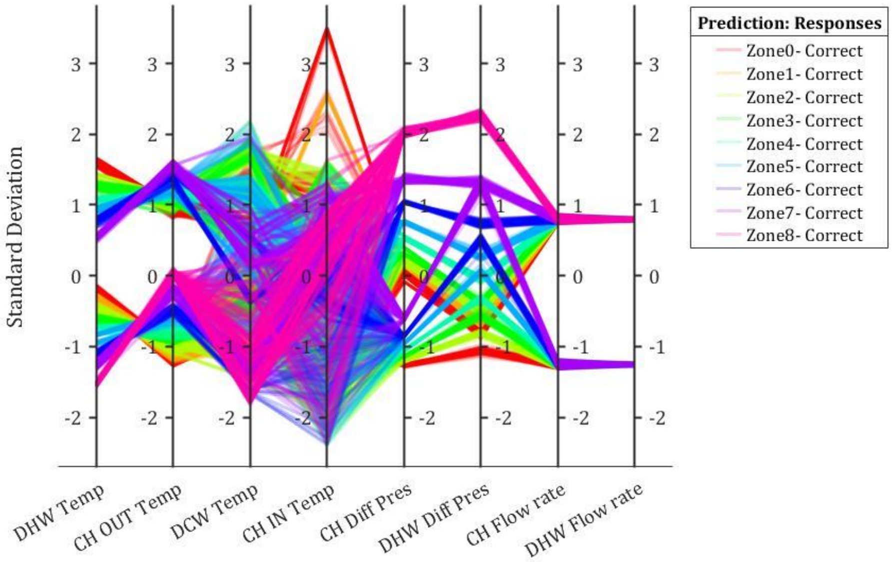

Fig. 7 – Parallel coordinates plot of experiment data shown in standard deviation scales for Naïve Bayes model.

The PHE indicated as 1st in reference study has similar geometric parameters with the ones used in this study. The Reynolds number range designated in the reference study is similar to the range in this study. 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 Fig. 6. Naïve Bayes is 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 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 Fig. 7. In the figure, the classes show more distinguishable distribution on the CH and DHW pressure difference data rather than CH inlet and DCW temperatures. This results in that the pressure difference is more convenient parameters to predict classes correctly rather than the other parameters. The DHW and CH outlet temperature data are also helpful to distinguish the classes according to the distribution shown in parallel coordinate plot (Fig. 7). With this representation of high dimensional experiment data as 2-dimensional visualization, the relation of standard deviation between the predictors can be seen.

4. Conclusion

In this study, an algorithm is developed to imply on combi-boiler appliances with the aim of generating a warning that indicates the fouling level of PHEs. Naïve Bayes, kNN and decision tree are used as the multi-classification machine learning 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. During creation of zones, 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 instead of the designed. The behaviours of the overall heat transfer coefficient and the fouling resistance in normalized time scale show the expected trends. The attempted models of machine learning 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 to generate a warning for current fouling level classification of PHEs in combi-boiler appliances by implying machine learning algorithms with high accuracy. Generation of fouling level warning results in the possibility to release 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 time-dependent dataset into account to assess optimum time schedule of maintenance in future studies.

5. Acknowledgement

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

6. References

1. T. Reg Bott, G. F. Hewitt, G. L. Shires. Process Heat Transfer. 1st ed. CRC-Press; 1994.

2. Lee E, Jeon J, Kang H, Kim Y. Thermal resistance in corrugated plate heat exchangers under crystallization fouling of calcium sulfate (CaSO4). International Journal of Heat and Mass Transfer. 2014 Nov 1;78:908–16.

3. Pääkkönen TM, Riihimäki M, Simonson CJ, Muurinen E, Keiski RL. Modeling CaCO3 crystallization fouling on a heat exchanger surface – Definition of fouling layer properties and model parameters. International Journal of Heat and Mass Transfer. 2015 Apr 1;83:84–98.

4. Pääkkönen TM, Riihimäki M, Simonson CJ, Muurinen E, Keiski RL. Crystallization fouling of CaCO3 – Analysis of experimental thermal resistance and its uncertainty. International Journal of Heat and Mass Transfer. 2012 Nov 1;55(23–24):6927–37.

5. Müller-Steinhagen HM, Middis J. Particulate fouling in plate heat exchangers. Heat Transfer Engineering. 1989;10(4):30–6.

6. Zhang GM, Li GQ, Li W, Zhang Z, Leng XL, Tian MC. Particulate fouling and composite fouling assessment in corrugated plate heat exchangers. International Journal of Heat and Mass Transfer. 2013 May 1;60(1):263–73.

7. Li W, Li HX, Li GQ, Yao SC. Numerical and experimental analysis of composite fouling in corrugated plate heat exchangers. International Journal of Heat and Mass Transfer. 2013 Aug 1;63:351–60.

8. Song KS, Lim J, Yun S, Kim D, Kim Y. Composite fouling characteristics of CaCO3 and CaSO4 in plate heat exchangers at various operating and geometric conditions. International Journal of Heat and Mass Transfer. 2019 Jun 1;136:555–62.

9. Sun L, Zhang Y, Saqi R. Research on the fouling prediction of heat exchanger based on Support Vector Machine optimized by Particle Swarm Optimization algorithm. In: 2009 International Conference on Mechatronics and Automation. IEEE; 2009. p. 2002–7.

10. Sun L, Zhang Y, Zheng X, Yang S, Qin Y. Research on the fouling prediction of heat exchanger based on support vector machine. In: Proceedings - International Conference on Intelligent Computation Technology and Automation, ICICTA 2008. 2008. p. 240–4.

11. Ardsomang T, Wesley Hines J, R. Upadhyaya B. Heat Exchanger Fouling and Estimation of Remaining Useful Life. Annual Conference of the PHM Society. 2021 May 29;5(1).

12. Nguyen H-P, Fauriat W, Zio E, Liu J. A Data-Driven Approach for Predicting the Remaining Useful Life of Steam Generators. In: 2018 3rd International Conference on System Reliability and Safety (ICSRS). IEEE; 2018. p. 255–60.

13. Riverol C, Napolitano V. Estimation of fouling in a plate heat exchanger through the application on neural networks. Journal of Chemical Technology and Biotechnology. 2005 May;80(5):594–600.

14. Mohanty DK, Singru PM. Fouling analysis of a shell and tube heat exchanger using local linear wavelet neural network. International Journal of Heat and Mass Transfer. 2014 Oct 1;77:946–55.

15. Wen X, Miao Q, Wang J, Ju Z. A multi-resolution wavelet neural network approach for fouling resistance forecasting of a plate heat exchanger. Applied Soft Computing. 2017 Aug 1;57:177–96.

16. Sundar S, Rajagopal MC, Zhao H, Kuntumalla G, Meng Y, Chang HC, et al. Fouling modeling and prediction approach for heat exchangers using deep learning. International Journal of Heat and Mass Transfer. 2020 Oct 1;159:120112.

17. Dheenamma M, Soman DP, Muthamizhi K, Kalaichelvi P. In pursuit of the best artificial neural network configuration for the prediction of output parameters of corrugated plate heat exchanger. Fuel. 2019 Mar 1;239:461–70.

18. Palsson OP, Jonsson GR, Desmet B, Lalot S, et al. Comparison of neural networks and Kalman filters performances for fouling detection in a heat exchanger [Internet]. Article in International Journal of Heat Exchangers. 2007. Available from: https://www.researchgate.net/publication/286315300

19. Wang W, Christer AH. Towards a general condition based maintenance model for a stochastic dynamic system. Journal of the

Operational Research Society. 2000;51(2):145–55.

20. SUN J, WANG F, NING S. Aircraft air conditioning system health state estimation and prediction for predictive maintenance. Chinese Journal of Aeronautics. 2020 Mar 1;33(3):947–55.

21. Shohet R, Kandil MS, McArthur JJ. Machine learning algorithms for classification of boiler faults using a simulated dataset. IOP Conference Series: Materials Science and Engineering. 2019 Sep 1;609(6):062007.

22. Cross-Validation - MATLAB & Simulink [Internet]. [cited 2022 Jan 3]. Available from: https://www.mathworks.com/discovery/cross-validation.html

23. Naive Bayes classification for multiclass classification - MATLAB [Internet]. [cited 2022 Jan 3]. Available from: https://www.mathworks.com/help/stats/classificationnaivebayes.html

24. Train Naive Bayes Classifiers Using Classification Learner App - MATLAB & Simulink [Internet]. [cited 2022 Jan 3]. Available from: https://www.mathworks.com/help/stats/Train-Naive-Bayes-Classifiers-Using-Classification-Learner-App.html

25. Classification Using Nearest Neighbors - MATLAB & Simulink [Internet]. [cited 2022 Jan 3]. Available from: https://www.mathworks.com/help/stats/classification-using-nearest-neighbors.html

26. Wang, Po-Wei C-JL. Data Classification: Algorithms and Applications [Internet]. Aggarwal CC, editor. C. Aggarwal eds. Data Classification: Algorithms and Applications,Data Mining and Knowledge Discovery Series. Chapman & Hall/CRC; 2014 [cited 2022 Jan 3]. 187–202. Available from: http://www.crcnetbase.com/doi/book/10.1201/b17320

Data Statement

The datasets generated during and/or analysed during the current study are not available due to company policy restrictions based on company profits.