Detection of the low ΔT syndrome using machine learning models

Anand Thamban [a], *Alet* van den Brink [b], *Shalika* Walker [a], *Wim* Zeiler [a]

[a] The Department of the Built Environment, Eindhoven University of Technology, Eindhoven, the Netherlands, a.thamban@tue.nl , s.s.walker@tue.nl , w.zeiler@tue.nl

[b] Afdeling O&T, Kropman Installatietechniek B.V., Nijmegen, the Netherlands, alet.van.den.brink@kropman.nl

Abstract. The low ΔT syndrome has been a prevalent issue in many chilled water systems, leading to an increase in the pump energy consumption, increase in the chiller energy consumption, and/or failure to meet the cooling loads. It is therefore important to detect the low ΔT syndrome using suitable fault detection and diagnosis methods. One such fault detection method is the data-based approach using machine learning algorithms. The main signs indicating the low ΔT syndrome include a reduced return water temperature from the cooling coil and an increased mass flow rate through the cooling coil. Since the mass flow rate of water is not measured in all chilled water installations, the cooling coil valve position is measured instead. This research aims to compare the performance of different machine learning regression models which predict the return water temperature and the cooling coil valve position, based on the R^2 score and root mean square error. The different machine learning algorithms compared for the study include Support Vector Regression, Artificial Neural Network and eXtreme Gradient Boosting. The data required for the analysis was obtained from fault-introduced experiments conducted in an office building. The different fault cases include stuck cooling coil valve at 50%, stuck cooling coil valve at 75%, reduced supply air temperature by 2K and reduced supply air temperature by 1K. The regression models are expected to predict the fault-free data ($X_{predicted}$) of the system such that faulty data (X_{actual}) can be identified with residuals ($X_{predicted} - X_{actual}$). The results showed that XGBoost was the best performing algorithm in terms of model accuracy. The XGBoost based prediction models for return water temperature and cooling coil valve position were able to successfully detect anomalies for 3 out of the 4 fault cases.

Keywords. Low ΔT syndrome, XGBoost, ANN, SVR, fault detection and diagnosis.
DOI: https://doi.org/10.34641/clima.2022.121

1. Introduction

The built environment contributes to about 35% of the total energy consumption in the Netherlands [1]. With an increasing trend of warming witnessed every year, the cooling demand is expected to increase in the European continent [2]. To cope with this rising cooling demand, the energy consumption for cooling is also expected to increase. Around 75% of this cooling energy use comes from Heating, Ventilation and Air-Conditioning (HVAC) systems [1]. The HVAC system aims to maintain thermal comfort and the required indoor air quality for human occupation. Research has shown that retro-commissioning of HVAC systems can lead to a 5-15% savings in energy use [3,4]. Therefore, there is a potential to reduce about 1.3 - 3.9% of the total energy consumption in the Netherlands if all systems are continuously monitored for faults, and frequent replacements of components are made to ensure that they run efficiently.

The low ΔT syndrome is an infamous phenomenon related to the chilled water system in a building/plant, and if left unaddressed, can have a negative impact on the energy consumption of the system. The problem has been widely discussed by many researchers, some of whom have identified its presence in almost all big distributed chilled water systems and suggested alternative design changes [5] and some who have identified its main causes, as well as solutions to mitigate the problem [6]. In a chilled water system, the capacity of the cooling coil is mainly attributed to the water-side temperature difference during part-load conditions. A smaller temperature difference between the supply water and return water will lead to an inefficient chilled water system, reducing cooling output and causing energy wastage to operate extra chillers and pumps to keep up with demand [5]. This phenomenon of a

Copyright ©2022 by the authors. This conference paper is published under a CC-BY-4.0 license.

reduced temperature difference across the cooling coil with an increased demand of flow to keep up with system demand is called the low ΔT syndrome [5].

The low ΔT syndrome can be caused by abrupt faults (control failure, physical failure etc.), incipient faults (fouling, equipment degradation) or design faults (three-way valves, improper coil selection) [6]. In this study, only the abrupt faults will be addressed.

One of the ways to identify and solve the low ΔT syndrome is by using fault detection and diagnosis (FDD) tools. FDD tools are used in the maintenance of building installations with the main purpose of detecting faults and diagnosing them, such that corrective measures can be taken to fix the system. The aim of an FDD tool is to detect a fault (in this case, the low ΔT syndrome) by observing certain signs (reduced return water temperature (RWT), increased mass flow rate etc.) and then diagnose the causes (stuck cooling coil valve, reduced supply air temperature (SAT)) leading to the fault. Few studies have been conducted which used grey box models to detect low ΔT syndrome [7,8] but they are limited to fouling related faults and/or just the low ΔT syndrome in large distributed chilled water plants.

FDD processes can be classified into data-driven and knowledge-driven methods [9]. Data-driven methods include classification and regression models which rely on large amounts of data with the capability of learning complex patterns from it. But these methods are largely black-box thus making it difficult to interpret what goes on behind the model. Knowledge-driven models include Bayesian networks, fuzzy logic etc., and are largely supported by expert knowledge in the particular field.

Using the data-driven methods, anomaly detection is a popular way of detecting faults in the data by identifying unexpected or abnormal data from normal fault-free data. This is usually done using supervised learning regression models which predict fault-free data and is then compared with the measured data. When large residuals are identified between the expected value (fault-free) and the measured value (faulty), it can be assumed that a fault exists in the system. Different kinds of machine learning (ML) algorithms have been used for regression purposes including Artificial Neural Networks (ANN), Support Vector Machines (SVM) or Support Vector Regression (SVR), Decision Trees, Random Forest and eXtreme Gradient Boosting (XGBoost). Each of these algorithms is advantageous over the other depending on the characteristics of the dataset.

SVM or SVR is an algorithm used for both classification as well as regression problems. It has the advantage of performing well with a limited amount of data compared to other models. But the computational time required for model development is considerably higher than other ML algorithms like ANN and Random Forest [10]. Decision trees are also regression algorithms that are based on the approach of splitting a dataset while evaluating certain conditions. Ensemble algorithms are based on the ML theory that a group of weak learners create a much stronger ensemble than a single strong learner [11]. XGBoost is one such ensemble algorithm that has proven to be a well-performing ML algorithm in several studies [12–14] and has been previously used for fault detection in HVAC systems [15]. Since previous research clearly showed the benefits of ensemble algorithms compared to individual Decision Trees [11], Decision Trees are not included in this study. ANN has also been used to develop regression models to predict continuous variables like energy consumption [10], temperature [16] and cooling coil valve position [17]. But ANN is more complex in nature compared to SVR and requires precise adjustment of its many hyper-parameters [18]. Neural networks also perform better with larger amounts of data, which could be a drawback if limited data is available [18].

Since each of the algorithms has its advantages and disadvantages, it is necessary to compare their performance to see which algorithm can make predictions with the least amount of error, which is an essential factor for anomaly detection. The purpose of the study is to develop an FDD tool to detect the low ΔT syndrome using a suitable ML algorithm which can be replicated in most kinds of chilled water systems. It is important to note that in this research only the fault detection aspect of the low ΔT syndrome is addressed and not the fault diagnosis part.

2. Methodology

The low ΔT syndrome is detected by two main phenomena: the increase in mass flow rate through the cooling coil and the decrease in the RWT from the cooling coil. Since many installations do not have mass flow rate meters, one way to get an idea of the flow demand from the system is to observe the cooling coil valve position (CCVP). Therefore, the low ΔT syndrome can be detected by monitoring the CCVP and the RWT.

The CCVP and RWT prediction models were developed by following a structured procedure shown in Fig. 1. The different steps are as follows:

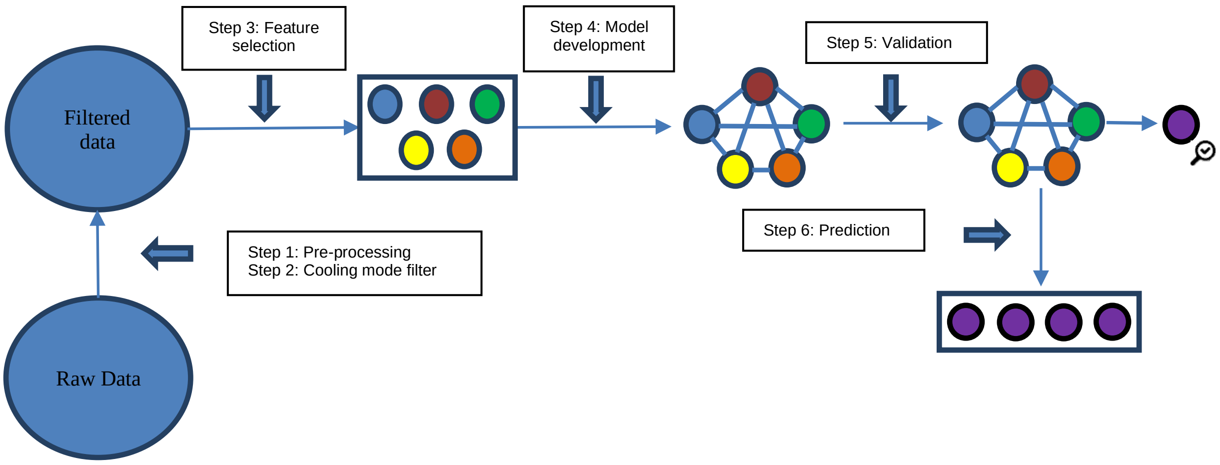

Fig. 1 - ML model development process

Step 1: Pre-processing of raw data to remove noise and missing data
Step 2: Filtering of data for cooling mode
Step 3: Feature selection
Step 4: Model development
Step 5: Validation and error calculation
Step 6: Prediction

If the CCVP prediction model showed a positive residual (residual = actual value – predicted value) and the RWT prediction model showed a negative residual, then it could be said that low ΔT syndrome was detected in the system.

2.1 Description of raw dataset

The analysis in this study was done using raw data collected from an office building located in Breda, the Netherlands. The office building is a living lab with multiple sensors placed both in the indoor environment as well as the HVAC installations. The building offers an ideal environment for conducting studies related to occupant comfort, energy demand predictions and HVAC fault simulations.

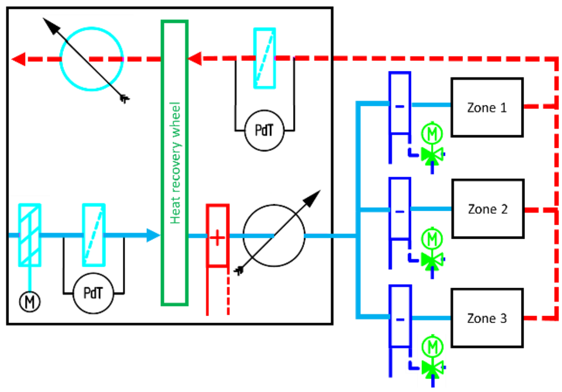

Fig. 2 - Schematic of the AHU in the use case building

The building has an air handling unit (AHU) in a constant air volume system which supplies conditioned air to three zones. The cooling coils are situated outside the AHU and act as after-coolers. Fig. 2 shows the schematic of the AHU and the zones supplied with conditioned air.

The raw data used in the study was collected from January 2017 up to September 2021. In this setup, extra sensors were installed near the cooling coil to measure the mass flow rate of water, inlet and outlet water temperatures and pressure difference of air over the cooling coil. These sensors were installed in March 2021, therefore the raw data used for RWT prediction was available only from April 2021 to September 2021. The mass flow rate sensor is not used in the fault detection process because it is not usually present in all chilled water installations. The temperature sensors used to measure the inlet and outlet water temperatures are Kamstrup Pt500 sensors which have a reaction time of 5s, and a maximum deviation of 0.1K for a temperature difference of 15K (i.e. ± 0.1K) between the inlet and outlet ports.

2.2 Fault simulation

Certain faults were introduced into the HVAC system from June 2021 to September 2021 in order to collect faulty data. These faults include:

1. Stuck cooling coil valve (50% and 75%)
2. Reduced SAT (1K and 2K)

These specific faults were introduced into the system because they were observed from computer simulations to be the abrupt faults that have the highest impact on the energy consumption of the pumps as well as occupant comfort.

2.3 Step 1: Pre-processing of raw data to remove noise and missing data

Before the model was developed, the data collected from the building management system (BMS) was cleaned to remove noisy and missing data points. Data for a particular timestamp where one or more features were absent were completely dropped. Interpolation was not done to fill in the missing data points because it could lead to incorrect values being fed into the dataset, leading to a faulty training dataset.

Moreover, some features in the dataset had incorrect values due to errors in the reporting system from the BMS, e.g., the CCVP signal was multiplied by a factor of 10. This had to be scaled down to be within the limits of 0 to 100%.

2.4 Step 2: Filtering of data for cooling mode

Once the data had been pre-processed, the next step was to ensure that the correct data was used for the ML model development. Since the focus of the study was to detect the low ΔT syndrome which occurs only when the AHU operates in cooling mode, it is more efficient to filter data when the AHU operates in cooling mode and then use the filtered data for the model development. The data was filtered using two conditions applied together:

1. The CCVP signal > 0
2. The chiller water outlet temperature < 8°C

Condition 1 is more generalizable since almost all HVAC installations have sensors to measure the cooling coil valve position. A value greater than 0 indicates that chilled water is demanded by the system and cooling is required.

Condition 2 is another criterion to see if the system is in cooling mode. The outlet water temperature from the chiller is a good indication of whether cooling is required by the system or not.

2.5 Step 3: Feature selection

To have a good quality prediction, it is necessary to choose the features which have the most impact on the predicted variable. In the case of fault detection, it is also necessary to avoid those features which can be influenced by the fault, e.g., a stuck valve fault can impact the SAT coming from the outlet of the cooling coil. If the SAT is used as a feature, then the prediction from the model would be biased to the fault hence giving a smaller residual than usual.

Feature selection can be done either manually by looking at a cross-correlation heat map and selecting the features with the highest cross-correlation score or by automatically selecting them using Recursive Feature Elimination using Cross-Correlation (RFECV). Both methods were used in this research.

The main features available for the study include inlet air temperature, inlet air relative humidity, inlet air absolute humidity, airflow rate, chilled water supply temperature, outdoor air temperature, SAT setpoint for a zone, measured and setpoint return air temperature, pressure across the supply and return filters, and fan speed.

2.6 Step 4: Model development

The purpose of the model is to predict the CCVP and cooling coil RWT in a fault-free mode such that when faults occur in the system, large residuals are generated to indicate faults in the system. The model was trained with fault-free data in the system to ensure the predictions are always fault-free. Three different ML models namely SVR, ANN and XGBoost were developed and compared in Python using the sci-kit learn library and XGBoost library.

2.7 Step 5: Validation

All the ML algorithms were validated using the k-fold cross-validation method [10]. K-fold cross-validation is a commonly used statistical method to determine the performance and accuracy of an ML model. In this method, the dataset is split into k folds (groups), where one group is held as a test set and the rest (k-1) groups are held as the training set. The model is then fit with the training set, evaluated with the test set and discarded while retaining the values. The process is then continued for the other k-1 test groups, eventually giving a summarized score of the model. In this study, 10-fold cross-validation is used.

The accuracy of the model was identified by two metrics [10]: the root mean square error (RMSE) and the coefficient of determination (R^2).

RMSE is used to measure the differences between values predicted by an estimator and the actual values observed. It is a measure of accuracy to compare forecasting errors of different models for a particular dataset and not between datasets, as it is scale-dependent. The RMSE is calculated using equation (1), where $\hat{y}_i$ is the estimated value, y_i is the actual value and N is the number of samples.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{N}} \tag{1}$$

The coefficient of determination is a statistical parameter that measures the proportion of variation in the dependent variable that is predictable from the independent variable. It is a measure of how well the model is able to replicate outcomes, based on the proportion of total variation of all the outcomes the model replicates. The R^2 value is determined using equation (2) where $\bar{y}_i$ is the mean value.

$$R^2 = \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y}_i)^2} \tag{2}$$

2.8 Step 6: Prediction

The models developed using the different algorithms generated predictions for the same data set. They were then evaluated and compared based on the average RMSE and R^2 value from k-fold cross validation.

The model with the best performing algorithm was then used to detect anomalies in the fault simulated experimental data.

3. Results

The results section is divided into two parts: the comparison of different ML algorithms for the prediction of CCVP and RWT, and the detection of low ΔT syndrome using the best performing algorithm. The comparison of the different ML algorithms is done based on the R^2 score and the RMSE.

3.1 Comparison of ML algorithms

The different ML algorithms were compared for two models – CCVP prediction and RWT prediction. Fig. 3 shows the comparison of R^2 values between different algorithms for the prediction of RWT.

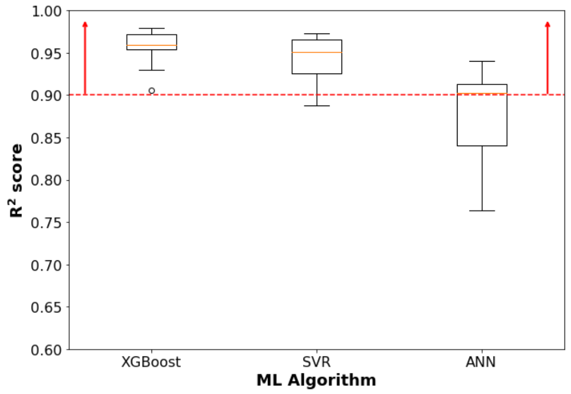

Fig. 3 - Comparison of R^2 values for RWT prediction

The boxplot shows the results obtained with the k-fold cross validation. For a good regression model with lower chances of false positives or true negatives, it is ideal to have R^2 values larger than 0.9 (red dotted line). It is evident that XGBoost and SVR perform well with their median values close to 0.95. ANN on the other hand has its interquartile range between 0.84 and 0.91. XGBoost, therefore, performed the best in terms of R^2 value for RWT prediction.

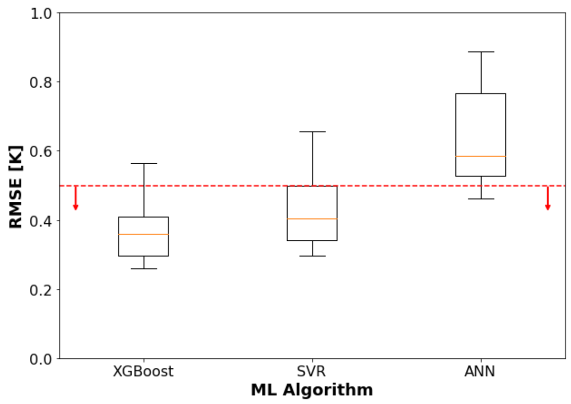

Fig. 4 - Comparison of RMSE values for RWT prediction

Fig. 4 shows the comparison of RMSE values between different algorithms for the prediction of RWT. XGBoost shows to produce the least amount of error, compared to the other algorithms. Since the RWT usually reduces by small values, it is essential to choose a model with relatively low error, so that the model can detect anomalies in the data if faults exist. Therefore, a threshold of 0.5K is chosen for selecting the appropriate algorithm.

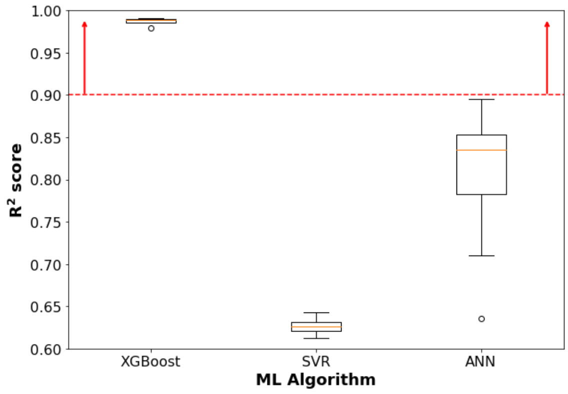

Fig. 5 - Comparison of R^2 values for CCVP prediction

Fig. 5 shows the comparison of R^2 values between different algorithms for the prediction of CCVP. In the figure, it is seen that XGBoost is the only model which performs well, with R^2 values above the threshold of 0.9, whereas the other models perform poorly.

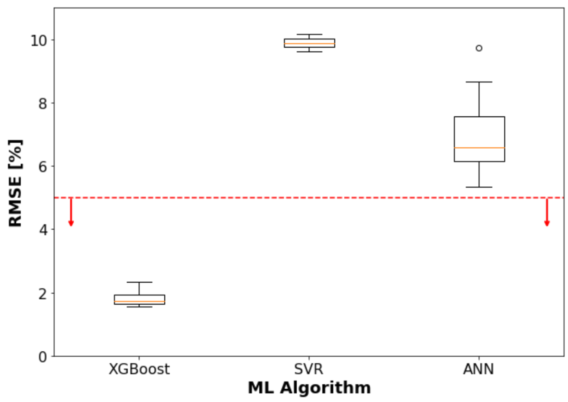

Fig. 6 - Comparison of RMSE values for CCVP prediction

Fig. 6 shows the comparison of RMSE values between different algorithms for the prediction of CCVP. XGBoost shows to have the least amount of error compared to the other models. A threshold of 5% valve position units is chosen to ensure a minimum number of false positives and true negatives among the predictions.

3.2 Fault detection

From the results shown in the previous subsection, it was found that XGBoost was the best performing model in terms of the R^2 value as well as the RMSE. The next step is to use the developed RWT and CCVP prediction model on the chilled water system to detect the low ΔT syndrome.

Fig. 7 shows the comparison between the predicted RWT (red line) and the actual RWT (blue line) for the fault use case where the CCVP is stuck at 75%. The generated negative error residual (green line) is greater in magnitude than the threshold of 0.5K and clearly shows the presence of a fault in the system.

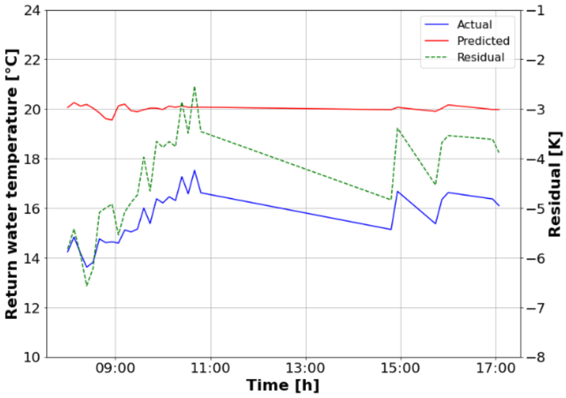

Fig. 7 - RWT prediction for CCVP stuck at 75%

Fig. 8 shows the comparison between the predicted CCVP and the actual CCVP for the same fault use case where the CCVP is stuck at 75%.

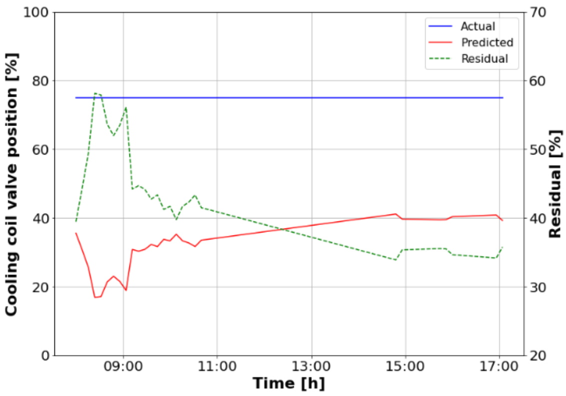

Fig. 8 - CCVP prediction for CCVP stuck at 75%

The generated positive error residual is greater than the threshold of 5% and clearly shows the presence of a fault in the system.

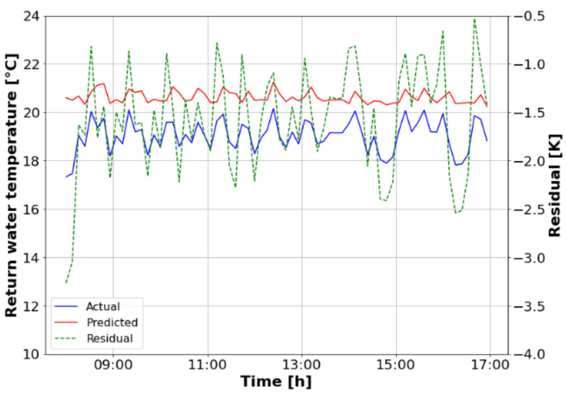

Fig. 9 - RWT prediction for CCVP stuck at 50%

Fig. 9 shows the comparison between the predicted RWT and the actual RWT for the fault use case where the CCVP is stuck at 50%. The generated negative error residual is on an average greater in magnitude than the threshold of 0.5K and therefore shows the presence of a fault in the system.

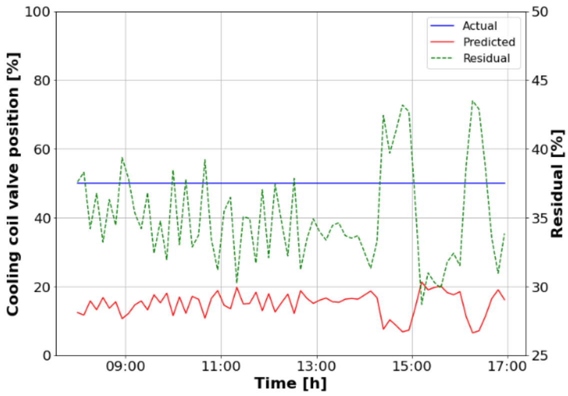

Fig. 10 - CCVP prediction for CCVP stuck at 50%

Fig. 10 shows the comparison between the predicted CCVP and the actual CCVP for the fault use case where the CCVP is stuck at 50%. The generated positive error residual is greater than the threshold of 5% and therefore clearly shows the presence of a fault in the system.

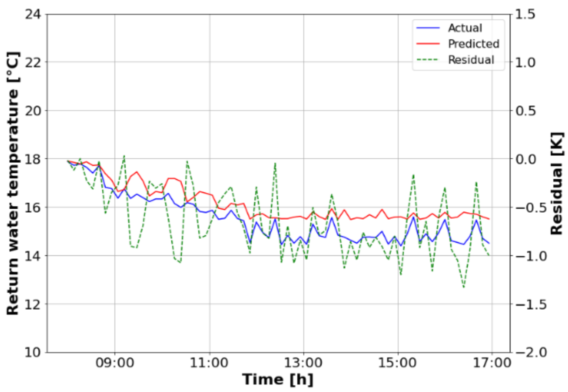

Fig. 11 - RWT prediction for reduced SAT by 2K

Fig. 11 shows the comparison between the predicted RWT and the actual RWT for the fault use case where the SAT setpoint is reduced by 2K. The generated negative error residual on an average is greater in magnitude than the threshold of 0.5K and therefore shows the presence of a fault in the data.

Fig. 12 shows the comparison between the predicted CCVP and the actual CCVP for the fault use case where the SAT setpoint is reduced by 2K. The generated positive error residual is greater than the threshold of 5% for certain periods (8 am to 12 pm and 3 pm to 5 pm) and therefore shows the presence of a fault in the above-mentioned time period, although the fault was present throughout whole period (7 am to 5 pm). The prediction (12 pm to 3 pm) becomes very close to the actual measured value reducing the chances of fault detection.

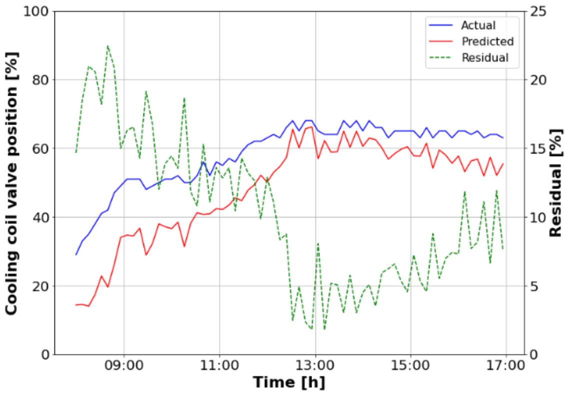

Fig. 12 - CCVP prediction for reduced SAT by 2K

Fig. 13 shows the comparison between the predicted RWT and the actual RWT for the fault use case where the SAT setpoint is reduced by 1K. It is observed that the predictions are quite close to the actual value and therefore sufficiently large residuals are not generated.

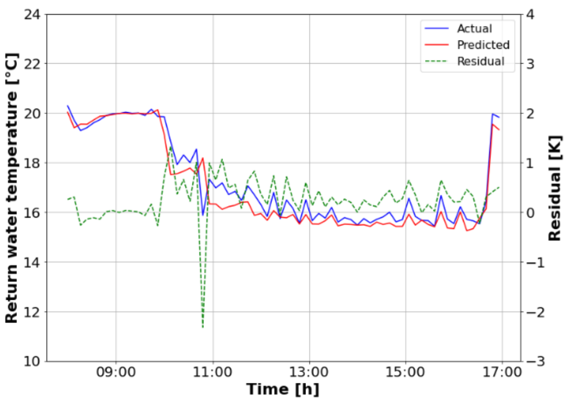

Fig. 13 - RWT prediction for reduced SAT by 1K

Fig. 14 shows the comparison between the predicted CCVP and the actual CCVP for the fault use case where the SAT setpoint is reduced by 1K.

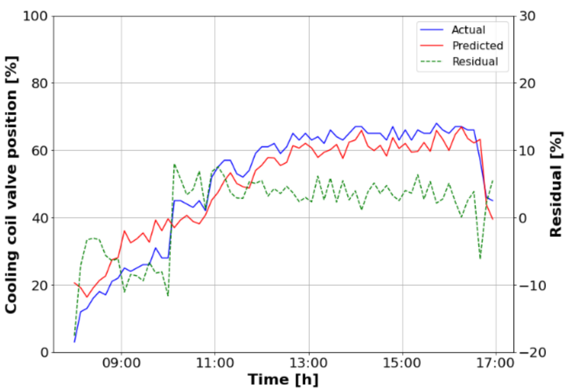

Fig. 14 - CCVP prediction for reduced SAT by 1K

Similar to the RWT prediction model, it is observed that the predictions are close to the actual value and therefore sufficiently large residuals are not generated. The reduction of the SAT by 1K does not seem to have much of an impact on the CCVP either.

4. Discussion

This study compared the capability of different ML algorithms in the detection of low ΔT syndrome using two different fault conditions. From the comparison, it is evident that XGBoost performed the best. SVR also performed well for a limited amount of data, which is seen in the RWT prediction. But its performance drops when large amounts of data are used. ANN did not perform well in all cases and did not cross the required thresholds for model acceptance. Apart from the model performance, the model development process for an ANN was also more complicated than XGBoost or SVR. ANN did not support RFECV thus requiring manual feature selection. This is a time-consuming and inefficient method of identifying the correct features. The automated feature selection capability supported by XGBoost and SVR is highly desirable, especially for commercial implementation of the FDD tool.

The XGBoost prediction model for RWT and CCVP was tested for different fault use cases conducted at the experiment site. The models show the ability to generate predictions (positive for CCVP and negative for RWT) with sufficiently large residuals which indicate the presence of the low ΔT syndrome in the system. The cooling coil stuck valve faults show a significant impact on the RWT (residual > 0.5K) and CCVP prediction (residual > 5% CCVP). The reduction of SAT by 2K fault on the other hand did not show very large residuals but were large enough for certain periods throughout the day indicating the presence of a fault. The reduction of SAT by 1K had little to no impact on the RWT or the CCVP and therefore did not show any anomalies. This is because the reduction of SAT by 1K did not lead to a significant increase in the mass flow rate.

It is important to note that during the experiments for reduced SAT, the outdoor air dry-bulb temperature did not go above 23°C. Therefore, there wasn't a very high cooling demand from the building. The reduction in SAT would have caused a small increase in the mass flow rate which would be difficult to detect. This could be a possible explanation for the smaller impact of the reduced SAT faults on RWT and CCVP.

5. Conclusions

This study shows that XGBoost is a suitable model for predicting RWT and CCVP in HVAC installations and can be confidently used for fault detection of low ΔT syndrome. Even though the low ΔT syndrome was detected in 3 of the 4 fault cases, the impact and intensity of the fault varied for each use case. The stuck cooling coil valve fault had a very high impact on RWT and CCVP and therefore was easily detected by the system. The reduced SAT fault on the other hand had a smaller impact.

6. Recommendations

It would be beneficial to conduct further experiments for reduced SAT when the outdoor air dry-bulb temperature is higher than 25 °C. This is also a more practical fault simulation since usually the SAT would be reduced only if the cooling demand is very high on hot days. The extension of the current study to another building would also be valuable to confirm the trend of the findings.

7. Acknowledgement

The research work is funded by the TKI Urban Energy project "Continuous commissioning of low-dt" and the Eindhoven Engine project "CM-HVAC-FDD". We would like to express our gratitude to all the other partners in the project. A special thanks to Roel Vos and John Verlaan of Kropman Breda for helping in setting up and coordinating with the experiments at the use case building.

8. References

[1] Agentschap NL. Monitor Energiebesparing Gebouwde Omgeving. 2019.

[2] Lhotka O, Kyselý J, Farda A. Climate change scenarios of heat waves in Central Europe and their uncertainties. Theor Appl Climatol 2018;131:1043–54.

[3] Friedman H, Piette MA. Comparative guide to emerging diagnostic tools for large commercial HVAC systems. Lawrence Berkeley Natl Lab 2001.

[4] Mills E. Building commissioning: A golden opportunity for reducing energy costs and greenhouse gas emissions in the United States. Energy Effic 2011;4:145–73.

[5] Kirsner W. The demise of the primary-secondary pumping paradigm for chilled water plant design. HPAC Heating, Piping, Air Cond 1996;68:73–8.

[6] Taylor ST. Degrading chilled water plant delta-T: Causes and mitigation. ASHRAE Trans 2002;108 PART 1:641–53.

[7] Gao D ce, Wang S, Sun Y, Xiao F. Diagnosis of the low temperature difference syndrome in the chilled water system of a super high-rise building: A case study. Appl Energy 2012;98:597–606.

[8] Gao D ce, Wang S, Shan K, Yan C. A system-level fault detection and diagnosis method for low delta-T syndrome in the complex HVAC systems. Appl Energy 2016;164:1028–38.

[9] Zhao Y, Li T, Zhang X, Zhang C. Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future. Renew Sustain Energy Rev 2019;109:85–101.

[10] Walker S, Khan W, Katic K, Maassen W, Zeiler W. Accuracy of different machine learning algorithms and added-value of predicting aggregated-level energy performance of commercial buildings. Energy Build 2020;209:109705.

[11] Zhou Z-H. Machine Learning. Springer; 2021.

[12] Pan B. Application of XGBoost algorithm in hourly PM2.5 concentration prediction. IOP Conf Ser Earth Environ Sci 2018;113.

[13] Mo H, Sun H, Liu J, Wei S. Developing window behavior models for residential buildings using XGBoost algorithm. Energy Build 2019;205:109564.

[14] Yao L, Cai M, Chen Y, Shen C, Shi L, Guo Y. Prediction of antiepileptic drug treatment outcomes of patients with newly diagnosed epilepsy by machine learning. Epilepsy Behav 2019;96:92–7.

[15] Chakraborty D, Elzarka H. Early detection of faults in HVAC systems using an XGBoost model with a dynamic threshold. Energy Build 2019;185:326–44.

[16] Montazeri A, Kargar SM. Fault detection and diagnosis in air handling using data-driven methods. J Build Eng 2020;31:101388.

[17] Wang S, Jiang Z. Valve fault detection and diagnosis based on CMAC neural networks. Energy Build 2004;36:599–610.

[18] Seyedzadeh S, Rahimian FP, Glesk I, Roper M. Machine learning for estimation of building energy consumption and performance: a review. Vis Eng 2018;6.

Data Statement

The datasets generated during and/or analysed during the current study are not publicly available because they are provided by a company and has to be kept private owing to privacy and security concerns but are/will be available if the concerned parties contact us and explain the reason for the data requirement, but the final decision lies with the company and access cannot always be guaranteed.