2025 European Conference on Computing in Construction

CIB W78 Conference on IT in Construction

Porto, Portugal
July 14-17, 2025

DEVELOPMENT OF A DYNAMIC GREY-BOX A CONDENSING BOILER EMULATOR

Marjan Fatehijananloo[1], J. J. McArthur [1]
[1]Dept. Architectural Science, Toronto Metropolitan University, Toronto, ON

Abstract

HVAC systems significantly impact building energy use and faults can significantly reduce efficiency. While data-driven fault detection and diagnosis techniques are well-established, fault emulation is critical, given the lack of available training data. This study develops a condensing boiler emulator using a hybrid gray-box approach combined with Bayesian optimization to estimate heat transfer coefficients under varying conditions. Testing demonstrates that the model accurately matches nominal data and can reliably simulate faulty conditions such as fouling, scaling, and excess air. This contribution provides a valuable tool for fault analysis, improving the optimization and reliability of HVAC systems.

Introduction

Heating, ventilation, and air conditioning (HVAC) systems account for more than one-third of total building energy consumption and emissions, making them a key focus of global energy and environmental efforts (IPCC, 2022). However, factors such as inadequate maintenance, component failures, and control malfunctions often reduce HVAC efficiency (Zhao, et al., 2019). As a result, fault detection and diagnosis (FDD) has become a crucial strategy for mitigating performance degradation, lowering maintenance costs, and extending the lifespan of HVAC components (Shohet, et al., 2020).

Condensing boilers are increasingly used for heating due to their high efficiency, 99%+ vs 75-80% for non-condensing boilers, arising from ability to recover latent heat from exhaust gases. However, they are particularly vulnerable to operational faults—such as fouling, scaling, and inadequate flow rates—that can disrupt the condensation process and negate efficiency gains (Baladi, et al., 2017). While data-driven FDD techniques are increasingly used to detect such faults (Singh, et al., 2022), their development and validation often require large fault datasets, which are rarely available in practice (Li & O'Neill, 2018). Consequently, modeling and simulation frameworks that can emulate both nominal and faulty states of condensing boilers have emerged as a cost-effective solution.

Over the past decade, several studies have advanced the modeling of condensing and related boiler systems, primarily under ideal or near-ideal conditions. For instance, Satyavada and Baldi (2016) introduced a hybrid dynamic model for condensing boilers that improved upon traditional steady-state efficiency estimates by incorporating thermal transients. While their approach enhanced performance modeling, it did not include fault simulations or a diagnostic framework. Similarly, Antonescu and Stănescu (2017) developed a thermodynamics-based computational model validated against a single experimental dataset. Their model demonstrated low error margins under nominal conditions but lacked the flexibility to simulate faulty scenarios. Beyond individual boiler modeling, other studies have explored broader energy recovery and district heating configurations. Koppauer et al. (2017) investigated waste heat recovery in diesel engines using an Organic Rankine Cycle, focusing on optimizing power generation rather than diagnosing faults in condensing heat exchangers. Similarly, Simic et al. (2021) developed a dynamic thermodynamic model for gas-fired condensing boilers using Modelica, assessing environmental impacts such as humidity and air temperature but without addressing fault simulation.

While these studies represent important advancements in condensing boiler modeling, they primarily emphasize steady-state or nominal performance and do not fully address real-world degradation and faults. Many building energy modeling tools also assume fault-free operation, which oversimplifies the reality of HVAC system malfunctions (Li & O'Neill, 2018; Qian, Li, Niu, & O'Neill, 2018). Consequently, there is a clear need for a condensing boiler emulator that not only captures complex thermal behavior — including latent heat recovery — but also accommodates fault simulation and diagnosis. Developing such a tool is essential for building robust FDD systems that can operate effectively under diverse and realistic conditions. The integration of model simulation with a gray-box modeling approach has been one of the most effective methods in recent times for the identification and prediction of operational issues within HVAC systems (Singh, et al., 2022). Gray-box models make use of physics-based principles combined with machine learning techniques in a balanced framework to provide an accurate representation of system dynamics, integrating empirical data for fault detection and diagnosis (Afram & Janabi-Sharifi, 2015).

This paper addresses these gaps by introducing a condensing heat exchanger emulator—developed using a hybrid gray-box framework and Bayesian optimization—that accurately models both normal and faulty boiler states. Built upon physics-based thermodynamic models, it incorporates fouling, scaling, and excess air deviations to simulate faulty operations. The model is validated against manufacturer performance data (Viessmann, 2008), demonstrating high accuracy across varying temperatures and operating modes. Furthermore, its ability to generate labeled fault data helps address the scarcity of fault datasets in the HVAC field. By providing a flexible, validated environment for FDD and system optimization, the proposed emulator contributes to the development of more reliable and energy-efficient condensing boiler systems.

Methodology

The condensing heat exchanger emulator was developed in Simulink/MATLAB. The emulator is structured to simulate both nominal and faulty operating conditions, providing a robust dataset for fault detection and diagnostic algorithm development. This paper outlines the design, implementation, and validation of the emulator, with each critical component described in detail below.

Simulation

The process of modeling the condensing heat exchanger began with analyzing performance curves provided by the manufacturer. These curves depicted the heat transfer rate of the heat exchanger as a function of the inlet water temperature, corresponding to flue gas inlet temperatures of 180°C and 200°C. Polynomial equations were fitted to the provided data to represent these performance curves.

For the wet (condensing) mode, the heat transfer rate, denoted as $Q_{wet,poly}$ was expressed using a cubic polynomial equation (Eq.1). Meanwhile, for the dry (non-condensing) mode, the heat transfer rate $Q_{dry,poly}$ was defined by a linear equation (Eq.2).

$$Q_w = -0.0008 * T_{w,i}^3 + 0.005 * T_{w,i}^2 - 1.43T_{w,i} + C_1 \qquad (1)$$

$$Q_d = -0.1717 * T_{w,i} + C_2 \qquad (2)$$

Where $T_{w,i}$ indicates the inlet water temperature to the heat exchanger and C_1 and C_2 are constant used to calibrate the model based on different flue gas temperature.

The wet mode occurs when the temperature of the water passing through the heat exchanger is less than 55°C, permitting flue gas condensation, while the dry mode occurs when it is greater than 55°C. These equations served as the baseline for nominal operation but were limited to the specific flue gas inlet temperatures of 180°C and 200°C. Since the heat exchanger operates over a broader range of flue gas temperatures, we interpolated the performance coefficients C_1 and C_2 to account for intermediate temperature values. The interpolation formulas used allowed us to generalize the heat exchanger performance for any flue gas inlet temperature within the specified range.

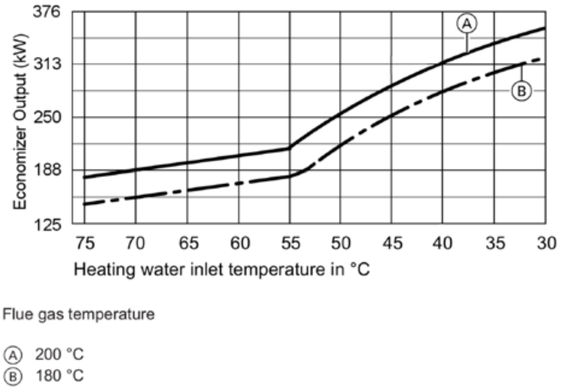

Figure 1: Performance Curve of the Condensing Heat Exchanger (Viessmann, 2008)

While these performance curves provided an idealized understanding of how the heat exchanger operates, they did not factor in variations in physical parameters, such as fouling resistances, nor did they allow for the flexibility required to simulate faulty conditions. To address this limitation and enable flexible modeling under faulty conditions, we adopted the effectiveness-NTU method.

In the dry mode, the effectiveness (ϵ_{dry}) for a counter-current heat exchanger was calculated using a specific formula. The heat capacity ratio C_r is a key parameter in this calculation, defined as the ratio of the smaller heat capacity rate to the larger one. Following (Portyanikhin & Shishov, 2023), the ϵ_{dry} is derived as illustrated in Eq. 3 and Eq. 4:

$$\epsilon_{dry} = \frac{1 - e^{-NTU_{dry}(1-C_r)}}{1 - C_r e^{-NTU_{dry}(1-C_r)}} \qquad (3)$$

$$C_r = \frac{C_{min,dry}}{C_{max,dry}} \qquad (4)$$

Where ϵ_{dry} is the Effectiveness of the heat exchanger in dry conditions, NTU_{dry} is the number of transfer unit in dry conditions, C_r is the ratio of the minimum $C_{min,dry}$ to maximum $C_{max,dry}$ heat capacity rate.

The heat transfer rate in the dry mode Q_{dry} was then calculated using the effectiveness value, the minimum heat capacity rate $C_{min,dry}$, and the temperature difference between the hot and cold fluid inlets.

$$Q_{dry} = \epsilon_{dry} * C_{min,dry} * (T_{h,i} - T_{c,i}) \qquad (5)$$

Here, $T_{h,i}$ represents the temperature of the hot fluid (flue gas), and $T_{c,i}$ is the temperature of the cold fluid (water) entering the heat exchanger. Modeling the condensing mode using the effectiveness-NTU method proved to be more complex. Limited research exists on applying this method to condensing heat exchangers, as most effectiveness formulas are designed for non-condensing, dry modes. However, a study (Portyanikhin & Shishov, 2023) provided an effectiveness formula applicable to heat exchangers with one condensing side. This formula was structurally similar to the dry mode formula but required a different approach to calculating the NTU_{wet} value due to the distinct heat transfer mechanisms in condensing and non-condensing modes.

$$\epsilon_{wet} = \frac{1 - e^{-NTU_{wet}(1-C_r)}}{1 - C_r e^{-NTU_{wet}(1-C_r)}} \tag{6}$$

$$C_r = \frac{C_{min,wet}}{C_{max,wet}} \tag{7}$$

Another significant difference between wet and dry modes lies in how the Q_{wet} is calculated. In the wet mode, heat transfer depends on the specific enthalpy difference between the flue gas at its inlet condition ($h_{h,1}$) and its dew point ($h_{s,c,1}$).

$$Q_{wet} = \epsilon_{wet} * C_{min,wet} * (h_{h,1} - h_{s,c,1}) \tag{8}$$

To calculate NTU values for each mode, the overall heat transfer coefficient-area product (UA) is divided by the minimum heat capacity rate. The UA value is derived using standard heat transfer empirical relations that incorporate convective resistances, fouling resistances, and tube material properties.

$$NTU = \frac{UA}{C_{min}} \tag{9}$$

$$UA = \left[\frac{1}{h_i * A_i} + \frac{R_{f,i}}{A_i} + \frac{\ln\left({}^{d_o}/{}_{d_i}\right)}{2\pi k_w L} + \frac{R_{f,o}}{A_o} + \frac{1}{h_o * A_o}\right]^{-1} \tag{10}$$

This equation accounts for the inner and outer heat transfer coefficients (h_i and h_o), inner and outer areas ($A_i$and A_o), fouling resistances ($R_{f,i}$ and $R_{f,o}$), and the geometric properties, including inner and outer diameters (d_i and d_o). Geometric properties such as the inside and outside surface areas, tube length (L), and material thermal conductivity (k_w) were provided by the manufacturer. These properties, combined with fouling resistances, were used to compute the UA value, which is overall heat transfer conductance and directly affects the heat transfer and effectiveness calculations. However, the critical challenge was calculating the heat transfer coefficients (h_i and h_o) for various operational conditions. Their accurate estimation was essential to ensure the reliability of the model, as they directly influence the UA value and, consequently, the heat transfer and effectiveness calculations for the heat exchanger.

Since the total heat transfer (Q) in wet and dry modes under nominal conditions was provided by the polynomial functions derived from the manufacturer's performance curves, we set these values equal to the respective Q_{dry} and Q_{wet} calculated using the effectiveness-NTU method. By doing so, we established a relationship that allowed us to estimate the heat transfer coefficients (h_i and h_o) for different operational conditions using gray-box modeling and optimization techniques.

For the dry mode, all parameters were available except for the specific heat capacity of the flue gas, which varies with temperature. Accurately determining the specific heat capacity was essential for calculating the minimum heat capacity rate ($C_{min,dry}$), NTU, and heat transfer rate. For the wet mode, calculations were more intricate, requiring not only the specific heat capacity but also the specific enthalpy of the flue gas at both its inlet condition ($h_{h,1}$) and dew point ($h_{s,c,1}$) to account for the condensation process.

Calculation of Specific Enthalpy and Heat Capacity of Flue Gas

To calculate the specific enthalpy of the flue gas at the inlet and its dew point, we relied on thermodynamic principles to ensure precision and consistency. These parameters are critical for accurately estimating the total heat transfer (Q) in the heat exchanger, which directly impacts its effectiveness and performance. While these enthalpies could theoretically be treated as variables in parameter optimization, this would increase the number of unknowns and unnecessarily complicate the model. Instead, we computed these values directly using Python's CoolProp library, which provides robust access to thermodynamic property tables.

The specific enthalpy of the flue gas depends on its composition, temperature, and pressure.

$$CH_4 + 2(1+\alpha)O_2 + 2(O_2 + 3.76N_2) \xrightarrow{yields} CO_2 + 2H_2O + (7.52 + 7.52\alpha)N_2 + 0.2\alpha O_2 \tag{11}$$

The composition was derived from the complete combustion of methane (CH_4) with excess air (α%), as described earlier. This equation accounts for varying levels of excess air, ensuring practical and complete combustion. The combustion products (CO_2, H_2O, N_2, and O_2) were calculated with their mole fractions based on the value of α.

Using the *CoolProp PropsSI* function (Bell, et al., 2014), we calculated the specific enthalpy of the flue gas by inputting its temperature, pressure, and composition. The mole fractions of the combustion products, determined by the combustion equation, were included to ensure accuracy. For the specific heat capacity of the flue gas at varying temperatures, we also used CoolProp. The temperature dependency of this property was critical for accurately modeling the system under different operational conditions. By inputting the composition derived from the combustion equation, we ensured accurate calculations for both dry and wet flue gas. Although the main heat exchanger model was implemented in *MATLAB* Simulink, CoolProp in Python was preferred for these calculations due to its straightforward implementation, open-source availability, and ease of use. While other tools, such as *REFPROP* (Lemmon, et al., 2018) could also perform these calculations, its licensing requirements and complexity made it less suitable for this application.

This approach supports the efficient calculation of thermodynamic properties needed for the heat exchanger model, ensuring compatibility with both real-time system measurements and manufacturer-provided nominal conditions. Additionally, this method enables the variation of excess air (α%) in the combustion process, allowing the simulation of faulty conditions caused by deviations in excess air levels. By adjusting α%, the impact of changes in the flue gas composition on heat transfer and overall heat exchanger performance can be analysed, further enhancing the flexibility and robustness of the model.

Parameter Optimization Using a Hybrid Gray-Box Model

To optimize the heat exchanger model parameters and incorporate operational variability, we adopted a hybrid gray-box approach that integrates physical modeling with data-driven optimization. This approach is particularly effective for systems like condensing heat exchanger, where complex physical phenomena are involved, but empirical data is also available to refine parameter estimates. The objective was to estimate the heat transfer coefficients ($h_{i,wet}, h_{o,wet}, h_{i,dry}, h_{o,dry}$) minimizing the differences between the manufacture performance curve and emulator heat transfer rates. The optimization process was performed using Bayesian optimization, a technique that systematically explores the parameter space to minimize the sum of squared errors (SSE) between the calculated heat transfer rate (Q_{eff}) and the polynomial approximation (Q_{poly}). The optimization objective can be expressed mathematically (Eq. 12).

$$f(\boldsymbol{h}) = SSE = argmin \sum_{n=1}^{N} (Q_{eff,n}(\boldsymbol{h}) - Q_{dry,n}(\boldsymbol{h}))^2 \tag{12}$$

Where:

$$\mathbf{h} = \{h_{i,wet}, h_{o,wet}, h_{i,dry}, h_{o,dry}\} \tag{13}$$

The algorithm iteratively evaluated the objective function, using probabilistic modeling to guide the search for the optimal solution. The best parameters were those that minimized the SSE, ensuring that the model closely aligned with the observed data. This hybrid gray-box approach effectively uses the strengths of both physical modeling and data-driven techniques, resulting in a robust parameter estimation process. The parameterized model developed through this process accurately reflects the heat exchanger's performance, enabling reliable simulations under a wide range of operating conditions, including both normal and faulty scenarios.

Validation

The validation process began by comparing the emulator's heat transfer predictions with polynomial equations derived from manufacturer data, which serve as a benchmark for nominal performance. The predicted heat transfer rates (Q_{eff}) were assessed against the reference values (Q_{poly}) over a range of inlet water and flue gas temperatures. A parity plot, shown in Figure 2, illustrates the agreement between the emulator performance and reference values, where an ideal 1:1 correspondence would indicate perfect accuracy.

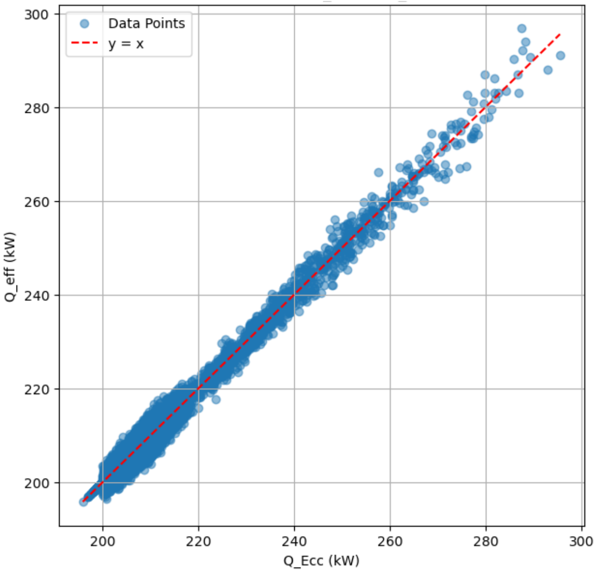

Figure 2:Parity plot comparing the estimated heat transfer rate with the reference value from manufacture's performance curve.

To quantify the model's performance, three primary error metrics were computed: the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), and the coefficient of determination (R^2). Table 1 presents these validation results for different operational conditions. The relatively low MAE and RMSE values, along with an R^2score close to unity, indicate that the emulator accurately captures the heat exchanger's thermal behavior under nominal condition. However, the wet mode exhibits higher RMSE and MAE values compared to the dry mode, which can be attributed to the complexities of condensation heat transfer. Unlike the dry mode, where heat transfer follows a well-defined linear relationship driven by convection, the wet mode involves phase change dynamics, introducing additional variability. The condensation process is influenced by factors such as latent heat release variations, the dependence of heat transfer on vapor concentration, and changes in flue gas composition due to excess air levels.

Table 1: Validation Results

Mode	RMSE	MAE	R^2
Wet	2.3019	1.7151	0.8276
Dry	1.6553	1.1317	0.9776

The validation process confirms that the emulator accurately replicates the heat exchanger's thermal behavior under various operating conditions. The close match between the model's performance and the manufacturer-provided data, as shown in Figure 2, supports the framework's accuracy. Additionally, the computed error metrics (Table Y) show low RMSE and MAE values, along with an R^2 score near 1, highlighting the model's strong predictive capability.

Since the model is built on thermodynamic principles, its reliability is well-founded. However, some differences between the emulator's predictions and the reference data may arise due to assumptions made in the combustion modeling. Specifically, the emulator assumes that the flue gas consists entirely of methane (CH_4) under complete combustion. In reality, natural gas used in actual systems often contains other hydrocarbons and trace gases, which can affect the flue gas composition and temperature **(Lan, et al., 2024)**. These variations may cause minor deviations in heat transfer predictions when compared to the manufacturer's performance data.

Despite these potential discrepancies, the results indicate that the model has been carefully calibrated across different operating conditions, effectively capturing the heat exchange's key thermal characteristics with high accuracy. This strong alignment across various scenarios demonstrates the emulator's suitability for further research on optimization, fault detection, and energy efficiency improvements in condensing boiler systems.

Application in Real-World Conditions

Beyond theoretical validation, the real strength of the emulator lies in its ability to analyze real-world data and support fault detection and optimization in operational boiler systems. To demonstrate its practical use, the emulator's predictions were compared with actual operational data from the Building Automation System (BAS) from a large hospital in ASHRAE CZ4C, as well as with manufacturer-provided performance data. Table 2 presents the results of this comparison, showing the heat recovery values predicted by the emulator, the manufacturer's expected values, and the actual measured data.

Table 2: Validation of Heat Recovery by the Condensing Heat Exchanger: Comparison of Emulator, Manufacturer, and Operational Data at Different Hot Water Return Temperature (HWRT) and a Flue Gas of 200 (C°)

HWRT (C°)	Manufacturer Data (kW)	Emulator (Kw)	Operational Data (kW)
60.13	209.08	207.27	204.79
62.84	204.43	203.87	201.53
63.84	202.43	205.15	203.64
60.04	209.25	210.13	208.43
61.5	206.73	203.23	205.63
55.45	217.12	217.81	213.57
62.22	208.5	207.68	206.58
58.29	212.24	212.81	211.88
52.77	235.85	238.83	233.12

The results show that while the emulator closely follows manufacturer specifications, discrepancies arise when compared to real operational data, which can be attributed to three main factors. First, the emulator assumes a constant flue gas temperature of 200°C due to the lack of real-time temperature data in the BAS records. However, in reality, flue gas temperatures fluctuate based on boiler operation and external conditions, directly impacting heat transfer efficiency. If these temperature variations were available and integrated into the emulator, the model's predictions would likely align more closely with operational data. Second, real-world heat losses—which the emulator does not explicitly model—reduce the actual amount of recovered heat. The emulator assumes an ideal heat exchange process, but in practice, systems lose energy to their surroundings, affecting overall efficiency. Incorporating heat loss as an additional parameter in future refinements could improve the model's predictive accuracy. Third, discrepancies may also indicate system faults such as fouling, scaling, or fluctuations in excess air levels, all of which can reduce heat exchanger efficiency. If real-world heat recovery is consistently lower than predicted, it could signal one of these underlying issues. Since the emulator is designed to simulate fault conditions, this comparison can serve as a diagnostic tool to help identify and analyze inefficiencies in the system.

Figure 3 provides further insight by illustrating the emulator's performance over time, comparing predicted and actual heat recovery values. While the emulator effectively captures overall trends, noticeable deviations highlight the impact of unmodeled heat losses and operational variability.

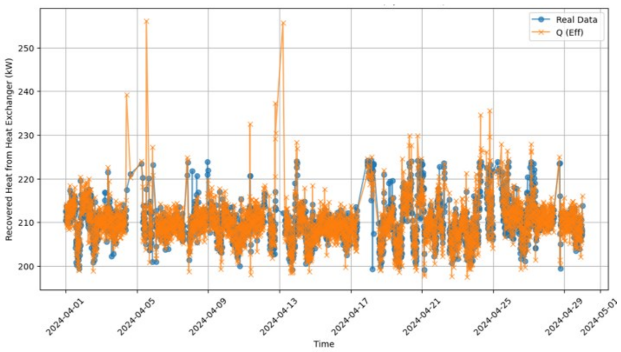

Figure 3: Time-Series Comparison of Recovered Heat (kW) from the Condensing Heat Exchanger - Emulator Predictions vs. Real Operational Data

To better understand these discrepancies, Figure 4 presents the residuals for the predicted model as a function of time. This analysis identifies periods of more significant deviations, which could be linked to transient effects, sensor inaccuracies, or system anomalies. Consistently negative residuals suggest systematic heat losses or unaccounted inefficiencies, while sudden deviations may indicate temporary faults or variations in combustion conditions.

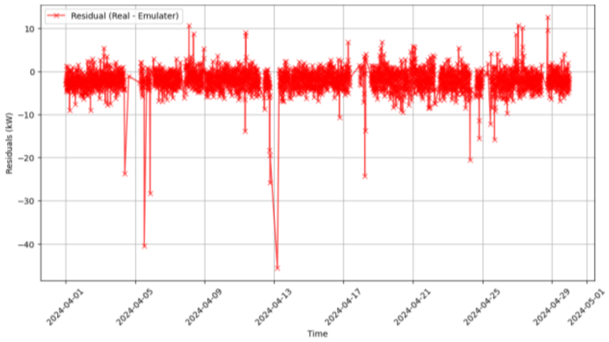

Figure 4: Residual Analysis of Heat Recovery Estimation (kW) – Discrepancies Between Emulator Estimates and Real Operational Data Over Time

From an operational perspective, these findings offer valuable insights. If the emulator predicts higher heat recovery than what's actually observed, it may indicate unexpected heat losses or declining efficiency. If the predictions match real-world data, the system is functioning normally. On the other hand, if the emulator underestimates heat recovery, it could point to sensor inaccuracies or external influences that the model does not currently account for. These results demonstrate the emulator's potential as both a predictive and diagnostic tool for condensing heat exchangers. By incorporating additional real-time data—such as flue gas temperature and heat loss estimates—the model's accuracy could be further improved, making it even more effective for fault detection, performance monitoring, and energy optimization in real-world heating systems.

Discussion and result analysis

The emulator developed in this work effectively captures both standard and condensing heat exchanger behavior under different operating conditions. By combining thermodynamic principles with data-driven methods, the model closely aligns with manufacturer performance curves across various inlet water and flue gas temperatures. The main source of deviation is from real-world complexities, such as unmodeled heat losses, fluctuations in flue gas temperatures, and variations in natural gas composition. These challenges highlight the importance of accurately defining system boundaries, combustion characteristics, and operational constraints. In the wider research field, there has been a clear gap in condensing heat exchanger emulators that explicitly model the condensation process.

Many previous studies relied solely on static manufacturer data tied to a single design condition (Shohet, et al., 2020; Antonescu & Paul-Dan, 2017). This work addresses that limitation by introducing a dynamic emulator that directly accounts for two-phase heat transfer and phase-change effects. This capability is especially crucial for systems where latent heat recovery plays a significant role in overall performance. In practical terms, the emulator assists building operators in understanding the relationship between a condensing boiler's performance and its direct impacts on energy efficiency, operating costs, and occupant comfort. By enabling early fault detection, the emulator empowers users to anticipate potential issues and schedule preventive maintenance proactively, thereby minimizing unexpected repair expenses. Furthermore, maintaining the boiler in optimal condition helps conserve energy and ensures consistent indoor comfort, balancing cost savings with occupant satisfaction.

Additionally, the emulator's ability to factor in fouling, scaling, and fluctuations in excess air makes it a versatile tool for fault detection and diagnosis (FDD). By simulating these faults in a controlled manner, researchers can generate labeled datasets—helping to address the data scarcity that often hinders FDD algorithm development.

Comparisons with real-world operational data suggest areas for improvement, such as refining transient flue gas temperature modeling or explicitly incorporating ambient heat losses. These refinements could further bridge the gap between simulated and actual performance. Overall, the results highlight the value of a physically grounded, integrated approach to modeling condensing heat exchangers. The emulator not only delivers highly accurate heat transfer predictions and fills a critical gap in existing research by offering a dynamic emulator focused on condensation. This platform is essential for proactive maintenance, energy optimization, and the continuous improvement of heating systems in large-scale applications. Moreover, the emulator's generalizable framework means it is not limited to a single boiler design; it can be extended to other similar systems and building types, highlighting its broad applicability across the HVAC domain.

Conclusions

The findings confirm that a hybrid gray-box emulator, calibrated using nominal data from manufacturer performance curves, can accurately replicate the thermal behavior of a condensing heat exchanger. By combining thermodynamic formulations with parameter optimization, this model achieves a high level of accuracy in both dry and wet operational modes. Importantly, this research fills a gap in the literature by introducing a dynamic model that explicitly captures condensing behavior—moving beyond the static, single-condition analyses commonly found in earlier studies. As a result, the emulator provides a more comprehensive and robust platform for studying the complex dynamics of latent heat recovery. Equally significant is the model's ability to simulate common faults, such as fouling, scaling, and improper excess air levels, generating synthetic datasets for training and validating fault detection and diagnosis (FDD) algorithms. This feature is particularly valuable for real-world HVAC applications, where authentic faulty data are often scarce. By creating a wide range of fault scenarios, the emulator supports the development of advanced fault detection and predictive maintenance strategies, ultimately contributing to more energy-efficient and reliable boiler systems. In practical terms, the emulator provides a safe, virtual environment to test and refine fault mitigation strategies, allowing engineers and facility managers to improve maintenance planning without risk to actual equipment. Furthermore, because the emulator is built on fundamental thermodynamic principles, it can be adapted to other condensing heating systems and building contexts, demonstrating broad utility beyond the specific system studied.

Future improvements could focus on incorporating transient flue gas profiles, modeling ambient heat losses, and refining combustion chemistry to enhance accuracy under real operating conditions. With these refinements, the emulator could serve as a powerful tool for optimizing performance, diagnosing efficiency losses, and guiding both research and operational best practices in condensing boiler technology.

Acknowledgments

This research was funded by the Natural Science and Engineering Research Council (Canada) [ALLRP-580958-22 and CREATE 510284-2018] and H. H. Angus and Associates Consulting Engineers Ltd.

References

Afram, A. & Janabi-Sharifi, F., 2015. Gray-box modeling and validation of residential HVAC system for control system design. *Applied Energy,* Volume 137, pp. 134-150.

Antonescu, N. & Paul-Dan, S., 2017. Computational model for a condensing boiler with finned tubes heat exchanger. *Energy Procedia,* pp. 555-562.

Baladi, S., Le Quang, T., Holub, O. & Endel, P., 2017. Real-time monitoring energy efficiency and performance degradation of condensing boilers. *Energy Conversion and Management ,* Volume 136, pp. 329-339.

Bell, I. H., Wronski, . J., Quoilin, S. & Lemort, V., 2014. Pure and Pseudo-pure Fluid Thermophysical Property Evaluation and the Open-Source Thermophysical Property Library CoolProp. *Industrial & Engineering Chemistry Research 53, no. 6,* pp. 2498-2508.

IPCC, 2022. *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change,* Cambridge: Intergovernmental Panel on Climate Change.

Koppaue, H., Kemmetmülle, . W. & Kugi, A., 2017. Modeling and optimal steady-state operating points of an ORC waste heat recovery system for diesel engines. *Applied Energy 206.*

Lan, Y., Wang, Z., Xu, J. & Yi, W., 2024. The Impact of Hydrogen on Flame Characteristics and Pollutant Emissions in Natural Gas Industrial Combustion Systems. *Energies,* pp. 17-19.

Lemmon, E., Bell, W., Hube, H. & McLinden, M. O., 2018. NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties-REFPROP. *Version 10.0. National Institute of Standards and Technology, Standard Reference Data Program..*

Li, Y. & O'Neill, Z., 2018. A critical review of fault modeling of HVAC systems in buildings. *Building Simulation, vol. 11,* pp. 953-975.

Pérez-Lombard, L., Ortiz, J. & Pout, C., 2008. A review on buildings energy consumption information. *Energy and buildings 40, no. 3,* pp. 394-398.

Portyanikhin, V. A. & Shishov, V. V., 2023. Adaptation of the ε-NTU Method to Calculations of Air Coolers in Conditions of Moisture Condensation or Frost Deposition. *Chemical and Petroleum Engineering,* pp. 966-979.

Satyavada, . H. & Baldi, S., 2016. A novel modelling approach for condensing boilers based on hybrid dynamical systems. *Machines.*

Shohet, R., Kandil, M. S. & McArthur, J., 2020. Fault detection for non-condensing boilers using simulated building automation system sensor data. *Advanced Engineering Informatics,* Volume 46, p. 101176.

Simić, K. et al., 2021. Dynamic modelling of gas-fired condensing boiler units for the use in residential buildings. *In HEFAT 2021: 15th International Conference on Heat Transfer, Fluid Mechanics and Thermodynamics,* pp. 688-693.

Singh, V., Mathur, J. & Bhatia, A., 2022. A comprehensive review: Fault detection, diagnostics, prognostics, and fault modeling in HVAC systems. *International Journal of Refrigeration,* Volume 144, pp. 283-295.

Viessmann, 2008. *VITOMAX 200-HW.* Allendorf: Viessmann.

Zhao, Y., Li, T., Zhang, X. & Zhang, C., 2019. Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future. *Renewable and Sustainable Energy Reviews 109* , pp. 85-101.