Operation optimization of multi-boiler district heating systems using artificial intelligence-based model predictive control: Field demonstrations

Etienne Saloux[a,*] Jason Runge[a,†] and Kun Zhang[a,b]

[a] CanmetENERGY, Natural Resources Canada, Varennes, Québec, Canada

[b] École de Technologie Supérieure, Montréal, Québec, Canada

[*] Corresponding author: e-mail etienne.saloux@NRCan-RNCan.gc.ca

[†] Jason Runge's present address: Vigilent, Oakland, California, USA

Abstract

District energy systems through fourth and fifth generations have shown great promises to help integrate renewable energy sources at large scale and to decarbonize the built environment. However, older generations using boilers represent a large proportion of systems currently in operation, and significantly contribute to greenhouse gas (GHG) emissions. Moreover, operational data has become increasingly available for these older generation systems and represents a suitable opportunity for Artificial Intelligence (AI)-based modelling and advanced controls. In this context, model predictive control (MPC) has appeared as a powerful control solution; however, field implementation remains relatively scarce. This paper aims to develop an AI-based MPC strategy for multi-boiler district heating systems. It relies on heating load forecasting machine learning models and data-driven boiler performance curves to optimize boiler thermal outputs. This strategy was implemented in two Canadian demonstration sites and showed GHG emissions reductions of 1.3% and 2.8%. Although relatively modest, statistical analysis confirmed the realization of these savings. In absolute terms, the strategy helped avoid 45 t and 77 t CO2eq emissions and save \$10,268 and \$19,975 CAD during the testing period of 2-3 months. The model and strategy developed in this work could be easily scalable to similar district heating systems.

Highlights:

- AI predictive control was developed for multi-boiler district heating systems
- The strategy relies on data-driven load forecasting and boiler performance curves
- The strategy was implemented in two Canadian demonstration sites
- Natural gas boiler consumption was reduced by 1.3 and 2.8%
- GHG emissions, energy costs were reduced by 45-77 t CO2 eq and \$ 10,268-19,975 CAD

Keywords: boiler efficiency; district heating; field demonstration; load forecasting; model predictive control; scheduling.

1. Introduction

1.1 Background and motivation

District heating and cooling (DHC) systems have shown high potential to decarbonize the built environment at the district level. With the advent of 4^{th} and 5^{th} generation systems (4GDHC [1] and 5GDHC [2], respectively), district energy systems (DES) can enable the optimal integration of a wide range of new technologies to provide heating and cooling to buildings. It includes among others: low greenhouse gas (GHG) emitter generation units (e.g. biomass boilers [3]), electrically driven generation units (e.g. electric boilers [4], heat pumps [2,5,6]), renewable energy-based systems (e.g. solar thermal collectors [7,8]), and short-term and long-term energy storage devices [9,10]. Nonetheless, although 4GDHC and 5GDHC systems are becoming increasingly present, older generation DES still represent a good proportion of systems currently in operation; specifically, steam-based systems. Performance improvements for steam-based systems seem relatively limited unless a major retrofit occurs, which can be costly. As of 2019 in Canada, 217 DES were in operation and supplied energy to a total of 3,203 buildings, mainly for heating applications, representing a total annual thermal energy delivery of 5.5 million MWh (2.2% of the total building energy usage for space heating, cooling and domestic hot water) [11]. Almost 60% of the total installed thermal capacity is associated to steam-based systems for heating purposes. Furthermore, approximately 50% of all installed systems in Canada rely solely on natural gas or oil and diesel [11].

Properly controlling DES becomes critical to reach high performance levels and to unlock the full potential of new but also well-established technologies [12,13]. In this context, advanced rule-based controls could already significantly improve system performance [14–16]. Model-based predictive control (MPC) has also emerged as a promising solution to better control DHC systems. It takes advantage of various forecasts such as weather conditions, occupancy patterns, price signals and/or grid carbon intensity to predict future system behaviour up to a few hours or days in advance, and to adjust the operation accordingly. An optimization routine allows to satisfy objectives tailored for a given application such as energy and economic savings [17], GHG emission reduction [18] or flexibility applications [19]. This procedure could be programmed to automatically control the DES; furthermore, it could also be used as a support decision system to help operators make informed decisions.

Although these advanced controls were investigated in depth in the past and have shown good promises, field demonstration studies of advanced controls in DES remain relatively scarce in the scientific literature. A few initiatives could still be mentioned. For instance, within the IEA DHC Annex TS2, the implementation of low-temperature district heating systems was investigated in depth for 15 demonstration sites in Europe and showed savings as well as increased efficiencies in every case study [20]. Furthermore, the STORM controller was developed to leverage flexibility from building thermal capacity and tested in two demonstration sites; it showed peak load reduction of 3.1% in Rottne (Sweden) and 7.5-34% in Heerlen (The Netherlands) [21]. A smart controller aiming at reducing peak loads and return temperatures in district heating networks was developed by Van Oevelen et al. [22] for the TEMPO project. They tested the controller in a peripheral branch of the existing high-temperature district heating system in Brescia (Italy) and obtained a 34% average reduction in their peak load.

The objective of this work is to develop a MPC strategy that optimizes the operation of steam-based DES using natural gas boilers, to deploy it in two demonstration sites, to report implementation results and lessons learned. The strategy targets the reduction of GHG emissions and aims to be replicable in similar systems while being adaptable for more advanced systems.

1.2 Literature review

Load forecasting represents a critical step in the development of a MPC strategy. Heating and cooling loads are usually the main drivers of DHC operation and it becomes of paramount importance to forecast them correctly. Research has extensively focused on this aspect in the past few years with the advent of artificial intelligence (AI) algorithms [23]. Saloux and Candanedo [24] compared the capabilities of linear regressions, decision trees (DT), artificial neural networks (ANN) and support vector machine (SVM) to forecast the heating demand of a solar district heating system in Canada; models were evaluated using measured weather conditions as well as weather forecasts. Powell et al. [25] tested several modelling techniques to forecast heating, cooling and electrical load of a district energy system in the U.S.A.; the non-linear autoregressive model with exogenous inputs (NARX) showed the best accuracy. Johansson et al. [26] investigated extra-trees regressors and extreme learning machine (ELM) techniques for a district heating system in Sweden. Kurek et al. [27] tested among others ANN, ridge regression and fuzzy logic to forecast the demand of a district heating network in Poland; the demand was evaluated at the city and at the substation levels. Potocnik et al. [28] explored several modelling techniques to estimate the heating demand of a large district heating system in Slovenia; Gaussian process regression (GPR) showed promising results. Leiprecht et al. [29] compared autoregressive approaches, adaptive boosting decision trees and long-short term memory neural network (LSTM) to predict heating demand of a district energy system in Germany. Runge and Saloux [23] explored different machine learning and deep learning algorithms and compared their capabilities as prediction and forecasting heating demand models for a Canadian district energy system; LSTM and XGBoost models gave encouraging results in terms of model accuracy and stability. Chung et al. [30] assessed the performance of a convolutional neural network (CNN) coupled with a LSTM for load prediction for a district heating system in Korea; this model outperformed XGBoost, LSTM and random forest techniques. More initiatives on load forecasting are reported in [23]. The vast number of existing modelling techniques and the continuous emergence of new algorithms make the forecasting model selection particularly challenging. Some models could be more suitable for occupancy behavior or for large networks with a high thermal inertia. Some others could perform better with different load types (space heating, space cooling, domestic hot water), or at different scales (district, neighborhood or substation levels). In this context, modelling competitions play a crucial role and enable model comparison under the same conditions and datasets. The M4 competition [31] has focused on 61 forecasting methods, which were tested on 100,000 time series data from various domains (e.g. industry, finance, demographic). Similarly, the ASHRAE Great Energy Predictor III Competition [32] was a Machine Learning competition targeting long-term prediction of building energy use for measurement and verification. The datasets consisted of over 20 million points of training data from 2,380 energy meters collected for 1,448 buildings from 16 sources.

Thermal load forecasting can bring invaluable information to better control central heating and cooling plants that supply energy to the district. Gunay et al. [33] explored the sequencing of natural gas boilers and chillers based on a day ahead peak load forecasting; they showed that an optimal sequencing could reduce heating energy use by 4% and cooling energy use by 25%. Labidi et al. [34] investigated a new strategy based on power demand forecasting to operate a district heating system using a wood boiler, hot water tanks and gas boilers. This strategy was able to reduce gas consumption and GHG emissions while being cost effective. Saletti et al. [35] evaluated the performance of a predictive controller for a small-scale district heating system equipped with a boiler; it yielded a reduction in fuel consumption of up to 7%. Aoun et al. [36] developed a predictive control strategy to exploit the thermal inertia of buildings connected to a district heating system for flexibility purposes; the strategy was found cost effective while preserving thermal comfort. Saloux and Candanedo [18] optimized the usage of solar collectors along with short-term and long-term thermal energy storage devices to decrease natural gas boiler consumption; results showed a 32% decrease in GHG emissions.

While such previous publications have been insightful, none of them actually performed field demonstrations. Thus, the contribution of this work is the field demonstration of an AI-based MPC for district heating systems. Although it has not been strictly explored for DES, the optimization of multi-energy systems has also raised high interest for commercial buildings, particularly for multi-chiller sequencing. Liao and Huang [37] (7% energy cost reduction), Thangavelu and Khambadkone [38] (20-40% energy savings), and Saloux and Zhang [39] (19% electric power decrease) tackled this topic, among others.

1.3 Paper organization

The paper is organized as follows: a generic representation of the steam-based DES is given in Section 2 along with more details about the two demonstration sites. Section 3 deals with the development of the predictive control strategy, whereas Section 4 tackles the benchmarking model for savings assessment and the implementation results. Section 5 discusses obtained results, potential improvements for the proposed approach, and lessons learned.

2. Field demonstration sites

2.1 Steam-based district energy system

The steam-based DES under study consists of multiple natural gas boilers, which generate steam at a predefined pressure. Steam is sent to the network to provide heating and domestic hot water to a district composed of buildings of various types. Boilers are installed in parallel and can be run individually or in combination with other boilers. Figure 1 shows a schematic of the system under consideration.

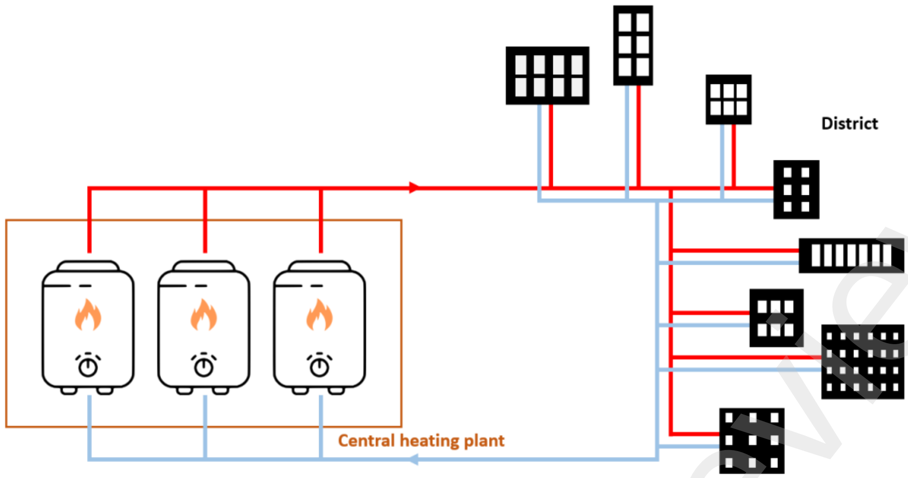


Figure 1. Schematic of the steam-based district energy system equipped with natural gas boilers. Note that the schematic includes only the relevant components for illustration purposes and omits common components (pumps, substations, etc.).

2.2 Site description and operational data

The first site is located in the province of Ontario (Canada). The central heating plant is equipped with 3 non-condensing natural gas boilers, whose capacity ranges between 42,000 and 84,000 MJ/h. The plant supplies space heating and domestic hot water to the district. Historical datasets contain weather conditions, system operating conditions, steam production and gas consumption at 1-min intervals. A unique control system records and stores the data; the memory of trend logs varies from one variable to another; therefore, each variable has its own historical data period. Weather conditions were also extracted from other sources to retrieve unmeasured data (e.g. solar radiation) and infer missing data in historical datasets, when required.

The second site is located in the province of Québec in Canada. The DES supplies space heating and domestic hot water to a total building surface area of 271,000 m^2 through an 11-km network. The central heating plant has 3 non-condensing natural gas boilers of relatively similar capacity (62,000 – 67,000 MBH). Historical data is recorded at 15-min intervals by two different control systems. The first control system deals with weather conditions, system operating conditions and steam production, and shows historical data for several years. The second one records gas consumption but shows historical data for 8 months only. As for the first site, weather conditions were also extracted from other sources.

For both sites, raw operational datasets were pre-processed to make them suitable for control applications. Datasets were retimed and synchronized at regular intervals. Variables were renamed and their units were identified and converted, if required. Time index variables were calculated; it includes among others weekday/weekend/holiday, hour of the day, day of the year, season. Outliers and anomalies were detected using statistical means and expert knowledge, and were removed. The data was then analysed and enabled to identify miscalibrated sensors, discover trends (e.g. daily and seasonal profiles), and evaluate system performance. In addition to operational data, datasets for daily steam production and gas consumption were available and were used during the implementation phase to evaluate the performance of the MPC strategy.

Table 1 summarizes the main characteristics for each demonstration site. Typical building schedules were extracted from the temporal analysis of the total district heating demand. Similarly, the estimated impact of COVID-19 pandemic on energy usage was evaluated by comparing the total district heating demand before and after March 2020.

Table 1. Main characteristics of the two demonstration sites under consideration.

Characteristics	Site #1 (Ontario)	Site #2 (Québec)
Estimated peak load from 15-min measurements (99.7% percentile)	16 MW	19 MW
Total building floor area	Unknown	271,000 m^2
Number of gas boilers	3	3
Typical building schedules	Winter: Mon-Fri, 4am-10pm Summer: Mon-Fri, 4am-10pm	Winter: Mon-Fri, 6am-6pm Summer: low variations
Estimated impact of COVID-19 pandemic on energy usage	Low	Low
Min, average, max outdoor air temperature	- 26°C, 6°C, 36°C	- 37°C, 5°C, 34°C
Historical data (steam, gas) availability	Steam: Mar 2019 – Mar 2021 Gas: Dec 2019 – Mar 2021	Steam: Apr 2015 – Apr 2021 Gas: Jun 2020 – Jan 2021
Historical data time-step	1 min	15 min

Figure 2 depicts typical heating demand profiles for site #1. Figure 2a shows the 15-min steam production as a function of the hour of the day during workdays (Mon-Fri) for the winter season (Nov 1, 2020 – Apr 1, 2021); the average value is shown in red. Figure 2b gives the average load according to the day of the week. From these figures, the heating demand shows a peak in the morning around 4:00 and drops in the evening at 20:00, indicating typical building mechanical system schedules. The load remains relatively high in the morning while it smoothly decreases in the afternoon. Similar behaviours were observed for both workdays and weekends. During summer (not shown in the figure), the load is lower and varies less but is non-zero due to domestic hot water needs. Figures 2c and 2d illustrate the daily heating demand as a function of outdoor air temperature for workdays (Mon-Fri) and weekends (Sat-Sun), respectively; the effect of COVID-19 pandemic seems relatively low, either during workdays or weekends.

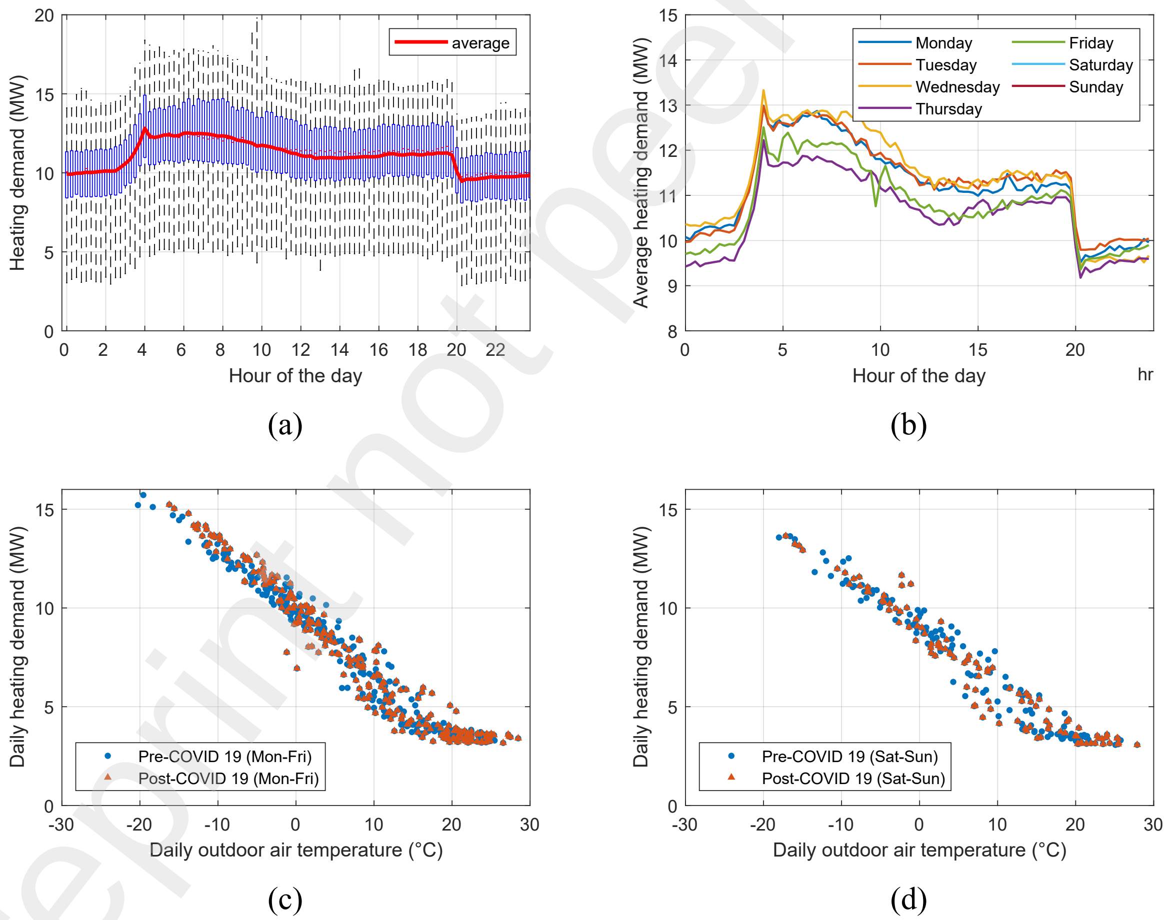


Figure 2. Typical heating load profile for site #1: (a) daily profile over the winter season 2020-21 using box plots (outliers excluded), (b) daily profiles of the average heating demand for each day of the week, (c) daily heating demand as a function of outdoor air temperature for workdays

(Mon-Fri), (d) daily heating demand as a function of outdoor air temperature for weekends (Sat-Sun).

In terms of controls, for both sites, the operators decide in the morning which boilers (usually one or two) to run for the rest of the day. This operation is adjusted throughout the day according to specific needs and operational issues (e.g. boiler vibration and noise, maintenance and repairs, chemicals control, power failure). In site #1, if two boilers are in operation, one boiler provides a base load and another one fulfills the peak; in site #2, the demand is equally distributed among the boilers. By Canadian law and regulations, operators must turn on and off the boilers manually. Therefore, boiler sequencing is not only based on outdoor conditions and heating demand, but also on operator's preferences. This also restricts the strategy and prevents the fully automated control of boiler systems. Based on these restrictions, the objective of the predictive control strategy is to help operators make informed decisions about which boilers to run and at which load.

3. Development of the model predictive control strategy

The overall approach for the field implementation of the predictive control strategy is summarized in Figure 3. The MPC strategy is described step-by-step in the following subsections and aims to maximize boiler efficiency at part load ratio to minimize gas consumption and thus, GHG emissions.

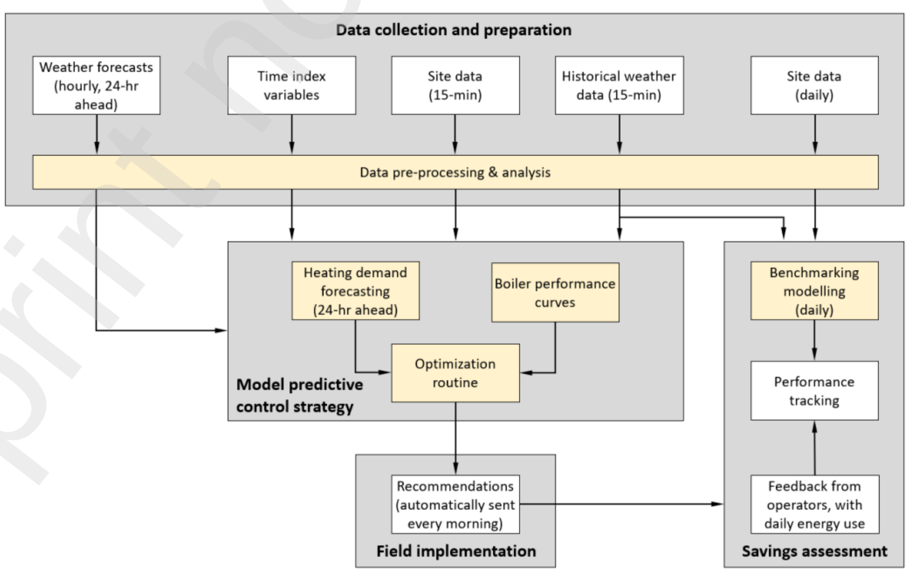


Figure 3. Schematic of the predictive control strategy approach: data collection and preparation, development of the model predictive control strategy, field implementation and savings assessment.

3.1 Control-oriented models

The control-oriented model of the system is the cornerstone of the predictive control strategy. In this study, two models were developed: one machine learning model forecasting the heating demand 24 hours in advance and one model targeting boiler modelling through performance curves at part-load ratio. These models are trained for each site, using their own datasets.

3.1.1 Heating demand forecasting model

Heating demand forecasting is crucial in the development of the predictive control strategy. The district heating load is the aggregated load of each individual building. Depending on the size of the district, modelling each building individually might be cumbersome. However, since the aggregated load is measured, it makes black-box modelling techniques suitable for such an application. Different machine learning techniques (feedforward ANN, CNN, LSTM, light gradient boosting model LGBM, random forests, SVM and XGBoost) along with different sets of inputs were investigated to evaluate the appropriate modelling level to forecast the district heating demand. These models were tested within two approaches: a forecasting approach that uses only past and current information, and a prediction approach that also leverages weather forecasts in addition to past and current information. More details can be found in Runge and Saloux [23].

Based on the findings from reference [23], XGBoost prediction models were selected in this work and used for both sites due to their high accuracy, good stability and low computational time for training. The district heating demand was estimated at hourly intervals for the next 24 hours. Table 2 provides heating demand forecasting model features and accuracy for each site. R^2 is the coefficient of determination, NMBE is the normalized mean bias error, RMSE is the root mean square error and CV-RMSE is the coefficient of variation of the RMSE. It is also worth mentioning that solar radiation was not used as an input since it did not significantly increase load forecasting accuracy. Furthermore, solar radiation is difficult to forecast and comes with high uncertainty. As shown in the table, the model is relatively fast to compute (a few seconds) while it also meets

ASHRAE Guideline 14 requirements for model calibration even when the model is used along with weather forecasts [40]. Figure 4 shows heating demand measurements against model results when inputs use actual weather conditions and weather forecasts 24 hours ahead. Results are given for two weeks in January and March, for both sites, and show good consistency between measurements and model estimations.

Table 2. Heating demand forecasting model features and accuracy.

Variable	Site #1 (Ontario)	Site #2 (Québec)
Model type	XGBoost	
Model inputs	Outdoor dry air temperature Heating cosine function Cosine hour of the day Hour of the day Work hour Weekend	
Model hyperparameters	Number of estimators: 350 Max depth of the trees: 9 Subsample size: 1 Learning rate: 0.3	
Training period	Mar 30, 2019 – Dec 31, 2020	Apr 30, 2016 – Jun 22, 2020
Testing period	Jan 1, 2021 – Mar 24, 2021	Jun 22, 2020 – Apr 30, 2021
Model training time	3.5 s	5.5 s
Model accuracy during training	R^2 = 0.999 NMBE = - 7.4e-05 % RMSE = 0.08 MW CV-RMSE = 1.2 %	R^2 = 0.998 NMBE = 7.2e-05 % RMSE = 0.2 MW CV-RMSE = 2.0 %
Model accuracy during testing *Inputs: measured weather*	R^2 = 0.800 NMBE = - 3.5 % RMSE = 0.9 MW CV-RMSE = 8.0 %	R^2 = 0.925 NMBE = 0.6 % RMSE = 0.9 MW CV-RMSE = 10.7 %
Model accuracy during testing *Inputs: weather forecasts*	R^2 = 0.760 NMBE = - 2.1 % RMSE = 1.0 MW CV-RMSE = 8.7 %	R^2 = 0.911 NMBE = 1.1 % RMSE = 1.0 MW CV-RMSE = 11.7 %

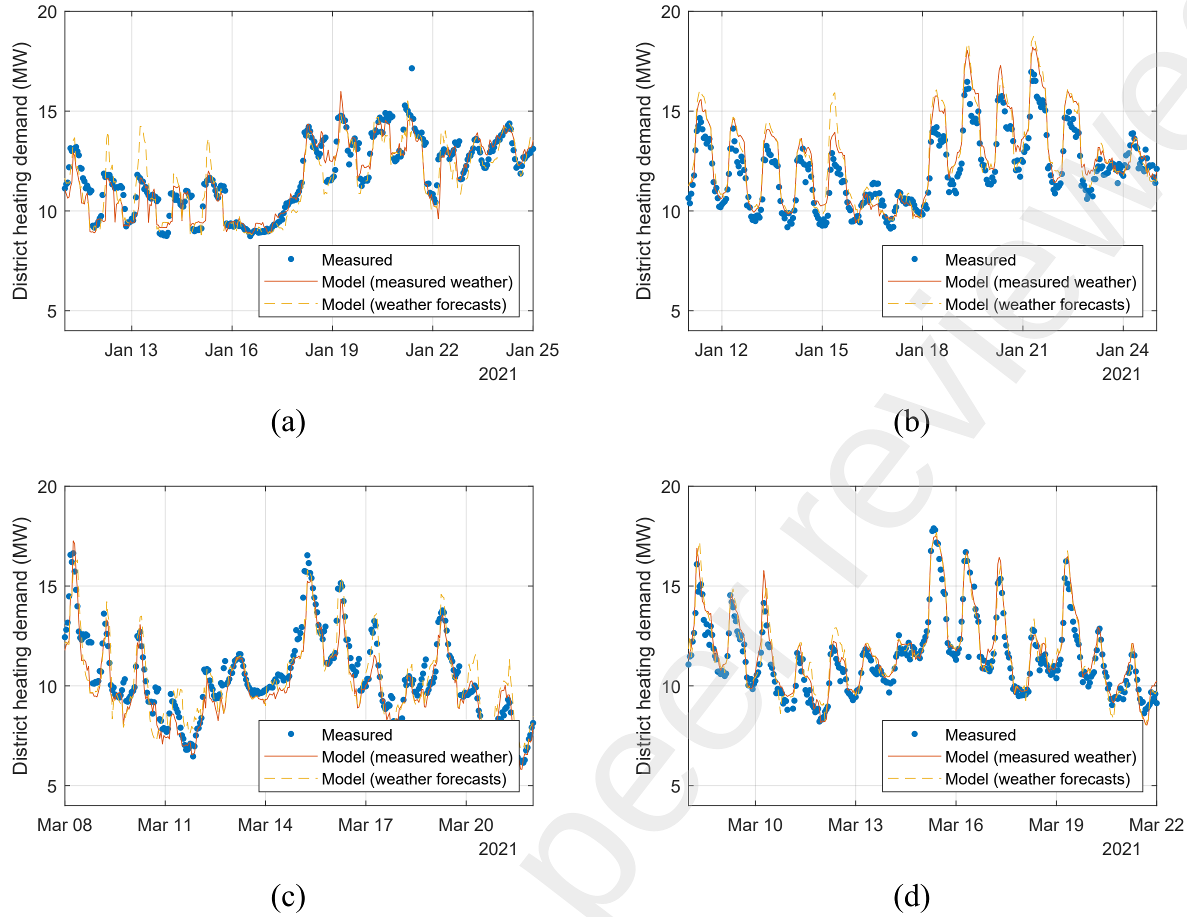


Figure 4. District heating demand measurements against models using measured weather and weather forecasts: (a) for two weeks in January 2021 for site #1, (b) for two weeks in January 2021 for site #2, (c) for two weeks in March 2021 for site #1, (b) for two weeks in March 2021 for site #2.

3.1.2 Boiler performance curves

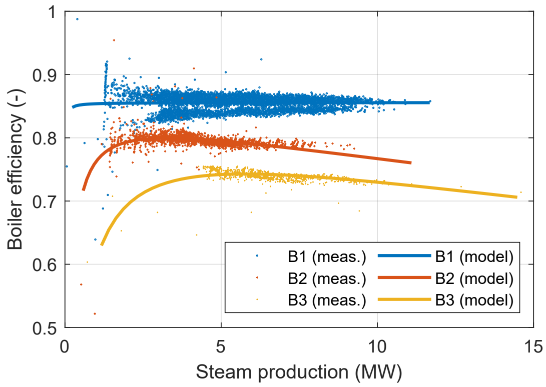
Once the demand is forecasted, this information can be used to optimize the operation of the central heating plant. In this study, the approach consists in developing part-load ratio performance curves for each boiler using operational data. These data-driven performance curves allow to capture the general boiler efficiency and to evaluate which boiler performs the best at a given load. Due to the lack of exhaustive datasets, for simplification, and generalization purposes; the proposed approach excludes among others make-up water, water mass flow rate and inlet temperature as well as chemicals considerations.

Different techniques can be used for modelling boilers. Labidi et al. [34] considered boilers as production units with varying capacity (minimum and maximum capacities). Dainese et al. [3] and Saletti et al. [35] calculated the boiler efficiency with a linear regression model using the nominal efficiency and the actual load or the load factor. Higher order polynomial fits could also be tested to draw boiler performance curves. However, the main drawback of these polynomial fits is the physical sense of results at low part-load ratio or during extrapolation. To overcome this issue, Gunay et al. [33] suggested a 3-parameter exponential model to calculate the mean daily efficiency as a function of the mean daily part-load ratio. In this work, to ensure robustness to extrapolation, the modelling was inspired from the National Energy Code of Canada for Buildings (NECB) [41] where the fuel consumption at part-load ratio is calculated using the fuel consumption at design conditions and a quadratic fit of the part-load ratio. Pre-defined coefficients are recommended for condensing and non-condensing boilers. This equation was modified as follows to leverage available operational data and explicitly express boiler efficiency:

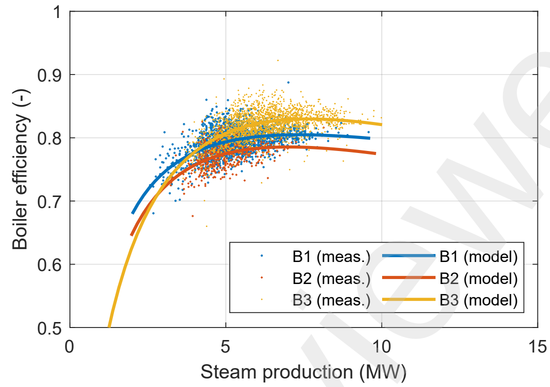
$$\eta_{\text{boil}} = \frac{a}{\frac{b}{PLR} + c + d \times PLR} \tag{1}$$

where a, b, c and d are parameters to calibrate with operational data. η_{boil} is boiler efficiency and PLR is the boiler part-load ratio (i.e. thermal load divided by nominal power). The boiler efficiency is estimated at hourly intervals. To remove transient operation [39], the efficiency is calculated only if the boiler has been in operation for 2 hours and remains in operation for a further 2 hours.

Figure 5 shows data-driven boiler performance curves obtained for both sites as a function of boiler steam production, and Table 3 gives model accuracy for each boiler and each site. Each site shows relatively different boiler performance curves: site #1 displays three distinct levels of performance (i.e. 85% for boiler #1, 80% for boiler #2, 75% for boiler #3) whereas the performance is relatively similar for site #2 (i.e. around 80% between 5 and 10 MW). Models show good consistency with measurements in terms of NMBE and CV-RMSE, but relatively poor R^2 values. This is mainly due to the fact that the boilers were already operated in a range around their maximum efficiency, and thus showed low efficiency variations, whereas R^2 is not suitable for nonlinear models.



(a)



(b)

Figure 5. Data-driven boiler performance curves for: (a) site #1 and (b) site #2.

Table 3. Boiler model accuracy.

Site	Boiler	R^2	NMBE	RMSE	CV-RMSE
#1 (Ontario)	#1	8.0e-04	- 1.9e-05 %	0.012	1.4 %
	#2	0.223	6.3e-05 %	0.009	1.1 %
	#3	0.484	- 5.4e-04 %	0.007	1.0 %
#2 (Québec)	#1	0.299	5.9e-05 %	0.019	2.3 %
	#2	0.213	5.8e-04 %	0.019	2.4 %
	#3	0.364	- 2.3e-04 %	0.019	2.2 %

3.2 Weather forecasts

The MPC strategy is based on weather forecasts obtained from the Canadian Meteorological Centre. Specifically, numerical weather prediction files (GRIB format) were used and provide tabulated values for multiple variables (temperature, cloud cover, wind speed, humidity, precipitations, etc.) at hourly intervals up to 48 hours ahead; a new forecast is generated every six hours. The software tool CanMETEO [42] was used to ease data retrieval and to calculate solar radiation from cloud cover forecasts. In order to test the effectiveness of the developed models with forecasts (see Section 3.1), a database of historical weather forecasts was created. Outdoor air temperature and solar radiation were the two main variables considered for modelling purposes. The accuracy of weather forecasts up to 24 to 48 hours ahead were investigated for site #2 [23] and for two other locations in Canada [43]. Overall, the outdoor air temperature is relatively well forecasted, with a standard deviation of 2.2-2.3°C. In contrast, solar radiation is more difficult to forecast due to its intermittent nature (standard deviation of 80-170 W/m^2), especially during partly cloudy days.

3.3 Optimization routine

The control-oriented model was used along with weather forecasts to optimize the multi-boiler district heating system operation. The control variables aim to evaluate the contribution of each boiler to fulfill the district heating demand at a given hour of the day; in other words, which boiler to run and at which power output for the next 24 hours:

$$\dot{Q}_{steam,boil} = \begin{bmatrix} \dot{Q}_{steam,B1,8\mathrm{am}} \\ \dot{Q}_{steam,B1,9\mathrm{am}} \\ \vdots \\ \dot{Q}_{steam,B1,7\mathrm{am}+1\mathrm{day}} \\ \dot{Q}_{steam,B2,8\mathrm{am}} \\ \vdots \\ \dot{Q}_{steam,B3,7\mathrm{am}+1\mathrm{day}} \end{bmatrix} \tag{2}$$

Since both sites show different boiler operational practices, the optimization routine needs to be adjusted accordingly.

- Site #1: one boiler operates for the base load (i.e. constant load over the next 24 hours), the second boiler operates for the peak load ("*base-peak*" *method*).
- Site #2: the load is equally distributed among the boilers; "*base-peak*" *method* can be manually applied by the operators if needed.

The objective function in the optimization routine is to decrease gas consumption of all boilers to reduce GHG emissions. Overall, the optimization problem can be summarized as follows:

$$\min_{\dot{Q}_{\text{steam,boil}}} J, \quad \text{where } J = \sum_{k=1}^{n_{\text{boil}}} \int_{t}^{t+\Lambda} \dot{Q}_{\text{gas,boil},k}\, dt = \sum_{k=1}^{n_{\text{boil}}} \int_{t}^{t+\Lambda} \frac{\dot{Q}_{\text{steam,boil},k}}{\eta_{\text{boil},k}(\dot{Q}_{\text{steam,boil},k})} dt$$

$$s.t.$$

$$\dot{Q}_{\text{district}} = \sum_{k=1}^{n_{\text{boil}}} \dot{Q}_{\text{steam,boil},k} \tag{3}$$

$$\dot{Q}_{\text{steam,boil},k}^{\min} \leq \dot{Q}_{\text{steam,boil},k} \leq \dot{Q}_{\text{steam,boil},k}^{\max}$$

$$n_{\text{boil}} \in [1, 2, 3]$$

$$mode_{\text{boil},k} \in [\text{base}, \text{peak}, \text{even}]$$

where $\dot{Q}_{\text{gas,boil},k}$, $\dot{Q}_{\text{steam,boil},k}$ and $\eta_{\text{boil},k}$ are gas consumption, steam production and efficiency of the k^{th} boiler in operation. n_{boil} is the number of boilers in operation (whether one, two or three). $\dot{Q}_{\text{steam,boil},k}^{\min}$ and $\dot{Q}_{\text{steam,boil},k}^{\max}$ are minimum and maximum capacity at which the k^{th} boiler is allowed to operate. $mode_{\text{boil},k}$ is the operation mode of the k^{th} boiler: base load, peak load or even distribution. $\dot{Q}_{\text{district}}$ is the district heating demand. t is the time at which the recommendation starts (8am) and Λ is the prediction horizon (24 hours).

The optimization problem was solved with Python using a parametric analysis. All possible boiler combinations were explored. The base load was incrementally adjusted for each boiler. Scenarios that do not respect constraints were removed and the results for the different feasible scenarios were ranked based on daily energy efficiency. This approach was selected rather than a formal optimization because we (and the building operators) were interested in the general optimum but also in other specific scenarios for reporting purposes (see Section 3.4).

3.4 MPC algorithm and recommendations

Since operators must turn on boilers manually and decide which boilers to run once a day (Section 2.2), the MPC strategy output is a PDF report containing operational recommendations that aims to help operators make an informed decisions about which boilers to operate and at which load. This report is automatically generated and sent around 7am before operators' daily meeting. The report contains recommendations from 8am to 7am the next day. To this purpose, the latest weather forecasts were retrieved for the next 48 hours using CanMETEO [42] and adjusted to match the recommendation period (8am – 7am the next day). Site operational data and weather data were retrieved to train a heating demand machine learning model and data-driven performance curves; this data was static during the implementation phase, but it allows adaptive learning in the future to make sure the MPC strategy relies on the most recent data. The optimization is then performed within Python and results are displayed in a PDF report, also generated with Python, before being sent to the central heating plant operators. The whole process takes around 6 minutes to run. Every day, during the implementation phase, operators filled out a feedback form including daily gas consumption and steam production, which was used to track if the recommendation was implemented and to evaluate the energy performance, while being aware of any technical issues that may have occurred.

The PDF report was adjusted for each site depending on the operators' needs and preferences. Overall, it displayed the estimated forecast information (outdoor air temperature, district heating demand) and the recommended boiler thermal power for the next 24 hours, along with total energy usage, system efficiency and expected savings (gas consumption, energy costs, GHG emissions). Additional scenarios were also provided in case one boiler could not be operated during a given day due to technical reasons. The last page was used for collecting the operators' feedback.

4. Performance evaluation following field implementation

4.1 Benchmarking modelling

Assessing savings requires a model to estimate how the district heating system would have performed if it was operated under normal conditions (i.e. without the MPC strategy). Different approaches have been investigated in the past and signature models have been developed to calculate energy usage or costs, normalized with weather conditions (e.g. heating degree days, average outdoor air temperature). To this purpose, data at sub-hourly to hourly [44], daily [45–48] and monthly [49] intervals can be used.

In this paper, daily values of gas consumption and steam production were used since they are readily available in operators' monthly reports or utility bills. The MPC strategy intends to improve the system energy efficiency by reducing gas consumption for a given district heating demand.

Therefore, to properly evaluate the performance, daily natural gas consumption of all boilers was plotted against daily district heating demand (i.e. total steam production of all boilers), rather than outdoor air temperature. A linear regression was used as a baseline model. Table 4 shows model accuracy for both sites; high accuracy was achieved in both cases.

Table 4. Benchmarking model accuracy.

Site	R^2	NMBE	RMSE	CV-RMSE
#1 (Ontario)	0.993	- 3.7e-14 %	34 GJ	4.4 %
#2 (Québec)	0.974	- 2.1e-14 %	58 GJ	5.3 %

4.2 Site #1: implementation results

The implementation phase took place from February 2nd to April 27th, 2023. After this date, the most efficient boiler was able to fulfill the heating demand for the next 24 hours and could be kept in constant operation. Therefore, sending MPC reports would not be required. Figure 6 shows the implementation results, which were divided into three categories depending on the operators' feedback. The recommendation was whether: 1) followed, 2) partially followed (e.g. for most of the day except for a few hours, with a base power slightly different from the recommendation) or 3) not followed (e.g. boilers were not operated as recommended due to technical issues with one or several boilers, recommendations were not followed during weekends). Figure 6 also displays the baseline model along with the historical data that was used for calibrating the model (September 2021 – November 2022). This figure shows that the gas consumption was consistently lower than the baseline when the recommendations were followed or partially followed. The higher gas consumption observed during the implementation phase occurred when the recommendation was not followed due to maintenance issue (boiler chimney repairs, boiler maintenance, high tension electrical works).

Table 5 gives the performance over the whole implementation period when the recommendations were either followed or partially followed (48 days). To this purpose, we assumed for natural gas, lower heating value of 37.6 MJ/m^3, an energy cost of 0.50 CAD/m^3 and a GHG emissions factor of 1,921 g CO2eq/m^3. The table shows that the system efficiency was increased by 2.28%, which yields to natural gas consumption reduction of 2.8 %. Although modest, this reduction helped save

$ 19,975 CAD and 77 t CO2eq over a period of 48 days; more savings could thus be expected when the MPC strategy is applied to the whole winter season. It represents on average daily reductions of $ 416 CAD and 1.6 t CO2eq. To confirm that savings were actually obtained, a statistical analysis of these results was conducted. Details can be found in Appendix A.

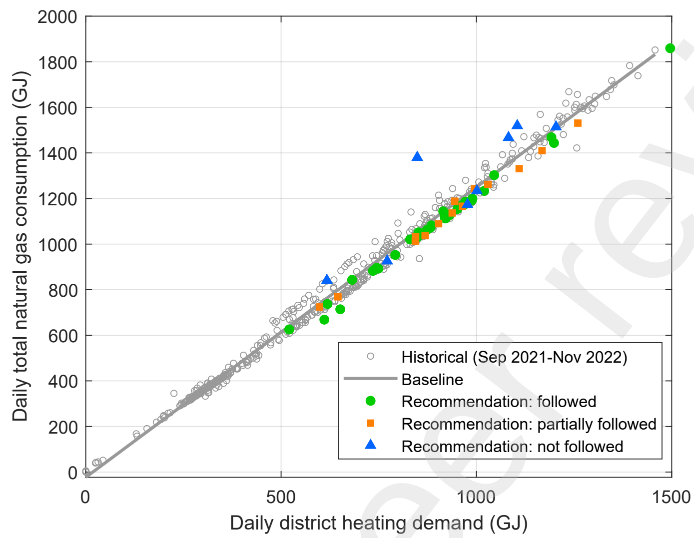


Figure 6. MPC implementation results for site #1: daily natural gas consumption against daily total steam production.

Table 5. MPC implementation results for site #1: performance over the implementation period (Feb 2 – Apr 27, 2023) when the recommendation was partially or fully followed (48 days).

Performance metric	Reference	MPC	Difference
Total district heating demand (GJ)	43,529	43,529	0
Total gas boiler consumption (GJ)	54,305	52,803	- 1,502 (- 2.8 %)
Average system efficiency (%)	80.16	82.44	2.28
Total energy cost (CAD)	722,140	702,165	- 19,975 (- 2.8 %)
Total GHG emissions (t CO2eq)	2,774	2,698	- 77 (- 2.8 %)

4.3 Site #2: implementation results

The implementation phase took place from February 15th to May 28th, 2023. After this date, a single boiler could be kept in operation to fulfill the heating demand. Figure 7 shows: the implementation results, the baseline model, and historical data used for calibrating the model (January 2022 – February 2023). Implementation results were divided into two categories depending on operators' feedback. The recommendation was whether: 1) followed, or 2) not followed (e.g. boilers were not operated as recommended due to technical issues with one or several boilers). Unlike the first site, Figure 7 shows that the gas consumption was in average lower than the baseline when the recommendation was followed, but some occurrences are higher than the baseline. Overall, the model is also a bit less accurate than for site #1.

Table 6 gives the performance over the whole implementation period when the recommendation was followed (67 days). We assumed for natural gas, lower heating value of 38.8 MJ/m^3, an energy cost of 0.44 CAD/m^3 and a GHG emissions factor of 1,921 g CO2eq/m^3. The table shows that the system efficiency was increased by 1.07%, which corresponds to a reduction of 1.3 % in natural gas consumption. This decrease is again modest, but it still helped save \$ 10,268 CAD and 45 t CO2eq for a period of 48 days. On a daily basis, it represents average daily reductions of \$ 153 CAD and 0.7 t CO2eq. As for site #1, more savings are expected if the MPC strategy is applied to the whole winter season. A statistical analysis of these results is given in Appendix A.

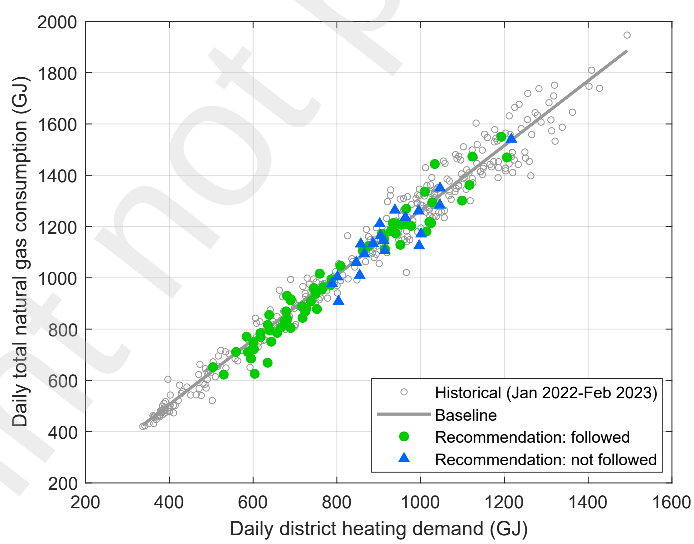


Figure 7. Results from MPC implementation for site #2: daily natural gas consumption against daily total steam production.

Table 6. MPC implementation results for site #2: performance over the implementation period (Feb 15 – May 28, 2023) when the recommendation was followed (67 days).

Performance metric	Reference	MPC	Difference
Total district heating demand (GJ)	53,378	53,378	0
Total gas boiler consumption (GJ)	67,463	66,560	- 903 (- 1.3 %)
Average system efficiency (%)	79.12	80.20	1.07
Total energy cost (CAD)	766,680	756,412	- 10,268 (- 1.3 %)
Total GHG emissions (t CO2 eq)	3,338	3,293	- 45 (- 1.3 %)

5. Discussion on performance results, potential improvements and lessons learned

5.1 MPC strategy performance comparison between sites

The MPC strategy allowed to improve system performance and decrease natural gas consumption, although the reduction was modest. Savings of 2.8% were obtained for site #1, for which the gas consumption under the MPC strategy was consistently lower than the baseline. For site #2, savings of 1.3% were obtained and gas consumption was on average, but not always, lower than the baseline.

An important aspect to highlight is that the effectiveness of the control strategy mainly depends on the original control strategy (baseline), the performance of each boiler, and the available boilers to operate. If the original control strategy shows several inefficiencies, the MPC strategy may provide more savings; conversely, if the central heating plant is already well operated, lower savings will be obtained. It is worth mentioning that the operators are experienced technicians and managers, who are proud in efficiently operating and maintaining their facility for years. Our approach is at least a transcript of their intangible experience and is more easily transferrable to a new technician.

Moreover, the performance of each boiler may limit the effectiveness of the approach. In site #1, the three boilers have different levels of efficiency, and the strategy aims to take advantage of the most efficient boiler and run it at the highest possible power. Therefore, this strategy can be quite effective if the most efficient boiler was not used as much. In site #2, the efficiency is relatively similar for all the three boilers and the optimization is thus limited, although it mainly avoids operating boilers at low part load ratio. In this case, the equal load distribution appears as the best approach. The accuracy of the heating demand forecasting model may have less impact on the overall savings in this case study, especially when the load is equally distributed, but it remains essential to build operators' confidence with the proposed recommendations.

5.2 Potential improvements

The proposed approach is a first step towards the development of predictive control strategies for district heating and cooling at a large scale. The models were successfully applied to two different demonstration sites, which might indicate a good potential for replication in similar sites. Although some aspects could still be improved such as supply steam pressure optimization or adaptive learning (which was not applied in this study due to network security restrictions), the approach is relatively simple, generic, and easily scalable to larger and more complicated systems, if required.

Nevertheless, steam-based district heating systems are smoothly transitioning to next generation systems and may undergo major retrofits: conversion of a portion of the network from steam to hot water, buildings and substations retrofits, integration of more complex systems to the central heating plant (electric boiler, heat pumps and renewable energy systems, thermal energy storage devices). Such modifications in the district heating system could be integrated into the overall approach: retraining the machine learning forecasting model could account for building retrofits; additional models could be developed to include renewable energy generation and thermal energy storage devices [50,51]. Nonetheless, the proposed methodology and current models could be the foundation of predictive control strategies for more advanced technologies.

5.4 Challenges encountered and lessons learned

One of the critical challenges of any data-driven approach is data collection and preparation, especially with older generation district heating systems. Operational data is not necessarily properly stored at the right frequency (either too frequent or too sporadic) or stored for a long period of time. Sensors might have been installed in a wrong location or might be miscalibrated or show questionable readings. Recorded variables are not necessarily correctly labelled, they might include internal calculations involving specific assumptions or units might be unknown. Retrieving the data was not a straightforward process; furthermore, access to real-time data was limited. However, this work has highlighted the importance of a proper data management system and showcased the added value of operational data for improving the performance, thus pushing towards an increased utilization of data in district heating and cooling. A data-driven approach gives an accurate picture of the current operation and could assist operators in their daily routine: it could confirm or disconfirm beliefs while a brief data analysis could already raise awareness to certain critical issues (peak demand, energy usage, schedules) and help identify technical issues for remediation.

Another key challenge of this field demonstration was to translate the optimization results from the MPC strategy into concrete actions for the operators, and how to incentivise them to change their daily habits and adopt a new tool. Automatic control of the district heating system was not allowed and engaging operators at the very beginning of the project was crucial. It was also important to keep them in the loop during the whole process, since it has generated fruitful discussions and facilitated the elaboration and adoption of the MPC strategy output.

In general, operators were open to use the MPC report and it helped them be more aware of the efficiency and run the plant in an effective way. Under some circumstances, it was however more difficult to follow, mainly due to technical reasons (e.g. maintenance, power failure). Nonetheless, in site #1 for instance, the most efficient boiler has been used more than ever compared to previous years. Operators were more aware of the importance of running it, although they were a bit outside of their comfort zone (the least efficient boiler works well and leads to fewer technical issues).

6. Conclusions

This paper presents the development of a predictive control strategy to optimize the operation of steam-based multi-boiler district heating systems. The approach is based on a machine learning model for district heating demand forecasting and data-driven part-load performance curves of natural gas boilers to recommend which boilers to operate at which power for the next 24 hours. This strategy was implemented in two Canadian demonstration sites for part of the 2023 winter season. Results showed that the proposed approach was able to reduce gas consumption by 2.8% in site #1 and 1.3% in site #2. Although modest, it yields to economic savings of $ 10,228 and $ 19,975 CAD and GHG emissions reduction of 45 t and 77 t CO2eq over the implementation period (2-3 months). The proposed strategy has helped investigate the system performance from a data-driven perspective, and it has showcased how operational data can be leveraged to improve district heating system operation to decarbonize the built environment. The proposed approach aims to be simple, generic, and scalable so it could be easily replicable to other sites while also allowing for adaptation for more complex systems including renewable energy systems and thermal energy storage devices.

CRediT author statement

Etienne Saloux: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Jason Runge**: Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Writing – review & editing. **Kun Zhang**: Software, Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the financial support of Natural Resources Canada through the Greening Government Fund. The authors would like to thank our external partners for sharing their experience, providing data and feedback, and implementing the proposed strategies. Internal and external reviewers are also acknowledged for their useful comments.

References

[1] Lund H, Østergaard PA, Chang M, Werner S, Svendsen S, Sorknæs P, et al. The status of 4th generation district heating: Research and results. Energy 2018;164:147–59. https://doi.org/10.1016/j.energy.2018.08.206.

[2] Buffa S, Cozzini M, D'Antoni M, Baratieri M, Fedrizzi R. 5th generation district heating and cooling systems: A review of existing cases in Europe. Renewable and Sustainable Energy Reviews 2019;104:504–22. https://doi.org/10.1016/j.rser.2018.12.059.

[3] Dainese C, Faè M, Gambarotta A, Morini M, Premoli M, Randazzo G, et al. Development and application of a Predictive Controller to a mini district heating network fed by a biomass boiler. Energy Procedia 2019;159:48–53. https://doi.org/10.1016/j.egypro.2018.12.016.

[4] Yu J, Guo L, Ma M, Kamel S, Li W, Song X. Risk assessment of integrated electrical, natural gas and district heating systems considering solar thermal CHP plants and electric boilers. International Journal of Electrical Power & Energy Systems 2018;103:277–87. https://doi.org/10.1016/j.ijepes.2018.06.009.

[5] Descamps MN, Leoncini G, Vallée M, Paulus C. Performance assessment of a multi-source heat production system with storage for district heating. Energy Procedia 2018;149:390–9. https://doi.org/10.1016/j.egypro.2018.08.203.

[6] Eslami Nejad P, Bastani A, Nguyen A, Saloux E. Borehole latent energy storage system integrated with solar thermal collectors and heat pumps. vol. 17, IBPSA; 2021, p. 207–14. https://doi.org/10.26868/25222708.2021.30418.

[7] Winterscheid C, Dalenbäck J-O, Holler S. Integration of solar thermal systems in existing district heating systems. Energy 2017;137:579–85. https://doi.org/10.1016/j.energy.2017.04.159.

[8] Mesquita L, McClenahan D, Thornton J, Carriere J, Wong B. Drake Landing Solar Community: 10 Years of Operation, Abu Dhabi, UAE, Oct 29 - Nov 2: 2017.

[9] Olsthoorn D, Haghighat F, Mirzaei PA. Integration of storage and renewable energy into district heating systems: A review of modelling and optimization. Solar Energy 2016;136:49–64. https://doi.org/10.1016/j.solener.2016.06.054.

[10] Rad FM, Fung AS. Solar community heating and cooling system with borehole thermal energy storage – Review of systems. Renewable and Sustainable Energy Reviews 2016;60:1550–61. https://doi.org/10.1016/j.rser.2016.03.025.

[11] Canadian Energy & Emissions Data Centre. District Energy in Canada. 2019. https://www.sfu.ca/content/dam/sfu/ceedc/publications/facilities/CEEDC%20-%20District%20Energy%20Report%202019.pdf

[12] Vandermeulen A, van der Heijde B, Helsen L. Controlling district heating and cooling networks to unlock flexibility: A review. Energy 2018;151:103–15. https://doi.org/10.1016/j.energy.2018.03.034.

[13] Buffa S, Fouladfar MH, Franchini G, Lozano Gabarre I, Andrés Chicote M. Advanced Control and Fault Detection Strategies for District Heating and Cooling Systems—A Review. Applied Sciences 2021;11:455. https://doi.org/10.3390/app11010455.

[14] Saloux E, Candanedo JA. Optimal rule-based control for the management of thermal energy storage in a Canadian solar district heating system. Solar Energy 2020;207:1191–201. https://doi.org/10.1016/j.solener.2020.07.046.

[15] Rehman H ur, Hirvonen J, Sirén K. Design of a Simple Control Strategy for a Community-size Solar Heating System with a Seasonal Storage. Energy Procedia 2016;91:486–95. https://doi.org/10.1016/j.egypro.2016.06.183.

[16] Saloux E, Candanedo JA. Sizing and control optimization of thermal energy storage in a solar district heating system. Energy Reports 2021;7:389–400. https://doi.org/10.1016/j.egyr.2021.08.092.

[17] Gonçalves D, Sheikhnejad Y, Oliveira M, Martins N. One step forward toward smart city Utopia: Smart building energy management based on adaptive surrogate modelling. Energy and Buildings 2020;223:110146. https://doi.org/10.1016/j.enbuild.2020.110146.

[18] Saloux E, Candanedo JA. Model-based predictive control to minimize primary energy use in a solar district heating system with seasonal thermal energy storage. Applied Energy 2021;291:116840. https://doi.org/10.1016/j.apenergy.2021.116840.

[19] Gravelsins A, Pakere I, Tukulis A, Blumberga D. Solar power in district heating. P2H flexibility concept. Energy 2019;181:1023–35. https://doi.org/10.1016/j.energy.2019.05.224.

[20] Schmidt D, Lygnerud K, Werner S, Geyer R, Schrammel H, Østergaard DS, et al. Successful implementation of low temperature district heating case studies. Energy Reports 2021;7:483–90. https://doi.org/10.1016/j.egyr.2021.08.079.

[21] Van Oevelen T, Vanhoudt D, Johansson C, Smulders E. Testing and performance evaluation of the STORM controller in two demonstration sites. Energy 2020;197:117177. https://doi.org/10.1016/j.energy.2020.117177.

[22] Van Oevelen T, Neven T, Brès A, Schmidt R-R, Vanhoudt D. Testing and evaluation of a smart controller for reducing peak loads and return temperatures in district heating networks. Smart Energy 2023:100105. https://doi.org/10.1016/j.segy.2023.100105.

[23] Runge J, Saloux E. A comparison of prediction and forecasting artificial intelligence models to estimate the future energy demand in a district heating system. Energy 2023:126661. https://doi.org/10.1016/j.energy.2023.126661.

[24] Saloux E, Candanedo JA. Forecasting District Heating Demand using Machine Learning Algorithms. Energy Procedia 2018;149:59–68. https://doi.org/10.1016/j.egypro.2018.08.169.

[25] Powell KM, Sriprasad A, Cole WJ, Edgar TF. Heating, cooling, and electrical load forecasting for a large-scale district energy system. Energy 2014;74:877–85. https://doi.org/10.1016/j.energy.2014.07.064.

[26] Johansson C, Bergkvist M, Geysen D, Somer OD, Lavesson N, Vanhoudt D. Operational Demand Forecasting In District Heating Systems Using Ensembles Of Online Machine Learning Algorithms. Energy Procedia 2017;116:208–16. https://doi.org/10.1016/j.egypro.2017.05.068.

[27] Kurek T, Bielecki A, Świrski K, Wojdan K, Guzek M, Białek J, et al. Heat demand forecasting algorithm for a Warsaw district heating network. Energy 2021;217:119347. https://doi.org/10.1016/j.energy.2020.119347.

[28] Potočnik P, Škerl P, Govekar E. Machine-learning-based multi-step heat demand forecasting in a district heating system. Energy and Buildings 2021;233:110673. https://doi.org/10.1016/j.enbuild.2020.110673.

[29] Leiprecht S, Behrens F, Faber T, Finkenrath M. A comprehensive thermal load forecasting analysis based on machine learning algorithms. Energy Reports 2021;7:319–26. https://doi.org/10.1016/j.egyr.2021.08.140.

[30] Chung WH, Gu YH, Yoo SJ. District heater load forecasting based on machine learning and parallel CNN-LSTM attention. Energy 2022;246:123350. https://doi.org/10.1016/j.energy.2022.123350.

[31] Makridakis S, Spiliotis E, Assimakopoulos V. The M4 Competition: 100,000 time series and 61 forecasting methods. International Journal of Forecasting 2020;36:54–74. https://doi.org/10.1016/j.ijforecast.2019.04.014.

[32] Miller C, Arjunan P, Kathirgamanathan A, Fu C, Roth J, Park JY, et al. The ASHRAE Great Energy Predictor III competition: Overview and results. Science and Technology for the Built Environment 2020;26:1427–47. https://doi.org/10.1080/23744731.2020.1795514.

[33] Gunay HB, Ashouri A, Shen W. Load Forecasting and Equipment Sequencing in a Central Heating and Cooling Plant: A Case Study. ASHRAE Transactions 2019;125:513–23.

[34] Labidi M, Eynard J, Faugeroux O, Grieu S. A new strategy based on power demand forecasting to the management of multi-energy district boilers equipped with hot water tanks. Applied Thermal Engineering 2017;113:1366–80. https://doi.org/10.1016/j.applthermaleng.2016.11.151.

[35] Saletti C, Gambarotta A, Morini M. Development, analysis and application of a predictive controller to a small-scale district heating system. Applied Thermal Engineering 2020;165:114558. https://doi.org/10.1016/j.applthermaleng.2019.114558.

[36] Aoun N, Bavière R, Vallée M, Aurousseau A, Sandou G. Modelling and flexible predictive control of buildings space-heating demand in district heating systems. Energy 2019;188:116042. https://doi.org/10.1016/j.energy.2019.116042.

[37] Liao Y, Huang G. A hybrid predictive sequencing control for multi-chiller plant with considerations of indoor environment control, energy conservation and economical operation cost. Sustainable Cities and Society 2019;49:101616. https://doi.org/10.1016/j.scs.2019.101616.

[38] Thangavelu SR, Myat A, Khambadkone A. Energy optimization methodology of multi-chiller plant in commercial buildings. Energy 2017;123:64–76. https://doi.org/10.1016/j.energy.2017.01.116.

[39] Saloux E, Zhang K. Data-Driven Model-Based Control Strategies to Improve the Cooling Performance of Commercial and Institutional Buildings. Buildings 2023;13:474. https://doi.org/10.3390/buildings13020474.

[40] American Society of Heating, Refrigeration and Air conditioning Engineers. ASHRAE Guideline 14-2014: Measurement Of Energy, Demand, And Water Savings. Atlanta, GA: 2014.

[41] National Research Council of Canada. National Energy Code of Canada for Buildings 2020, Fifth Edition. 2020.

[42] Natural Resources Canada. CanMETEO, English / French. 2017. https://www.nrcan.gc.ca/energy/software-tools/19908.

[43] Candanedo JA, Saloux E, Hardy J-M, Platon R, Raissi-Dehkordi V, Côté A. Preliminary Assessment of a Weather Forecast Tool for Building Operation, Purdue, July 9-12: 2018.

[44] Hobson BW, Gunay HB, Ashouri A, Newsham GR. Occupancy-based predictive control of an outdoor air intake damper: A case study. eSim 2020: 11th Conference of IBPSA-Canada, Vancouver, Canada, June 14-16: 2020, p. 8.

[45] Cotrufo N, Saloux E, Hardy JM, Candanedo JA, Platon R. A practical artificial intelligence-based approach for predictive control in commercial and institutional buildings. Energy and Buildings 2020;206:109563. https://doi.org/10.1016/j.enbuild.2019.109563.

[46] Saloux E, Cotrufo N, Candanedo J. A Practical Data-Driven Multi-Model Approach to Model Predictive Control: Results from Implementation in an Institutional Building. 6th International High Performance Buildings Conference 2021.

[47] De Coninck R, Helsen L. Practical implementation and evaluation of model predictive control for an office building in Brussels. Energy and Buildings 2016;111:290–8. https://doi.org/10.1016/j.enbuild.2015.11.014.

[48] Arroyo J, Spiessens F, Helsen L. Comparison of Model Complexities in Optimal Control Tested in a Real Thermally Activated Building System. Buildings 2022;12:539. https://doi.org/10.3390/buildings12050539.

[49] Freund S, Schmitz G. Implementation of model predictive control in a large-sized, low-energy office building. Building and Environment 2021;197:107830. https://doi.org/10.1016/j.buildenv.2021.107830.

[50] Saloux E, Candanedo JA. Control-oriented model of a solar community with seasonal thermal energy storage: development, calibration and validation. Journal of Building Performance Simulation 2018:1–23. https://doi.org/10.1080/19401493.2018.1523950.

[51] Saloux E, Candanedo JA. Modelling stratified thermal energy storage tanks using an advanced flowrate distribution of the received flow. Applied Energy 2019;241:34–45. https://doi.org/10.1016/j.apenergy.2019.02.075.

APPENDIX A: Statistical analysis of implementation results

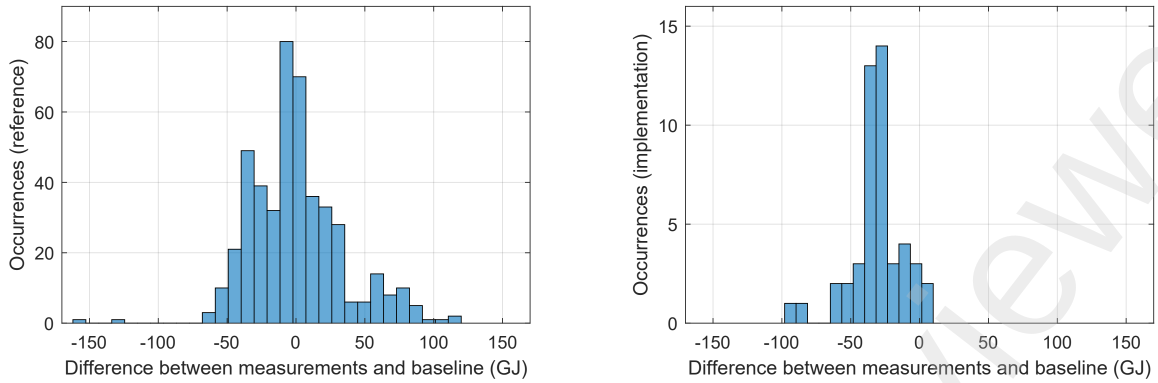
Implementation results showed savings of 2.8 % for site #1 and 1.3 % for site #2. Since these reductions are modest, a statistical analysis has been conducted to confirm that savings were actually obtained. To this purpose, we calculated the difference between measurements and the baseline model for two datasets: a) the training dataset used to derive the baseline model ($\Delta E_{\text{reference}}$), which is typical of model accuracy, and b) the implementation dataset typical of the implemented MPC strategy (ΔE_{impl}):

$$\Delta E_{\text{reference}} = \dot{Q}^{\text{hist}}_{\text{gas,tot,meas}} - \dot{Q}^{\text{hist}}_{\text{gas,tot,calc}} \tag{A-1}$$

$$\Delta E_{\text{impl}} = \dot{Q}^{\text{impl}}_{\text{gas,tot,meas}} - \dot{Q}^{\text{impl}}_{\text{gas,tot,calc}} \tag{A-2}$$

where $\dot{Q}^{\text{hist}}_{\text{gas,tot,meas}}$ and $\dot{Q}^{\text{hist}}_{\text{gas,tot,calc}}$ are the total daily gas consumption measured with the training dataset and calculated with the model, respectively. $\dot{Q}^{\text{impl}}_{\text{gas,tot,meas}}$ and $\dot{Q}^{\text{impl}}_{\text{gas,tot,calc}}$ are the total daily gas consumption measured and calculated with the model during the implementation phase. Comparing the statistical distribution of variables $\Delta E_{\text{reference}}$ and ΔE_{impl} allows to grasp changes in probability distributions and indicate whether savings fell within model accuracy or were truly obtained.

Figure A-1 shows the statistical distribution of the difference between measurements and baseline model for site #1. From this figure, implementation results are almost consistently lower than the baseline (mainly negative values in Figure A-1b), which is clearly not the case for the training data for the baseline model. Table A-1 gives the probability distribution of the difference between measurements and baseline model for site #1. μ (= 2.8e-13 GJ) and σ (= 34 GJ) are the mean and standard deviation of the distribution for training the baseline model for site #1. In this table, energy savings are indicated with a higher distribution during the implementation phase compared to the baseline model training when $p < \mu$, and a lower distribution when $p > \mu$. For site #1, most of the values during the implementation phase are comprised between μ–2σ and μ, which confirms obtained savings. The mean of the distribution for implementation for site #1 is -31 GJ.



(a) (b)

Figure A-1. Statistical distribution of the difference between measurements and baseline model for site #1 (Ontario): (a) for historical data used for benchmarking, (b) during MPC implementation.

Table A-1. Probability distribution of the difference between measurements and baseline model for site #1 (Ontario).

Probability distribution (p)	Benchmarking	MPC
$p < \mu\text{-}2\sigma$	0.4 %	4.2 %
$\mu\text{-}2\sigma \leq p < \mu\text{-}\sigma$	14.9 %	37.5 %
$\mu\text{-}\sigma \leq p < \mu$	40.8 %	54.2 %
$\mu \leq p < \mu+\sigma$	32.0 %	4.2 %
$\mu+\sigma \leq p \leq \mu+2\sigma$	6.8 %	0.0 %
$p > \mu+2\sigma$	5.0 %	0.0 %

The same exercise was conducted for site #2; Figure A-2 and Table A-2 show the results. μ (= 2.3e-13 GJ) and σ (= 58 GJ) are the mean and standard deviation of the distribution for training the baseline model for site #2. In this case, observations are less evident but still tend to confirm obtained savings. Except for $\mu \leq p < \mu+\sigma$ (34.3% vs. 31.7%), the probability distribution shows energy savings: higher distribution when $p < \mu$ and lower distribution when $p > \mu$, compared to the training data for the baseline model. For site #1, most of the values during the implementation phase are comprised between μ–2σ and μ+σ, which confirms obtained savings. The mean of the distribution for implementation for site #1 is -14 GJ.

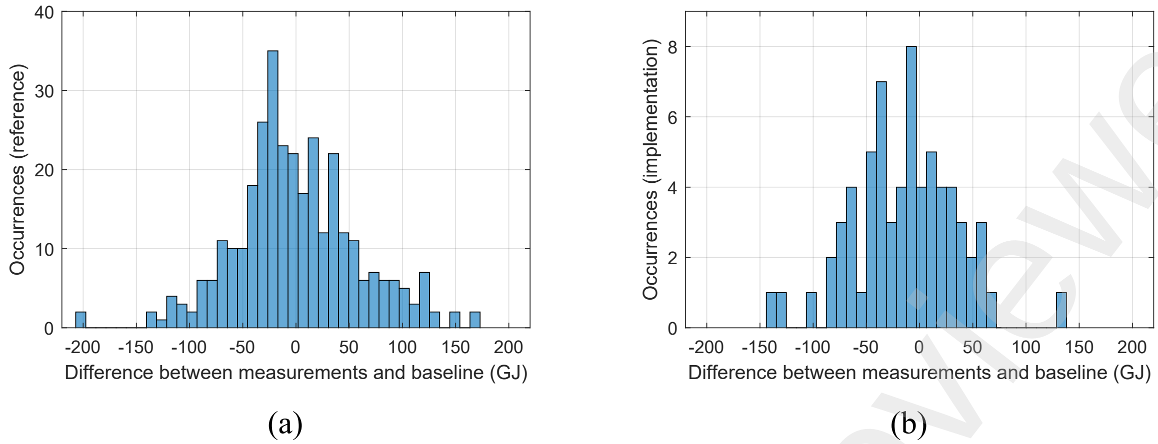


Figure A-2. Statistical distribution of the difference between measurements and baseline model for site #2 (Québec): (a) for historical data used for benchmarking, (b) during MPC implementation.

Table A-2. Probability distribution of the difference between measurements and baseline model for site #2 (Québec).

Probability distribution (p)	Benchmarking	MPC
$p < \mu\text{-}2\sigma$	1.5 %	3.0 %
$\mu\text{-}2\sigma \leq p < \mu\text{-}\sigma$	12.3 %	14.9 %
$\mu\text{-}\sigma \leq p < \mu$	40.3 %	43.3 %
$\mu \leq p < \mu\text{+}\sigma$	31.7 %	34.3 %
$\mu\text{+}\sigma \leq p \leq \mu\text{+}2\sigma$	10.2 %	3.0 %
$p > \mu\text{+}2\sigma$	4.0 %	1.5 %