Downloaded from orbit.dtu.dk on: Feb 07, 2025

DTU Library

Analysis of the impact of predictive models on the quality of the model predictive control for an experimental building

Erfani, Arash ; Yu, Xingji; Kull, Tuule Mall; Bacher, Peder; Jafarinejad, Tohid ; Roels, Staf; Saelens, Dirk

Publication date:
2021

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Erfani, A., Yu, X., Kull, T. M., Bacher, P., Jafarinejad, T., Roels, S., & Saelens, D. (2021). *Analysis of the impact of predictive models on the quality of the model predictive control for an experimental building*. Paper presented at Building Simulation 2021 Conference, Bruges, Belgium.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Analysis of the impact of predictive models on the quality of the model predictive control for an experimental building

Arash Erfani[1], Xingji Yu[2], Tuule Mall Kull[3], Peder Bacher[4], Tohid Jafarinejad[1], Staf Roels[1], Dirk Saelens[1,5]
[1]KU Leuven, Department of Civil Engineering, Building Physics & Sustainable Design Section, Leuven, Belgium
[2]NTNU, Trondheim, Norway
[3]Tallinn University of Technology, Tallinn, Estonia
[4]DTU, Lyngby, Denmark
[5]EnergyVille, Genk, Belgium

Abstract

To increase energy efficiency of the building sector, many measures have been suggested which often require a predictive model of the building to function. Developing these models is one of the crucial challenges hampering pervasive use of these measures. Therefore, this study aims at assessing the impact of using different predictive models in an energy optimization application for an experimental building. First step in achieving this goal is developing various data-driven models for the investigated building in this study. Afterwards, a framework has been developed in which the performance of predictive models in the optimization strategy namely Model Predictive Control (MPC) could be evaluated. The results reveal that common indicators in the literature do no suffice to score the performance of models used in MPC, but another state-of-the-art indicator; multi-step ahead prediction error is more suitable for evaluating predictive models deployed in MPC.

Key innovations

- Finding a proper Key Performance Indicator (KPI) for evaluating various predictive models in an MPC framework
- Assessing the impact of one-step ahead prediction and multi-step ahead prediction accuracy on model's quality in MPC
- Applying Support Vector Machine as a powerful AI tool for building behavior identification

Practical implications

This paper could guide practitioners who work on building energy optimization in choosing a suitable model in their optimization algorithm .In addition; we suggest an appropriate criterion to assess the predictive models in terms of their performance in MPC, which could be instrumental for both researchers and practitioners.

Introduction

Surveys have shown that building stock has the highest potential in terms of energy saving to achieve *well below 2 ºC* target by 2050 set in Paris agreement (EU Commission, 2018). Approximately 71% of all the final energy use in residential sector in Europe is used for space heating alone (EU Commission, 2018). Hence, there is a considerable potential of energy saving which could be activated by optimizing performance of existing HVAC systems. Several strategies have been proposed to increase existing building's energy efficiency such as RES integration (Haddadi et al., 2019), shading control (Da Silva et al., 2012), optimal control of HVAC systems (De Coninck & Helsen, 2016), heat recovery (Jafarinejad, et al., 2019), glazing improvement (Djamel & Noureddine, 2017), smart houses and etc. (Eicker et al., 2015; Guerra-Santin & Tweed, 2015). RES integration is one of the promising options. However, their uncertain nature affects all the energy users such as buildings (Reynders et al., 2017). One of these impacts is that the integration of RES in buildings renders performance of traditional control strategies non-optimal (Sangi et al., 2019). Hence, substantial attention has been paid to advanced control strategies recently (Afram & Janabi-Sharifi, 2014).

Amongst various control strategies suggested for optimization of building's thermal performance, Model Predictive Control (MPC) is one of the most promising ones. MPC is an active control strategy which optimizes the performance of a system over a given time horizon (Drgoňa et al., 2020). It has shown a considerable potential in optimizing the performance of HVAC systems along with facilitating the integration of RES in buildings (Atam & Helsen, 2016). MPC has the ability to handle slow moving dynamics, which matches the requirements for a good optimization strategy for buildings. To this end, MPC uses a model of the building to predict its thermal behavior in the future. This prediction feature gives MPC a crucial edge compared to other controllers (Reynders et al., 2014; Sourbron et al., 2013). Myriad studies have been carried out on application of MPC in buildings (Drgoňa et al., 2020). In spite of the abundance of such studies, MPC is still not used prevalently in buildings. One of the main issues hampering easy and cost-efficient implementation of MPC in buildings is developing a predictive model of the building (Sourbron et al., 2013).

This study aims at comparing different modeling techniques, which are used to identify a building's thermal characteristic and are integrated into the MPC framework as the predictive model. In this study, we place our focus on data-driven methods used for characterizing building's thermal behavior since there is an ever-increasing interest in employing data-driven techniques for building energy optimization applications (Sepasgozar et al., 2020).

In general, data-driven methods could be divided into two categories, black box and grey-box models. In black box modeling techniques, the mapping between the input and the output is a mathematical one. These models have a wide range from a simple regression to complex AI-based methods such as deep learning methods (Drgoňa et al., 2020). As for grey-box models, they could be defined as a hybrid between mathematical (black-box) models and physics-based (white-box) models. In the most popular type of grey-box models used in buildings, structure of the model is determined by simplified physical laws governing building's dynamics. Next, , parameters of the model are estimated based on datasets (Afroz et al., 2018).

An important question, which comes up in the process of model selection and training, is how do I know which model would perform the best in my MPC? In other words, how to quantify the quality of different predictive models in the context of MPC. In this work, we endeavor to answer the above question by applying different KPIs to different models and assessing their suitability to score those models. We look into the one-step ahead prediction and multi-step-ahead prediction error (MSPE) of the models as two different KPIs for predictive models.

MSPE has been considered for quantifying predictive models in MPC before. The concept of Model predictive control Relevant Identification (MRI) for buildings was first introduced by (Žáčeková & Prívara, 2012) in which they developed a grey-box model based on MSPE minimization. Thereafter, some research studies in this field reported MSPE of their models in their work (Zhan & Chong, 2021). In one of the most relevant of these studies, (Picard et al., 2016) developed a detailed Modelica model for an office building. This Modelica model is then linearized into a state space model. Two grey-box models were developed in their study as well. One was identified based on measurements and the other one based on the proxy data obtained from the emulator. Performance of these three models are evaluated for 1 hour and 24 hours ahead. They showed that the most accurate model (linearized state space model) used 50% less energy while providing better thermal comfort. In a similar study, (Picard et al., 2017) applied model order reduction techniques to a white box model of a residential building and reported their model's quality based on both one-step ahead error and MSPE. They concluded that such models should be of higher order compared to their peer data-driven models to yield an MPC with good performance.

Although the concept of MSPE has been used before in the context of MPC design for a building, but a thorough analysis on its suitability for scoring different predictive models is lacking. In other words, previous studies did not consider various data-driven models in their structures. In addition, they did not distinguish between the impact of one-step and multi-step ahead prediction accuracy on the models and its impact on the controller's performance. In this paper, we take into account the MSPE associated with each model as a KPI and compare it to one-step ahead prediction error. In addition, Support Vector Machine (SVM) as a powerful tool in machine learning field has been applied for in the context of MPC.

The models developed for this study are AutoRegressive with eXogenous inputs (ARX), grey-box RC models with different orders, black-box State Space (SS) models with different orders, SVM and Artificial Neural network (ANN) with a Non-linear Auto-Regressive with eXogenous inputs (NARX) structure. The results of applying MPC with different models for 2 weeks in the heating season are presented. To be able to evaluate performance of these models in MPC, a framework has been considered. First, a simulation model of an experimental building equipped with an underfloor heating system has been developed. This simulation model replaces the real building in our simulations. This experimental building is one of the experimental buildings in the context of IEA Annex 71 project. MPC has been developed in MATLAB SIMULINK environment. The simulation model has been coupled to the controller using an Application Programming Interface (API). Furthermore, MPC results are compared to the ones of a well-tuned Rule-Based Controller (RBC) to show its superiority over traditional control methods.

First, framework of the study is described. Subsequently, predictive models developed for this study are described. Then, we proceed by presenting and analyzing the results of different MPCs. Last section concludes the paper.

Framework

In this section, various parts required for evaluating MPC for a building are described; starting with the building itself. Then, the API used in this study is briefly explained. Afterwards, structure of the MPC itself is explained. General schematic of the framework applied in this study is shown in Figure 1, which is explained in the rest of this section from top to bottom.

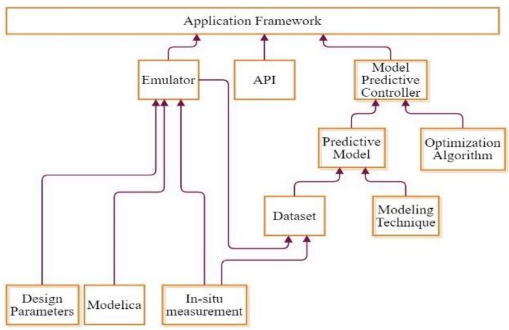

Figure 1: MPC framework for data-driven models

As it could be seen in Figure 1, one of the important components in simulating the performance of MPC in a building is to have a simulation model, which mimics the behaviour of the real building. Henceforth, this model is called the *emulator*. The emulator model is usually a complex white-box model, which due to its high computational load is not suitable to be deployed in real-time optimization applications (Afram & Janabi-Sharifi, 2014). Therefore, there is a need for simple and efficient model embedded in the MPC which is responsible for predicting building's dynamics over a given time horizon. These models are called *predictive models.* The rest of this section is allocated to detailed description of different components of the framework shown in Figure 1.

Building

We first give a short description of the real building, which has been modelled using Modelica language in Dymola software using the OpenIDEAS library (Baetens et al., 2015). This simulation model serves as the emulator in this study. The building under study is one of the test cases of the IEA ANNEX 71 project titled: *"Building energy performance assessment based on in situ measurements."* Test building in this study is a two-storey experimental dwelling located in Holzkirchen, Germany. (See Figure 2)

Figure 2: Test Building

This building is equipped with multiple instruments for measuring and storing time-series data of different variables. Heating is provided by means of an air-to-water heat pump, which provides hot water for an underfloor heating system that is installed for both floors. (Figure 3). Occupants are introduced to the building using electrical heaters based on a pre-defined schedule. Thermometers are installed in all rooms to measure room's temperature. Ventilation system functions based on a pre-defined schedule and is equipped with measurement instruments both in exhaust and supply terminals.

Figure 3: Heat pump and the underfloor heating system

API

The controller in this study has been developed in SIMULINK, whereas the emulator is developed in Dymola. Hence, a way of communication is required to make the co-simulation between the two softwares possible. To tackle this issue, we use an interface, which facilitates the connection between Dymola and Matlab, which is called Functional Mock-up Interface (FMI).

The developed building model in Modelica is essentially composed of equations derived from physical laws. FMI translates these equations into binary format, which is supported by many simulation tools including MATLAB and Python. This binary file could be loaded and run by these softwares (Modelica Association Project "FMI," 2013). In this work, the emulator model of the building is compiled as a Functional Mock-up Unit (FMU). Then it is imported into SIMULINK using the FMU block of MATLAB. From there on, the FMU serves as the emulator model in our MPC framework and easily communicates with the controller in the Simulink environment.

Model Predictive Controller

MPC is composed of two main parts. A predictive model and an optimizer. Interested readers can refer to (Drgoňa et al., 2020) for further details on MPC formulation for buildings. The objective of the MPC here is to minimize the electricity use of the heat pump while minimizing indoor thermal discomfort. Total discomfort is calculated as Kelvin hours outside thermal comfort band. Heat pump's electricity use is obtained from the emulator model (FMU). A day-ahead electricity price profile from real-life implementation has been chosen as a way to reflect the integration of RES in the building load profile. Furthermore, to illustrate the suitability of MPC with respect to other common control methods, a well-tuned RBC has been designed and applied to our case study and the results of this RBC are compared to the ones of MPC. In addition, we investigate the propriety of two important KPIs in scoring predictive models in the context of MPC; namely one-step ahead and multi-step ahead prediction error. In the following, constraints acting on the system along with the optimizer used in this study are described.

Constraints

Constraints acting on the system are divided into two different types. First type of constraints are the ones imposed on the inputs. The manipulated inputs considered in this study are the heat pump's on/off status (u_1) and heat pump's supply water temperature (u_2). It should be mentioned that in this study the mass flow rate of the supply water is considered constant (when the heat pump is on) during the whole simulation. The issue with the constraints on the heat pump operation is the fact that if heat pump operates with low loads, it would have a low COP, which should be avoided. Determining the point that the efficiency of the heat pump deteriorates depends on many factors, including the modulation rate of the compressor, ambient temperature and supply water temperature. Hence, imposing an accurate bound for lower modulation of the heat pump is not straightforward and would complicate the optimization problem (Verhelst et al., 2012). To avoid these complexities, a lower band for supply water temperature is imposed to avoid performing with a low COP level. The upper limit of supply water temperature is extracted from the datasheet provided by the manufacturer. To wrap up, Equations (1)-(2) show the constraints on input signals of the heating system:

$$u_1 \epsilon [0,1] \quad (1)$$

$$28 \le u_2 \le 45 \quad (2)$$

In which u_1 is heat pump's on/off status and is a boolean variable, 1 means heat pump is on and 0 denotes that it is off. Here, u_2 is the supply water temperature provided by the heat pump. Second type of constraint applied in this study is indoor thermal comfort bands. Comfort band considered in this study is [20 24] from 7:00 to 23:00 with a night setback of [18 22] from 23:00 to 7:00.

$$T_{low,t} \leq T_{in,t} \leq T_{up,t} \tag{3}$$

$T_{low,\,t}$ is equal to the lower limits of the bands defined above and $T_{up,\,t}$ corresponds to the upper limit of comfort bands. In this study, building is seen as one thermal zone and one temperature is used to represent the whole building, which is the volume-averaged temperature of all 10 thermal zones in Figure 2.

OCP formulation

Now having defined constraints and the objective function, the Optimal Control Problem (OCP) can be fully formulated. To avoid the high switching frequency of heat pump, time step chosen for this study is one hour. Furthermore, a control horizon of 12 hours (N=12) has been applied in this study which is sufficiently larger than the time constant of the building and the computational load does not become too large.

$$\underset{\vec{u}_1,\vec{u}_2,\ldots\vec{u}_{k+N-1}}{\text{Min}} \sum_{k=0}^{N-1} Lv_{k+i+1} + C_{el,k+i+1}\hat{P}_{el,k+i+1} \tag{4}$$

$$\hat{T}_{in,k+i+1} = f(\vec{x}_{k+i}, \vec{u}_{k+i}, \vec{d}_{k+i}) \tag{5}$$

$$\hat{P}_{el,k+i+1} = g(\vec{u}_{k+i}, \vec{d}_{k+i}) \tag{6}$$

$$\hat{T}_{in,k+i+1} + v_{k+i+1} \geq T_{low,k+i+1} \tag{7}$$

$$\hat{T}_{in,k+i+1} - v_{k+i+1} \leq T_{up,k+i+1} \tag{8}$$

$$u_1 \in [0,1] \tag{9}$$

$$28 \leq u_2 \leq 45 \tag{10}$$

$$\vec{u} = [u_1, u_2] \tag{11}$$

$$v_{k+i+1} \geq 0 \;\; , i=0,1,\ldots.N-1 \tag{12}$$

In these equations, u_1 and u_2 indicate heat pump's status and its supply water temperature, d_k and x_k represent disturbances acting on the system and the systems states at time step k respectively. Equation (4) describes the objective function which is composed of two terms, one for the electricity cost and the other one penalizes thermal discomfort level, C_k is the electricity cost at each time step and $\hat{P}_{k+i}$ represents the estimation of heat pump's electricity use for *i* steps forward in time. In equation (5), $\hat{T}_{in,k+i}$ denotes the estimation of indoor temperature *i* steps ahead in time, f(.) is essentially the predictive model which provides the estimated temperature profile of the building throughout the control horizon (N) while g(.) represents the estimation of heat pump's electricity use. Heat pump's electricity consumption is estimated using a quadratic function of the supply water temperature (u_2) , return water temperature ($T_{w,ret}$) and ambient temperature (T_{amb}) which is multiplied by the status of the heat pump (u_1):

$$\hat{P}_{el,k+1} = u_1 * g_{quad}(u_2, T_{ret}, T_{amb}) \tag{13}$$

g_{quad} (.) represents the quadratic function on its three arguments. Equations (7-8) show the soft constraints on thermal comfort bands designated in Equation (3). These soft constraints help the solver in coming up with a feasible solution by allowing thermal comfort bands to be violated. The latter is achieved by introducing a slack variable (v_k). Its value is penalized in the objective function given a weight of L.

Solver

Now with the OCP defined we can choose a suitable type of solver for this case study. Looking at equations (13) and (4) we realize that the second term in the objective function is a non-linear function of decision variables ($u_{1,k}$, $u_{2,k}$). Hence, even in case that linear predictive models are deployed (equation (5)), we are dealing with a mixed integer non-linear programming problem, which is most likely to be non-convex. Therefore, the solver has to be able to handle non-convex mixed integer programming problems. In this study, we used Genetic Algorithm (GA) as the solver since it has proven to be able to deal with such programming problems (Afram & Janabi-Sharifi, 2014). To achieve the latter, Matlab's GA function has been deployed (*Global Optimization Toolbox*, 2021).

Predictive Model

As it could be seen in Figure 1, a data-driven predictive model has two important attributes: dataset and modelling technique. In this section, we are going to describe these two components of predictive models.

Dataset

To train and test the data-driven models datasets are essential. If we use data from in-situ measurements for training the models, quality of the models in the simulation environment would be influenced not only by the accuracy of the predictive model but also by the accuracy of the simulation model. Thus, it will not be possible to distinguish between the impact of the model quality and emulator's accuracy on the MPC results. Hence, with respect to the goal of this study, which is investigating the effect of different modelling techniques, proxy data generated from the emulator has been used for training data-driven models instead of in-situ measurements.

To create data for training models, we generated two random sequences for the heat pump's status and heat pump's supply temperature. The resulting temperature of the emulator is shown in Fig. 4. As it could be seen in Fig. 4, the indoor temperature varies between 15.5 and 25, which fully covers the full range of thermal comfort assigned in Equation (3). Throughout this paper, this dataset is used to train and validate data-driven models. To avoid the impact of dataset bias on modelling techniques, another dataset is used to test all the models, which has not been used in training process (explained later on).

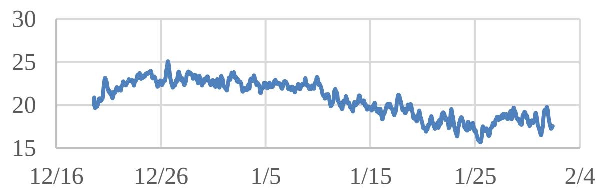

Figure 4: Temperature profile generated by feeding random sequence of inputs to emulator

Modelling technique

In this section, different modelling techniques applied for the purpose of this study are explained. Table 1 shows the general structure of these models.

Table 1- General description of the models

Model name	Model order	Inputs	R^2 (%) 1-step ahead
Grey-box 1	1	T_e, *GHI*, H_{in}	97.5
Grey-box 2	2	T_e, *GHI*, T_s, S_{hp}	97.8
State Space 1	1	T_e, *GHI*, T_s, S_{hp}, *IHG*,	97.5
State Space 2	7	T_e, *GHI*, T_s, S_{hp}, *IHG*, VFR_{av}	98
ARX	3	T_e, *GHI*, T_s*S_{hp}, *DHI*, VFR_{living}	97.5
NARX (ANN)	7	T_e, *GHI*, T_s, S_{hp}, *IHG*, VFR_{av}	98.8
SVM	---	T_e, *GHI*, T_s, S_{hp}, *IHG*, VFR_{av}	99.5

Input variables in the table are as follows:

T_e: Ambient temperature (°C)

GHI: Global horizontal irradiance (W/m^2)

T_s: Supply water temperature (u_2) (°C)

S_{hp}: Heat pump status (u_1)

H_{in}: Heat injected to building by underfloor heating (W)

IHG: Internal heat gains (W)

VFR: Volumetric flow rate of ventilation system (m^3/h)

DHI: Diffuse horizontal irradiance (W/m^2)

Grey-box Model 1

As a popular building identification method, grey-box models are included in this study. For the purpose of this study, we start with a simple structure for grey-box models (Figure 5) and we build up complexity onward. For each of the grey box models, first, the structure of the model is determined and then the parameters of the model are identified based on the training dataset. Interested readers are referred to Reynders et al., (2014) for more details on grey-box models. This grey-box model has only one state, which represents the average temperature of the indoor air.

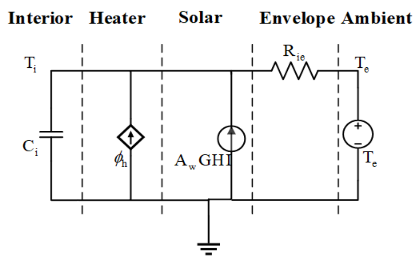

Figure 5: Grey-box model with one state

Grey-box Model 2

Another grey-box model used in this study has two states, one for the air temperature, while the other one represents the floor temperature (as the heating medium) of the building (Figure 6). Interested readers are referred to (Bacher & Madsen, 2011) for more details on this model.

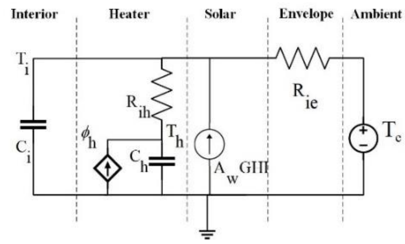

Figure 6: Grey-box model with two states

Autoregressive with exogenous inputs

One of the most common black box methods applied for building behaviour identification is Auto-Regressive with eXogenous input (ARX) models (Bourdeau et al., 2019). To develop this model, a Principle Component Analysis (PCA) has been carried out which led to the selection of optimal set of inputs (Table 1) as well as the number of output lags used for the building behaviour identification, which is three. The general structure of ARX models for identifying a multi input single output system is given in equation (14).

$$A(z)y(z) = B_1(z)u_1(z) + \ldots B_{n_u}(z)u_{n_n}(z) + C(z)e(z) \quad (14)$$

In this Equation, n_u stands for the number of input signals, which is five in this case (Table 1). A(), B() and C() are polynomials representing the parameter of the ARX model which are estimated using the training dataset.

State Space

Another popular modelling technique in the category of black-box models is state space identification, which has been successfully deployed for optimization of HVAC systems as well (Bourdeau et al., 2019). One of the advantages of linear state space models is the fact that most linear systems could be described using this formulation and most of the notations and theorems developed regarding MPC are based on state space representation of the systems (Maciejowski, 2002).

In this study, we focus on the Linear Time-Invariant (LTI) state space models. Two different state space models are deployed using Matlab's system Identification toolbox (*System Identification Toolbox*, 2021) .One only has one state, which is the simplest state space model possible; as for the other model, the number of states has been determined based on Singular Value Decomposition (SVD) of the Henkel Matrix for which 7 states is selected as the optimum number of states (Drgoňa et al., 2018). It should be noted that to identify this model 'Focus' is put on 'Simulation' rather than 'Prediction', which is an option provided in Matlab's system identification toolbox.

Artificial Neural Network

Artificial Neural networks (ANNs) are known as a powerful tool in machine learning. They are inspired by the structure of the brain (Abu-Mostafa, 1992). There is an ever-increasing interest in applying ANNs for HVAC system optimization applications. There are various architectures of ANNs available. One of the architectures deemed suitable for the application of building characterization is NARX which has proven successful in capturing dynamics of HVAC systems and it is selected here as well (Bourdeau et al., 2019). These models have essentially the same input-output structure as ARX models. The main difference is that ANN-based NARX models use neurons for capturing system's dynamics instead of linear mapping in the ARX case. Interested readers can refer to (Erfani et al., 2018; Jafarinejad et al., 2019) for further details on NARX model.

Support Vector Machine

Support Vector Machine is a powerful method originally suggested for classification applications. Recently, it has been successfully applied in many regression applications as well, which is called Support Vector Regression (SVR). Like other non-linear regression techniques, SVM tries to find the function between the input and the output (*f(.)).* To carry out this task, SVM transforms the input-output space to a higher dimension space, which is called feature space. Function *f()* then would be in the form:

$$f(x) = < \omega, \Phi(x) > + b \quad (15)$$

In which x is the input vector, Φ represents the higher dimension mapping and ω and b are parameters that are estimated by solving a convex optimization problem called the primal objective function. Operator < , > describes the kernel function in the feature space. Interested readers are referred to (Kumar & Kar, 2009) for more details on SVM.

Test Dataset

Since different combinations of the training dataset were used to train and validate the models, a second dataset was generated solely with the purpose of testing the models. As stated earlier, MPC solves an optimization problem over a given time horizon. Hence, predictive model used in the MPC should be able to provide acceptable predictions not only for one-step ahead in time but also throughout the whole control horizon. Therefore, here we are going to investigate whether one-step ahead prediction accuracy is a good enough indicator to reflect the quality of predictive models or should we look into multi-step ahead prediction accuracy. The results obtained by running the models against test dataset are presented in Figure 7.

This figure provides the boxplot accuracy of different modelling techniques used in this study. The maximum in each box corresponds to the one-step ahead prediction accuracy while the minimum corresponds to N (Control horizon) steps ahead prediction accuracy. As could be seen in Figure 7, NARX model and the SVR are the best performing models in terms of one-step ahead prediction accuracy but they are not the best models when looking into the multi-step ahead prediction.

Results

In this section, the results of deploying different predictive models in the context of MPC are brought out. The goal of this study is twofold. First, showing that integration of RES into the building energy structure is plausible by utilizing MPC. Integration of RES is considered here as a variable electricity pricing structure. The other goal of this study is finding a suitable KPI to score different predictive models, which are used in the MPC. The simulations have been carried out for a total duration of two weeks from 19th of December to second of January. The weather data used for this study is from in-situ measurements of the building. Perfect forecast is considered for weather as well.

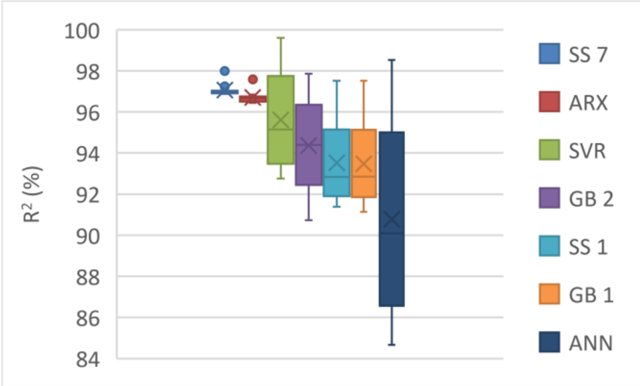

Figure 7: R^2 (%) of models against test dataset

KPIs considered in this study based on which the MPCs are compared are total thermal discomfort level and the total electricity cost of the heating system. Attempting to analyze the results of different MPCs, we come across an impediment, which blocks the way of a straightforward comparison of the controllers. This barrier arises from the fact that the MPC aims at optimizing two objectives (thermal discomfort and electricity cost) which are not physically related to each other. Therefore, by changing the weight (L in Equation (4)) optimal performance of the controllers are obtained in a way that they yield similar discomfort levels as could be seen in Figure 8. By employing this method, we ensure that all controllers have a similar thermal discomfort so that we can compare the controllers only based on electricity cost. As could be seen in Figure 7, NARX model and the SVR are the best performing models in terms of one-step ahead prediction accuracy. Nevertheless, these two models are not the best performing models in our framework. This statement is especially more significant in the case of the NARX model since it leads to the highest electricity cost compared to the other models. Looking at MSPE, one can easily realize that, although the NARX model has the second highest one-step ahead R^2 (See Table 1), its multi-step ahead prediction performance is the poorest amongst all the models (see Figure 7). The reason for this observation is explained by the fact that ANNs easily become over-fit to training data if no regularization of some sort is used (Afroz et al., 2018). This issue should be tackled when using ANNs as predictive models otherwise one might end up with an ANN model, which is highly accurate for one-step ahead prediction but provides poor forecasts for multi-step ahead prediction.

Analyzing the results as illustrated in Figure 8, it could be concluded that the best performing MPC (deploying state space model with 7 states) compared to the RBC, reduces electricity cost from 11€ to 8.5 € which corresponds to 22.7%. Comparing different MPCs using Figure 8 we can deduce that the difference between electricity cost resulted from using different predictive models in the MPC is 7% (Electricity cost of 8.5 € in the SS7 model compared to 9.1 € achieved by using the NARX model). Considering the

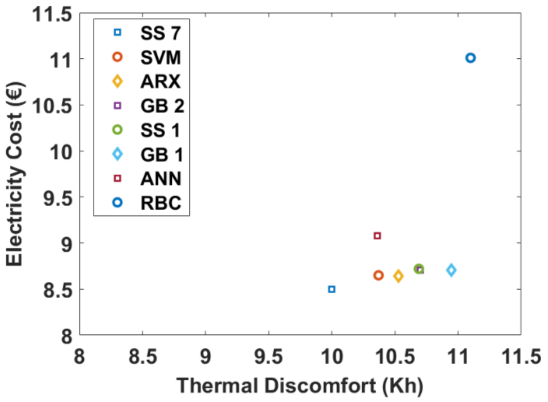

Figure 8: KPIs deploying different predictive models

22.7% as the highest potential of MPC achieved by our models for this case study, it could be inferred that the models used here vary by 24% in terms of activating the potential energy savings achieved by MPC, which demonstrates the importance of using models with high multi-step ahead prediction accuracy in the MPC.

Results obtained by applying state space model with 7 states, are presented in Figure 9 and Figure 10. It is illustrated in Figure 9 that the controller is able to maintain the temperature within the thermal comfort band although there are some minor violations. These violations could have two main causes. First, the magnitude of weight (L) scalar in the objective function, which allows thermal discomfort to some extent especially when the electricity cost is relatively high. The second reason behind the minor thermal discomfort could be the mismatch between the predictive model and the emulator. Electricity price shown in Figure 10 is based on time of use pricing structure from a supplier in Belgium. As seen in Figure 10, the load profile does not completely correspond with the time-of-use price. This observation is expected since the MPC does not optimize the building's behaviour only for one time-step but for the whole control horizon.

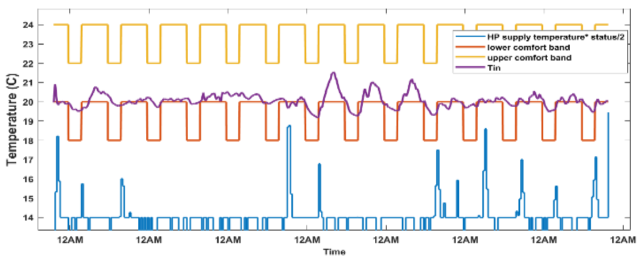

Figure 9: Building's temperature profile due to MPC

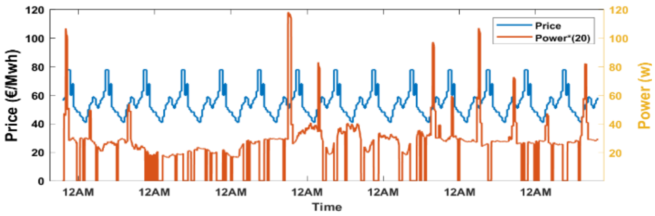

Figure 10: Electricity use against electricity price

Conclusion

Application of different data-driven models to serve as the predictor in a Model Predictive Control (MPC) are assessed. To score the predictive models in MPC, one-step ahead and Multi-Step ahead Prediction Error (MSPE) of the models are compared. Comparing performance of MPCs using different models shows that MSPE reflects the suitability of predictive models better; compared to one-step ahead accuracy. It has also been shown that models with similar one-step ahead accuracy could lead to 24% difference in terms of activating the potential cost savings achieved by MPC. On the other hand, ANN-based NARX model yielded the highest electricity cost, which is due to its poor multi-step ahead prediction performance. Furthermore, MPC is compared to a well-tuned Rule Based Controller (RBC). Best performing MPC (using state space model with 7 states) yielded 22.7% decrease in energy cost compared to the RBC.

It would be interesting to compare these models for longer simulation time and on other case studies to see whether the findings are valid or not. Another suggestion for future work is to train models based on MSPE and then check the suitability of each model. The impact of state estimator in case of grey-box and state space models have not been addressed yet and combining the dynamics of the estimator and the model might yield a better KPI for comparing these models.

Acknowledgement

This work has emerged from IEA-EBC Annex 71 project, an international project aimed at assessing building's energy performance based on in-situ measurements.

References

Abu-Mostafa, Y. S. (1992). Neural networks and learning. In *Institute of Physics Conference Series* (Vol. 127).

Afram, A., & Janabi-Sharifi, F. (2014). Theory and applications of HVAC control systems - A review of model predictive control (MPC). *Building and Environment*, *72*, 343–355.

Afroz, Z., Shafiullah, G. M., Urmee, T., & Higgins, G. (2018). Modeling techniques used in building HVAC control systems: A review. *Renewable and Sustainable Energy Reviews*, *83* (October 2017), 64–84.

Atam, E., & Helsen, L. (2016). Control-Oriented Thermal Modeling of Multizone Buildings: Methods and Issues. *IEEE Control Systems*, *36*(3), 86–111.

Bacher, P., & Madsen, H. (2011). Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings*, *43*(7), 1511–1522.

Baetens, R., De Coninck, R., Jorissen, F., Picard, D., Helsen, L., & Saelens, D. (2015). Openideas - An open framework for integrated district energy simulations. *14th International Conference of IBPSA - Building Simulation 2015, BS 2015, Conference Proceedings*, 347–354.

Bourdeau, M., Zhai, X. qiang, Nefzaoui, E., Guo, X., & Chatellier, P. (2019). Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustainable Cities and Society*, *48*(April), 101533.

Da Silva, P. C., Leal, V., & Andersen, M. (2012). Influence of shading control patterns on the energy assessment of office spaces. *Energy and Buildings*, *50*, 35–48.

De Coninck, R., & Helsen, L. (2016). Practical implementation and evaluation of model predictive control for an office building in Brussels. *Energy and Buildings*, *111*, 290–298.

Djamel, Z., & Noureddine, Z. (2017). The Impact of Window Configuration on the Overall Building Energy Consumption under Specific Climate Conditions. *Energy Procedia*, *115*, 162–172.

Drgoňa, J., Arroyo, J., Cupeiro Figueroa, I., Blum, D., Arendt, K., Kim, D., Ollé, E. P., Oravec, J., Wetter, M., Vrabie, D. L., & Helsen, L. (2020). All you need to know about model predictive control for buildings. *Annual Reviews in Control*, *August*.

Drgoňa, J., Picard, D., Kvasnica, M., & Helsen, L. (2018). Approximate model predictive building control via machine learning. *Applied Energy*, *218*(March), 199–216.

Eicker, U., Demir, E., & Gürlich, D. (2015). Strategies for cost efficient refurbishment and solar energy integration in European Case Study buildings. *Energy and Buildings*, *102*, 237–249.

Erfani, A., Rajabi-Ghahnaviyeh, A., & Boroushaki, M. (2018). Design and construction of a non-linear model predictive controller for building's cooling system. *Building and Environment*, *133*(November 2017), 237–245.

EU Commission. (2018). *IN-DEPTH ANALYSIS IN SUPPORT OF THE COMMISSION COMMUNICATION* (Issue November).

Global Optimization Toolbox. (2021). https://nl.mathworks.com/help/gads/genetic-algorithm.html

Guerra-Santin, O., & Tweed, C. A. (2015). In-use monitoring of buildings: An overview of data collection methods. *Energy and Buildings*, *93*, 189–207. 2

Haddadi, M., Jafarinejad, T., & Badpar, F. (2019). Solar-hydrogen renewable supply system optimisation based on demand side management. *International Journal of Ambient Energy*, *0*(0), 1–10.

Jafarinejad, T., Erfani, A., Fathi, A., & Shafii, M. B. (2019). Bi-level energy-efficient occupancy profile optimization integrated with demand-driven control strategy: University building energy saving. *Sustainable Cities and Society*, *48*(November 2018).

Jafarinejad, T., Shafii, M. B., & Roshandel, R. (2019). Multistage recovering ventilated air heat through a heat recovery ventilator integrated with a condenser-side mixing box heat recovery system. *Journal of Building Engineering*, *24*(March), 100744.

Kumar, M., & Kar, I. N. (2009). Non-linear HVAC computations using least square support vector machines. *Energy Conversion and Management*, *50*(6), 1411–1418.

Maciejowski, J. M. (2002). *Predictive Control: With Constraints* (2nd ed.). Pearson Education.

Modelica Association Project "FMI." (2013). *Functional Mock-up Interface for Model Exchange and Co-Simulation*.

Picard, D., Drgoňa, J., Kvasnica, M., & Helsen, L. (2017). Impact of the controller model complexity on model predictive control performance for buildings. *Energy and Buildings*, *152*, 739–751.

Picard, D., Sourbron, M., Jorissen, F., Cigler, J., & Váňa, Z. (2016). Comparison of Model Predictive Control Performance Using Grey-Box and White-Box Controller Models of a Multi-zone Office Building. *4th International High Performance Buildings Conference*, 4.

Reynders, G., Diriken, J., & Saelens, D. (2014). Quality of grey-box models and identified parameters as function of the accuracy of input and observation signals. *Energy and Buildings*, *82*, 263–274.

Reynders, G., Diriken, J., & Saelens, D. (2017). Generic characterization method for energy flexibility: Applied to structural thermal storage in residential buildings. *Applied Energy*, *198*, 192–202.

Sangi, R., Kümpel, A., & Müller, D. (2019). Real-life implementation of a linear model predictive control in a building energy system. *Journal of Building Engineering*, *22*(January), 451–463.

Sepasgozar, S., Karimi, R., Farahzadi, L., Moezzi, F., Shirowzhan, S., Ebrahimzadeh, S. M., Hui, F., & Aye, L. (2020). A systematic content review of artificial intelligence and the internet of things applications in smart home. *Applied Sciences*

Sourbron, M., Verhelst, C., & Helsen, L. (2013). Building models for model predictive control of office buildings with concrete core activation. *Journal of Building Performance Simulation*, *6*(3), 175–198.

System Identification Toolbox. (2021). https://nl.mathworks.com/help/ident/index.html

Verhelst, C., Logist, F., Van Impe, J., & Helsen, L. (2012). Study of the optimal control problem formulation for modulating air-to-water heat pumps connected to a residential floor heating system. *Energy and Buildings*, *45*, 43–53.

Žáčeková, E., & Prívara, S. (2012). Control relevant identification and predictive control of a building. *Proceedings of the 2012 24th Chinese Control and Decision Conference, CCDC 2012*, 246–251.

Zhan, S., & Chong, A. (2021). Data requirements and performance evaluation of model predictive control in buildings : A modeling perspective. *Renewable and Sustainable Energy Reviews*, *January*, 110835.