Investigation of the Model Structure for Low-Order Grey-Box Modelling of Residential Buildings

Xingji Yu^{1*}, Laurent Georges¹, Michael Dahl Knudsen², Igor Sartori³, Lars Imsland⁴

¹NTNU, Dept. of Energy and Process Engineering, Trondheim, Norway

²AU, Dept. of Indoor Climate and Energy, Aarhus, Denmark

³SINTEF Building and Infrastructure AS, Oslo, Norway

⁴NTNU, Dept. of Engineering Cybernetics, Trondheim, Norway

* corresponding author: xingji.yu@ntnu.no

Abstract

Model Predictive Control (MPC) is a key technology to activate the building energy flexibility. A reliable control-based model should be developed for each specific building. The structure of grey-box models is usually based on the physical knowledge of the building. Firstly, it is not certain that this information will be available for all buildings, especially for small residential buildings. Secondly, developing a specific model structure for each building is most probably not affordable. Therefore, the paper investigates the dependency on the model structure to create reliable control-oriented model for the thermal mass of residential buildings. Using a test case, the performance of grey-box models based on the physical knowledge of the building are compared to grey-box models based on a generic structure taken from building standards (EN 13790 and VDI 6007) as well as black-box models where no knowledge of the building is required.

Introduction

Renewable energy plays an increasingly important role in economy growth and to limit CO₂ emissions. However, the increasing penetration of renewables in the grid poses a challenge for balancing the demand and supply of electricity. Electricity generated from renewables may not be consumed optimally and this mismatch can also cause challenges in the power system. Thus, multiple types of flexibility are needed in the energy sectors for smart grid integration (You, Jin, Hu, Zong, & Bindner, 2015). Buildings accounts for more than one-third of the total final energy consumption worldwide (*Transition to Sustainable Buildings*, 2013). A considerable portion of the energy consumed by buildings is used for heating, especially in Nordic countries where the space-heating season is long and cold. It has been shown that the thermal mass of buildings can be a significant heat storage (Le Dréau & Heiselberg, 2016; Glenn Reynders, 2015; Zong et al., 2017) to perform demand response (DR). Model predictive control (MPC) is often considered to be a key technology to activate the thermal energy flexibility. The control will take the predictions of future disturbances and system constraints into the optimization so that an optimal control decision could be made at each time step to perform DR. The typical disturbances taken into account are the ambient temperature, solar radiation and internal heat sources. The objective of the optimization using MPC is usually to minimize energy use, power, energy

costs or CO₂ emissions while subjected to thermal comfort constraints (Dahl Knudsen & Petersen, 2016).

An effective implementation of a MPC requires a specific control model of the thermal dynamics of the building. Methods are typically divided into white-, black- and grey-box models. White-box models are almost entirely based on physical laws. Therefore, they require detailed knowledge of the system to be modelled and its parameters (such as the geometry). It is often difficult and very time-consuming to obtain this information in practice. Some parameters may also change during operation and deviate from the original design. In addition, the mathematical complexity of white-box models makes them unsuited for MPC due to the computational cost to optimize a large non-linear system of equations in real time. Model reduction techniques can nonetheless be applied to white-box models. Black-box modelling is a data-driven method only considering system inputs and outputs. It can be applied even if a limited physical knowledge of the system is available. Since the data is the only information for the modelling process, the quality and amount of data will significantly influence the precision of black-box models. Their ability to predict the system dynamics outside operating conditions considered during the model training is also critical. Grey-box modelling is a combination of the previous two approaches. The physical knowledge of the system is used to determine a general model structure (a low-order model) and the model parameters are identified using experimental data. Due to the model structure based on physical grounds, grey-box models require less experimental data than black-box models and are less sensitive to the data quality. In addition, they should be more robust to extrapolate outside the operating conditions used during the period of the model training.

A single residential building will not provide a large amount of energy flexibility to the grid. In the context of smart grids, a large number of residential buildings needs to be considered. However, the thermal dynamics is different for each single building. Creating a suitable control-oriented model is also acknowledged as the most time-consuming part of the MPC implementation (Atam & Helsen, 2016), especially when physical knowledge specific to the building is required (such as in grey-box models). For grey-box models, the typical approach is to progressively increase the complexity of the model structure and to perform a forward selection process to identify the optimal configuration.

Some research work has been recently done to develop a tool that can automatically identify the grey-box models based on a BIM (Andriamamonjy, Klein, & Saelens, 2019). However, it is still worth investigating if generic model structures can be an acceptable option to lower the cost of modelling in MPC. For this purpose, three modelling approaches are compared on a test case. It is here assumed that model should be identified using indoor temperature measurements only. Firstly, a traditional grey-box modelling approach is used where models of increasing complexity are created based on the physical knowledge of the building (i.e. forward approach). Secondly, the model structures of building standards EN 13790 and VDI 6007 are considered. These models are low-order white-box models that are able to successfully predict space-heating needs for various building types. For instance, some studies have also been done to compare the simulation results of the two standard white-box models with TRNSYS simulation (Bruno, Pizzuti, & Arcuri, 2016; Vivian, Zarrella, Emmi, & De Carli, 2017). Their structures, even not optimal for a specific building, may nonetheless be a good candidate for a generic structure of grey-box models. Finally, these two methods to create grey-box models are compared to black-box models identified using a subspace method and then refined using numerical minimization of simulation errors. The long-term predictions (i.e. simulation without a disturbance model) of these models are evaluated as well as their estimates of some major building thermal characteristics (such as overall heat loss coefficient and thermal capacities).

Dataset and platform description

The ZEB Living Lab is a zero-emission single-family house located in the campus of the Norwegian University of Science and Technology (NTNU) in Trondheim. The total floor area of the building is about 100 m². The envelope is a wooden frame insulated with 35-40 cm mineral wool and with a glazing ratio of 0.2. Photovoltaic panels installed on the roof has been designed to provide enough onsite renewable energy production to reach a zero CO_{2eq} emission balance over the building lifetime. The water-based heating system consists of a ground source heat pump. The space-heating can be either performed by floor heating, a central radiator or the ventilation air. However, the current study is based on measurement data from a previous experiment where space heating was performed using an electrical heater (P. Vogler-Finck, Clauß, & Georges, 2017; P. J. C. Vogler-Finck, Clauß, Georges, Sartori, & Wisniewski, 2018). The electric heater was located in the middle of the building while a Pseudo-Random Binary Signal (PRBS) was used to excite the thermal dynamic of the building in a large spectrum of frequencies. The floor plan of the building is shown in Figure 1 along with the location of the electric radiator and the temperature sensors.

The dataset contains three successful experiments, which are named experiment 2, 3 and 4 (experiment 1 being omitted). The data includes measurements every 5 minutes of the ambient temperature, indoor air operative

temperatures as well as the global solar radiation on a horizontal plane and the electricity consumption. In this study, the ventilation losses are not identified but directly introduced in the model as heat losses. These losses have been evaluated using the measured temperature difference between the supply and exhaust ventilation air combined with the measured airflow rate (constant during experiments). This research focuses on data from experiments 2 and 4 since both experiments were conducted with the building unoccupied and internal doors opened. This leads to an almost uniform spatial distribution of the air temperature inside the building. However, there is some temperature stratification and all air temperature sensors are therefore volume-averaged to represent the mean indoor air temperature T_i . Experiment 2 is used to train the model while experiment 4 is used for validation.

Figure 1: Picture and Floor plan of the ZEB Living Lab (P. J. C. Vogler-Finck et al., 2018)

Grey-box models

As mentioned in the introduction section, grey-box modelling is a combination of measurement data and physical knowledge. The thermal dynamics of the ZEB Living Lab is assumed to be linear and time invariant which is a common approximation for building envelopes. It can then be approximated by low-order resistance-capacitance (RC) networks. The model order is defined by the number of heat capacitances included in the model. This paper considers two categories of structure for the grey-box model: seven structures derived from our prior knowledge of the Living Lab and two

generic structures derived from standards for energy calculations (EN 13790) and (VDI 6007).

Knowledge-based models

The simplest model structure is a 1st order model. The other models are created by progressively adding more components, the most complex structure being a 3rd order model. Only the most complicated 3rd order model is presented in Figure 2 since the other models are simplifications of this structure as described in Table 1.

Figure 2: Most complex model structure (3C).

The 1st order model has one heat capacitance (C_i), the 2nd order model has two heat capacitances (C_i and C_e) while the 3rd order model has three heat capacitances (C_i , C_e and C_h). The physical meaning of these components is described here below.

Table 1: Structure of the different knowledge-based grey-box models.

MODEL	1C	2Cs	2C	3Cs	3C	3Csd	3Cd
R_{ih}	x	x	x	✓	✓	✓	✓
R_{ie}	x	✓	✓	✓	✓	✓	✓
R_{ea}	x	✓	✓	✓	✓	✓	✓
R_{ia}	✓	x	✓	x	✓	x	✓
C_e	x	✓	✓	✓	✓	✓	✓
C_i	✓	✓	✓	✓	✓	✓	✓
C_h	x	x	x	✓	✓	✓	✓
A_{wi}	✓	✓	✓	✓	✓	✓	✓
A_{we}	x	✓	✓	✓	✓	x	x

- T_i Temperature of interior heat capacity [°C].
- T_e Temperature of the building envelope [°C].
- T_h Temperature of the electric heater [°C].
- T_a The outdoor/ambient temperature [°C].
- C_i Heat capacity of the interior which is assumed to be the combination of the thermal mass of the air, the furniture, internal walls and the first centimeters of the internal surface of external walls [kWh/K].
- C_e Heat capacity of the building envelope, (external walls and windows) [kWh/K].
- C_h Heat capacity of the heater which is assumed to the combination of its thermal mass and some air around the heater [kWh/K].
- R_{ie} ($1/UA_{ie}$) Heat resistance between the building envelope and the interior of the building [K/kW].

- R_{ea} ($1/UA_{ea}$) Heat resistance between the ambient and the building envelope [K/kW].
- R_{ia} ($1/UA_{ia}$) Heat resistance between the ambient and the interior of the building [K/kW].
- P_h Heat gain from the electric heater [kW].
- P_s Global solar irradiation on a horizontal plane [kW].
- Q_{app} Heat gain from internal loads (appliances) [kW].
- Q_{vent} Heat gains from the ventilation [kW].
- A_{wi} Effective window area for the solar gain that enters directly the interior node [m²].
- A_{we} Effective wall area for the solar gain directly applied to the envelope of the building [m²].

Standard models

The structure of the two RC-models is taken from the standards EN 13790 and VDI 6007 (Vivian et al., 2017). The EN 13790 model has originally five resistances and one capacitance, as shown in Figure 3. The heat capacitance represents the heat capacity of the building envelope. Since the ventilation heat loss is directly injected as a heat gain in our model, the specific resistance of EN 13790 related to ventilation is removed. The resulting grey-box model has therefore four resistances. Detailed physical explanation of the RC components in EN 13790 are described below.

Figure 3: Model structure of EN 13790.

- T_i Interior temperature (as previously defined) [°C].
- T_e Temperature of the building envelope [°C].
- T_s The temperature of the internal surface of the building envelope [°C].
- T_a The outdoor/ambient temperature [°C].
- C_e Heat capacity of the building envelope [kWh/K].
- R_{es} ($1/UA_{es}$) Heat resistance between the building envelope and the internal surface of the building envelope [K/kW].
- R_{ea} ($1/UA_{ea}$) Heat resistance between the ambient and the building envelope [K/kW].
- R_{as} ($1/UA_{as}$) Heat resistance between the ambient and the internal surface of the building envelope [K/kW].

- R_{is} ($1/UA_{is}$) Heat resistance between the internal surface of the building envelope and the internal node [K/kW].
- P_h Heat gain from the electric heater [kW].
- P_s Global solar irradiation on a horizontal plane [kW].
- Q_{app} Heat gain from internal loads (appliances) [kW].
- Q_{vent} Heat gains from the ventilation [kW].
- A_w Effective window area for the solar gain [m²].

According to EN 13790, half of the internal gains Q_{app} is directly entering the internal node (T_i) while the heat emitted by the radiator is fully entering this node. Thus, a coefficient of 1/2 is applied to Q_{app} and a coefficient of 1 to the space-heating power P_h . In EN 13790, the fraction of internal and solar gains entering the internal surface (T_s) and the envelope nodes (T_e) should be evaluated using the detailed geometry of the building. Therefore, one model parameter (a) is added and should be identified. In order to guarantee energy conservation, the sum of the two fractions of Q_{app} applied respectively to nodes T_s and T_e is constrained to 1/2. Solar radiation only enters at nodes T_s and T_e , so the sum of the two fractions of solar gains at T_s and T_e is constrained to 1.

The VDI 6007 model originally has seven resistances and two capacitances. Like EN 13790, the heat resistance related to the ventilation heat losses is substituted by the measured ventilation heat gains. Detailed physical explanations of the RC components of VDI 6007 are shown below.

Figure 4: Model structure of VDI 6007.

- T_i Interior air and furniture temperature [°C].
- T_v Interior “star” node [°C].
- T_{s1} The internal surface temperature of building envelope (meaning external walls and windows) [°C].
- T_{s2} The internal surface temperature of internal walls [°C].
- T_{e1} Temperature of the building envelope [°C].
- T_{e2} Temperature of internal walls [°C].
- T_a The outdoor/ambient temperature [°C].
- C_{e1} Heat capacitance of the building envelope [kWh/K].
- C_{e2} Heat capacitance of internal walls [kWh/K].

- R_{es1} ($1/UA_{es1}$) Resistance between the building envelope and the internal surface of the envelope [K/kW].
- R_{es2} ($1/UA_{es2}$) Resistance between the internal walls and their surface [K/kW].
- R_{is1} ($1/UA_{is1}$) Heat resistance between the internal surface of the envelope and the star node [K/kW].
- R_{is2} ($1/UA_{ei1}$) Heat resistance between the ambient and the building envelope [K/kW].
- R_{ea1} ($1/UA_{ei1}$) Heat resistance between the ambient and the building envelope [K/kW].
- R_{iv} ($1/UA_{ei1}$) Heat resistance between the star node and the indoor air and furniture node [K/kW].
- P_h Emitted heat from the electric heater [kW].
- P_s Global solar irradiation on a horizontal plane [kW].
- Q_{app} Heat gain from internal loads (appliances) [kW].
- Q_{vent} Heat gains from the ventilation [kW].
- A_w Effective window area for the solar gain [m²].

The coefficient for internal heat gains (Q_{app}) and the emitted power by the radiator (P_h) are estimated by the grey-box modelling approach. It results in six parameters a_1 , a_2 , b_1 , b_2 , d and e . To guarantee the conservation of energy, additional constraints for these variables are applied: the sum of the coefficients of a_1 , a_2 and d and the sum of coefficients of b_1 , b_2 and e are set to 1.

Black-box models

In black-box modelling, data is used to train mathematical models with parameters that cannot be given an immediate physical interpretation. In this study linear time-invariant state-space models with model orders from one to three are examined:

$$x^{k+1} = \mathbf{A} x^k + \mathbf{B} u^k \quad (1)$$

$$y^k = \mathbf{C} x^k \quad (2)$$

where x is the state vector and $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are system matrices. u is the input (outdoor temperature, solar radiation and heat gains) and y is the output (indoor temperature). Note that no disturbance term $\mathbf{K}$ (Kalman gain) is included because we focus on long-term predictions. All coefficients in the system matrices are free and unconstrained which gives the optimizer maximum freedom (for a given model order) to adjust the coefficients to fit to the training data. The downfall is that the model cannot be guaranteed to always comply with physical laws (i.e. conservation of energy). In case of poor training data, the model might have a very poor performance on new data.

This study has trained the black-box models in two steps. Firstly, subspace system identification estimates an initial model using the *n4sid* function in MATLAB (Ljung, 2011). Secondly, the model is refined through numerical simulation-error minimization using the *pem* function of MATLAB.

It should be noted that, even though the black-box parameters do not have an immediate physical meaning, it is in fact possible to extract some insight on the thermal characteristics of the building. For instance, the overall heat loss coefficient can be estimated as the inverse of the steady-state output (i.e. the room temperature) for a step response of the heat input. This value is also shown in Tables 2 and 3 for the 2nd order black box model (2B).

Results

Tables 2 and 3 show some of the most important physical parameters identified and the simulation results for all models, respectively for the original sampling time of 5 minutes and a data subset based on 15-minutes averaging. Since this paper mainly focuses on the long-term stable prediction of the model, no disturbance model is included to correct the current state based on prediction errors. Other studies on grey-box modelling often focus on the one-step ahead predictions (Bacher & Madsen, 2011), but for MPC implementation it is important to have good performance for a longer prediction horizon. Since there is no human occupancy during experiments and since energy consumption and the indoor temperature are accurately measured, the data quality is considered to be relatively high compared to what can be expected in real building operation: the signal-to-noise ratio (SNR) is expected to be high. The model performance is evaluated using the NRMSE (fitting), RMSE and the range of magnitude taken by the physical parameters (which should be related to the building physics to a reasonable degree). AIC (Akaike information criterion) and BIC (Bayesian information criterion) are used as complementary performance criteria to judge if the model is over-fitting.

Knowledge-based models are first analysed. The fitting of the 1st order model is relatively low: the NRMSE fitting is 58% for the training dataset while the fitting decreases to only 23% for the validation dataset. This implies that only one state is not enough to describe the thermal dynamics of this building. The 2nd order model has a new state related to the building envelope which can significantly improve the model prediction performance compared to the 1st order model. This is confirmed by the results reported in Table 2. The fitting of the 2nd order model 2Cs increases significantly to 81% for the training dataset and 79% for the validation dataset, RMSE values confirms this trend with decreasing values (from 0.5502 to 0.2489 in the training dataset). The 2nd order model 2C has one additional resistance (R_{ia}) compared to the model 2Cs. R_{ia} represents the heat losses from infiltrations and heat conduction through components of the building envelope with negligible thermal mass (such as windows or external doors of the building). However, results show that the value of $1/R_{ia}$ is zero so that the model collapses into the 2Cs models (with parameters being exactly the same). The AIC and BIC values of 2C are slightly higher than the 2Cs. The ZEB Living Lab is a super-insulated building envelope with limited infiltrations. It was therefore anticipated that the contribution of infiltrations should be negligible. However, the building has a large

amount of glazing. An infinite R_{ia} was therefore unexpected based on prior physical knowledge. In conclusion, for this test case, the resistance R_{ia} is not necessary for the 2nd order model. Based on the simulation performance and the value of the parameters, the model 2Cs could be a reliable control model for the ZEB Living Lab. Comparing the 2Cs model for 5- and 15-minutes time intervals, it shows very similar physical parameters and simulation performance. The 2nd order black-box model 2B has an 83% fit on training data and 88% on validation data. It is thus the best performing 2nd order model.

In order to investigate whether the 2nd order model could be further improved, a new state corresponding to the heater is introduced in the grey-box model. The physical reason for adding this additional state is to compensate for the potential time delay related to the thermal dynamics of the electric heater. Four 3rd order models are tested in this paper. Model 3Cs shows a good fitting of 81% for the training dataset (with a RMSE value of 0.248) while the AIC and BIC values are lower than for the 2Cs model. However, the heat capacity of the interior (C_i) is extremely low while the heat capacity of the heater (C_h) is estimated to be 1.496 kWh/K, which is too high for an electric radiator. As the value of these parameters has limited physical meaning, the model 3Cs is discarded. The fitting of model 3C is low compared to the others. Its AIC and BIC values are also much higher so that this model is also discarded. As for 3Cs, model 3Csd has a relatively good fitting of 79% on the training dataset but the values for the parameters related to the building envelope and heater are far from reasonable. This model is also rejected. Model 3Cd has a fit of 81% and an RMSE value of 0.253 on the training data. However, the fit for the validation dataset is 70%, which is not as good as the 2nd order model 2Cs. The parameter values of model 3Cd are within a reasonable range. However, the value of the AIC and BIC criteria increases noticeably. Therefore, the model 3Cd could also be kept for further investigations. Nevertheless, when analyzing results for the 15-minutes time interval, the values of the heat capacitances completely change. The parameters of model 3Cd vary significantly with the time step which can be considered as a lack of reliability. The overall conclusion regarding 3rd order models is that none of them are completely satisfactory. Two possible reasons are that the heat capacitance (C_h) was probably not required for the electric radiator and the increasing the number of parameters leads the model to be over-fitting. The 3rd order black-box model has a fit of 88% on the validation data which is equivalent to the 2nd order model (2B). The 2B model is therefore the appropriate order for a linear model of this building based on the current dataset.

Secondly, the performance of the two standard models is investigated. The 1st order EN 13790 model has much better fit compared to the knowledge-based 1st order model: the NRMSE fitting can reach 71% for the training dataset but the fitting drops to 37% for the validation dataset. Accordingly, the RMSE value increases drastically to 1.509 for the validation dataset. This

confirms the previous conclusion that one state is not enough to describe the thermal dynamics of this building. From the Figure 5, it is very clear that the EN 13790 model can predict the general trend of the indoor air temperature. However, there are obvious excessive fluctuations in the indoor temperature prediction. This can be easily explained. In the EN 13790 model, the heat emitted by the radiator (P_h) is directly injected in the node T_i while no capacitance is allocated to this node. Therefore, the indoor air will immediately respond to the heat injection of the electric heater (P_h), without delay. It again proves that only one capacitance is not enough for capturing all the dynamics of the building. The conclusion regarding EN 13790 is that it could be a good candidate for a generic model but it needs to be adapted to capture the faster thermal dynamics of the air, furniture and the first centimeters of the walls. For instance, an extension of the EN13790 model to a 5R2C model has been recently proposed (Hedegaard, Kristensen, Pedersen, Brun, & Petersen, 2019) showing much better prediction performance compared to original EN13790.

Figure 5: Comparison of the simulation performance of the most representative models

The 2nd order model VDI 6007 has decent fitting results (i.e. 71% for training and 81% for validation). Nevertheless, the values of the two heat capacitances and the overall UA-values are extremely small. These parameters do not have physical meaning. In conclusion, the VDI 6007 structure contains many physical phenomena but is too complicated to be correctly identified only using indoor air temperature measurements. This is confirmed by the high AIC and BIC values evaluated for this model. To be a potential candidate for a generic model structure, the number of parameters needs to be decreased to increase the identifiability of the model (Hedegaard & Petersen, 2017). Alternatively, additional measurements could be made inside the building in order to make the model identifiable (such as wall surface temperature and/or heat flux) (G. Reynders, Diriken, & Saelens, 2014). However, it is here assumed that these measurements will not be made available for all existing residential buildings.

Conclusion and future work

This paper investigates the performance of grey-box models based on a physical knowledge of the building to grey-box models using a generic model structure (based on the building standards EN 13790 and VDI 6007) as well as black-box models. Their performance is evaluated on the long-term prediction of the thermal dynamics, here using a single-family house as a test case (the ZEB Living Lab). The model identification is based on the measurement of the indoor air temperature resulting from the space heating using an electric radiator controlled by a PRBS. The building was unoccupied during experiments.

For knowledge-based grey-box models, results show that the 2nd order model has reasonable parameter estimates and the prediction error is small (within the range of $\pm 1^\circ\text{C}$). The prediction performance varies significantly between the investigated 3rd order models. However, in all cases, the estimated parameters do not have reasonable values. In addition, none of the 3rd order models investigated were able to significantly improve the prediction performance of the 2nd order model. Therefore, the 2nd order model is considered as a good candidate for this test case.

Regarding grey-box models based on standards, results show that the 1st order EN 13790 model is able to follow the general evolution of the indoor temperature and provides meaningful values of the parameters. However, the simulated temperature has significantly higher fluctuations directly corresponding to the start and stop cycles of the electric radiator. As recently proposed by Hedegaard et al., the EN 13790 is a potential candidate for a generic model structure but it should be adapted by adding a state corresponding to the fast dynamics of the building ($< 1\text{h}$). The 2nd order grey-box model based on VDI 6007 has good prediction performance, but it generates parameter estimates that cannot be explained from a physical point of view. The number of parameters of this model needs to be reduced to make it identifiable. In future work, it should be investigated whether this

simplification can be done without impairing significantly the model universality.

The 2nd order black-box model shows a good performance equivalent to the 2nd order grey-box model. Nevertheless, with black-box models, the physical meaning of the states is unknown. However, the estimate of the overall heat transfer coefficient is similar between the 2nd order black-box and the best grey-box models. It is worth mentioning that these investigations were performed with high-quality input-output data. In addition, experiments corresponding to the validation data set took place a few days after the training period. The relative performance of black-box and grey-box models could be different if these experimental conditions were not fulfilled. This research nonetheless suggests that black-box models deserve to be investigated in detail to create control-oriented model with a limited knowledge of the building and limited amount of time.

Acknowledgement

The authors would like to acknowledge the valuable input from Pierre Vogler-Finck and John Clauß. They also would like to thank the Norwegian Research Centre on Zero Emission Neighbourhoods in Smart Cities (ZEN) and its industry partners.

References

- Andriamamonjy, A., Klein, R., & Saelens, D. (2019). Energy & Buildings Automated grey box model implementation using BIM and Modelica. *Energy & Buildings*, 188–189, 209–225. <https://doi.org/10.1016/j.enbuild.2019.01.046>
- Atam, E., & Helsen, L. (2016). Control-Oriented Thermal Modeling of Multizone Buildings: Methods and Issues: Intelligent Control of a Building System. *IEEE Control Systems*, 36(3), 86–111. <https://doi.org/10.1109/MCS.2016.2535913>
- Bacher, P., & Madsen, H. (2011). Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings*, 43(7), 1511–1522. <https://doi.org/10.1016/j.enbuild.2011.02.005>
- Bruno, R., Pizzuti, G., & Arcuri, N. (2016). The prediction of thermal loads in building by means of the EN ISO 13790 dynamic model: A comparison with TRNSYS. *Energy Procedia*, 101, 192–199.
- Dahl Knudsen, M., & Petersen, S. (2016). Demand response potential of model predictive control of space heating based on price and carbon dioxide intensity signals. *Energy and Buildings*, 125, 196–204. <https://doi.org/10.1016/j.enbuild.2016.04.053>
- Hedegaard, R. E., Kristensen, M. H., Pedersen, T. H., Brun, A., & Petersen, S. (2019). Bottom-up modelling methodology for urban-scale analysis of residential space heating demand response. *Applied Energy*, 242, 181–204.
- Hedegaard, R. E., & Petersen, S. (2017). ScienceDirect ScienceDirect Evaluation of Grey-Box Model Parameter Estimates Intended for The 15th International Symposium on District Heating and Cooling Thermal Characterization of Buildings Assessing the feasibility of using the heat temperature function for a long-term district heat demand forecast. *Energy Procedia*, 132, 982–987. <https://doi.org/10.1016/j.egypro.2017.09.692>
- Le Dréau, J., & Heiselberg, P. (2016). Energy flexibility of residential buildings using short term heat storage in the thermal mass. *Energy*, 111, 991–1002. <https://doi.org/10.1016/j.energy.2016.05.076>
- Ljung, L. (2011). System identification toolbox. *The Matlab User's Guide*, 1, 237.
- Reynders, G. (2015). Quantifying the impact of building design on the potential of structural storage for active demand response in residential buildings, (September), 266. <https://doi.org/10.13140/RG.2.1.3630.2805>
- 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. <https://doi.org/10.1016/j.enbuild.2014.07.025>
- Transition to Sustainable Buildings. (2013). <https://doi.org/10.1787/9789264202955-en>
- Vivian, J., Zarrella, A., Emmi, G., & De Carli, M. (2017). An evaluation of the suitability of lumped-capacitance models in calculating energy needs and thermal behaviour of buildings. *Energy and Buildings*, 150, 447–465. <https://doi.org/10.1016/j.enbuild.2017.06.021>
- Vogler-Finck, P., Clauß, J., & Georges, L. (2017, October). A dataset to support dynamical modelling of the thermal dynamics of a super-insulated building. <https://doi.org/10.5281/ZENODO.1034820>
- Vogler-Finck, P. J. C., Clauß, J., Georges, L., Sartori, I., & Wisniewski, R. (2018). Inverse Model Identification of the Thermal Dynamics of a Norwegian Zero Emission House. In *Cold Climate HVAC Conference* (pp. 533–543). Springer.
- You, S., Jin, L., Hu, J., Zong, Y., & Bindner, H. W. (2015). The Danish perspective of energy internet: From service-oriented flexibility trading to integrated design, planning and operation of multiple cross-sectoral energy systems. *Zhongguo Dianji Gongcheng Xuebao*, 35(14), 3470–3481.
- Zong, Y., Böning, G. M., Santos, R. M., You, S., Hu, J., & Han, X. (2017). Challenges of implementing economic model predictive control strategy for buildings interacting with smart energy systems. *Applied Thermal Engineering*, 114, 1476–1486. <https://doi.org/10.1016/j.applthermaleng.2016.11.141>

Table 2: Results of the grey- and black-box identification using 5 minutes time interval.

Model	1C	2Cs	2C	3Cs	3C	3Csd	3Cd	EN13790	VDI6007	2B
C_e	-	8.032	8.032	8.224	0.303	10.000	7.638	8.451	-	-
C_i	7.115	1.366	1.366	$1.24 \cdot 10^{-7}$	$1.48 \cdot 10^{-7}$	1.781	2.158	-	-	-
C_h	-	-	-	1.496	6.445	4.822	0.028	-	-	-
C_{el}	-	-	-	-	-	-	-	-	$1.77 \cdot 10^{-5}$	-
C_{e2}	-	-	-	-	-	-	-	-	$3.49 \cdot 10^{-5}$	-
A_{we}	-	3.158	3.158	3.185	3.815	11.806	-	-	1.928	-
A_{wi}	8.375	4.163	4.163	4.071	5.955	0.186	4.923	9.152	1.656	-
UA_{tot}	0.125	0.115	0.115	0.114	0.117	0.186	0.102	0.125	$5.17 \cdot 10^{-7}$	0.114
RMSE _t	0.55	0.25	0.25	0.25	0.40	0.27	0.25	0.36	0.28	0.23
RMSE _v	1.84	0.49	0.49	0.40	1.33	1.07	0.70	1.51	0.47	0.29
NRMSE _t	58%	81%	81%	81%	69%	79%	81%	73%	79%	83%
NRMSE _v	23%	79%	79%	83%	44%	55%	71%	37%	81%	88%
AIC	$2.847 \cdot 10^3$	113.4166	115.4166	102.8692	$1.775 \cdot 10^3$	450.2877	177.9489	$1.371 \cdot 10^3$	522.9770	-
BIC	$2.869 \cdot 10^3$	157.0543	164.5091	162.8711	$1.841 \cdot 10^3$	504.8349	237.9508	$1.414 \cdot 10^3$	599.3430	-

t: training data

v: validation data

Table 3: Results of the grey- and black-box identification using 15 minutes time interval.

Model	1C	2Cs	2C	3Cs	3C	3Csd	3Cd	EN13790	VDI6007	2B
C_e	-	8.038	8.038	8.154	3.481	10.000	7.802	8.339	-	-
C_i	7.122	1.346	1.346	$2.87 \cdot 10^{-5}$	$4.88 \cdot 10^{-8}$	1.764	$2.04 \cdot 10^{-7}$	-	-	-
C_h	-	-	-	1.5978	5.1958	4.8375	2.2863	-	-	-
C_{el}	-	-	-	-	-	-	-	-	$1.55 \cdot 10^{-5}$	-
C_{e2}	-	-	-	-	-	-	-	-	$2.38 \cdot 10^{-5}$	-
A_{we}	-	3.213	3.213	2.776	14.380	-	-	-	2.042	-
A_{wi}	8.394	4.152	4.152	4.197	1.617	11.780	4.893	8.872	1.511	-
UA_{tot}	0.125	0.115	0.115	0.113	0.108	0.186	0.102	0.124	$3.99 \cdot 10^{-7}$	0.114
RMSE _t	0.55	0.25	0.25	0.24	0.37	0.27	0.25	0.39	0.28	0.23
RMSE _v	1.84	0.51	0.51	0.31	0.79	1.06	0.73	1.45	0.43	0.29
NRMSE _t	58%	81%	81%	81%	72%	79%	81%	70%	79%	83%
NRMSE _v	23%	79%	79%	87%	67%	56%	70%	39%	82%	88%
AIC	953.9709	51.3763	53.3763	34.8207	519.5818	163.0977	41.9178	598.7925	201.7279	-
BIC	971.3953	86.2252	92.5813	82.7379	571.8551	206.6588	93.8350	633.6413	262.7134	-

t: training data

v: validation data