



Comparison of Data-Driven Thermal Building Models for Model Predictive Control

Gernot Steindl^{*1}, Wolfgang Kastner², Verena Stangl³

¹Institute of Computer Engineering, Treitlstrasse 3, A-1040 Wien, Vienna, Austria
e-mail: gernot.steindl@tuwien.ac.at

²Institute of Computer Engineering, Treitlstrasse 3, A-1040 Wien, Vienna, Austria
e-mail: wolfgang.kastner@tuwien.ac.at

³University of Applied Sciences Burgenland, Steinamangerstraße 21, A-7423 Pinkafeld, Austria
e-mail: 1510267021@fh-burgenland.at

Cite as: Steindl, G., Kastner, W., Stangl, V., Comparison of Data-Driven Thermal Building Models for Model Predictive Control, *J. sustain. dev. energy water environ. syst.*, 7(4), pp 730-742, 2019, DOI: <https://doi.org/10.13044/j.sdewes.d7.0286>

ABSTRACT

Energy flexible buildings in combination with demand response will play a key role in the future smart grid. To implement control strategies, which enable demand response, like model predictive control, thermal building models are necessary. Therefore, three lumped capacitance models, are compared with a k-Nearest Neighbor regression model. All models show accurate prediction results, if the operating condition of the building is similar during parameter identification or rather during training and the validation period. Parameter identification of lumped capacitance models is a time-consuming task. Especially for complex lumped capacitance models, the search space for certain parameters has to be reduced to avoid local minima. The investigated k-Nearest Neighbor algorithm has the advantage of easy implementation, very fast training and minimal effort for parameter identification in combination with accurate predictions. But its seasonal dependency is very strong, which can be easily overcome with periodically data update, as it is an instance-based learning algorithm.

KEYWORDS

Data-driven, Black-box model, Gray-box model, Model development, Machine learning.

INTRODUCTION

As a result of the increasing share of Renewable Energy Sources (RES), accompanied with a volatile production, today's energy system is facing serious challenges. To further increase the share of RES in the smart grid, flexible loads have to be utilized. Therefore, demand becomes an active player within the electrical energy system [1], and Demand Response (DR) is considered as one of the main components in the smart grid [2]. DR can be described as a change of the normal consumption pattern by the end-use customer in reaction to some signal [3].

As European households have a share of about 25% on the final energy consumption, and more than 65% of this energy is used for heating and cooling [4], these applications seem to be first candidates for DR actions. They are suitable because of their thermal

^{*} Corresponding author

inertia, which allows buildings to shift heating and cooling demand to periods in time where energy production from RES is available. The objective of residential demand response programs is, therefore, the reduction of peak load or rather carbon dioxide (CO₂) and cost reduction [5].

The DR potential of buildings depends on various factors like the type of heating, cooling, or air ventilation systems, structure and layout of the building, and the user behavior. It is limited by thermal comfort criteria and air quality constraints which must not be violated. A number of case studies have been carried out to investigate potentials and building performance implications [6].

Model Predictive Control (MPC) fits the requirements for implementing residential DR perfectly as it enables load curtailment strategies as well as balance conflicts between user comfort, energy consumption and costs [7]. The energy flexible building can actively support the electrical grid by reacting on DR signals. MPC is one of the most suitable control strategies to integrate such energy flexible buildings into the smart grid. Moreover, such control strategies become more important in perspective of the amendment to the European Directive on the Energy Performance of Buildings (EPBD) which was approved in May 2018 by the council of the European Parliament [8]. The main objective of this proposal is to improve the energy efficiency of existing buildings. Therefore, strategies and methods are mentioned in the proposal, like using built-in intelligent devices and smart technologies which will guarantee an efficient building operation. Information and Communication Technologies (ICT) should be implemented to interact with the user behavior and provide DR management.

DR in combination with MPC could be used to satisfy such requirements [9], but precise and robust thermal building models are crucial for MPCs to operate as expected [10]. In this context, MPC have been applied for price-based DR programs to optimize costs [11], reacting to fast DR events [12], as well as using other performance indices, like flexibility indicators, in their objective function [13]. The used building models can be divided into four categories: white-box, calibrated simulation models, gray-box and black-box models [14].

White-box models need detail knowledge about the system and the underlying physical process. Thus, they are usually too complex to be used in an MPC.

Calibrated simulation models are basically white-box models, but they were calibrated with measured data. They are often used as a surrogate model during development of new control strategies.

Gray-box models are data-driven models, but they are based on physical principles. Their parameters have to be identified, based on historical data. A common way to create Linear, Time Invariant (LTI) gray-box models, is applying an RC-formalism, also known as lumped capacitance model (RC models). The lumped capacitance model is a common technique to model the thermal condition of a building in analogy to an electrical circuit [15]. RC models have also been extended to incorporate humidity into the model, to increase comfort inside the building [16]. Creating RC models is a time consuming and error prone task. To reduce development costs, information from the Building Information Model (BIM) has been utilized to setup a tool chain, which is able to automatically generate gray-box models [17]. BIM data is mostly only available for new construction, because creating a BIM is also very time-consuming process, which is usually done at the planning phase of the construction.

Black-box models are purely data-driven. Thus, knowledge about the system is not necessarily needed. The model is based on historical data of the behavior of the system. Various machine learning algorithms, like k-Nearest Neighbor (k-NN), can be applied to create such regression models. A comprehensive review of data-driven modeling techniques for building energy consumption is given in Bourdeau *et al.* [18].

The objective of our work was to investigate the performance of three different RC models with different complexity in comparison with a k-NN model. The seasonal

dependency of the identification or rather training periods is investigated, to examine the usability of the identified RC models as well as the k-NN at different operating conditions (heating and cooling). Therefore, the seasonal model performances were evaluated.

MATERIALS AND METHODS

Measurements were carried out at a detached house in Austria. This single-family house was built in 2015 and has a living area of about 296 m². The house has a basement with a heated floor area of about 63 m². The ground floor (112.3 m²) and the first floor (85.5 m²) are totally conditioned. The external walls are made of brick, insulation and plasterboard. Load-bearing walls in the first floor are capped through a pitched roof and in the ground floor through a flat roof. The room used for this simulation is located in the ground floor and has a total volume of 191.7 m³. This living room consists of a high window area of 28.9 m² and has ceilings connected to external and adjacent thermal zones. The heat demand is covered by a floor heating system with a geothermal heat pump, operating in a monovalent mode. The data acquisition of the individual room temperatures, the outdoor ambient temperature, the solar radiation as well as the mass flow, supply and return temperature of the floor heating system was performed with calibrated mobile measurement equipment. The accuracy of the temperature sensors was in a range of ± 0.2 °C. A portable ultrasonic mass flow meter was installed to measure the supply flow of the heating system with an accuracy of ± 0.01 m/s. The solar irradiation was measured by a pyranometer, classified as so called “First Class” after ISO-9060. The sensor data were collected by a central aggregator over wireless network. Two temporary measurements were carried for two months in summer (July and August 2017) and four months in winter and spring (January till May), respectively.

This data was used to calibrate a building simulation model in TRNSYS. The geothermal heat pump system is realized through the single-stage heat pump type 927 combined with manual equations and a two-point controller. The minimum air change rate for a living area is defined by the national standard with 0.5 h^{-1} [19]. As only temporary presence is considered, the ventilation and natural infiltration are taken with a constant air change rate of 0.33 h^{-1} in the TRNSYS model. Internal gains produced by artificial lighting, persons or electrical equipment are neglected. Also, automatic shading is turned off for the simulation. The simulation data surrogate the measured data to focus on the thermal behavior of the building itself. This was done to eliminate disturbances on the model identification through user behavior or unknown control actions. Data for one thermal zone of this TRNSYS simulation was used for further analysis. With the data, three different RC models were parameterized and the k-NN model was trained. The whole model development process is explained in the following subsection.

Data-driven model development process

Figure 1 shows the main steps in the data-driven model development process which is inspired by Ge [20] and was applied in this work. The first step was a process decomposition. In the case of a thermal building model, the heat transfer process was analyzed and influencing parameters, like adjacent zone conditions, heat distribution system, window areas, infiltration as well as orientation of the room, were identified based on domain knowledge and construction documents of the building. Afterwards, the available measurements were assigned to their thermal zones.

Usually, the available raw data have to be pre-processed. This means, outliers have to be identified and removed, missing data have to be filled or estimated and the sampling rates have to be synchronized. This is usually a manual and time-consuming task. As a surrogated model was used for the investigation, this task could be skipped.

The next step is a sample and variable selection. The variable selection was performed based on domain knowledge and information from the construction

documents of the building. With the help of statistical analysis, like scatter plots and cross-correlation coefficients, a better insight into the building dynamics could be achieved by determining the linear dependency of the ambient air temperature and solar irradiance on the operational zone temperature. To take the different seasonal operating conditions of the building into account, a training sample of 18 winter days was selected. To validate the performance of the trained models, seven days in every of the four seasons were used.

As already mentioned, three RC gray-box models as well as a k-NN model were selected and trained. Training means an optimization process, where the error between the model output and the training data is minimized. For the gray-box models this means a variation of the given model parameters. A detailed description of the models is given in the next subsection.

The trained models were deployed and forecasts for the validation period were produced. The errors of the forecasts were used for model evaluation. Usually, if a model does not perform well, a new model or additional samples are used to perform a new training step. This iterative cycle was not performed in this work, because the purpose is to compare different models under the same training conditions.

The described model development process requires a lot of domain knowledge which is often not explicitly stated or scattered over domain experts. To reduce the effort for data driven model development and facilitate the re-use of knowledge, a systematic data and information management would be beneficial.

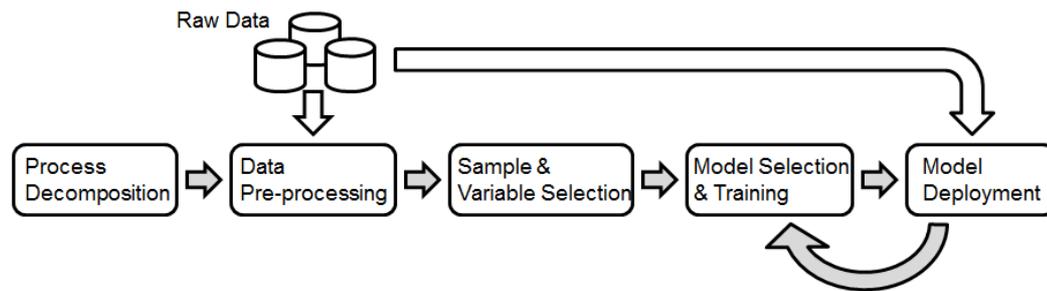


Figure 1. Data-driven modeling process

Investigated lumped capacitance models

Three different RC models with different complexity were investigated. For all models under investigation, the inputs of the model are the outside ambient temperature (T_A), the solar irradiance (q_{SOL}) as well as the heat flow into the thermal zone (ϕ_H), which is delivered by the heating system. The output of the model is the operative room temperature (T_R). The amount and their physical meaning of the model parameters differ for each RC model, due to the lumped network-structure.

The simplest RC model would be one lumped thermal capacity (C_R) for the whole room as well as one thermal resistance ($R_{R,A}$) for the envelope of the building. As the house under investigation uses a floor heating system, this simple model was extended with an additional resistor (R_H) and capacitor (C_H), to consider the high thermal inertia of that heating system. This 2R2C-model is shown in Figure 2. The state variable (T_H) represents the temperature of the floor heating system. The parameter A_w represents an effective window area factor, which is a simplification of various factors, like orientation of the building, shading by surroundings, etc. The differential equations of the simple 2R2C-model are stated in eq. (1) and eq. (2):

$$\dot{T}_R = \frac{1}{C_R} \times q_{SOL} \times A_w + \frac{1}{R_H \times C_R} \times (T_H - T_R) - \frac{1}{R_{R,A} \times C_R} \times (T_R - T_A) \quad (1)$$

$$\dot{T}_H = \frac{1}{R_H \times C_H} \times (T_R - T_H) + \frac{1}{C_H} \times \Phi_H \quad (2)$$

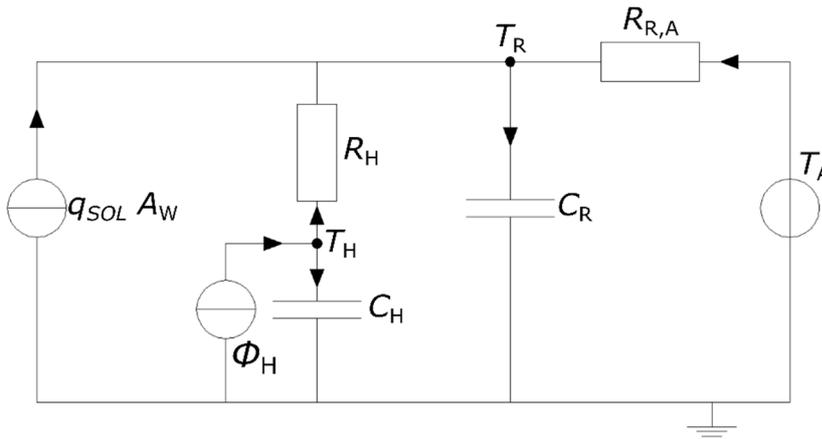


Figure 2. 2R2C building model

The second model is extended with an additional resistor ($R_{R,E}$), representing the thermal resistance between the room temperature and the building envelope, and a capacitor (C_E), representing the heat capacity of the building envelope. The thermal resistance between the envelope and the ambient air is represented by $R_{E,A}$. This leads to an additional internal temperature of the envelope (T_E). This 3R3C-model is presented in Figure 3. The corresponding differential equations are given in eq. (3-5).

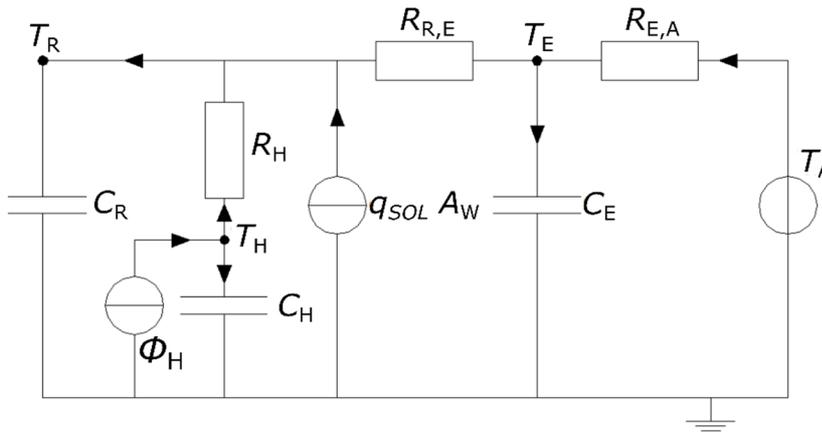


Figure 3. 3R3C building model

$$\dot{T}_R = \frac{1}{R_H \times C_R} \times (T_H - T_R) + \frac{1}{C_R} \times q_{SOL} \times A_W + \frac{1}{R_{R,E} \times C_R} \times (T_E - T_R) \quad (3)$$

$$\dot{T}_H = \frac{1}{R_H \times C_H} \times (T_R - T_H) + \frac{1}{C_H} \times \Phi_H \quad (4)$$

$$\dot{T}_E = \frac{1}{R_{R,E} \times C_E} \times (T_R - T_E) + \frac{1}{R_{E,A} \times C_E} \times (T_A - T_E) \quad (5)$$

The third and most complex RC model of this study is a 5R3C model, shown in Figure 4. The temperature node (T_R) has no thermal capacity contrary to the other models. Moreover, a temperature node (T_{IN}), representing the interior, and a thermal resistance

($R_{R,A}$), representing infiltration, are included. Heat gains through solar radiation have a direct impact to the room temperature based on the following eq. (6-9).

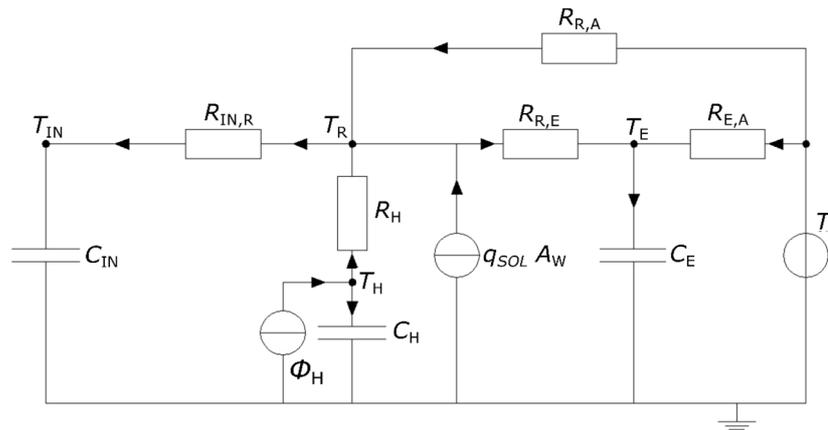


Figure 4. 5R3C building model

$$\dot{T}_{IN} = \frac{1}{R_{IN,R} \times C_{IN}} \times (T_R - T_{IN}) \quad (6)$$

$$\dot{T}_E = \frac{1}{R_{R,E} \times C_E} \times (T_R - T_E) + \frac{1}{R_{E,A} \times C_E} \times (T_A - T_E) \quad (7)$$

$$0 = -\frac{1}{R_{IN,R}} \times (T_R - T_{IN}) + \frac{1}{R_{R,E}} \times (T_E - T_R) + \frac{1}{R_{R,A}} \times (T_A - T_R) + \frac{1}{R_H} \times (T_H - T_R) + q_{SOL} \times A_W \quad (8)$$

$$\dot{T}_H = \frac{1}{R_H \times C_H} \times (T_R - T_H) + \frac{1}{C_H} \times \Phi_H \quad (9)$$

The identification of the RC model parameters was implemented as an optimization problem, where the minimum of the model output error has to be found. The Mean Squared Error (MSE) was used as output error function for that optimization. Eq. (10) shows the calculation of the MSE, where \hat{y}_i is the predicted load, y_i is the observed load, and n is the number of samples forecasted. The parameters of the RC models have to be tuned to find the minimum of the MSE. The identification period was chosen for 18 days in winter (2016-01-18 till 2016-02-04). The time resolution of the training and validation data was one hour:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (10)$$

Because of the non-convex and multidimensional parameter search space, a Genetic Algorithm (GA) was used to identify the optimal parameter set. The advantage of GA's over gradient decent techniques is that they are not biased towards locally optimal solutions, but still have the ability to direct the search towards a prospective region in the search space [21].

The GA was implemented by the use of the DEAP framework [22], with a population size of 100, a one-point crossover, and tournament selection which uses a tournament size of three. The mutation was implemented by a normal distributed random variation which range depends on the type of parameter. The cross over and mutation probability have been chosen by a sensitivity analysis. The cross over probability is set to a value

between 0.4 and 0.6, and the mutation probability is reduced from 0.3 to 0.1 during the search to reach better convergence. To evaluate the fitness of an individual, the MSE for the training data is calculated.

Investigated k-Nearest Neighbor model

In contrast to the RC-models, a black-box model was also investigated. Because of the short training period, an instance based shallow learning algorithm was chosen. As shown in previous work, a k-NN regression model is suitable for such kind of thermal modeling applications [23].

As mentioned before, k-NN is an instanced-based learning algorithm. Predictions are made, based on a comparison of the input data with the k closest existing instances. The output is calculated as an interpolation of those k closest neighbors. The principles of k-NN for time series prediction is explained in more detail in Yakowitz [24]. For multivariate time series regression, more details can be found in Al-Qathani and Crone [25]. The implementation uses distance weighting which considers close instances more than other. The k -parameter was empirically chosen with a value of 5. Other values of k have been tested, but delivered worse results. An additional lagged input of the heating power (ϕ_H) was used, to consider the thermal inertia of the heating system. The training and test set for the k-NN is the same as for the RC model identification. A sample of 18 days in winter and a resolution of 1 hour is used as training data. To test the performance, 7 days in each seasonal period are chosen, which are not in the training set.

Model validation

The model validation was done for all models in different seasons. A period of 7 days was used as shown in Table 1.

Table 1. Validation periods

	Winter	Spring	Summer	Fall
Start	2016-02-05	2016-04-22	2016-07-23	2016-11-01
End	2016-02-11	2016-04-28	2016-07-29	2016-11-07

To compare the performance of the models in different seasons, the Mean Absolute Error (MAE) was calculated, as shown in eq. (11), where \hat{y} represents the model output and y represents the observed value for a time step i , and n is the number of samples forecasted:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (11)$$

The MAE was calculated for every season to investigate the impact of the model identification period on the model performance.

RESULTS

The scatter plots in Figure 5 show the correlation between the operative room temperature (T_R) and the outside ambient air temperature (T_A) as well as the solar irradiance (q_{SOL}). As the power of the heating system cannot be regulated, this input is not shown in the plots.

As the cross-correlation coefficients show, there is not a very high linear dependency between the inputs and the operative room temperature (T_R). The low dependency of (T_A)

can be explained by the implemented two-point temperature control without shading. That control strategy holds T_R between two boundaries but is not able to prevent overheating during warm days with high solar gains. This can be seen in the higher correlation of q_{SOL} in Figure 5. The solar gain is increasing the room temperature, but depends on the sun position, the room's window size and its orientation.

Usually, the solar irradiance causes very high and short fluctuations in the operative temperature profile. As the time resolution is reduced to one hour, mean values for the radiation are used. This can explain the relatively low cross-correlation between T_R and q_{SOL} .

To characterize the operating condition of the building for different validation periods, the input data distribution is analyzed and shown in Figure 6. As the heat pump is only used for heating and not for cooling, the heat flow into the thermal zone (ϕ_H) is zero during the summer period. It can be visualized, that the operating conditions of the winter and the fall period are most similar with respect to mean ambient temperature, solar irradiance, and their variance. This is reflected in the model validation process, as described below.

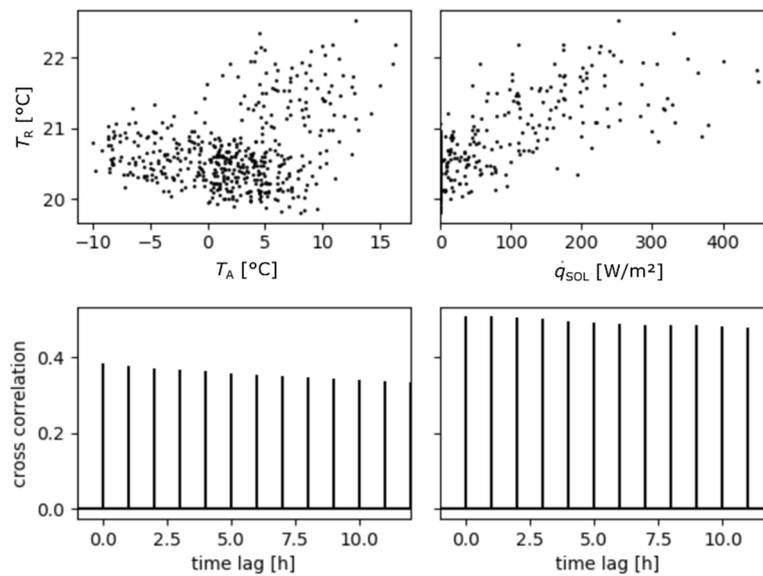


Figure 5. Cross-correlation between input data and room temperature

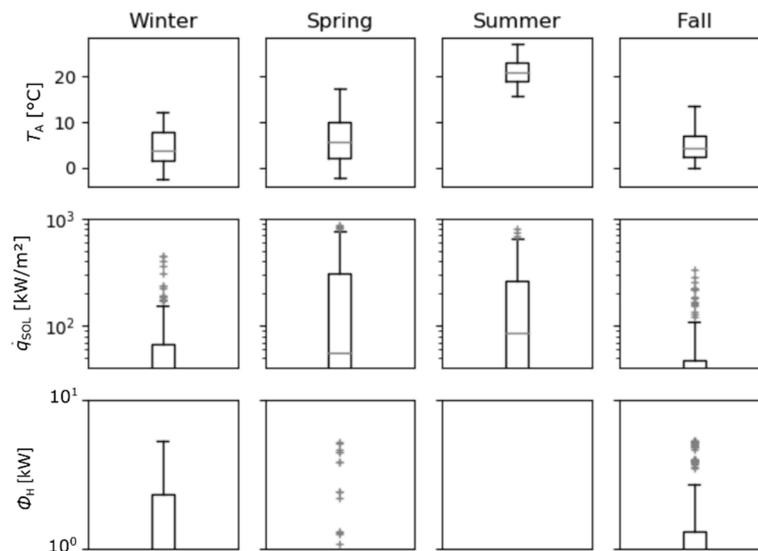


Figure 6. Distribution of input data

Parameter identification of the lumped capacitance models

The parameter identification was made with data from 18 days and a resolution of 1 h. As the search space increases with the complexity or rather the number of parameters of the model, the runtime of the GA took more than two hours on an Intel i5 dual-core with 4 GB RAM. Especially for the 5R3C model, the boundaries of the search space have to be well defined to avoid getting stuck in local minima. This needs some experience and knowledge about the building itself and the system dynamics. Thus, the whole parameter identification process was a quite complex and time-consuming task.

To reduce the parameter search space, parameters are often pre-defined based on physical information about the building structure (e.g. wall or window structure) [26]. This was not done in the present work, because this information is not always available (i.e., in particular, for existing buildings). The identified parameters are shown in Table 2.

Table 2. Identified model parameters

Parameters	Unit	2R2C	3R3C	5R3C
R_H	[°C/W]	$9,897 \times 10^{-7}$	$5,211 \times 10^{-7}$	$6,498 \times 10^{-5}$
$R_{A,R}$	[°C/W]	$1,260 \times 10^{-5}$	-	-
$R_{R,E}$	[°C/W]	-	$1,098 \times 10^{-5}$	1.655
$R_{E,A}$	[°C/W]	-	7.028×10^{-7}	$3,095 \times 10^{-6}$
$R_{R,A}$	[°C/W]	-	-	$1,222 \times 10^{-5}$
$R_{IN,R}$	[°C/W]	-	-	65.19
C_H	[Wh/°C]	13,847	18,937	15,396
C_R	[Wh/°C]	$9,699 \times 10^{-5}$	-	-
C_{IN}	[Wh/°C]	-	$6,938 \times 10^{-6}$	28,623
C_E	[Wh/°C]	-	4,788	3,612
i_w	[m ²]	5.454	7.741	0.0046

Model validation

As a performance metric for the model validation, the MAE was chosen. For a seven days prediction, the obtained results are shown in Table 3 for all four seasons. As expected, the MAE is good for all types of models. Also, the simplest 2R2C model shows good results, if the operation condition of the building is quite similar during the identification and validation process. This is the case for the winter and fall period, as shown in Figure 6. The simple 2R2C and 3R3C models show better performance for the other two periods (spring and summer), but the MAE of about 2 °C is still very high.

The k-NN performed best in winter and allowed fast training, as it is an instance-based learning algorithm. The MAE in spring and especially in summer is very high, because these conditions were not present in the training period.

Table 3. MAE of the models in different seasons

Model	Winter [°C]	Spring [°C]	Summer [°C]	Fall [°C]
2R2C	0.33	2.02	2.77	0.36
3R3C	0.31	2.08	2.44	0.31
5R3C	0.33	6.72	8.12	0.81
k-NN	0.20	2.29	12.48	0.37

Figure 7 shows the residuals of the model validation for all four seasons. For the 2R2C and the 3R3C model, the error distribution looks very similar. These mismatches of the models can be explained by the non-linearity of the system, caused by seasonal changes in building dynamics [27].

The 5R3C model seems to get stuck in a local minimum during the parameter identification. This could be the reason for the high deviations during the spring and

summer period. Despite the empirically change of parameter boundaries to force the identification process into a certain direction, no better parameter set could be found, as it was not intended to predefine parameters to reduce the search space. This dispersion of parameter estimates is a common problem in RC gray-box model identification [28].

The k-NN model is a data-driven model. Thus, it is only able to make accurate predictions under conditions where the input data is similar to the training data set. Therefore, the k-NN model produces a very high MAE in the summer period, but for the winter period the k-NN model outperforms the others.

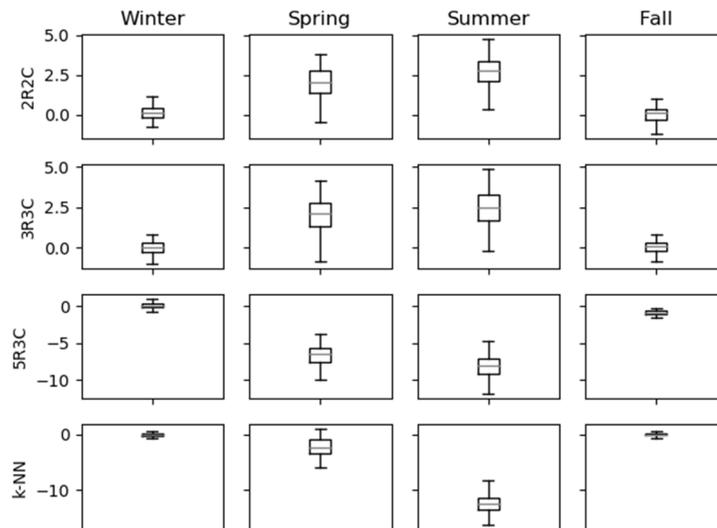


Figure 7. Distribution of residuals [°C]

Figure 8 illustrates the time series of model predictions during the validation period in winter. It shows the reference data (T_{R_ref}), which is the actual simulated data and the results of the 3R3C as well as the k-NN model for that period. The k-NN is able to reproduce the fast temperature fluctuation better than the 3R3C model.

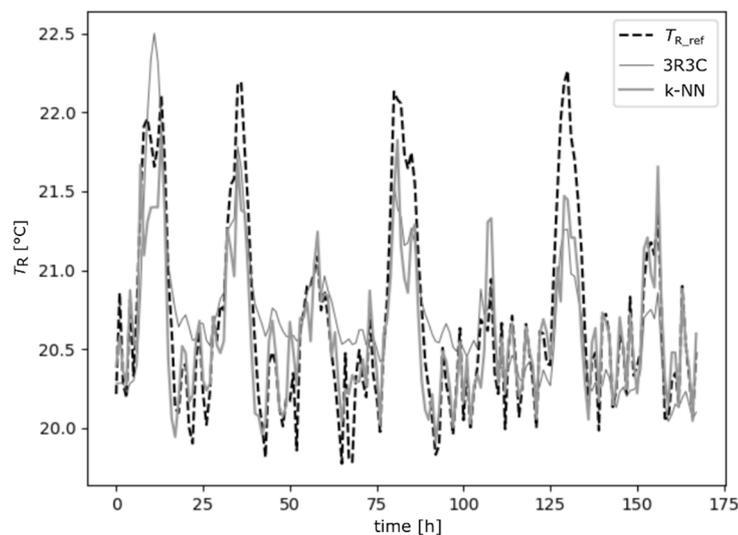


Figure 8. Model prediction for winter period

CONCLUSIONS

Even simple RC model structures are well suited for heat load prediction and intelligent model predictive control. As a drawback, the RC structure has to be adapted

considering the property of the heating systems. As well as black-box models, these RC models have the downside of seasonal dependency. Also, the parameter identification is a quite challenging task, because of the non-convex search space. Thus, a heuristic search method, such as a genetic algorithm was necessary to identify the model parameters. For complex gray-box models, domain knowledge should be applied to reduce the search space and avoid local minima during parameter identification. Also, the model identification should be performed under the same operating conditions as the model will be used, because of the seasonal impact or rather the non-linearity caused by seasonal changes in building dynamics. This is reflected by the error investigation of the models. Inner gains were not considered as the identification data was produced by a TRNSYS simulation.

For intelligent building control, the used models should be accurate and easy to create. It has been shown that machine learning algorithms, like k-NN are as accurate as RC models, but have less effort to create. Also, shallow learning algorithms do not need more training data as necessary for parameter identification, as shown in this work. The seasonal dependency of the machine learning models can be avoided, if more data is collected and online learning or periodical updates are implemented. Further investigations have to be done, how various machine learning algorithms perform and how the seasonal and day time influences can be considered by using adaptive models. Also, user behavior should be investigated and considered for these models. Furthermore, the time resolution of the prediction has to be increased and checked if the same accuracy can be achieved.

During the application of the presented data-driven model development process, potential of optimization was discovered in the context of data and knowledge management. A systematic support would help to facilitate heterogeneous data integration and increase the quality, consistency and expressiveness of data. In addition, a model reuse could be achieved. This could help to partly automate the model-based development process. Further investigations in this area will be conducted.

NOMENCLATURE

A_W	effective window area	[m ²]
C_E	thermal capacitance of the envelope	[Wh/°C]
C_H	thermal capacitance of the heating system	[Wh/°C]
C_{IN}	thermal capacitance of the interior	[Wh/°C]
C_R	thermal capacitance of the room	[Wh/°C]
$R_{A,R}$	thermal resistance of the ambient air to the room	[°C/W]
$R_{E,A}$	thermal resistance of the envelope to the ambient air	[°C/W]
R_H	thermal resistance of the floor heating system	[°C/W]
$R_{IN,R}$	thermal resistance of the interior to the room	[°C/W]
$R_{R,A}$	thermal resistance of the room to the ambient air	[°C/W]
$R_{R,E}$	thermal resistance of the room to the envelope	[°C/W]
q_{SOL}	solar irradiance	[W/m ²]
T_A	ambient air temperature	[°C]
T_E	temperature of the envelope	[°C]
T_H	temperature of the floor heating system	[°C]
T_{IN}	temperature of the interior	[°C]
T_R	operative room temperature	[°C]
T_{R_ref}	reference room temperature	[°C]

Greek letters

ϕ_H	heat flow produced by the heating system	[W]
----------	--	-----

REFERENCES

1. European Commission, Directorate-General for Energy, Standardization Mandate to European Standardisation Organisations (ESOs) to Support European Smart Grid Deployment, 2011, https://ec.europa.eu/energy/sites/ener/files/documents/2011_03_01_mandate_m490_en.pdf, [Accessed: 05-May-2019]
2. Losi, A., Mancarella, P. and Vicino, A., *Integration of Demand Response into the Electricity Chain: Challenges, Opportunities, and Smart Grid Solutions*, Wiley-ISTE, Hoboken, New Jersey, USA, 2015, <https://doi.org/10.1002/9781119245636>
3. Siano P., Demand response and smart grids - a survey. In: *Renewable and Sustainable Energy Reviews*, Vol. 30, pp 461-478, 2014, <https://doi.org/10.1016/j.rser.2013.10.022>
4. Eurostat, Electricity Consumption by Industry, Transport Activities and Households/Services, Technical Report, 2016.
5. Priya Esther, B. and Sathish Kumar, K., A Survey on Residential Demand Side Management Architecture, Approaches, Optimization Models and Methods, *Renewable and Sustainable Energy Reviews*, Vol. 59, pp 342-351, 2016, <https://doi.org/10.1016/j.rser.2015.12.282>
6. Aduda, K. O., Labeodan, T., Zeiler, W., Boxem, G. and Zhao, Y., Demand Side Flexibility: Potentials and Building Performance Implications, *Sustainable Cities and Society*, Vol. 22, pp 146-163, 2016, <https://doi.org/10.1016/j.scs.2016.02.011>
7. Killian, M. and Kozek, M., Ten Questions Concerning Model Predictive Control for Energy Efficient Buildings, *Building and Environment*, Vol. 105, pp 403-412, 2016, <https://doi.org/10.1016/j.buildenv.2016.05.034>
8. European Union, EUR-Lex – Access to European Union Law, Procedure 2016/0381/COD – COM (2016) 765: Proposal for a Directive of the European Parliament and of the Council Amending Directive 2010/31/EU on the Energy Performance of Buildings, EUR-Lex – 52016PC0765 – EN, 2016.
9. European Commission, Proposal for a Directive of the European Parliament and of the Council Amending Directive 2010/31/EU on the Energy Performance of Buildings, p 2 & 17, 2016.
10. Prívará, S., Cigler, J., Vána, Z., Oldewurtel, F., Sagerschnig, C. and Žáčková, E., Building Modelling as a Crucial Part for Building Predictive Control, *Energy and Buildings*, Vol. 56, pp 8-22, 2013, <https://doi.org/10.1016/j.enbuild.2012.10.024>
11. Cao, Y., Du, J. and Soleymanzadeh, E., Model Predictive Control of Commercial Buildings in Demand Response Programs in the Presence of Thermal Storage, *Journal of Cleaner Production*, Vol. 218, pp 315-327, 2019, <https://doi.org/10.1016/j.jclepro.2019.01.266>
12. Shengwei Wang, R. T. and Xu, L., An MPC-Based Optimal Control Strategy of Active Thermal Storage in Commercial Buildings During Fast Demand Response Events in Smart Grids, *Energy Procedia*, Vol. 158, pp 2506-2511, 2019, <https://doi.org/10.1016/j.egypro.2019.01.395>
13. Finck, C., Li, R. and Zeiler, W., Economic Model Predictive Control for Demand Flexibility of a Residential Building, *Energy*, Vol. 176, pp 365-379, 2019, <https://doi.org/10.1016/j.energy.2019.03.171>
14. Serale, G., Fiorentini, M., Capozzoli, A., Bernardini, D. and Bemporad, A., Model Predictive Control (MPC) for Enhancing Building and HVAC System Energy Efficiency: Problem Formulation, Applications and Opportunities, *Energies*, Vol. 11, No. 3, p 631, 2016, <https://doi.org/10.3390/en11030631>
15. Crabb, J. A., Murdoch, N. and Penman, J., A Simplified Thermal Response Model, *Building Services Engineering Research & Technology*, Vol. 8, No. 1, pp 13-19, 1987, <https://doi.org/10.1177/014362448700800104>
16. Yang, S., Wan, M. P., Ng, B. F., Zhang, T., Babu, S., Zhang, Z., Chen, W. and Dubey, S., A State-Space Thermal Model Incorporating Humidity and Thermal

- Comfort for Model Predictive Control in Buildings, *Energy and Buildings*, Vol. 170, pp 25-39, 2018, <https://doi.org/10.1016/j.enbuild.2018.03.082>
17. Andriamamonjy, A., Klein, R. and Saelens, D., Automated Grey Box Model Implementation Using BIM and Modelica, *Energy and Buildings*, Vol. 188-189, pp 209-225, 2019, <https://doi.org/10.1016/j.enbuild.2019.01.046>
 18. Bourdeau, M., Zhai, X.-Q., Nefzaoui, E., Guo, X. and Chatellier, P., Modelling and Forecasting Building Energy Consumption: A Review of Data-Driven Techniques, *Sustainable Cities and Society*, Vol. 48, 2019, <https://doi.org/10.1016/j.scs.2019.101533>
 19. ÖNORM EN 12831, Heating Systems in Buildings, Method for Calculation of the Design Heat Load (in German), Austrian Standards, 2003.
 20. Ge, Z., Review on Data-Driven Modeling and Monitoring for Plant-Wide Industrial Processes, *Chemometrics and Intelligent Laboratory Systems*, Vol. 171, pp 16-25, 2017, <https://doi.org/10.1016/j.chemolab.2017.09.021>
 21. Srinivas, M. and Patnaik, L. M., Adaptive Probabilities of Crossover and Mutation in Genetic Algorithms, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 24, No. 4, pp 656-667, 1994, <https://doi.org/10.1109/21.286385>
 22. Fortin, F.-A., De Rainville, F.-M., Gardner, M.-A., Parizeau, M. and Gagné, C., DEAP: Evolutionary Algorithms Made Easy, *Journal of Machine Learning Research*, Vol. 13, pp 2171-2175, 2012.
 23. Steindl, G. and Pfeiffer, C., Comparison of Black Box Models for Load Profile Generation of District Heating Networks, *Proceedings of the 12th Sustainable Development of Energy, Water and Environment Systems*, Vol. 12, Dubrovnik, Croatia, 2017.
 24. Yakowitz, S., Nearest-Neighbour Methods for Time Series Analysis, *Journal of Time Series Analysis*, Vol. 8, No. 2, pp 235-247, 1987, <https://doi.org/10.1111/j.1467-9892.1987.tb00435.x>
 25. Al-Qathani, F. H. and Crone, S. F., Multivariate k-Nearest Neighbour Regression for Time Series Data – A Novel Algorithm for Forecasting UK Electricity Demand, *Proceedings of International Joint Conference on Neural Networks*, pp 228-235, Dallas, Texas, USA, 2013, <https://doi.org/10.1109/IJCNN.2013.6706742>
 26. Ogunsola, O. T. and Song, L., Application of a Simplified Thermal Network Model for Real-Time Thermal Load Estimation, *Energy and Buildings*, Vol. 96, pp 309-318, 2015, <https://doi.org/10.1016/j.enbuild.2015.03.044>
 27. Killian, M., Mayer, B. and Kozek, M., Cooperative Fuzzy Model Predictive Control for Heating and Cooling of Buildings, *Energy and Buildings*, Vol. 112, pp 130-140, 2016, <https://doi.org/10.1016/j.enbuild.2015.12.017>
 28. Brastein, O. M., Perera, D. W. U., Pfeifer, C. and Skeie, N.-O., Parameter Estimation for Grey-Box Models of Building Thermal Behaviour, *Energy and Buildings*, Vol. 169, pp 58-68, 2018, <https://doi.org/10.1016/j.enbuild.2018.03.057>

Paper submitted: 07.02.2019
Paper revised: 05.05.2019
Paper accepted: 22.05.2019