11th International Workshop on Simulation for Energy, Sustainable Development & Environment
20th International Multidisciplinary Modeling & Simulation Multiconference

2724-0061 © 2023 The Authors.
doi: 10.46354/i3m.2023.sesde.004

Evaluating Machine Learning and Heuristic Optimization Based Surrogates as a Replacement for a Complex Building Simulation Model

Kathrin Kefer[1,*], Samuel Haijes[1], Michael Mörth[2], Richard Heimrath[2], Thomas Mach[2], Valentin Kaisermayer[3,4], Christopher Zemann[4], Daniel Muschick[4], Bogdan Burlacu[5], Stephan Winkler[5] and Michael Affenzeller[5]

[1]Fronius International GmbH, Günter-Fronius-Straße 1, Thalheim, 4600, Austria
[2]Institute of Thermal Engineering, Graz University of Technology, Inffeldgasse 25b, Graz, 8010, Austria
[3]Institute of Automation and Control, Graz University of Technology, Inffeldgasse 21b, Graz, 8010, Austria
[4]BEST - Bioenergy and Sustainable Technologies GmbH, Inffeldgasse 21b, Graz, 8010, Austria
[5]Heuristic and Evolutionary Algorithms Lab, University of Applied Sciences Upper Austria, Campus Hagenberg, Softwarepark 11, Hagenberg, 4232, Austria

*Corresponding author. Email address: kefer.kathrin-maria@fronius.com

Abstract

Intelligent energy management systems can play a vital role in supporting the much needed energy transition. However, in order to train machine learning models for this task, often very complex and detailed simulation models are needed. This can make the overall training process very slow or even impossible, which is why using resource efficient surrogates of the original simulation model during the training can be a possible solution. This work therefore focuses on the training of surrogates of a very detailed building simulation model using three different algorithms (k-Nearest Neighbour, Random Forest and Genetic Algorithm) and evaluates and compares them for their prediction capabilities, learned behaviours as well as execution time. Results show that the Random Forest algorithm achieves the best overall performance for 28 of the 35 surrogates, can learn the expected behavior and improves the execution speed by a factor of up to 664 compared to the original IDA ICE simulation model.

Keywords: Building Simulation Model Surrogates; Machine Learning; Heuristic Optimization; Energy Management System

1. Introduction

In the last years, climate change has proven to bring a lot of challenges for the future due to heating up the earth and by that facilitating enhanced natural disasters like floods or draughts. In order to push the energy transition, the world tries to limit the global warming and agreed in the Paris Agreement (par, 2015) to keep the rise of the global average temperature well below 2°C. In addition to that, according to the Global Status Report for Buildings and Construction (environment programme, 2021) done by the UN environment program, buildings (residential and non-residential) are responsible for 27% of the CO2 emissions. Therefore, intelligent systems that can optimize the energy flows in buildings, so that as much renewable produced energy as possible is used, become increasingly important. Until now, such systems either follow very simple, rule-based approaches (e.g. (Salpakari and Lund, 2016)) or take a long time during their execution due to needing a simulation model for predicting the future system behaviour (e.g. (Chen et al., 2013; Godina et al., 2018)).

In order to avoid these drawbacks, new approaches which use heuristic optimization based algorithms have come up (Morganti et al., 2009; Soares et al., 2016). However, as the training with such algorithms takes a lot of iterations and especially when using a simulation model for the training also a lot of time, it is essential to keep these simulation models as fast as possible. If the models are very detailed, thus resulting in longer execution times per iteration, one option is to replace them with surrogates. These surrogates should approximate the behaviour of the simulation model as well as possible while simultaneously keeping the execution time as small as possible. This is why in this work three different algorithms are used for the training of such surrogate models, including the k-Nearest Neighbour (kNN) as the most basic machine learning algorithm, the Random Forest (RF) as well-known benchmarking algorithm and also the Genetic Algorithm (GA) in combination with Symbolic Regression due to its ability to be extremely performant during the execution.

Therefore, this work contributes to the presented requirements in the following ways:

- In total, 7735 surrogates are trained for 35 outputs of a detailed building simulation model using three algorithms and different parameter settings: k-Nearest Neighbour, Random Forest and Genetic Algorithm.
- The surrogates are evaluated and compared with each other for their ability to match the predicted outputs with the simulation outputs on a held-back test set.
- The speed up of the execution is tested in comparison to the original simulation model.

The remaining work is structured as follows: chapter 2 gives an overview over related work, followed by the description of the method in chapter 3. Chapter 4 gives an overview on the results and chapter 5 concludes this work with a short summary and an outlook on future work.

2. Related Work

There are different approaches on how to train surrogates, and also a wide variety of applications and use cases available. Examples can be found for spot welding sequence optimization, where surrogates are trained using Neural Networks (Tabar et al., 2020), the approximation of a reservoir simulation based on deep learning surrogates (Jin et al., 2020), or the approximation of the low voltage energy grid also using artificial neural networks (Balduin et al., 2020).

The use case for this work is the approximation of a building simulation model in order to speed up the later training of an energy management system, whose objective is to optimally control the energy flows in a building in order to minimize its energy costs. Therefore, this chapter summarizes mainly related works where surrogates are trained by machine learning models and used to approximate building simulation models. The used algorithms range from classic machine learning e.g. by using support vector machines up to sophisticated deep learning models.

Classic machine learning algorithms used to train the surrogates are mainly focused on support vector machines (SVM) and mostly use the simulation environment EnergyPlus (Crawley et al., 2000) as a basis for the training. For example, already in 2012, Eisenhower et al. (Eisenhower et al., 2012) simulated an *EnergyPlus* building model and then trained a Support Vector Regression model on this data. This surrogate is then used for the optimization of the building in regard to a cost function which penalizes thermal comfort and energy consumption (Eisenhower et al., 2012). In 2017, Chen and Yang (Chen and Yang, 2017) published their work on a surrogate-based multi-stage optimization of passively designed high-rise residential buildings. Just as Eisenhower et al., also Chen and Yang use an *EnergyPlus* simulation model as basis for the surrogates, but train different surrogate model types: one based on multiple linear regression (MLR), one on multivariate adaptive regression splines (MARS) and one also on support vector machines. The SVM surrogate achieved the best prediction performance and is therefore used to optimize the design of the building using the multi-objective NSGA-II (Non-Dominated Sorting Genetic Algorithm II). With that, the computational efficiency of the trainings with the NSGA-II could be greatly improved (Chen and Yang, 2017). One approach using the k-Nearest Neighbour algorithm was proposed by Liang et al. in 2022 (Liang et al., 2022). Using simulations done in *EnergyPlus*, they created an electric load database with seven building parameters as inputs and the respective hourly energy consumption as output. With that, they trained five k-Nearest Neighbour surrogates using five different spatial metrics and evaluated them for their ability to predict hourly heating/cooling loads for hotel, office and retail buildings. With an accuracy of more than 90%, their approach proved to achieve very good results (Liang et al., 2022).

Another early work, but one that uses an artificial neural network (NN) to train surrogates, was presented by Magnier and Haghighat in 2010 (Magnier and Haghighat, 2010). They use the simulation environment *TRNSYS* and validate the created simulation model with measured data. Once the training of the artificial neural network with the validated model is done, the prediction results from the network are also validated. Finally, similar to the work done by Chen and Yang (Chen and Yang, 2017), also an NSGA-II is used to optimize the thermal comfort and the energy consumption in the residential building (Magnier and Haghighat, 2010). In the same year, Wong, Wan and Lam (Wong et al., 2010) published their work on an artificial neural network surrogate based on an *EnergyPlus* simulation model. From this simulation model, they extract weather and time data, the electric load for heating, cooling, lighting and the total building energy consump-

Table 1. An overview of the related works presented in chapter 2 for the used algorithms in comparison to this work and ordered by their year of publishing.

Authors	Year	Simulation Environment	Algorithm							
			kNN	SVM	RF	NN	MLR	MARS	GPM	GA
Magnier and Haghighat	2010	TRNSYS				x				
Wong et al.	2010	EnergyPlus				x				
Eisenhower et al.	2012	EnergyPlus		x						
Chen and Yang	2017	EnergyPlus		x			x	x		
Westermann and Evins	2021	Net-Zero navigator project				x			x	
Liang et al.	2022	EnergyPlus	x							
Kefer et al.	2023	IDA ICE	x		x					x

tion. Using this data, the surrogate model is trained and tested for its efficiency on four electrical target values with the *Nash-Sutcliffe coefficient*. With this approach, the authors achieved excellent prediction power and only small error rates (Wong et al., 2010). A more recent work which uses Bayesian Neural Networks to train surrogates was presented by Westermann and Evins in 2021 (Westermann and Evins, 2021). In their work, they try to approximate twelve energy performance metrics of a complex, high dimensional building using 35 input values. With that, they try to optimize the building's energy performance. In addition, they also train a stochastic variational Gaussian Process model (GPM) as surrogate and compare its performance to the one trained with the Bayesian Neural Network and find that both approaches achieved competitive results (Westermann and Evins, 2021).

Summing up and as shown in table 1, it can be stated that there are similar approaches already available in literature using classic machine learning (Eisenhower et al., 2012; Chen and Yang, 2017; Liang et al., 2022). Except for one work that uses the k-Nearest Neighbour from Liang et al. (Liang et al., 2022) in a similar way as it is done in this work, the related works mainly focus on the usage of support vector machines, which is not done in this work due to performance reasons during the training. Despite that, to the best of our knowledge, there are no previous works using a Random Forest algorithm and a Genetic Algorithm to train building simulation model surrogates. In literature, heuristic optimization algorithms are mainly used as the optimization algorithm in a surrogate assisted optimization approach, but not to train the surrogates themselves as it is done in this work. This work is also not training neural networks to create surrogates again due to performance reasons and uses an *IDA ICE* simulation model as a basis instead of an *EnergyPlus* model. This work also focuses more on a comparison of the three different algorithms used to create the surrogates using error metrics, a behaviour analysis and by measuring the execution time of the final surrogates.

3. Method

In order to create the surrogates, first all necessary data is extracted from the simulation model (section 3.1) by running it for a total of four years starting at the beginning of 2018. During that process, the input data used for the simulation model as well as all relevant output values are recorded and stored in csv files for later usage. This data basis and the additional features calculated for the surrogate trainings are described in more detail in section 3.2. Using this data, the surrogates are trained with three different algorithms (k-Nearest Neighbour, Random Forest and Genetic Algorithm) in the optimization framework HeuristicLab (Wagner et al., 2010) with multiple hyperparameter settings (section 3.3). Finally, the trained surrogates are evaluated for their prediction capabilities, behaviour and time saving capabilities as explained in section 3.4.

3.1. Building Simulation Model

The building in focus is located in the Innovation District Inffeld at Inffeldgasse 19, Graz, Austria (see figure 1). It was built in 2012 and is heated and cooled by two heat pumps and a coupled geothermal probe field. Heating and cooling are provided by underfloor heating and three central ventilation systems. The net floor area (2 216.84 m²) is divided into the following uses: 43% offices, 26% circulation areas, 8% storage, 7% recreation rooms, 5% lecture halls and libraries, 5% technical facilities, 4% sanitary and other areas and 3% laboratory and workshop. The sanitary and technical areas are connected to an exhaust air system. In addition, the circulation and sanitary areas are continuously supplied with fresh air by a central ventilation system. Table 2 shows the physical properties of the building envelope, as well as the net volume (NV), area/volume ratio (A/V), window/area ratio (W/A) and the infiltration rate at a pressure difference of 50 Pa (n_{50}) as input variables to the simulation model.

Based on the described parameters, a multi-zone simulation model was created in the IDA ICE simulation environment (AG, 2022) and calibrated using measurement data. The thermal zoning of the model is based on ÖNORM EN ISO 52016-1 (International, 2018) and considers the influencing factors solar radiation, orientation, occupancy, schedules and function as identified by Shin et al (Shin and Haberl, 2019). The use profiles are based on SIA 2024:2015 (und Architektenverein, 2015). The validation and calibration processes are done by evaluating the deviations of the simulation model from the real system. In addition, the annual heating and cooling energy in $kWh/(m^2*a)$ and the monthly heating and cooling energy in kWh/m^2 of the simulated and the real system are compared.

The year 2019 is used for the validation, as full occupancy (pre-corona) can still be assumed. The first step is

Table 2. Properties of the simulation model in terms of building physics.

Building Envelope Designation	Area [m²]	U-Value [W/(m²·K)]	UA-Value [W/K]	Share [%]	NV [m³]	A/V [m²/m³]	W/A [%]	n50 [1/h]
Walls in contact with air	907,46	0,2022	183,51	18,65%				
Walls in contact with the ground	400,87	0,2092	83,85	8,52%				
Roof	515,94	0,1389	71,66	7,28%				
Floor in contact with the ground	562,47	0,1144	64,34	6,54%				
Glazing	552,31	0,9605	530,51	53,90%				
Thermal bridging	-	-	50,34	5,11%				
Total	2939,05	0,3349	984,21	100%	7954	0,3695	18,79%	1,2

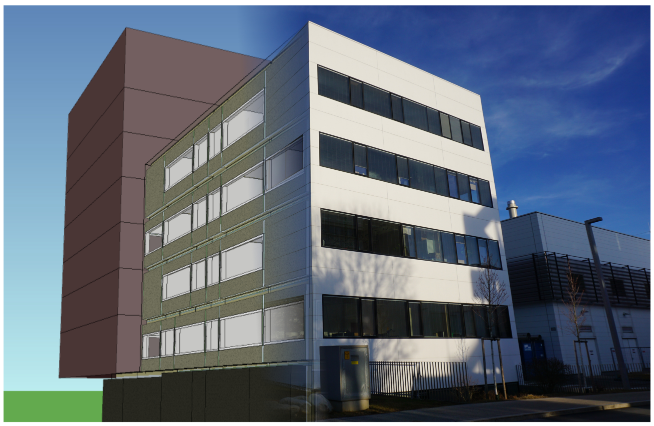


Figure 1. Inffeldgasse 19 in its real state in the picture on the right and the simulation model in IDA ICE on the left (Mörth, 2022)

to define the parameters and their bandwidths to adapt the simulation to reality and reduce the "performance gap". For the calibration process, energy transmittance (g-value), infiltration (n_{50}), thermal bridging coefficient (ψ-value) and internal heat loads for occupancy, lighting, and equipment, and finally the efficiencies of the latent heat exchangers in the ventilation system are used. For that, the released parameters are adjusted, and the deviation is reduced with the help of Automatic Multi-Objective Optimization (AutoMOO), which is an internal optimization algorithm of IDA ICE. These steps are repeated until the requirements for the model are met or a maximum number of iterations is reached. A very good agreement of the calibrated simulation results with reality could be achieved with that process. From November to March there is a maximum relative deviation of -3.4% (-0.3 kWh/m²). Large relative deviations occur only in the transitional months, for example up to 62.7% (0.4 kWh/m²) in September. A detailed description of the modelling method and the full set of boundary conditions and results can be found in the master thesis of Michael Mörth (Mörth, 2022).

3.2. Data Basis

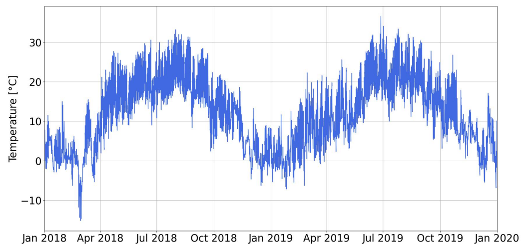
As data basis, four full years (2018-2021) are exported from the simulation model by running it in the IDA ICE (AG, 2022) simulation environment with an interval of 300 seconds in open-loop, i.e. no thermal controllers active, with an amplitude modulated pseudo-random bit sequences (APRBS), where the valve position was constant for six hours at a time. During the run, the input data as well as 104 relevant output values are logged and exported to a csv file once the simulation is finished. As input data for the simulation, weather data for the city of Graz, Austria, is fetched from the Geosphere data hub (zam) and contains the ambient temperature in °C (figure 2 (a)), the relative humidity (figure 2 (b)), wind speed in x and y direction (figure 2 (c)) and the solar irradiance values direct normal, diffuse horizontal and global horizontal irradiance (figure 2 (d)). The output values relevant for the surrogate trainings include the timestamp, the electric load for heating and cooling (figure 3 (a)) and the room temperatures (figure 3 (b)) as well as the valves for heating and cooling for each of the 34 controllable rooms in the building.

As preprocessing steps, the valves for heating and cooling are combined into one value with a range of $[-1;1]$ (figure 3 (c)), where positive values denote heating and negative values denote cooling. Then, the two output values for the electric load for heating and cooling of the building are combined by summing them up, which results in one electric load value. Based on the timestamp, seven additional time-based features are calculated for the training of the surrogates: the hour of the day, the day of the week as well as the month of the year each in sine/cosine representations (figure 3 (c)) so that the cyclic nature of these values is reflected appropriately, and a boolean indicating whether it is a working day (represented as 1) or a non-working day like saturday, sunday or public holiday in Styria, Austria (represented as 0). The sine and cosine representations are calculated as shown in equations 1 and 2.

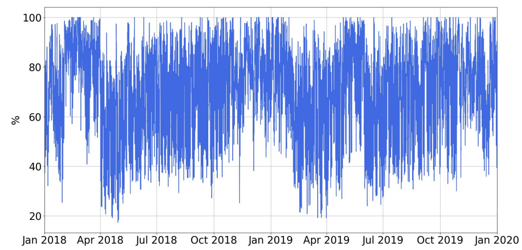
$$y = sin(2 * \pi * ParameterValue/maxValueOfParam) \quad (1)$$

$$y = cos(2 * \pi * ParameterValue/maxValueOfParam) \quad (2)$$

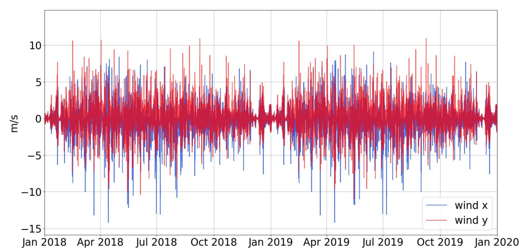
Once the preprocessing is finished, the full four year dataset is split up into two separate ones. The first one includes the first three years of the data from 2018 until the end of 2020. It is used for the training of the surrogates by using the first two years directly as training data and the year 2020 for testing. The remaining year 2021 is used as held-back test set for the final evaluation of the surrogates as described in section 3.4.



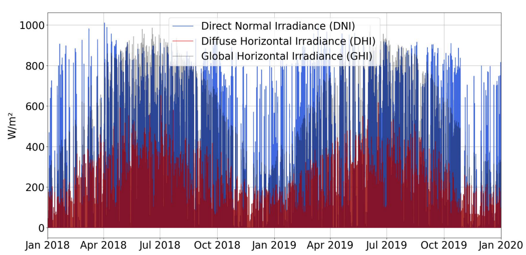
(a) The ambient temperature for 2018 and 2019 in °C.



(b) The relative humidity for 2018 and 2019 in %.



(c) The wind in x and y direction for 2018 and 2019 in m/s.

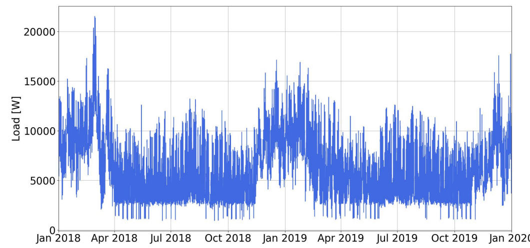


(d) The global irradiance values for 2018 and 2019 in W/m².

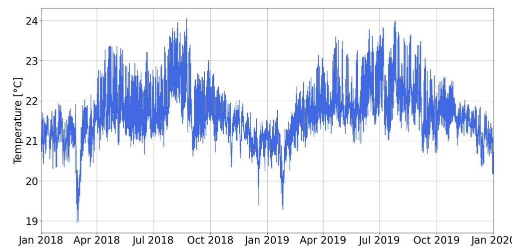
Figure 2. The weather data used as input for the simulation of the IDA ICE building model for 2018 and 2019.

3.3. Building Simulation Model Surrogates

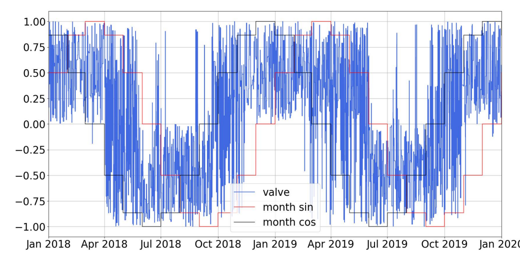
The training of the simulation model surrogates is done with three different algorithms, all implemented in the optimization framework HeuristicLab (Wagner et al., 2010), and include the k-Nearest Neighbour, the Random Forest and the standard Genetic Algorithm (Srinivas and Patnaik, 1994). The years 2018 and 2019 of the dataset described in section 3.2 are used as training data and



(a) The summed electrical load for heating and cooling of the building for the years 2018 and 2019 in Watts.



(b) The temperature course of a room (ground floor corridors and halls) for the years 2018 and 2019 in °C.



(c) The valve for the room *ground floor corridors and halls* as well as the sine and cosine representation of the month for the years 2018 and 2019.

Figure 3. An example of a room temperature, the building load for heating and cooling and the room valve for the training data years 2018 and 2019.

2020 is used as test data during the training. Using this dataset, one surrogate is trained to predict the energy consumption of the building i.e. the load for heating and cooling in Watts. For that, the heating/cooling valve of all 34 rooms, the described weather data as well as the calculated time features are used, summing up to a total of 48 input values. In addition to that, one surrogate is trained for each of the 34 controllable rooms of the building. For that, 15 input values are used, including the described weather data, the calculated time features as well as the heating/cooling valve for the one room for which the surrogate is trained. With that, the surrogates are trained to predict the exact room temperature in °C.

As hyperparameter settings $k = 1, 3, 5, 10, 20, 50, 100, 200, 500, 1000, 2000$ is chosen for the k-Nearest Neighbour. For the Random Forest, a batch size of five runs,

R = 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and M = 0.4, 0.5, 0.6, 0.7 are chosen as parameters. The same batch size of five runs is also set for the Genetic Algorithm. As parameters, mutation rates of 0.2 and 0.3, a population size of 100, 250 and 500 as well as 50, 100 and 200 generations are chosen. Additionally, the selector is defined to be the *Proportional Selector* with *windowing* set to *true*, as crossover the *Subtree Swapping Crossover* and as mutator the *Multi Manipulator* are set in HeuristicLab. As settings for the symbolic regression, the maximum tree depth is set to 50, the maximum tree length is set to 100 and as grammer the following operators are allowed: arithmetic (+-*/, avg, min, max), trigonometric (sin, cos, tan), exponential function, logarithm, power functions (square, power, squareroot and root) and the conditional symbols if-then-else, greater than, less than, and, or, xor and not. This results in a total of 11 k-Nearest Neighbour, 120 Random Forest and 90 Genetic Algorithm surrogates that are trained for the 34 controllable rooms and the electrical load, summing up to a total of 7735 trained surrogates.

3.4. Evaluation

The evaluation of the trained surrogates is done in a multi-step process. First, the best surrogate for each room and the building energy consumption is chosen among all algorithms and hyperparameter configurations based on the R^2 metric for the training test. These 35 surrogates are then tested on the held-back test set for the year 2021. The values for the R^2 and the Mean Absolute Error (MAE) metrics for the training, the training test and the held-back test as well as the estimated values are extracted from HeuristicLab. After that, the results are analysed and evaluated towards their prediction capabilities, the learned behaviour and the reduction of the execution time in comparison to the original IDA ICE simulation model.

The execution time comparison is done by running the IDA ICE simulation model and the surrogates with four different timespans (31, 90, 181 and 365 days) five times each with the held-back test set for 2021 in order to mitigate possible side effects of other programs running on the computing device. As computing resource, a Lenovo ThinkPad P15 Pro with an Intel Core i7vPro 10th Gen and 32GB RAM is used. First, the simulations are done directly in IDA ICE by running the simulation five times for the respective timespan and then reading the duration parameter provided by the simulation environment. Before being able to run the surrogates, they are extracted from HeuristicLab as C-Code, which is then integrated into a MATLAB Simulink simulation model (The Mathworks, 2022). Using the code generation functionality provided by MATLAB's Embedded Coder toolbox and a slightly adapted version of the energy management controller training process developed by Kefer et al. (Kefer et al., 2022), a DLL that can be executed from C#/.NET code, is generated. This DLL is then run by a Visual Studio 2022 project for five times with the four timespans described above and with the same input data as the IDA ICE simulation model. Finally, for every timespan, the average execution duration is calculated from the five consecutive runs for both, the IDA ICE simulation model as well as the surrogates DLL.

4. Results and Discussion

This section is split up in four different parts: first, the results for the surrogates' prediction capabilities are described in section 4.1. Then, the behaviour of some of the trained surrogates is analysed in more detail in section 4.2, followed by the description of the execution speed evaluation in section 4.3. Finally, the results are discussed in reference to the proposed contributions in section 4.4.

4.1. Prediction Capabilities

The prediction capabilities of the best 35 trained surrogates are evaluated based on the R^2 and Mean Absolute Error (MAE) metrics for the training, the test during the training and also on the held-back test set holding the data from 2021. For the electric load, the Random Forest with R = 0.6 and M = 0.5 performed best and achieved an R^2 of 0.9995 for training, 0.9355 for the training test and 0.9180 for the held-back test set. The mean absolute error showed a slight overfitting by achieving an error of 22.21W during the training and 513.07W respectively 585.16W for the training test and the held-back test set.

The results for the error metrics as well as the best performing algorithm for each of the 34 controllable room surrogates are shown in table 3. For 27 of these 34 rooms including all the office rooms, meeting rooms and common areas also the Random Forest algorithm performs best. However, the results show a higher overfitting compared to the electric load surrogate, which is in this case already visible when looking at the R^2 metric. Additionally, the results are varying between the different types of rooms. However, this can be easily explained by looking at the specific room types where the worst prediction results are occurring: Sanitary Rooms, Corridors and Halls as well as some Meeting rooms. It can be assumed that the temperatures of these rooms are harder to predict most likely just because there are more manual room temperature modifications happening e.g. in form of window openings done by the people in the building. Nevertheless, the results are very promising and show, that a maximum and average deviation of 0.73°C and 0.36°C respectively, are achieved for the held-back test set of 2021.

4.2. Behaviour Analysis

When analysing the behaviour of the different surrogates, we first take a closer look on the results achieved by the Random Forest for the prediction of the building load. Predicting the load of a building is generally a challenging task due to a lot of variability and the nearly unpredictable behaviour of the people in the building. However, the orig-

Table 3. The training, test and held-back test set results for all 34 room surrogates trained for this work sorted in decreasing order for R^2 Test 2021.

Room	Best Surrogate	R^2 Training	R^2 Test	R^2 Test 2021	MAE [°C] Training	MAE [°C] Test	MAE [°C] Test 2021
Staircase	Genetic Algorithm (MutProb=0.2, PopSize=500, MaxGens=200)	0.9516	0.9462	0.9583	0.2928	0.3064	0.2949
3rd Floor Corridors&Halls	Genetic Algorithm (MutProb=0.2, PopSize=250, MaxGens=200)	0.9234	0.9196	0.9369	0.2685	0.2660	0.2763
3rd Floor Meeting Rooms	Random Forest (R=0.6, M=0.6)	0.9983	0.9024	0.9022	0.0143	0.1811	0.1996
3rd Floor Common Area	Random Forest (R=0.5, M=0.7)	0.9971	0.9055	0.8896	0.0222	0.2126	0.2492
3rd Floor Office 4	Random Forest (R=0.4, M=0.6)	0.9959	0.8918	0.8818	0.0263	0.2038	0.2369
Ground Floor Office	Random Forest (R=0.7, M=0.5)	0.9993	0.8842	0.8752	0.0133	0.3921	0.4075
3rd Floor Office 1	Random Forest (R=0.4, M=0.7)	0.9957	0.9070	0.8698	0.0244	0.1700	0.2047
Ground Floor Office 1	Random Forest (R=0.4, M=0.7)	0.9967	0.8635	0.8650	0.0328	0.3615	0.3716
Basement Laboratories	Random Forest (R=0.3, M=0.7)	0.9926	0.9036	0.8548	0.0182	0.0978	0.0996
Ground Floor Office 2	Random Forest (R=0.7, M=0.4)	0.9992	0.8601	0.8472	0.0175	0.5352	0.5652
Ground Floor Meeting Rooms	Random Forest (R=0.5, M=0.5)	0.9975	0.8458	0.8316	0.0264	0.3775	0.4181
3rd Floor Office	Random Forest (R=0.6, M=0.7)	0.9980	0.8630	0.8284	0.0162	0.2507	0.2992
Basement Library	k-Nearest Neighbour (K=500)	0.9208	0.8367	0.8237	0.2056	0.2615	0.2915
Basement Sanitary Rooms	Random Forest (R=0.4, M=0.4)	0.9969	0.8281	0.8194	0.0140	0.1816	0.1934
Ground Floor Office 3	Random Forest (R=0.8, M=0.4)	0.9994	0.8438	0.8118	0.0104	0.4495	0.5113
1st Floor Office 4	Random Forest (R=0.6, M=0.7)	0.9959	0.8211	0.8063	0.0224	0.3782	0.4465
Ground Floor Office 4	Random Forest (R=0.7, M=0.6)	0.9988	0.8188	0.7895	0.0111	0.3254	0.3781
2nd Floor Corridors&Halls	Genetic Algorithm (MutProb=0.3, PopSize=500, MaxGenes=200)	0.7326	0.7172	0.7880	0.2408	0.2342	0.2639
1st Floor Meeting Rooms	Random Forest (R=0.6, M=0.5)	0.9978	0.7843	0.7645	0.0176	0.3807	0.4305
Ground Floor Corridors&Halls	Random Forest (R=0.5, M=0.4)	0.9974	0.7640	0.7556	0.0121	0.2454	0.2559
1st Floor Office 2	Random Forest (R=0.6, M=0.6)	0.9948	0.7687	0.7501	0.0190	0.3180	0.3548
1st Floor Corridors&Halls	Genetic Algorithm (MutProb=0.3, PopSize=100, MaxGens=200)	0.6809	0.6735	0.7455	0.2282	0.2196	0.2463
1st Floor Common Area	Random Forest (R=0.8, M=0.5)	0.9979	0.7714	0.7398	0.0099	0.3494	0.4095
1st Floor Office	Random Forest (R=0.8, M=0.5)	0.9990	0.7850	0.7242	0.0091	0.4031	0.4852
2nd Floor Office	Random Forest (R=0.8, M=0.6)	0.9983	0.7698	0.7175	0.0109	0.4333	0.4871
2nd Floor Office 3	Random Forest (R=0.3, M=0.6)	0.9760	0.7534	0.7157	0.0830	0.4615	0.5534
2nd Floor Office 1	Random Forest (R=0.6, M=0.6)	0.9921	0.7774	0.7039	0.0231	0.3257	0.3901
Ground Floor Sanitary Rooms	Random Forest (R=0.8, M=0.4)	0.9994	0.7213	0.7025	0.0025	0.1734	0.2053
2nd Floor Office 2	Random Forest (R=0.5, M=0.5)	0.9947	0.7355	0.7019	0.0446	0.6580	0.7292
Basement General Rooms	Random Forest (R=0.4, M=0.6)	0.9939	0.7607	0.6981	0.0385	0.4361	0.5031
2nd Floor Office 5	Random Forest (R=0.7, M=0.6)	0.9956	0.7057	0.6652	0.0153	0.3564	0.4376
Basement Corridors&Halls	k-Nearest Neighbour (K=500)	0.8467	0.6435	0.6496	0.1426	0.1988	0.2084
2nd Floor Meeting Rooms	Random Forest (R=0.6, M=0.4)	0.9959	0.6664	0.6268	0.0183	0.3431	0.3991
1st Floor Sanitary Rooms	k-Nearest Neighbour (K=500)	0.7960	0.5790	0.5861	0.1774	0.2525	0.2468

inal load behaviour can be approximated well even though there are some inaccuracies when predicting the consumption peaks (figure 4). One reason for that might be that the building is relatively well known with a regularly reoccurring behaviour: less load on non-working days and the transitional periods between summer and winter while there is higher load on working days or during summer and winter due to an increased heating and cooling effort.

The best result for the room temperature surrogates trained with the Random Forest is achieved for the 3rd floor meeting rooms with a R^2 on the held-back test set of 0.9022 and a MAE of 0.1996°C. Also the yearly course of the temperature is matched quite well (figure 5 (a)). In comparison to that, figure 5(b) shows the room where the Random Forest achieved the worst results with a R^2 of 0.6268 and a mean absolute error of 0.3991°C on the held-back test set, but still performs better than the two other algorithms. When taking a closer look on the temperature course, it becomes obvious that the major temperature characteristics of the room are still covered.

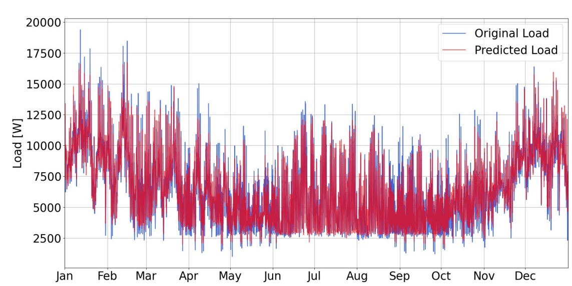
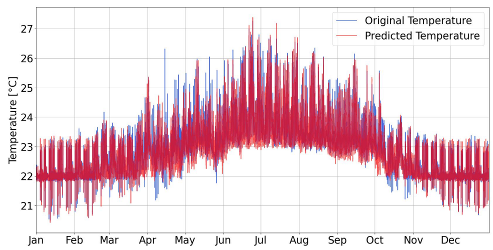
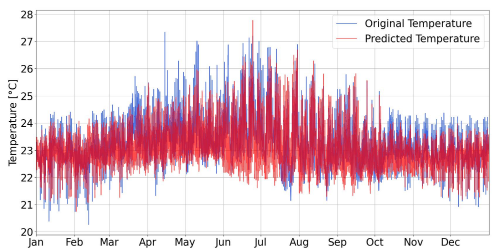


Figure 4. The original and predicted values achieved by the best-performing Random Forest surrogate on the held-back test set for the load of the building.

Taking a closer look on the results shown in table 3, it can be found that the Genetic Algorithm performed best in training nearly all *Corridors and Halls* surrogates and additionally also in training the surrogate for the staircase, where the overall best results among all rooms are



(a) 3rd Floor Meeting Rooms - Best Result



(b) 2nd Floor Meeting Rooms - Worst Result

Figure 5. The temperature course in °C of the two rooms for the held-back test set, where the Random Forest achieved the best and worst result.

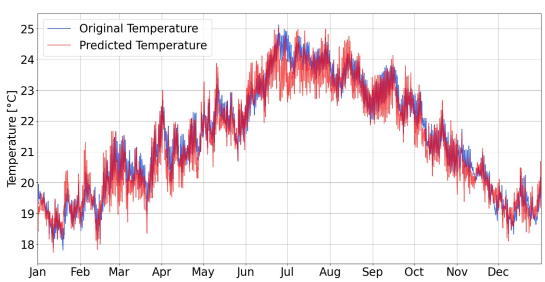
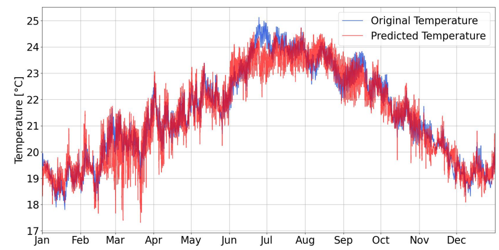


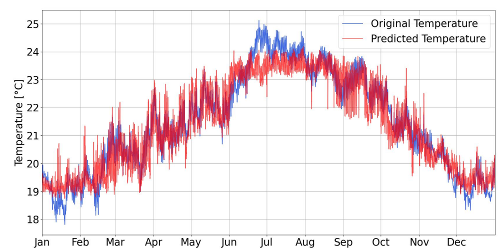
Figure 6. The temperature course in °C of the staircase for the held-back test set of 2021 achieved by the Genetic Algorithm.

achieved. When plotting the temperatures for the held-back test set for 2021, it is found that all rooms where the Genetic Algorithm performed best, have a very distinctive curve. As shown in figure 6, it starts with lower temperatures in the beginning of the year, higher ones during summer and then again lower temperatures towards the end of the year, while other rooms do not show these characteristics that much (e.g. the rooms shown in figure 5).

Comparing the results on the staircase from the Genetic Algorithm (figure 6) with the results achieved by the two other algorithms as shown in figure 7, it can be found that the general course of the temperature is learned well by all three algorithms. However, the k-Nearest Neighbour and the Random Forest are unable to predict some of the peaks occurring during the summer and winter time, which the Genetic Algorithm manages to approximate better.



(a) Random Forest - Staircase



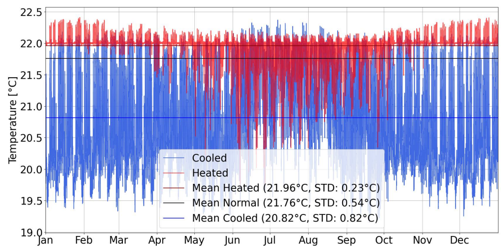
(b) k-Nearest Neighbour - Staircase

Figure 7. The temperatures in °C of the Random Forest and the k-Nearest Neighbour surrogates on the Staircase for the held-back test set for 2021.

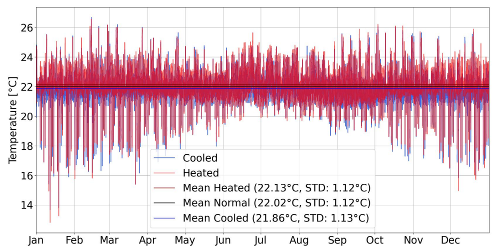
In order to check whether the trained temperature surrogates really learned the correlation between the valve and the temperature, e.g. a negative valve means cooling and therefore also lower room temperatures, two artificially alternated datasets are created based on the held-back test set. There, the valves for the rooms are manually set to -1 or +1 for the whole year, simulating maximum cooling or heating. These datasets are then applied on the surrogates just like the original held-back test set. The best visible effect can be found for the surrogate trained with the Random Forest for the *Basement Laboratories* (figure 8 (a)). Here, it becomes obvious that setting the valve manually to cooling drops the average room temperature by nearly 1°C. When setting it to heating, it does not have such a big effect but still increases the room temperature by 0.2°C on average. Similar, but not as big temperature changes have also been found for most of the other rooms. As example, the temperature course of an office on the 2nd floor is shown in figure 8 (b). Even though these temperature changes are not significant, they still indicate that the surrogates have learned the right behaviour.

4.3. Execution Speed Comparison

As shown in tables 4 and 5, a massive time reduction can be achieved when executing the surrogates instead of the original simulation model. On average, simulating one day with IDA ICE takes 20.2 seconds (not including the build process of the model and the initialization phase), while it takes only 0.035 seconds to run the surrogates with the DLL. This means, that the original simulation model is up



(a) Basement Laboratories.



(b) 2nd Floor Office.

Figure 8. The predicted values from the best-performing Random Forest surrogate using the datasets with the valves set to maximum heating (+1) and cooling (-1) for the year 2021. The mean for the predicted values of the original held-back test set is also plotted.

Table 4. The time measurements in seconds for different simulation time-spans of the year 2021 with the original IDA ICE simulation model.

	31 days	90 days	191 days	365 days
Run 1	547	1682	3918	7587
Run 2	688	1807	4284	7511
Run 3	590	1788	3863	7787
Run 4	564	1766	3948	7602
Run 5	569	1898	4018	7645
Average	591.6	1788.2	4006.2	7626.4

to 664 times slower in execution than the surrogates. Also the time needed to train the surrogates is negligible, as the training with the k-Nearest Neighbour and the Random Forest only take on average 1min 32s for one run. Only the training with the Genetic Algorithms takes some time and on average finishes after approximately six hours.

4.4. Discussion

Summing up, results show that machine learning algorithms can learn exact and fast surrogates of a complex simulation model. Especially training with the Random Forest resulted in good accuracies and the correct behaviour, while also being able to speed up the execution by a factor of up to 664.

5. Conclusions

In this work a comparison of three different algorithms, two classic machine learning algorithms in form of the k-Nearest Neighbour and the Random Forest and a Genetic Algorithm in combination with Symbolic Regression based on heuristic optimization, are used to train surrogates of a complex building simulation model. The building has a total of five floors and is modelled in the simulation environment IDA ICE. By running the simulation with an interval of 300 seconds for the four years from 2018 - 2021, the needed input data and the 104 relevant output values (including two different electric loads of the building and 34 room temperatures and valves) are logged and stored in csv files for later usage. Using this data, the surrogates are trained with the three different algorithms and multiple hyperparameter settings. Then, they are compared with each other for their prediction performance, their correct behaviour and also the speed up during execution using the held-back test set for 2021.

Table 5. The time measurements in seconds for different simulation time-spans of the year 2021 with the surrogate models encapsulated in the DLL.

	31 days	90 days	191 days	365 days
Run 1	1.168	2.535	6.679	10.658
Run 2	1.213	2.563	6.978	12.105
Run 3	1.557	2.168	6.841	11.433
Run 4	1.532	3.135	6.731	12.122
Run 5	1.085	3.057	7.215	11.713
Average	1.311	2.692	6.889	11.606

The results show that the Random Forest algorithm performed best in training the surrogates by achieving the overall best results for 28 of the 35 trained surrogates. For the electric load surrogate, an R^2 error of 0.918 could be achieved on the held-back test set, which means also a mean absolute error of 585.16W. Taking into account, that predicting the load of a building is a challenging task in general and that the building has an energy consumption of 6349.67W on average, this is a good result. The results for the 34 controllable rooms in the building are similarly good. The best results on the held-back test set for 2021 are achieved for the staircase with an R^2 score of 0.9583 and a mean absolute error of 0.295°C. However, especially for the rooms where a lot of unpredictable ventilation due to manually opened windows is happening, the prediction results can drop down to an R^2 score of 0.5851 while still having a mean absolute error of 0.2468°C. The Random Forest also proves to be able to learn the correct behaviour and can also reduce the time of execution by a factor of up to 664 compared to the original simulation with IDA ICE.

The limiting factors of this work are the use of only three very specific algorithms for training the surrogates as well as the used simulation environment. The given setup in this work limits therefore the comparability to the other related works, where none used a Random Forest and a Genetic Algorithm as it is done in this work and only one used a k-Nearest Neighbour model. Additionally, also most of the other works use the EnergyPlus simulation environment while this work uses IDA ICE. Despite that, the surrogate approach presented in this work is only the first step in a bigger project setup. For future work, the presented surrogates will be used in the process of training an energy management system, which should optimize the described building for minimal energy costs while keeping the user comfort as high as possible.

6. Funding

The research leading to these results has received funding from the Austrian Climate and Energy Fund, in the framework of the RTI-initiative "Flagship region Energy" under Grant No. 880792 (UserGRIDs - User-Centered Smart Control and Planning of Sustainable Microgrids), in cooperation with Green Energy Lab, an innovation laboratory for green energy.

References

Geosphere austria data hub. https://data.hub.zamg.ac.at/. Accessed: 2023-07-07.

(2015). The paris agreement. https://unfccc.int/process-and-meetings/the-paris-agreement. Accessed: 2023-04-10.

AG, E. S. (2022). Ida ice 5.0 beta 22. https://equa.se/de/ida-ice. Accessed: 2023-04-12.

Balduin, S., Westermann, T., and Puiutta, E. (2020). Evaluating different machine learning techniques as surrogate for low voltage grids. *Energy Informatics*, 3(1):1–12.

Chen, C., Wang, J., Heo, Y., and Kishore, S. (2013). Mpc-based appliance scheduling for residential building energy management controller. *IEEE Transactions on Smart Grid*, 4(3):1401–1410.

Chen, X. and Yang, H. (2017). A multi-stage optimization of passively designed high-rise residential buildings in multiple building operation scenarios. *Applied Energy*, 206:541–557.

Crawley, D., Pedersen, C., Lawrie, L., and Winkelmann, F. (2000). Energyplus: Energy simulation program. *Ashrae Journal*, 42:49–56.

Eisenhower, B., O'Neill, Z., Narayanan, S., Fonoberov, V. A., and Mezić, I. (2012). A methodology for meta-model based optimization in building energy models. *Energy and Buildings*, 47:292–301.

environment programme, U. (2021). 2021 global status report for buildings and construction. https://globalabc.org/sites/default/files/2021-10/GABC_Buildings-GSR-2021_BOOK.pdf. Accessed: 2023-04-13.

Godina, R., Rodrigues, E. M. G., Pouresmaeil, E., Matias, J. C. O., and Catalão, J. P. S. (2018). Model predictive control home energy management and optimization strategy with demand response. *Applied Sciences*, 8(3).

International, A. S. (2018). Energy performance of buildings — energy needs for heating and cooling, internal temperatures and sensible and latent heatloads: Part 1: Calculation procedures, en iso 52016-1.

Jin, Z. L., Liu, Y., and Durlofsky, L. J. (2020). Deep-learning-based surrogate model for reservoir simulation with time-varying well controls. *Journal of Petroleum Science and Engineering*, 192:107273.

Kefer, K., Hanghofer, R., Kefer, P., Stöger, M., Hofer, B., Affenzeller, M., and Winkler, S. (2022). Simulation-based optimization of residential energy flows using white box modeling by genetic programming. *Energy and Buildings*, 258:111829.

Liang, Y., Pan, Y., Yuan, X., Jia, W., and Huang, Z. (2022). Surrogate modeling for long-term and high-resolution prediction of building thermal load with a metric-optimized knn algorithm. *Energy and Built Environment*.

Magnier, L. and Haghighat, F. (2010). Multiobjective optimization of building design using trnsys simulations, genetic algorithm, and artificial neural network. *Building and Environment*, 45(3):739–746.

Morganti, G., Perdon, A., Conte, G., Scaradozzi, D., and Brintrup, A. (2009). Optimising home automation systems: A comparative study on tabu search and evolutionary algorithms. In *2009 17th Mediterranean Conference on Control and Automation*, pages 1044–1049. IEEE.

Mörth, M. (2022). Thermisch-elektrische modellierung und validierung der komponenten des quartier energiesystems innovation district inffeld. Master Thesis.

Salpakari, J. and Lund, P. (2016). Optimal and rule-based control strategies for energy flexibility in buildings with pv. *Applied Energy*, 161:425–436.

Shin, M. and Haberl, J. S. (2019). Thermal zoning for building hvac design and energy simulation: A literature review. *Energy and Buildings*, 203:109429.

Soares, A., Gomes, Á., Antunes, C. H., and Oliveira, C. (2016). A customized evolutionary algorithm for multiobjective management of residential energy resources. *IEEE Transactions on Industrial Informatics*, 13(2):492–501.

Srinivas, M. and Patnaik, L. M. (1994). Genetic algorithms: A survey. *computer*, 27(6):17–26.

Tabar, R. S., Wärmefjord, K., and Söderberg, R. (2020). A new surrogate model–based method for individualized spot welding sequence optimization with respect to geometrical quality. *The International Journal of Advanced Manufacturing Technology*, 106:2333–2346.

The Mathworks, I. (2022). *MATLAB version R2022a*. Natick, Massachusetts.

und Architektenverein, S. I. (2015). Raumnutzungsdaten für die energie- und gebäudetechnik, sia 2024:2015.

Wagner, S., Beham, A., Kronberger, G., Kommenda, M., Pitzer, E., Kofler, M., Vonolfen, S., Winkler, S., Dorfer, V., and Affenzeller, M. (2010). Heuristiclab 3.3: A unified approach to metaheuristic optimization. In *Actas del séptimo congreso español sobre Metaheurísticas, Algoritmos Evolutivos y Bioinspirados (MAEB'2010)*, page 8.

Westermann, P. and Evins, R. (2021). Using bayesian deep learning approaches for uncertainty-aware building energy surrogate models. *Energy and AI*, 3:100039.

Wong, S., Wan, K. K., and Lam, T. N. (2010). Artificial neural networks for energy analysis of office buildings with daylighting. *Applied Energy*, 87(2):551–557.