Buildingenergy.ninja: A web-based surrogate model for instant building energy time series for any climate.

Paul Westermann, Guillaume Rousseau, Ralph Evins
Energy and Cities group, Department of Civil Engineering, University of Victoria, BC, Canada
E-mail: pwestermann@uvic.ca

Abstract. Machine learning-based surrogate models are trained on building energy simulation input and output data. Their key advantage is their computational speed allowing them to produce building performance estimates in fractions of a second. In this work we showcase the use of deep convolutional neural network surrogate models embedded into a web application, allowing users to rapidly explore building performance at high spatio-temporal resolution. Users can pick any climate on an interactive map, customize a building design with thirteen decisive design parameters, and the surrogate model allows them to retrieve hourly heating and cooling load time series data in fractions of a second.

In this work, we further show that the surrogate model reaches an accuracy of $R^2 > 0.93$ ($MAE < 0.27$ kWh) for unseen design specifications and climates. These results motivate the use of computationally cheap surrogate models to replace building energy simulation for a wide variety of tasks in the future.

1. Introduction

Energy performance prediction of buildings is a key discipline to model the built environment, the sector with the largest energy demand. Predictions are essential for the sustainable and energy efficient design from individual buildings to entire cities; and for the design of stable energy networks around the globe.

Machine learning models can be used to augment well established physics-based building energy simulation [1]. In this work we focus on the use of machine learning models that are trained on simulation input and output data to serve as a surrogate of the underlying simulation model [2]. The key advantage of surrogate models is their low computational cost to produce simulation output estimates, usually in fractions of a second [3]. This fundamentally changes modelling processes from one-at-a-time sample simulations to interactive exploration of spaces of design options [4][5]. Recent research has shown (a) surrogate models’ high accuracy when estimating various building performance metrics, (b) that novel machine learning model architectures can process highly complex simulation inputs like building geometry data [6] or climate time series data making them suitable for various design problems [7], and (c) that surrogate model-based interactive building design processes are favoured by building designers over traditional simulation-based and also lead to better performing designs [11]. Alongside academia, these findings have sparked new initiatives in industry to develop surrogate model-based building, neighbourhood and city modelling tools for scalable, performance-driven design [8][9][10].
Figure 1. buildingenergy.ninja: A website hosting surrogate models to provide hourly building energy time series data for any location around the globe based on the use of sequence-to-sequence CNN surrogate models (compare [17]).

While recent results have set the trajectory towards an uptake in the use of surrogate models significant shortcomings remain.

- Epistemic uncertainty is introduced by the surrogate model layer and may be large for individual samples [12].
- Surrogate models showed excellent performance for aggregated performance metrics but lower accuracy when estimating dynamic outputs [13].
- They have proven to generalize well over various geometries and a large variety of design parameters [14] but still require retraining when used for predicting building performance at locations and with building systems they weren’t trained for.

In this study, we focus on the use of sequence-to-sequence surrogate models [15][16] that are able to process time series data inputs, here entire annual hourly climate files, and estimate hourly building performance time series, here heating and cooling loads. After training the surrogate models on simulation data from 20’000 EnergyPlus simulations at only 120 locations around the globe, the surrogate model is capable to produce accurate building time series data for any location in the world.

We foresee the application of the trained surrogate model for various disciplines. Therefore, we make the trained surrogate model accessible via a containerized web app, buildingenergy.ninja, based on [17]. It is built around an interactive map of the world to pick climates (taken from [18]), and further, lets users customize building designs of currently two archetypes (office building, single-family home) for which the time series are generated.

2. Methodology
Training a surrogate model involves 3 steps. First a simulation model, which is aimed to train the surrogate model on, and its design parameters are specified. Second, simulations are sampled,
Figure 2. Overview of the building simulation models and design parameters.

There for various climates, to generate the training (and testing) dataset. Third, a neural network architecture was specified, that is capable to process both the climate data (around 150’000 values) and the building design parameters, and trained on the dataset. The model is then tested on simulation ground truth data generated with climates not contained in the training data.

2.1. Building simulation model
Here we built a surrogate model for two buildings and two outputs to demonstrate that the process generalizes over various design problems. We modelled an H-shaped office building with significant self-shading that needs to be captured by the surrogate model and a 220 m² single-family home with three thermal zones. The design parameters that we used as surrogate model inputs were aligned for both buildings and are listed in Figure 2. The runtime of one EnergyPlus simulation using two cores and 4GB of RAM is around 30 seconds.

2.2. Simulation data set
The selection of training and testing simulation samples involves two steps (see Figure 4). Selecting (a) building parameter combinations and (b) climates for which the EnergyPlus simulations are performed. (a) was done by creating a latin hypercube with 20’000 levels of the thirteen design parameters within their range of values [19]. For (b), we initially took a random selection of 500 samples from all climates available on [18]. However, this introduced a strong sample selection bias towards climates in the northern hemisphere and towards mild climates. Instead, we used the Koeppen climate classification scheme and hand-picked 4 locations per Koeppen climate instead (totalling 120 training locations) [20]. The climate files consist of 17 variables with 8760 hourly values. The list of climate variables is given in [7]. After each simulation run, we stored the hourly heating and cooling load time series (8760 values each). For testing the model performance, we repeated that process with 30 testing locations, where we produced 1000 building building heating simulations for each of them. All the simulation samples were run on computing resources from Compute Canada [2].

1 The Koeppen-based sample selection drastically improved the model performance compared to the random climate selection.

2 www.computecanada.ca
Figure 3. Architecture of the convolutional neural network. The inputs are (a) an EnergyPlus weather file (.epw file) and (b) values for thirteen design parameters. The network is trained twice to estimate two hourly time series for heating load and cooling load.

2.3. Neural network architecture: ResNet
The architecture is displayed in Figure 3. It is based on the convolutional residual neural network (ResNet) used in [21] to translate a set of input time series (hourly time series of various weather variables) to produce an output time series, here building energy loads. We concatenate the time series data with building design parameters. The final model has 5 blocks of 3 convolutional layers (kernel sizes: 8,5,3) plus one dense layer that up-samples the design parameters to 8760 dynamic values (see block 2). Each layer is followed by a batch normalization to prevent a covariate shift and an activation layer (rectified linear unit). As it was found the location of concatenating the up-sampled design data with the processed weather data is essential. When concatenating design parameters and the processed weather data after block 2, the network outputs were less dynamic leading to lower accuracy. The architecture was implemented in Tensorflow using the Keras API [22] and trained using a computing instance provided by a Google Research Grant.

2.4. Web hosting
The trained neural network, including the network architecture (stored in .json-file), approximately 650’000 weights (.h5-file) and the data preprocessing pipeline that standardizes the inputs and outputs (.p) has the advantage of having low storage requirements (≈ 3 MB) and is easily transferable among machines. Here, we host the model on a small webserver as part of a Flask application written in Python [23] (buildingenergy.ninja). The core element of the website is a map of the world that allows to pick the location for which heating or cooling load time series can be downloaded. The time series for a customized building design are made available as a .csv file.

3. Results
We quantify the performance of the models using the coefficient of determination, $R^2$ and the mean absolute error, $MAE$ [kWh] as it was used in other research on surrogate models [2]. We

3 An explanation why CNNs are specifically suitable for modelling the impact of climate data on building energy demand can be found in [7].
Figure 4. (a) The Koeppen climates based selection of training locations (red), as well as 30 randomly picked testing locations (blue). (b) Heating (top) and cooling load prediction accuracy for the office building broken down into each of the testing locations sorted by climate (cold on the left; warm on the right). The $R^2$ score is shown as circles, the $MAE$ shown as stars.

Table 1. Accuracy for hourly heating and cooling load predictions.

|                   | $R^2_{\text{train}}$ | $R^2_{\text{test}}$ | $MAE_{\text{train}}$ [kWh] | $MAE_{\text{test}}$ [kWh] |
|-------------------|----------------------|----------------------|-----------------------------|-----------------------------|
| Heating load (Office) | 0.9421               | 0.9365               | 0.27                        | 0.28                        |
| Cooling load (Office)  | 0.9603               | 0.9547               | 0.27                        | 0.27                        |
| Heating load (Single-family home) | 0.9623               | 0.9724               | 0.08                        | 0.07                        |
| Cooling load (Single-family home) | 0.9511               | 0.9431               | 0.16                        | 0.18                        |

reach an overall accuracy of $R^2 > 0.93$ and $MAE < 0.28$ kWh on the test data (see Table 1). To the best of our knowledge this is currently the highest accuracy of hourly building energy surrogate models published, e.g. it outperforms [13] although the same climates were used for training and testing. We find that an increase in model complexity, i.e. more thermal zones in the office building in comparison to the single-family home, leads to a small decrease in accuracy. In Figure 4 we further break down the model performance to each location of the test data for the office building and find that the absolute errors are the highest where also the loads are the highest (e.g. larger cooling load $MAE$ in warm places). In future we will account for this behaviour with a normalized error metric.
4. Conclusion
In this work we present building energy surrogate models with high spatio-temporal resolution. They allow to produce hourly building heating and cooling load estimates for any location around the globe for two reference buildings. Further, we showcase how surrogate models are easy to make building simulation accessible for everyone. Here, we host the lightweight and fast surrogate models on a webserver and included into an app that allows users to retrieve building time series data for customized building designs globally. In upcoming publications we will show how the process can be extended to any kind of time series simulation inputs (e.g. occupancy time series data) and how it can be merged with existing surrogate models that can predict building performance for a large variety of building geometries [6]. Another shortcoming is that the output is limited to heating and cooling loads, i.e. no building system was modelled. Future research is required on how surrogate models can generalize over various building systems.

4.1. Acknowledgements
This research was funded by Canarie through the BESOS research grant (RS-327). Further, it highly benefited from computing resources provided through the Google Cloud Platform (GCP) research credits program.

References
[1] Sutherland B R 2020 Joule 4 2256–2258
[2] Westermann P and Evins R 2019 Energy and Buildings 198 170–186
[3] Ostergard T, Jensen R L and Maagaard S E 2018 Applied Energy 211 89–103 URL [GotoISI]:://WOS:000425075600008
[4] Wortmann T 2017 International Journal of Architectural Computing 15 38–53
[5] Brown N C, Jusiega V and Mueller C T 2020 Automation in Construction 118 103252
[6] Han J M, Chang C K and Malkawi A 2020 Arinet: Using 3d convolutional neural networks to estimate annual radiation intensities on building facades 2020 Building Performance Analysis Conference and SimBuild co-organized by ASHRAE and IBPSA-USA
[7] Westermann P and Evins R 2020 Applied Energy 264 114715
[8] URL https://vizcab.io/
[9] URL https://hello.delve.sidewalklabs.com/
[10] Paul Westermann, David Rulff, Kevin Cant, Gaëlle Faure and Ralph Evins URL http://www.enerarxiv.org/page/thesis.html?id=1975
[11] Brown N C 2020 Design Studies
[12] Westermann P, David N and Evins R 2018
[13] Vazquez-Canteli J, Demir A D, Brown J and Nagy Z 2019 Deep neural networks as surrogate models for urban energy simulations Journal of Physics: Conference Series vol 1343 (IOP Publishing) p 012002
[14] Geyer P and Singaravel S 2017 Applied Energy 228 1439–1453 ISSN 0306-2619
[15] Gehring J, Auli M, Grangier D, Yarats D and Dauphin Y N 2017 Convolutional sequence to sequence learning Proceedings of the 34th International Conference on Machine Learning-Volume 70 (JMLR. org) pp 1243–1252
[16] Fawaz H I, Forestier G, Weber J, Idoumghar L and Muller P A 2019 Data Mining and Knowledge Discovery 33 917–963
[17] Staffell I and Pfenninger S 2018 Renewables. ninja
[18] Lawrie L K and Crawley D B 2019 Development of global typical meteorological years (tmyx). URL http://climate.onebuilding.org
[19] Stein M 1987 Technometrics 29 143–151
[20] Wortmann T 2018 Journal of Computational Design and Engineering ISSN 2288-4300 URL http://www.sciencedirect.com/science/article/pii/S2288430018300721
[21] Wang Z, Yan W and Oates T 2017 Time series classification from scratch with deep neural networks: A strong baseline 2017 international joint conference on neural networks (IJCNN) (IEEE) pp 1578–1585
[22] Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving G, Isard M et al. 2016 Tensorflow: a system for large-scale machine learning. OSDI vol 16 pp 265–283
[23] Grinberg M 2018 Flask web development: developing web applications with python (" O'Reilly Media, Inc.")