Applied Energy 228 (2018) 1439–1453

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Component-based machine learning for performance prediction in building design

Philipp Geyer*, Sundaravelpandian Singaravel

Architectural Engineering, Faculty of Engineering Science, KU Leuven, Belgium

HIGHLIGHTS

- A component-based method of machine learning for performance prediction in engineering.
- Components instead of one monolithic model extend reusability and generalization.
- Flexible design support by machine learning for early design phases.
- Internal parameters between components allow insights in the "black box" of machine learning.
- Good prediction accuracies (error < 3.9%) for test cases different from the training model.

ARTICLE INFO

Keywords:
Component-based machine learning
Systems engineering
Parametric systems modeling
Building performance prediction
Building simulation

ABSTRACT

Machine learning is increasingly being used to predict building performance. It replaces building performance simulation, and is used for data analytics. Major benefits include the simplification of prediction models and a dramatic reduction in computation times. However, the monolithic whole-building models suffer from a limited transfer of models and their data to other contexts. This imposes a vital limitation on the application of machine learning in building design. In this paper, we present a component-based approach that develops machine learning models not only for a parameterized whole building design, but for parameterized components of the design as well. Two decomposition levels, namely construction level components (wall, windows, floors, roof, etc.), and zone-level components, are examined. Results in test cases show that, depending on how far the cases deviate from the training case and its data, high prediction quality may be achieved with errors as low as 3.7% for cooling and 3.9% for heating.

1. Introduction

The challenge of a sustainable built environment requires the early integration of performance in design processes. This leads to a complexity never seen before in building design. To manage this complexity, designers, planners and engineers need to quickly obtain an overview of the overall performance of a building, including the systemic dependencies. Systemic dependencies often cross disciplinary boundaries, and thus lead to multidisciplinary interdependencies that play a vital role in overall performance. To manage this complexity, information is needed on performance, on these interdependencies, on the emerging design space, and on specific well-performing regions in a process of design space exploration (DSE) [1,2] in order to be available quickly enough and easily enough for the design process.

Machine learning (ML) provides a solution in this situation, offering the advantages of fast prediction and simplified parameter structures matching early design phases. This, allows designers and engineers to change designs quickly and to observe the consequences for performance in a DSE process. However, current ML approaches are developed too specifically for design situations, and require redevelopment for new cases.

The paper will therefore develop a component-based approach supporting such a systemic performance prediction. Based on a component-oriented structure, and on machine learning models, the thermal energy performance of buildings is predicted as an example of performance-based design. The aim is to prove the novel component-based approach of ML with components that serve prediction in multiple cases. Although the paper is limited to thermal performance, other disciplines of building performance can be treated in a similar manner, thus allowing a link to be created to observe systemic interdependencies and multidisciplinary building performance.

* Corresponding author.
E-mail address: p.geyer@kuleuven.be (P. Geyer).

https://doi.org/10.1016/j.apenergy.2018.07.011
Received 26 October 2017; Received in revised form 20 April 2018; Accepted 3 July 2018

1.1. Background

Two approaches currently exist to predict performance in building design: firstly, physical modeling and simulation and, secondly, machine learning models. In terms of modeling categorization, the first approach is also known as white box modeling, whereas the second is known as black box modeling.

The methods of physical white box modeling of buildings are quite well examined, as is shown by the fundamental literature such as the books by Clarke [3] or the book by Hensen and Lamberts [4]. These approaches consider dynamic effects as they are required to provide quite precise results. However, they have a high information demand and involve relatively long computation times. Multi-domain simulation (co-simulation) can be carried out by coupling energy simulation with other simulation models, and has been shown in several examples [5–7]. This allows the consideration of discipline-crossing dependencies, thus permitting the complexity described above to be addressed. However, the effort involved in terms of modeling and computation time is significant, and exceeds what is possible in normal design situations.

Black box methods of statistical surrogate or meta modeling and machine learning offer effective energy performance prediction in this situation of complex design and engineering [8–10]. Such methods, also known as black box models, typically act in an analogous way to a physical simulation. Surrogate models based on the response surface method (RSM) were an early approach towards reducing the number of experiments by making good predictions [11,12].

The approach of black box modeling based on ML and surrogate modeling is used for building energy performance prediction, as is shown by reviews of the field [13–17]. Table 1 breaks the selected exemplary studies down into categories. Model Category 1 uses ML and other surrogate models to give static feedback, which is basically one number per feedback variable, such as yearly total heating energy demand in kWh/a. This is carried out for application at subsystem level, at building level and at urban level. For the latter two subcategories, parametric black box models of buildings are used to quickly predict energy consumption. This supports either one design, or the optimization of one building (Section 1.2). Furthermore, black box models have been applied to the design and optimization of subsystems (Section 1.1). Besides these applications with a static response, models with a dynamic response have been developed (Section 2). On the one hand, they are applied for the purposes of design and optimization, i.e. to identify peaks in operation quickly and to reduce them in designing (Section 2.1). The second purpose of developing dynamic-response models is application in control (Section 2.2). All these models are monolithic black box models, which means that they cannot interpret what is happening between input and output models in a physical way. For instance, monolithic building models predict the heating energy consumption by input parameters, such as physical properties of windows and wall. However, internal properties that lead to the energy prediction, e.g. the share of heat flows through the windows and walls, are not known due to the monolithic characteristic of the model. This missing internal model information is a significant disadvantage of black box models that is overcome by taking a component-based approach.

When it comes to analyzing and designing energy systems, the application of ML takes place in a system engineering-based way that provides internal information and produces reusable system components (Section 3 in Table 1). This is founded on the discipline of systems engineering [52–55]. This is very valuable for designing engineering artifacts, helping to understand dependencies between its system modules, and for managing the related complexity, as it is performed by the design structure matrix method [56,57]. This method has been transferred to building construction and its adaptability [58,59], as well as to information flows during design [56,60,61]. Modeling approaches following this modular method have a high potential for understanding the complexity caused by sustainability, as is shown by the system modeling approach that has been developed for building design and urban contexts [62,63]. As is described in the next section, we follow this systems approach and use it as a basis to develop prediction models for the domain of building design.

Table 1
Overview of relevant ML modelling approaches.

Model type	Application area	Studies
Monolithic model of static energy prediction (total yearly energy consumption)	Subsystem: Zone/HVAC/ Building component	[18–25]
	Building level: Design, Optimization	
	Building stock: Urban, Regional, National	[26–30]
Monolithic model of dynamic energy consumption patterns	Building level: Design, Optimization	[31–35]
	Building level: Control	[36–46]
Systems models	Energy systems/Systems engineering with ML models	[47–51]

1.2. Component-based approach

On this basis, we propose a component-based approach using machine learning for the prediction of performance and the management of complexity arising in the design and planning of energy-efficient and sustainable buildings. The component is a subordinate model of a building part or its technology, such as walls, windows, roof, floor slab as well as heaters, chillers, etc. It is defined by input and output parameters. In the interest of rapid prediction, the component-based approach uses surrogate models (black box models) to connect these parameters, and combines them with the paradigms of systems engineering to manage the complexity that arises. In this paper, we develop a component-based model using ML.

We expect four advantages to ensue from the component-based approach:

- In contrast to the monolithic use of surrogate models, it allows the potential of system engineering to be exploited for complexity management.
- By building models for general components, such as walls, windows, roofs, etc., and zones, we expect a much greater degree of generalization, i.e., transferability to other new cases not included in the training model structure, something which is currently a problem for black box methods in building performance prediction in design.
- The component-based approach links very well to building information modeling (BIM) [64–66] as an upcoming method of future design and planning.
- The component-based models make quantities available in the analysis between components. This design-supporting insight is made available by a detailed performance simulation, but not by monolithic black-box models. It enables designers and engineers to analyze and understand systemic interdependencies much better.

Basically, from the perspective of machine learning, component-based ML is an engineering-based application of deep learning [67,68] that includes transfer learning, something which is discussed in detail in an accompanying paper [69]. This also includes multilevel models within components, which in turn depend on the availability of data or on the need to achieve more accurate predictions.

The component-based method has been developed for energy performance prediction. However, the method is applicable to all types of simulation results to build surrogate models and to form a systems model predicting multidisciplinary performance for design space exploration. Within systems that model schemes, the components form