Applied Energy 211 (2018) 89–103

Contents lists available at ScienceDirect

Applied Energy

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

A comparison of six metamodeling techniques applied to building performance simulations

Torben Østergård[b,a,*], Rasmus Lund Jensen[b], Steffen Enersen Maagaard[a]

[a] Aalborg University, Department of Civil Engineering, Thomas Manns Vej 23, DK-9220 Aalborg Ø, Denmark
[b] MOE Consulting Engineers, Mariane Thomsens Gade 1C, DK-8000 Aarhus, Denmark

HIGHLIGHTS

- Linear regression (OLS), support vector regression (SVR), regression splines (MARS).
- Random forest (RF), Gaussian processes (GPR), neural network (NN).
- Accuracy, time, interpretability, ease-of-use, model selection, and robustness.
- 13 problems modelled for 9 training set sizes spanning from 32 to 8192 simulations.
- Methodology for comparison using exhaustive grid searches and sensitivity analysis.

ARTICLE INFO

Keywords:
Gaussian process regression (kriging)
Random forest
Neural network
Support vector regression
Sensitivity analysis
Supervised learning

ABSTRACT

Building performance simulations (BPS) are used to test different designs and systems with the intention of reducing building costs and energy demand while ensuring a comfortable indoor climate. Unfortunately, software for BPS is computationally intensive. This makes it impractical to run thousands of simulations for sensitivity analysis and optimization. Worse yet, millions of simulations may be necessary for a thorough exploration of the high-dimensional design space formed by the many design parameters. This computational issue may be overcome by the creation of fast metamodels. In this paper, we aim to find suitable metamodeling techniques for diverse outputs from BPS. We consider five indicators of building performance and eight test problems for the comparison six popular metamodeling techniques – linear regression with ordinary least squares (OLS), random forest (RF), support vector regression (SVR), multivariate adaptive regression splines, Gaussian process regression (GPR), and neural network (NN). The methods are compared with respect to accuracy, efficiency, ease-of-use, robustness, and interpretability. To conduct a fair and in-depth comparison, a methodological approach is pursued using exhaustive grid searches for model selection assisted by sensitivity analysis. The comparison shows that GPR produces the most accurate metamodels, followed by NN and MARS. GPR is robust and easy to implement but becomes inefficient for large training sets compared to NN and MARS. A coefficient of determination, R^2, larger than 0.9 have been obtained for the BPS outputs using between 128 and 1024 training points. In contrast, accurate metamodels with R^2 values larger than 0.99 can be achieved for all eight test problems using only 32–256 training points.

1. Introduction

1.1. Motivation for metamodeling

The building sector accounts for roughly 40% of the total energy consumption and 38% of the CO_2 emissions in the European Union [1]. On a global scale, the energy savings potential is estimated to 53 exajoules each year by 2050 [2]. Building designers play a vital role in realizing this enormous energy savings potential. Architects and engineers use building performance simulations (BPS) to assess and reduce the environmental impact of buildings and, at the same time, meet strict requirements related to indoor climate. Examples of performance objectives are energy demand, CO_2 footprint, thermal comfort, daylight availability, and construction costs. To find possible solutions, the design team may vary a large number of design parameters such as building geometry, insulation thickness, glazing properties, and HVAC systems. The variations of these parameters constitute an enormous multi-dimensional “design space”. The many design parameters

* Corresponding author at: Aalborg University, Department of Civil Engineering, Thomas Manns Vej 23, DK-9220 Aalborg Ø, Denmark.
E-mail address: to@civil.aau.dk (T. Østergård).

https://doi.org/10.1016/j.apenergy.2017.10.102
Received 4 August 2017; Received in revised form 16 October 2017; Accepted 29 October 2017
0306-2619/ © 2017 Elsevier Ltd. All rights reserved.

(model inputs) and requirements (model outputs) makes it difficult and time-consuming to explore the design space efficiently and find favorable solutions that meet all requirements. To address this multivariate problem, it is becoming increasingly popular to perform a large number of simulations using Monte Carlo methods or optimization routines [3–7]. Unfortunately, most BPS software are computational demanding. A single simulation often takes minutes to compute or even hours in the case of CFD simulations. For real-life applications, the task of running thousands or millions of simulations is an obstacle for widespread adoption of design space exploration, uncertainty analysis, sensitivity analysis, and optimization.

The computational obstacle of BPS may be overcome by supercomputers, cloud computing, or metamodeling. Supercomputers are expensive if not managed efficiently to avoid downtime. Cloud computing is presumably a cheaper alternative, and several popular BPS tools provide this feature for optimization or uncertainty analysis [3]. Still, if the design team wish to perform sensitivity analysis to identify important inputs or interaction effects, such analysis easily requires thousands of simulations. This is likely to take hours or days – even with access to cloud computers [8,9]. An extensive set of Monte Carlo simulations allows the design team to explore a high-dimensional design space under various constraints and immediately observe the consequences of different design choices [10]. A useful tool for such analysis is the interactive parallel coordinate plot, which enables rapid and visual exploration of multivariate data (see Fig. 5) [11]. In terms of computational effort, Østergård et al. [10] demonstrated that 5000 simulations were insufficient when applying five constraints representing building legislation and architects' ambitions. Similarly, in the development of a design tool for thermal comfort evaluation, it was necessary to run millions of simulations to cover the design space sufficiently, even though the model only contained nine design parameters and two objectives [12]. Conclusively, this "curse of dimensionality" advocates the use of fast metamodels to overcome the computational challenges.

1.2. Performance requirements of metamodels

Above, we have argued for a potential of using metamodeling in the context of BPS. The type of metamodeling, addressed in this paper, is also referred to as supervised learning, which overlap with regression analysis. Hence, the metamodels are constructed from a set of input and output data to enable predictions of future outputs. That means we need to run a number of building simulations from which we construct the metamodel to be used for faster predictions of building performance. The metamodels may be constructed in a wide variety of ways using many different techniques. These may differ substantially with respect to predictive accuracy, computational efficiency, ease-of-use, transparency, and robustness (see Section 4.1). Unfortunately, there is no "perfect" and one-fits-all method. Thus, in searching for the most suitable metamodel for BPS, we have to describe the characteristics of our building simulations and define requirements, and optionally desirable features, of the metamodels. Emphasis is on applicability for novel users and the applied algorithms are all available with Matlab.

Building simulations are typically performed with complex "black box" models with many variable design parameters. Therefore, the users normally have no clear knowledge of the underlying equations and how the inputs interact with each other. In addition, the design team need to assess quite different building performance indicators, such as energy, thermal comfort, and daylight. These may vary substantially in complexity and in the shape of the output distribution (see Section 4.2). Thus, the metamodeling technique must be *robust* to the number of inputs and the type of output. We have identified three separate cases for which metamodeling can be applied to enable design space exploration, optimization, and sensitivity analysis. The three cases denoted A, B, and C can be distinguished by their diverse requirements, which may be characteristic for many other applications across scientific disciplines:

A. Expert with considerable time (~days) and emphasis on accuracy.
The task is to develop a generic, reusable tool for early design support based on a predefined room types that often occur in buildings, e.g. open-offices and meeting rooms. Time for training and construction of the metamodel is not critical.

B. Non-expert with limited time (~hours) and need for ease-of-use and robustness
For each building project, an engineer (or architect) with presumably limited knowledge of metamodeling needs to construct metamodels, which represent the particular building and its desired performance. The applied technique must be robust towards different objectives and easy to use.

C. Automated metamodeling requiring a minimum of training points (obtained in minutes) and high robustness
These conditions apply to the development of a joint CAD and BPS framework, in which the design team selects specific rooms in a CAD environment. For the selected room, BPS and metamodels are automatically performed with no user interaction and within a limited time frame.

For all cases, a large number of new predictions must be performed rapidly for real-time design-space exploration during meetings with multiple stakeholders [10].

The purpose of this study is to perform a comprehensive comparison of metamodeling techniques and thereby identify the techniques most suitable to accommodate the requirements listed above. We have strived for an extensive comparison by considering diverse building performance metrics and well-known mathematical test functions. The number of training data points has been varied substantially, i.e. from 2^5 to 2^{13} (32–8192). To improve transparency and reproducibility, we show all hyperparameter variations and data online [13]. In addition, we use the unit-less R^2 values to report accuracies, which makes it easier to compare the accuracies for diverse problems encountered in different disciplines. With these intensions, we strongly believe that this study is relevant to all research areas and industries, which apply metamodeling in the form of supervised learning.

2. Literature review

First, we investigate earlier uses of metamodeling in the context of building simulations. Afterwards, we look for promising metamodeling techniques based on comparisons made across scientific disciplines. Emphasis is on accurate, rapid techniques that can be applied by a practitioner of building performance simulations with limited knowledge of metamodeling. Moreover, non-Gaussian distributions of aggregated outputs (see Section 4.2) and interaction effects must be captured without knowing the underlying equations, which are mostly hidden in the commercial "black box" software.

2.1. Metamodeling in the field of building performance simulations

Metamodeling has been applied to building simulations for a variety of reasons, which include early design decision-making [14–17], uncertainty and sensitivity analysis [15,17–22], design optimization [16,18,19], and model calibration [21,23,24]. Most research addresses energy consumption, though metamodels have also been used to emulate thermal comfort [18,25,22], daylight [16,22], and financials costs [19]. A wide range of metamodeling techniques have been applied in the reviewed studies: linear regression (OLS) [14,5,22,24], polynomial regression (PR) [16,17,24], multivariate adaptive regression splines (MARS) [22], support vector regression (SVR) [18], neural network (NN) [20], and Gaussian processes regression (GPR) [21,25]. Additional methods, such as step-wise linear regression, decision trees (CART) and random forests (RF), have been found in works that compare metamodel methods in relation to BPS [26–28]. To sum up, a wide variety of metamodeling methods have been applied for diverse

Download English Version:

https://daneshyari.com/en/article/6681121

Download Persian Version:

https://daneshyari.com/article/6681121

Daneshyari.com