Automation in Construction 72 (2016) 279–293

Contents lists available at ScienceDirect

Automation in Construction

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

Data-driven approximation algorithms for rapid performance evaluation and optimization of civil structures

Stavros Tseranidis *, Nathan C. Brown, Caitlin T. Mueller

Building Technology Program, Department of Architecture, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

ARTICLE INFO

Article history:
Received 1 September 2015
Received in revised form 20 January 2016
Accepted 6 February 2016
Available online 24 March 2016

Keywords:
Surrogate modeling
Machine learning
Approximation
Structural design

ABSTRACT

This paper explores the use of data-driven approximation algorithms, often called surrogate modeling, in the early-stage design of structures. The use of surrogate models to rapidly evaluate design performance can lead to a more in-depth exploration of a design space and reduce computational time of optimization algorithms. While this approach has been widely developed and used in related disciplines such as aerospace engineering, there are few examples of its application in civil engineering. This paper focuses on the general use of surrogate modeling in the design of civil structures and examines six model types that span a wide range of characteristics. Original contributions include novel metrics and visualization techniques for understanding model error and a new robustness framework that accounts for variability in model comparison. These concepts are applied to a multi-objective case study of an airport terminal design that considers both structural material volume and operational energy consumption.

1. Introduction

The engineering design process for civil structures often requires computationally expensive analysis and simulation runs within a limited timeframe. When the assessment of a design's performance takes hours or days, the potential for exploring many solutions and significantly improving design quality is limited. This paper addresses these limitations by investigating the application of surrogate modeling, a data-driven approximation technique, to civil engineering in order to empower designers to achieve more efficient and innovative solutions through rapid performance evaluation.

1.1. Design optimization for civil structures

Optimization is an established and widely used method for high-performance engineering design. However, unlike in other engineering disciplines, the optimization objectives and constraints in civil and architectural structures such as buildings and bridges are not always easily quantified and expressed in equations. Instead, they often require human intuition and initiative to materialize. For this reason, design optimization in civil engineering has yet to reach its full potential. There are a few examples where it has been successfully applied to buildings, such as in braced frame systems for tall buildings by Skidmore, Owings, and Merrill (SOM) [1], but these cases remain exceptional. In contemporary architectural design, a wide variety of braced frame systems have been employed recently, including both the example in [1], as well as the more irregular design by Neil M. Denari Architects (High Line 23, New York City [2]). While High Line 23 is not structurally optimized, its architectural success is closely linked to the visual expression of its structural system and geometry. As demonstrated by these buildings, there often needs to be a balance between quantitative and qualitative objectives in design optimization.

Both quantitative and qualitative design goals pose challenges for the use of optimization in terms of computational speed. First, simulations such as structural analysis and predictions of building energy consumption often require significant computational power, increasing with the complexity and size of the project. Thus, optimization algorithms can require substantial execution time, slowing down or impeding the design process. On the other hand, the qualitative nature of civil engineering design, with its hard-to-quantify considerations as described above, can require many iterations, and the final design is unlikely to be selected from a single optimization run. In practice, the combination of problem formulation challenges and long simulation times means that optimization is rarely used in the design of architectural and civil structures. Even quantitatively comparing several design alternatives can be too time-consuming, resulting in limited exploration of the design space and likely a poorly performing design.

1.2. Need for computational speed

The exploration of a design problem and various optimal solutions should ideally happen in real time, so that the designer is most productive. Research has shown that rapid response time can result in significant productivity and economic gains [3]. The upper threshold for

---

* Corresponding author at: 77 Massachusetts Avenue, Room 5-418, Cambridge, MA 02139, USA. Tel.: +1 617 852 1737.
E-mail address: stser@mit.edu (S. Tseranidis).

http://dx.doi.org/10.1016/j.autcon.2016.02.002
0926-5805/

computer response time for optimal productivity has been estimated at 400 ms and is commonly referred to as the Doherty threshold [3]. This threshold was originally developed in the 1980s for system response of routine tasks such as typing. Today, however, many software users expect similarly rapid response for any interactions with the computer, even those that require expensive calculations like performance simulation. In addition, immediate response from the computer can benefit not just rote productivity but also stimulate creative thinking. The concept of flow is used in cognitive science to describe “completely focused motivation,” when a person becomes fully immersed in a task and is most productive and mentally engaged. Among the key requirements for achieving creative flow, as characterized by Csikszentmihalyi [4], is “immediate feedback.”

The first way to implement the “immediate feedback” effect in computer response is by increasing the available computational power. This can be achieved through either increased processing power or by harnessing parallel and distributed computing capabilities. The second way to generate immediate feedback is to use different or improved algorithms. This paper focuses on the second approach, investigating algorithms that improve computational speed for design-oriented simulation through approximation.

1.3. Surrogate modeling

Although a number of possible approximation algorithms exist, this paper considers surrogate modeling algorithms, a class of machine learning algorithms, and their use in making computation faster and enabling more productive exploration and optimization in the design of buildings. Machine learning typically deals with creating models about the physical world based only on available data. The data are either gathered through physical experiments and processes or by computer-generated samples. Those samples are then fit to an approximation mathematical model, which can then directly generate new representative data samples. In surrogate modeling specifically, the data are collected from simulations run on the computer.

Similar techniques are being implemented successfully in many other engineering disciplines but have not been studied and applied extensively in the fields of civil and architectural engineering. This research investigates surrogate modeling techniques and evaluates them on various related case studies. Focus is given to the development of a holistic framework that is generalizable. Fig. 1 displays a core concept in surrogate modeling. The circles represent the available data, with many different models being able to fit them. The art in surrogate modeling is to choose the one that will also fit new data well.

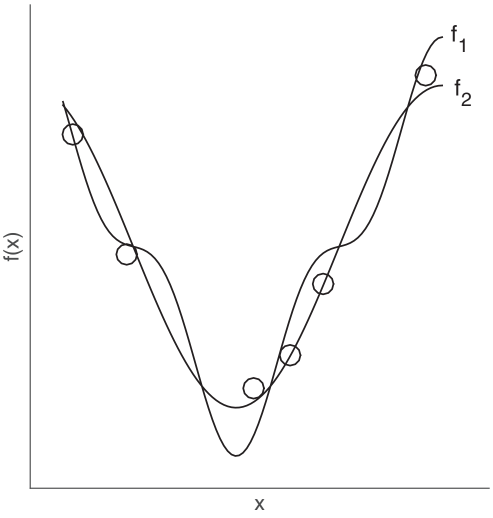

Fig. 1. “*Many surrogates may be consistent with the data*” (Figure inspired from [15]).

1.4. Big data approach

The recent surge of available information is reshaping the existing methodologies in several scientific and engineering fields. It has been argued that current methodologies should shift toward data-driven approaches. As described by Denning [5], engineers can build algorithms that can “recognize or predict patterns in data without understanding the meaning of the patterns”. Furthermore, Mayer-Schönberger and Cukier [6] analyze this shift from causality to correlation in the methodologies in the “Big Data” era. Therefore, the term *big data* extends beyond the generation and accumulation of large amounts of data to describe a new methodological paradigm based purely on data. Applications of this fascinating emerging field can be found in architecture and structural engineering. This paper presents a data-driven approach for the design exploration of architectural and civil structures. By harnessing powerful computational tools and rigorous methodologies based on the aforementioned ideas of Big Data science, designers are going to be able to explore concepts and ideas more rapidly and produce more diverse, cost-efficient, and high-performing structures.

1.5. Organization of paper

First, Section 2 outlines a literature review of the existing research in surrogate modeling, its application in structural engineering problems, model types, and error assessment methods. Section 3 gives an overview of the main features of the methodology framework, including assumptions and the models used. Sampling and normalization techniques are also discussed. Section 4 explains the new error assessment and visualization methods used throughout the rest of the paper. Section 5 introduces the proposed method for robust model comparison, which is then illustrated through the case study in Section 6. In turn, Section 6 introduces the case study problem, the parameters examined, and the analysis assumptions, while also presenting all the numerical results obtained from the approximation. The original contributions, findings, and future considerations are summarized in Section 7.

2. Literature review

To avoid a computationally expensive simulation, one historical approach has been to construct a physical model that is simpler and includes more assumptions than the original. This process is very difficult to automate and generalize, and it requires a high level of expertise and experience in the respective field. A more general approach is to substitute the analytical simulation with an approximate (surrogate) model that is constructed based purely on data. This approach is referred to as data-driven or black-box simulation. These terms signify that the constructed approximation model is invariant to the inner details of the actual simulation and analysis. The model has only “seen” data samples that have resulted from an unknown process, thus the name black-box. This paper addresses data-driven surrogate modeling.

The two main areas in which surrogate modeling can be applied are optimization and design space exploration. Specifically, an approximation model can be constructed as the main evaluation function for an optimization routine, or in order to explore a certain design space in its entirety to better understand variable trade-offs and performance sensitivity. For optimization, surrogate modeling is often used when more than one optimization objective is present, thus called multi-objective optimization (MOO), and the computational cost of computing them is significant.

2.1. Surrogate modeling for structural designs

In previous decades, when computational power was significantly less than that of today, scientists began exploring the possibility of adapting approximation model techniques in intensive engineering