Data Driven Design of Load Bearing Components

Report

on

Bachelor of Technology Project Thesis (Semester V)

Submitted

By

Devanuj Baruah (180103023) Atharva Malwadkar (180103015)

DEPARTMENT OF MECHANICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

(Nov 2020)

Bachelor's Thesis Project- Semester V Report

Project Title: Data Driven Design of Load Bearing Components

Students: 1. Devanuj Baruah (180103023)

2. Atharva Malwadkar (180103015)

Supervisors: Prof. Debabrata Chakraborty and Prof. S.M. Hazarika

Table of Contents

1. Introduction

Design, in engineering, generally refers to problem solving process. A proposed design provides solutions to some problem leading to a product or process. Engineering design is the process of devising a system, component or process to meet desired needs, using our knowledge of mathematics and natural sciences.

In mechanical design, the focus is mainly on the design of load bearing components with an objective of determining optimum size, shape and material ensuring safety of the component under the load. Therefore, historically, a mechanical design process has always been undertaken using relevant constitutive models. These constitutive models have been important for structural design and material certification, and are traditionally achieved through mostly uniaxial tests aiming in the determination of elastic material properties primarily.

These models are the mathematical descriptions of relations between physical quantities which don't follow directly from physical laws. Example: how materials respond to various loadings. These relations include the elasticity-plasticity theory, solution of non-linear equilibrium equations etc. Analytic solutions are replaced by numerical methods in the modern age due to their computational prowess and accuracy.

However, constitutive models have their own limitations. When dealing with complex problems for systems with high dimensional design variables, these models are complex and slow to process. In certain instances, the design domain is far too big to be processed by a constitutive model. The nonexistence of knowledge of a design problem also makes getting to the correct constitutive model almost impossible. Constitutive models bring with them material modelling empiricism, modelling error and uncertainty.

2. Motivation

Traditional experiments and computational modelling not only consume tremendous time and resources but also are limited by their experimental conditions and theoretical foundations leading to uncertainties in the design. Especially for new and novel materials it becomes more so important to develop new methods of accelerating the discovery and design process.

Our motivation is to propose Data-Driven design as a way to eliminate these epistemic uncertainties linked to traditional constitutive models in mechanical design. By eschewing empirical models, data driven design can eliminate modelling error and uncertainty, and no loss of experimental information is incurred. A model free data driven approach can not only eliminate the explicit postulation of a specific constitutive model, it also generates databases of material states respecting some important sampling. Further, for orthosis and prostheses, the designs that are data –driven are more engaging and tailored to users' preferences. Figure 1 shows the important steps in a data driven design where a large number of relevant data is collected and processed to be used as the model (guiding the behavior of the component instead of constitutive model) for design.

Figure 1. Data driven design

3. Literature Review

In the recent years, data driven design in in general has gain momentum in research and a number of research works have been published in open literature in the broad area of data driven design. Some of those have been discussed here to understand the state of the art in the area of data driven design in general and its potential application in mechanical design in particular.

Liu et al. [1] presented a detailed review of applying machine learning in materials science where the main algorithms have been classified and compared. They have also reviewed the current research status with regard to applications of machine learning in material property prediction, in new materials discovery. Finally, they discussed the problems related to machine learning in materials science and proposed possible solutions along with the potential directions of future research.

Charlotte et al. [2] addressed the problem of identifying an osteoarthritic glenoid from a wide range of variability among population. The main focus of their work was to create a model with existing mathematical and statistical modeling approaches to present the structure of a Glenoid in a different way to critically analyze it and contribute to the counter the complications in Total Surgical Arthroplasty. A set of landmark locations were selected to represent the topology of the glenoid. These landmark locations were transformed into a PC frame of reference to find the variability that spans across glenoid geometries in intact shoulders. By manipulating the PC scores, a suite of shoulder geometries was created within the observed variability from existing shoulders.

Hongzhan et al. [3] discuss in detail about the transformation of product usage data, including product time-dependent performance feature data and field data, into valuable information to guide product design improvement, so that design and manufacturing engineers can better identify design defectives and improve product reliability. With embedded sensors, the performance and the working field data is collected from the product. To prevent severe and rapid product performance degradation, the existing products are improved based on the gathered information along with its time-dependent performance. A case study is done to modify the design parameters and redesign a Large Tonnage Crawler Crane (LTCC). Due to the complexity to process the product usage data, design improvement in the proposed approach is conducted in three main steps:

1. The data collected from various operational parameters have a large number of features. For efficient degradation assessment and to reduce processing time feature reduction is done using *Kernel Principal Component Analysis* (KPCA).
2. One of the primary difficulties to effectively assess performance degradation is the highly stochastic nature of the collected data such as multimodal and nonlinear time-dependent performance features. *Gaussian Mixed Model* (GMM) is used for modeling complicated distributions of performance feature data.
3. Both the methods mentioned above are used for performance degradation assessment. To identify the abnormal field data which contribute to the performance degradation, *Data Clustering* is done to identify to-be modified design parameters.

To demonstrate the effectiveness of the proposed model, the results are compared with Failure Mode and Effects Analysis (FMEA) and Function-Failure Design Method (FFDM) approach which uses the failure data. Although the functions identified by the integrated approach was not exactly similar, but we can identify the critical parameters to be re-designed just from the performance feature data.

Du and Zhu [4] undertook a study with the aim of developing a novel methodology for system with high-dimensional design variables, based on data mining theory. It is implemented through designing a crashworthy passenger car, which is a multi-level (system – components) complicated system. The design alternatives designed by Finite Element analysis when subjected to pre-defined constraints. To achieve an efficient and effective design, methods to identify inter-relation between design parameters such as load path, structural deformations, component collapse sequence, weight of the vehicle, geometry, etc from a complex high dimensional data set. To overcome the issues related high dimensional data, two strategies are employed, namely:

1. Critical Parameter Identification (CPI), to identify those design variables which have significant effect on the system performance.
2. Design Domain Reduction (DDR), to reduce the range of values of the variables to further decrease the design space.

Decision tree technique was used to mine the crash simulation datasets to identify the key design parameter with most significant effect (CPI) on the vehicular energy absorption response and determine the range of their values (DDR) simultaneously. In this way, the further design can focus on these “high impact” variables with smaller range of values. Without CPI and DDR, the simulation dataset would be of huge size and the computational cost is not affordable. To evaluate the performance of this new method, the conventional Response Surface Method (RSM) based optimization was done on the original as well as the new reduced design data. The result obtained from both were compared with each other. It was concluded that the new approach can be used, as long as it is used to design a high dimensional system, over the conventional method with efficiency and effectiveness.

Stainier et al [5] propose an integrated model-free data-driven approach to solid mechanics and show how principles and techniques of data driven computational mechanics leads to elimination of epistemic uncertainty linked to traditional constitutive models. Qualitative as well as quantitative results obtained using a data-driven approach were more accurate. The methodology proposed in the paper is to generate a material database using the Data-Driven Identification method, which is used to solve the classical boundary value problem. This is done by employing a distance minimization algorithm to find a compatible mechanical state (stress-strain field) in equilibrium. Acquired data is used to train a model to predict responses from different geometric structures of the same material. This method thus eliminates the postulation of a constitutive model entirely, thus getting rid of its uncertainties.

Tien-Thinh Le [6] discusses a study where a surrogate Machine Learning (ML)-based model was developed, to predict the load-bearing capacity (LBC) of concrete-filled steel square hollow section (CFSS) members, considering loading eccentricity. The proposed Artificial Neural Network (ANN) model was trained and validated against experimental data using the following error measurement criteria: coefficient of determination (R2), slope of regression, root mean square error (RMSE) and mean absolute error (MAE). A parametric study was conducted to calibrate the parameters of the ANN model, including the number of neurons, activation function, cost function and training algorithm, respectively. Sensitivity analysis showed that the geometric parameters of the steel tube (width and thickness) and the compressive strength of concrete were the most important variables. Finally, the effect of eccentric loading on the LBC of CFSS members is presented and discussed, showing that the ANN model can assist in the creation of continuous LBC maps, within the ranges of input variables adopted in this study. The results showed that the ANN model can provide reliable and effective prediction of LBC (R2 = 0.975, Slope = 0.975, RMSE = 294.424 kN and MAE = 191.878 kN)

Zhang et al. [7] presents a systematic function recommendation process to suggest new functions to an existing product and service. Different from the conventional approaches where new functions are largely formulated by experienced designers, the proposed approach builds upon recommendation systems that dynamically catch the trendy requirements from targeted users that are not recognized by existing product and service yet. A detailed case study reveals the merits of the proposed approach

4. Gap in Literature and Objectives of the present work

Literature review reveals that data driven design methodology has been an important research area in recent years especially with the availability of high capacity digital computers. However, there has been a scope of using data driven design in mechanical design in particular. With this back ground, this project proposes to explore machine learning algorithms to put forward an integrated model-free data-driven approach to mechanical design problems, without postulating a specific constitutive model. Traditional design problems will be addressed, particularly those lacking in well-established constitutive models and attempt will be made to design analyze using data driven approach.

5. Methodology

In this project, the aim at hand is to explore machine learning techniques for parameter identification and a model-free material data identification, resulting in a machine element design methodology. We shall now discuss such relevant algorithms and processes.

5.1 Critical Parameter Identification

In mechanical systems with high dimensional design variables, to achieve an effective and efficient design it is imperative to identify complex coupling effects of the numerous inter-dependent variables. At such instances when constitutive models fail, data-driven design is employed to resolve the aforementioned issue. Critical parameters identification (CPI) followed by design domain reduction can be applied. CPI is to identify those design variables which have a strong effect on the system performance. Based on these key variables, design domain reduction is then performed quantitatively to reduce the range of values of the variables to further decrease the design space.

For critical parameter identification, a machine learning algorithm is employed where most of the time a decision tree – based data mining approach is applied to the engineering product design. Decision tree is a tree-like graph to identify the key variables in a large design space or dataset, and classify their effect on the system performance. Following this, according to the design objectives, the range of the values of these key parameters can be determined quantitatively. In this way, the further design can focus on these “high impact” variables with a smaller range of values.

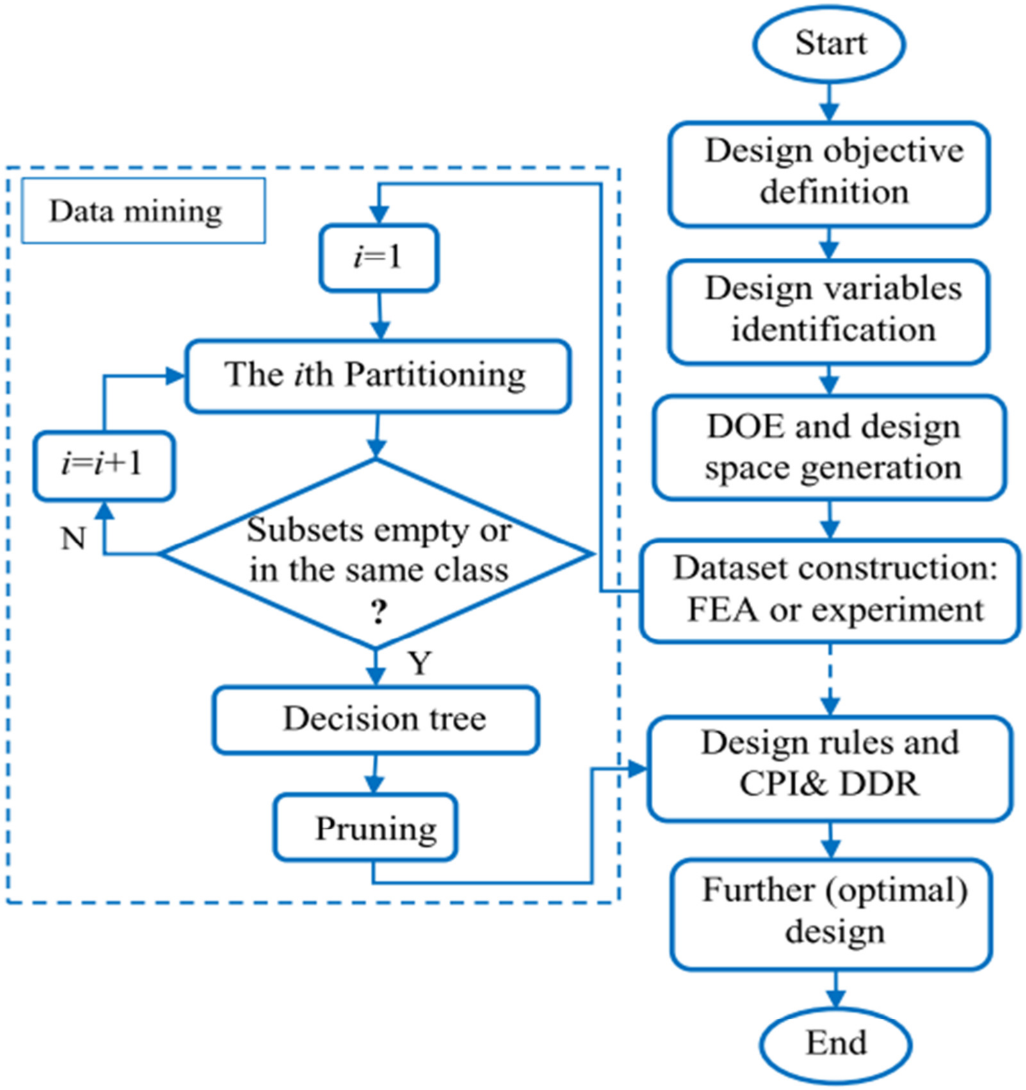

Figure 2. Design of a system with high dimensional design variables

Figure 2 showcases the flowchart of the data mining method to design a system with high-dimensional design variables. In the first step, the design problem is analyzed and the design objective is determined. Likewise, the design variables are further identified and the range of their values are determined. Then, in the design space, a large number of design of experiments (DOEs) are built. Each design is computed through FEA and the simulation results form a large design dataset. After that, data mining is performed on the simulation dataset to construct a decision tree, with which CPI and DDR can be implemented simultaneously to reduce the design space. Likewise, useful inter-relationships among design variables and decision making rules can be derived. Based on this information and reduced design space, the further optimal design can be achieved faster and at lower cost.

In the data mining process, the recursive partitioning on the simulation dataset (the initial design space) is conducted to construct the decision tree. The gain ratio of entropy is first calculated for the whole dataset to determine the first (i = 1) partitioning. If all subsets after this operation are empty or belong to the same class, the decision tree construction is complete. Otherwise, a new round of partitioning is performed (i = i + 1) on the impure subsets and each of these impure subsets includes more than one class. The procedure is repeated in this recursive way until all subsets are empty or in the same class. After the decision tree is built, it is pruned to obtain more concise results. Using this decision tree, design rules are derived, and then CPI and DDR are carried out to achieve the final design.

5.2 Model-free Material Data Identification

For classical problems of mechanical state analysis of load bearing members, we see high discrepancies while solving for the initial-condition boundary-value problems, for different geometries of the same material. Such uncertainties can be successfully removed by completely bypassing the empirical material modelling step altogether. We propose to achieve that through a model-free data driven technique, which would be used for the identification of the material data-set. Using ML algorithms such as K-means, for each material point we converge on a mechanical state under equilibrium.

Our aim is to subsequently apply a Euclidean distance minimization algorithm to converge on a compatible material state space (of stress-strain), thus formulating a boundary value problem directly in terms of material data.

Material behavior is known through a material data set $\boldsymbol{D_e}$ of points obtained experimentally or by some other means as

$$z_e = (\varepsilon_e, \sigma_e) \in Z_e \tag{1}$$

We regard $z_e = (\varepsilon_e, \sigma_e)$ as a point in a local phase space $\boldsymbol{Z_e}$ -

$$\boldsymbol{Z_e = R^{me} \times R^{me}} \tag{2}$$

and

$$z = \{(\varepsilon_e, \sigma_e)\}^{M}_{e=1} \tag{3}$$

as a point in the global phase space $\boldsymbol{Z = Z_1 \times\cdots\times Z_M}$. These phase spaces correspond to stress and strain fields, which are our primary point of interest in analyzing a load-bearing member. A class of Data-Driven problems consists of finding the compatible and equilibrated internal state $z \in \boldsymbol{E}$ that minimizes the distance to the global material data set $\boldsymbol{D = D_1 \times\cdots\times D_M}$, the associated distance '$\boldsymbol{d}$' being-

$$d\,(z, y) = |z - y| \quad \text{for } y, z \in Z \tag{4}$$

The distance-minimizing Data-Driven problem then becomes:

$$\min_{y \in D}\ \min_{z \in E}\ d(z, y) = \min_{z \in E}\ \min_{y \in D}\ d(z, y) \tag{5}$$

i.e., we wish to find the point $y \in \boldsymbol{D}$ in the material data set that is closest to the constraint set $\boldsymbol{E}$ of compatible and equilibrated internal states or, equivalently, we wish to find the compatible and equilibrated internal state $z \in \boldsymbol{E}$ that is closest to the material data set $\boldsymbol{D}$.

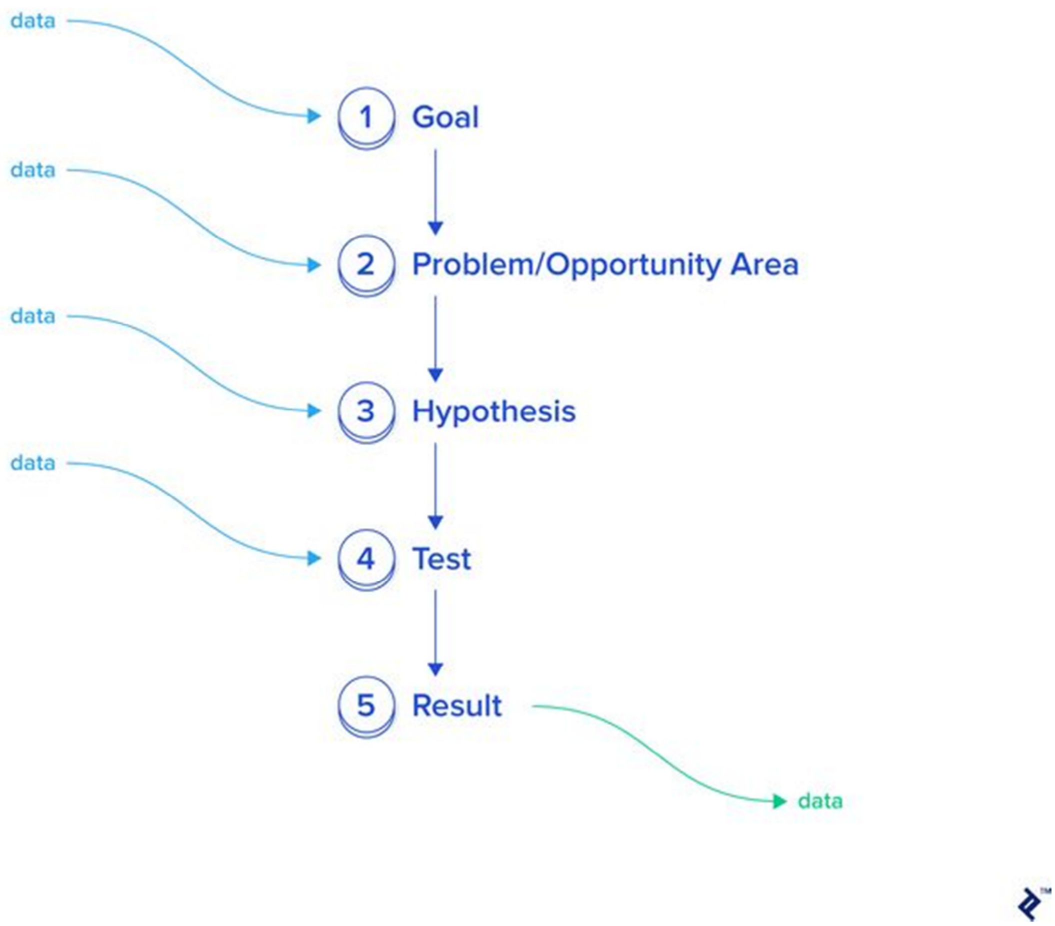

Figure 3 Data flow in design

In this discussion we have seen how a data-driven approach to design is devoid of any form of traditional modelling. Instead, as shown in Fig. 3 data is sequentially used to process and give design insight. The actualization of a goal, identification of the and formulation of the hypothesis, use algorithms to manipulate experimental or synthetic data, which provides us design optimization based on data analysis, rather than having to invoke empirical relations and conventional models.

The aforementioned algorithms will be critically evaluated under limited constraints and checking their viability when expanded to a typical design problem. Proper parameter identification and domain reduction are important for a seamless evaluation of a methodology.

6. Conclusion

Literature review has been performed to understand the data driven design concept in general and to understand the state of the art in the data driven design approach with special focus on the data driven mechanical design. From the thorough review of literature, the potential of data driven methodology as more efficient approach in obtaining the desired solutions bypassing the constitutive models could be observed. It is clear that such methodologies can be applied for numerous other mechanical design problems where the constitutive models have limitations in the form of not producing results under our desired constraints.

7. Work Ahead

Next Semester:

1 Generate large amount of synthetic data for a mechanical design problem using FEM
2 Apply data driven methodology to extract with force deformation response.
3 Compare different algorithms to understand the best methodology
4 Formulate a precise design methodology by the relevant integration of aforementioned algorithms, suitably devised to meet the design requirements

Based on the outcome of the next semester, details work plan for next two semesters (7^{th} and 8^{th}) will be planned.

8. REFERENCES

[1] Yue Liu, Tianlu Zhao, Wang wei,Ju, Siqi Shi, “Materials discovery and design using machine learning”, Journal of Materiomics, Volume 3, Issue 3, September 2017, Pages 159-1

[2] Charlotte M. de, Matthew B. Parkinson, “Modeling the Variability of Glenoid Geometry in Intact and Osteoarthritic Shoulders” [DOI: 10.1115/1.4037408]

[3] Ma Hongzhan, Xuening Chu, Guolin Lyu, Deyi Xue, “An Integrated Approach for Design Improvement Based on Analysis of Time-Dependent Product Usage Data” July 2017 Journal of Mechanical Design 139(11), DOI: 10.1115/1.4037246

[4] Xianping Du, Feng Zhu, “A new data-driven design methodology for mechanical systems with high dimensional design variables”, Department of Mechanical Engineering, Embry-Riddle Aeronautical University, Daytona Beach, FL 32114, USA

[5] Laurent Stainier, Adrien Leygue, Michael Ortiz, “Model-free data-driven methods in mechanics: material data identification and solvers” Published online: 4 June 2019, © Springer-Verlag GmbH Germany

[6] Tien-Thinh Le, “Surrogate Neural Network Model for Prediction of Load-Bearing Capacity of CFSS Members Considering Loading Eccentricity” Published: 16 May 2020

[7] Zhinan Zhang, Ling Liu, Wei Wei, Fei Tao, Tianmeng Li, Ang Liu, "A Systematic Function Recommendation Process for Data-Driven Product and Service Design", *J. Mech. Des*. Nov 2017