Design of ultra-thin shell structures in the stochastic post-buckling range using Bayesian machine learning and optimization

M.A. Bessa[a], S. Pellegrino[a,*]

[a] *Graduate Aerospace Laboratories, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA*

Abstract

A data-driven computational framework combining Bayesian machine learning for imperfection-sensitive quantities of interest, uncertainty quantification and multi-objective optimization is developed for the analysis and design of structures. The framework is used to design ultra-thin carbon fiber deployable shells subjected to two bending conditions. Significant increases in the ultimate buckling loads are shown to be possible, with potential gains on the order of 100% as compared to a previously proposed design. The key to this result is the existence of a large load reserve capability after the initial bifurcation point and well into the post-buckling range that can be effectively explored by the data-driven approach. The computational strategy here presented is general and can be applied to different problems in structural and materials design, with the potential of finding relevant designs within high-dimensional spaces.

Keywords: ultra-thin composite structures, buckling, post-buckling, design charts, data mining, heteroscedastic Gaussian process, evolutionary optimization

*Corresponding author
Email address: sergiop@caltech.edu (S. Pellegrino)

1. Introduction

The recent resurgence of interest in buckling of slender structures (Hu and Burgueño, 2015; Reis, 2015) can be largely explained by the advent of new manufacturing techniques (Kalpakjian et al., 2014) for complex shapes, as well as improved modeling capabilities (Benson et al., 2010; Bai et al., 2015) that leverage the extensive theoretical understanding of buckling (Hutchinson and Koiter, 1970; Bažant and Cedolin, 2010). These developments have spawned a myriad of creative solutions for a wide range of applications such as energy harvesters (Chen et al., 2010; Wang et al., 2014), sensors (Elvin et al., 2006; Kiyono et al., 2012), dampers and absorbers (Dong and Lakes, 2013; Kim et al., 2013), actuators (Loukaides et al., 2014; Lazarus and Reis, 2015), morphing structures (Diaconu et al., 2008; Daynes et al., 2014), and deployable structures (Pellegrino, 2014; Mallikarachchi and Pellegrino, 2014), all of which exploit the geometrically non-linear behavior of thin shell structures.

These structures are designed to capture the benefits of the first bifurcation point (initial buckling). In these cases, post-buckling is mostly viewed as a sudden behavior that leads to large configuration changes, often occurring as a form of snap-through or to a lesser extent snap-back. Hence, post-buckling behavior and subsequent bifurcations are not usually a design target (Hu and Burgueño, 2015).

However, there are structures where the post-buckling behavior assumes particular importance. Thin-walled structures provide some of the most relevant examples, as can be seen in studies showing the effects of cutouts in composite shells (Tafreshi, 2002) or of nonuniform wall thickness in steel shells (Aghajari et al., 2006), as well as investigations on functionally graded carbon nanotube-reinforced shells undergoing thermal post-buckling (Shen, 2012) and functionally graded shallow plates (Woo et al., 2005). Recently, Leclerc et al. (2017) noted that an ultra-thin composite Triangular Rollable And Collapsible (TRAC) boom is able to carry significantly increased loads well into the post-buckling regime. The article here presented focuses on designing TRAC booms to improve their buckling and post-buckling behavior through a data-driven computational framework.

In data-driven approaches (Bisagni and Lanzi, 2002; Yvonnet and He, 2007; Ning and Pellegrino, 2015; Bessa et al., 2017) a new model or design is found by collecting enough data about the response of the structure or material under multiple input conditions. In principle, data can be collected by experimental testing, analytical or computational predictions. Yet, in most engineering applications experimental characterization of previously untested designs is too time-consuming to gather enough data in a timely manner, and most applications are too complex to be predicted by probabilistic analytical models (Elishakoff, 2014). Hence, computational predictions are often the only viable resource to explore the design space and generate enough data to use machine learning and/or optimization.

Multiple authors have been exploring the use of data-driven approaches for different scientific and engineering applications. Notable examples exist in computer science with artificial intelligence algorithms that master the game of GO (Silver et al., 2016), materials science with data mining of first principle calculations leading to discoveries of new material compounds (Curtarolo et al., 2003; Fischer et al., 2006; Saal et al., 2013; Gautier et al., 2015), fluid mechanics in the characterization of flows with high Reynolds numbers (Ling and Templeton, 2015), and different solid mechanics applications (Bisagni and Lanzi, 2002; Yvonnet and He, 2007; Bessa et al., 2017).

This article extends a recently developed data-driven framework for materials and structures (Bessa et al., 2017) to the design of optimized structures with uncertain response. The proposed approach is illustrated for the design of TRAC booms, in order to increase their ultimate buckling limit, but it can be applied to any other structure or material. The extended framework includes two significant contributions: 1) machine learning for noisy observations with uncertainty quantification; and 2) introducing a multi-objective optimization step after the machine learning procedure to determine optimal designs. The first extension is crucial for the design and analysis of imperfection-sensitive structures (with noisy or uncertain response). The second extension is relevant when the goal is not just establishing the relationship between input design descriptors and output performance of the structure (or material), but also to find the set of input descriptors that leads to an optimal response within given constraints.

The paper is laid out as follows. Section 2 presents the ultra-thin composite TRAC boom structure to which the data-driven framework is applied. Section 3 discusses the data-driven framework for noiseless applications in 3.1, and for noisy applications with multi-objective optimization goals in 3.2. Concluding remarks are included in Section 4. Details on the structural imperfections of TRAC booms are provided in Appendix A, and a discussion of discontinuous Pareto frontiers is given in Appendix B.

2. Behavior of ultra-thin TRAC booms

Figure 1 shows a schematic of a TRAC boom, partially wrapped on a spool. This type of structure was initially developed by Murphey and Banik (2011) and can be viewed as two tape springs bonded along one edge and thus forming a flat region (web) with twice the thickness of the flanges. In the packaged configuration, the TRAC boom is flattened (flat cross-section) and wrapped around a spool of radius R. The deployed geometry is then fully characterized by the TRAC boom length L and its cross-section parameters: web height h (thickness $2t$), flange radius r, angle θ, and thickness t.

Figure 2 shows the predictions obtained from nonlinear finite element analyses using the arc length method to determine the post-buckling response of an ultra-thin composite TRAC boom manufactured by Leclerc et al. (2017). The structure was subjected to two separate boundary conditions: (a) bending moment around X leading to compression at the outer edge of the web; and (b) bending moment around Y. The same nominal geometric parameters reported by Leclerc et al. (2017) are used herein: total length $L = 504$ mm, and cross-section parameters $r = 10.6$ mm, $\theta = 105^o$, and $h = 8$ mm. The material

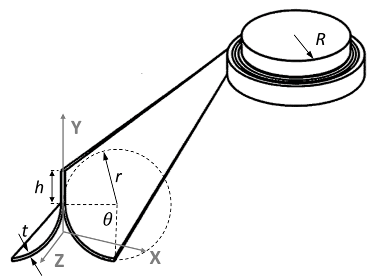
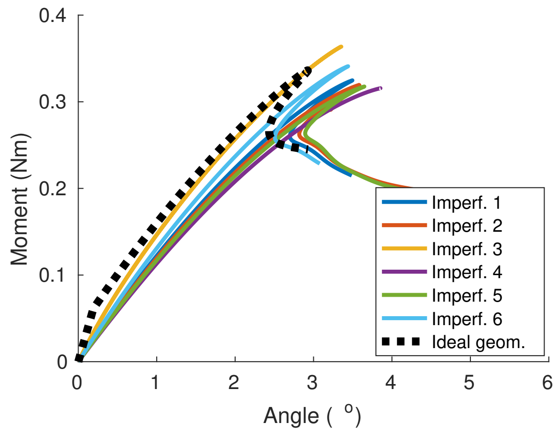


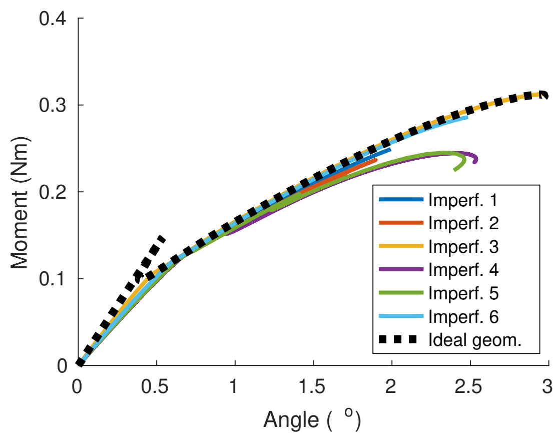
Figure 1: Schematic of the TRAC boom architecture (modified from Murphey and Banik (2011)).

is a composite laminate with stacking sequence $[0^o, 90^o]_S$ and nominal post-cure thickness of $t = 71\ \mu$m, where the four composite plies are stacked from a 17GSM unidirectional tape supplied by North Thin Ply Technology (T800 fibers and ThinPreg 120EPHTg-402 epoxy resin). The orthotropic elastic properties of each ply are considered as $E_1 = 128.0$ GPa, $E_2 = 6.5$ GPa, $\nu_{12} = 0.35$, $G_{12} = G_{13} = G_{23} = 7.5$ GPa.

The dashed lines in Figure 2 show the responses predicted for the idealized geometry, where a negligibly small imperfection based on the first buckling mode was seeded to numerically resolve the first bifurcation point. Details about the six imperfect cases shown in the figure are discussed later in Appendix A. The focus at this point should be on noting that due to the ultra-thin nature of the structure the first bifurcation point occurs prematurely for both loading conditions, but is followed by a further, significant increase in the applied moment until the ultimate buckling limit of the structure is reached. Here the *ultimate buckling limit* is defined as the analytical maximum in the moment-angle response of the structure. Due to the complexity and stochasticity of the buckling and post-buckling behavior of the TRAC boom, the particular geometry considered in Figure 2 is likely not optimal for given quantities of interest, e.g. maximizing both buckling limits. Since closed form solutions to find such optimal geometries do not exist, the viable alternative is to use a data-driven approach applicable to stochastic responses.



(a) Response for applied moment around X



(b) Response for applied moment around Y

Figure 2: Responses predicted for a TRAC boom with $h = 8$ mm and $\theta = 105^{\circ}$, and considering the first 6 imperfection points shown in Figure A.1.

3. Data-driven computational framework

The data-driven framework conjugates the following steps: 1) Design of Experiments[1] (DoE); 2) computational analyses; 3) machine learning; and 4) multi-objective optimization. The following subsections discuss the application of the framework to the design of TRAC booms, for two different cases. A simplified case where the machine learning process does not need to be probabilistic, and where the obtained design charts may be easily interpreted without an optimization step. This case represents an application of the previously developed framework (Bessa et al., 2017) without the proposed extensions. The other case illustrates the need for a Bayesian machine learning process including uncertainty quantification, as well as the advantage of introducing an optimization step to find particular designs after machine learning. Each step is summarized in Figure 3 and explained in the following sections for each of the above mentioned cases.

3.1. A design case without uncertainty nor optimization: initial buckling of idealized TRAC booms

This first case focuses on finding the combination of cross-sectional parameters h and θ that maximize the initial buckling moments of the TRAC boom considering idealized geometries, i.e. without imperfections. For this first example only the first bifurcation point occuring at the end of the linear elastic regime is of interest – recall Figure 2 and observe the first kink in the dashed line for each boundary condition applied.

In this investigation the length $L = 504$ mm and thickness $t = 71$ μm are kept constant, and the same composite material described previously with stacking sequence $[0^{\mathrm{o}}, 90^{\mathrm{o}}]_{\mathrm{S}}$ is considered. The volume (mass) of the structure is also kept

[1] The term "Design of Experiments" (DoE) is extensively used in the literature (Morris and Mitchell, 1995) and refers to finding an optimal sampling of the design space without *a priori* knowledge of the regions of interest. Note that "experiments" does not only refer to physical experiments but also to "computational experiments".

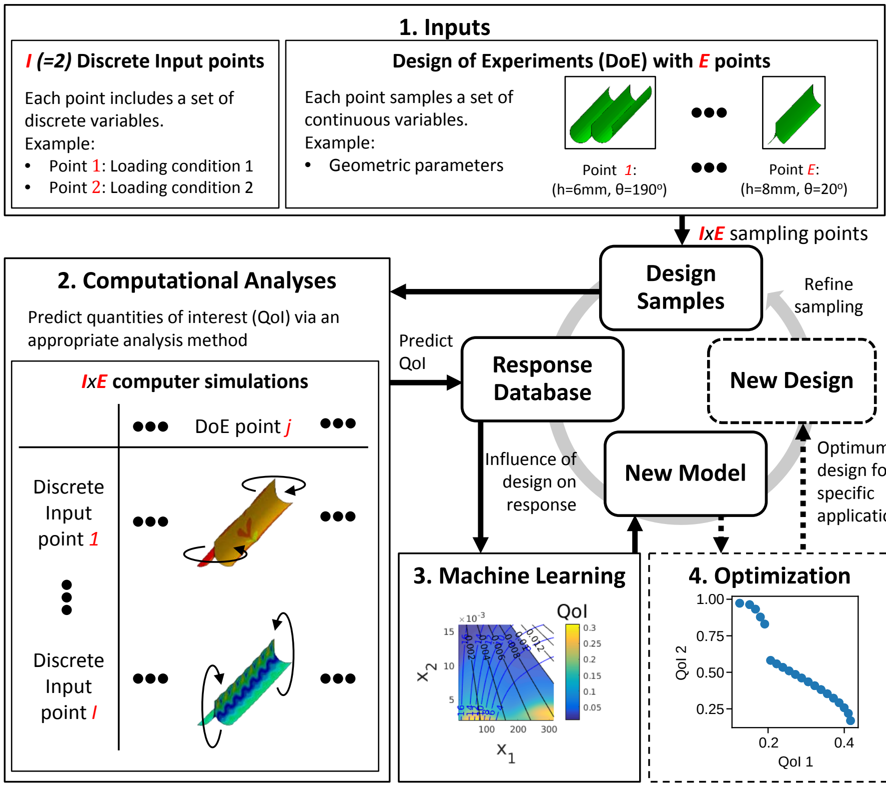


Figure 3: Data-driven framework applied to structural modeling and design (modified from Bessa et al. (2017)).

constant for all TRAC booms, in order to establish a fair comparison between different geometries. This volume constraint introduces a relationship between the radius of the flanges r and the two independent parameters considered in this study, h and θ:

$$r = \frac{V - 2thL}{2\theta tL} \tag{1}$$

with $V \approx 1963$ mm^3, $t = 71$ μm, and $L = 504$ mm, while the two design descriptors θ (in radians, above) and h assume different values for different cross-sections.

An additional constraint is introduced due to the need of flattening the boom for packaging. This causes a transverse strain $\varepsilon_{\text{flat}}$ (Murphey et al., 2017):

$$\varepsilon_{\text{flat}} = \frac{t}{2r} \tag{2}$$

that typically needs to be below 1% to avoid failure of the composite material. The strain caused by wrapping the flattened structure around a spool could also be included as a constraint, but this strain is usually not the limiting factor since the radius of the spool R used to wrap the boom is large.

The first step (DoE) in the data-driven framework is the sampling of the design space without *a priori* knowledge of the relationship between the input variables and the output quantities of interest. Different methods can be used for the DoE, with two common options being the Sobol sequence (Sobol, 1967, 1976) and the Latin Hypercube sampling (McKay et al., 1979). According to previous investigations (Bessa et al., 2017), the Sobol sequence is used herein. For this study the bounds for the input design variables h and θ are defined as:

$$h = [2, 16] \text{ mm} \quad , \quad \theta = [10^{\text{o}}, 315^{\text{o}}] \tag{3}$$

The Sobol sequence DoE method produces a nonuniform space-filling design where different hyperplane projections do not lead to coincident points – properties that have been shown to facilitate the machine learning process (Simpson et al., 2001). Figure 4 presents the 1,000 DoE points obtained from a random Sobol sequence of the two design descriptors.

The subsequent step in the data-driven framework after the DoE is the computational analysis of each DoE point. Note that variations in the geometry of the structure are expected to lead to competing effects for the two boundary conditions applied. For example adding/removing material in the web has significant impact on the bending stiffness and buckling of the structure for bending around X, but less impact for bending around Y because the web is in the neutral plane in this case. However, this influence is less intuitive when considering structures with constant volume, because the material that is added in the web needs to be respectively removed in the flanges, which alters the

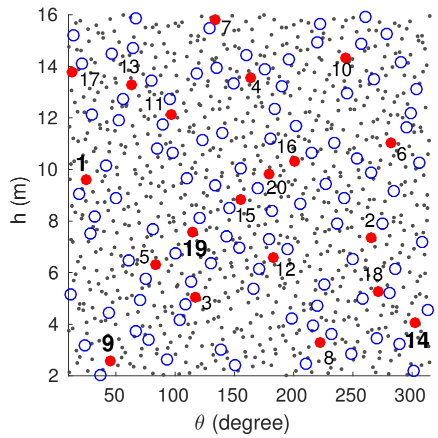


Figure 4: Design of Experiments to determine the influence of the TRAC boom geometries on the initial buckling behavior of the ultra-thin composite structure. The solid red circles correspond to the first 20 points of the Sobol sequence and are labeled with the corresponding sequence number. The black dots correspond to the remaining 880 points, and the last 100 points are shown as blue circumferential markers.

bending stiffness and local curvature of the structure. Recall that the actual influence of r, θ and h is not trivial to predict analytically due to the fact that the structures are ultra-thin, leading to highly localized buckling modes instead of global modes.

The following procedure is automated for computationally predicting the first bifurcation point of the structures:

1. Linear bifurcation analysis of the undeformed configuration of the structure. This provides an initial prediction for the first bifurcation point and buckling modes;

2. Static analysis under displacement control without stabilization until the simulation stops or the previously determined bifurcation point is reached;

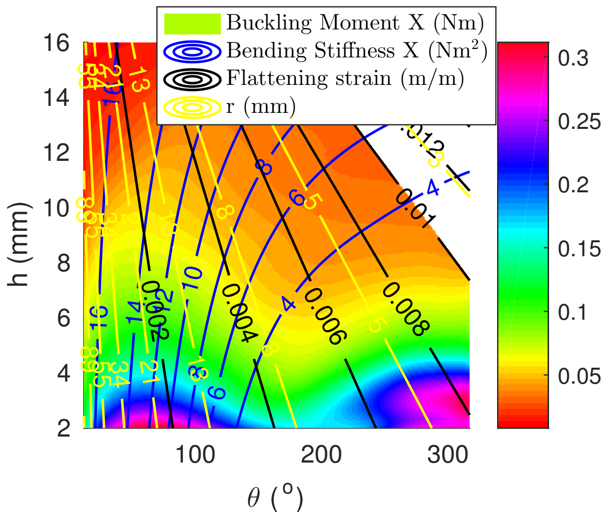
3. Sequential linear bifurcation analyses starting from the last available increment of the static analysis until reaching an increment where the bi-

furcation point can be predicted, i.e. an increment where the structure is not on the post-buckling regime but where it is close to the bifurcation point.

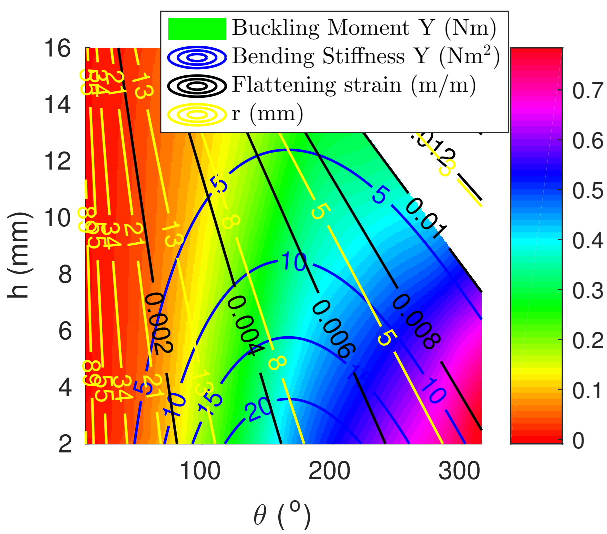
The outlined procedure was implemented such that it could be performed automatically and in parallel for different input DoE points. The output quantities of interest obtained from the simulations, e.g. moments and angles at the two ends, are stored in a database that is then used for the machine learning process.

The machine learning process for this subsection is simplified because the quantities of interest have negligible noise (no geometric imperfections). In this article small databases are appropriate for all the applications considered, favoring the use of the Gaussian process method (Krige, 1951; Matheron, 1963) as opposed to methods such as artificial neural networks (Rosenblatt, 1958; Widrow and Hoff, 1960; Hopfield, 1982; Rumelhart et al., 1986). This was shown in (Bessa et al., 2017) with a comparison of the data-driven framework for higher dimensional design spaces using both Gaussian process (also known as kriging) and neural networks. Gaussian process regression is discussed in detail in Section 3.2.1 where the less common case of noisy observations is included. The interested reader can consult the work of Rasmussen and Williams (2005) for additional details. If interested in artificial neural networks, the reader is referred to Demuth et al. (2014).

Construction of design charts for the quantities of interest is then a trivial outcome of the machine learning process. Figure 5 shows the variation of the first bifurcation point of the perfect structures with different cross-sections. The figure also includes contour lines of the bending stiffness, radius of the flanges, and the transverse strain $\varepsilon_{\text{flat}}$ caused by flattening the structure for packaging – see equation (2). The score based on the mean squared error of the predictions of the last 100 points of the Sobol sequence (see Figure 4) is evidently very high (above 0.99) because this is a simple regression problem (low-dimensional), unlike previously considered cases (Bessa et al., 2017). In fact, the score is still



(a) Initial buckling moment in X



(b) Initial buckling moment in Y

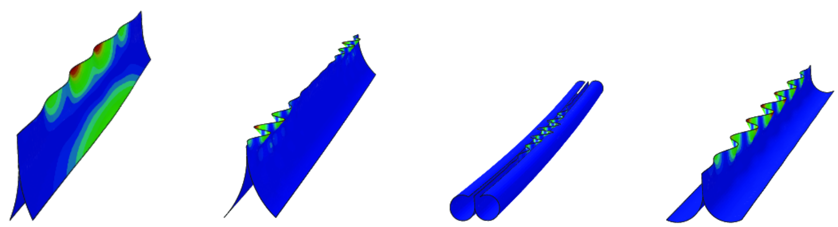
Figure 5: Design charts obtained for the variation of the initial buckling moments in X and Y as a function of two cross-section parameters (h and θ).

above 0.98 when considering 300 DoE points for the learning task instead of 900.

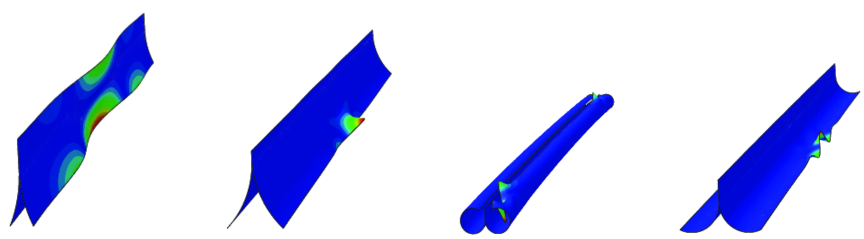
Observing Figure 5 one can identify that the maximum buckling moment does not occur for regions with maximum bending stiffness for both boundary conditions. This behavior occurs due to the presence of local buckling modes. From the figure it is also clear that constraining the flattening strain below 1% does not significantly limit the design space. Interestingly, there is a common region of the design space that maximizes the initial buckling moment for both loading conditions at large flange angles and small web heights, despite the bending stiffness along X being low at that region.

Figure 6 shows the buckling modes corresponding to 4 DoE points shown in Figure 4. Focusing first on the modes obtained by bending the structures around the X axis (top row of the figure), one can see that DoE point 9 (Figure 6b), and DoE point 14 (Figure 6c) show localized buckling modes, as opposed to the more global modes seen around the top of the web for the other geometries. The buckling behavior is improved in point 9 by localizing the deformation at the ends of the structure, while for point 14 it is improved by localizing it in the center after a more compliant behavior in this bending condition. Focusing now on the modes obtained by bending the structure around the Y axis (bottom row of the figure), one can see that only point 14 (Figure 6g) shows the localization of deformation at the ends of the structure, justifying the fact that only this geometry is in the region of the design chart with a higher buckling moment in Y. Therefore, DoE point 14 corresponds to the best design of the 4 geometries shown in Figure 6, and is within the optimal region found in the design charts.

In general, one can conclude that designs with large flange angles and small web heights are the most effective when aiming to increase both initial buckling loads for constant mass of the structure. If the design goal is different, then the design charts can be used to optimize the structure for other applications, e.g. maximum bending stiffness achieved for a desired minimum buckling strength.



(a) DoE point 1 for bending around X **(b)** DoE point 9 for bending around X **(c)** DoE point 14 for bending around X **(d)** DoE point 19 for bending around X



(e) DoE point 1 for bending around Y **(f)** DoE point 9 for bending around Y **(g)** DoE point 14 for bending around Y **(h)** DoE point 19 for bending around Y

Figure 6: First buckling mode observed for different geometries subjected to two loading conditions: bending around X (top row); and bending around Y (bottom row). From left to right the geometries shown correspond to DoE points 1, 9, 14 and 19 shown in Figure 4. The modes are scaled such that the maximum displacement component is 5 mm for all cases.

3.2. A design case with uncertainty and optimization: imperfection-sensitive TRAC boom ultimate buckling

The previous approach of designing idealized structures for maximizing the first bifurcation point included three important simplifications: 1) each computational analysis was relatively inexpensive, allowing the generation of a sufficiently large database; 2) the observations were noiseless, facilitating the machine learning process; and 3) there was no need for including an optimization algorithm after the machine learning process due to the low-dimensionality of the problem and the obvious position of the optimum in the design charts. Nonetheless, there are problems where these simplifications are not possible. This subsection discusses the design of imperfect ultra-thin TRAC booms targeting improved ultimate buckling moments for the two loading conditions considered above.

Ultimate buckling of ultra-thin carbon fiber TRAC booms occurs after post-buckling, as shown in Figure 2. The computational analyses required in this case have a higher computational cost (each simulation requires approximately 4 CPU-hours). This limits the exploration of the design space – see Remark 1.

Remark 1. *Data-driven approaches are limited by the generation of a sufficiently large database through efficient computational analyses – Box 2 in Figure 3. Most engineering applications require costly computational analyses to generate each data point. Therefore, new reduced order models (Ladevèze et al., 2010; Chinesta et al., 2011; Liu et al., 2016) that efficiently and accurately simulate such applications are one of the most critical needs in data-driven approaches for new designs. However, nonlinear buckling and post-buckling phenomena are yet to be demonstrated to be efficiently and accurately predicted by a reduced order model, to the authors' knowledge.*

An additional challenge to apply the data-driven framework is in dealing with uncertainty, i.e. when the quantities of interest are stochastic such as imperfection-sensitive phenomena. Post-buckling and subsequent ultimate buckling of a structure are often strongly dependent on the presence of imperfections.

This is visible in Figure 2. Furthermore, the imperfection-sensitivity is not usually the same for different geometries of the structure. Therefore, characterizing the uncertainty for the entire design space involves answering a two-part question: what imperfections and how many imperfect samples per DoE point should be considered?

In particular to the ultra-thin composite TRAC booms considered herein, there is currently not enough experimental evidence to devise a strategy to model realistic imperfections – see Remark 2. The adopted strategy is then to seed geometric imperfections in the form of combinations of buckling modes obtained from linear perturbation analysis of the first bifurcation point. This strategy has been successfully used by different authors, e.g. Riks (1979); Bisagni (2000); Ning and Pellegrino (2015).

Remark 2. *Modeling geometric imperfections can be performed with higher-fidelity for problems with additional experimental information. For these cases, imperfections can be modeled directly from three-dimensional scans of the geometry or by finding combinations of frequency and buckling modes until the local imperfections are statistically equivalent to the scanned geometries.*

The number of buckling modes used to seed imperfections is estimated by the proximity of the first eigenvalues determined from the nonlinear buckling analyses of the previous section. The first two eigenvalues were found to be similar for a large number of TRAC boom designs, hence the first two buckling modes were chosen to seed the imperfections. An estimation of the amplitudes of these modes can be extracted from a preliminary experimental investigation (Leclerc et al., 2017) of various specimens of a single ultra-thin TRAC boom design, reporting deviations from the idealized structure as large as 2 mm with clear localized kinks. The amplitude bounds for the first two buckling modes were then considered as:

$$\lambda_1 = [-30t, 30t] \quad , \quad \lambda_2 = [-30t, 30t] \tag{4}$$

where $t = 71\ \mu$m, as previously referred.

Specific details of the approach used to quantify the uncertainty of each TRAC boom design are included in Appendix A. Based on the initial design considered in Figure 2 it was determined that the first 20 imperfection sampling points highlighted in Figure A.1 would provide a reasonable estimate for the uncertainty of each design point. This means that the same combination of buckling mode amplitudes were used to generate 20 imperfect TRAC booms per nominal design. This does not mean that the same geometric imperfections are being seeded because each nominal geometry of the TRAC boom has different first and second buckling modes, as seen in the previous section. The statistical distribution of the quantities of interest (ultimate buckling moments) was approximated as Gaussian – see Appendix A. Hence, the two metrics needed to approximate the distribution at each design point are the mean and standard deviation of the ultimate buckling moments computed from the 20 imperfect TRAC booms for each nominal design.

Including uncertainty quantification for this example increases 20 times the number of simulations needed to subsequently perform the machine learning process. Also, each analysis of the post-buckling and ultimate buckling of these structures has a considerably higher computational cost when compared to the analyses used to determine the first bifurcation points, as referred previously. Thus, the bounds of the DoE were shrunk to:

$$h = [2, 6] \text{ mm} \quad , \quad \theta = [10^{\text{o}}, 315^{\text{o}}] \tag{5}$$

so that the number of DoE points could be reduced to 150, while keeping a similar density of points in the same region of the design space that led to a good approximation of the response in the previous example. The same method was used to perform the DoE as in Figure 4, so no additional figure is shown here.

The database with the quantities of interest is created by conducting the respective computational analyses, which consist of the finite element solution of an arc length method using a commercial finite element software. A total of 6000 computational analyses were conducted (2 boundary conditions, and 20 imperfect TRAC booms for each of the 150 DoE points defining the different nominal designs). This corresponds to approximately 3 weeks of continuous computations in a computer cluster using 48 CPUs on average.

Subsequently, the machine learning process for these noisy observations is required to establish the relationship between the uncertain ultimate buckling moment and the input variables defining the nominal geometry of the TRAC booms. The specific method used in this article is the Gaussian process regression, as alluded in the previous section and detailed next.

3.2.1. Gaussian process regression with heteroscedastic noise

Contrary to the noiseless cases discussed in previous illustrations of the data-driven framework – see Section 3.1 and previous materials design examples (Bessa et al., 2017) – the ultimate buckling moments of the TRAC booms are strongly imperfection-sensitive, as seen in Figure 2 and in Appendix A. Moreover, different nominal designs of TRAC booms are expected to have different imperfection sensitivity, as widely reported in the buckling literature, e.g. Ning and Pellegrino (2015). In statistics this phenomenon is called *heteroscedasticity* (Goldberg et al., 1998), i.e. input-dependent noise (variance).

When possible, considering noiseless cases or at least considering independent and identically distributed (i.i.d.) Gaussian noise with constant variance is attractive. In these cases the Gaussian process regression has fast implementations because the likelihood and marginal likelihood can be integrated analytically (Rasmussen and Williams, 2005). However, a general treatment where noise is allowed to be heteroscedastic (input-dependent) and/or non-Gaussian requires numerical integration of the likelihood and marginal likelihood.

Gaussian process regression (GPR) with heteroscedasticity was first discussed by Goldberg et al. (Goldberg et al., 1998), while GPR with non-Gaussian noise was introduced by Neal (1997). Both articles considered a fully Bayesian framework using Markov chain Monte Carlo to perform the integration of the likelihood and the marginal likelihood. Unfortunately, this approach is computationally demanding, so alternative solutions have been proposed (Snelson et al., 2004; Kersting et al., 2007; Titsias and Lázaro-Gredilla, 2011). Snelson et al. (2004) proposed a warping scheme for Gaussian processes, Kersting et al. (2007) considered the most likely value using a maximum *a posteriori* (MAP) approach, while Titsias and Lázaro-Gredilla (2011) presented a variational approach with only twice the computational expense of a standard Gaussian process.

Here the implementation provided by Pedregosa et al. (2011) is followed, where noise is assumed to be i.i.d. Gaussian but allowed to be heteroscedastic. The Gaussian assumption has implications illustrated in Figure A.3 for the TRAC boom design problem, since the statistical distribution for the ultimate buckling moment around X is not Gaussian if using the Sobol sequence to seed imperfections. However, considering the current lack of experimental data for this problem and the reported computational limitations arising from the non-Gaussian noise treatment, this simplifying assumption is considered henceforth.

The previously determined database with $N_{DoE} = 150$ DoE points is given by:

$$\{(\mathbf{x}^{(1)}, \mathbf{q}^{(1)}), ..., (\mathbf{x}^{(I)}, \mathbf{q}^{(I)}),\ \dots, (\mathbf{x}^{(N_{DoE})}, \mathbf{q}^{(N_{DoE})})\} \tag{6}$$

or in index notation,

$$\{(x_j^{(1)}, q_i^{(1)}), ..., (x_j^{(I)}, q_i^{(I)}),\ \dots, (x_j^{(N_{DoE})}, q_i^{(N_{DoE})})\} \quad \text{for } j = 1, ..., d_{in} \text{ and } i = 1, ..., d_{out} \tag{7}$$

where d_{out} corresponds to the number of output quantities of interest $\mathbf{q}$, and d_{in} is the number of input design variables $\mathbf{x}$. In the TRAC boom design example there are $d_{in} = 2$ inputs $x_1 \equiv \theta$ and $x_2 \equiv h$, and $d_{out} = 2$ outputs $q_1 \equiv M_X^f$ and $q_2 \equiv M_Y^f$ corresponding to the ultimate buckling moments obtained for applied bending around X and Y, respectively.

Separate Gaussian processes will be used for each output. This way, the notation can be simplified by dropping the index in q_i because q can be separately associated to quantity q_1 or q_2. In the heteroscedastic Gaussian process considered here the noisy observations are approximated as,

$$q^{(I)} = f[\mathbf{x}^{(I)}] + \epsilon^{(I)} \tag{8}$$

where $f[\mathbf{x}^{(I)}]$ is the unknown function value at $\mathbf{x}^{(I)}$ to be approximated by the Gaussian process, and $\epsilon^{(I)}$ is the additive i.i.d. Gaussian noise with standard deviation $\sigma_\epsilon^{(I)}$ that also depends on the input point $\mathbf{x}^{(I)}$.

The Gaussian process establishes its foundations by defining a prior that depends on the proximity of the data points weighted by a kernel function k with noise added to its diagonal,

$$\mathrm{cov}[q^{(I)}, q^{(J)}] = k[\mathbf{x}^{(I)}, \mathbf{x}^{(J)}] + \left(\sigma_\epsilon^{(I)}\right)^2 \delta_{IJ} \tag{9}$$

that can be arranged as a sum of two matrices,

$$\mathrm{cov}[\mathbf{q}] = \mathbf{K} + \mathbf{R} \tag{10}$$

where $\mathbf{K}$ is called the kernel matrix or covariance matrix with element (I, J) as $K_{IJ} = k[\mathbf{x}^{(I)}, \mathbf{x}^{(J)}]$, and $\mathbf{R}$ is the diagonal noise matrix with each term including the variance of the quantity of interest q at the respective input point. This representation of noise can be viewed as a type of Tikhonov regularization (Tikhonov, 1963).

If one desires to predict the quantity of interest $q^{(*)}$ at a new input point $\mathbf{x}^{(*)}$, then the Gaussian process is written as a multivariate Gaussian distribution:

$$\begin{bmatrix} \mathbf{q} \\ \hat{q}^{(*)} \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{K} & \mathbf{k}_* \\ \mathbf{k}_*^T & k[\mathbf{x}^{(*)}, \mathbf{x}^{(*)}] \end{bmatrix}\right) \tag{11}$$

where $\mathbf{k}^T = \{k(\mathbf{x}^{(1)}, \mathbf{x}^{(*)}), ..., \mathbf{x}^{(N_{DoE})}, \mathbf{x}^{*})\}$ is the vector of kernel functions evaluated at all the pairs composed by the N_{DoE} training points and the new point $\mathbf{x}^{(*)}$.

The predicted mean and variance of the quantity of interest $q^{(*)}$ at the new point is then written as:

$$\text{mean}[\hat{q}^{(*)}] = \mathbf{k}_*^{\mathrm{T}} \left(\mathbf{K} + \mathbf{R}\right)^{-1} \mathbf{q} \tag{12}$$

$$\text{cov}[\hat{q}^{(*)}] = k[\mathbf{x}^{(*)}, \mathbf{x}^{(*)}] - \mathbf{k}_*^{\mathrm{T}} \left(\mathbf{K} + \mathbf{R}\right)^{-1} \mathbf{k}_* \tag{13}$$

In this article the kernel function k is chosen to be the squared exponential not just due to its ubiquitous use in the literature (Goldberg et al., 1998; Neal, 1997; Rasmussen and Williams, 2005; Bessa et al., 2017) but also due to the fact that the Tikhonov regularization is directly related to the variance at the input values used for training. The squared exponential kernel function is then written as:

$$k[\mathbf{x}^{(I)}, \mathbf{x}^{(J)}] = \eta^2 \exp\left[\sum_{k=1}^{d_{in}} -\frac{1}{2\lambda_k^2}(x_k^{(I)} - x_k^{(J)})^2\right] \tag{14}$$

where η and λ_k are called hyperparameters to be determined by integration of the likelihood and marginal likelihood.

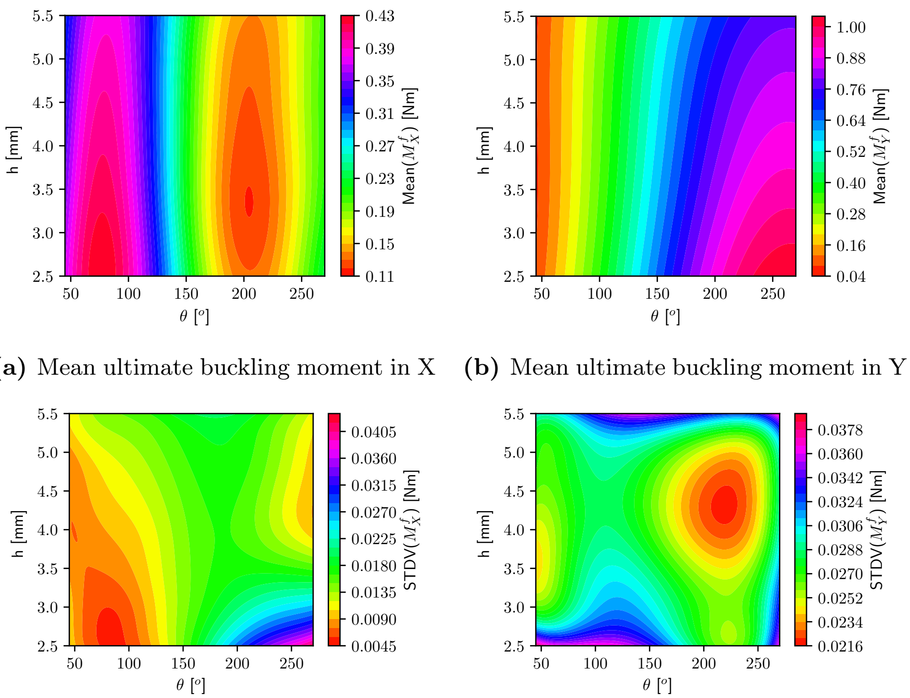
Under the assumption that the prior is i.i.d Gaussian and heteroscedastic the likelihood and marginal likelihood can still be integrated exactly. The exact integration of the marginal likelihood leads to the log marginal likelihood:

$$\log\left(p[\mathbf{q}|\mathbf{X}, \mathbf{\Xi}]\right) = -\frac{1}{2}\mathbf{q}^{\mathrm{T}} \left(\mathbf{K} + \mathbf{R}\right)^{-1} \mathbf{q} - \frac{1}{2}\log|\mathbf{K} + \mathbf{R}| - \frac{N_{DoE}}{2}\log(2\pi) \tag{15}$$

where $\mathbf{X}$ is the matrix with all the input points $\mathbf{x}^{(I)}$, and $\mathbf{\Xi}$ is the vector with all the hyperparameters considered in the Gaussian process.

Finally, finding the values of the hyperparameters is achieved by maximizing the log marginal likelihood. In this article that maximization is performed using the Limited Broyden-Fletcher-Goldfarb-Shanno algorithm for Bound constraints (L-BFGS-B) (Zhu et al., 1997).

The predicted mean and standard variation of the ultimate buckling moments obtained when bending the TRAC booms around X or Y are shown in Figure 7. Note that even though the GPR was conducted for the bounds given



(a) Mean ultimate buckling moment in X

(b) Mean ultimate buckling moment in Y

(c) Standard deviation of ultimate buckling moment in X

(d) Standard deviation of ultimate buckling moment in Y

Figure 7: Distribution of mean and standard deviation of the ultimate buckling moments obtained by bending the structure around X and Y.

in equation (5), the bounds of the figure were shrunk for two reasons: 1) self-contact was not included in the simulations, so the TRAC booms with very large flange angles could suffer from artificial inter-penetration of surfaces; and 2) uncertainty typically increases close to the boundaries of the design space due to the decrease in point density in those regions, so by excluding the region close to the boundaries one is certain that there are no artificial increases of uncertainty. In general shrinking the bounds is not necessary, but the authors wanted to reinforce that the increase in uncertainty for high angles and high web heights seen in Figure 7d is not explained by boundary effects.

3.2.2. Optimization

Machine learning allows the creation of a non-parametric model that can be used for regression as shown in the previous sections. These non-parametric relationships between inputs and outputs can be the end goal for several applications, such as finding the constitutive behavior of heterogeneous materials (Yvonnet and He, 2007; Bessa et al., 2017), or even when looking for global optima in low dimensional data where the maxima or minima are in obvious locations, as seen in Section 3.1. However, in other applications the goal is not creating a model relating inputs and outputs, but finding non-obvious optima for the quantities of interest. In these cases, optimization algorithms need to be considered.

A myriad of optimization strategies is available in the literature. Typically they can be classified as derivative-based or derivative-free optimization methods. Derivative-based methods usually have faster convergence, see e.g. Zhu et al. (1997); Gill et al. (2005). Derivative-free methods are mainly reserved for applications where direct access to the derivatives of the quantities of interest does not exist or when the derivatives are just too time-consuming to compute (Zitzler and Thiele, 1999; Zitzler et al., 2000; Konak et al., 2006).

Optimization of the ultimate buckling moments of TRAC booms is an example where derivatives are unavailable (the derivatives are only available after machine learning). The TRAC boom problem also includes additional complications: each computational analysis is expensive, and the observations are noisy. Under these conditions direct derivative-free optimization is therefore prohibitively expensive, and surrogate-based modeling becomes particularly advantageous. In the latter approach a machine learning model is developed from a limited number of computational analysis of the process of interest and a subsequent optimization step is conducted. The interested reader on surrogate-based analysis and optimization is referred to the following literature reviews and respective references (Queipo et al., 2005; Forrester and Keane, 2009).

Including machine learning in the data-driven approach to optimization provides a framework for uncertainty quantification, facilitates optimization of noisy observations, allows clear data visualization and sensitivity analyses, as well as quick evaluation of any response and respective derivatives within the bounded design space. One of the most important advantages is also the possibility of performing various optimizations for different applications of the same structure or material. For example, one may be interested in maximizing the TRAC boom ultimate buckling moments for a given application, but for other cases one may be interested in maximizing the initial and post-buckling stiffnesses, or the toughness, among other possibilities. This can be achieved by running multiple optimization problems on the same machine learning model, without performing additional computational analysis.

The challenge in this surrogate-based optimization approach is in keeping a balance between exploration and exploitation (Forrester and Keane, 2009). Exploration meaning sampling refinement (more DoE points) to improve the accuracy of the machine learning model, and exploitation meaning the fast evaluation of the surrogate model to find its global optimum, at the expense that the true global optimum of the physical process may not necessarily coincide with the surrogate.

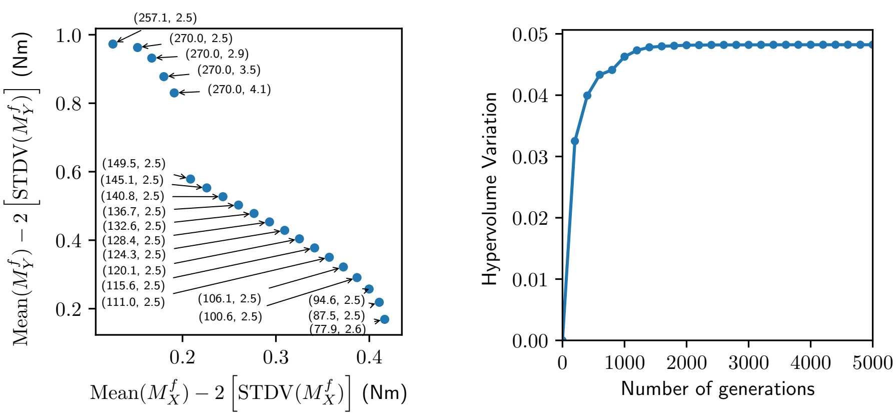
Optimization applied to the the previously determined GPR model shown in Figure 7 can be done by a large class of optimization algorithms. For example, the L-BFGS-B algorithm used previously for finding the maximum of the log marginal likelihood function could be easily applied here. Nevertheless, in order to establish a fair comparison between the surrogate-based optimization and the results one would obtain by directly performing the optimization task without machine learning, derivative-free optimization algorithms are considered instead.

Evolutionary multi-objective optimization algorithms have been a major research area in the past two decades, as can be observed from multiple comprehensive reviews (Zitzler and Thiele, 1999; Zitzler et al., 2000; Konak et al., 2006). Recently the topic of many-objective optimization (Ishibuchi et al., 2008) is gaining particular attention in the literature in an attempt to solve problems where more than four objectives are simultaneously being optimized (dealing with issues related to visualization of the Pareto frontier, convergence, curse of dimensionality, etc.).

In this article, the authors focused on simultaneously maximizing the two imperfection-sensitive ultimate buckling moments of TRAC booms. Within the possible evolutionary algorithms adequate to perform this task, the authors selected four state-of-the-art algorithms as implemented by Izzo (2012) and initialized with the recommended default parameters: the Strength Pareto Evolutionary Algorithm 2 (SPEA-2) by Zitzler et al. (2001), the Non-dominated Sorting Genetic Algorithm II (NSGA-II) by Deb et al. (2002), the Non-dominated Sorting Particle Swarm Optimizer for Multi-objective Optimization (NSPSO) by Li (2003), and the more recent $\mathcal{S}$ Metric Selection Evolutionary Multi-objective Optimization Algorithm (SMS-EMOA) by Beume et al. (2007) that is appropriate for many-objective optimization problems (more than 4 objectives) due to improved scalability.

The TRAC boom design problem is low-dimensional and the decision space is relatively small, so it is expected that all 4 algorithms should be able to determine the Pareto frontier without difficulties. The calculation of the Pareto frontier hypervolume (Beume et al., 2009) relative to a reference point located at $M_X^f = 0$ Nm and $M_Y^f = 0$ Nm was used as convergence metric. Note that in this application the hypervolume is in fact an area since there are only two objectives (area obtained from the origin of the decision space delimited by the Pareto frontier – see Figure 8). By definition, maximizing the Pareto frontier leads to its successive growth until full convergence is reached; thus, the hypervolume (area) tends to increase as the number of generations increases. As an additional note, the efficient calculation of the hypervolume for high-dimensional Pareto frontiers is currently still an active area of research (Bader and Zitzler, 2011).

Figure 8a shows the Pareto frontier obtained from the SMS-EMOA with an initial population of 20 individuals after 5000 generations. This Pareto frontier was computed considering the mean value of the ultimate buckling moments



(a) Pareto frontier

(b) Convergence metric

Figure 8: Figure (a) shows the Pareto frontier obtained from SMS-EMOA by maximizing the lower bound of the 95% confidence interval of the ultimate buckling moments in X and Y. The labels refer to the input points located in the Pareto frontier (θ, h). Figure (b) presents the hypervolume variation with increasing number of generations. Markers in (b) represent the hypervolume every 200 generations, and the hypervolume variation is calculated by comparing the current value to the initial hypervolume of 0.22745 $(\mathrm{Nm})^2$.

minus two standard deviations, in order to penalize designs sensitive to imperfections. In other words, the lower bound of the 95% confidence interval of the ultimate buckling moments is considered for optimization – see Figure A.3.

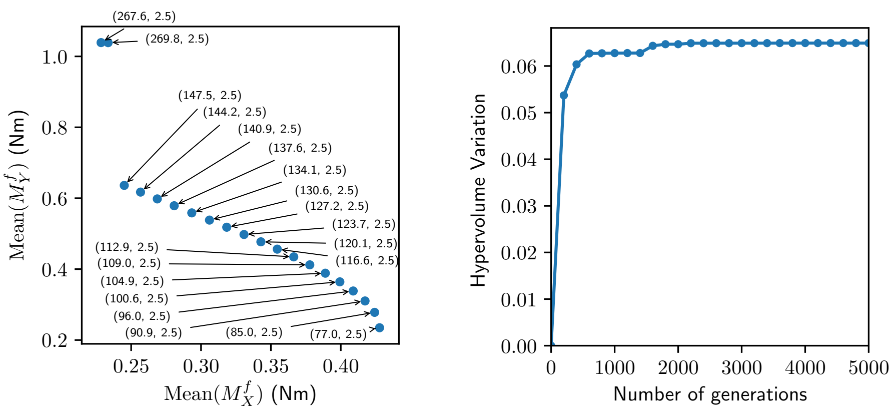
Figure 8b shows the variation of the hypervolume (area) as the number of generations increases compared to the hypervolume computed from the initial population of individuals. From this figure one can observe that after approximately 1000 generations the result has converged, which corresponds to a total of $40,000$ function evaluations. This large number of function evaluations would be the approximate number of actual computational simulations of the TRAC booms[2] required to determine the Pareto frontier if the surrogate model was not used. Clearly this number of simulations would be infeasable considering the computational expense of each simulation, so the need for the machine learning step is evident.

The Pareto frontiers obtained from the other algorithms are similar, so the respective figures are not included in this article. The SPEA-2 showed excellent performance requiring only 100 generations of an initial population of 20 individuals, also leading to a uniform distribution of the objective functions values as in Figure 8a. This means that "only" $4,000$ function evaluations were needed, which is still significantly higher than the 150 DoE points used for machine learning. This algorithm was already expected to be more efficient than the SMS-EMOA for two-dimensional decision spaces, but its performance degrades for higher dimensional ones (Ishibuchi et al., 2008), so the authors preferred to focus on the most general algorithm.

The NSGA-II algorithm showed similar convergence to the SPEA-2 but it required longer times to compute each generation and had the additional disadvantage that the quantities of interest located in the Pareto frontier were not so uniformly distributed. Finally, the NSPSO showed even slower convergence than

[2]In fact, since the observations are noisy the number would be higher because the estimation of the standard deviation is needed at each point! Alternatively, one could use fuzzy optimization algorithms, although they have other limitations not discussed herein.

the SMS-EMOA and highest nonuniformity than the NSGA-II for the Pareto frontier sampling. For this last algorithm a larger initial population (50) was needed due to the nonuniformity of the Pareto frontier sampling. Uniformity in sampling the Pareto frontier is desirable because it facilitates the hypervolume estimation and the identification of possible discontinuities as the one observed in Figure 8a.



(a) Pareto frontier of mean values

(b) Convergence metric

Figure 9: Similar to Figure 8 but considering the mean ultimate buckling moments as maximization objectives instead of their lower bound estimates. In figure (b) markers are also placed every 200 generations, and the initial hypervolume value is 0.27059 (Nm)^2.

The four evolutionary algorithms above described require a large number of function evaluations which impairs their direct use for optimization in this application due to the computational cost of simulating the ultimate buckling moments for different TRAC booms. The Pareto frontier obtained when penalizing imperfection-sensitive designs (Figure 8a) shows a discontinuity close to a 0.20 Nm for the lower bound of the ultimate buckling moment around X. This finding provides a reasonable argument to select the TRAC boom geometry with $\theta = 270^{\text{o}}$ and $h = 4.1$ mm because a small increase of the lower bound of the ultimate buckling moment around X would lead to an undesirably sharp decrease in the lower bound of the ultimate buckling moment around Y.

Furthermore, different optimization objectives can be defined without additional computational analyses, as previously mentioned. For instance, if one is not too concerned with imperfection-sensitivity one could simply optimize for the mean value of the ultimate buckling moments without taking into account the uncertainty of the response. Figure 9 includes the respective results obtained in this case. As can be seen the optimal values are evidently larger than the values in Figure 8, but interestingly the discontinuity in the Pareto frontier occurs earlier.

In the exceptional circumstances provided by considering a low-dimensional problem and a relatively small design space, the shape of the decision space can be seen without performing optimization and just evaluating a large initial population of input points. Appendix B includes the two informative figures for optimization considering the lower bounds or the mean values of the ultimate buckling moments. Comparing Figures 8a with A.5a, and Figures 9a with A.5b, one can clearly understand the occurrence of the discontinuities in the Pareto frontiers.

4. Conclusion

A data-driven computational framework is proposed for the analysis and design of a large class of structures and materials addressing complications such as noisy observations through machine learning with uncertainty quantification, and multi-objective optimization. The challenging problem of designing ultrathin deployable carbon fiber composite structures for simultaneously improving their imperfection-sensitive ultimate buckling moments under two loading conditions is provided as an illustration of the framework.

The combination of machine learning for imperfection-sensitivity quantities of interest and multi-objective optimization is shown to be critically important in order to reach viable solutions with a reasonable effort, due to the computational expense involved in analyzing the ultimate buckling of structures. In particular, for TRAC booms, different designs were found by defining a Pareto frontier of the mean value of the ultimate buckling moments with large potential improvements: 1) possibility of simultaneous increases for the ultimate buckling moments larger than 25%; 2) separate increases above 90% and 27%, without decreasing the other respective ultimate buckling; and 3) an increase of above 300% for one of the ultimate buckling moments with a respective decrease of 28% for the other, which results from exploring the presence of a discontinuity in the Pareto frontier.

Future applications of the framework to multi-physics applications of different materials and structures are expected to reveal innovative solutions to otherwise complex problems. The main limitations of data-driven approaches when designing new materials or structures under untested conditions remain to be the computational expense of predicting the behavior for each design point, as well as the curse of dimensionality occurring when the input design space and/or the decision space are high dimensional. Advanced reduced order models such as the "self-consistent clustering analysis" (Bessa et al., 2017; Liu et al., 2016) are expected to address the former challenge, while further improvements in machine learning and optimization algorithms will address the latter.

Acknowledgement

The authors acknowledge financial support from the Northrop Grumman Corporation.

Appendix A. Uncertainty quantification for TRAC boom design

Computationally generating structural imperfections as combinations of buckling modes requires the definition of a statistical distribution for the respective amplitudes within given bounds – see equation 4. Due to lack of experimental data, a statistical distribution of the amplitudes needs to be assumed. Here a Sobol sequence (Sobol, 1967, 1976) was used to create 300 sampling points of the imperfection amplitudes that were normalized by the thickness of the TRAC booms – see Figure A.1.

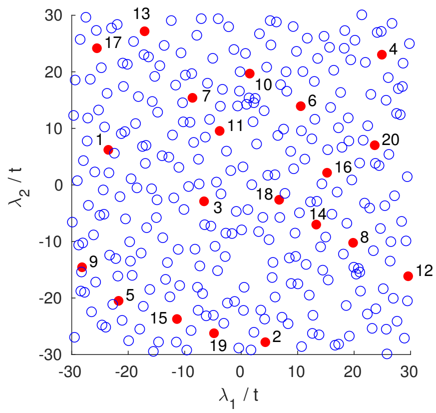
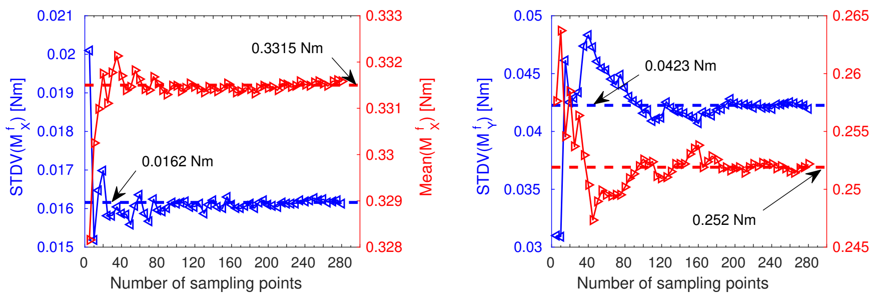


Figure A.1: Three hundred points from a Sobol sequence used to define the amplitudes of the first two buckling modes that create imperfections on the different TRAC boom geometries. The amplitude of the first buckling mode is λ_1, while the amplitude of the second is λ_2. Both amplitudes are normalized by the flange thickness $t = 71$ μm. The figure highlights the first 20 points by respective labels and solid red circles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Note that assuming a Sobol sequence to seed imperfection amplitudes may not be realistic. In fact, the quasi-uniformity of the distribution shown in Figure A.1 implies that higher imperfection amplitudes are increasingly more likely to occur than lower ones. In practice when manufacturing the TRAC booms this is not expected because different manufacturing processes introduce different systematic and random errors that can substantially skew and bias the distribution of imperfections. Nevertheless, due to the current lack of experimental evidence this distribution was assumed henceforth.

Estimation of uncertainty for the quantities of interest can be achieved by predicting the response of each design of the structure for the various sampled imperfections. But how many sampling points are needed to estimate the uncertainty for each design point?



(a) Standard deviation (blue) and mean (red) of the ultimate buckling moment around X

(b) Standard deviation (blue) and mean (red) of the ultimate buckling moment around Y

Figure A.2: Mean and standard deviation of the ultimate buckling moments obtained considering different amounts of imperfection sampling points. These results are for the initial geometry of the TRAC boom with nominal cross-section parameters of $h = 8$ mm and $\theta = 105^{\mathrm{o}}$.

Figures A.2a and A.2b show the variation of the mean and standard deviation of the ultimate buckling moments of a single design of the TRAC boom ($h = 8$ mm and $\theta = 105^{\mathrm{o}}$). The mean and standard deviation computed in these figures are obtained for an increasing number of sampling points. Note that the design of the TRAC boom used in this analysis coincides with the initial design that has the responses shown in Figure 2 for the first 6 imperfection points labeled in Figure A.1. Figures A.2a and A.2b demonstrate that both metrics (mean and standard deviation) tend to a constant value for a large enough number of sampling points. As expected, estimating the standard deviation requires more sampling points than estimating the mean – see also Table A.1.

Observing Figure A.2 and Table A.1 it may seem that there is a need for more than 80 sampling points to accurately estimate the standard deviation of a single TRAC boom design. This would be true if one were to accurately characterize only one TRAC boom design, but the data-driven framework discussed

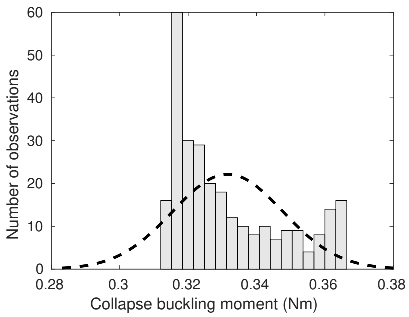
Table A.1: Error of the mean and standard deviation of the ultimate buckling moment obtained for the initial geometry of the TRAC boom (h = 8 mm and $\theta = 105^{\circ}$) for different numbers of sampling points.

	Number of sampling points					
	5	10	20	40	80	160
mean(M_X^{f}) error	+2.3%	+4.7%	+2.6%	−0.6%	−0.7%	+0.8%
STDV(M_X^{f}) error	−26.7%	−26.8%	+0.6%	+14.3%	+3.6%	−3.7%
mean(M_Y^{f}) error	−1.0%	−0.4%	+0.1%	+0.1%	0.0%	0.0%
STDV(M_Y^{f}) error	+24.4%	−6.0%	+5.1%	+1.9%	−1.2%	−0.3%

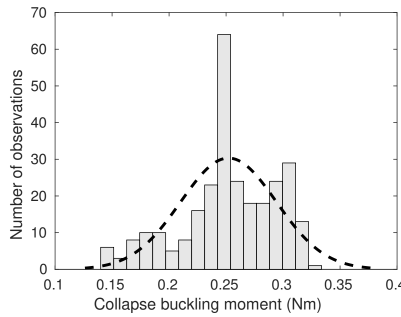
in this article is used to characterize multiple designs within a Bayesian machine learning framework (Gaussian process regression). Gaussian process regression provides a global approximation to the quantities of interest, as opposed to a local one, which in practice means that each design point only requires a reasonable estimate of the uncertainty. An intuitive way to understand this is by observing that the presence of other design points around each point improves the local predictions of uncertainty at every point of the domain.

As can be seen in Figure A.2 and Table A.1, considering 20 sampling points provides reasonable estimates of uncertainty of the initial design. Therefore, this amount of sampling points is expected to be reasonable for other design points as well.

An important point should be made about the statistical characterization of the TRAC boom ultimate buckling moments. In this article a Gaussian distribution was used to approximate their statistical distribution due to imperfections in each design. As shown in Figure A.3 assuming a Gaussian distribution is more reasonable for the ultimate buckling moment obtained when bending around Y than around X. The choice of a Gaussian distribution simplifies the implementation and lowers significantly the computational cost of the Gaussian process regression because the likelihood function is Gaussian (Rasmussen and



(a) Histogram and corresponding Gaussian for the ultimate buckling moment around X



(b) Histogram and corresponding Gaussian for the ultimate buckling moment around Y

Figure A.3: Histogram and corresponding Gaussian approximation for the distribution of the TRAC boom ultimate buckling moments. Same TRAC boom considered in Figure A.2 and considering 280 imperfection points.

Williams, 2005). Section 3.2 further elaborates this reasoning.

As a final comment about the statistical distributions of the ultimate buckling moments, it is noted that a small and yet non-negligible amount of computational analyses of TRAC booms subjected to bending around Y showed convergence issues. This was not observed for bending around X. Figure A.4 illustrates the convergence issue by plotting the moment–angle response of the TRAC boom subjected to bending around Y for two very similar imperfections. As can be seen in the figures one of the responses is incomplete because the equilibrium path after a second bifurcation point is not resolved, even though this does not correspond to the ultimate buckling. Such small imperfection difference is unlikely to cause the suggested premature ultimate buckling of the TRAC boom (see dashed line in the figure). This has implications on the distribution seen in Figure A.3b, since it is likely that the minimum ultimate buckling moment around Y is closer to 0.25 Nm than 0.15 Nm – see Remark 3.

Remark 3. *The detailed uncertainty analysis summarized herein for both loading conditions applied to the TRAC booms intends to demonstrate two important*

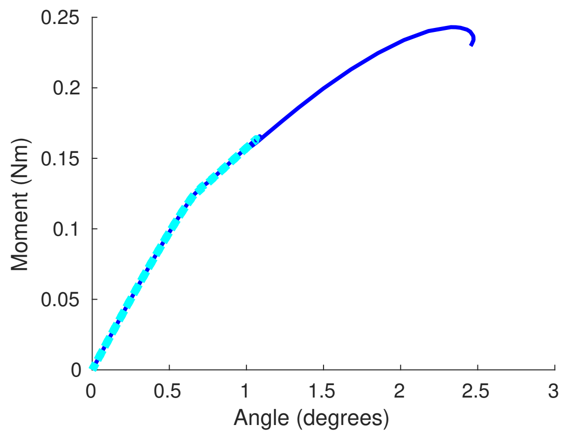
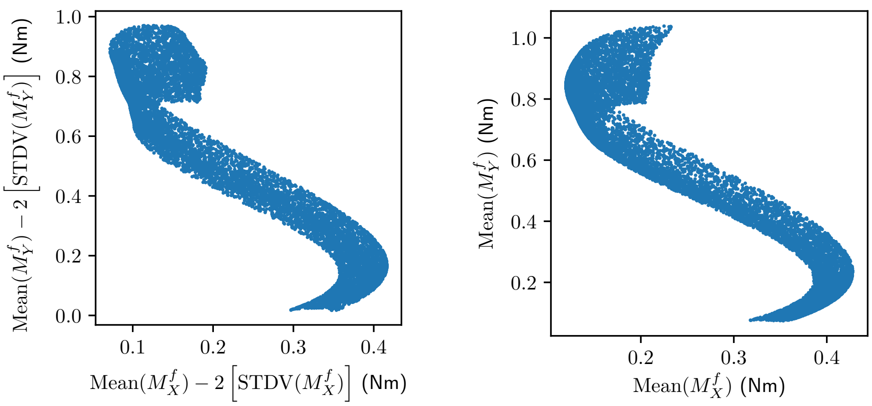


Figure A.4: The moment–angle response of a TRAC boom ($h = 8$ mm and $\theta = 105^{\mathrm{o}}$) subjected to bending around Y and considering two similar imperfections: solid line corresponds to ($\lambda_1 \approx 16.7t\,, \lambda_2 \approx -28.4t$), while the dashed line corresponds to ($\lambda_1 \approx 17.4t\,, \lambda_2 \approx -28.3t$). These results illustrate that some analyses stop prematurely due to convergence issues for this loading condition.

issues. First, the importance of the statistical distribution assumed for seeding the imperfections. This process should be guided by experiments or physical arguments that justify the distribution of imperfections, which are dependent on the manufacturing and testing conditions. Secondly, the relevance of modeling errors on the estimation of uncertainty. In fact, the statistical distribution of the ultimate buckling moments of the TRAC booms around Y is probably skewed, as in Figure A.3a, and certainly the standard deviation should be lower if no convergence issues were observed. Yet, the standard deviation for bending around Y would still be higher than the one for bending around X for this particular design.

Appendix B. Comment on shape of Pareto frontier

The particular problem of interest in this article is low-dimensional, i.e. it only has 2 input variables and two output quantities of interest. Since the machine learning model can be evaluated for any combination of input values within the appropriate bounds, then there is the possibility of exhaustively evaluating the responses for a large number of different designs without using optimization strategies.



(a) Decision space considering lower bounds of ultimate buckling moments and a large initial population.

(b) Decision space considering mean ultimate buckling moments and a large initial population.

Figure A.5: Exhaustive exploration of a large number of possible values of the output quantities of interest within the input bounds of $h = [2.5, 5.5]$ mm, $\theta = [45^o, 270^o]$. Note that this figure is not generated with any optimization algorithm. Instead, it provides the ultimate buckling moments of $10,000$ different designs, effectively showing the entire output space to be optimized.

Figure A.5 shows this exhaustive evaluation of a population of $10,000$ designs from which the entire output space can be seen as a collection of the responses for those points. This demonstrates why the Pareto frontier has a discontinuity around $M_X^f = 0.20$ in Figure 8a and around $M_X^f = 0.25$ in Figure 9a, where there is no combination of points that simultaneously improves both ultimate buckling moments. Evidently, Figure A.5 is included here for illustrative purposes. Under normal circumstances the results in Figure A.5 are not available, nor necessary.

References

Aghajari, S., Abedi, K., Showkati, H., 2006. Buckling and post-buckling behavior of thin-walled cylindrical steel shells with varying thickness subjected to uniform external pressure. Thin-Walled Structures 44 (8), 904 – 909.

Bader, J., Zitzler, E., 2011. Hype: An algorithm for fast hypervolume-based many-objective optimization. Evolutionary Computation 19 (1), 45–76, pMID: 20649424.

Bai, X., Bessa, M. A., Melro, A. R., Camanho, P. P., Guo, L., Liu, W. K., 2015. High-fidelity micro-scale modeling of the thermo-visco-plastic behavior of carbon fiber polymer matrix composites. Composite Structures 134, 132 – 141.

Bažant, Z. P., Cedolin, L., 2010. Stability of structures: elastic, inelastic, fracture and damage theories. World Scientific.

Benson, D., Bazilevs, Y., Hsu, M., Hughes, T., 2010. Isogeometric shell analysis: The reissner–mindlin shell. Computer Methods in Applied Mechanics and Engineering 199 (5–8), 276 – 289, computational Geometry and Analysis.

Bessa, M., Bostanabad, R., Liu, Z., Hu, A., Apley, D. W., Brinson, C., Chen, W., Liu, W., 2017. A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality. Computer Methods in Applied Mechanics and Engineering 320, 633 – 667.

Beume, N., Fonseca, C. M., Lopez-Ibanez, M., Paquete, L., Vahrenhold, J., Oct 2009. On the complexity of computing the hypervolume indicator. IEEE Transactions on Evolutionary Computation 13 (5), 1075–1082.

Beume, N., Naujoks, B., Emmerich, M., 2007. Sms-emoa: Multiobjective selection based on dominated hypervolume. European Journal of Operational Research 181 (3), 1653 – 1669.

Bisagni, C., 2000. Numerical analysis and experimental correlation of composite shell buckling and post-buckling. Composites Part B: Engineering 31 (8), 655 – 667.

Bisagni, C., Lanzi, L., 2002. Post-buckling optimisation of composite stiffened panels using neural networks. Composite Structures 58 (2), 237 – 247.

Chen, S., Gonella, S., Chen, W., Liu, W. K., 2010. A level set approach for optimal design of smart energy harvesters. Computer Methods in Applied Mechanics and Engineering 199 (37–40), 2532 – 2543.

Chinesta, F., Ammar, A., Leygue, A., Keunings, R., 2011. An overview of the proper generalized decomposition with applications in computational rheology. Journal of Non-Newtonian Fluid Mechanics 166 (11), 578 – 592.

Curtarolo, S., Morgan, D., Persson, K., Rodgers, J., Ceder, G., Sep 2003. Predicting crystal structures with data mining of quantum calculations. Phys. Rev. Lett. 91, 135503.

Daynes, S., Grisdale, A., Seddon, A., Trask, R., 2014. Morphing structures using soft polymers for active deployment. Smart Materials and Structures 23 (1), 012001.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., Apr 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6 (2), 182–197.

Demuth, H. B., Beale, M. H., De Jess, O., Hagan, M. T., 2014. Neural Network Design, 2nd Edition. Martin Hagan, USA.

Diaconu, C. G., Weaver, P. M., Mattioni, F., 2008. Concepts for morphing airfoil sections using bi-stable laminated composite structures. Thin-Walled Structures 46 (6), 689 – 701.

Dong, L., Lakes, R., 2013. Advanced damper with high stiffness and high hysteresis damping based on negative structural stiffness. International Journal of Solids and Structures 50 (14–15), 2416 – 2423.

Elishakoff, I., 2014. Resolution of the twentieth century conundrum in Elastic Stability. World Scientific.

Elvin, N. G., Lajnef, N., Elvin, A. A., 2006. Feasibility of structural monitoring with vibration powered sensors. Smart Materials and Structures 15 (4), 977.

Fischer, C. C., Tibbetts, K. J., Morgan, D., Ceder, G., 2006. Predicting crystal structure by merging data mining with quantum mechanics. Nature Materials 5 (8), 641–646.

Forrester, A. I., Keane, A. J., 2009. Recent advances in surrogate-based optimization. Progress in Aerospace Sciences 45 (1), 50 – 79.

Gautier, R., Zhang, X., Hu, L., Yu, L., Lin, Y., L., S. O., Chon, D., Poeppelmeier, K. R., Zunger, A., Apr. 2015. Prediction and accelerated laboratory discovery of previously unknown 18-electron abx compounds. Nat Chem 7 (4), 308–316.

Gill, P. E., Murray, W., Saunders, M. A., 2005. Snopt: An SQP algorithm for large-scale constrained optimization. SIAM Review 47 (1), 99–131.

Goldberg, P. W., Williams, C. K., Bishop, C. M., 1998. Regression with input-dependent noise: A gaussian process treatment. Advances in Neural Information Processing Systems, 493–499.

Hopfield, J. J., 1982. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences 79 (8), 2554–2558.

Hu, N., Burgueño, R., 2015. Buckling-induced smart applications: recent advances and trends. Smart Materials and Structures 24 (6), 063001.

Hutchinson, J., Koiter, W., 1970. Postbuckling theory. Appl. Mech. Rev 23 (12), 1353–1366.

Ishibuchi, H., Tsukamoto, N., Nojima, Y., June 2008. Evolutionary many-objective optimization: A short review. In: 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence). pp. 2419–2426.

Izzo, D., 2012. Pygmo and pykep: Open source tools for massively parallel optimization in astrodynamics (the case of interplanetary trajectory optimization). In: Proceedings of the Fifth International Conference on Astrodynamics Tools and Techniques, ICATT.

Kalpakjian, S., Schmid, S. R., Sekar, K. V., 2014. Manufacturing engineering and technology. Pearson Upper Saddle River, NJ, USA.

Kersting, K., Plagemann, C., Pfaff, P., Burgard, W., 2007. Most likely heteroscedastic gaussian process regression. In: Proceedings of the 24th International Conference on Machine Learning. ICML '07. ACM, New York, NY, USA, pp. 393–400.

Kim, S. Y., Mechefske, C. K., Kim, I. Y., 2013. Optimal damping layout in a shell structure using topology optimization. Journal of Sound and Vibration 332 (12), 2873 – 2883.

Kiyono, C., Silva, E., Reddy, J., 2012. Design of laminated piezocomposite shell transducers with arbitrary fiber orientation using topology optimization approach. International Journal for Numerical Methods in Engineering 90 (12), 1452–1484.

Konak, A., Coit, D. W., Smith, A. E., 2006. Multi-objective optimization using genetic algorithms: A tutorial. Reliability Engineering & System Safety 91 (9), 992 – 1007, special Issue - Genetic Algorithms and ReliabilitySpecial Issue - Genetic Algorithms and Reliability.

Krige, D. G., 1951. A statistical approach to some mine valuation and allied problems on the witwatersrand. Ph.D. thesis, Master's thesis of the University of Witwatersrand.

Ladevèze, P., Passieux, J.-C., Néron, D., 2010. The {LATIN} multiscale computational method and the proper generalized decomposition. Computer Methods in Applied Mechanics and Engineering 199 (21–22), 1287 – 1296, multiscale Models and Mathematical Aspects in Solid and Fluid Mechanics.

Lazarus, A., Reis, P. M., 2015. Soft actuation of structured cylinders through auxetic behavior. Advanced Engineering Materials 17 (6), 815–820.

Leclerc, C., Wilson, L. L., Bessa, M. A., Pellegrino, S., Jan. 2017. Characterization of ultra-thin composite triangular rollable and collapsible booms. In: AIAA SciTech Forum. American Institute of Aeronautics and Astronautics, AIAA-2017-0172.

Li, X., 2003. A Non-dominated Sorting Particle Swarm Optimizer for Multiobjective Optimization. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 37–48.

Ling, J., Templeton, J., 2015. Evaluation of machine learning algorithms for prediction of regions of high Reynolds averaged Navier Stokes uncertainty. Physics of Fluids 27 (8), 085103.

Liu, Z., Bessa, M., Liu, W. K., 2016. Self-consistent clustering analysis: An efficient multi-scale scheme for inelastic heterogeneous materials. Computer Methods in Applied Mechanics and Engineering 306, 319 – 341.

Loukaides, E., Smoukov, S., Seffen, K., 2014. Magnetic actuation and transition shapes of a bistable spherical cap. International Journal of Smart and Nano Materials 5 (4), 270–282.

Mallikarachchi, H. M. Y. C., Pellegrino, S., May 2014. Design of ultrathin composite self-deployable booms. Journal of Spacecraft and Rockets 51 (6), 1811–1821.

Matheron, G., 1963. Principles of geostatistics. Economic geology 58 (8), 1246–1266.

McKay, M. D., Beckman, R. J., Conover, W. J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21 (2), 239–245.

Morris, M. D., Mitchell, T. J., 1995. Exploratory designs for computational experiments. Journal of Statistical Planning and Inference 43 (3), 381 – 402.

Murphey, T. W., Banik, J., Mar. 1 2011. Triangular rollable and collapsible boom. US Patent 7,895,795.

Murphey, T. W., Turse, D., Adams, L., Jan. 2017. Trac boom structural mechanics. In: AIAA SciTech Forum. American Institute of Aeronautics and Astronautics, pp. –.

Neal, R. M., Jan. 1997. Monte carlo implementation of gaussian process models for bayesian regression and classification. Technical Report 9702, Department of Statistics, University of Toronto.

Ning, X., Pellegrino, S., 2015. Imperfection-insensitive axially loaded thin cylindrical shells. International Journal of Solids and Structures 62, 39 – 51.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Pellegrino, S., 2014. Deployable structures. Vol. 412. Springer.

Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., Tucker, P. K., 2005. Surrogate-based analysis and optimization. Progress in Aerospace Sciences 41 (1), 1 – 28.

Rasmussen, C. E., Williams, C. K. I., 2005. Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning). The MIT Press.

Reis, P. M., Sep. 2015. A perspective on the revival of structural (in)stability with novel opportunities for function: From buckliphobia to buckliphilia. Journal of Applied Mechanics 82 (11), 111001–111001–4.

Riks, E., 1979. An incremental approach to the solution of snapping and buckling problems. International Journal of Solids and Structures 15 (7), 529 – 551.

Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review 65 (6), 386.

Rumelhart, D. E., McClelland, J. L., PDP Research Group, C. (Eds.), 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations. MIT Press, Cambridge, MA, USA.

Saal, J. E., Kirklin, S., Aykol, M., Meredig, B., Wolverton, C., 2013. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (oqmd). Jom 65 (11), 1501–1509.

Shen, H.-S., 2012. Thermal buckling and postbuckling behavior of functionally graded carbon nanotube-reinforced composite cylindrical shells. Composites Part B: Engineering 43 (3), 1030 – 1038.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al., 2016. Mastering the game of go with deep neural networks and tree search. Nature 529 (7587), 484–489.

Simpson, T., Poplinski, J., Koch, N. P., Allen, J., 2001. Metamodels for computer-based engineering design: Survey and recommendations. Engineering with Computers 17 (2), 129–150.

Snelson, E., Ghahramani, Z., Rasmussen, C. E., 2004. Warped gaussian processes. In: Advances in Neural Information Processing Systems. pp. 337–344.

Sobol, I. M., 1967. On the distribution of points in a cube and the approximate evaluation of integrals. Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki 7 (4), 784–802.

Sobol, I. M., 1976. Uniformly distributed sequences with an additional uniform property. USSR Computational Mathematics and Mathematical Physics 16 (5), 236–242.

Tafreshi, A., 2002. Buckling and post-buckling analysis of composite cylindrical shells with cutouts subjected to internal pressure and axial compression loads. International Journal of Pressure Vessels and Piping 79 (5), 351 – 359.

Tikhonov, A., 1963. Solution of incorrectly formulated problems and the regularization method. Soviet Meth. Dokl. 4, 1035–1038.

Titsias, M. K., Lázaro-Gredilla, M., 2011. Variational heteroscedastic gaussian process regression. In: Proceedings of the 28th International Conference on Machine Learning (ICML-11). pp. 841–848.

Wang, P., Casadei, F., Shan, S., Weaver, J. C., Bertoldi, K., Jul 2014. Harnessing buckling to design tunable locally resonant acoustic metamaterials. Phys. Rev. Lett. 113, 014301.

Widrow, B., Hoff, M. E., 1960. Adaptive switching circuits. IRE WESCON Convention Record 4, 96–104.

Woo, J., Meguid, S., Stranart, J., Liew, K., 2005. Thermomechanical postbuckling analysis of moderately thick functionally graded plates and shallow shells. International Journal of Mechanical Sciences 47 (8), 1147 – 1171.

Yvonnet, J., He, Q.-C., 2007. The reduced model multiscale method (r3m) for the non-linear homogenization of hyperelastic media at finite strains. Journal of Computational Physics 223 (1), 341–368.

Zhu, C., Byrd, R. H., Lu, P., Nocedal, J., 1997. Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. ACM Transactions on Mathematical Software (TOMS) 23 (4), 550–560.

Zitzler, E., Deb, K., Thiele, L., 2000. Comparison of multiobjective evolutionary algorithms: Empirical results. Evolutionary Computation 8 (2), 173–195.

Zitzler, E., Laumanns, M., Thiele, L., 2001. Spea2: Improving the strength pareto evolutionary algorithm.

Zitzler, E., Thiele, L., Nov 1999. Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. IEEE Transactions on Evolutionary Computation 3 (4), 257–271.