Eleventh U.S. National Conference on Earthquake Engineering
Integrating Science, Engineering & Policy
June 25-29, 2018
Los Angeles, California

A FRAMEWORK FOR PERFORMANCE-BASED ANALYTICS DRIVEN DESIGN

H. Burton[1]

ABSTRACT

Despite its effectiveness in enabling structural engineers to target specific stakeholder-driven building performance objectives, the second generation performance-based seismic design (PBSD) framework has not been widely adopted in practice. This is partly due to the computational expense associated with performing iterative nonlinear response history analyses and damage, loss and downtime assessments. To address this challenge, a performance-based analytics-driven (PBAD) design framework is proposed as an alternative to the traditional PBSD methodology. Central to the PBAD design framework is the use of surrogate models to guide preliminary design iterations. The surrogate models, which serve as a statistical link between key design variables and seismic performance outcomes, quantified in terms of nonlinear response demand parameters (e.g. story draft ratios, component deformations, floor accelerations), can be used to target the optimal design region. Once the target region has been established, the design can be assessed, revised and finalized using the mechanistic approach that is used in the traditional design process. The expectation is that the PBAD design approach will minimize the number of iterations and computational expense needed to target specific stakeholder-driven performance objectives.

[1]Assistant Professor and Englekirk Presidential Chair in Structural Engineering, Dept. of Civil and Environmental Engineering, University of California, Los Angeles, CA 90095 (email: hvburton@ucla.edu)

Burton, HV. A framework for performance-based analytics driven design. *Proceedings of the 11th National Conference in Earthquake Engineering*, Earthquake Engineering Research Institute, Los Angeles, CA. 2018.

A Framework for Performance-Based Analytics Driven Design

J. Lord[1] and R. Wright[2]

ABSTRACT

Despite its effectiveness in enabling structural engineers to target specific stakeholder-driven building performance objectives, the second generation performance-based seismic design (PBSD) framework has not been widely adopted in practice. This is partly due to the computational expense associated with performing iterative nonlinear response history analyses and damage, loss and downtime assessments. To address this challenge, a performance-based analytics-driven (PBAD) design framework is proposed as an alternative to the traditional PBSD methodology. Central to the PBAD design framework is the use of surrogate models to guide preliminary design iterations. The surrogate models, which serve as a statistical link between key design variables and seismic performance outcomes, quantified in terms of nonlinear response demand parameters (e.g. story draft ratios, component deformations, floor accelerations), can be used to target the optimal design region. Once the target region has been established, the design can be assessed, revised and finalized using the mechanistic approach that is used in the traditional design process. The expectation is that the PBAD design approach will minimize the number of iterations and computational expense needed to target specific stakeholder-driven performance objectives.

Introduction

Advancements in the performance-based earthquake engineering (PBEE) framework provide rigorous probabilistic descriptions of seismic performance, using metrics such as economic losses, fatality estimates and downtime [1]. In addition to the framework itself, guidelines [2] and computing tools (e.g. SP3 and PACT) have been developed to facilitate its implementation. As shown in Figure 1, performance-based seismic design (PBSD) involves a series of steps, the first of which is defining a set of performance objectives using a metric of interest (e.g. life cycle costs). After performing a preliminary design, ideally, the performance should be assessed using the PBEE framework, which typically involves performing nonlinear response history analyses (NRHAs) on a structural model and using the generated engineering demand parameters to conduct damage and loss assessment using a computing tool such as the Seismic Performance Prediction Program (SP3) to assess physical damage, economic losses, the probable number of fatalities and recovery time. Based on the results of this initial assessment, the design is revised as needed and the assessment is repeated.

While it has been shown to be effective for target specific performance outcomes, PBSD has not been widely adopted by practicing structural engineers. This is partly because the state-of-

[1]Professor, Dept. of Civil Engineering, University of Somwherehill, Somewhere, XY 12345 (email: XX@abc.edu)
[2]Graduate Student Researcher, Dept. of Civil Engineering, University of Somwherehill, Somewhere, XY 12345

Lord J, Wright R. Template to illustrate paper format for the 11NCEE. *Proceedings of the 11th National Conference in Earthquake Engineering*, Earthquake Engineering Research Institute, Los Angeles, CA. 2018.

practice in structural engineering is such that most designers rely on linear analysis models to estimate response demands, which is generally unsuitable for PBEE assessments. Even when NRHA is performed, the iterative process of conducting damage and loss assessments, revising designs and doing repeat NRHAs, loss and damage assessments is expensive, both computationally and in terms of labor, which often leads to unjustifiable design fees.

To address the challenges associated with multiple design iterations and evaluations, a performance-based analytics-driven (PBAD) seismic design framework is being proposed. Central to the application of the PBAD design approach is the use of surrogate models to estimate nonlinear response demands, damage, economic losses and recovery time for a given iteration during the schematic design phase. The role of the surrogate model is to replicate the results of the structural response simulation and risk assessments as closely as possible while being computationally cheaper. Recent advances in prediction-analytics such as Machine Learning algorithms have facilitated the development of data-driven surrogate models that can replicate mechanistic (numerical models that explicitly simulate the phenomena under consideration) simulation models results to a high degree of accuracy. The use of a surrogate model during the "assessment" phase of PBAD design makes previously computationally intractable tasks such as design optimization, design space exploration, sensitivity analysis and what-if analyses, much more feasible. The long-term vision is that generalized PBEE assessment-based surrogate models for different lateral force resisting systems and building occupancies are developed and made available to practicing structural engineers. This vision has certainly been made possible by the increased availability of computational and data storage resources. This paper presents an overview of the proposed PBAD seismic design framework and illustrates the development of a nonlinear response demand surrogate model for a 6-story controlled rocking braced frame (CRBF) system.

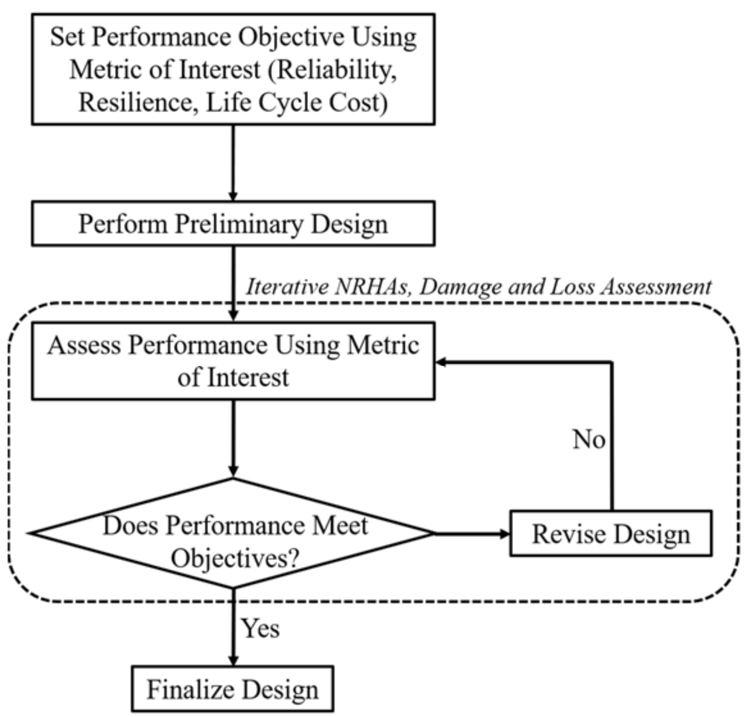

Figure 1. Overview of performance-based seismic design procedure highlighting the need for iterative NRHAs and impact (damage, loss and downtime) assessments

PBAD Design Framework Overview

Figure 2 shows an overview of the proposed PBAD seismic design framework. The first step involves defining the design variables, constraints and performance criteria. Examples of design constraints include the number of stories, the floor area and architectural and programmatic layout of the space, which limit the configuration and size of lateral force resisting system (LRFS) elements. The performance criteria will vary based on the project and constraints on the type of mechanistic assessment that can be used by the structural engineer (usually due to labor costs). For example, most new building design projects in high seismic regions of the United States are performed using linear elastic (response spectrum or equivalent static) analyses. In these cases, "performance" is measured by the ability of the design to meet the basic requirements of the code in terms of strength and stiffness. Strength requirements are met by designing the LFRS using some minimum base shear and stiffness is controlled by the drift limits set by the code. In other cases, the same code-based performance criteria are evaluated using NRHAs. A third case, which represents the overwhelming minority, involves using performance metrics such as damage, losses and recovery (in addition to safety). As described later, in either of these three design contexts, surrogate models can provide a means of efficiently exploring the design space to meet the desired performance objective.

In the PBAD seismic design process, the primary variables of interest are the ones that are known to play a major role in achieving the desired performance outcome. Strength, stiffness and ductility are three variables that are used to control seismic response demands. These "high-level" design variables are a function of the sizes and global and local (detailing) configuration of the members that comprise the LFRS. If the performance goal is to limit economic losses and downtime, variables that are explicitly linked to controlling structural and non-structural damage (e.g. component-damage fragility parameters) must also be considered. Note that there are several types of design variables that can be used to produce the desired performance outcome. Seismic design variables can be continuous (or quantitative) and defined on some interval (e.g. strength, stiffness, median value of damage fragility function) or categorical (e.g. introduction of a damage control mechanism). The dependence of design variables is also an important consideration, especially in the surrogate model development stage.

The second step of the PBAD design framework uses the surrogate models to rapidly explore the variable space and produce designs that meet the desired performance outcome. Generally, surrogate models are compact representations of the approximate multivariate input/output behavior a complex system, which should ideally be developed from a limited number of simulations. They represent the statistical link between relevant design variables and the performance outcome of the system under consideration. In the context of PBAD seismic design, their role is to mimic the complex structural response behavior that is simulated using NRHAs as well as the damage, financial loss and downtime assessed as part of the PBEE framework. Once developed, surrogate models can be used to efficiently conduct parametric studies, design space exploration, optimization and sensitivity analyses. Ultimately, the designer will be able to quickly find the optimal design region, or, more specifically, the ranges (for quantitative) or assignment (for categorical) of design variables that achieve the desired performance objective. More details of the surrogate model development as well as an illustrative example are provided later in the paper.

Once the optimal region (variable ranges) are obtained, the LFRS design can be performed in accordance to the governing criteria (third step in Figure 2). Except, in contrast to the traditional design procedure, the structural engineer has both quantitative and qualitative (for categorical variables) guidance on how to optimally size and configure the LRFS elements. Recall that three design/assessment contexts are considered: (1) code-based design performed using linear elastic analysis with the goal of achieving life safety performance, which is targeted by meeting the EDP-based requirements (minimum strength and non-exceedance of drift limit), (2) code-based design using NRHA and (3) performance-based design with the desired outcome described in terms of safety and reducing economic losses and recovery time. Although NRHA and damage-, loss- and downtime-assessment are not part of the first design context, the surrogate model will provide implicit information related to these metrics of performance and can be used in the design process. Similarly, in the second design context, the surrogate model will provide some insights into the probable damage, loss and recovery time.

As shown in Figure 2 (fourth step), the PBAD design framework does not preclude the application of the mechanistic performance assessment that would typically be done as part of the design process. The key difference is that the surrogate model allows the designer to efficiently target the desired performance outcome before conducting this assessment. In the case of the first design context, this will involve developing a linear elastic structural model based on the surrogate-model-informed optimal design and conducting equivalent static or dynamic (response spectrum analyses). However, as noted earlier, the designer will now have access to implicit knowledge about the nonlinear response and stakeholder-driven performance outcomes. In the second design/assessment context, the mechanistic assessment will include conducting NRHAs to ensure the relevant demand limits (drift ratios, component ductility demands, forces in capacity-designed components) are not exceeded. This will be done with the added benefit of having insights into damage, losses and recovery times. The mechanistic assessment in the third design context is a complete PBEE evaluation (NRHA + SP3 or PACT assessment). For the second and third design/assessment contexts, this stage of the PBAD design framework also serves an evaluation of the surrogate models themselves, which should ideally give performance outcomes that are comparable to that of the mechanistic models. As discussed later, an added layer of surrogate model evaluation is included as part of the design process. The mechanistic assessment is used to determine if the surrogate-based design meets the performance criteria. If so, the design can be finalized. In the case where the desired performance outcomes are not achieved, the design must be revised accordingly and the mechanistic assessment repeated. However, the expectation is that the PBAD approach reduces the number of design iterations needed to achieve optimal performance.

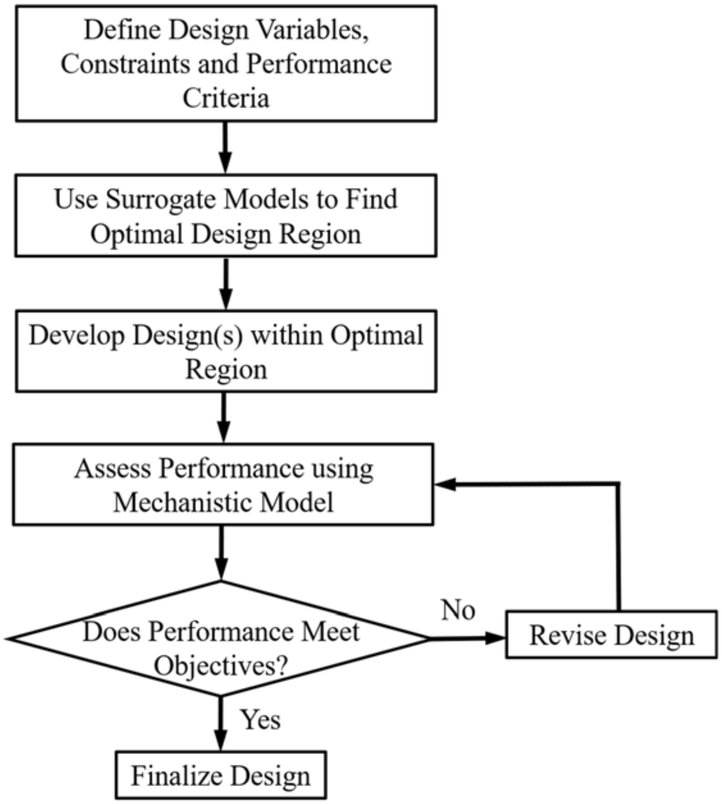

Figure 2 Overview of PBAD seismic design methodology

Surrogate Model Development

Figure 3 shows an overview of the procedure for developing the surrogate model, which involves solving multiple sub-problems. The main goal is to establish an approximation model that is as accurate as possible over the complete domain (design variable ranges and combinations) of interest while minimizing the simulation cost (Gorissen et al. 2010).

The first steps in the surrogate model development process are to identify relevant design variables and determine which performance criteria will be considered. Examples of performance measures to be considered in PBAD seismic design include nonlinear structural response demands and PBEE metrics such as physical damage, economic losses and downtime. Ideally, any surrogate model would incorporate multiple categories of metrics. However, in the illustrative example described later in the paper, only nonlinear structural response demands are considered.

The next step of the surrogate model development process is identifying which design variables are least important to the performance outcome and can therefore be removed. This step is important because it is directly related to the dimensionality and complexity of the surrogate model. However, it should be noted that there are regressing techniques that have been developed specifically to deal with issues of high-dimensionality, which can be employed in the construction of the surrogate model. Non-influential design variables can be identified by performing global sensitivity analyses. For application to PBAD seismic design, model developers will be able to draw on the wealth of research literature on the sensitivity of building seismic response and impacts to various design parameters. Otherwise, the sensitivity analyses can be performed as part of the surrogate model development process.

As shown in Figure 3, the primary stage of the surrogate model development involves multiple sub-steps, the first of which is the applicability domain definition, which refers to the design space, knowledge or information used to develop the surrogate model, and for which it is applicable to make predictions for new designs. For a given PBAD design problem, the applicability domain permits an evaluation of whether the model's assumptions are met and whether the model is reliably applicable to that problem. Establishing the applicability domain requires both quantitative and qualitative considerations. An example of a quantitative consideration is the possible ranges of the design variables that will be considered. Whether or not a surrogate model can be generalized for applications to multiple LRFSs is an example of a qualitative consideration.

The design space comprises all possible combinations of the considered variables. However, performing a mechanistic assessment of all possible design cases is likely to be intractable. Therefore, a choice must be made about which designs or variable combinations will be evaluated to construct the surrogate model. This process is referred to as the sampling plan development or a design of experiments (DOE). In the context of surrogate model development, these can be both physical and computational experiments, with the latter being used to support the PBAD seismic design process. The research literature provides several alternative approaches to developing sampling plans. Some researchers have recommended using a uniform, but not regular, spread of points across the design space. An example of such an approach is the space filling Latin Hypercube sampling technique, which uses a sample of points whose projections onto each variable axis are uniform, the logic being that it is wasteful to sample a variable more than once at the same value [4]. The central composite design is an alternative approach, where each variable is associated with a discrete value (or factor level) and a subset of the complete set of factor combinations is used as the sampling points [5].

Once the sampling points have been established, each design is evaluated using the mechanistic approach (NRHAs and SP3 assessments) and a performance point is obtained for each of the considered metrics. The selection and training of the surrogate model is the next step in the process. Model selection involves choosing the type of model that will be used to statistically represent the relationship between the design variables and the performance points. In choosing an appropriate surrogate model, a primary consideration is the accuracy of the prediction of the function landscape to be emulated and, moreover, in the context of optimization, this prediction should be most accurate in the region of the optimum. Other important considerations include the efficiency, ease-of-use, interpretability and complexity of the model. An efficient surrogate model is compact and allows for fast design space exploration. A surrogate model that is embedded in a software tool (e.g. excel, MATLAB or custom software) will facilitate ease-of-use. In surrogate model development, there is often a tradeoff between model complexity and interpretability. For example, under the right conditions, advanced learning techniques such as Support Vector Machines and Artificial Neural Networks have powerful prediction algorithms. However, since the surrogate model is not in the form of an analytical solution, designers may have a hard time interpreting the results. Whereas, surrogate models based on simple linear and polynomial functions are generally easier to digest.

The surrogate models should be trained using a subset of the sampling points with the remaining points used for model validation, which enables an assessment of its predictive capabilities away from the training data. Several measures of prediction capability can be used to evaluate the surrogate model. These measures vary primarily based on how they combine the errors at the various sampling points, which is taken as the difference in actual and predicted value of the performance measure. Examples of such measures include the adjusted root mean square error, coefficient of multiple determination [7] and median absolute relative deviation [7].

In the development of the surrogate model, the design space is defined before the relationship between the design variables, constraints and performance outcome is known. Once the surrogate model has been developed and this relationship has been established, a refinement of the design space is sometimes necessary [7]. For example, the bounds on the design variables are usually set based on empirical data from past design cases, which can lead to a larger than needed design space. Moreover, it may be discovered that some sampling points are infeasible, meaning that they represent unrealistic or poorly performing designs. The size of the design space may also affect the quality of the surrogate model. For example, a design space that is too large may lead to a surrogate model with poor prediction capabilities. An intelligent examination and reduction of the design space can alleviate these issues. Refinement of the design space can involve reducing the range of design variables, eliminating unrealistic or poorly performing variable combinations or by using functions to define boundaries that are based on a combination of more than one variable. The surrogate model development step (including all sub-steps) must be repeated once the new design space has been defined.

Once the surrogate model is finalized, the strategies for generating optimal design outcomes must be formulated. One such strategy is to apply traditional constrained optimization algorithms to the surrogate model. In cases where there are multiple competing performance objectives (e.g. minimizing upfront costs and maximizing life cycle performance) a weighting scheme can be adopted. Note that the goal of this stage is not to obtain a single optimal design since this will be very much dependent on the constraints of the specific PBAD design context. However, the expectation is that it will provide insights on the regions of the design space that will lead to optimal performance under the right set of constraints. Once these optimal regions have been identified, the accuracy of the predictions in these locations take on a greater priority compared to other parts of the design space. This may require further refinement of the design space by increasing the density of the design points in these optimal regions.

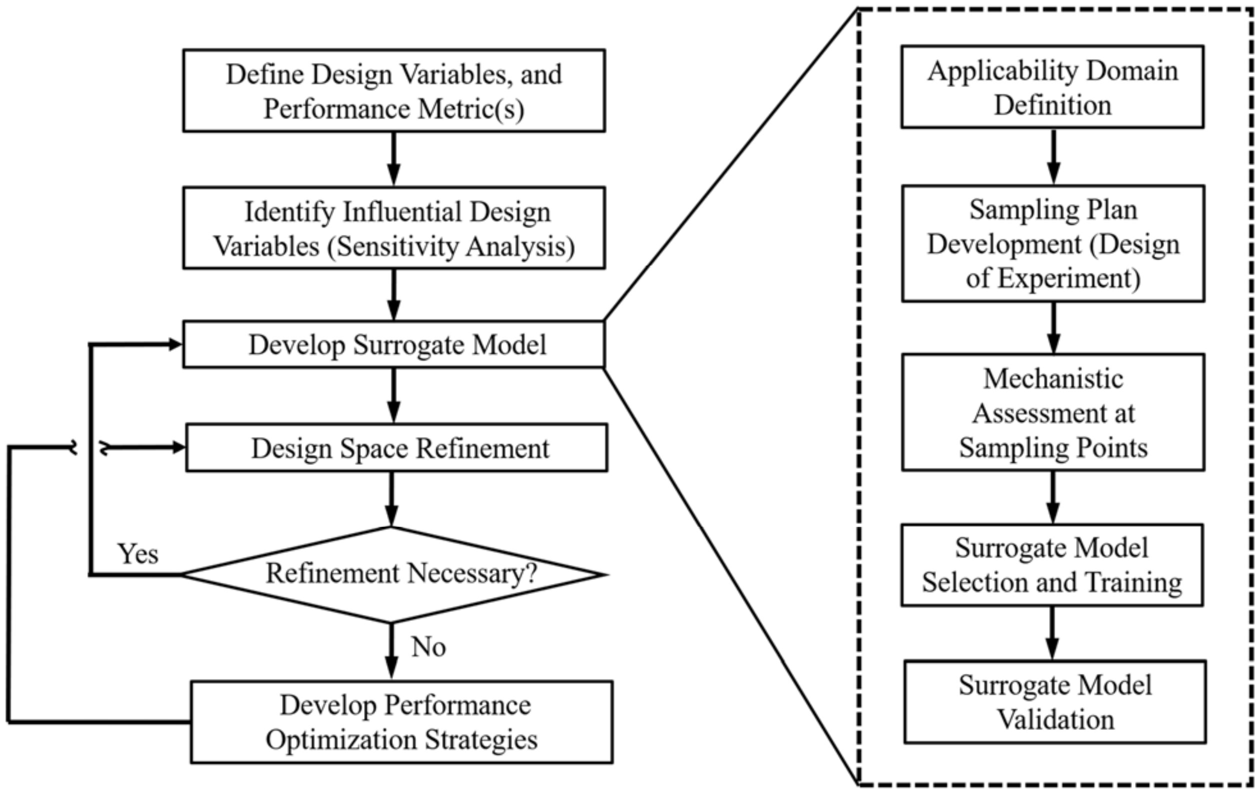

Figure 3 Overview of surrogate model development process

Example Application for Controlled Rocking Braced Frame System

This section presents an illustration of the development of a surrogate model to predict the seismic response demands in a 6-story controlled rocking braced frame system (CRBF). Local and global seismic demand parameters including peak story drift ratio (*PSDR*) and peak residual story drift ratio (*PRDR*), PT strain (ε_{PT}) and fuse shear deformation (δ_f), are considered. Five factors, which have been identified in previous studies as having a significant influence on the seismic demand of rocking frames are considered as the design variables, including the dead load on the rocking frame (P_D), initial PT force (F_{pt}), fuse yield strength (F_{yf}), the frame aspect ratio (i.e., the bay width-to-height ratio) (B/H), and the ground motion intensity level. Details of the building and structural model are described in [8].

The central composite design (CCD) approach is used to define the design sampling points, which refers to the combinations of design variables. The experiment involves performing nonlinear response history analysis of a structural model with a single combination of design variables. From each experiment, the mean and standard deviation of the natural logarithm of the seismic demands (ε_{PT}, δ_f, *PSDR* and *PRDR*) are computed as performance outcomes. CCD is applied to generate 43 sampling point combinations for normalized design variables (P_D/W, F_{pt}/W, F_{yf}/W and B/H). Using the basic (fixed) building design variables (e.g. building height, plan layout, tributary seismic weight) and the assigned values for each variable in each combination, structural models are constructed in OpenSees and nonlinear response history analyses are performed using thirty ground motion records. The mean and standard deviation of the natural logarithm of each seismic demand parameter are used as the response variables associated with the analysis of each model. Surrogate models (predictive equations) describing the statistical relationship between the response demands and input parameters are developed.

The surrogate models in this study are predictive equations that approximate the relationship between the response demand, y, (e.g. *PSDR*) and the predictors, x_i, or input parameters (e.g. P_D/W). After evaluating several functional forms, the quadratic polynomial relationship shown in Eq. (1) was adopted and used for the surrogate model.

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \sum_{j=1}^{k} \beta_{ij} x_i x_j + \varepsilon \qquad \text{Eq. (1)}$$

where the $\beta's$ are unknown regression coefficients and ε is a random error term (residual) that represents the difference between the actual (y_l) and predicted response, $(\hat{y_l})$; that is $\varepsilon_l = y_l - \hat{y_l}$ for each factor combination or design point $(l = 1, 2, \dots, n)$. The regression coefficients are determined using least squares method, seeking the $\beta's$ that minimize the sum of the squares of the errors $(\sum_{l=1}^{n} \varepsilon_l^2)$. The surrogate models are generated from the results of the nonlinear response history analyses using structural models that incorporate the input factor combinations from the experimental design. The $\beta's$ obtained for the predictive equations for the estimated log-mean $(\hat{\mu}_{\ln(y)})$ and log-standard deviation $(\hat{\sigma}_{\ln(y)})$ of the four response variables are summarized in [8]. The resulting surrogate models can be used to predict the seismic demand in the CRBFs for a given combination of influential design parameters and ground motion intensity. Using $\hat{\mu}_{\ln(y)}$ and $\hat{\sigma}_{\ln(y)}$ from the surrogate model, a single realization of the log-transformed variables is obtained from Eq. (2):

$$ln(y) = \hat{\mu}_{ln(y)} + \hat{\sigma}_{ln(y)} N(0,1) \qquad \text{Eq. (2)}$$

where $N(0,1)$ is a normally distributed random variable with mean of zero and standard deviation of 1. By incorporating $\hat{\sigma}_{\ln(y)}$, the surrogate model considers earthquake record-to-record variability and can be used for CRBFs of varied combinations of influential design parameters.

The prediction accuracy of the surrogate model is verified by developing structural models corresponding to ten additional sampling point combinations for the building and analyzing them using the thirty ground motions. Prediction accuracy is quantified based on the coefficient of determination (R^2), which indicates the goodness of fit and the mean (μ_χ) of the predicted-to-actual response ratio (χ). Accurate surrogate models that have higher coefficients of determination and a close-to-one μ_χ value are desirable. Table 1 provides details about the prediction capability of each of the surrogate models where $\hat{\mu}_{(y)}$ and $\hat{\sigma}_{(y)}$ are the mean and standard deviation of the response variable, respectively. The μ_χ values in parenthesis are those obtained for the verification runs (or testing dataset), which are different from the original experimental design runs (training dataset). Table 1 shows that the R^2 values are essentially 1.0 for the log-mean values and range from 0.76 to 0.93 for the log-standard deviation. The minimum value of μ_χ is 0.94 and the maximum is 1.18. Based on these results, the accuracy of the surrogate models for predicting the CRBF seismic response demands is deemed acceptable.

Table 1 Surrogate model verification details

Response Variable	Log-mean, $\mu_{ln(y)}$		Log-standard deviation, $\sigma_{ln(y)}$	
	R^2	μ_χ	R^2	μ_χ
PSDR	0.99	1.18	0.76	1.11
PRDR	0.99	1.15	0.90	1.01
ε_{PT}	0.99	1.02	0.93	1.11
δ_f	0.98	1.00	0.82	0.94

Conclusions

The need to perform iterative nonlinear response history analyses (NRHAs) and damage, loss and recovery assessments is a major hurdle in the implementation of the current performance-based seismic design (PBSD) methodology. This has effectively created a barrier against widespread adoption in structural engineering practice. A performance-based analytics-driven (PBAD) design methodology is presented as an alternative, with the specific goal of eliminating this barrier. The main departure from traditional PBSD is the use of surrogate models that enable efficient explorations of the design space to target stakeholder-driven performance objectives including safety, damage control and reduced direct and indirect economic losses. The expectation is that the proposed framework will empower engineers to target optimal performance-based designs without (or prior to) conducting detailed mechanistic (NRHAs, damage, loss and recovery) assessments.

Acknowledgments

This research is supported by National Science Foundation Award No. 1554714.

References

1. Krawinkler, H., and Deierlein, G. G., 2014. Challenges towards achieving earthquake resilience through performance-based earthquake engineering. In Performance-Based Seismic Engineering: Vision for an Earthquake Resilient Society (pp. 3-23). Springer Netherlands.
2. FEMA, 2012. Seismic performance assessment of buildings Volume 1: Methodology. FEMA P58-1, Applied Technology Council, Redwood City, CA.
3. Gorissen, D., Couckuyt, I., Demeester, P., Dhaene, T., and Crombecq, K. (2010) A Surrogate Modeling and Adaptive Sampling Toolbox for Computer Based Design. Journal of Machine Learning Research 11 (2010) 2051-2055.
4. Forrester A. and Keane A. (2009). Recent advances in surrogate-based optimization. Progress in Aerospace Sciences 45, 50-79.
5. Montgomery, D. C. (2013). Design and analysis of experiments. John Wiley & Sons, Inc.
6. Mack, Y., Goel, T., Shyy, W., & Haftka, R. (2007). Surrogate model-based optimization framework: a case study in aerospace design. Evolutionary computation in dynamic and uncertain environments, 323-342.
7. Sun, H., Burton, H. V., Zhang, Y. and Wallace, J. W. (2017). Inter-building interpolation of peak seismic response demands using spatially correlated demand parameters. Earthquake Engineering Structural Dynamics, (accepted for publication).
8. Moradi, S., Burton, H.V. and Kumar, I. (2017). Parameterized Fragility Functions for Controlled Rocking Steel Braced Frames, ASCE Journal of Structural Engineering (in review).