Structural Health Monitoring and Uncertainty Quantification in Large Steel Frame Structures

Gbandi Nikabou

Graduate Student, Dept. of Civil Engineering, Vanderbilt University, Nashville, TN, USA & Research Engineer, ERDC, Vicksburg, MS, USA

Pranav M. Karve

Assistant Professor, Dept. of Civil Engineering, Vanderbilt University, Nashville, TN, USA

Sankaran Mahadevan

Professor, Dept. of Civil Engineering, Vanderbilt University, Nashville, TN, USA

Brian Eick

SHM Program Manager & Research Engineer, ERDC, Champaign, IL, USA

ABSTRACT: Large structural systems, under extreme natural hazard loadings, exhibit strongly nonlinear behavior that challenges commonly used diagnosis and modeling techniques. In such situations, the initial model may be quite inadequate for structural health monitoring. This paper presents the first step toward a digital twin approach that continuously updates the model with health monitoring data, focusing on extreme wind events, and includes uncertainty quantification in both sensor data and model prediction. A two-step modeling methodology is developed for diagnostic and prognostic analysis and uncertainty quantification. The first step consists of Principal Components Analysis (PCA) on strain data from the undamaged structure to compute the Q_Statistic, a damage index vector based on the PCA reconstruction error. The obtained damage index vector is used to compute anomaly/damage thresholds. The strain values, obtained from the finite elements analysis simulation of are projected onto the constructed PCA baseline model for damage/anomaly detection. The damage index vector obtained from the PCA projection is used as the output in constructing a Bayesian predictive surrogate model of the structure along with environmental variables of temperature, wind speed, and wind direction as model inputs. The surrogate model is verified, calibrated, and validated with real sensor data from the field, and the uncertainty in the surrogate model prediction is quantified. The validated model is then used for probabilistic prognosis to inform proactive decision making ahead of critical events. The proposed methodology is illustrated for a large steel frame and rolling door used in aircraft hangars in Florida. The results of the proposed methodology show that both the PCA and its projection-based predictive model are respectively successful in detecting damage on the door structure and predicting damage indicator for the door structure.

1 INTRODUCTION

The development of machine learning (ML) has greatly benefited the structural health monitoring (SHM) community by making SHM data analysis and interpretation much easier and faster. In most of ML-SHM literature, the focus is generally on the structural state diagnosis such as damage identification (detection and/or localization) [1 - 4], estimation of risk of failure [5], or time series analysis [6 &7]. A particularly widespread usage of ML-SHM is for autonomous damage identification. Various techniques, mostly unsupervised learning, are used in the literature. The following methods have been studied [2]: principal components analysis (PCA), Kernel principal components analysis (KPCA), Gaussian mixture models (GMMs), Mahalanobis squared distance (MSD), and the auto-associative neural network (AANN). These methods frequently compute a damage index [1&2], which often is the Euclidean norm of the residual features, i.e., the difference between the original data and the reconstructed data through the chosen method. This is known as the Q-statistic (or Q-index or Q score). Other damage indicators such as T^2-statistic (D score) are also used. It is worth noting that all these methods exploit the structure's response data such as strain, stress, acceleration, or deflection without considering the excitations (loadings on the system) causing these responses. The excitation data is usually missing from the data collection either because these are difficult to measure or the instrumentation cost is too high; so, prognosis analysis based on environmental loads on the system is not often considered.

This work proposes a two-step diagnosis and prognosis analysis for damage detection and prediction with the latter based on the applied loading. The first step consists of extracting the damage indicator Q-statistic using PCA on undamaged structural numerical simulation output data, and choosing a threshold based on a probability interval or percentage cutoff, then proceed to damage detection using numerical simulation data for damaged structure. The second step consists of building a predictive model using the damage index as the output and external excitations as the input variables. The predictive model is to forecast damage growth into the future given a set of environmental variables. A surrogate model is used to map the relationship between the baseline Q-index from the first step and environmental variables. The surrogate model is then calibrated with sensor data from real in-service structure. The calibration is carried out using the Kennedy O'Hagan (KOH) approach [10], to Bayesian calibration. The calibrated model is then used to make predictions about the damage index of the structure given the environmental loadings on the real in-service structure, with quantified uncertainties about the predictions.

2 METHODOLOGY

2.1 Principal Component Analysis

Principal component Analysis (PCA) is a standard multivariate statistical technique, also known as orthogonal decomposition, commonly used as unsupervised learning for dimension reduction and anomaly/outlier detection [1&2]. The method consists of computing the eigen-decomposition of the covariance matrix; where the sample covariance matrix of the data $\boldsymbol{Y}$ is calculated as $Cov(\boldsymbol{Y}) = Y^T Y$.. From the decomposition, a set of p eigenvectors is maintained. The reduced dimension is obtained by projecting the original data $\boldsymbol{Y}$ onto the matrix formed by the selected eigenvectors.

$$\boldsymbol{d} = \boldsymbol{Y}\boldsymbol{a_p} \tag{1}$$

where $\boldsymbol{d}$ is the new data with reduced number of variables through PCA and $\boldsymbol{a_p}$ is the matrix of p eigenvectors retained. $\boldsymbol{d_i} = \boldsymbol{Y}\boldsymbol{a_{p_i}}$ is the ith vector of the new dataset.

The number of retained eigenvectors depends on the desired cumulative explained variance of the data. Data information explained by each eigenvectors is defined $\%explained = \frac{\lambda_i}{\sum_1^n \lambda_j}$. Summing the explained variance percentage of the total retained PCs provide the cumulative variance explained of the constructed model. For outlier or anomaly detection, there are two approaches, using the Q-statistic or T^2-statistic damage indicators. The Q-statistic or Q-index is of interest in this paper; it is a residual-based method where the original data is reconstructed.

$$\boldsymbol{Y}_{rec} = \boldsymbol{d} * \boldsymbol{a}_p^T \tag{2}$$

$$\boldsymbol{E} = \boldsymbol{Y} - \boldsymbol{Y}_{\boldsymbol{rec}} \tag{3}$$

Taking the Euclidian norm of matrix E gives the Q-index. The outlier detection threshold is chosen based on the experience in the domain of study. Often, statistical thresholds between 95% and 99% cutoffs are considered, or 3σ cutoff are used in defining the threshold. While equation 3 represents the residual error as a function of original data and reconstructed data, that residual error can also be expressed as a function of data noise ϵ and the PCA model discrepancy δ_{pc}.

$$\boldsymbol{E} = \boldsymbol{Y} - \boldsymbol{Y}_{\boldsymbol{rec}} = \delta_{pc} + \epsilon \tag{4}$$

Consider a set independent input variables $\boldsymbol{x}$ such that $\mathrm{y}_i = f(\mathrm{x}_i)$ where y_i is the i^{th} measurement vector and x_i is the vector of independent input variables for the i^{th} measurement. While the noise ϵ might be not be input-dependent, the PCA model discrepancy is input-dependent. Equation 4 hence can be written as follows:

$$\boldsymbol{E} = \boldsymbol{\delta}_{\boldsymbol{pc}} + \boldsymbol{\epsilon} = \boldsymbol{g}_{\boldsymbol{pc}}(\boldsymbol{x}) + \boldsymbol{\epsilon} \tag{5}$$

The consequence of equation 5 is that it makes Q-index dependent on the input variable x; $Q = h(x)$. This deduction is helpful in building predictive models for damage detection based on system input variables and the damage indicator Q.

2.2 *Surrogate Modeling and Bayesian Calibration*

The next step consists of defining the nature of the function $h(x)$. Because the true solution for the problem is unknown, and the simulation data is limited and uncertain, a surrogate model is used to map the relationship between the Q and x. The surrogate is trained with simulation data. Since the surrogate model is trained using simulation data over a small number of input points, and the trained surrogate model has fitting error, there is uncertainty in the prediction of the surrogate model. In addition, the model needs to be applicable to a real-world problem; hence the need for model calibration that is able to account for the various epistemic uncertainty sources in the model and field data. The considered approach is based on the KOH Bayesian calibration framework [10]. This framework uses Bayesian inference to establish the relationship between the computer simulation h and field measurement (observation) z and in the process account for observation error ϵ, model discrepancy δ (difference between true solution and computer model), and surrogate model uncertainty:

$$\boldsymbol{z}(\boldsymbol{x}) = \boldsymbol{\rho\eta}(\boldsymbol{x}, \boldsymbol{\theta}) + \boldsymbol{\delta}(\boldsymbol{x}) + \boldsymbol{\epsilon} \tag{6}$$

where $\eta(x, \theta)$ is the unknown true solution. The prior information about both $\eta(.,.)$ and $\delta(.)$ are represented as Gaussian processes: $\eta(.,.) \sim \mathcal{N}[m_1(.,.), c_1\{(.,.), (.,.)\}]$ and $\delta(.) \sim \mathcal{N}[m_2(.,.), c_2(.,.)]$ with the mean function $m_i(.,.)$, and variance function $c_i\{..\}$ for both $\delta(.)$ and $\eta(.,.)$ defined as in Kennedy and O'Hagan [10] Thus in the KOH framework, both the physics model parameters and the Gaussian process hyper-parameters of the discrepancy term are calibrated. The process consists of sampling the posterior of the model parameters given observation. The prior information is obtained from training the model with simulated data.

3 IMPLEMENTATION RESULTS

Rolling-hangar doors are usually designed by a third party independent from hangar design contractor. These doors mostly have minimal to no actual physical ties to the hangar building. For instance, the rolling hangar door considered in this work has practically no physical tie to the hangar building. The door is simply seated on a rolling rail with the only contact being its two wheels on the rail. In addition, a C-shape beam, as shown in Figure 1, guides the upper end of the door. The door is in coastal environment subject to wind load and temperature variations. Located in Florida, the door and its hangar system are potentially subject to hurricanes. During the October 2018 Category 5 Hurricane Michael, the hangar explored for this study suffered enormous amounts damage which included the rolling doors used for aircraft ingress/egress to the hangar. Failure of the rolling doors led to hurricane-force winds entering the hangar, leading to cladding failure. That event led to the SHM instrumentation of the door for to inform proactive decision-making for future events.

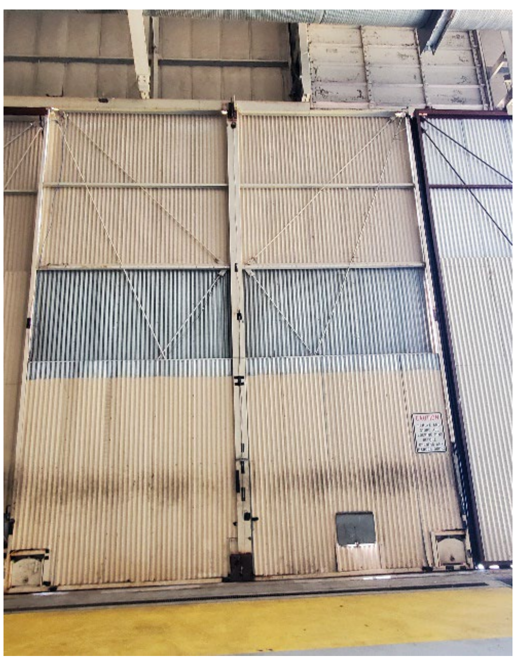

Figure 1: Rolling-hangar door used in this study.

Two FEA models of the hangar-door were constructed and simulated in ABAQUS [Figure 2]. The first model is for the undamaged structure. The model was constructed with mostly quadratic shell elements (S4R) with general contact. The boundary conditions are placed on the rail as fixed and on the top guiding beam as pinned. The door as described above is sitting on the rail is defined by simple contact. This model setup tried as much as possible to reflect the actual door boundary condition. Wind speed, wind direction, and temperature are used as model inputs. random combinations of the three inputs were used to generate 100 input sets. For the random combination, the wind speed was taken within the interval [0 mph, 250mph], [0^o, 360^o] for the direction and [20K, 115K] for temperature. A dynamic simulation of the door was carried for each input set and strain and acceleration data were collected at nodes [Figure 4]. The second model is for a damaged structure [Figure 5], with damage represented by the reduction in Young's modulus of the material in the region of interest. Stiffness reduction is considered as damage because it is the consequence for many damage types in steel structures. Whether the damage is a crack (mostly large cracks), corrosion, or section loss, the immediate effect of these types of damages is stiffness reduction of the material in the affected part or area of the structure.

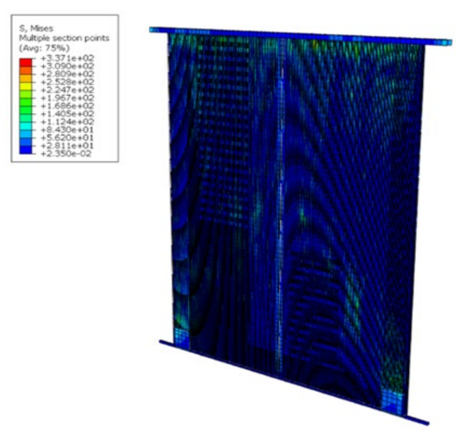

Figure 2: ABAQUS FEA model of rolling-hangar door used in this study

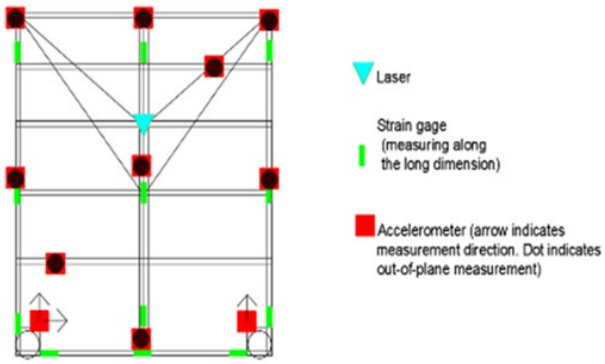

Figure 3: Sensors Locations on the door used in this study.

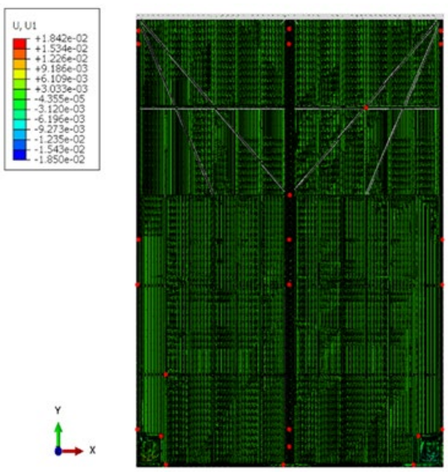

Figure 4: Sensors locations approximation in ABAQUS

The selection of stiffness reduction area of the doorframe is based on inspection observations of areas with heavy corrosion and section losses. The door frame [Figure 1] shows retrofit at the connection of the middle column and the bottom beam. Inspection of the other doors on the building revealed heavy corrosion and section loss at that connection and its immediate surrounding areas. The damaged FEA model [Figure 5] attempts to be realistic by modeling the damage in the observed area. Strain and acceleration data are collected at specific nodes [Figure 4] in a manner to approximate these locations to the actual sensors locations on the door at the hangar [Figure 3]. For the study in this paper, only strain data is used.

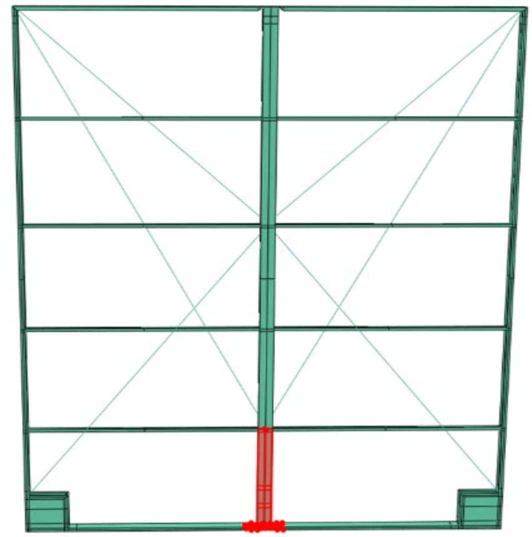

Figure 5: Damage Frame in ABAQUS FEA.

Using PCA as explained in section 2.1*Principal Component Analysis*, the FEA data from the undamaged case is reduced to five principal components with a cumulative explanation of 99% of the variance. The baseline damage indicator Q-statistic is computed through reconstruction error. A 99% cutoff limit is assumed for the threshold [Figure 7].

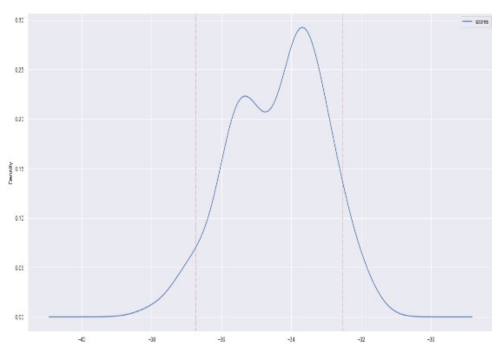

Figure 6: Baseline Q indicator distribution with Thresholds

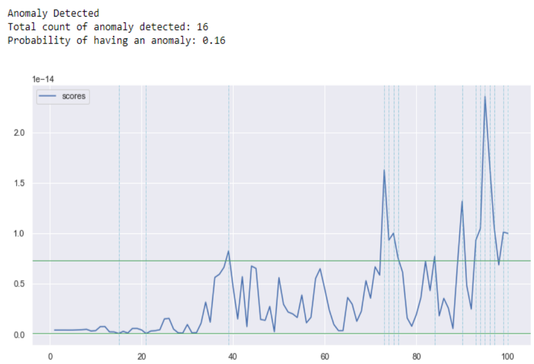

Figure 7: Anomaly detected on data from the damaged structure model

For damage detection, the damaged case data is projected on the baseline PCA model and reconstructed. The Q indicator is computed and compared to the threshold for outlier detection [Figure 6. With the detection step completed, the baseline Q indicator is used as output with wind speed, wind direction, and temperature as independent input variables to fit a surrogate model. A Gaussian process regression model is used to fit the data with R-square of 0.87 [Figure 9].

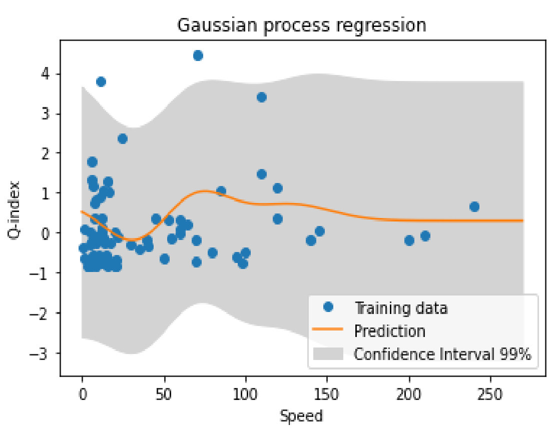

Figure 9: Surrogate model fit plot vs wind speed

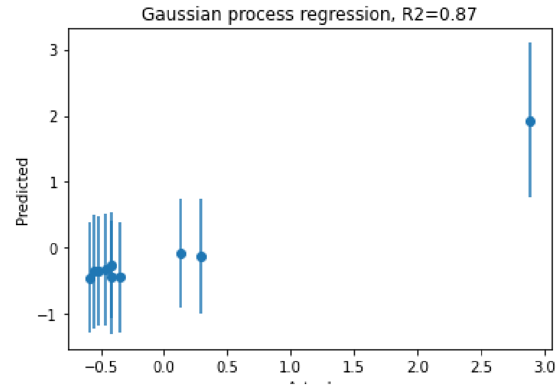

Figure 10: Surrogate model test with R^2

The field data collected is projected onto the baseline PCA model. The Q indicator of the observed data is computed through the reconstruction error. The Gaussian process regression surrogate model is then calibrated with the Q indicator following the Bayesian calibration framework described in Section 2.3 and sampling is carried out using the Hamiltonian Monte Carlo (HMC) in python pystan package. The sampled posterior is plotted for all parameters hyper parameters [Figure 10].

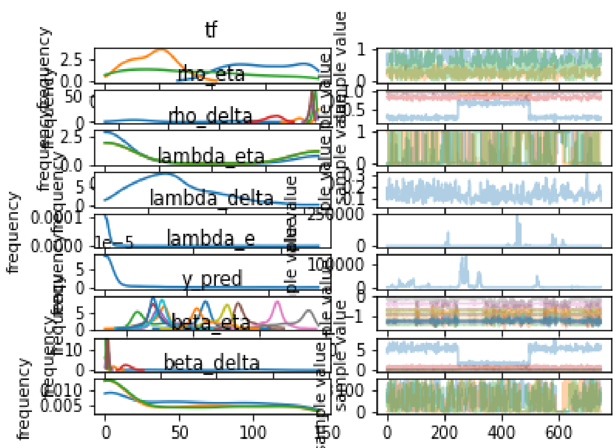

Figure 8: Bayesian Calibration Parameters Posterior Distributions

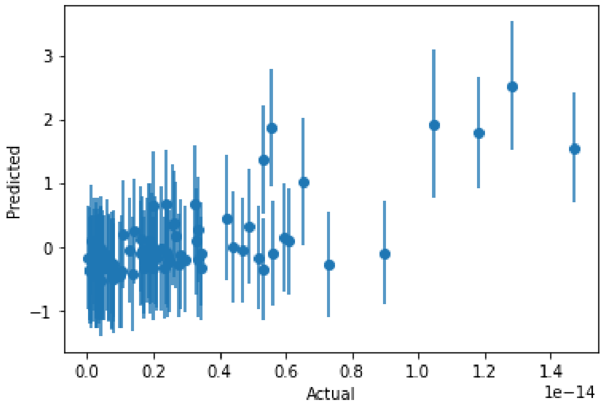

Figure 11: Prediction vs. Observation Plot with uncertainty

With the posterior now constructed, a prediction is carried out for the field data set aside for the purpose. The model in general did well in predicting the damage indicator [Figure 11]. The expectation is that data from the undamaged structure should fit within the boundaries of the calibrated model uncertainty and if it is not the case then the model should be either retained or updated; because the structure does go through a fatigue cycle which can change its Q-statistic overtime. A retrained or updated model will account for these changes.

4 CONCLUSION AND DISCUSSION

The basic idea in this work is to design to build a two-step diagnosis and prognosis model, using PCA for the diagnosis and constructing a calibrated surrogate model with quantified uncertainty for prognosis. With no real-world damaged structure for validation of the detection (diagnosis) step, a FEA model was used to verify the damage detection using Euclidean norm of the PCA reconstruction error as damage indicator. Because strain data is dependent on input variables such as wind speed, wind direction and ambient temperature, the damage index is deduced to be dependent on the input variables as well. The prognosis step successfully mapped that dependence using Gaussian process regression and quantified in the process the epistemic uncertainty of the model given the FEA model and field data.

With a model that projects the diagnosis into the future give values of input variables, further work consists of incorporating the entire diagnosis and prognosis process into a digital twin of damage detection and prediction in the structure by updating the PCA threshold values and predictive model with new field data, and forecasting the input. The digital twin will support informed decision making about maintenance and proactive mitigation measures for specific hazardous events.

5 REFERENCES

1. Mujica, L., Rodellar, J., A Fernandez, A., Guemes, A. “Q-statistic and T^2-statistic PCA-based measures for damage assessment in structures” (2010) *Structural Health Monitoring*, 539–553
2. Figueiredo, E., Santos, A. “Machine Learning Algorithms for Damage Detection”(2010) *Vibration-Based Techniques for Damage Detection and Localization in Engineering Structures*, 1 1-39
3. Golinval, J. C. “Damage Detection in Structures Based on Principal Component Analysis of Forced Harmonic Responses” (2017) *Procedia Engineering*, 199 1912-1918
4. Villamizar, R., Leonardo, J. Q., Camacho, J., Mujica, L. E., Ruiz, M. L. “Structural Damage Detection Algorithm Based on Principal Component Indexes and Embedded on a Real Time Platform” (2014) *EWSHM - 7th*)
5. Ozer, E., Feng, M. Q. "Structural reliability estimation with participatory sensing and mobile cyber-physical structural health monitoring systems" (2019) *Applied Sciences* 9.14 - 2840.
6. Kumar, K., Biswas, P. K., Dhang, N. “Time series-based SHM using PCA with application to ASCE benchmark structure” (2020) *Journal of Civil Structural Health Monitoring* 10.5 -899-911.
7. Gul, M., Catbas, F. N. "Statistical pattern recognition for Structural Health Monitoring using time series modeling: Theory and experimental verifications" (2009) *Mechanical Systems and Signal Processing* 23.7 -2192-2204.

8. Sankararaman. S., Mahadevan, S. “Integration of model verification, validation, and calibration for uncertainty quantification in engineering systems” (2015) *Reliability Engineering & system Safety*, 138- 194-209
9. Sankararaman, S., Mahadevan, S. “Comprehensive Framework for Integration of Calibration, Verification and Validation” (2012) *Collection of Technical Papers - AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*
10. Marc C. Kennedy, M. C. O'Hagan, A. “Bayesian calibration of computer models, Journal of the Royal Statistical Society” (2001) Series B (*Statistical Methodology*) 63(3)-425-464
11. Higdon, D., and Kennedy, M. C., and Cavendish, J. C., Cafeo, J. A., Ryne, R., D., “Combining Field Data and Computer Simulations for Calibration and Prediction” (2004) *SIAM Journal on Scientific Computing,26,2, -448-466*
12. White, A., Mahadevan, S., Grey, Z., Schmucker, J., Karl, A. “Efficient Calibration of a Turbine Disc Heat Transfer Model Under Uncertainty” (2020). *Journal of Thermophysics and Heat Transfer*. 35. 1-11. 10.2514/1.T6047.
13. Chong, A., Menberg, K., “Guidelines for the Bayesian calibration of building energy models” (2018) *Energy and Buildings*, 174 -527-547,
14. Diestmann, T., Broedling, N., Götz, B., Melz, T. “Surrogate Model-Based Uncertainty Quantification for a Helical Gear Pair” (2021) In: Pelz, P.F., Groche, P. (eds) *Uncertainty in Mechanical Engineering*.
15. Sudret, B., Marelli, S., Joe Wiart, J. "Surrogate models for uncertainty quantification: An overview" (2017) 11th *European conference on antennas and propagation* (EUCAP). IEEE
16. Sankararaman, S., Ling, Y., Shantz, C., Mahadevan, S. "Uncertainty quantification in fatigue damage prognosis." (2009) *Annual Conference of the PHM Society*. Vol. 1. No. 1