Gaussian Process Regression-Based Material Model for Stochastic Structural Analysis

Baixi Chen[1]; Luming Shen[2]; and Hao Zhang[3]

Abstract: Data-driven material models can capture the constitutive relationship directly from the data without involving any material-dependent mathematical expressions. But most data-driven approaches, such as artificial neural networks, only estimate the deterministic relations and do not consider the material uncertainty. In this paper, the constitutive relation is taken as a stochastic function following the Gaussian process, where a probability-based nonparametric method, called Gaussian process regression (GPR), is used to capture the constitutive relation with the uncertainty being included. Both one-dimensional (1D) and two-dimensional (2D) material data are used to validate the GPR-based constitutive model (GPR model). The obtained GPR model shows higher accuracy than other data-driven approaches, particularly when the data set size is small. When compared with the assumed true model, the GPR-based model has an average relative error of <2.3%. Finally, with the help of the material uncertainty identified by the GPR-based model from the material data, a data-driven stochastic structural analysis procedure is developed. The relative errors of the expected deflection and probability of failure given by the GPR model are smaller than 2% and 3%, respectively. **DOI: 10.1061/AJRUA6.0001138.**

Author keywords: Data-driven material model; Machine learning (ML); Gaussian process regression (GPR); Material uncertainty; Stochastic structural analysis.

Introduction

Modern numerical simulation methods, such as the finite-element method (FEM), have gained successful application in many areas over the decades. The solutions of most numerical simulations are based on conservation laws and material laws (Kirchdoerfer and Ortiz 2016). Different from conservation laws inferred from some universal principles, materials laws should be derived from material data, such as stress and strain data. Conventional material modeling approaches often need to construct a specific postulated mathematical expression for each material and adjust its parameters by incorporating experimental data. However, most conventional material models are only suitable for a limited type of materials, and a new mathematical form is often required to ensure the accuracy of analysis when a new type of material is involved. Recent advances in machine learning (ML) technologies provided an opportunity to establish the material model directly from the data. This kind of data-based approach does not need postulating a mathematic expression.

ML is a general framework that could map inputs to the outputs directly by training the data set and has wide application in many field of civil engineering, such as the reliability-based design optimization (Fei et al. 2020) and structural probabilistic analysis (Lu et al. 2019). Because ML does not have any assumption of the mathematical form of the underlying relation and is driven directly by the data, it could be applied to any material. Additionally, the constitutive model driven by a ML algorithm is usually called a model-free constitutive model or data-driven constitutive model.

The earliest application of ML in material modeling was conducted by Ghaboussi et al. (1991). They used the artificial neural network (ANN), one type of ML algorithm, to capture the static material behavior (Ghaboussi et al. 1991) and rate-dependent behavior (Jung and Ghaboussi 2006) of concrete. They also proposed the autoprogressive training method and self-learning algorithm for training an ANN-based material model (Ghaboussi et al. 1998) and applied the obtained model into the finite-element analysis via ABAQUS (Hashash et al. 2004). In the following decades, the application of ANNs in material modeling became popular, and ANN-based constitutive models for viscoelastic material, rheological material, sand, and gravel, among others, were established as well (Ellis et al. 1995; Kessler et al. 2007; Penumadu and Zhao 1999; Settgast et al. 2019; Stoffel et al. 2018; Yang et al. 2020). Despite ANNs being a powerful tool that could provide the material model with high accuracy, their interpretability is poor, and their architecture needs to be selected a prior by trial and error.

In order to bypass these drawbacks, many researchers tried to apply other ML algorithms in constitutive modeling, such as support vector regression (SVR) by Zhao et al. (2015), symbolic regression by Versino et al. (2017), and evolutionary polynomial regression (EPR) (Javadi and Rezania 2009a, b; Javadi et al. 2009). Besides modeling the constitutive relation explicitly, some implicit material modeling methods were also implemented by some researchers. For example, Kirchdoerfer and Ortiz (2016) used the distance-minimization scheme, which is similar to the k-nearest algorithm, and maximum-entropy scheme (Kirchdoerfer and Ortiz 2017) to relate the material data set and conservation laws directly in both elastic problems (Kirchdoerfer and Ortiz 2016) and inelastic problems (Eggersmann et al. 2019). Ibanez et al. (2018) extracted material manifold from the data set by manifold learning and replaced the conventional constitutive model by material manifold

[1]Ph.D. Candidate, School of Civil Engineering, Univ. of Sydney, NSW 2006, Australia. ORCID: https://orcid.org/0000-0002-9503-2851. Email: baixi.chen@sydney.edu.au

[2]Professor, School of Civil Engineering, Univ. of Sydney, NSW 2006, Australia (corresponding author). ORCID: https://orcid.org/0000-0001-6243-4304. Email: luming.shen@sydney.edu.au

[3]Associate Professor, School of Civil Engineering, Univ. of Sydney, NSW 2006, Australia. Email: hao.zhang@sydney.edu.au

Note. This manuscript was submitted on May 19, 2020; approved on February 15, 2021; published online on May 26, 2021. Discussion period open until October 26, 2021; separate discussions must be submitted for individual papers. This paper is part of the ***ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering***, © ASCE, ISSN 2376-7642.

in both elastic problems (Ibañez et al. 2018) and plastic problems (Chinesta et al. 2017) as well.

Because there exist inevitable uncertainties in structural systems, stochastic structural analysis is often used to provide sufficient knowledge on the variability of the structural response (Schuëller and Pradlwarter 2006). There are many approaches to achieve the stochastic structural analysis, such as the stochastic finite-element method (Aldosary et al. 2018) and the Monte Carlo (MC) methods (Papadopoulos and Giovanis 2017). As a key step of stochastic structural analysis approaches, the quantification of material uncertainties (i.e., uncertainty of the material constitutive relation/stress-strain relation) has significant influence on the obtained stochastic structural responses. Without accurate evaluation of the material uncertainties, the obtained stochastic structural responses are not reliable. Conventionally, material uncertainty is represented by the random parameters of the constitutive model. When there is no suitable constitutive model available, especially for new materials, the material uncertainty cannot be determined accurately. ML-based approaches could generate the constitutive relation directly from the data and offer higher accuracy than the conventional parametric approaches (Freitag 2012; Ghaboussi et al. 1991; Javadi and Rezania 2009a; Zhao et al. 2015).

However, most of the current ML-based constitutive models, such as ANN-based (Ghaboussi et al. 1991) and SVR-based (Zhao et al. 2015) constitutive models, are only represented in the deterministic form without considering the material uncertainty, which limits the application of the ML-based constitutive models in the stochastic structural analysis. Thus, how to take advantage of ML-based constitutive models in stochastic structural analysis needs further investigation.

For considering the material uncertainty via ML technologies, the recurrent neural network was incorporated with the fuzzy theory to predict material behaviors by stress-strain data (Freitag et al. 2011; Graf et al. 2012) or load-displacement data (Freitag 2012), and the obtained model was applied to fuzzy stochastic FEM (Freitag et al. 2013). However, developing the fuzzy system requires much expertise, and the system is required to be redeveloped for each individual material (Godil et al. 2011). Ayensa-Jiménez et al. (2018) also considered the material intrinsic uncertainty and extended the data-driven scheme of Kirchdoerfer and Ortiz (2016) to the stochastic form. In this paper, a new approach, i.e., the Gaussian process regression (GPR)-based material model, is proposed to extract the material law with the material uncertainty included from the data set and can be directly used for stochastic structural analysis.

GPR is a nonparametric supervised learning method first introduced by Williams and Rasmussen (1996). It has been applied in various areas of civil engineering in recent years. For example, Hoang et al. (2016) used a GPR model to predict the concrete compressive strength based on the concrete ingredients. Su et al. (2017) replaced FEM by the GPR surrogate model to reduce the computational cost of the traditional MC method for the structural reliability analysis. However, very few researchers used the GPR model to represent the uncertain material constitutive relation. Because the relation captured by GPR is in the form of a stochastic function distributed as a Gaussian process, the constitutive relation given by GPR, that is the GPR-based constitutive model (GPR model), is a stochastic function as well and has both the expectation and corresponding uncertainty included. Instead of extracting the information from the data and incorporating them into parameters like ANN and EPR, GPR uses the Bayesian inference method to make the prediction directly from the training data (e.g., experimental data).

Hence, the GPR model has a direct relation to the experimental data. This feature gives the GPR model the capacity to model the constitutive behavior and uncertainty of novel materials for which there is no existing constitutive model. Besides, the transparent random expression of the GPR-based constitutive model makes it possible to be incorporated into the numerical simulation directly for stochastic structure analysis, which is hard to be achieved by other ML models because there are many parameters needed to be evaluated in other ML models, and it is difficult to determine the stochastic behavior of all parameters. Because this kind of stochastic structure analysis is achieved by a data-driven constitutive model, i.e., the GPR model, it is referred to the data-driven stochastic structure analysis here. Compared with the conventional material model–based stochastic analysis procedure, the GPR material model–based procedure can use the material uncertainty directly captured from the data and does not need to assume the mathematical format of the model and determine the probability distribution of the model parameters for different materials.

The remaining sections of the paper are organized as follows. The theories of GP and GPR are given below, which is followed by the implementation of the GPR to establish the material model. Then, a data-driven stochastic structural analysis procedure is demonstrated. Finally, the concluding remarks are given in the section "Summary and Conclusions."

Theory of Gaussian Process Regression

GPR is one of the supervised machine learning algorithms. It describes the underlying relation as a random function following the Gaussian process (GP) and use a Bayesian framework to infer the GP. Both the expected relation and corresponding uncertainty could be included in the obtained GP. The GPR algorithm and the model training procedure are explained below.

Gaussian Process Regression

GPR is conducted by three steps: (1) assume that the underlying relation distributes as a prior GP; (2) update a prior GP to a posterior GP in the light of the data set and Bayesian inference; and (3) use the obtained posterior GP to describe the distribution of the underlying relation.

A prior GP could be set up by including some known information about the underlying relation. Appropriate prior knowledge could increase the accuracy of a posterior GP. But most of time, there is not sufficient knowledge about the underlying relation. In this situation, zero is often chosen as the mean function, and the squared exponential covariance function [Eq. (1)] is a popular choice for the covariance function. More options for the covariance function have been discussed by Williams and Rasmussen (2006)

$$k(x, x') = \sigma_f^2 \exp\left[\frac{-(x - x')^2}{2\sigma_l^2}\right] \tag{1}$$

where σ_f = signal standard deviation; and σ_l = characteristic length scale. Both σ_f and σ_l are the hyperparameters of GPR and determined in the training process.

In the single input variable to single-output response problems, because both the training data set $\mathcal{D}(\mathbf{x}^*, \mathbf{y}^*)$ and the arbitrary point (x, y) are collected from a same unknown stochastic function, in other words a same unknown GP, both training data and the arbitrary point should comply with the multivariable Gaussian distribution (Rasmussen 2003)

$$\begin{bmatrix} \mathbf{y}^* \\ y \end{bmatrix} \backsim \mathcal{N}\left(\begin{bmatrix} \mathbf{m}(\mathbf{x}^*) \\ m(x) \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{**} & \mathbf{K}_*^{\mathbf{T}} \\ \mathbf{K}_* & \mathbf{K} \end{bmatrix} \right) \tag{2}$$

$$\mathbf{K}_{**} = \begin{bmatrix} k(x_1^*, x_1^*) & k(x_1^*, x_2^*) & \cdots & k(x_1^*, x_n^*) \\ k(x_2^*, x_1^*) & k(x_2^*, x_2^*) & \cdots & k(x_2^*, x_n^*) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n^*, x_1^*) & k(x_n^*, x_2^*) & \cdots & k(x_n^*, x_n^*) \end{bmatrix} \tag{3}$$

$$\mathbf{K}_* = [\, k(x_1^*, x) \quad k(x_2^*, x) \quad \cdots \quad k(x_n^*, x) \,] \tag{4}$$

$$\mathbf{K} = k(x, x) \tag{5}$$

where $\mathbf{x}^*$ is the training inputs vector $\{x_1^*, x_2^*, \ldots, x_i^*, \ldots, x_n^*\}$, and each element of $\mathbf{x}^*$ is one input sample of the training data set; $\mathbf{y}^*$ is the training output vector $\{y_1^*, y_2^*, \ldots, y_i^*, \ldots, y_n^*\}$, and each element of $\mathbf{y}^*$ is one output sample of the training data set; x = single arbitrary input; and y = unknown output or prediction at the arbitrary inputs x.

By Bayesian inference, the unknown output y follows the normal distribution as follows:

$$y \sim \mathcal{N}(m(x) + \mathbf{K}_* \mathbf{K}_{**}^{-\mathbf{1}}(\mathbf{y}^* - \mathbf{m}(\mathbf{x}^*)), k(x, x) - \mathbf{K}_* \mathbf{K}_{**}^{-\mathbf{1}} \mathbf{K}_*^{\mathbf{T}}) \tag{6}$$

The obtained y is just a specific case of the posterior GP. After generalizing this case, the posterior GP could be accessed easily and expressed as $\mathcal{GP}(m_D, k_D)$ with mean function m_D and covariance function k_D

$$m_D(x) = m(x) + \mathbf{K}_* \mathbf{K}_{**}^{-\mathbf{1}}(\mathbf{y}^* - \mathbf{m}(\mathbf{x}^*)) \tag{7}$$

$$k_D(x, x') = k(x, x') - \mathbf{K}(\mathbf{x}^*, x) \mathbf{K}_{**}^{-\mathbf{1}} \mathbf{K}(\mathbf{x}^*, x')^{\mathbf{T}} \tag{8}$$

$$\mathbf{K}(\mathbf{x}^*, x) = [\, k(x_1^*, x) \quad k(x_2^*, x) \quad \cdots \quad k(x_n^*, x) \,] \tag{9}$$

$$\mathbf{K}(\mathbf{x}^*, x') = [\, k(x_1^*, x') \quad k(x_2^*, x') \quad \cdots \quad k(x_n^*, x') \,] \tag{10}$$

Because the underlying relation obeys the posterior GP, each sample of the posterior GP is one possibility of the underlying relation. The mean function m_D represents the expected relationship captured from the data sets and could be expressed as the weighted average of the training output data with the weight $(\mathbf{K}_* \mathbf{K}^{-\mathbf{1}})$ reflecting the correlation between the test points and training points. In addition, $k_D(x, x)$ is the variance of the expected relationship at input x. Furthermore, the posterior GP has a transparent expression.

Although the aforementioned explanation is based on the single input, it can be extended to multiple inputs easily by some transformation, e.g., by changing the covariance function into the vector function. The multi-input counterpart of Eq. (1) is given in Eq. (11)

$$\begin{aligned} k(\mathbf{x}, \mathbf{x}') &= \sigma_f^2 \exp\left[\frac{-(\mathbf{x} - \mathbf{x}')(\mathbf{x} - \mathbf{x}')^T}{2\sigma_l^2} \right] \\ &= \sigma_f^2 \exp\left[\frac{-(\mathbf{x} - \mathbf{x}')^2}{2\sigma_l^2} \right] \end{aligned} \tag{11}$$

where $\mathbf{x}$ and $\mathbf{x}'$ are row vectors containing multi-inputs and represent two different positions as well but in vector form.

As suggested by Williams and Rasmussen (1996), it is feasible to develop an independent GPR model for each output separately in the multioutputs problem, and this approach is also known as a multikriging approach (Boyle and Frean 2005). If higher accuracy is required, some advanced multioutput approaches, such as intrinsic coregionalization model and linear model of coregionalization (Liu et al. 2018), could be used to consider the correlation between the outputs. For simplicity, the independent GPR model is constructed for each output in this paper.

When there is the noise or uncertainty in the data, constant Gaussian noise can be implemented into the GPR model simply by replacing $\mathbf{K}_{**}$ by $(\mathbf{K}_{**} + \sigma_n^2 \mathbf{I})$, where σ_n represents the intensity of noise or uncertainty of the data. For material data, σ_n represent the material uncertainty; σ_n is usually a constant or a function of inputs. For describing the constitutive relation, the exponential noise model [Eq. (12)] is recommended and its parameters, a and t, could be determined by the training processes

$$\sigma_n = a(1 - \exp(-t|\mathbf{x}|)) \tag{12}$$

where $|\mathbf{x}|$ = Euclidean norm of $\mathbf{x}$.

Training of GPR

With the advantage of Bayesian inference, there is no need to modify the parameters during the regression procedures, and the data sets are combined into the model directly. But the reliability of the model depends on the hyperparameters $\boldsymbol{\theta}$, which consist of the parameters of the mean function, parameters of the covariance function, and parameters of noise model. The process of optimizing the hyperparameters $\boldsymbol{\theta}$ is referred as training the GPR model.

If enough knowledge on the hyperparameters is known, the training process of the GPR model could be omitted. However, this kind of situation is rare. The optimization of the hyperparameters is conducted by maximizing the marginal likelihood of the training data. For convenience, the logarithm of the marginal likelihood is often considered (Ebden 2015) and is expressed

$$\begin{aligned} L &= \log(p(\mathbf{y}^*|\mathbf{x}^*, \boldsymbol{\theta})) \\ &= -\frac{1}{2}\log(|\mathbf{K}_{**}|) - \frac{1}{2}(\mathbf{y}^* - \mathbf{m}(\mathbf{x}^*))^T \mathbf{K}_{**}^{-1}(\mathbf{y}^* - \mathbf{m}(\mathbf{x}^*)) \\ &\quad - \frac{n}{2}\log(2\pi) \end{aligned} \tag{13}$$

The optimization could be conducted by any numerical optimization routines, such as the gradient descent method. The whole regression procedure of the GPR model is summarized as follows:

1. Collect the training data set $\mathcal{D}(\mathbf{x}^*, \mathbf{y}^*)$.
2. Select the covariance function $k(x, x)$ and mean function $m(x)$ according to the prior information.
3. Determine the hyperparameters $\boldsymbol{\theta}$ by maximizing the marginal likelihood [Eq. (13)] of the training data.
4. Use the mean function [Eq. (7)] and the covariance function [Eq. (8)] of the posterior Gaussian process to predict the expected value and corresponding variance at any arbitrary input points.

The rigorous proof of the GPR algorithm will not be provided here, and further details have been given elsewhere (Ebden 2015; Rasmussen 2003; Williams and Rasmussen 1996, 2006).

Application of GPR in Material Modeling

When GPR is applied to model the constitutive relation from the material data, the obtained constitutive relation is regarded as a random function distributed as a posterior GP. Because the GPR-based constitutive model (GPR model) is a stochastic relation, the mean function $m_D(x)$ of the stochastic relation is used to represent the expected constitutive relation, and the material uncertainty is inherently reflected in the covariance function of the posterior GP. All possible constitutive relations caused by the material uncertainty could be accessed by sampling the posterior GP.

Table 1. Inputs and outputs for the GPR model

Model type	Inputs	Outputs
1D material model	$\{\varepsilon\}$	$\{\sigma\}$
2D material model	$\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$	$\{\sigma_1, \sigma_2, \sigma_3\}$
3D material model	$\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6\}$	$\{\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6\}$

Note: 1D = stress and strain components represented in one direction; 2D = stress and strain components represented in the plane; 3D = stress and strain components represented in a real space; and ε_3 and σ_3 in the 2D model, and $\varepsilon_4, \varepsilon_5, \varepsilon_6$ and $\sigma_4, \sigma_5, \sigma_6$ in the 3D model = shear strain and stress, respectively, and $\varepsilon_1, \varepsilon_2$ and σ_1, σ_2 in the 2D model and $\varepsilon_1, \varepsilon_2, \varepsilon_3$ and $\sigma_1, \sigma_2, \sigma_3$ in the 3D model = normal strain and stress, respectively.

Input and Output Parameters

The process of selecting input and output parameters for GPR usually depends on the problem types. In mechanic problems, the constitutive relation is often established between stress and strain. In this study, the strain ε is chosen as the input and the stress σ is chosen as output to capture the mechanical behavior of a material. The inputs and outputs for problems of different dimensions are given in Table 1. The inversion with the stress as the input and strain as outputs is also acceptable, although it is not used here.

Developed GPR Model

The data set is split into two parts, training and testing. The training set takes up 85% of the data points, and the testing set takes up 15% of the data points. Because there is usually no information available about the underlying material behavior, zero is often chosen as the prior mean function $m(x)$, and the squared exponential covariance function [Eqs. (1) and (11)] is considered here. Following the training process described in the section "Training of GPR."

The material data points from a single test are usually not independent from each other, but the GPR model does not directly take this correlation into consideration. As a result, the possible constitutive relation given by GPR model would fluctuate and cannot represent the possible constitutive relation, which is usually smooth. In order to resolve this problem, it is here assumed that all samples should have the same trend as the mean function of the posterior Gaussian process with the same variance as the posterior Gaussian process. Under this assumption, the constitutive relation can be expressed as Eq. (14). Each possible stress-strain relation can be sampled from Eq. (14) by randomly sampling the random variable ξ from $\mathcal{N}(0, 1)$ and substituting it into Eq. (14). By this treatment, the correlation coefficient of the stresses at different strains is 1, which means the posterior Gaussian process is restricted by the perfect linear correlation condition

$$\sigma(\boldsymbol{\varepsilon}) = m_D(\boldsymbol{\varepsilon}) + \xi \cdot \sqrt{k_D(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon})} \tag{14}$$

with

$$m_D(\boldsymbol{\varepsilon}) = \mathbf{K}_* \mathbf{K}_{**}^{-1} \boldsymbol{\sigma}^{\mathbf{d}} = \sum_{i=1}^{n} s_i \sigma_f^2 \exp\left[\frac{(\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{d}} - \boldsymbol{\varepsilon})^2}{-2\sigma_l^2}\right]$$

$$k_D(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}) = k(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}) - \mathbf{K}_* \mathbf{K}_{**}^{-1} \mathbf{K}_*^{\mathbf{T}} = \sigma_f^2 + \sigma_n^2 - \sum_{i=1}^{n} \sum_{j=1}^{n} o_{ij} \sigma_f^4 \exp\left[\frac{(\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{d}} - \boldsymbol{\varepsilon})^2 + (\boldsymbol{\varepsilon}_\mathbf{j}^{\mathbf{d}} - \boldsymbol{\varepsilon})^2}{-2\sigma_l^2}\right]$$

where ξ = standard Gaussian random variable complying $\mathcal{N}(0, 1)$; s_i = elements of a constant vector $\mathbf{s}$ ($\mathbf{s} = \mathbf{K}_{**}^{-1} \cdot \boldsymbol{\sigma}^{\mathbf{d}}$); $\boldsymbol{\sigma}^{\mathbf{d}}$ is the vector containing n stress data in the training data set; $\boldsymbol{\varepsilon}^{\mathbf{d}}$ is the matrix (each row of the matrix represents one training strain vector, and each element in the row represents one strain component) containing n strain data in the training data set; $\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{d}}$ = ith row of the training inputs matrix $\boldsymbol{\varepsilon}^{\mathbf{d}}$; σ_i^d = ith element of training stress vector $\boldsymbol{\sigma}^{\mathbf{d}}$; and o_{ij} = element of the constant matrix $\mathbf{K}_{**}^{-1}$.

Model Evaluation Methods

In order to evaluate the GPR model and compare it with other methods, two indexes, namely the mean square error (MSE) and coefficient of determination (COD), are used in this study to quantify the accuracy of the mean function of the GPR model. The two indexes are calculated based on the testing set. The one-dimensional (1D) forms of the two indexes are given here, and the high-dimensional form can be expressed in a similar way

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (\sigma_i^D - \sigma(\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{D}}))^2 \tag{15}$$

$$\text{COD} = 1 - \frac{\sum_{i=1}^{n} (\sigma_i^D - \sigma(\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{D}}))^2}{\sum_{i=1}^{n} (\sigma_i^D - \frac{1}{n} \sum_{i=1}^{n} \sigma_i^D)^2} \tag{16}$$

where σ_i^D = output stress of the data set; $\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{D}}$ = input strain of the data set; σ = mean function of GPR model; and n = number of data.

If the underlying constitutive model is known, the average absolute true error (AAE) and average relative true error (ARE) between the GPR model and true model could be calculated. They could be evaluated by selecting m test inputs uniformly in the domain covered by the training data and averaging the absolute error and relative error between the true model and the GPR model at this m test inputs. The expressions of AAE and ARE are

$$\text{AAE} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (\sigma_t(\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{t}}) - \sigma(\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{t}}))^2} \tag{17}$$

$$\text{ARE} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} \left(\frac{\sigma_t(\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{t}}) - \sigma(\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{t}})}{\sigma_t(\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{t}})}\right)^2} \tag{18}$$

where σ_t = true relation; σ = mean function of the GPR model; and $\boldsymbol{\varepsilon}_\mathbf{i}^{\mathbf{t}}$ = ith test input.

Numerical Examples

In order to illustrate the behavior of GPR models, two examples are reported next. In the first example, a 1D GPR material model is established from the experimental data of the aluminum samples under the unidirectional loading, and the effect of the data size is discussed. The obtained GPR material models are also compared with other machine learning methods, including ANN and SVR. In the second example, the two-dimensional (2D) synthetic data on unidirectional loading generated from a hypothetical underlying constitutive model are used to train the 2D GPR model. The obtained 2D material model is compared with the true material model, which is assumed a prior for data generation. A three-dimensional (3D) GPR model is not given here, but it could be developed following the same way as the 2D model. In addition, this paper mainly focuses on the static unidirectional loading situation, and dynamic problems with hysteretic characteristics will be tackled in future work.

One-Dimensional Example

The training data for the 1D GPR model are from the uniaxial tensile test on aluminum alloy 7075-T651, where T651 means the

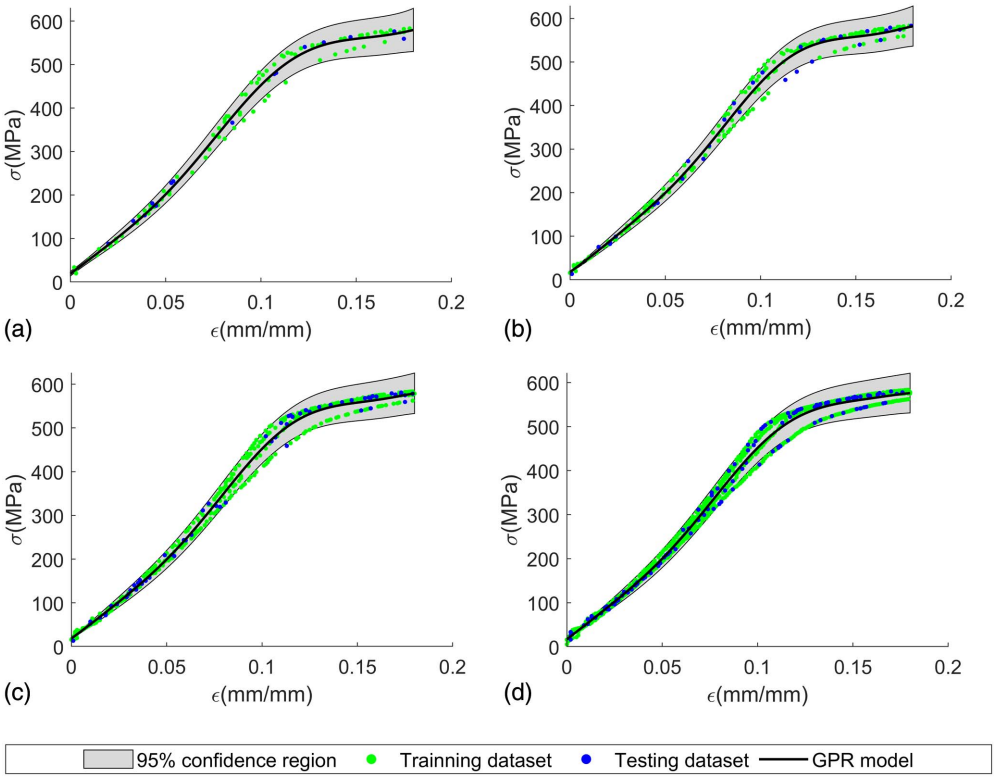

Fig. 1. Four different GPR models obtained by using the four data sets of (a) 100 data points; (b) 200 data points; (c) 400 data points; and (d) 800 data points.

peak strength of the alloy (Børvik et al. 2010). Three plate-type specimens are manufactured for the test according to ASTM E8 (ASTM 2011) and do not have any loading history. The test is conducted by a universal testing machine [MTS Criterion (Eden Prairie, Minnesota), with a maximum loading capacity of 50 kN] at room temperature. The loading process is quasi-static with crosshead velocity of 1.0 mm/min, i.e., a strain rate of 0.0007 s^{-1}.

A total of 800 stress-strain data points are obtained from three tests. To evaluate the influence of the data set size on the performance of the 1D GPR model, three data sets are created by randomly selecting 100, 200, and 400 data points from the experimental results. The four data sets with 100, 200, 400, and 800 data pairs are split into training set (85%) and testing set (15%) and are shown in Fig. 1.

The GPR models for the four data sets (Fig. 1) are developed by the procedure described in the section "Developed GPR Model." The optimal hyperparameters obtained by the training processes are presented in Table 2. Because the GPR model is a random function, the expected relation and its 95% confidence region are used to represent the captured constitutive relation and the material uncertainty, and both are plotted in Figs. 1(a–d). Table 3 also presents the MSE and COD of the GPR model on the testing set.

It can be observed from Fig. 1 that the four developed GPR models all can capture both the material behavior and the corresponding uncertainty very well. Both the training data and testing data are covered by the 95% confidence region given by the GPR model. As indicated in Table 2, the optimal hyperparameters are could be obtained in small data set sizes. All hyperparameters stay in a small interval, even though the data set size increases threefold. The COD of GPR model is above 99% for different data set sizes.

Because the testing set is small, the MSE of different data set sizes is influenced by the randomly selected testing data and varies greatly. It can be concluded that small data set size is sufficient here, and even 100 data points can lead to a GPR model of high accuracy with COD of above 99%. Besides, the computational cost of the GPR model increases with the data set size. Because the objective function Eq. (13) in the training process needs the inverse of the matrix, the computational complexity will increase cubically with the data set size. Hence, for problems with large data set size, some sparse approximation technologies, such as online sparse matrix GP (Ranganathan et al. 2010) and sparse spectrum technology (Lázaro-Gredilla et al. 2010), could be used to improve the efficiency, although they are not adopted in this study.

Table 2. Hyperparameters of GPR models developed by training the four data sets of different sizes

	σ_n			
Data size	a	t	σ_f	σ_l
100	30.05	8	354.99	0.0661
200	28.69	8	301.06	0.0504
400	28.59	9	324.46	0.0638
800	28.20	9	301.13	0.0615

Table 3. MSE and COD of the three ML-based material models developed by training four data sets of different sizes on the testing set

	MSE (MPa^2)			COD (%)		
Data size	GPR	SVR	ANN	GPR	SVR	ANN
100	85.58	93.44	111.55	99.70	99.71	99.66
200	267.77	339.09	357.05	99.04	99.00	98.95
400	151.13	270.54	272.61	99.63	99.32	99.32
800	249.59	378.34	260.20	99.32	98.97	99.29

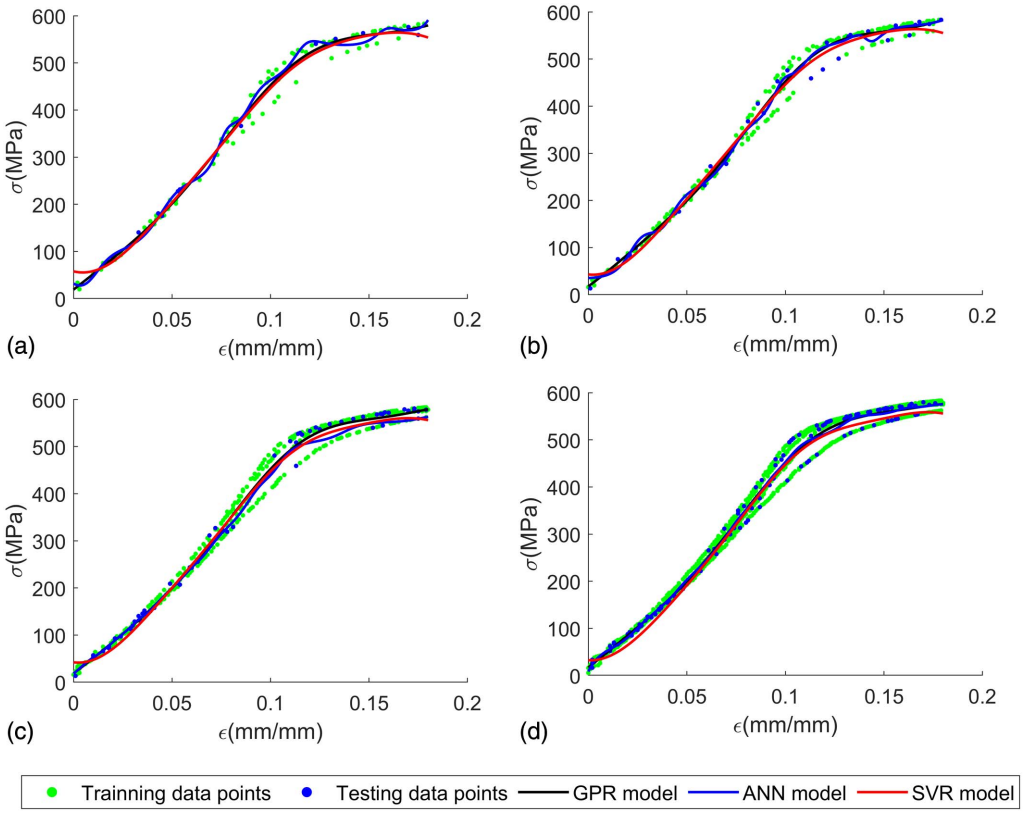

Fig. 2. ML-based material models obtained by using the four data sets of (a) 100 data points; (b) 200 data points; (c) 400 data points; and (d) 800 data points.

To compare the performance of the GPR models with other ML approaches, two additional popular ML algorithms, ANN and SVR, are used to develop constitutive models as well. The architecture of the ANN-based material model, suggested by Ghaboussi and Sidarta (1998) and Sidarta (1998), contains two hidden layers with six nodes each layer, and the Levenberg-Marquardt method is used to train it. SVR based on the Gaussian kernel function is used to establish the material model and is trained by the iterative single data algorithm. The data should be normalized before training the three ML-based models, and 15% of the data should be separated for testing.

The material models obtained using the ANN, SVR, and GPR model, are shown in Figs. 2(a–d). Table 3 illustrates the MSE and COD of the three ML-based material models, respectively.

As shown in Fig. 2, the ANN-based material model is significantly influenced by the data set size, and its curve is not smooth when the sample size is smaller than 800. The fluctuation is obvious over the whole domain when the data set size is 100, and local fluctuations still exit when the data set size increases to 200 and 400. By contrast, both the expected relation of the GPR model and SVR model present smooth behavior and high robustness to the variation of the data set size. The GPR model, however, has slightly lower MSE and higher COD than the SVR model and ANN model on the testing data set, as illustrated in Table 3. Although all three models have the close accuracy with COD higher than 99%, the ANN and SVR models can only provide the deterministic constitutive relation, which makes it hard to apply them to stochastic structural analysis. On the other hand, the GPR model can capture both constitutive relation and material uncertainty and can be directly used in stochastic structural analysis.

Two-Dimensional Example

In order to achieve the comparison between the GPR model and the underlying true constitutive relation, which is always unknown in the real world, a hypothetical biaxial test under the material isotropy and plane stress condition is conducted to generate the 2D stress-strain data. The Ramberg-Osgood (RO) model (Ramberg and Osgood 1943), as expressed in Eq. (19), is assumed as the true material model

$$\varepsilon = \frac{\sigma}{E} + \frac{2\beta\sigma_0}{3E}\left(\frac{\sigma}{\sigma_0}\right)^n \tag{19}$$

where model parameter $\beta = 1.25$ and $n = 5$; E has a mean of 10 GPa and a standard deviation of 0.5 GPa; and σ_0 has a mean of 90 MPa and a standard deviation of 4.5 MPa. Both E and σ_0 are normal random variables with coefficient of variation (COV) of 0.05.

In the biaxial test, only a limited number of loading paths are available, so nine different loading paths ($\sigma_1/\sigma_2 = -1/-1$, $-1/-0.5$, $-1/0$, $-1/0.5$, $-1/1$, $1/-0.5$, $1/0$, $1/0.5$, and $1/1$) and a Poisson ratio of 0.3 are considered here. For each loading path, 100 stress σ_1 values between 0 and 60 MPa with equal interval were sampled to obtain 100 stress σ_2 values by using σ_1/σ_2 ratios, and the stress σ_3 values are set to 0 MPa. After that, 100 stress sets [σ_1, σ_2, σ_3] are used to generate 100 strain sets [ε_1, ε_2, ε_3] by using the RO model with random parameters, and 100 synthetic stress-strain sets [σ_1, σ_2, σ_3, ε_1, ε_2, ε_3] for each loading path are obtained.

The data set is then extended by the data set enhancement strategy via two stages (Shin and Pande 2002). The first stage is a two-fold expansion by exchanging two axes of material data points. As a result, a total of 1,800 data points is achieved. In the second stage,

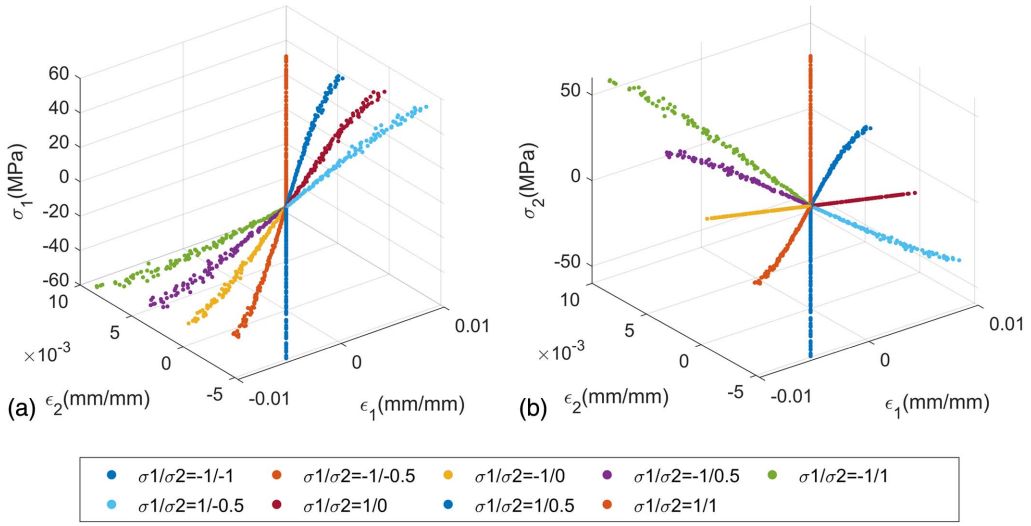

Fig. 3. Stress-strain sample data sets under different biaxial loading conditions: (a) data set with respect to σ_1 (ε_3, $\sigma_3 = 0$); and (b) data set with respect to σ_2 (ε_3, $\sigma_3 = 0$).

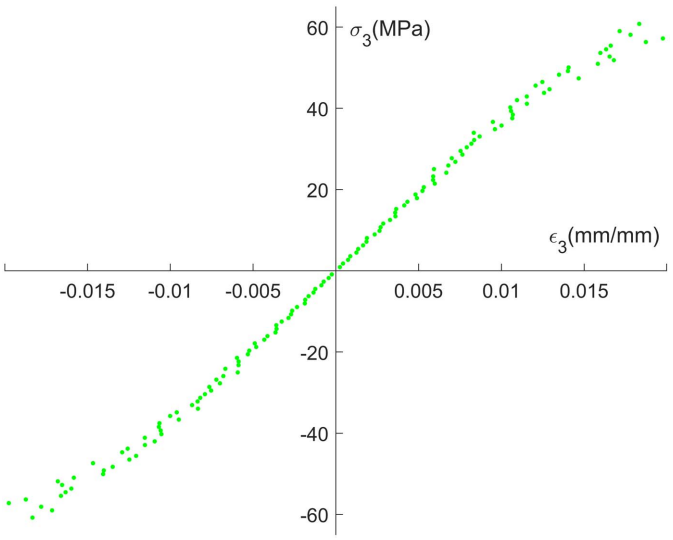

Fig. 4. Stress-strain sample data points under the pure shear condition.

the biaxial data are rotated by 15°, 30°, 45°, 60°, and 75° so that the shear stress and shear strain (including pure shear) are considered in the data set. As a result, the size of data set is extended to 10,800 (six sets of 1,800 data points). The pure shear condition has been included when the data under loading path $\sigma_1/\sigma_2 = -1/1$ are rotated by 45°. Identical data are deleted from the expanded data set, and the final size of the data set is 8,600. No extra noise is added to the synthesized data. The uncertainty of the synthesized data comes from the variation of the parameters in the RO model. Some parts of the data are shown in Figs. 3(a and b) and 4.

The data set is split into a training set (85%) and testing set (15%). After feeding the training data to the GPR model, the 2D GPR model can be obtained, and the hyperparameters are given in Table 4. The expected relation given by the 2D GPR model is shown in Figs. 5(a and b). Similar to the 1D example, the 2D GPR model has a COD over 99.8% (Table 5) and is highly consistent with the data, which can be observed directly from Figs. 5(a and b).

The optimal hyperparameters of the GPR models for σ_1 and σ_2 are close, implying that the correlations between the test points and

Table 4. Optimum hyperparameters of 2D GPR model

Outputs	σ_n		σ_f	σ_l
	a	t		
σ_1	2.35	147	36.90	0.0086
σ_2	2.35	147	36.45	0.0084
σ_3	1.40	100	27.86	0.0083

training points should be similar in both the GPR models. As illustrated in the section "Training of GPR," the prediction can be regarded as the weighted average of the training output data, with the weight reflecting the correlation between the test points and training points, and a similar correlation can lead to a similar weight. So, when the correlation is the similar, the similar training data can give the similar prediction at the test points, and the symmetric training data give the symmetric prediction. This dramatic property of the GPR models means that the material symmetry embedded in the training data could be reflected in the GPR models.

Besides, the obtained 2D GPR model can be generalized to any arbitrary loading paths, even though those paths are not included in the training data. To show this property, three additional true loading paths ($\sigma_1/\sigma_2 = 1/-0.3$, $\sigma_1/\sigma_2 = 1/0.7$, and $\sigma_2 = -20$ MPa), which are not used for training, are tested here. As shown in Figs. 6(a and b), the prediction of the 2D GPR model for the untrained loading paths is highly consistent with the true relation, which is derived based on Eq. (19).

Because the true material behavior is assumed a priori, it is possible to evaluate the true error of the GPR model. AAE and ARE of the GPR model are calculated here and presented in Table 6 for this purpose. The AAE over the domain is smaller than 0.5 MPa, and the AREs for all stress components are smaller than 2.3%. In addition, like other ML methods, the GPR model is poor at extrapolation, so the error of the prediction will increase greatly if the maximum strain or stress is not covered by the training data sets.

Data-Driven Stochastic Structural Analysis

Because the GPR-based constitutive model (GPR model) is a stochastic function whose distribution follows the posterior GP, which

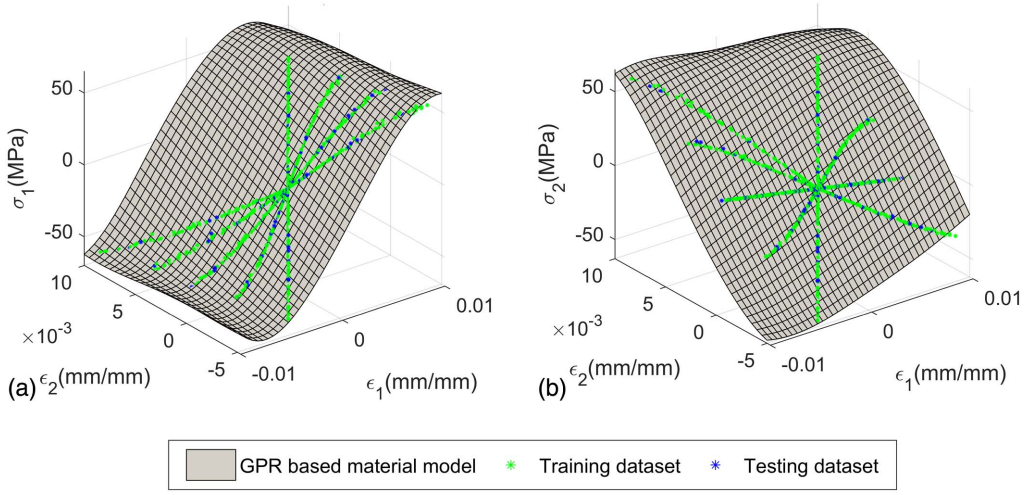

Fig. 5. GPR model under biaxial loading conditions: (a) GPR model with respect to σ_1 (ε_3, $\sigma_3 = 0$); and (b) GPR model with respect to σ_2 (ε_3, $\sigma_3 = 0$).

Table 5. MSE and COD of 2D GPR models on the testing set

Output	MSE	COD (%)
σ_1	1.046	99.82
σ_2	1.034	99.82
σ_3	0.376	99.83

has the transparent expression, it is tractable to access the stochastic structural behavior, which is caused by the material uncertainty, via incorporating the GPR model into structural analysis. Due to the application of the data-driven constitutive model, this kind of stochastic structural analysis is called data-driven stochastic structure analysis. Different from the conventional stochastic structural analysis procedure, whose material uncertainty is represented by the random parameters of the constitutive model, the material uncertainty in the proposed procedure is captured directly from the data. The proposed data-driven stochastic structure analysis is suitable for a structure where there is no proper constitutive model to describe the material uncertainty. In this section, the implementation of the GPR material model into FEM is illustrated first, then the data-driven stochastic structural analysis and a numerical example are explained.

Implementation of GPR Model in FEM

The GPR model can be encoded into ABAQUS version 6.14 by a user-defined subroutine UMAT for finite-element analysis. In the UMAT, the Jacobian matrix, i.e., material stiffness matrix, is required to update the stresses for each increment. As suggested in the ABAQUS user manual and by Faramarzi et al. (2013), the Jacobian matrix for the static problem is expressed

$$\mathbf{J} = \frac{\partial(d\boldsymbol{\sigma})}{\partial(d\boldsymbol{\varepsilon})} = \frac{d\boldsymbol{\sigma}}{d\boldsymbol{\varepsilon}} \tag{20}$$

Because the GPR model is expressed in a random form [Eq. (14)], the corresponding Jacobian matrix is also expressed in a random form. The 1D random Jacobian matrix (1×1) can

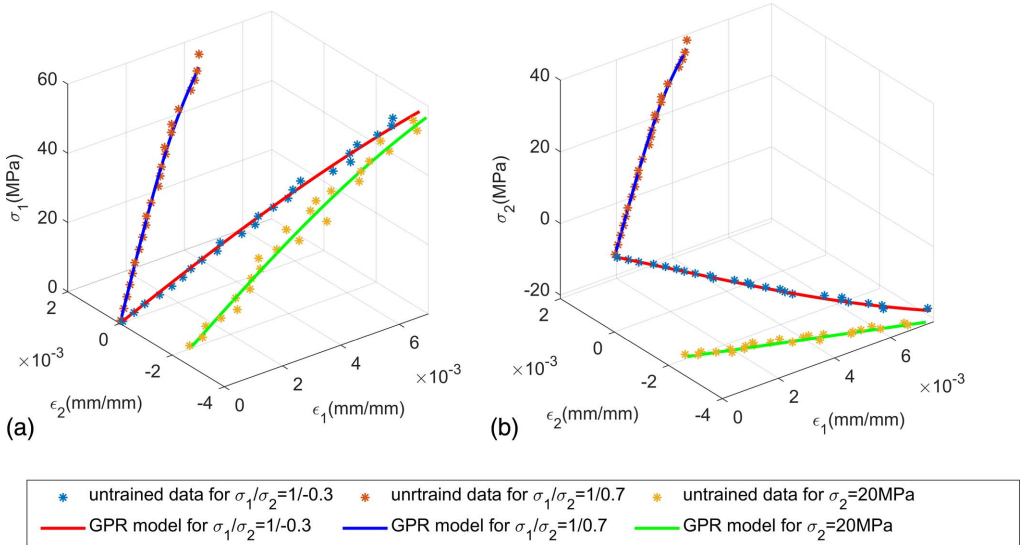

Fig. 6. Prediction of the GPR model for the untrained loading paths: (a) loading paths with respect to σ_1; and (b) loading paths with respect to σ_2.

Table 6. AAE and ARE of 2D GPR model

Evaluation methods	σ_1	σ_2	σ_3
AAE (MPa)	0.4841	0.4684	0.2335
ARE (%)	2.24	2.14	0.91

be obtained by substituting Eq. (14) into Eq. (20) and can be expressed as Eq. (21). The 2D and 3D random Jacobian matrixes could be achieved by following the similar approach. After encoding the corresponding random Jacobian matrix into the UMAT, the GPR model could be implemented by ABAQUS for finite-element analysis

$$\mathrm{J}=\frac{d\sigma}{d\varepsilon}=\sum_{i=1}^{n}\left\{s_i\sigma_f^2\exp\left[\frac{-(\varepsilon_i^d-\varepsilon)^2}{2\sigma_l^2}\right]\times\frac{(\varepsilon_i^d-\varepsilon)}{\sigma_l^2}\right\}+\xi\cdot\frac{1}{2}\cdot\left(\sigma_f^2+\sigma_n^2-\sum_{i=1}^{n}\sum_{j=1}^{n}o_{ij}\sigma_f^4\exp\left[\frac{-(\varepsilon_i^d-\varepsilon)^2-(\varepsilon_j^d-\varepsilon)^2}{2\sigma_l^2}\right]\right)^{-0.5}\cdot\left(2\sigma_n\cdot\frac{d\sigma_n}{d\varepsilon}-\sum_{i=1}^{n}\sum_{j=1}^{n}o_{ij}\sigma_f^4\exp\times\left[\frac{(\varepsilon_i^d-\varepsilon)^2+(\varepsilon_j^d-\varepsilon)^2}{-2\sigma_l^2}\right]\frac{(\varepsilon_i^d-\varepsilon)+(\varepsilon_j^d-\varepsilon)}{\sigma_l^2}\right) \tag{21}$$

Implementation of Data-Driven Stochastic Structural Analysis

It is always hard to find the influence of material uncertainty on the structural response in closed form. Hence, the data-driven stochastic structural analysis is based on the MC simulation. The detailed implementation procedure is given next:

1. Use the experimental or synthetic material data to develop the GPR model.
2. N possible stress-strain relations are sampled from the GPR model by using Eq. (14). If there are any other sources of uncertainty, such as the loads, they can also be sampled at this stage.
3. Run N structural analysis via FEM by using the N sampled stress-strain relations and the N sampled other uncertainty sources. After that, the N possible structural responses are obtained $\{R_1, R_2, \ldots, R_i, \ldots, R_N\}$.
4. The probability of failure and the structural reliability can be estimated based on the N possible structural responses.

   The probability of failure is

$$P_f=f(R>R_c)=\frac{N_H}{N} \tag{22}$$

where N_H = number of structural responses R_i that exceed R_c; and R_c = critical or allowable response of the structure, such as the drift limit or the deflection limit.

Structural reliability is given

$$P_r=1-P_f=f(R<R_c)=1-\frac{N_H}{N} \tag{23}$$

Numerical Example

To demonstrate the data-driven stochastic structural analysis, a planar frame is analyzed here as an example. The 2D frame is adopted from Nie and Ellingwood (2005) with slight modifications to the dimensions and member section. The modified dimensions of the frame are given in Fig. 7. The modified section of beams is a 200×400-mm rectangular section, and the modified section of columns is a 400×400-mm rectangular section. The loads applied on the frame are random variables with the distribution presented in Table 7. In order to consider the influence of the data set sizes and uncertainty levels, nine synthetic 1D material data sets with three different data set sizes (200, 400, and 800) and three different uncertainty levels (COVs of E and $\sigma_0 = 0.05$, 0.10, and 0.15) are generated by using the RO model [Eq. (19), with E having a mean of 80 GPa, σ_0 having a mean of 50 MPa, $\beta = 1.25$, and $n = 5$]. The generated data sets are shown in Figs. 8(a–c). The RO model with random parameters (E and σ_0) is the true relation under the data.

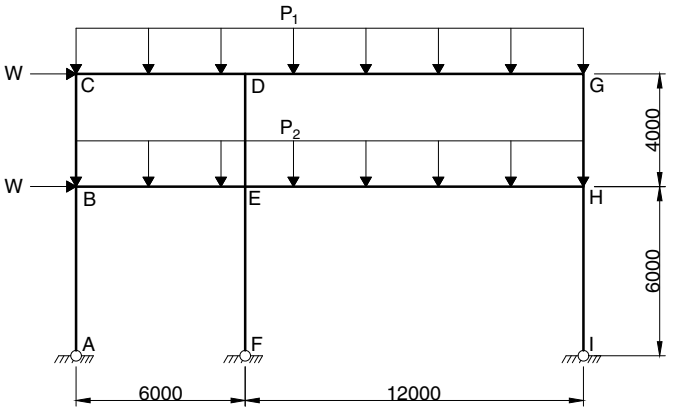

Fig. 7. Dimensions of the plane frame structure.

Table 7. Distributions of the loads

Description	Distribution	Mean	Coefficient of variation
P_1	Normal	15 kN/m	0.10
P_2	Normal	15 kN/m	0.10
W	Gamma	20 kN	0.37

Without loss of generality, the maximum horizontal deflection of the structure, that is the horizontal displacement at Node G, is studied here and other structural responses could be analyzed if one interested. The critical value (allowable value) of the deflection R_c is $H/500$, that is 20 mm, and the number of simulations N is set to 3,000. The ANN approach described by Papadopoulos and Giovanis (2017) is used to reduce the time cost of MC simulation for GPR model. Following the procedure given before, the expected value, variance, and probability of failure of the structure are presented in Table 8. In order to evaluate the performance of the proposed data-driven stochastic structural analysis, the reference structural stochastic response (true value) is determined by using true relation (the RO model). Here, 3,000 Monte Carlo simulations were carried out to determine the true response with the values of the parameters (E and σ_0) in the RO model and the loads being assigned according to their probability distributions in each simulation.

As indicated in Table 8, a larger data set size is required to obtain a good estimation of the expected deflection for the high uncertainty level, whereas the expectation is accurate on a small data set size for the low uncertainty level. This is consistent with the observation from one-dimensional example, namely that the small data set is sufficient when the uncertainty level is relatively low. Because the GPR model achieved the high accuracy at the data set of 400 points when the COV of the data set is 0.05, the data set of 800 points is not tested here for the COV of 0.05. In the light of the sufficient data, the expected deflection given by the GPR model has a relative error with respect to the true value smaller than 2%. On the other hand, the estimated standard deviation exhibits high accuracy with the relative error lower than 1%.

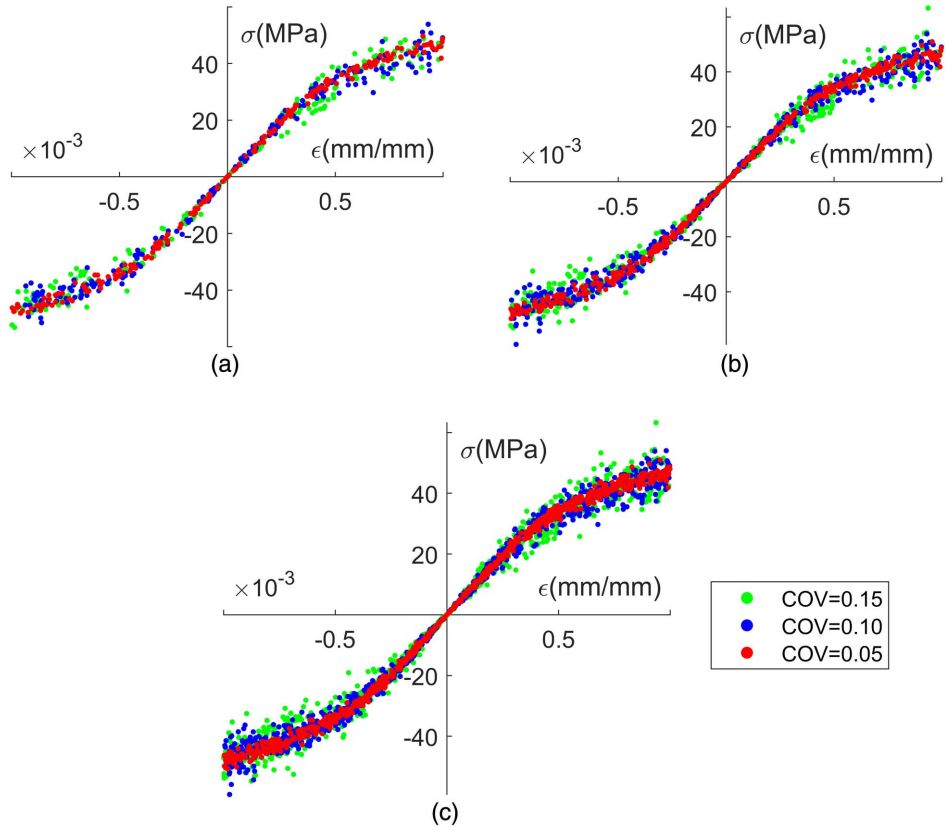

Fig. 8. Data sets containing (a) 200; (b) 400; and (c) 800 data points with three different uncertainty levels (COV = 0.05, 0.10, and 0.15).

Table 8. Statistical characteristics of the calculated structural horizontal deflection at Node G

COV	Data set size	Expectation (mm)		Standard deviation (mm)		Probability of failure (%)	
		True	GPR	True	GPR	True	GPR
0.15	200	12.50	12.97	5.82	6.05	10.7	12.8
	400		12.66		6.02		11.1
	800		12.73		5.79		11.0
0.10	200	12.40	12.26	5.46	5.63	9.3	9.5
	400		12.21		5.49		9.3
	800		12.33		5.46		9.3
0.05	200	12.09	12.19	5.19	5.29	8.1	8.2
	400		12.20		5.23		8.1

As for the probability of failure, its value increases with rising uncertainty levels, which suggests that the structural reliability could be low without the appropriate consideration of the material uncertainty. The probability of failure predicted by using the GPR model is conservative and will converge to the reference value with increasing numbers of data points. When a large data set size is available, the probability of failure has the relative error lower than 3%. In addition, for the low material uncertainty level, the GPR model could produce the probability of failure exactly.

Summary and Conclusions

Data-driven approaches are often used in the material modeling because they could be used to capture the mechanical behavior of new materials where no suitable mathematical expressions are available. A new approach called Gaussian process regression is introduced to model the material behavior in this paper. Different from other data-driven approaches that can only estimate the expectation, GPR captures the underlying relation using a stochastic function. As a result, both the underlying relation and corresponding uncertainty could be included.

In order to demonstrate the capacity of the GPR model, 1D experimental data and 2D synthetic data are applied. In the 1D example, the GPR model presents better performance than ANN and SVR models, with COD reaching 99% even though the data set size is small. When compared with the true constitutive relation underlying the 2D synthetic data, the GPR model also gives high accuracy with a relative true error smaller than 2.3%. Besides, the proposed GPR model focuses on the static unidirectional loading situation. Some materials may have hysteric behavior, and the proposed GPR model is not able to handle it at this stage. Dynamic problems with hysteretic characteristics will be tackled in future work.

After incorporating the data-driven constitutive model, that is the GPR model, into the structural analysis, the data-driven stochastic structural analysis is conducted. By comparing with the reference structural stochastic behavior determined by using the assumed true constitutive relation with the Monte Carlo simulations, it shows that the GPR model could give a good estimation of the expected response, variance, probability of failure, and the structural critical load when the data set size is large enough.

Data Availability Statement

All data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

This work was supported in part by the Australian Research Council through the Discovery Projects Scheme (Grant No. DP190102954).

References

Aldosary, M., J. Wang, and C. Li. 2018. "Structural reliability and stochastic finite element methods." *Eng. Comput.* 35 (6): 2165–2214. https://doi.org/10.1108/EC-04-2018-0157.

ASTM. 2011. *Standard test methods for tension testing of metallic materials*. ASTM E8/E8M-11. West Conshohocken, PA: ASTM.

Ayensa-Jiménez, J., M. H. Doweidar, J. A. Sanz-Herrera, and M. Doblaré. 2018. "A new reliability-based data-driven approach for noisy experimental data with physical constraints." *Comput. Methods Appl. Mech. Eng.* 328 (Jan): 752–774. https://doi.org/10.1016/j.cma.2017.08.027.

Børvik, T., O. S. Hopperstad, and K. O. Pedersen. 2010. "Quasi-brittle fracture during structural impact of AA7075-T651 aluminium plates." *Int. J. Impact Eng.* 37 (5): 537–551. https://doi.org/10.1016/j.ijimpeng.2009.11.001.

Boyle, P., and M. Frean. 2005. *Multiple output Gaussian process regression*. Wellington, New Zealand: Victoria Univ. of Wellington. https://doi.org/10.1.1.114.3898.

Chinesta, F., P. Ladeveze, R. Ibanez, J. V. Aguado, E. Abisset-Chavanne, and E. Cueto. 2017. "Data-driven computational plasticity." *Procedia Eng.* 207 (Jan): 209–214. https://doi.org/10.1016/j.proeng.2017.10.763.

Ebden, M. 2015. "Gaussian processes: A quick introduction." Preprint, submitted May 12, 2015. http://arxiv.org/abs/1505.02965.

Eggersmann, R., T. Kirchdoerfer, S. Reese, L. Stainier, and M. Ortiz. 2019. "Model-free data-driven inelasticity." *Comput. Methods Appl. Mech. Eng.* 350 (Jun): 81–99. https://doi.org/10.1016/j.cma.2019.02.016.

Ellis, G., C. Yao, R. Zhao, and D. Penumadu. 1995. "Stress-strain modeling of sands using artificial neural networks." *J. Geotech. Eng.* 121 (5): 429–435. https://doi.org/10.1061/(ASCE)0733-9410(1995)121:5(429).

Faramarzi, A., A. A. Javadi, and A. Ahangar-Asr. 2013. "Numerical implementation of EPR-based material models in finite element analysis." *Comput. Struct.* 118 (Mar): 100–108. https://doi.org/10.1016/j.compstruc.2012.10.002.

Fei, C.-W., H. Li, H.-T. Liu, C. Lu, B. Keshtegar, and L.-Q. An. 2020. "Multilevel nested reliability-based design optimization with hybrid intelligent regression for operating assembly relationship." *Aerosp. Sci. Technol.* 103 (Aug): 105906. https://doi.org/10.1016/j.ast.2020.105906.

Freitag, S. 2012. "Artificial intelligence for identification of material behaviour using uncertain load and displacement data." In *Proc., Int. Conf. on Scalable Uncertainty Management*, 606–611. Berlin: Springer.

Freitag, S., W. Graf, and M. Kaliske. 2011. "FE analysis using recurrent neural networks for uncertain stress-strain-time dependencies." *Proc. Appl. Math. Mech.* 11 (1): 211–212. https://doi.org/10.1002/pamm.201110097.

Freitag, S., W. Graf, and M. Kaliske. 2013. "A material description based on recurrent neural networks for fuzzy data and its application within the finite element method." *Comput. Struct.* 124 (Aug): 29–37. https://doi.org/10.1016/j.compstruc.2012.11.011.

Ghaboussi, J., J. Garrett, and X. Wu. 1991. "Knowledge-based modeling of material behavior with neural networks." *J. Eng. Mech.* 117 (1): 132–153. https://doi.org/10.1061/(ASCE)0733-9399(1991)117:1(132).

Ghaboussi, J., D. A. Pecknold, M. Zhang, and R. M. Haj-Ali. 1998. "Autoprogressive training of neural network constitutive models." *Int. J. Numer. Methods Eng.* 42 (1): 105–126. https://doi.org/10.1002/(SICI)1097-0207(19980515)42:1<105::AID-NME356>3.0.CO;2-V.

Ghaboussi, J., and D. Sidarta. 1998. "New nested adaptive neural networks (NANN) for constitutive modeling." *Comput. Geotech.* 22 (1): 29–52. https://doi.org/10.1016/S0266-352X(97)00034-7.

Godil, S. S., M. S. Shamim, S. A. Enam, and U. Qidwai. 2011. "Fuzzy logic: A "simple" solution for complexities in neurosciences?" *Surg. Neurol. Int.* 2. https://doi.org/10.4103/2152-7806.77177.

Graf, W., S. Freitag, J. U. Sickert, and M. Kaliske. 2012. "Structural analysis with fuzzy data and neural network based material description." *Comput.-Aided Civ. Infrastruct. Eng.* 27 (9): 640–654. https://doi.org/10.1111/j.1467-8667.2012.00779.x.

Hashash, Y. M. A., S. Jung, and J. Ghaboussi. 2004. "Numerical implementation of a neural network based material model in finite element analysis." *Int. J. Numer. Methods Eng.* 59 (7): 989–1005. https://doi.org/10.1002/nme.905.

Hoang, N.-D., A.-D. Pham, Q.-L. Nguyen, and Q.-N. Pham. 2016. "Estimating compressive strength of high performance concrete with Gaussian process regression model." *Adv. Civ. Eng.* 2016 (Jan): 1–8. https://doi.org/10.1155/2016/2861380.

Ibañez, R., E. Abisset-Chavanne, J. V. Aguado, D. Gonzalez, E. Cueto, and F. Chinesta. 2018. "A manifold learning approach to data-driven computational elasticity and inelasticity." *Arch. Comput. Methods Eng.* 25 (1): 47–57. https://doi.org/10.1007/s11831-016-9197-9.

Javadi, A. A., and M. Rezania. 2009a. "Applications of artificial intelligence and data mining techniques in soil modeling." *Geomech. Eng.* 1 (1): 53–74. https://doi.org/10.12989/gae.2009.1.1.053.

Javadi, A. A., and M. Rezania. 2009b. "Intelligent finite element method: An evolutionary approach to constitutive modeling." *Adv. Eng. Inf.* 23 (4): 442–451. https://doi.org/10.1016/j.aei.2009.06.008.

Javadi, A. A., T. P. Tan, and A. Elkassas. 2009. "Intelligent finite element method and application to simulation of behavior of soils under cyclic loading." In *Foundations of computational intelligence*, 317–338. Berlin: Springer.

Jung, S., and J. Ghaboussi. 2006. "Neural network constitutive model for rate-dependent materials." *Comput. Struct.* 84 (15–16): 955–963. https://doi.org/10.1016/j.compstruc.2006.02.015.

Kessler, B. S., A. S. El-Gizawy, and D. E. Smith. 2007. "Incorporating neural network material models within finite element analysis for rheological behavior prediction." *J. Pressure Vessel Tech.* 129 (1): 58–65. https://doi.org/10.1115/1.2389004.

Kirchdoerfer, T., and M. Ortiz. 2016. "Data-driven computational mechanics." *Comput. Methods Appl. Mech. Eng.* 304 (Jun): 81–101. https://doi.org/10.1016/j.cma.2016.02.001.

Kirchdoerfer, T., and M. Ortiz. 2017. "Data driven computing with noisy material data sets." *Comput. Methods Appl. Mech. Eng.* 326 (Nov): 622–641. https://doi.org/10.1016/j.cma.2017.07.039.

Lázaro-Gredilla, M., J. Quiñonero-Candela, C. E. Rasmussen, and A. R. Figueiras-Vidal. 2010. "Sparse spectrum Gaussian process regression." *J. Mach. Learn. Res.* 11 (Aug): 1865–1881.

Liu, H., J. Cai, and Y.-S. Ong. 2018. "Remarks on multi-output Gaussian process regression." *Knowl.-Based Syst.* 144 (Mar): 102–121. https://doi.org/10.1016/j.knosys.2017.12.034.

Lu, C., Y.-W. Feng, C.-W. Fei, and S.-Q. Bu. 2019. "Improved decomposed-coordinated kriging modeling strategy for dynamic probabilistic analysis of multicomponent structures." *IEEE Trans. Reliab.* 69 (2): 440–457. https://doi.org/10.1109/TR.2019.2954379.

Nie, J., and B. R. Ellingwood. 2005. "Finite element-based structural reliability assessment using efficient directional simulation." *J. Eng. Mech.* 131 (3): 259–267. https://doi.org/10.1061/(ASCE)0733-9399(2005)131:3(259).

Papadopoulos, V., and D. G. Giovanis. 2017. *Stochastic finite element methods*. Berlin: Springer.

Penumadu, D., and R. Zhao. 1999. "Triaxial compression behavior of sand and gravel using artificial neural networks (ANN)." *Comput. Geotech.* 24 (3): 207–230. https://doi.org/10.1016/S0266-352X(99)00002-6.

Ramberg, W., and W. R. Osgood. 1943. *Description of stress-strain curves by three parameters*. Washington, DC: National Advisory Committee for Aeronautics.

Ranganathan, A., M.-H. Yang, and J. Ho. 2010. "Online sparse Gaussian process regression and its applications." *IEEE Trans. Image Process.* 20 (2): 391–404. https://doi.org/10.1109/TIP.2010.2066984.

Rasmussen, C. E. 2003. "Gaussian processes in machine learning." In *Proc., Summer School on Machine Learning*, 63–71. Berlin: Springer.

Schuëller, G. I., and H. J. Pradlwarter. 2006. "Computational stochastic structural analysis (COSSAN)—A software tool." *Struct. Saf.* 28 (1–2): 68–82. https://doi.org/10.1016/j.strusafe.2005.03.005.

Settgast, C., M. Abendroth, and M. Kuna. 2019. "Constitutive modeling of plastic deformation behavior of open-cell foam structures using neural networks." *Mech. Mater.* 131 (Apr): 1–10. https://doi.org/10.1016/j.mechmat.2019.01.015.

Shin, H., and G. Pande. 2002. "Enhancement of data for training neural network based constitutive models for geomaterials." In *Proc., 8th Int. Symp. on Numerical Models in Geomechanics-NUMOG VIII*, 141–146. London: CRC Press. https://doi.org/10.1201/9781439833797.

Sidarta, J. G. D. E. 1998. "New nested adaptive neural networks (NANN) for constitutive modeling." *Comput. Geotech.* 22 (1): 29–52. https://doi.org/10.1016/S0266-352X(97)00034-7.

Stoffel, M., F. Bamer, and B. Markert. 2018. "Artificial neural networks and intelligent finite elements in non-linear structural mechanics." *Thin-Walled Struct.* 131 (Oct): 102–106. https://doi.org/10.1016/j.tws.2018.06.035.

Su, G., L. Peng, and L. Hu. 2017. "A Gaussian process-based dynamic surrogate model for complex engineering structural reliability analysis." *Struct. Saf.* 68 (Sep): 97–109. https://doi.org/10.1016/j.strusafe.2017.06.003.

Versino, D., A. Tonda, and C. A. Bronkhorst. 2017. "Data driven modeling of plastic deformation." *Comput. Methods Appl. Mech. Eng.* 318 (May): 981–1004. https://doi.org/10.1016/j.cma.2017.02.016.

Williams, C. K., and C. E. Rasmussen. 1996. "Gaussian processes for regression." In *Proc., Advances in Neural Information Processing Systems*, 514–520. Cambridge, MA: MIT Press.

Williams, C. K., and C. E. Rasmussen. 2006. *Gaussian processes for machine learning*. Cambridge, MA: MIT Press.

Yang, Y., B. Chen, Y. Su, Q. Chen, Z. Li, W. Guo, and H. Wang. 2020. "Concrete mix design for completely recycled fine aggregate by modified packing density method." *Materials* 13 (16): 3535. https://doi.org/10.3390/ma13163535.

Zhao, W., J. K. Liu, and Y. Y. Chen. 2015. "Material behavior modeling with multi-output support vector regression." *Appl. Math. Modell.* 39 (17): 5216–5229. https://doi.org/10.1016/j.apm.2015.03.036.