



OPEN

Heteroscedastic sparse Gaussian process regression-based stochastic material model for plastic structural analysis

Baixi Chen, Luming Shen[✉] & Hao Zhang

Describing the material flow stress and the associated uncertainty is essential for the plastic stochastic structural analysis. In this context, a data-driven approach-heteroscedastic sparse Gaussian process regression (HSGPR) with enhanced efficiency is introduced to model the material flow stress. Different from other machine learning approaches, e.g. artificial neural network (ANN), which only estimate the deterministic flow stress, the HSGPR model can capture the flow stress and its uncertainty simultaneously from the dataset. For validating the proposed model, the experimental data of the Al 6061 alloy is used here. Without setting a priori assumption on the mathematical expression, the proposed HSGPR-based flow stress model can produce a better prediction of the experimental stress data than the ANN model, the conventional GPR model, and Johnson Cook model at elevated temperatures. After the HSGPR-based flow stress model is implemented into finite element analysis, two numerical examples with synthetic material properties are performed to demonstrate the model's capability in stochastic plastic structural analysis. The results have shown that with sufficient data, the distribution of the structural load carrying capacity at elevated temperatures and the variation of load-displacement curves during the loading and unloading processes can be accurately predicted by the HSGPR-based flow stress model.

The flow stress behavior takes an important role in plastic structural analysis. Many constitutive models have been proposed to describe the flow stress of material and they can mainly be divided into two types, i.e., phenomenological constitutive models and physics-based constitutive models. Phenomenological constitutive models include the Johnson Cook model^{1,2}, Hockett-Sherby model³ and Arrhenius model⁴, etc. For physical-based constitutive modelling, there are Zerilli–Armstrong model⁵ and Rusinek–Klepaczko model⁶ etc. The above-mentioned conventional flow stress models could achieve good description in many situations, but the applications are limited by their mathematical expression. When there is a new alloy or a complex non-linear flow behavior, the original expression will become unreliable, and a new mathematical expression often needs to be developed. For example, Samantaray et.al⁷ modified the expression of the Zerilli–Armstrong model to make it suitable for the modified 9Cr-1Mo steel. However, the development of a new mathematical expression is non-trivial and can be time-consuming⁸.

To overcome the abovementioned limitations of the conventional constitutive approaches, machine learning has been introduced to model the flow stress behavior by many researchers^{8–14}. Quan et al.⁸ used one of the popular machine learning algorithms, the artificial neural network (ANN), to evaluate the flow stress of Al7050 aluminum alloy and the obtained ANN-based flow stress model showed higher accuracy than the Arrhenius model. Besides the ANN model, the support vector regression was also used by Tang et al. to predict the flow stress of Al2519 aluminum alloy¹⁵. The high agreement between the experimental data and the machine learning-based flow stress model was observed in the above-mentioned studies. Unlike the phenomenological constitutive models and physical-based constitutive models, machine learning-based flow stress models do not require any pre-postulated mathematical model and can predict the flow behavior more efficiently and more conveniently, as long as sufficient amount of experimental data are available⁸.

Although the machine learning based flow stress models showed good performance in the previous studies, most of the existing machine learning based models are deterministic and do not consider the material uncertainty. Due to the manufacture error of the material and the variation of the chemical compositions, even

School of Civil Engineering, The University of Sydney, Sydney, NSW 2006, Australia. [✉]email: luming.shen@sydney.edu.au

the same material will show different behavior under the same conditions. For example, Aakash et.al¹⁶ reported that the flow stress-plastic strain curve of Al6061 aluminum alloys had large variation even under the same temperature. Without considering the material uncertainty, the reliability of the constitutive model cannot be guaranteed, and the simulation given by this constitutive model can be unreliable.

By keeping the advantage of machine learning and taking the material uncertainty into account, a Bayesian based machine learning algorithm, Gaussian process regression (GPR), has been adopted to model material behavior¹⁷. Different from other machine learning methods only providing the deterministic estimation, the GPR model can capture both the underlying relation and corresponding uncertainty of the data simultaneously via the Bayesian approach¹⁸. In recent years, GPR model has achieved wide applications in several areas, including structural reliability analysis¹⁹, material strength estimation²⁰, and deformation forecasting²¹. However, the time complexity of the conventional GPR model is the cube of the data points ($O(n^3)$). When more factors are included and the dataset become large, it will be hard to train and use the conventional GPR model. Hence, in order to improve the model efficiency, an enhanced GPR model, i.e. heteroscedastic sparse Gaussian process regression (HSGPR)²², is introduced to capture the flow stress behavior in the present work. Compared with the conventional GPR model, the HSGPR model uses the sparse technology to achieve higher efficiency and reduce the time complexity from $O(n^3)$ to $O(nmd + nm^2)$ ^{22,23}, where n is the number of samples, d is the dimension of the input and m is the number of the basis functions with $d < n$ and $m < n$. Besides, different from the conventional GPR model, which usually uses the constant uncertainty model, the HSGPR model encodes a more advanced input-dependent uncertainty model in its structure. The advanced uncertainty model enables the HSGPR model to produce more accurate uncertainty estimation than the conventional GPR model²².

In this paper, the HSGPR model which takes the temperature and plastic strain as input variable and flow stress as output variable is used to identify the stochastic temperature-dependent elasto-plastic behavior. Without loss of generality, the experimental data of the Al6061 aluminum alloy at elevated temperatures are used in the present work to verify the HSGPR-based flow stress model. The obtained HSGPR-based flow stress model is then compared with the Johnson Cook model, which was often used to describe the Al6061 aluminum alloy in the previous studies^{24–27}, ANN model and the conventional GPR model. For the purpose of completeness, the methodology of numerically implementing the HSGPR model-based flow stress model into finite element analysis and stochastic plastic structural analysis is introduced in this work as well. Finally, two numerical examples with the synthetic flow stress data are used to demonstrate the capability of the HSGPR-based flow stress model in stochastic plastic structural analysis.

HSGPR-based flow stress model

Basic theory of HSGPR. To achieve the sparse Gaussian process regression, the basis functions approach is used here^{22,23}. As shown in the equations below, the output y is regarded as a linear combination of m nonlinear basis functions $\phi(\mathbf{x})$ of the inputs \mathbf{x} :

$$y = \phi(\mathbf{x})\mathbf{w} + \epsilon \quad (1)$$

$$\phi(\mathbf{x}) = [\phi_1(\mathbf{x}), \dots, \phi_j(\mathbf{x}), \dots, \phi_m(\mathbf{x})] \quad (2)$$

$$\phi_j(\mathbf{x}) = \exp\left[-\frac{1}{2}\left(\mathbf{x} - \mathbf{p}_j\right)^T \boldsymbol{\Gamma}_j^T \boldsymbol{\Gamma}_j \left(\mathbf{x} - \mathbf{p}_j\right)\right] \quad (3)$$

where $\mathbf{x} \in \mathbb{R}^{1 \times d}$ is the input vector and d is the dimension of the input, $y \in \mathbb{R}^1$ is the output, $\phi_j(\mathbf{x})$ is the basis function. In this paper, the radial basis function kernel is chosen as the basis function for the HSGPR model and its expression is presented in Eq. (3). m is the number of the basis functions, \mathbf{w} is the weight of basis functions and has the prior distribution $\mathcal{N}(0, \mathbf{A})$ where \mathbf{A} is an $m \times m$ constant diagonal matrix and represents the covariance matrix of the prior distribution, $\mathbf{p}_j \in \mathbb{R}^{1 \times d}$ and $\boldsymbol{\Gamma}_j \in \mathbb{R}^{d \times d}$ are the parameters of the basis function and $\boldsymbol{\Gamma}_j$ is a diagonal matrix. ϵ is the uncertainty of the output and is assumed to follow the normal distribution $\mathcal{N}(0, \sqrt{\beta}^2)$ and β reflects the intensity of the uncertainty. To consider the input dependent uncertainty and enhancing the uncertainty estimation, the HSGPR model regards β as the linear combination of the nonlinear basis functions as given below.

$$\beta(\mathbf{x}) = \exp(\phi(\mathbf{x})\mathbf{u} + b) \quad (4)$$

where $\phi(\mathbf{x})$ is same as the basis functions in Eq. (2), \mathbf{u} is the weight of the basis functions and has the prior distribution $\mathcal{N}(0, N)$ where N is an $m \times m$ diagonal matrix, and b is the constant part of Eq. (4) and is used to consider the situation of constant uncertainty intensity. The exponential function used here is to make β positive.

In terms of training dataset $(\mathbf{X}^*, \mathbf{y}^*)$, $\mathbf{X}^* = \{\mathbf{x}_i^*\}_{i=1}^n \in \mathbb{R}^{n \times d}$ is the $n \times d$ matrix storing n training inputs vectors and $\mathbf{y}^* = \{y_i^*\}_{i=1}^n \in \mathbb{R}^{n \times 1}$ is the vector storing n training outputs. The log marginal likelihood $\ln p(\mathbf{y}^*)$ of the training data can be expressed as

$$\begin{aligned} \ln p(\mathbf{y}^*) = & -\frac{1}{2}(\Phi \bar{\mathbf{w}} - \mathbf{y}^*)^T \mathbf{B}(\Phi \bar{\mathbf{w}} - \mathbf{y}^*) \\ & + \frac{1}{2} \ln |\mathbf{B}| - \frac{n}{2} \ln 2\pi - \frac{1}{2} \bar{\mathbf{w}}^T \mathbf{A} \bar{\mathbf{w}} + \frac{1}{2} \ln |\mathbf{A}| \\ & - \frac{1}{2} \ln |\Phi^T \mathbf{B} \Phi + \mathbf{A}| - \frac{1}{2} \mathbf{u}^T \mathbf{N} \mathbf{u} + \frac{1}{2} \ln |\mathbf{N}| - \frac{m}{2} \ln 2\pi \end{aligned} \quad (5)$$

where $\Phi = \{\phi(\mathbf{x}_i^*)\}_{i=1}^n \in \mathbb{R}^{n \times m}$ is the $n \times m$ matrix storing the basis functions for each training inputs, \mathbf{B} is the $n \times n$ diagonal matrix with diagonal elements $B(i,i) = \beta(\mathbf{x}_i^*)$ and represents the uncertainty intensity at the training inputs X^* , and $\bar{\mathbf{w}} = (\Phi^T \mathbf{B} \Phi + \mathbf{A})^{-1} \Phi^T \mathbf{B} \mathbf{y}^*$.

The parameters mentioned above can be determined by the training process. The training process is conducted by maximizing the log marginal likelihood of the training data. The optimization is achieved by using the quasi-Newton approach. After training the model, the parameters and weights in the basis functions can be updated to the optimal values. The output y at the unseen input \mathbf{x} is normally distributed as $\mathcal{N}(\mu, \delta^2)$ with

$$\mu = \bar{\phi}(\mathbf{x}) \bar{\mathbf{w}} \quad (6)$$

$$\delta^2 = v + \beta \quad (7)$$

$$v = \bar{\phi}(\mathbf{x}) \Sigma^{-1} \bar{\phi}(\mathbf{x})^T \quad (8)$$

$$\beta = \exp(\bar{\phi}(\mathbf{x}) \bar{\mathbf{u}} + \bar{b}) \quad (9)$$

where $\bar{\phi}(\mathbf{x})$ is the updated basis functions; $\bar{\mathbf{w}}, \bar{\mathbf{u}}$ and \bar{b} are the updated parameters and $\Sigma = \Phi^T \mathbf{B} \Phi + \mathbf{A}$ is a $m \times m$ matrix. μ and δ are the expectation and standard deviation of the prediction of the unseen input \mathbf{x} .

HSGPR-based flow stress model. When the HSGPR model is used to model the material flow stress σ_p at elevated temperatures, the plastic strain ε_p and the temperature T (K) are used as the input variables and the flow stress σ_p (MPa) is used as the output variable. The number of basis functions can be determined by trial and error. A larger number of basis functions can give higher accuracy but will increase the model complexity. Based on Eq. (6) to Eq. (9) and the Gaussian prior distribution assumed before, the obtained HSGPR-based stochastic flow stress model $\sigma_p(\varepsilon_p, T)$ follows the Gaussian distribution with the mean of $\bar{\phi}(\varepsilon_p, T) \bar{\mathbf{w}}$ and the variance of $\delta^2(\varepsilon_p, T)$, as expressed below

$$\sigma_p(\varepsilon_p, T) = \bar{\phi}(\varepsilon_p, T) \bar{\mathbf{w}} + \delta(\varepsilon_p, T) \cdot R \quad (10)$$

$$\delta^2(\varepsilon_p, T) = \bar{\phi}(\varepsilon_p, T) \Sigma \bar{\phi}(\varepsilon_p, T)^T + \exp(\bar{\phi}(\varepsilon_p, T) \bar{\mathbf{u}} + \bar{b}) \quad (11)$$

where R is a standard normal distribution of $\mathcal{N}(0, 1^2)$. The possible curve of the HSGPR based stochastic flow stress model can be sampled by sampling standard normal random variable R first and then substituting the sampled random variable into Eq. (10). The procedure of the HSGPR-based flow stress modelling is illustrated as a flow chart in Fig. 1.

Model evaluation. To evaluate the accuracy of the obtained flow stress models, the average absolute relative error (AARE) is used here. The model with lower AARE has higher accuracy.

The expression of AARE is presented below:

$$AARE = \frac{1}{N} \sum_{i=1}^N \frac{|\sigma_{pi}^{exp} - \sigma_{pi}^{pre}|}{\sigma_{pi}^{exp}} \quad (12)$$

where σ_{pi}^{exp} and σ_{pi}^{pre} are the i th experimental flow stress and the i th predicted mean flow stress of the model respectively, and N is the number of the data.

To evaluate the uncertainty of the flow stress predicted by the models, the mean log likelihood (MLL) of the experimental data is introduced here. The MLL reflects the average likelihood of obtaining the experimental data under the given model. The model with a higher MLL means that it is more likely to reproduce the data under the given model and the material uncertainty predicted by this model has higher accuracy.

The expression of MLL is given below:

$$MLL = \frac{1}{N} \sum_{i=1}^N \left(-\frac{(\sigma_{pi}^{exp} - \sigma_{pi}^{pre})^2}{2\delta_i^2} - \frac{1}{2} \ln(\delta_i^2) \right) - \frac{1}{2} \ln(2\pi) \quad (13)$$

where δ_i is the standard deviation of the i th flow stress predicted by the model.

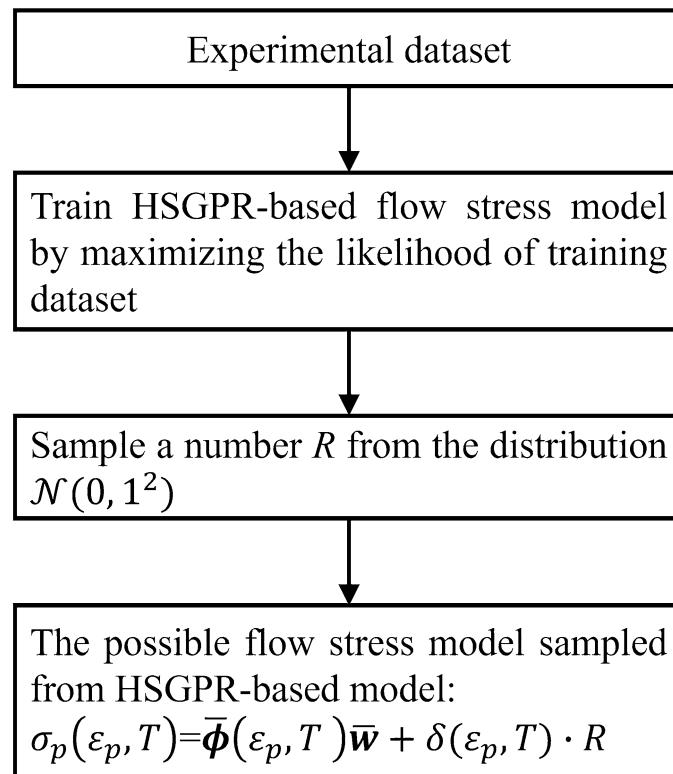


Figure 1. The flow chart of establishing HSGPR-based stochastic flow stress model.

	Temperature (°C)					
	20	100	150	200	250	300
Number of curves	19	18	12	18	12	19
Number of data points	4513	5194	3910	4168	2488	3738

Table 1. Number of the stress-strain curves for each lot and each temperature.

Application for material modelling

Material data. Without loss of generality, the experimental flow stress-plastic strain data of Al6061 aluminum alloy from the literature²⁸ are applied in the present work. There were six temperatures (20, 100, 150, 200, 250, 300 °C) considered in the experiments. The experiment data obtained by the uniaxial tension test conducted on a standard “dog bone” type coupon by MTS 810 (10-kip capacity). Design of the uniaxial tension test followed the guidelines in ASTM E8. A total of 98 stress-strain curves (totally 24,011 data points) are used here. The detailed breakdown of the obtained curves is illustrated in Table 1 and all data curves are presented in Fig. 2.

Kolmogorov-Smirnov hypothesis test is conducted on the flow stress data at different temperature and different plastic strain to examine if the flow stress follows the Gaussian distribution. The p-values of the hypothesis test on the flow stress are presented in Table 2. As demonstrated in Table 2, the test does not reject the hypothesis, indicating that the flow stress obeys the Gaussian distribution, at the 5% significance level. Hence, it is reasonable to use the normal distribution for the uncertainty \in in the HSGPR model and Eq. (10) to describe the stochastic behavior of the flow stress.

In addition, the experimental data at temperature of 250 °C are used as the testing dataset for the HSGPR model, while the rest of data (21,523 data points) are used as the training dataset for the ANN model and HSGPR model and conventional GPR model.

Results and discussion

Four models, namely, the Johnson Cook (JC) model²⁹, the ANN model, the HSGPR model, and the conventional GPR model are used to model the flow stress of the Al6061 aluminum alloys at elevated temperatures.

The Johnson-Cook flow stress model was used by many researchers to model the flow behavior of Al6061 aluminum^{24–27}. The formulation of JC model is expressed in Eq. (14).

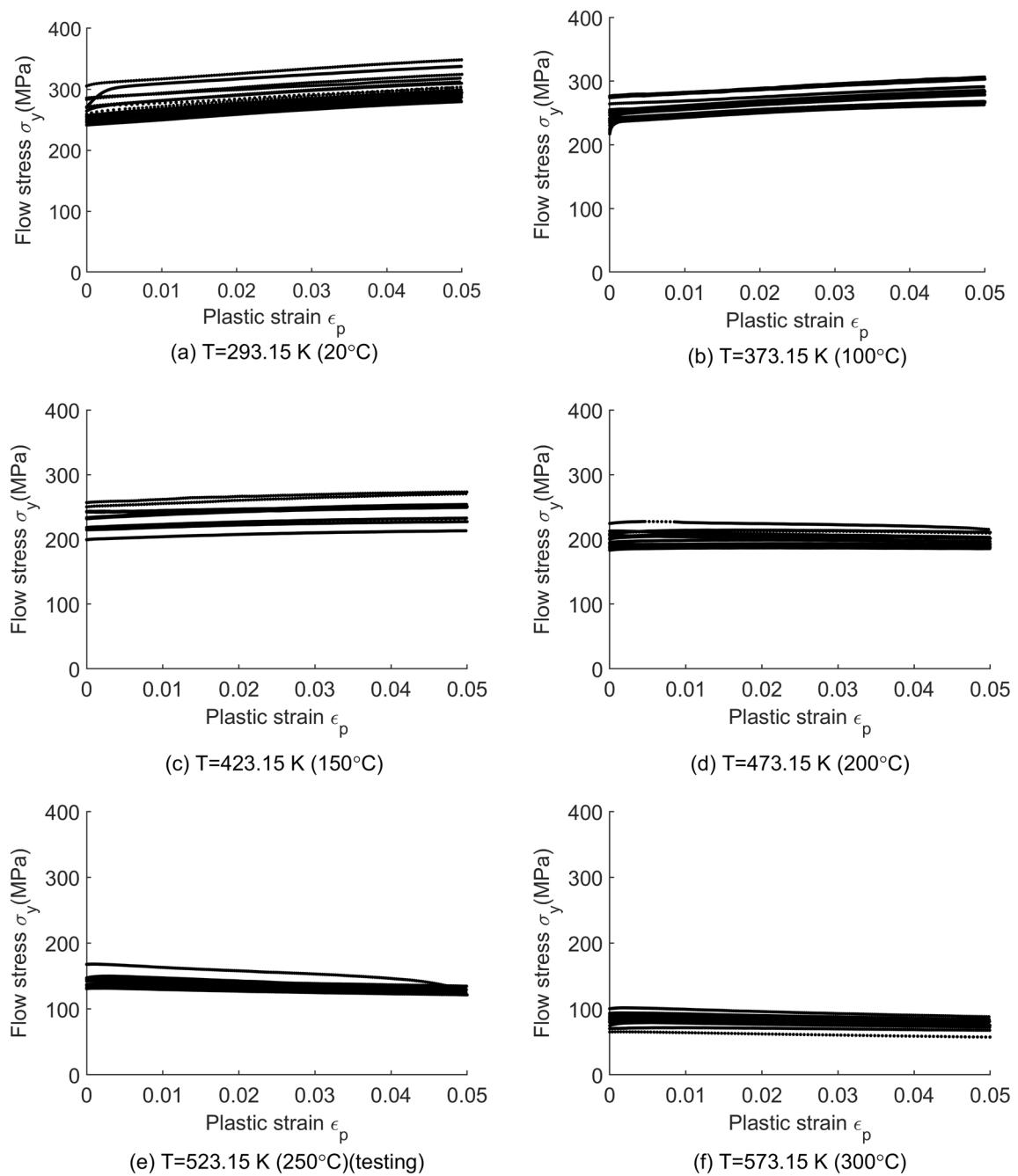


Figure 2. The experimental flow stress-plastic strain data of Al6061 aluminum alloys at the temperature of (a) 20 °C, (b) 100 °C, (c) 150 °C, (d) 200 °C, (e) 250 °C and (f) 300°C²⁸.

P-value	Temperature (°C)						
	20	100	150	200	250	300	
Plastic strain ϵ_p	0.00	0.61	0.92	0.84	0.57	0.99	0.79
	0.01	0.51	0.52	0.95	0.78	0.94	0.85
	0.02	0.65	0.76	0.93	0.99	0.98	0.92
	0.03	0.62	0.89	0.94	0.98	0.98	0.95
	0.04	0.58	0.98	0.60	0.96	0.99	0.79
	0.05	0.54	0.98	0.52	0.62	0.96	0.71

Table 2. The P-value of the Kolmogorov-Smirnov hypothesis test on the flow stress data at various temperatures and plastic strains.

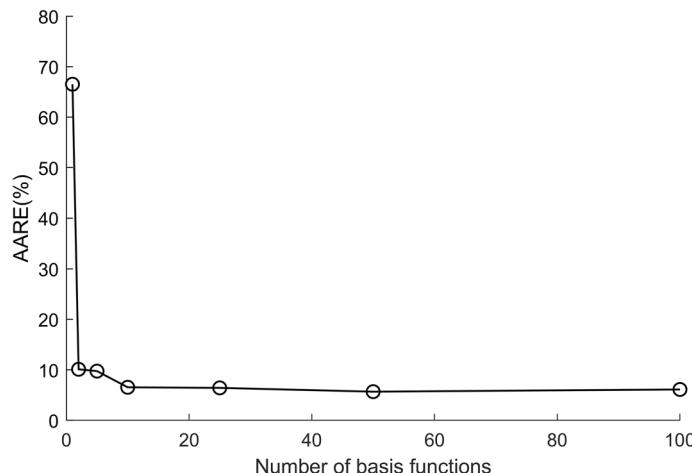


Figure 3. The relationship between the AARE on the testing dataset and the number of basis functions m .

$$\sigma_p = \left[c_1 + c_2 \varepsilon_p^{c_3} \right] \left[1 - \left(\frac{T - T_r}{T_m - T_r} \right)^{c_4} \right] \quad (14)$$

where σ_p (MPa) is the flow stress, ε_p is the plastic strain, T (K) is the temperature of material, c_1 , c_2 , c_3 and c_4 are the model parameters, T_r is the room temperature of 293.15 K (i.e. 20 °C) and T_m is the melt temperature of 925 K for the Al6061 aluminum alloy.

The parameters c_1 , c_2 , c_3 of JC model are calibrated by the stress-strain curves at room temperature (i.e. 20 °C) first. The Pearson correlation coefficients between c_1 and c_2 , and between c_1 and c_3 are -0.002 and -0.021 respectively, which means that the correlation is negligible and c_1 is independent of c_2 and c_3 ³⁰. As the Pearson correlation coefficients between c_2 and c_3 is 0.96, c_2 and c_3 are assumed to be linearly correlated with each other here. By conducting Kolmogorov-Smirnov test, the parameters c_1 , c_2 and c_3 follow normal distribution $\mathcal{N}(260.72, 17.32^2)$ MPa, $\mathcal{N}(504.26, 106.99^2)$ MPa and $\mathcal{N}(0.82, 0.14^2)$ respectively at the 5% significance level. The parameter c_4 is calibrated by the data of zero plastic strain. c_4 obeys the normal distribution $\mathcal{N}(1.277, 0.069^2)$ at the 5% significance level.

The HSGPR-based flow stress model for the Al6061 aluminum alloys is established by the approach described in “Section HSGPR-based flow stress model”. In order to determine the suitable number of basis functions m , the relationship between the AARE on the testing dataset and m is drawn in Fig. 3. It can be observed from Fig. 3 that the improvement of the model accuracy is negligible when m is larger than 10. Hence, 10 basis functions are sufficient to produce a good estimate and more basis functions will only increase the model complexity without obvious contribution to the model accuracy. For the sake of the model efficiency and accuracy, $m=10$ is adopted in the rest of the paper. With $m=10$, the optimal parameters of the HSGPR model can be obtained by the training process and presented in “Appendix A” of Supplementary Information.

The conventional GPR model and the ANN model are established by using the MATLAB toolbox as well for comparison. For the GPR model, the squared exponential kernel function is used as the kernel function, and the subset of the data points approximation is set as the fitting method to address the efficiency issue. The quasi-Newton method is used to optimize the parameters in the GPR model and the default value in MATLAB is used for all other settings of the GPR model. The optimal parameters of the GPR model (i.e., the signal standard deviation, the characteristic length scales and the noise standard deviation) are 228.8175, 244.4820 and 17.8537 respectively. The ANN structure is adopted from literature⁹ directly. As demonstrated in the reference⁹, the network with one hidden layer consisting of 20 nodes is the suitable setting for estimating the flow stress.

Since the flow stress model needs to be called to predict the flow stress at each integration point and each increment step in the finite element analysis (FEA), the prediction time of the data-driven flow stress model will significantly influence the efficiency of the model in the FEA. The prediction time of the ANN model, HSGPR model and GPR model for 10,000 input points in the MATLAB platform are 0.019 s, 0.011 s and 26.78 s, respectively. Compared with the conventional GPR model, the HSGPR model and ANN model have lower time consumption during the prediction, which means they have higher efficiency in the FEA.

The AARE and MLL of the three machine learning models (namely, the HSGPR model, GPR model and ANN model) on the training and testing dataset are presented in Table 3. Since only deterministic prediction is given by the ANN model, only the AARE of the ANN model is calculated. The ANN model, HSGPR model and GPR model have similar AARE on the training stage, while the performance of ANN model on the testing dataset is the worst among the three models with the AARE of 12.4%. In addition, with the advanced uncertainty model, the HSGPR model presents obviously larger MLL and smaller AARE than the GPR model on the testing dataset. The larger MLL means the testing dataset is more likely to be reproduced by the HSGPR model than the GPR model.

In order to further evaluate the generalization ability of the three machine learning approaches, the cross-validation procedure is performed here as well. The data points at one selected temperature are excluded for validation, while the rest of the data are used for training. Hence, there are totally 6 cross-validation tests. The

Dataset	Model	AARE (%)	MLL
Training	HSGPR	5.2	- 4.04
	GPR	5.7	- 4.23
	ANN	5.2	-
Testing	HSGPR	6.5	- 3.74
	GPR	8.7	- 7.08
	ANN	12.4	-

Table 3. AARE and MLL of ANN model, HSGPR-based flow stress model and GPR-based flow stress model on the training and testing dataset.

Validation No	Excluded temperature for validation (°C)	AARE (%)		
		HSGPR	GPR	ANN
1	20	9.9	30.0	17.2
2	100	8.6	8.7	8.5
3	150	4.2	4.4	7.2
4	200	4.5	6.6	4.5
5	250	6.5	8.7	12.4
6	300	8.1	33.1	23.6

Table 4. AARE of ANN-based flow stress model, HSGPR-based flow stress model, GPR-based flow stress model on cross validation set.

Model	AARE (%)	MLL
HSGPR	5.3	- 4.01
GPR	6.0	- 4.21
ANN	6.0	-
JC	31.1	- 10.18

Table 5. AARE and MLL of the ANN model, HSGPR-based flow stress model, GPR-based flow stress model and JC model on the whole dataset.

AARE of the three models on the excluded validation set is demonstrated in Table 4. In the validation test 1 and 6, the validation temperature is not covered by the training dataset, so the model extrapolation ability can be tested. Notably, the HSGPR model have superior accuracy over the GPR model and ANN model on extrapolation. Besides, the performance of the HSGPR model on the interpolation validation set (No. 2–5) is also better than the other two models. It can be concluded that the HSGPR model has better generalization ability than the conventional GPR model and ANN model on flow stress prediction.

As a conventional constitutive model, the JC model is directly evaluated using the whole dataset (including all the experimental data). For comparison, the three machine learning approaches are also evaluated on the whole dataset and the results are listed on Table 5. Due to the limitation of the mathematical expression, the performance of the JC model is worse than the machine learning approach. The HSGPR model has the smallest AARE and the largest MLL among all the models, which means that the HSGPR model shows better performance than the other models.

As for the material uncertainty predicted by the models, it can be visualized by the 95% confidence region (i.e. mean $\pm 2 \times$ standard deviation) and the coefficient of variation (COV, i.e. standard deviation/mean) of the flow stress model and is plotted in Figs. 4, 5, 6, 7. Since the ANN cannot provide the material uncertainty, only the HSGPR-based flow stress model, GPR-based flow stress model and JC model are plotted and discussed in the next.

At the low temperatures (20 °C to 150 °C), it can be observed from Figs. 4, 5, 6, 7 that the 95% confidence region given by the three models could cover most of the data at low temperatures. Besides, as shown in Fig. 4, the COVs given by the three models are similar and are close to the COV given by the experimental data at the temperatures of 20 °C to 150 °C.

At the high temperatures (200 °C to 300 °C), the mean curves and material uncertainty estimated by the HSGPR model are consistent with the experimental data, while the material uncertainty is overestimated by the GPR model and the COV given by GPR model is much larger than the experimental value. It is notable to mention that the advanced uncertainty model within HSGPR model results in the better uncertainty estimation

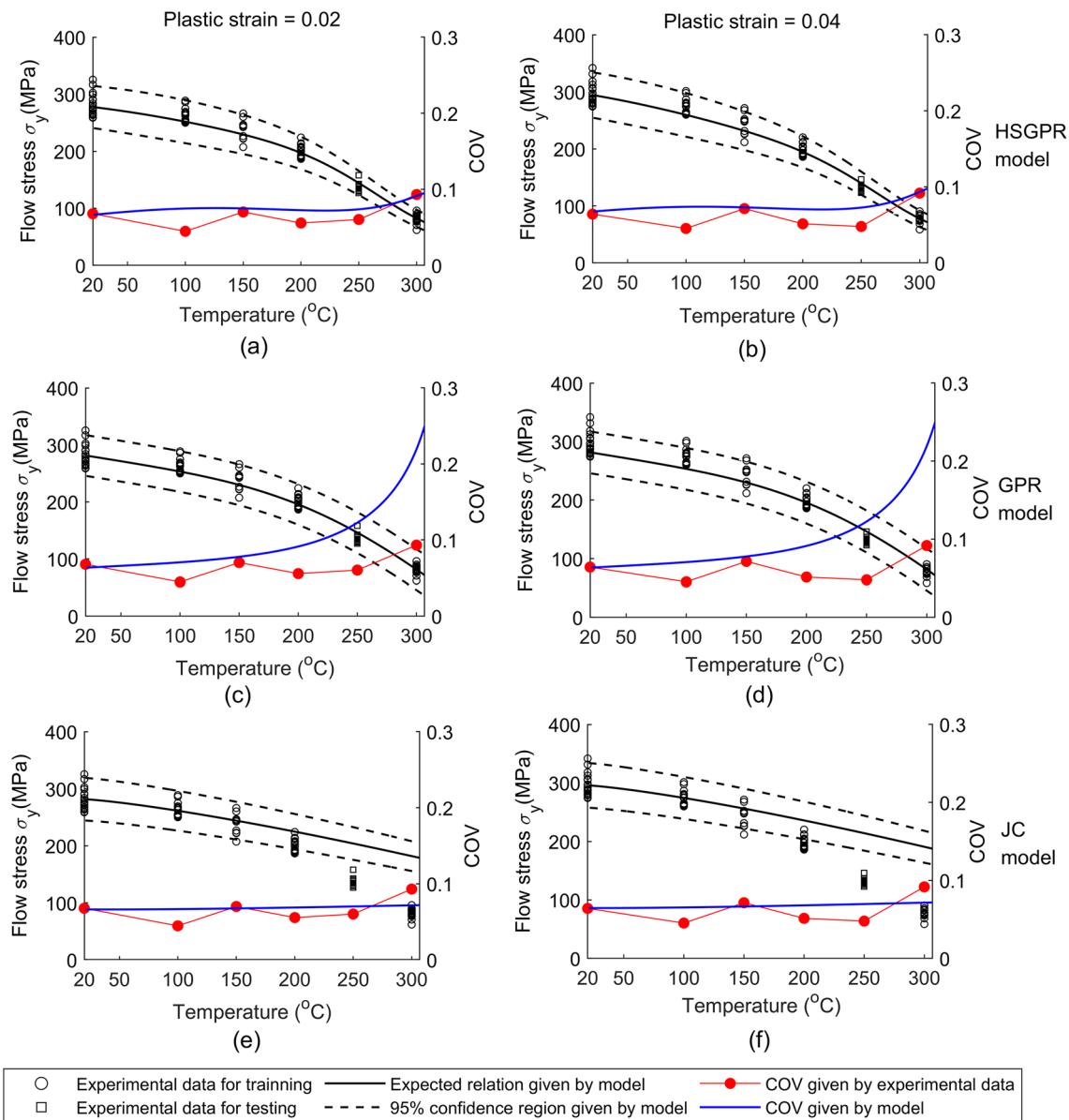


Figure 4. The obtained HSGPR-based flow stress model, GPR-based flow stress model and JC model with respect to the temperatures at the plastic strains of 0.02 and 0.04. (a) HSGPR model at the plastic strains of 0.02, (b) HSGPR model at the plastic strains of 0.04, (c) GPR model at the plastic strains of 0.02, (d) GPR model at the plastic strains of 0.04, (e) JC model at the plastic strains of 0.02, (f) JC model at the plastic strains of 0.04.

than the GPR model. In addition, as limited by its mathematic expression, the JC model cannot describe the severe temperature softening of Al6061 alloys at high temperatures. As shown in Fig. 7, the estimated flow stress given by the JC model deviates from the experiment data at high temperature, which causes the high AARE of the JC model in Table 5. The poor performance of the JC model for the Al6061 alloy at high temperatures is consistent with the results in the work of Fan et al.²⁵. Although it is possible to modify the JC model to improve the accuracy, those modifications are material-dependent and need to postulate the mathematical expression as well³¹. Establishing a new mathematic expression of the modified JC model for each new material is always a time-consuming work and sometimes a suitable modification is hard to find. Different from the JC model, the HSGPR-based flow stress model is a material-independent data-driven flow stress model. Without postulating the specific mathematical form, the HSGPR framework can be used to establish a flow stress model directly from the material data with high accuracy.

In addition, it should be admitted that the current HSGPR based model use the uncertainty model ϵ with the normal distribution to approximate the uncertainty of material data. If the data deviated from the normal distribution, the uncertainty ϵ with other distribution needs to be used. The new distribution of ϵ can be selected by trial and error, and the distribution, which can produce the maximum MLL, is regarded as the suitable one.

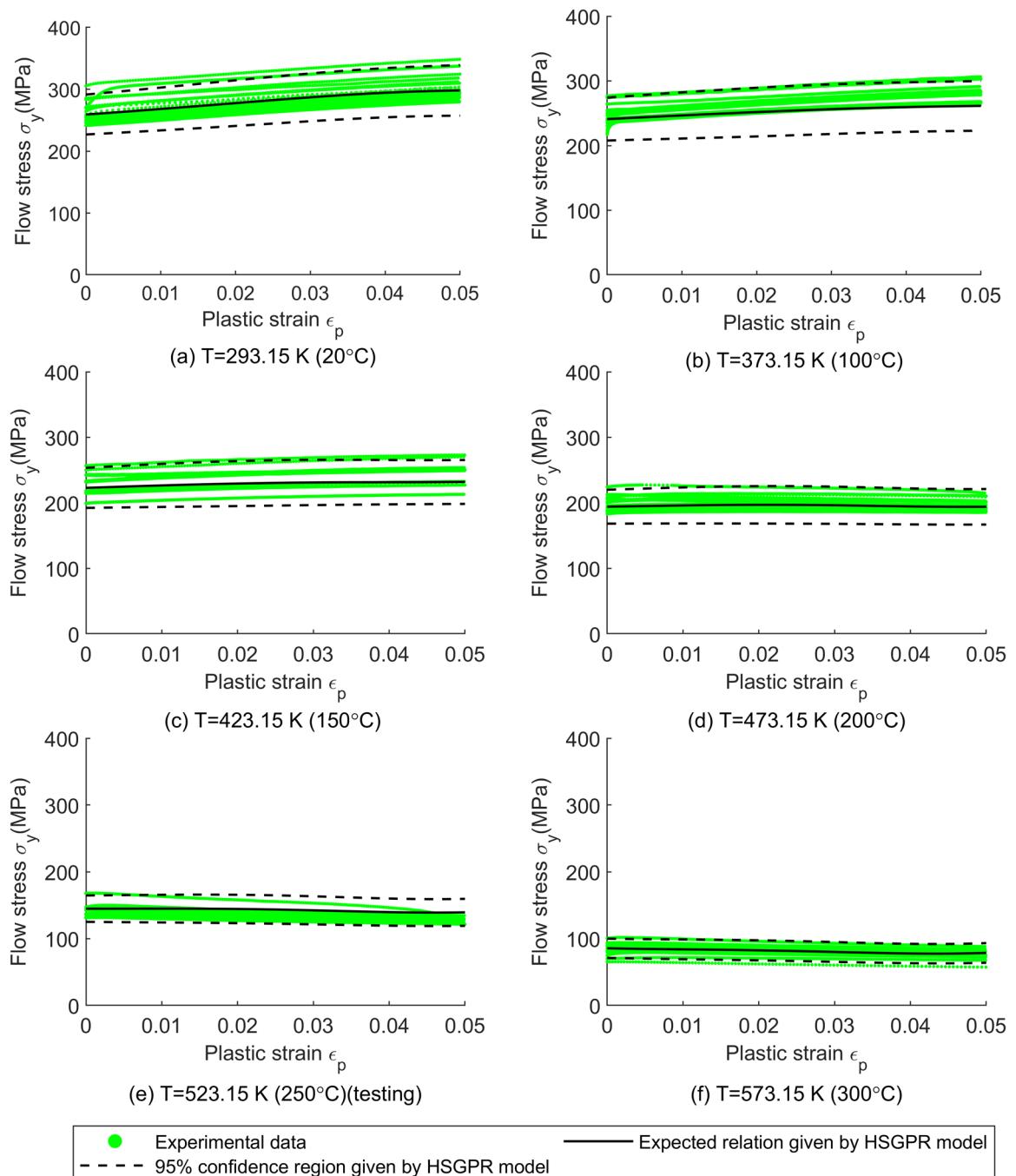


Figure 5. The flow stress obtained using the HSGPR-based model with respect to the plastic strain at the temperature of (a) 20°C , (b) 100°C , (c) 150°C , (d) 200°C , (e) 250°C and (f) 300°C .

Application in plastic stochastic structural analysis

Since the HSGPR-based stochastic flow stress model is introduced for the first time in the current work and has not been implemented in the finite element analysis (FEA) for plastic stochastic structural analysis previously, the numerical implementation of the HSGPR-based flow stress model into the FEA for plastic stochastic structural analysis is proposed in this section. And the model is then validated by two numerical examples with the synthetic material data.

Numerical implementation in FEA. The obtained HSGPR-based stochastic flow stress model can be combined with the von Mises yield criteria to describe the plastic behavior of the material in the FEA. The yield function with the HSGPR-based flow stress model can be expressed as

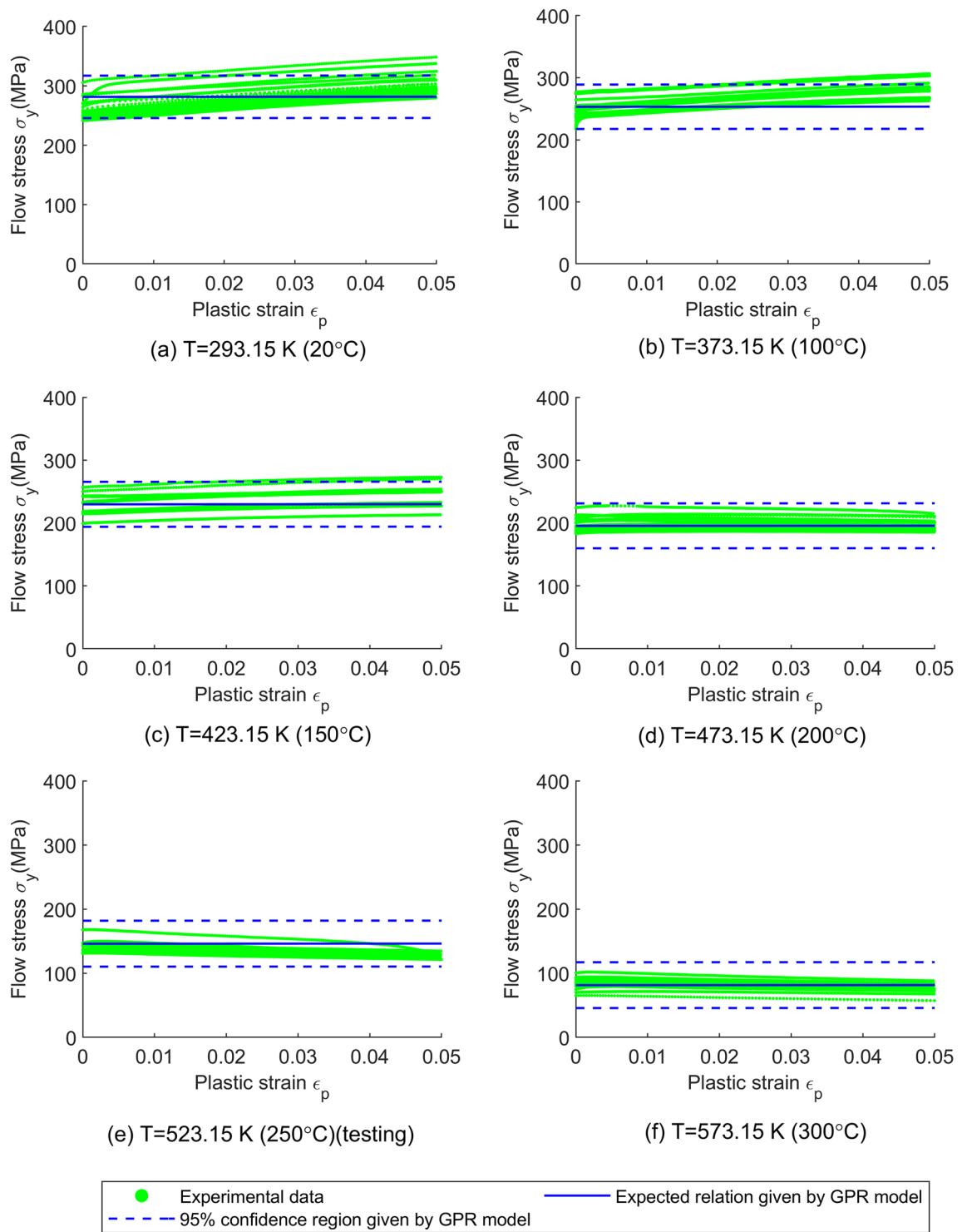


Figure 6. The flow stress obtained using GPR-based model with respect to the plastic strain at the temperature of (a) 20°C , (b) 100°C , (c) 150°C , (d) 200°C , (e) 250°C and (f) 300°C .

$$f(\bar{\sigma}, \bar{\epsilon}_p, T) = \bar{\sigma} - \sigma_p(\bar{\epsilon}_p, T) \quad (15)$$

$$\bar{\sigma} = \sqrt{\frac{(\sigma_{11} - \sigma_{22})^2 + (\sigma_{33} - \sigma_{22})^2 + (\sigma_{11} - \sigma_{33})^2 + 6(\sigma_{12}^2 + \sigma_{13}^2 + \sigma_{23}^2)}{2}} \quad (16)$$

where $\bar{\sigma}$ is the equivalent von Mises stress, σ_p is the HSGPR-based flow stress, $\bar{\epsilon}_p$ is the von Mises equivalent strain, T is temperature, and $[\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{13}, \sigma_{23}]$ is the Cauchy stress vector.

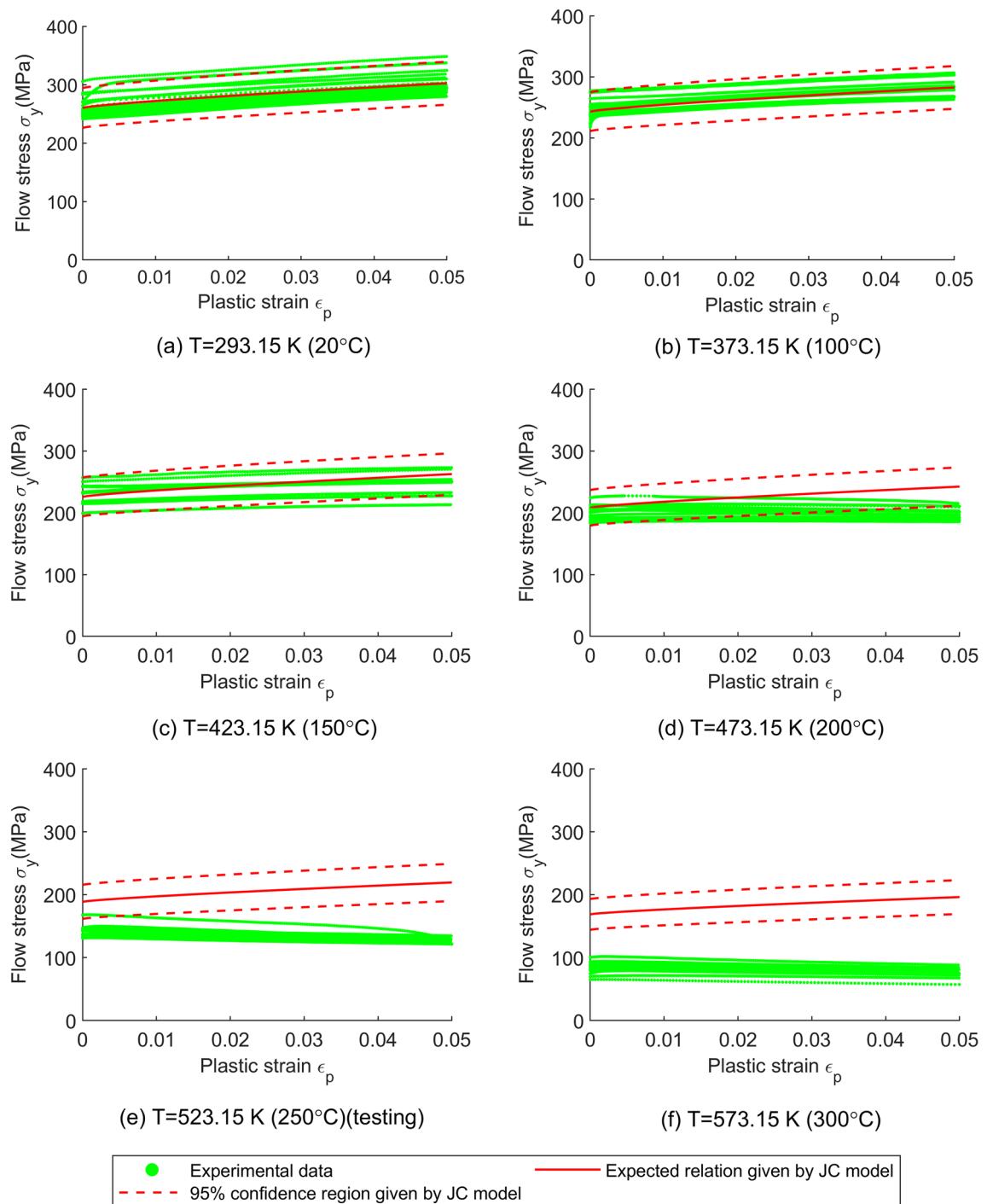


Figure 7. The flow stress obtained using the JC model with respect to the plastic strain at the temperature of (a) 20 °C, (b) 100 °C, (c) 150 °C, (d) 200 °C, (e) 250 °C and (f) 300 °C.

The yield function with the HSGPR-based flow stress model is called at each increment of the FEA to check the yield state of each element integration point and update the flow stresses of the element integration points which have reached plasticity. In order to update the flow stresses, the stiffness of the HSGPR-based flow stress model, that is the gradient of the flow stress model σ_p with respect to the equivalent plastic strain $\bar{\epsilon}_p$ and temperature T , respectively, is queried in the FEA³². The corresponding stiffness (i.e. $\partial \sigma_p / \partial \bar{\epsilon}_p$ and $\partial \sigma_p / \partial T$) of the HSGPR-based flow stress model is expressed as:

$$\frac{\partial \sigma_p}{\partial \bar{\epsilon}_p} = \frac{\partial \phi(\bar{\epsilon}_p, T)}{\partial \bar{\epsilon}_p} \bar{w} + \frac{\partial \delta(\bar{\epsilon}_p, T)}{\partial \bar{\epsilon}_p} \cdot \mathcal{N}(0, 1^2) \quad (17)$$

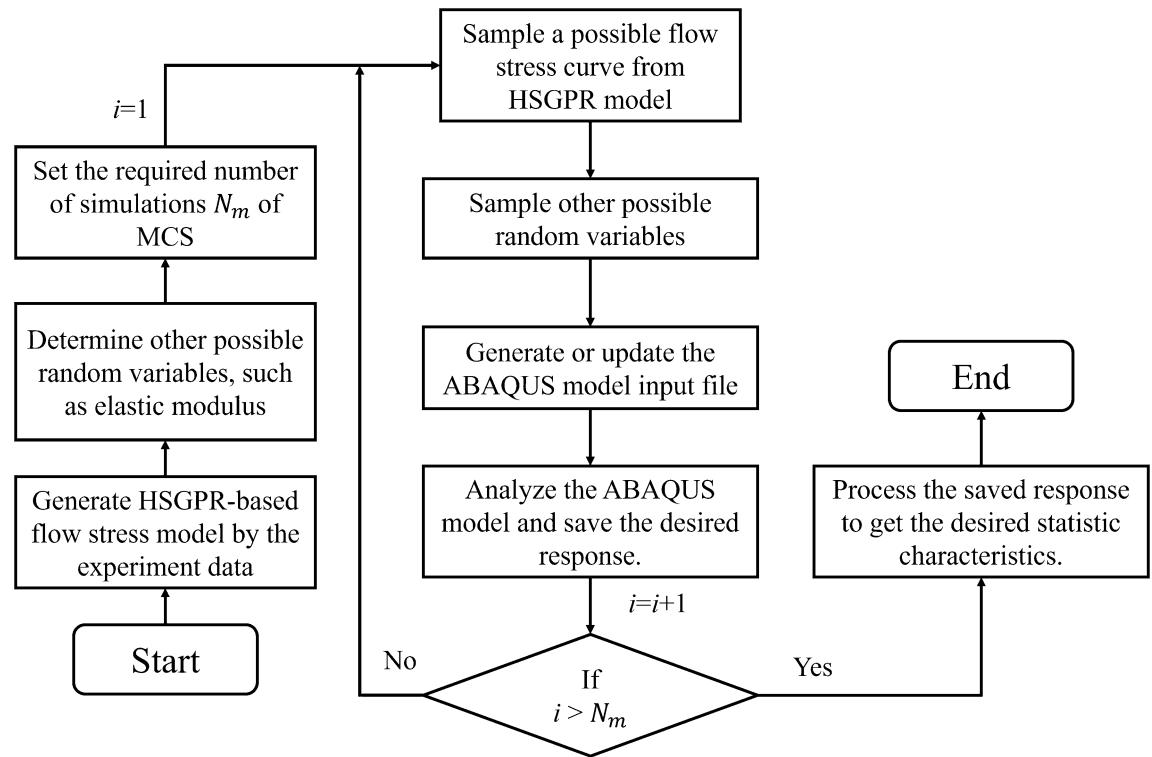


Figure 8. Flowchart of the stochastic plastic structural analysis steps with the HSGPR-based model.

$$\frac{\partial \sigma_p}{\partial T} = \frac{\partial \bar{\phi}(\bar{\varepsilon}_p, T)}{\partial T} \bar{w} + \frac{\partial \delta(\bar{\varepsilon}_p, T)}{\partial T} \cdot \mathcal{N}(0, 1^2) \quad (18)$$

where the detailed expressions and inferences of the gradients $\partial \bar{\phi}(\bar{\varepsilon}_p, T)/\partial \bar{\varepsilon}_p$, $\partial \bar{\phi}(\bar{\varepsilon}_p, T)/\partial T$, $\partial \delta(\bar{\varepsilon}_p, T)/\partial \bar{\varepsilon}_p$ and $\partial \delta(\bar{\varepsilon}_p, T)/\partial T$ are presented in “Appendix B” of the Supplementary Information.

This procedure can be achieved in the general-purpose finite element software Abaqus³³ via its subroutine UHARD. The numerically implemented HSGPR-based stochastic flow stress model can then be used to solve the temperature-dependent plastic problems.

Stochastic plastic structural analysis procedure. The Monte Carlo Simulation (MCS) method is applied to achieve the stochastic plastic structural analysis using the HSGPR-based model. The possible flow stress relations estimated by the HSGPR model are generated by following the flow chart in Fig. 1 and samples of other possible basic random variables, such as elastic modulus (E), can also be generated according to their probability distribution. For each sample, a finite element model is set up and a complete analysis is performed using Abaqus. The procedure of performing the plastic stochastic structural analysis with the HSGPR-based model is presented schematically in Fig. 8. During each iteration, the sampled flow stress and other possible basic random variables are used to update the corresponding value in the Abaqus input files. The number of the MCS, which can result in a converged estimation, is applied in the procedure for the stochastic structural analysis.

Numerical examples. *Beam problem.* As a data-driven material model, the HSGPR-based flow stress model should be compatible with both well-known and unknown (or new) constitutive laws. Here, a well-known flow stress model will be used to validate the developed HSGPR-based flow stress model in the stochastic plastic structural analysis. Instead of using the experimental data, the training data for this example are obtained from the Abaqus built-in Johnson Cook model (Eq. 14), which is the target for the HSGPR-based flow stress model to learn. T_r and T_m of the Johnson Cook model are the deterministic values of 293.15 K and 925 K respectively. The other parameters of the Johnson Cook model are set as the normal random variables to include the effect of the material uncertainty. The means of the parameters c_1 , c_2 , c_3 and c_4 of the JC model are 324 MPa, 114 MPa, 0.42 and 1.34, respectively³⁴. The coefficients of variation (COV) of four parameters are set as 0.05, 0.10 and 0.15 to consider three different uncertainty levels. Six temperatures (20, 100, 150, 200, 250, 300 °C) are considered in the training dataset. For each temperature, 20 stress-strain curves (24,000 data points), 100 stress-strain curves (120,000 data points), and 200 stress-strain curves (240,000 data points) are generated for studying the effect of the dataset size. Hence, there are totally nine synthetic flow stress datasets with three uncertainty levels and three dataset sizes generated for training the HSGPR model. The Poisson ratio of the material is set as 0.33 and the elastic modulus is set as a temperature-dependent normal random variable³⁵ and is expressed as

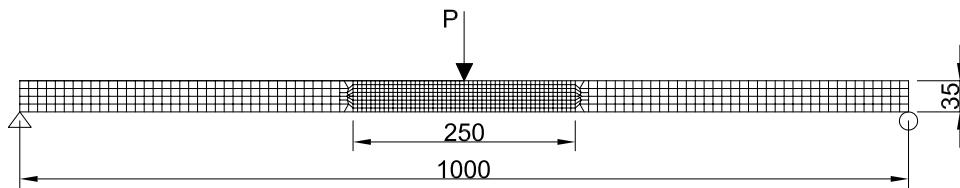


Figure 9. Boundary condition, loading condition, geometrical size (mm), and mesh distribution of the example model.

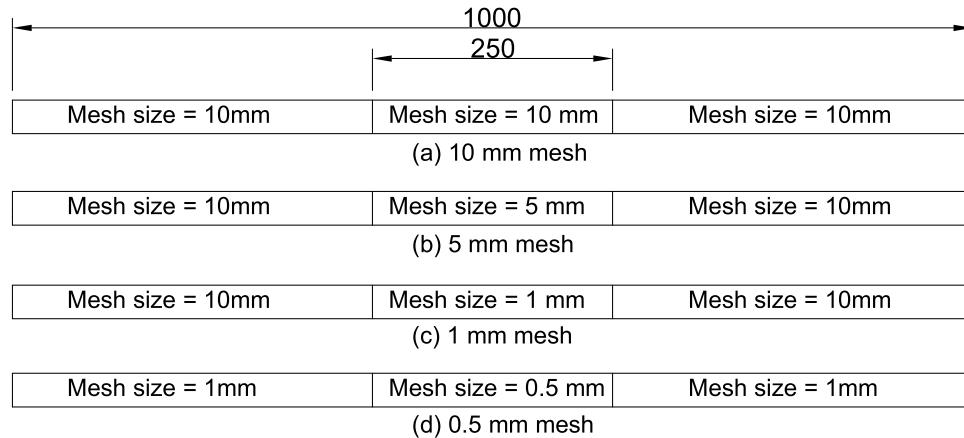


Figure 10. Four different mesh size strategies for the FEA model of the beam. (a) 10 mm mesh, (b) 5 mm mesh, (c) 1 mm mesh, (d) 0.5 mm mesh.

$$E_T = \left(1 - \frac{T - T_r}{1600} \right) E_0 \quad (19)$$

where T (K) is the temperature, T_r is the room temperature 293.15 K and the room temperature elastic modulus E_0 is a normal random variable with a mean of 69,500 MPa and a standard deviation of 8,548.5 MPa³⁵.

After the training process, the HSGPR-based flow stress model is implemented into the FEA and the model is validated by the beam problem adopted from³⁶. The geometry, boundary conditions and loading conditions of the beam are depicted in Fig. 9 and the thickness of the beam is 35 mm. The failure strain is set as 0.1¹⁶. Except for the elastic modulus and flow stress, other uncertainty sources are not taken into account in the current validation process. Abaqus is used to conduct the finite element analysis. The displacement-controlled loading is applied at the middle point of the beam model. The mesh convergence study is conducted with all the parameters being set to their mean values. The linear 4-node doubly curved shell element (S4R) is used in the Abaqus model. As depicted in Fig. 10, four different mesh size strategies are studied in the present work. The load–displacement curve of each case is presented in Fig. 11. The relative errors of the load capacity (i.e. peak load) of the cases with 10 mm, 5 mm and 1 mm mesh relative to the case with 0.5 mm mesh are 1.62%, 0.35% and 0.32% respectively. The case with 5 mm mesh, which could provide sufficient accuracy for estimating the load capacity and acceptable computational efficiency, is used in the current study. The mesh distribution of the beam is presented in Fig. 9.

In order to determine the number of MCS, the stochastic structural capacity of the beam is evaluated by the HSGPR model at the different number of iterations. The expectation and standard deviation of the structural capacity with respect to the number of MCS is presented in Fig. 12. It can be noted from the figure that both the expectation and standard deviation given by the MCS converge at the number of 1000. Hence, the number of MCS is set as 1000 for the structure analysis.

The probability density distributions of the structural load capacities given by the HSGPR-based flow stress model and the target model at the room temperature and the elevated temperatures are presented in Fig. 13. The corresponding mean and standard deviation of the load capacities are illustrated in Tables 6 and 7, respectively.

It can be observed from Table 6 that the expected capacity given by the HSGPR-based flow stress model have high accuracy with the relative error smaller than 4% even under the high uncertainty level and small dataset size. When the dataset size is large enough, the influence of increasing dataset has little effect on the accuracy of expected capacity. The slight error variation may be due to the model itself. For the variability of the load capacity, the probability density distribution obtained by the HSGPR model coincides with the distribution obtained by the target model at the low uncertainty level ($COV = 0.05$), as shown in Fig. 13. For the high uncertainty level, the HSGPR model requires larger dataset size to achieve good estimation of the COV of load capacity. With the sufficient material data, the relative error of the standard deviation given by the HSGPR model is smaller

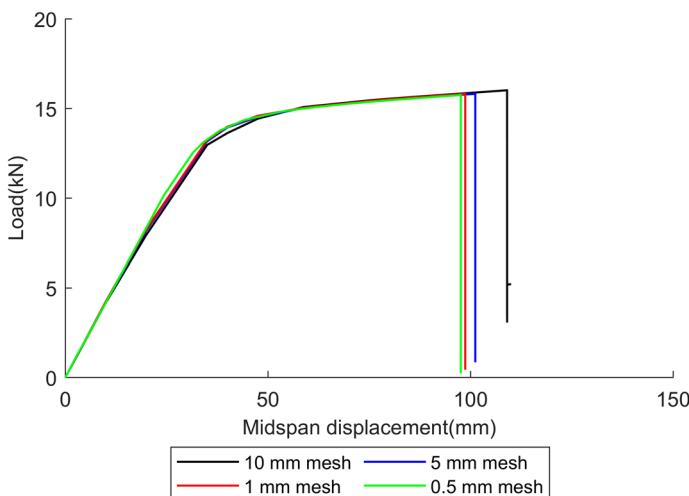


Figure 11. The load–displacement curves of four mesh strategies.

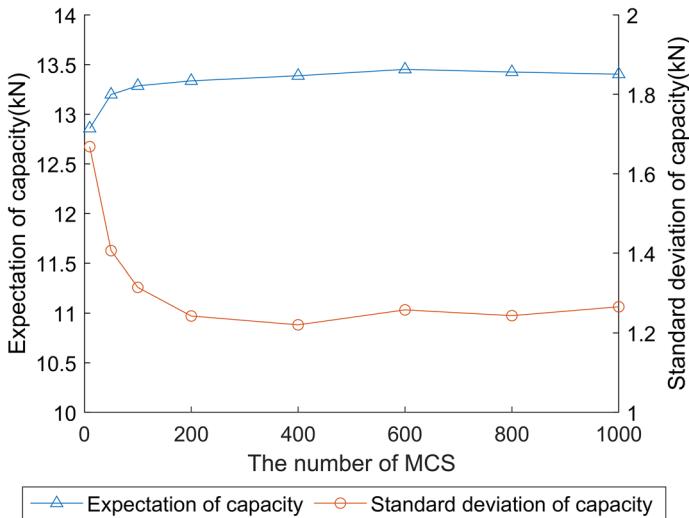


Figure 12. The expectation and standard deviation evaluated by the HSGPR model at the different number of MCS. (Temperature = 175 °C, COV = 0.10, the number of data curves = 100).

than 3%. As presented in Tables 6 and 7, the effect of temperature on the accuracy of the HSGPR model is not obvious at all uncertainty levels, which means that the HSGPR-based flow stress model is suitable for the stochastic plastic structural analysis at elevated temperatures with sufficient dataset. In addition, it is worth noting that the HSGPR model can give good estimation on the load capacity at the temperature of 175 °C which is not considered in the synthetic dataset.

Punch problem. To evaluate the performance of the HSGPR-based model in the loading and unloading process, the punch problem given by Huang et al.³⁷ is used as the benchmark test. The exponential flow stress law $\sigma_p = e_1 + e_1(e_2 + \varepsilon_p)^{e_3}$ is set as the target model. To consider the material uncertainty, the parameters e_1 , e_2 and e_3 are set as the normal random variables with the mean of 0.05 MPa, 0.00002 and 0.3 respectively and the COV of 0.05. A total of 200 stress–strain curves (40,000 data points) at the room temperature are generated for training the HSGPR model. The Poisson ratio is set as 0.33 and the elastic modulus is set as the normal random variable with the mean of 1 MPa and the standard deviation of 0.05 MPa.

The HSGPR-based flow stress model obtained after the training process is evaluated by the punch problem³⁷. The geometry of the problem is depicted in Fig. 14. The left side and the top right corner of the block are only fix in the horizontal direction and the bottom of the block is only fix in the vertical direction. During the loading stage, the vertical displacement of the top of the block u_0 is gradually increased to 0.07 mm, which can be regarded as a quasi-static process. During the unloading stage, the direction of the displacement is reversed until $u_0 = 0.02$ mm. Similar to Huang et al.³⁷, the block is divided into 100 quadratic quadrilateral elements for the FEA analysis. The load–displacement curve of the node A is recorded.

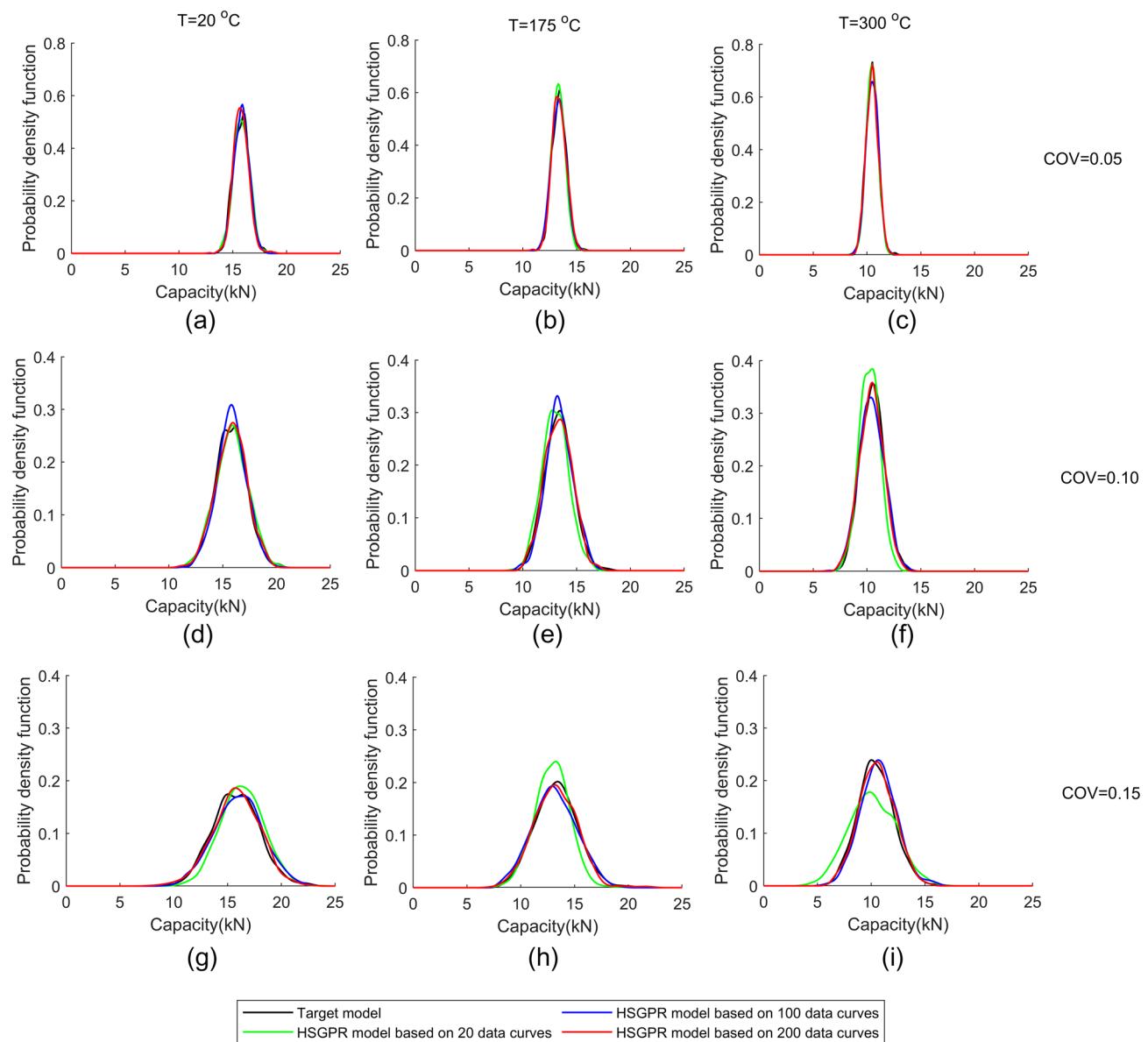


Figure 13. The probability density distribution of the structural capacity given by the HSGPR-based flow stress model and the target model at different dataset size, temperatures and coefficient of variation (COV). (a) COV = 0.05, T = 20 °C, (b) COV = 0.05, T = 175 °C, (c) COV = 0.05, T = 300 °C, (d) COV = 0.10, T = 20 °C, (e) COV = 0.10, T = 175 °C, (f) COV = 0.10, T = 300 °C, (g) COV = 0.15, T = 20 °C, (h) COV = 0.15, T = 175 °C, (i) COV = 0.15, T = 300 °C.

For the deterministic analysis, the obtained load–displacement curve is plotted in Fig. 15. The load displacement simulated by using the feed forward neural network (FNN)-based plasticity model by Huang et al.³⁷ is also presented in Fig. 15. It can be observed from Fig. 15 that the HSGPR model and FNN model given by Huang et al.³⁷ have the similar performance with both being close to the target model during the loading and unloading processes. And the average relative error of the HSGPR model is 0.75%. However, unlike the HSGPR model, the FNN model cannot capture the material uncertainty and be used in the stochastic plastic structural analysis.

Similar to the beam problem, 1000 Monte Carlo simulations are sufficient to provide the converged estimation for the punch problem as well. By conducting 1000 Monte Carlo simulations using the HSGPR model, the 95% confidence region (i.e. mean $\pm 2 \times$ standard deviation) of the load–displacement curve can be obtained and is plotted in Fig. 16. As depicted in Fig. 16, the confidence region given by the HSGPR model is in agreement with the target model and the variation of the load–displacement curve is depicted accurately by the HSGPR model in both the loading and unloading stages.

In order to investigate the influence of material uncertainty on the plastic deformation, the equivalent plastic strain of the block after loading is analyzed by using the HSGPR model. Two situations corresponding to the upper bound and the lower bound, as labeled in Fig. 16, are shown in Fig. 17. It can be observed from Fig. 17 that the plastic deformation of the upper bound situation in the lower left corner of the block is smaller than

COV	Number of data curves	T = 20 °C			T = 175 °C			T = 300 °C		
		Target (kN)	HSGPR (kN)	Relative error (%)	Target (kN)	HSGPR (kN)	Relative error (%)	Target (kN)	HSGPR (kN)	Relative error (%)
0.05	20	15.80	15.82	0.1	13.38	13.30	0.6	10.47	10.47	0.0
	100		15.82	0.1		13.34	0.3		10.49	0.2
	200		15.76	0.3		13.37	0.1		10.47	0.0
0.10	20	15.77	15.84	0.4	13.35	13.02	2.5	10.49	10.24	2.4
	100		15.84	0.4		13.40	0.4		10.43	0.6
	200		15.81	0.3		13.32	0.2		10.40	0.9
0.15	20	15.80	16.36	3.5	13.32	12.91	3.1	10.42	10.11	3.0
	100		16.05	1.6		13.12	1.5		10.68	2.6
	200		15.81	0.1		13.22	0.8		10.52	1.0

Table 6. The expectation of the structural capacity given by the HSGPR-based flow stress model and the target model at different dataset sizes, temperatures and coefficient of variation (COV).

COV	Number of data curves	T = 20 °C			T = 175 °C			T = 300 °C		
		Target (kN)	HSGPR (kN)	Relative error (%)	Target (kN)	HSGPR (kN)	Relative error (%)	Target (kN)	HSGPR (kN)	Relative error (%)
0.05	20	0.72	0.74	2.8	0.64	0.61	4.7	0.56	0.53	5.4
	100		0.70	2.8		0.66	3.1		0.56	0.0
	200		0.71	1.4		0.64	0.0		0.56	0.0
0.10	20	1.41	1.51	7.1	1.30	1.24	4.6	1.08	0.95	12.0
	100		1.34	5.0		1.27	2.3		1.16	7.4
	200		1.43	1.4		1.27	2.3		1.10	1.9
0.15	20	2.13	1.99	6.6	1.95	1.65	15.4	1.64	2.13	29.9
	100		2.21	3.8		2.03	4.1		1.68	2.4
	200		2.18	2.3		2.00	2.6		1.63	0.6

Table 7. The standard deviation of the structural capacity given by the HSGPR-based flow stress model and the target model at different dataset sizes, temperatures and coefficient of variation (COV).

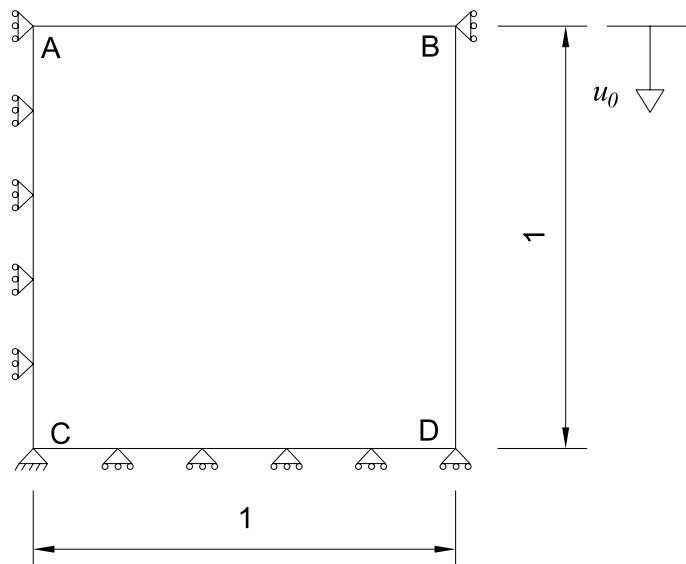


Figure 14. Punch problem given by Huang et al.³⁷ (Unit: mm).

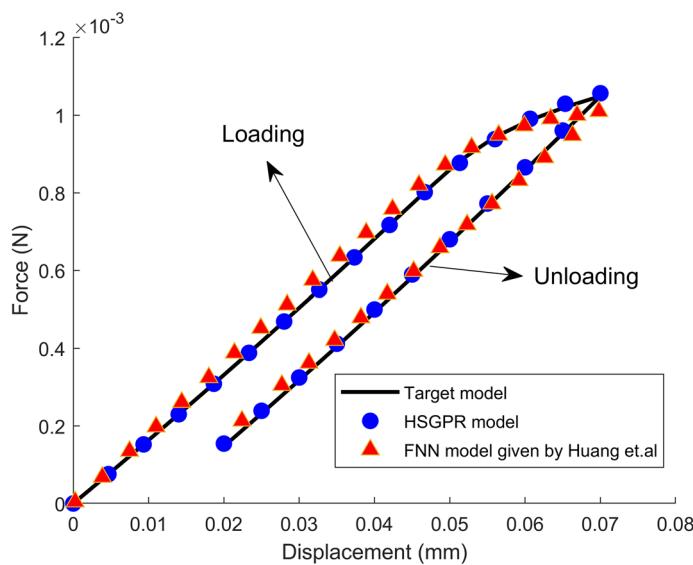


Figure 15. Deterministic load–displacement curves obtained by the target model, HSGPR model and FNN model given by Huang et. al.³⁷

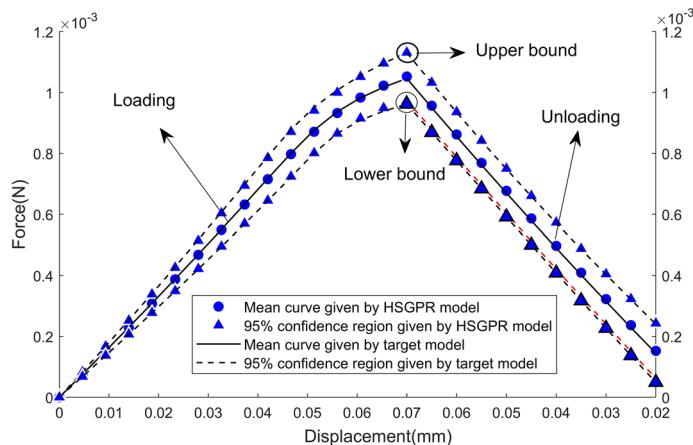


Figure 16. Stochastic load displacement curves obtained by target model, HSGPR model.

the lower bound situation. However, both the upper bound and the lower bound are the possible situations of the structural responses. The variation of the deformation is caused by the uncertainty inherent in the material. Hence, it can be concluded that the deformation of the structure cannot be described completely without sufficiently considering the flow stress uncertainty.

Conclusions

Understanding the uncertainty of material flow behavior is important in the stochastic plastic structural analysis. In this paper, a data-driven approach, the heteroscedastic sparse Gaussian process regression (HSGPR)-based flow stress model, is proposed to capture the flow stress behavior and the associated uncertainty directly from the available material data. The established HSGPR-based flow stress model is verified by the experimental data of the Al6061 aluminum alloy and is compared with the conventional GPR model, the ANN model and the Johnson Cook (JC) model. The flow stress of the Al 6061 aluminum alloy at elevated temperatures predicted by the HSGPR model are more accurate than those by the ANN model, the GPR model and the JC model, with the AARE of 6.5% and the MLL of -3.74 on the testing dataset. Besides, unlike the ANN model, which provides only the deterministic flow stress, the flow stress uncertainty is successfully estimated by the HSGPR based model as well.

The HSGPR-based stochastic flow stress model is then implemented into finite element method for stochastic plastic structural analysis. Two numerical examples with synthetic flow stress data, that is the beam problem and the punch problem, are used to verify the accuracy of the HSGPR model in stochastic plastic structural analysis. With the sufficient material data, the expectation and standard deviation of the structural capacity at elevated temperatures have the relative errors less than 1% and 3% respectively in the beam problem. In the punch problem, the variation of the loading and unloading paths given by the HSGPR model is consistent with that given

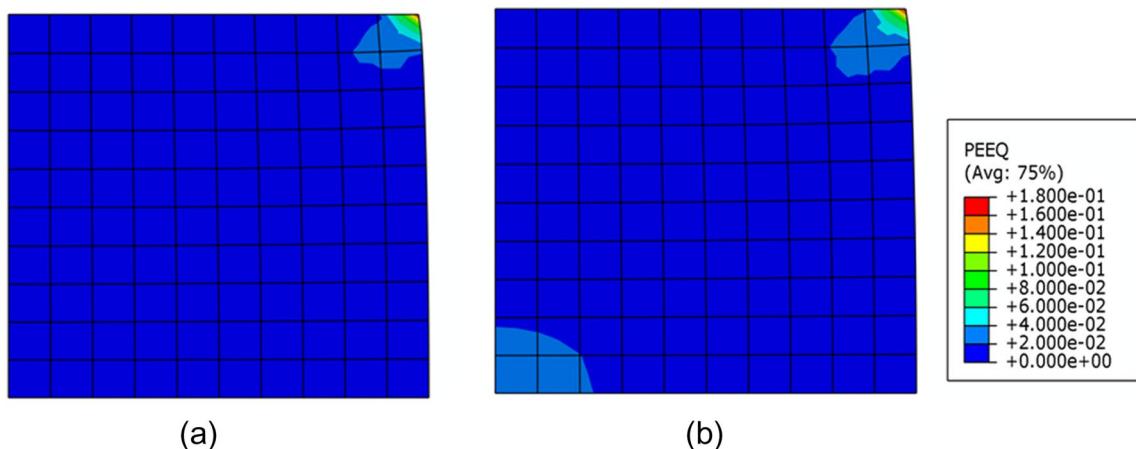


Figure 17. The equivalent plastic strains of the block after loading corresponding to (a) the upper bound and (b) the lower bound in Fig. 16.

by the target model. The influence of material uncertainty on the structure plastic deformation is successfully identified by the HSGPR model.

Data availability

The raw/processed data required to reproduce these findings cannot be shared at this time due to technical or time limitations. However, these data will be shared upon request to the corresponding author.

Received: 11 November 2021; Accepted: 3 February 2022

Published online: 22 February 2022

References

- Geng, G., Ding, D., Duan, L. & Jiang, H. A modified Johnson-Cook model of 6061-T6 aluminium profile. *Aust. J. Mech. Eng.* **1**, 1–11. <https://doi.org/10.1080/14484846.2020.1721966> (2020).
 - Zhang, Y.-B., Yao, S., Hong, X. & Wang, Z.-G. A modified Johnson-Cook model for 7N01 aluminum alloy under dynamic condition. *J. Cent. South Univ.* **24**, 2550–2555. <https://doi.org/10.1007/s11771-017-3668-5> (2017).
 - Rahmaan, T. *et al.* Anisotropic plasticity characterization of 6000- and 7000-series aluminum sheet alloys at various strain rates. *Int. J. Impact Eng.* **135**, 103390. <https://doi.org/10.1016/j.ijimpeng.2019.103390> (2020).
 - Qin, F., Qi, H., Kang, Y. & Liu, C. Study on constitutive characteristic of as-cast AA6061 alloy under plane strain compression based on orthogonal analysis. *Adv. Mater. Sci. Eng.* **2019**, 9328630. <https://doi.org/10.1155/2019/9328630> (2019).
 - Li, J. *et al.* Comparative investigation on the modified Zerilli-Armstrong model and Arrhenius-type model to predict the elevated-temperature flow behaviour of 7050 aluminium alloy. *Comput. Mater. Sci.* **71**, 56–65. <https://doi.org/10.1016/j.commatsci.2013.01.010> (2013).
 - Močko, W., Janiszewski, J. & Grążka, M. Application of an extended Ruseinek-Klepaczko constitutive model to predict the mechanical behavior of 6082-T6 aluminum under Taylor impact test conditions. *J. Strain Anal. Eng. Des.* **48**, 364–375. <https://doi.org/10.1177/0309324713488885> (2013).
 - Samantaray, D., Mandal, S. & Bhaduri, A. K. A comparative study on Johnson Cook, modified Zerilli-Armstrong and Arrhenius-type constitutive models to predict elevated temperature flow behaviour in modified 9Cr-1Mo steel. *Comput. Mater. Sci.* **47**, 568–576. <https://doi.org/10.1016/j.commatsci.2009.09.025> (2009).
 - Quan, G.-Z., Wang, T., Li, Y.-L., Zhan, Z.-Y. & Xia, Y.-F. Artificial neural network modeling to evaluate the dynamic flow stress of 7050 aluminum alloy. *J. Mater. Eng. Perform.* **25**, 553–564. <https://doi.org/10.1007/s11665-016-1884-z> (2016).
 - Haghidian, N., Zarei-Hanzaki, A., Khalesian, A. R. & Abedi, H. R. Artificial neural network modeling to predict the hot deformation behavior of an A356 aluminum alloy. *Mater. Des.* **49**, 386–391. <https://doi.org/10.1016/j.matdes.2012.12.082> (2013).
 - Sheikh, H. & Serajzadeh, S. Estimation of flow stress behavior of AA5083 using artificial neural networks with regard to dynamic strain ageing effect. *J. Mater. Process. Technol.* **196**, 115–119. <https://doi.org/10.1016/j.jmatprotec.2007.05.027> (2008).
 - Merayo Fernández, D., Rodríguez-Prieto, A. & Camacho, A. M. Prediction of the bilinear stress-strain curve of aluminum alloys using artificial intelligence and big data. *Metals* **10**, 904 (2020).
 - Forcellese, A., Gabrielli, F. & Simoncini, M. Prediction of flow curves and forming limit curves of Mg alloy thin sheets using ANN-based models. *Comput. Mater. Sci.* **50**, 3184–3197. <https://doi.org/10.1016/j.commatsci.2011.05.048> (2011).
 - Yang, Y. *et al.* Utilization of completely recycled fine aggregate for preparation of lightweight concrete partition panels. *Int. J. Concr. Struct. Mater.* **15**, 1–11 (2021).
 - Yang, Y. *et al.* Concrete mix design for completely recycled fine aggregate by modified packing density method. *Materials* **13**, 3535 (2020).
 - Tang, J. L., Liu, H. Y., Gui, M. H. & Tang, J. Y. SVR-based predictive model for 2519 aluminum alloy. *Key Eng. Mater.* **719**, 41–45. <https://doi.org/10.4028/www.scientific.net/KEM.719.41> (2017).
 - Aakash, B. S., Connors, J. & Shields, M. D. Variability in the thermo-mechanical behavior of structural aluminum. *Thin-Walled Struct.* **144**, 106122. <https://doi.org/10.1016/j.tws.2019.01.053> (2019).
 - Chen, B., Shen, L. & Zhang, H. Gaussian process regression-based material model for stochastic structural analysis. *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civil Eng.* **7**, 04021025. <https://doi.org/10.1061/AJRUA6.0001138> (2021).
 - Rasmussen, C. E. *Summer School on Machine Learning* 63–71 (Springer, 2003).
 - Su, G., Peng, L. & Hu, L. A Gaussian process-based dynamic surrogate model for complex engineering structural reliability analysis. *Struct. Saf.* **68**, 97–109. <https://doi.org/10.1016/j.strusafe.2017.06.003> (2017).
 - Hoang, N.-D., Pham, A.-D., Nguyen, Q.-L. & Pham, Q.-N. Estimating compressive strength of high performance concrete with Gaussian process regression model. *Adv. Civil Eng.* **1–8**, 2016. <https://doi.org/10.1155/2016/2861380> (2016).

21. Lin, C. *et al.* Gaussian process regression-based forecasting model of dam deformation. *Neural Comput. Appl.* **31**, 8503–8518. <https://doi.org/10.1007/s00521-019-04375-7> (2019).
22. Almosallam, I. A., Jarvis, M. J. & Roberts, S. J. GPz: non-stationary sparse Gaussian processes for heteroscedastic uncertainty estimation in photometric redshifts. *Mon. Not. R. Astron. Soc.* **462**, 726–739. <https://doi.org/10.1093/mnras/stw1618> (2016).
23. Almosallam, I. *Heteroscedastic Gaussian Processes for Uncertain and Incomplete Data* PhD thesis, University of Oxford, (2017).
24. Scapin, M. & Manes, A. Behaviour of Al6061-T6 alloy at different temperatures and strain-rates: experimental characterization and material modelling. *Mater. Sci. Eng. A* **734**, 318–328. <https://doi.org/10.1016/j.msea.2018.08.011> (2018).
25. Fan, X., Suo, T., Sun, Q. & Wang, T. Dynamic mechanical behavior of 6061 al alloy at elevated temperatures and different strain rates. *Acta Mech. Solida Sin.* **26**, 111–120. [https://doi.org/10.1016/S0894-9166\(13\)60011-7](https://doi.org/10.1016/S0894-9166(13)60011-7) (2013).
26. Zhu, D., Mobasher, B., Rajan, S. D. & Peralta, P. Characterization of dynamic tensile testing using aluminum alloy 6061-T6 at intermediate strain rates. *J. Eng. Mech.* **137**, 669–679. [https://doi.org/10.1061/\(ASCE\)EM.1943-7889.0000264](https://doi.org/10.1061/(ASCE)EM.1943-7889.0000264) (2011).
27. Manes, A., Peroni, L., Scapin, M. & Giglio, M. Analysis of strain rate behavior of an Al 6061 T6 alloy. *Procedia Eng.* **10**, 3477–3482. <https://doi.org/10.1016/j.proeng.2011.04.573> (2011).
28. Aakash, B. S., Connors, J. & Shields, M. D. Stress-strain data for aluminum 6061-T651 from 9 lots at 6 temperatures under uniaxial and plane strain tension. *Data Brief* **25**, 104085. <https://doi.org/10.1016/j.dib.2019.104085> (2019).
29. Johnson, G. R. A constitutive model and data for materials subjected to large strains, high strain rates, and high temperatures. In *Proceedings of 7th International Symposium Ballistics*, 541–547 (1983).
30. Schober, P., Boer, C. & Schwarte, L. A. Correlation coefficients: appropriate use and interpretation. *Anesthesia Analgesia* **126**, 1005 (2018).
31. Lin, Y. C. & Chen, X.-M. A critical review of experimental results and constitutive descriptions for metals and alloys in hot working. *Mater. Des.* **32**, 1733–1759. <https://doi.org/10.1016/j.matdes.2010.11.048> (2011).
32. Jansen van Rensburg, G. & Kok, S. In *Eighth South African Conference on Computational and Applied Mechanics (SACAM2012)* (Johannesburg, 2012).
33. Systèmes, D. *Abaqus 6.14 Documentation* (Dassault Systèmes, 2014).
34. Lesuer, D. R., Kay, G. & LeBlanc, M. *Modeling Large-Strain, High-Rate Deformation in Metals*. (Lawrence Livermore National Lab., 2001).
35. Su, M.-N. & Young, B. Material properties of normal and high strength aluminium alloys at elevated temperatures. *Thin-Walled Struct.* **137**, 463–471. <https://doi.org/10.1016/j.tws.2019.01.012> (2019).
36. Gillie, M. Analysis of heated structures: Nature and modelling benchmarks. *Fire Saf. J.* **44**, 673–680. <https://doi.org/10.1016/j.firesaf.2009.01.003> (2009).
37. Huang, D., Fuhg, J. N., Weißenfels, C. & Wriggers, P. A machine learning based plasticity model using proper orthogonal decomposition. *Comput. Methods Appl. Mech. Eng.* **365**, 113008. <https://doi.org/10.1016/j.cma.2020.113008> (2020).

Acknowledgements

This work was supported in part by the Australian Research Council through Discovery Projects (DP200101919 and DP190102954). The authors acknowledge the use of the National Computational Infrastructure (NCI) which is supported by the Australian Government and accessed through the Sydney Informatics Hub HPC Allocation Scheme, which is supported by the Deputy Vice-Chancellor (Research), University of Sydney and the ARC LIEF, 2019: Smith, Muller, Thorner et al., Sustaining and strengthening merit-based access to National Computational Infrastructure (LE190100021).

Author contributions

B.C. conceptualized the research, developed the methodology, performed the investigation and validation, and wrote the original manuscript. L.S. conceptualized the research, acquired the funding, supervised the research, and reviewed the manuscript. H.Z. conceptualized the research, supervised the research, and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-06870-9>.

Correspondence and requests for materials should be addressed to L.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022