Machine-learning-enhanced tail end prediction of structural response statistics in earthquake engineering

Denny Thaler | Marcus Stoffel | Bernd Markert | Franz Bamer

Institute of General Mechanics, RWTH Aachen University, Aachen, Germany

Correspondence
Franz Bamer, Institute of General Mechanics, RWTH Aachen University, Aachen 52062, Germany.
Email: bamer@iam.rwth-aachen.de

Abstract
Evaluating the response statistics of nonlinear structures constitutes a key issue in engineering design. Hereby, the Monte Carlo method has proven useful, although the computational cost turns out to be considerably high. In particular, around the design point of the system near structural failure, a reliable estimation of the statistics is unfeasible for complex high-dimensional systems. Thus, in this paper, we develop a machine-learning-enhanced Monte Carlo simulation strategy for nonlinear behaving engineering structures. A neural network learns the response behavior of the structure subjected to an initial nonstationary ground excitation subset, which is generated based on the spectral properties of a chosen ground acceleration record. Then using the superior computational efficiency of the neural network, it is possible to predict the response statistics of the full sample set, which is considerably larger than the initial training sample set. To ensure a reliable neural network response prediction in case of rare events near structural failure, we propose to extend the initial training sample set increasing the variance of the intensity. We show that using this extended initial sample set enables a reliable prediction of the response statistics, even in the tail end of the distribution.

HIGHLIGHTS
- A new Monte Carlo method is developed that provides the response statistics of a nonlinear system in the tail end of the distribution.
- A nonstationary filter is applied to a Gaussian white noise to generate realistic artificial earthquake records.
- An extended training strategy using neural networks is proposed to improve the reliability of the method in the tail end of the distribution.
- The new strategy reveals a significant speedup as well as a prediction of the response statistics in the tail end with high confidence.

KEYWORDS
earthquake generation, elastoplastic structures, Kanai–Tajimi filter, machine learning, Monte Carlo method, neural networks

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. Earthquake Engineering & Structural Dynamics published by John Wiley & Sons Ltd.
INTRODUCTION

The evaluation of the probability of failure in earthquake-prone regions is an important topic for the design of engineering structures.\textsuperscript{1-5} While statistical methods are very effective in the response prediction of linear structures, the efficient treatment of nonlinear complex structures is still a crucial issue.

The limit function between structural save and failure states can be defined in terms of a function $g(x) = 0$, whereby the vector $x$ is composed of $n$ random variables $x_1, \ldots, x_n$.\textsuperscript{6} The probability of structural failure $P_f$ can then be expressed by a multidimensional integral over an arbitrarily complex random distribution $f_X(x)$ within the defined failure region $g(x) < 0$:

$$P_f = \int \cdots \int_{g(x)<0} f_X(x) \, dx_1 \cdots dx_n. \quad (1)$$

In general, the solution of this integral is not straightforward. A promising strategy is to apply the Monte Carlo method, i.e., one deals with this integral by performing a series of computer experiments based on artificially generated distributed random inputs for $x$.\textsuperscript{7} In particular, regarding nonlinear systems, the Monte Carlo method has proven useful in providing reliable response statistics.\textsuperscript{8,9} From a statistical point of view, it is obvious that Monte Carlo sampling is more frequent around the mean of the distribution. Therefore, a disproportionately high number of samples is required to ensure reliable response statistics in the tail end of the distribution, i.e., for the reproduction of very rare events. However, it is generally the case in engineering that these very low probabilities of failure are desirable not only from a monetary but also from a social point of view.\textsuperscript{10} If the mechanical system is complex and reveals a large number of degrees of freedom, so that a single-sample computation requires a considerable amount of computational effort, the computation of a whole sample set, especially for a reliable response prediction in the tail region, is not feasible. Various strategies have been proposed to reduce this computational burden. In particular, there are two options to reduce the costs of the evaluation of the response statistics. On the one hand, one can speed up every single simulation in terms of model order reduction.\textsuperscript{11-13} This is particularly useful if the same reduced-order basis can be applied throughout the whole Monte Carlo process. In this regard, one is able to perform a large number of Monte Carlo sample runs in order to obtain reliable response statistics of the system. On the other hand, one can use more sophisticated sampling methods to increase the performance by decreasing the required number of samples to be calculated.\textsuperscript{14,15} Furthermore, an equivalent linearization of the system has proven effective in the domain near the expansion point.\textsuperscript{16-18} Since the limit states should be exceeded during very rare events only, the tail equivalent linearization method has been proposed for the design of engineering structures.\textsuperscript{19,20}

Neural networks have been found to be useful in multiple areas of mechanics and structural design.\textsuperscript{21} Based on certain input quantities, they have been used as surrogate models of nonlinear structures providing the numerical output values of interest.\textsuperscript{22-26} More elaborate recurrent neural networks have shown to be useful for the description of time history-dependent nonlinear behavior\textsuperscript{27} as well as anisotropic elastoplasticity.\textsuperscript{28}

The incorporation of machine learning for structural design has been the focus of research for decades.\textsuperscript{29} In that respect, neural network approaches have been used for seismic induced damage assessment and classification.\textsuperscript{30-32} Furthermore, structural reliability analyses have been performed by using machine learning approaches.\textsuperscript{33-38} Neural-network-enhanced estimation techniques of aftershock collapse vulnerability\textsuperscript{39} and postearthquake safety\textsuperscript{40,41} have shown auspicious results. A highly relevant topic constitutes the seismological prediction of earthquake events by neural networks.\textsuperscript{42} However, in this paper, the authors point out that deep learning does not offer more insights into aftershock prediction and should rather be incorporated in problems dealing with statistical seismology.

Neural networks can only predict complex dynamical systems as long as they have been provided with the appropriate data during training, i.e., extensive input–output relationships. They have been shown to be ineffective in extrapolating. One might say that they are not equipped with the required amount of “creativity” in that regard. Following this line of thought, the neural network generally does not succeed in the prediction of very rare events within the tail region of the distribution, although a quantification of these events is essential for reliable engineering decisions. Thus, in this paper, we present an efficient neural network that is able to reliably predict events in this region. Given that the evaluation of every training sample is computationally expensive, we create samples with a higher proportion of unlikely events. The significant characteristics from each generated artificial record will be used as input for a feedforward neural network. This enables the neural network architecture to be kept small, and therefore, to be trained fast. It is shown that the enhanced strategy is able to predict accurate response statistics and, in particular,
This paper is organized as follows: In Section 2, the basic theory of feedforward neural networks is outlined. In Section 3, the Monte Carlo simulation is discussed briefly, along with its general shortcomings. The nonlinear Kanai–Tajimi filter is introduced in that section, and the machine learning enhanced Monte Carlo simulation strategy is proposed. Section 4 shows a numerical example, comparing the new proposed strategy with the Monte Carlo benchmark solution as well as with the standard neural network procedure. Finally, in Section 5, conclusions are drawn.

2 FEEDFORWARD NEURAL NETWORKS IN A NUTSHELL

A feedforward neural network consists of at least three layers: input, hidden, and output layers, as shown in Figure 1. The number of hidden layers has to be chosen due to the level of complexity of the problem. Each layer $l$ ($l = 1, \ldots, L$) consists of neurons, each of which is—in the case of fully connected feedforward neural networks—connected with every neuron in the next layer. Thus, every neuron in layer $l$ receives the input signal $x^{(l)}$ from all the neurons in the previous layers. These inputs are weighted with matrix $W^{(l)}$ and combined with the bias vector $b^{(l)}$. The weighted input $z^{(l)}$ is then applied to an activation function $f(z^{(l)})$. It is written as:

$$\hat{y}^{(l)} = f(z^{(l)}) \text{ with } z^{(l)} = W^{(l)}x^{(l)} + b^{(l)}. \tag{2}$$

Hereby, $\hat{y}^{(l)}$ is the output of a layer $l$. As “feedforward” already indicates, the data in this network type only flow forward. Starting with the input layer, the data are hand over to the output layer via a variable number of hidden layers. The outcome of the last layer $L$ of the neural network is denoted by $\hat{y} = \hat{y}^{(L)}$, which is then the prediction of the neural network.

The training of the neural network is done by an error back propagation, using the gradient descent algorithm. This algorithm updates the weights $W^{(l)}$ for each layer to minimize a chosen error function (e.g., normalized mean squared error $\varepsilon = \frac{1}{N_y}(y - \hat{y})^2$), based on $N_y$ given targets within the vector $y$. The weight matrices $W^{(l)}$ are then evaluated by:

$$\Delta W^{(l)} = -\alpha \frac{\partial \varepsilon}{\partial W^{(l)}} \text{ with } \frac{\partial \varepsilon}{\partial W^{(l)}} = \frac{\partial z^{(l)}}{\partial W^{(l)}} \frac{\partial \hat{y}^{(l)}}{\partial W^{(l)}} \frac{\partial \hat{y}^{(l)}}{\partial z^{(l)}}, \tag{3}$$

where $\alpha$ is the learning rate, which controls the size of the adjustment in each step. Thus, the learning rate influences the time until convergence. However, if it is too large, it can result in suboptimal or diverging solutions. The learning rate can be updated during training and calculated individually for each entry. The activation function $f$ must be chosen according to the network type and the problem at hand. This choice affects the training speed as well as the performance of the neural network. The weight update of the layers, using the mean squared error and the gradient descent backpropagation,
is calculated by:

\[ \delta^{(l)} = \frac{\partial \epsilon}{\partial \dot{y}^{(l)}} = \begin{cases} \frac{2}{N_y} (\dot{y} - y) & \text{for } l = L \\ W^{(l+1)} \frac{\partial \dot{y}^{(l+1)}}{\partial \dot{z}^{(l+1)}} & \text{for } l < L \end{cases}, \tag{4} \]

\[ \frac{\partial z^{(l)}}{\partial W^{(l)}} = \mathbf{I} \otimes x^{(l)}, \tag{5} \]

with the activation function

\[ \frac{\partial \dot{y}^{(l)}}{\partial z^{(l)}} = \text{diag} \left\{ f'(z^{(l)}) \right\}. \tag{6} \]

The bias \( b^{(l)} \) is updated equivalently. Additional information about the theory of neural networks can be found in the existing literature.\textsuperscript{44, 47} For the implementation of neural networks in this paper, the script language Python\textsuperscript{48} has been used together with the library TensorFlow.\textsuperscript{49}

### 3 MACHINE LEARNING ENHANCED MONTE CARLO SIMULATION

#### 3.1 Estimation of the probability of failure using the Monte Carlo simulation

Extending Equation (1) by an indicator function \( I_g(x_1, \ldots, x_n) \), the probability of failure can be rewritten as the multiple integral:\textsuperscript{6}

\[ p_f = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} I_g(x_1, \ldots, x_n) f_{X_1, \ldots, X_n}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n, \text{with } \begin{cases} \text{if } g(x_1, \ldots, x_n) \leq 0 \rightarrow I_g(x_1, \ldots, x_n) = 1 \\ \text{else } I_g(x_1, \ldots, x_n) = 0 \end{cases}. \tag{7} \]

Performing numerical experiments by introducing a set of randomly generated ground excitations \( \dot{x}_g^{(k)}(t) \) \( (k = 1, \ldots, m) \), one can numerically evaluate the structural response set \( \dot{x}^{(k)}(t) \) \( (k = 1, \ldots, m) \) and, furthermore, obtain the indication set of structural failure:

\[ \langle \dot{x}_g^{(1)}(t), \ldots, \dot{x}_g^{(m)}(t) \rangle \longrightarrow \langle \dot{x}^{(1)}(t), \ldots, \dot{x}^{(m)}(t) \rangle \longrightarrow \langle I_g(\dot{x}^{(1)}), \ldots, I_g(\dot{x}^{(m)}) \rangle. \tag{8} \]

Thus, the Monte Carlo method allows one to define the probability of failure in terms of an expected value:

\[ p_f = \frac{1}{m} \sum_{k=1}^{m} I_g(\dot{x}^{(k)}). \tag{9} \]

#### 3.2 Generating sample excitations

The Kanai-Tajimi\textsuperscript{50, 51} filter has successfully been applied to model the movement of the ground in terms of the response of a linear single degree of freedom oscillator to a white noise excitation. The relevant input parameters are then \( \omega_g \) and \( \zeta_g \) as an estimation of natural circular frequency and damping ratio of the ground. Thus, one is able to take site-dependent ground conditions into account. The power spectral density of the ground response has a significant frequency
dependence.\(^6\)

\[
S_{aa}(\omega) = S_0 \frac{4\zeta^2 \omega_0^2 \omega^2 + \omega_g^4}{(\omega_g^2 - \omega^2)^2 + 4\zeta^2 \omega_g^2 \omega^2}.
\] (10)

The power spectral density \(S_{aa}(\omega)\) is multiplied by two independently generated stationary white noise samples, so that one real and one imaginary part is generated. Inverse Fourier transformation reveals the filtered white noise excitation in the time domain. Following this procedure, stationary artificial earthquake records can be generated and used for the Monte Carlo simulation procedure. However, the properties of an earthquake generally change with time. This concerns mostly intensity and frequency content. These two time- and site-dependent parameters are extracted from a real earthquake record, based on which a nonlinear Kanai–Tajimi filter\(^{12, 52, 53}\) is introduced that provides a set of more realistic ground excitation samples \(\ddot{x}_g^{(k)}(k = 1, \ldots, m)\). This strategy preserves the relevant properties of the record, and therefore, generates the desired site-dependency.\(^52\) The site- and time-dependent frequency content \(\omega_g(t)\), as well as intensity \(e(t)\), are identified and used to generate a large number of artificial records by variation of those two parameters within their allowable range.\(^52\)

In this paper, the frequency content \(\omega_g(t)\) and the intensity \(e(t)\) are extracted from one real earthquake record, chosen as a representative event. We hereby note that, using this strategy, it is also possible to introduce more measured records, preferably from one measurement site, to serve as a total set of representative events.

### 3.2.1 Identification of frequency and intensity

A moving time window is introduced that scans the desired properties of a measured record at certain time instants, as illustratively depicted in Figure 2. The constant time window, whose size is chosen by a parameter \(t_w\), moves through the record to extract statistically measured data within the time frame \([t - \frac{t_w}{2}, t + \frac{t_w}{2}]\). The choice of \(t_w\) is crucial for a meaningful extraction of the desired quantities. If it is too small, large oscillation will be observed. If it is too large, important features could get lost. A parameter study has shown that, for the problem in this paper, a value of 1.0 second for this time period is a good choice to avoid strong oscillations, while preserving sufficiently accurate time histories.

First, we extract the frequency content by counting the number of zero-crossings \(\hat{n}\) in the chosen time window within the record. The raw time-dependent ground frequency \(\hat{\omega}_g(t)\) is then written as:

\[
\hat{\omega}_g(t) = \frac{\hat{n}(t)2\pi}{t_w}.
\] (11)

Second, we evaluate the raw intensity of the record \(\hat{e}(t)\) by integrating the squared ground acceleration over the moving time window:

\[
\hat{e}(t) = \int_{t-\frac{t_w}{2}}^{t+\frac{t_w}{2}} (\dddot{x}_g^{(r)})^2 dt,
\] (12)

where \(\dddot{x}_g^{(r)}\) denotes the recorded ground acceleration.
The earthquake record, based on which the artificial excitations are generated in this paper, has been measured in Kobe Takarazuka in 1995. The time-dependent frequency content $\hat{\omega}_g(t)$, evaluated based on the extracted number of zero-crossings $\hat{n}(t)$ using relation (11), is presented in Figure 3(A). The extracted time-dependent intensity is presented in Figure 3(B).

Finally, polynomial functions are fitted to the extracted data, as shown in Figures 3(A) and 3(B). This results in the representation of the frequency data $\hat{\omega}_g(t)$ by a polynomial function $\omega_g(t)$. Equivalently, this results in the representation of the intensity data $\hat{e}(t)$ by a polynomial function $e(t)$. Both curves are depicted as red lines in Figures 3(A) and 3(B).

3.2.2 The nonstationary Kanai–Tajimi filter

The nonlinear formulation of the Kanai–Tajimi filter is represented by a nonlinear single degree of freedom system excited by a stationary Gaussian white noise $n(t)$:

$$\ddot{x}_f^{(k)} + 2\zeta_g \omega_g(t) \dot{x}_f^{(k)} + \omega_g^2(t) x_f^{(k)} = n(t).$$

(13)

This nonlinear single degree of freedom system is solved numerically using the explicit central difference integration scheme obtaining the filter displacement $x_f$ and the filter velocity $\dot{x}_f$. The filter is then obtained by the relation:

$$\ddot{x}_g^{(k)} = -2\zeta_g \omega_g(t) \dot{x}_f^{(k)} - \omega_g^2(t) x_f^{(k)}.$$

(14)

Finally, $\ddot{x}_g$ is multiplied by the extracted function $e(t)$, which describes the intensity of the earthquake over time:

$$\ddot{x}_g^{(k)} = \ddot{x}_g e(t).$$

(15)

The effective ground damping is chosen as $\zeta_g = 0.1$, and the spectral density is chosen as $G_0 = 1.0 \left[ \frac{\text{m}^2}{\text{s}} \right]$ for the white noise generation. The real Kobe earthquake record is shown in Figure 4(A), while one sample using the proposed generation strategy is shown in Figure 4(A). Clearly, the randomization of this approach comes from the white noise excitation on the right-hand side of Equation (13), while the seismic target properties are contained in the left-hand side of this equation.

3.3 Evaluation of the response function

We introduce the implicit formulation of the nonlinear set of equations of motion for a structure subjected to ground motion:

$$M \ddot{x}^{(k)} + C \dot{x}^{(k)} + K \Delta x^{(k)} = -M f \ddot{x}_g^{(k)} - R(x^{(k)}, \dot{x}^{(k)}),$$

(16)
where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ denote mass, damping, and the tangential stiffness matrix, respectively. The restoring force vector $\mathbf{R}(\mathbf{x}^{(k)}, \dot{\mathbf{x}}^{(k)})$ is responsible for nonlinear material behavior and corresponds, in case of application of the finite element method, to the internal stresses on Gauss integration point level.\(^{12}\) The components of the influence vector $\mathbf{f}$ amount to 1, if the corresponding degrees of freedom are affected by the ground motion $\ddot{\mathbf{x}}^{(k)}_g$, and amount to 0 otherwise. For the present paper, only the horizontal degrees of freedom are subjected to ground acceleration. To solve the nonlinear set of equations of motion (16), we used the implicit version of the Newmark equations.\(^{58,59}\)

Choosing values of $\frac{1}{4}$ and $\frac{1}{2}$ for the parameters $\beta$ and $\gamma$ of the Newmark equations leads to an unconditionally stable algorithm for linear elastic problems. In that regard, the stability analysis is most critical for zero damping.\(^{68}\) Additionally, stability and accuracy of the solution are verified by the solution obtained using the central difference integration scheme.\(^{12,55,56}\)

In either case, numerical strategies for the evaluation of the structural time-dependent response $\mathbf{x}(t)$ and its derivatives are generally computationally expensive. In particular, the required computational effort increases disproportionately with the number of degrees of freedom of the system. One can imagine that a whole set of response calculations during a Monte Carlo simulation ($k = 1, \ldots, m$) leads to even more unfeasible computation times, especially, if the number of samples to be calculated is high.

### 3.4 Machine learning enhanced evaluation of tail-end probabilities

The Monte Carlo sampling, defined in Equations (7) and (9), turns out to be effective for nonlinear, unpredictable systems, as it provides a consistent and unbiased estimate of $p_f$. Obviously, the probability of failure of engineering structures has to be very low in order to ensure infrastructural environments with the required level of safety. The standard deviation of this estimate of the failure probability, $\sigma_{p_f}$, is written as\(^6\):

$$\sigma_{p_f}^2 = \frac{p_f^2}{m} - \frac{p_f^2}{m} \approx \frac{p_f}{m} \Rightarrow \sigma_{p_f} = \sqrt{\frac{p_f}{m}}.$$  

(17)

The variance of the estimate of the probability of failure, $\sigma_{p_f}^2$, increases with a decreasing number of samples. In other words, for a small number of samples $m$, the reliability of the estimate of a small probability of failure is very low, and a disproportionately large number of samples must be evaluated so that reliable response statistics are obtained in the tail region of the distribution. Figure 5(A) demonstrates this scenario on a simple illustrative example with two random variables. Only one sample is randomly generated here that results in structural failure, i.e., $g(\mathbf{x}) \leq 0$. A significantly higher number of samples is necessary to estimate the expectation in Equation (9) with high confidence, as shown in Figure 5(B). Taking into consideration that numerical algorithms in structural dynamics are already time-consuming for a single sample, a reliable estimation of $p_f$ becomes unfeasible if a high number of degrees of freedom is involved.\(^8\)

Following this line of thought, cheap surrogate models of the complex nonlinear systems would enable us to evaluate a significantly larger number of samples, and therefore, provide a reliable estimate of the probability of failure for...
Naturally, one would train the neural network using a smaller initial set of earthquake samples that have the seismic properties taken from the extracted data \( \omega_g(t) \) and \( e(t) \). In this paper, we used the Kobe earthquake record for the evaluation of the seismic properties, as discussed in Section 3.2.

Figure 5 illustrates the following ideas on two random variables for demonstrative purposes. Looking at Figure 5(A), it can easily be seen that a neural network will be trained to reconstruct the response behavior around the mean of the distribution with most of the samples. Thus, the neural network will have higher accuracy in predicting events based on values from the mean region compared to low probability events. Accordingly, it will not learn to model the response behavior in the region of interest, i.e., in the region around structural failure \( g(\mathbf{x}) \leq 0 \), as neural networks are powerful in interpolation but not in extrapolation. A significant increase in the number of samples for the initial training set, as shown in Figure 5(B), would, if computationally possible, improve the situation. However, the whole proposed strategy in this paper would obviously be obsolete.

Instead, we propose an enhanced strategy that ensures a training set with a significantly higher share of extreme events. Therefore, the training earthquake samples are generated with an additional variance parameter for the intensity. For that, a factor \( \lambda \) is uniformly distributed. While the white noise generation is multiplied by \( f_0 = \sqrt{2\pi G_0} \) for the standard sampling, the factor \( f_1 = \lambda \sqrt{2\pi G_0} \) is used to generate samples for the extended sample set.

This allows us to cover significantly larger ranges, as illustratively shown for two variables in Figure 5(C). In doing so, the neural network will be trained sufficiently in the region of structural failure \( g(\mathbf{x}) \leq 0 \). Thus, the neural network is able to predict the structural response in this region with higher accuracy, as will be shown in Section 4.3 in this paper.

We propose to use the input layer of the neural network to include excitation quantities and the output layer to provide the structural response quantities necessary for engineering decisions. The peak story drift ratio is chosen for the latter calculated on the basis of the structural response time history. The first crucial decision to accurately provide the output quantity of interest by the neural network surrogate model is the amount of data provided as the input. Choosing the whole acceleration time history of an earthquake has the potential to result in very accurate output values. However, it will lead to huge neural network architectures and inefficient training as well as prediction procedures. In light of that, we decided to keep the neural network small. Therefore, we selected input quantities from earthquake intensity and frequency data instead. As shown in Figure 6, some selected intensity measures from the acceleration record, the deformation, the velocity, and the acceleration response spectra are used as inputs.

There are several intensity measures that come into consideration as neural network input parameters used to predict selected structural response quantities due to a seismic ground excitation. It has been shown that the use of five or more input parameters can already lead to a high level of accuracy. In that regard, Housner intensity, peak ground acceleration, and effective peak acceleration (EPA) have been shown to be promising. Furthermore, in several studies Arias intensity, characteristic intensity and cumulative absolute velocity (CAV) have also been found useful. Based on those studies, we chose to consider input parameters from the set of quantities listed in Table 1. The choice of the best parameters depends
on the structure and the data available. We present the choice of the input parameters considered for the problem at hand in the table below.

We note that, applying the proposed neural network enhanced strategy, the incorporation of further intensity measures as input parameters can lead to an improvement regarding the accuracy of the response predictions. In particular, the geometric mean of the spectral acceleration has shown to be a promising intensity measure and should be considered in future studies in this regard.

As soon as the neural network training reaches convergence with the target values, calculated by the finite element method \( \varepsilon = \frac{1}{N}(x_{\text{max}} - \hat{x}_{\text{max}})^2 \rightarrow 0 \), a validation set is evaluated to verify the neural network predictions. To reach satisfying results, a trial and error process—as described in Section 4.2 for the chosen numerical example—is necessary to find the neural network architecture that performs well on the given task. Additionally to the decision regarding the neural network type, the number of input variables, hidden layer, and neurons per layer must be determined.

The trained neural network can then be used to predict the response statistics. The input values for the neural network must still be calculated for every seismic excitation generated. However, the computational effort for the calculation of the quantities shown in Table 1 is small, if compared to the effort for a whole finite element calculation of the dynamic response when structures with a high number of degrees of freedom are involved. Therefore, the novel neural network enhanced Monte Carlo simulation method enables us to obtain quick results by evaluating a huge amount of sample ground accelerations, as schematically shown in Figure 5(B).
TABLE 1  Intensity measures considered as possible input parameters for the neural network

| Intensity measure               | Abbreviation | Formula       |
|--------------------------------|--------------|---------------|
| Peak ground acceleration       | PGA          | max |̄x|g|   |
| Arias intensity                | I_a          | \( \frac{1}{2g} \int_0^T \ddot{x}_g dt \) |
| Cumulative absolute velocity   | CAV          | \( \int_0^T |\ddot{x}_g| dt \) |
| Acceleration spectrum intensity| \( S_a \)  | \( \int_{0.1}^{0.5} S_a(T) dT \) |
| Velocity spectrum intensity    | \( S_v \)  | \( \int_{0.1}^{0.5} S_v(T) dT \) |
| Housner intensity              | \( S_d \)  | \( \int_{0.1}^{0.5} S_d(T) dT \) |
| Effective peak ground acceleration | EPA      | \( \frac{1}{2.5} \int_{0.1}^{0.5} S_a(T) dT \) |
| Maximum spectral acceleration  | max \( S_a \) | max \( S_a(T) \) |
| Spectral acceleration at \( T_1 \) | \( S_a(T_1) \) | \( S_a(T_1) \) |
| Spectral velocity at \( T_1 \)  | \( S_v(T_1) \) | \( S_v(T_1) \) |
| Spectral displacement at \( T_1 \) | \( S_d(T_1) \) | \( S_d(T_1) \) |

FIGURE 7  Three-story-two-bay frame structure used as a numerical example to demonstrate the proposed strategy, applied cross-sections, and elastoplastic material behavior with kinematic hardening

4  | NUMERICAL EXAMPLE

In this section, we present a numerical example of the neural network enhanced Monte Carlo simulation. For the implementation of this example, we used a python and C++ based, in-house finite element tool for the nonlinear structural dynamical calculations. A nonlinear frame structure is subjected to a set of artificial nonstationary earthquake excitations, as described in Section 3.2.

4.1  | The structure

The frame structure consists of three stories and two bays, with a total width of 12.0 m and a total height of 12.5 m, as shown in Figure 7. A height of 4.5 m is chosen for the first story and 4.0 m for the remaining stories. An elastoplastic material with kinematic hardening is used for the whole frame structure. The Young’s modulus of the material is \( E_1 = 210 \) MPa.
After exceeding the yield criterion of $\sigma_y = 235$ GPa, the stiffness is reduced to 10% of the initial stiffness $E_1 = 0.1 E_1$. The density of the material is $\rho = 7850$ kg/m$^3$. A rectangular hollow cross-section ($0.3 \text{ m} \times 0.4 \text{ m}$), with a thickness of 0.03 m, is chosen for the beams. In order to model the mass of the slabs, point masses of 3000 kg are added to the nodes of each beam. The columns are modeled as a hollow square cross-section with a width of 0.3 m and a thickness of 0.03 m.

The structure is subjected to the generated earthquakes, while the structural response is calculated using the Newton–Raphson algorithm. Analyzing the structural response reveals that maximum plastic stress–strain relations occur within the cross-sections of the beam in the region of the frame corners. Based on this observation, we introduce the story drift ratio as a measure of the level of damage. A peak story drift ratio equal to 4.5% corresponds to a full plastification in the frame corner cross-section of the structure and is in line with the existing literature.

### 4.2 Input parameters and hyperparameter search

In this section, the choice of the input parameters for the neural network estimation is briefly discussed. However, the extent of the study carried out within this project does not necessarily ensure the highest neural network performance. Investigations based on different parameters for the inputs using a classical neural network approach can be found in the existing literature. Notably, Arias intensity, characteristic intensity, and CAV are chosen more often as intensity measures as input for the neural network prediction.

As mentioned in Section 3.4, earthquake intensity measures are used to characterize the generated accelerations. These measures can be derived from the generated accelerations and from response spectra. In order to keep the neural network small and fast, we want to use few input parameters, and therefore, only the most relevant features. This input parameter search is started by using only one input measure. The input parameter that leads to the best performance of the neural network is then the first fixed input quantity. Furthermore, the choice of the next input could be made by ranking the performance of all the remaining input parameters. However, we want to ensure that the neural network is not fed by too much redundant data. This is based on the assumption that highly correlated input features will not result in beneficial output predictions. Thus, we also pay attention to the correlations between the input values, as shown in Figure 8.

The spectral displacement at the natural period of the structure $S_d(T_1)$ is observed to be the best single input parameter. This is not surprising because it has the highest correlation with the failure criterion chosen from the structural response. Therefore, this intensity measure is chosen as the fixed input parameter. Due to the correlations shown in Figure 8 and their performance in combination with the spectral displacement at the natural period $S_d(T_1)$, further parameters are chosen. Stepwise increasing of the number of input parameters and taking into account previous studies, the
THALER et al.

**Figure 9** Input parameter distributions visualized as violin plots; distributions of the intensity measures from the generated accelerations without additional intensity variation (A) and with additional intensity variation (B)

The following intensity measures are chosen: spectral acceleration at the natural period $S_a(T_1)$, velocity spectrum intensity $\int S_v$, spectral displacement at the natural period $S_d(T_1)$, CAV, peak ground acceleration (PGA) and EPA.

Additionally to the choice of the input parameters, the distribution of these chosen features is important for the simulation strategy proposed in Section 3.4. The violin plots, as shown in Figure 9, provide the information about these distributions. The values have been standardized, $z = \frac{x - \mu}{\sigma}$, to compare all the input features. Hereby, $\mu$ denotes the expected value and $\sigma$ is the standard deviation. The standardization of these values uses the same mean and standard deviation calculated from the distributions of the generated accelerations without additional variation. In Figure 9(A), these distributions are shown, whereas Figure 9(B) shows the distributions of the intensity measures of the generated accelerations with additional variation. This figure reveals that the intensity measures chosen behave as desired to reach the objectives of this paper. They are stretched by the use of the additional factor $\lambda$ in the white noise generation, as described in Section 3.4. The input parameters are scattered over a larger range, and therefore, cover a larger prediction range regarding possible inputs for the Monte Carlo simulation later.

We applied the training strategy, as shown in Figure 6, using a training input of 400 randomly chosen samples of the classical simulation sample set. A hyperparameter search is performed on this data set. The neural network architecture is found by varying the hidden layer between one and four layers, with up to 30 units per layer. The parameters are modified until a satisfying convergence is achieved. In order to avoid overfitting, a validation set of 100 further samples is used. Next to the neural network architecture, the most common activation functions have been tested. This includes logistic sigmoid $f(z) = \frac{1}{1+e^{-z}}$, hyperbolic tangent $f(z) = \frac{e^{2z} - 1}{e^{2z} + 1}$, and rectified linear units $f(z) = \max(0, z)$. We note that the neural network, found by the hyperparameter tuning procedure, performs well in respect of the given task. The architecture of the neural network consists of three hidden layers with 10 neurons in each layer. Rectified linear activation functions are used for the neurons of the hidden layers, whereas the output unit has a linear activation function.

**4.3 Results**

For the numerical demonstration of the new strategy, a neural network is trained using 400 scattered samples with higher intensity variance, as illustratively shown in Figure 5(C) and discussed in Section 3.4. The factor $\lambda$, introduced in Section 3.4, is uniformly randomized between 0.8 and 1.5 and multiplied by the generated white noise. This obviously results in a higher number of extreme events, as shown in Figure 9 and discussed in Section 4.2. The method is compared with the full benchmark Monte Carlo solution, calculated by the finite element method and the neural network method following the classical training procedure.

To obtain a feasible benchmark result for the Monte Carlo simulation, the peak story drift ratios are evaluated for a number of $10^4$ earthquake samples. Using the in-house tool, the evaluation of the structural response takes between 80 and 160 seconds for one earthquake sample, dependent on the severity of plastic deformation. The mean simulation time...
TABLE 2 Comparison of computational efforts between the crude Monte Carlo simulation and the new proposed Monte Carlo strategy

| Simulation strategy       |            | Full evaluation |            |            | Response spectra |            |            | Total time (seconds) |
|---------------------------|------------|-----------------|------------|------------|------------------|------------|------------|---------------------|
|                           |            | Samples         | Simulation time (seconds) | Samples | Simulation time (seconds) |            |            |                     |
| Crude Monte Carlo         |            | $10^4$          | 120        |            |                  |            |            | $1.2 \times 10^6$   |
| Proposed new strategy     |            | 400             | 120        | $10^4$     | 1                |            |            | $5.8 \times 10^4$   |

**FIGURE 10** Structural response to one sample excitation. Story drift ratio of the first, second, and third floors depicted in black, red, and blue color, respectively.

for $10^4$ samples is 124 seconds. The calculation of the response spectra and the evaluation of the inputs for the neural network, i.e., the chosen set of intensity measures, takes around 1 second.

For the neural network prediction, the input features need to be extracted from the acceleration histories and the corresponding response spectra. Furthermore, the training targets need to be evaluated. Thus, the structural response is calculated with the finite element method for the training samples, and the peak story drift ratios are extracted. Once both neural networks perform well on the training and validation sets, they are used to predict the full response statistics of the $10^4$ samples. Even though the response spectra must still be calculated for the input of the neural network, the speedup of this simulation compared to the crude Monte Carlo is outstanding, as presented in Table 2. The time needed to predict the response using neural networks is much smaller, and therefore, neglected in the comparison of the computational efforts. For the numerical example in this paper, the whole neural network approach is 20 times faster, if 400 samples are used for the initial training subset.

**FIGURE 11** Results of Monte Carlo simulation with $10^4$ samples: (A) probability density function and (B) complementary cumulative distribution function.
line with the standard Monte Carlo method. However, as expected, the complementary cumulative distribution function (Figure 11B) reveals that the tail region is badly predicted by the Monte Carlo strategy with the classically trained neural network.

Using more samples of unlikely earthquakes for the training of the neural network covers a wider range of inputs, which leads to better predictions in the tail region, as shown in Figure 11(B). The probability density function in Figure 11(A) shows that the neural network with classical training outperforms the neural network with extended training in the mean region. Comparing the value of the peak story drift ratios that exceeded a certain value by the probability of $p = 50\%$, we find that the standard neural network approach has a percentage error of only 0.3\%, whereas the proposed method deviates by 6.2\% from the solution of the crude Monte Carlo simulation. However, in the tail region, the neural network with the extended training strategy considerably outperforms the neural network with classical training.

In this paper, we define structural failure if the peak story drift ratio exceeds a value of 4.5\%, as shown by the vertical lines in Figures 11(A) and 11(B). This assumption is based on the plastification of the structure, as discussed in Section 4.1, and in line with the relevant literature. One can observe that the classical neural network training strategy fails to predict the probability of failure. However, the extended strategy improves the quality of the estimation significantly. The prediction of the probability of failure using the proposed strategy reveals a percentage error of 1.4\%, whereas the classical approach fails completely, and therefore, reveals an error of 100\%. From another perspective, one can also be interested in the peak story drift, which is exceeded with a probability of 0.1\% or 1\%. The peak story drift of the crude Monte Carlo simulation exceeds the values of $\text{PSDR} (p = 1\%) = 3.90\%$ and $\text{PSDR} (p = 0.1\%) = 4.67\%$, with a probability of 1\% and 0.1\%, respectively. The standard neural network approach estimates the values $\text{PSDR} (p = 1\%) = 3.44\%$ and $\text{PSDR} (p = 0.1\%) = 3.80\%$, whereas the extended approach predicts the values $\text{PSDR} (p = 1\%) = 3.87\%$ and $\text{PSDR} (p = 0.1\%) = 4.62\%$. This results in a percentage error of 11.8\% for the classical method and 0.6\% for the extended method, if the values with a probability of 1\% are compared. This is even more visible if the chosen limit is the probability of 0.1\%. The errors are found to be 18.6\% for the classical training and 0.9\% for the extended training.

We conclude that a feedforward neural network with three hidden layers and 10 neurons in each layer is able to predict the peak story drift ratios for the problem in this paper. Furthermore, the training of the neural network is efficient, since it contains a rather small number of parameters. Therefore, the number of time-consuming finite element calculations for the training set can be kept small.

5 | CONCLUSION

In this paper, we developed a neural network enhanced Monte Carlo simulation that provides the response statistics of a ground accelerated structure. The new strategy reveals an outstanding efficiency when compared to the crude Monte Carlo method and ensures a reliable prediction in the tail end of the distribution.

The main benefit of using neural networks to predict the response statistics is computational savings. However, in order to train a neural network, a certain number of calculated samples are necessary. To ensure that a neural network is able to perform well on a given task, it is crucial to provide the necessary information. The proposed method enables the neural network to predict low probability events in the tail region of the distribution with higher accuracy compared to a standard neural network approach. The computational effort to calculate the full response sample set decreases significantly compared to the standard Monte Carlo procedure.

We intentionally use a small feedforward neural network, which requires a rather smaller number of training samples, and therefore, a short training time period is sufficient to obtain accurate response statistics. Six input parameters, chosen from possible earthquake intensity measures, are sufficient to provide accurate response statistics.

By contrast with conventional machine learning approaches, we propose an adaptation to problems in earthquake engineering and used an extended sample range for the neural network training, taking site-dependent behavior into account. In that respect, we concentrate on accurately predicting the response statistics near structural failure and ensure that the neural network is sufficiently trained in this particular region. The major advantage of this approach is that the neural network does not have to extrapolate the predictions in case of very rare events, which would lead to unacceptably inaccurate predictions in the tail end of the distribution. A side effect of the proposed method is the accuracy loss in high probability events. The conventional approach, using the standard sample range, is able to predict this region better. However, to evaluate the probability of structural failure, this trade-off can be neglected. This idea makes our method particularly effective for the prediction of events near structural failure.
REFERENCES

1. Der Kiureghian A. Structural reliability methods for seismic safety assessment: a review. *Eng Struct.* 1996;18(6):412-424.
2. Kafali C, Grigoriu M. Seismic fragility analysis: application to simple linear and nonlinear systems. *Earthq Eng Struct Dyn.* 2007;36(13):1885-1900.
3. Fajfar P, Dolšek M. A practice-oriented estimation of the failure probability of building structures. *Earthq Eng Struct Dyn.* 2012;41(3):531-547.
4. Pinto P. Reliability methods in earthquake engineering. *Progr Struct Eng Mater.* 2001;3(1):76-85.
5. Ghobareh A. Performance-based design in earthquake engineering: state of development. *Eng Struct.* 2001;23(8):878-884.
6. Bucher, C. *Computational Analysis of Randomness in Structural Mechanics.* London: Taylor & Francis; 2009.
7. Shinozuka M. Monte Carlo solution of structural dynamics. *Comput Struct.* 1972;2(5):855-874.
8. Hurtado J, Barbat H. Monte Carlo techniques in computational stochastic mechanics. *Arch Comput Methods Eng.* 1998;5:3-30.
9. Ellingwood B. Earthquake risk assessment of building structures. *Reliab Eng Syst Saf.* 2001;74:251-262.
10. Elnashii, A, Di Sarno, L. *Fundamentals of Earthquake Engineering.* Chichester, UK: Wiley; 2008.
11. Bamer F, Amiri A, Bucher C. A new model order reduction strategy adapted to nonlinear problems in earthquake engineering. *Earthq Eng Struct Dyn.* 2017;46(4):537-559.
12. Bamer F, Markert B. An efficient response identification strategy for nonlinear structures subject to nonstationary generated seismic excitations. *Mech Based Des Struct Machine.* 2017;45:313-330.
13. Bamer F, Bucher C. Application of the proper orthogonal decomposition for linear and nonlinear structures under transient excitations. *Acta Mech.* 2012;222:2549-2563.
14. Bucher C. Asymptotic sampling for high-dimensional reliability analysis. *Probabilistic Eng Mech.* 2009;24(4):504-510.
15. Breitung K. Asymptotic approximations for multinormal integrals. *J Eng Mech.* 1984;110(3):357-366.
16. Caughey T.K., Equivalent linearization techniques. *J Acoust Soc Am.* 1963;35(11):1706-1711.
17. Atalik TS, Utku S. Stochastic linearization of multi-degree-of-freedom non-linear systems. *Earthq Eng Struct Dyn.* 1976;4(4):411-420.
18. Casciati F, Faravelli L, Hasofer A. A new philosophy for stochastic equivalent linearization. *Probabilistic Eng Mech.* 1993;8(3):179-185.
19. Yi S, Wang Z, Song J. Bivariate Gaussian mixture–based equivalent linearization method for stochastic seismic analysis of nonlinear structures. *Earthq Eng Struct Dyn.* 2018;47(3):678-696.
20. Fujimura K, Der Kiureghian A. Tail-equivalent linearization method for nonlinear random vibration. *Probabilistic Eng Mech.* 2007;22(1):63-76.
21. Alanis, AY, Arana-Daniel, N, López-Franco, C. *Artificial Neural Networks for Engineering Applications.* St. Louis, Missouri, Academic Press: Elsevier; 2019.
22. Stoffel M, Bamer F, Markert B. Artificial neural networks and intelligent finite elements in non-linear structural mechanics. *Thin-Walled Struct.* 2018;131:102-106.
23. Stoffel M, Bamer F, Markert B. Neural network based constitutive modeling of nonlinear viscoplastic structural response. *Mech Res Commun.* 2019;95:85-88.
24. Stoffel M, Gulakala R, Bamer F, Markert B. Artificial neural networks in structural dynamics: a new modular radial basis function approach vs. convolutional and feedforward topologies. *Comput Methods Appl Mech Eng.* 2020;364:112989.
25. Stoffel M, Bamer F, Markert B. Deep convolutional neural networks in structural dynamics under consideration of viscoplastic material behaviour. *Mech Res Commun.* 2020;108:103565.
26. Koepppe A, Bamer F, Markert B. An intelligent nonlinear meta element for elastoplastic continua: deep learning using a new time-distributed residual U-net architecture. *Comput Methods Appl Mech Eng.* 2020;366:113088.
27. Koepppe A, Bamer F, Markert B. An efficient Monte Carlo strategy for elasto-plastic structures based on recurrent neural networks. *Acta Mech.* 2019;230:3279-3293.
28. Heider Y, Wang K, Sun WC,. SO(3)-invariance of informed-graph-based deep neural network for anisotropic elastoplastic materials. *Comput Methods Appl Mech Eng.* 2020;363:112875.
29. Sun H, Burton HV, Huang H. Machine learning applications for building structural design and performance assessment: state-of-the-art review. *J Build Eng.* 2021;33:101816.
30. Molas G, Yamazaki F. Neural networks for quick earthquake damage estimation. *Earthq Eng Struct Dyn.* 1995;24(4):505-516.
31. Huang CS, Hung SL, Wen CM, Tu TT,. A neural network approach for structural identification and diagnosis of a building from seismic response data. *Earthq Eng Struct Dyn.* 2003;32(2):187-206.
32. Mangalathu S, Sun H, Nweke CC, Yi Z, Burton HV,. Classifying earthquake damage to buildings using machine learning. *Earthq Spectra.* 2020;36(1):183-208.
33. Hurtado JE, Alvarez DA. Neural-network-based reliability analysis: a comparative study. *Comput Methods Appl Mech Eng*. 2001;191(1):113-132.

34. Chojaczyk A, Teixeira A, Neves L, Cardoso J, Guedes Soares C. Review and application of Artificial Neural Networks models in reliability analysis of steel structures. *Struct Saf*. 2015;52:78-89.

35. Thaler D, Bamer F, Markert B. A machine learning enhanced structural response prediction strategy due to seismic excitation. *Proc Appl Math Mech*. 2020;20:e20200294.

36. Vazirizade S, Nozhati S, Zadeh M. Seismic reliability assessment of structures using artificial neural network. *J Build Eng*. 2017;11:230-235.

37. Kiani J, Camp C, Pezeshk S. On the application of machine learning techniques to derive seismic fragility curves. *Comput Struct*. 2019;218:108-122.

38. Gomes H, Awruch A. Comparison of response surface and neural network with other methods for structural reliability analysis. *Struct Saf*. 2004;26(1):49-67.

39. Burton HV. Estimating aftershock collapse vulnerability using mainshock intensity, structural response and physical damage indicators. *Struct Saf*. 2017;68:85-96.

40. Zhang Y, Burton HV, Sun H, Shokrabadi M. A machine learning framework for assessing post-earthquake structural safety. *Struct Saf*. 2018;72:1-16.

41. Zhang Y, Burton HV. Pattern recognition approach to assess the residual structural capacity of damaged tall buildings. *Struct Saf*. 2019;78:12-22.

42. Mignan A, Broccardo M. Neural network applications in earthquake prediction (1994–2019): meta-analytic and statistical insights on their limitations. *Seismol Res Lett*. 2020;91(4):2330-2342.

43. Hornik K. Approximation capabilities of multilayer feedforward networks. *Neural Netw*. 1991;4(7):251-257.

44. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: London: MIT Press; 2016. http://www.deeplearningbook.org.

45. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986;323:533-536.

46. Kingma D, Ba J. Adam: a method for stochastic optimization. In: *International Conference on Learning Representations*. 2015.

47. Kruse R, Borgelt C, Braun D, Klawonn F, Moewes C, Steinbrecher M. *Computational Intelligence, Eine methodische Einführung in Künstliche Neuronale Netze, Evolutionäre Algorithmen, Fuzzy-Systeme und Bayes-Netze*. Wiesbaden: Springer Vieweg; 2015. ISBN 978-3-658-10904-2.

48. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. https://www.tensorflow.org; 2015.

49. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving G, Isard M et al. *Python Reference Manual*. 2015.

50. Tajimi H. A statistical method of determining the maximum response of a building structure during an earthquake. In: *Proceedings of the 2nd World Conference on Earthquake Engineering*. Tokio. 1960:781-798.

51. Fan F, Ahmadi G. Nonstationary Kanai-Tajimi models for El Centro 1940 and Mexico City 1985 earthquakes. *Probabilistic Eng Mech*. 1990;5(4):171-181.

52. Jough FK, Şensoy S. A visco-elastic pounding damage formulation. *Eng Mech*. 2016;15(4):747-757.

53. Parameswaran R, Kannan A, Préfontaine J, Pande V. Application of machine learning techniques for predicting seismic fragility. *Earthq Eng Struct Dyn*. 2018;47(12):2490-2495.

54. Fan F, Ahmadi G. Nonstationary Kanai-Tajimi models for El Centro 1940 and Mexico City 1985 earthquakes. *Probabilistic Eng Mech*. 1990;5(4):171-181.

55. Jough FK, Şensoy S. A visco-elastic pounding damage formulation. *Eng Struct*. 2019;197:109373.

56. Bathe KJ. *Finite Element Procedures*. Englewood Cliffs, NJ: Prentice Hall; 1996.

57. Chopra A. *Dynamics of Structures, Theory and Applications to Earthquake Engineering*. Englewood Cliffs, NJ: Prentice Hall; 2007.

58. Morfis K, Kostinakis K. Seismic parameters’ combinations for the optimum prediction of the damage state of R/C buildings using neural networks. *Adv Eng Softw*. 2017;106:1-16.

59. Jough FK, Şensoy S. Prediction of seismic collapse risk of steel moment frame mid-rise structures by meta-heuristic algorithms. *Earthq Eng Vib*. 2016;15(4):743-757.

60. Metin Bamer F, Markert B. A nonlinear visco-elastoplastic model for structural pounding. *Earthq Eng Struct Dyn*. 2018;47(12):2490-2495.

61. Bamer F, Strubel N, Shi J, Markert B. A visco-elastoplastic pounding damage formulation. *Eng Struct*. 2019;197:109373.

62. Mitropoulou CC, Papadrakakis M. Developing fragility curves based on neural network IDA predictions. *Eng Struct*. 2011;33(12):3409-3421.

63. Bianchini SA, Dostal M, Baker JW. Prediction of inelastic structural response using an average of spectral accelerations. In: *Proceedings of the 10th International Conference on Structural Safety and Reliability (ICOSSAR 09)*. 2009.

64. Eads L, Miranda E, Lignos DG. Average spectral acceleration as an intensity measure for collapse risk assessment. *Earthq Eng Struct Dyn*. 2015;44(12):2057-2073.

65. Hintze JL, Nelson RD. Violin plots: a box plot-density trace synergism. *Am Stat*. 1998;52(2):181-184.

66. Liang X, Mosalam KM. Lyapunov stability and accuracy of direct integration algorithms applied to nonlinear dynamic problems. *J Eng Mech*. 2016;142:04016022.
69. FEMA 356. *Prestandard and Commentary for the Seismic Rehabilitation of Buildings*. Washington, DC: Federal Emergency Management Agency; 2000.

**How to cite this article:** Thaler D, Stoffel M, Markert B, Bamer F. Machine-learning-enhanced tail end prediction of structural response statistics in earthquake engineering. *Earthquake Engng Struct Dyn*. 2021;50:2098–2114. [https://doi.org/10.1002/eqe.3432](https://doi.org/10.1002/eqe.3432)