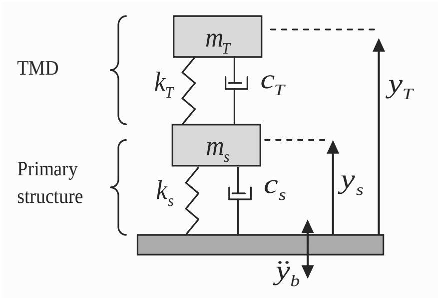

Figure 1 Basic structure of TMD system

3 The proposed method

The basic procedures of the proposed methodology is schematically described in Figure 2. We assume the uncertainties in the base structure are characterized by normal distributions. The structure failure probability in the proposed method is defined as the risk that the structure responses exceed a predefined limit within a certain interval of time [0, T].

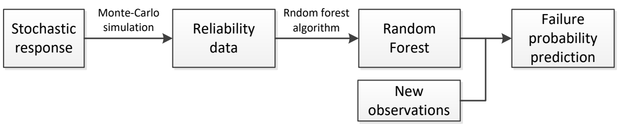

Figure 2 The schematic of the proposed method

3.1 Reliability evaluation by MCS

To apply MCS, the stochastic excitation is specified as a certain number of input random variables [6], $X=(X_1, X_2,\ldots, X_n)$ which can be simulated as a generation of the excitation. N samples X are obtained from a multi-variable Gaussian distribution. The Monte Carlo estimator of the structure failure probability is [7]

$$\hat{P}_f = \frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} I_f(\mathbf{Z}^{(i)}) \tag{3}$$

where $\mathbf{Z}^{(i)}$ is the ith sample of the excitation process; N_{mc} is the number of samples used in MCS, and $I_f(\cdot)$ $I_f(\cdot)$ satisfies $I_f(\cdot)=1$, if $y_{\max} > y_{\text{limit}}$; $I_f(\cdot)=0$, if $y_{\max} \le y_{\text{limit}}$. Here, $y_{\max} = \max(|\mathbf{y}_s|)$ is the maximum response of the base structure.

2.2 Reliability modeling and prediction by RF

A Random Forest (RF) is a meta-learner comprised of many individual trees called classification and regression tree (CART). A tree has a binary recursive structure. Its learning process is actually the node selecting and splitting process. Each split results in two subsets of the data that falls into this node so that the resulting child nodes are the "purest".

The splitting criterion is used to find the optimal split point. At node t, the best split s is chosen to maximize the goodness of split Δi(s, t).

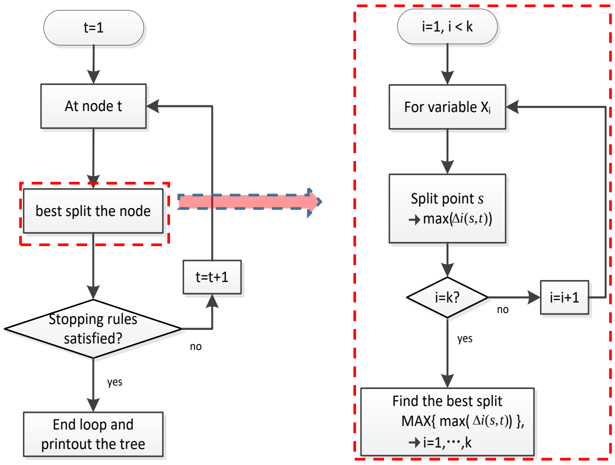

Figure 3 Tree learning process

The goodness of split corresponds to a decrease in impurity. We use Least Squares Deviation (LSD) to measure the impurity. Assume $h(t)$ is the sample set that falls into node t, a is a sample in $h(t)$, then at node t the impurity is defined as

$$i(t) = \sum_{a\in h(t)} (y_a - \bar{y}(t))^2 \tag{4}$$

where y_a is the decision value for sample a, $\bar{y}(t)$ is the average of all decision values in $h(t)$. At node t, the best splitting point s maximizes

$$\Delta i(s,t) = i(t) - i(t_L) - i(t_R) \tag{5}$$

where t_L and t_R are the left child node and right child node produced by the splitting point s. The splitting procedures should be stopped if the defined stopping rules are satisfied.

After training the RF, we determine the prediction value by averaging all predicted values of the individual trees on a new sample $\mathbf{x}'$:

$$\tilde{f} = 1/K \cdot \sum_{b=1}^{K} \tilde{f}_b(\mathbf{x}') \tag{6}$$

where K is the size of RF.

4 Numerical test

In this section, the RF performance in reliability modeling and predictions of an exemplary structure are shown and discussed. The structure parameter configurations come from Elyes [8], see Table 1. Besides, the natural frequency and damping ratio of the seismic model are $\omega_f = 25.224\ rad/s$, $\xi_f = 0.4$. The power spectral density of the white noise process is $S_0 = 0.031W / Hz$. The failure criterion is 4.3e^{-2}m.

Table 1. Parameters of the structure

Parameters	$m\,(kg)$	$c\,(N \cdot m / s)$	$k\,(N/m)$
Base structure	1	0.03	696.4
TMD	0.02	0.0695	12.725
standard dev.	0. 1	0.001	5.0

In this example, the uncertainties exist in the base structure properties. Different sizes of data of the uncertain property values are sampled according to their uncertain characteristics. The sizes of data range from 1000 cases to 19000 cases.

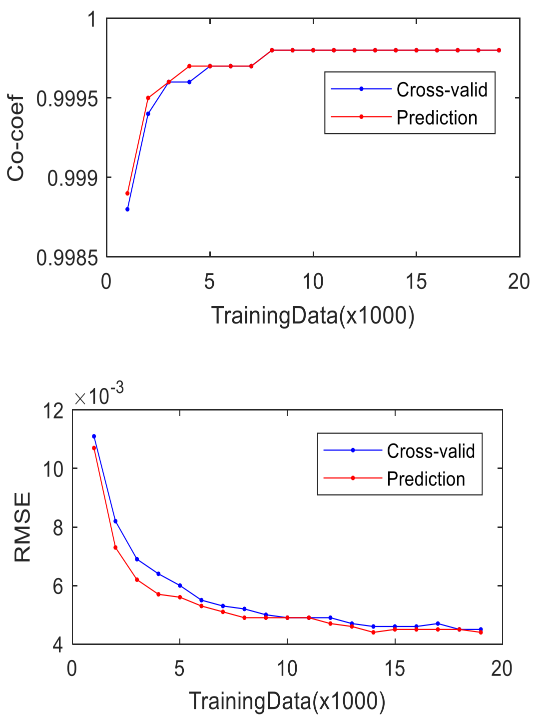

Figure 3 Simulation results based on RF model

We evaluate the model itself by cross-validation on the datasets. Another 1000 cases are sampled as a test set to study the model performance in predictions. From Figure 5, we see that with the increase of the training data, the Co-coef (correlation coefficients) increase no matter in the cross-validation or prediction procedures. This means that the RF model fits the data very well, and more importantly it doesn't overfit the data. Meanwhile, the RMSEs (root mean square error) decrease drastically, which indicates that the model becomes more accurate when more samples are available.

5 Conclusions

In this paper, the Random Forest method is investigated on reliability modeling and prediction of passive controlled structures. Results from the numerical example showed that the RF model behaves well in failure modeling and prediction of passive controlled structures. The RF model is an alternative way to study structure reliability and more attention should be paid to this approach.

References

1. Xu K, Xie M, Tang L C, et al. Appl. Soft Compu., 2003, 2(4): 255-268.
2. Das Chagas Moura M, Zio E, Lins I D, et al. Reliab. Eng. & Sys. Saf., 96, 11(2011).
3. Kutyłowska M. Eng. Failure Ana., 47(2015).
4. Nitze I, Schulthess U, Asche H. Proc. of the 4th GEOBIA (2012).
5. Raczko E, Zagajewski B. European Journal of Remote Sensing, , 50,1(2017).
6. Au S K, Beck J L. Prob. Eng. Mech., 2001, 16(3): 193-207.
7. Yu H, Gillot F, Ichchou M. Jour. of Sound and Vib., 2013, 332(9): 2222-2238.
8. Mrabet E, Guedri M, Ichchou M N, et al. Mech. Sys. and Sig. Proce., 2015, 60: 437-451.