Earthquakes and Structures, *Vol. 16, No. 6 (2019) 679-690*
DOI: https://doi.org/10.12989/eas.2019.16.6.679

Application of machine learning in optimized distribution of dampers for structural vibration control

Luyu Li* and Xuemeng Zhao[a]

School of Civil Engineering, State Key Laboratory of Coastal and Offshore Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China

(Received March 3, 2019, Revised April 2, 2019, Accepted April 3, 2019)

Abstract. This paper presents machine learning methods using Support Vector Machine (SVM) and Multilayer Perceptron (MLP) to analyze optimal damper distribution for structural vibration control. Regarding different building structures, a genetic algorithm based optimization method is used to determine optimal damper distributions that are further used as training samples. The structural features, the objective function, the number of dampers, etc. are used as input features, and the distribution of dampers is taken as an output result. In the case of a few number of damper distributions, multi-class prediction can be performed using SVM and MLP respectively. Moreover, MLP can be used for regression prediction in the case where the distribution scheme is uncountable. After suitable post-processing, good results can be obtained. Numerical results show that the proposed method can obtain the optimized damper distributions for different structures under different objective functions, which achieves better control effect than the traditional uniform distribution and greatly improves the optimization efficiency.

Keywords: damper distribution; optimization; support vector machine; multilayer perceptron; genetic algorithm; machine learning

1. Introduction

Passive energy dissipation devices are widely used in building structures to reduce the dynamic response of structures and avoid serious damage. In the previous study of energy dissipation systems, dampers were usually arranged uniformly. As the number of structural layers and the number of dampers increase, conventional distribution methods cannot meet the safety and economic requirements. In order to reduce the cost of energy dissipation system and improve energy consumption efficiency, how to allocate the number of dampers between different layers of structures to achieve the best control effect under certain economic conditions has become an important issue to be considered in arranging dampers. Under certain numbers of dampers, the global optimal solution of objective functions can only be obtained by exhaustive means. Considering the increasing number of structural layers and the number of dampers, it is impractical to exhaust many of these situations.

During the last decades, many researchers have made efforts to study on the optimal distribution of dampers. An optimal damper distribution using sequential optimization procedure based on the concept of controllability was proposed (Zhang and Soong 1992). Worst-Out-In (WOI) and Exhaustive Single Point Substitution method (ESPS) were used to locally optimize the position of the exciter (Haftka and Adelman 1985). A distribution of dampers at the position which can maximize the damping ratio of the basic vibration mode was advised (Ashour and Hanson 1987). A topology optimization method to study the position optimization of dampers came up (Natke and soong 1993). The simulated annealing method was used to search for the optimal position of dampers (Milman and Chu 1994). A method based on steepest directions search algorithm was used to obtain the optimal distribution of dampers that minimizes the system transfer function (Takewaki *et al*. 1999). A method was proposed to determine the required additional damping ratio (Kim *et al*. 2003 and 2006). With the wide application of genetic algorithms, researchers have introduced genetic algorithms into the field of structural control, and have done a lot of optimization research work (Takewaki 1997, Wu *et al*. 1997). Furuya *et al*. (1998) considered economic issues and determined a suitable damper distribution for vibration control of 40-storey buildings subject to various seismic excitations. Singh and Moreschi (2002) determined the optimal number and optimal distribution of dampers for seismic response control of a 10-storey linear building structure. Problems of optimally arranging the dampers under different optimization objectives and evaluation criteria have been proposed (Qu and Li 2008, Li and Qu 2010). It was shown that once the building frames and energy dissipation systems are modeled appropriately, the optimal quantity and placement of passive control and energy dissipation systems can be determined automatically and simultaneously (Takewaki 2009). A gradient-based evolutionary optimization methodology was presented for finding the optimal design of both dampers and their

*Corresponding author, Associate Professor
E-mail: liluyu@dlut.edu.cn
[a]Graduate Student
E-mail: zimizxm@mail.dlut.edu.cn

http://www.techno-press.com/journals/eas&subpage=7
ISSN: 2092-7614 (Print), 2092-7622 (Online)

supporting members to minimize an objective function of a linear multi-storey structure subjected to the critical ground acceleration (Fujita *et al*. 2010). An application was presented to use a robustness function for the design and evaluation of passive damper systems (Fujita and Takewaki 2012). Optimal tuned mass damper was designed for nonlinear frames by distributed genetic algorithms in 2012 (Mohebbi and Joghataie 2012). Semi-active fuzzy control of MR damper was proposed based on genetic algorithm (Huang *et al*. 2009). Hybrid genetic algorithm for optimal placement of transmission tower dampers was proposed (Guo *et al*. 2009). Adachi *et al*. proposed a nonlinear optimal oil damper design method in a multi-storey building frame (2013). Yan *et al*. (2014) proposed an optimal distribution of high-rise structural viscous dampers based on relative fitness genetic algorithm. The optimal design problem was solved through a numerical approach to a constrained optimization problem by minimizing some performance criteria that are representative measures of the system response (Domenico *et al*. 2019). A new method for optimal viscous damper placement was proposed for elastic-plastic multi-degree-of-freedom (MDOF) structures subjected to the critical double impulse as a representative of near-fault ground motions. (Akehashi and Takewaki 2019).

Recently machine learning methods are attracting widespread interest in the fields of civil engineering. Many researchers have conducted extensive and in-depth research in the fields of damage identification, pattern recognition, response prediction, image processing, reliability evaluation, program optimization and cost estimation. In this paper, Support Vector Machine (SVM) and Multilayer Perceptron (MLP) are used to predict the optimization scheme based on genetic algorithm, and quickly obtain optimization results for a large number of different structures under different objective functions in a short time, which greatly improves optimization efficiency. The method can solve the following problems: when studying the optimization of damper distribution, the objective functions proposed by many scholars may be different so that the obtained optimization results are also different, which may not give a generalized reference results for other structures. Moreover, in practical applications genetic algorithms are prone to premature convergence (Li *et al*. 2010). With the increase of the number of building layers and the number of dampers, in order to avoid falling into the local optimum problem, the number of initial populations cannot be selected below a certain standard, which leads to an upper limit on the optimization efficiency. Therefore, it is difficult to optimally arrange a large number of different structures efficiently. Using machine learning methods, a large number of damper optimization distributions obtained by different structures under different optimization targets can be used as training samples to train a learning model in advance. Then the trained model can quickly and efficiently generate a large number of feasible optimization schemes for different structures and target conditions according to input structural features and optimization target features, which will greatly improve optimization efficiency and achieve good results.

Considering the problem as a discrete optimal damper distribution problem, this paper makes multi-class prediction and multi-output regression prediction separately. Genetic algorithm is selected as the optimization method and the corresponding optimization results are used as samples. In the first example, classifiers are built using SVM and MLP, respectively. For a three-layer structure, the category of damper distribution is predicted by changing the weighting index of objective function. In the second example, MLP is used to establish a regression model. For different structures, the number of dampers in each layer is analyzed under different cases, which is regarded as a multi-output regression problem. The paper is organized as follows: Part 2 simplifies engineering problems into mathematical models; Part 3 introduces sample optimization methods, optimization objective functions, and optimization models for making sample sets; Part 4 is a brief introduction to the principles of SVM and MLP; Part 5 introduces numerical simulation and analysis of the proposed method; Part 6 is conclusions of this paper.

2. Mathematical model

Considering the damping force of dampers, the equation of motion of structure can be expressed as Eq. (1)

$$[M][\ddot{u}(t)]+([C]+[C_d])[\dot{u}(t)]+[K][u(t)]=-[M][I]\ddot{u}_g(t) \tag{1}$$

where [M], [C], [K] represent the mass matrix, damping matrix, and stiffness matrix of the structure, respectively; $[\ddot{u}(t)],[\dot{u}(t)]$, $[u(t)]$ are the acceleration vectors, velocity vectors, and displacement vectors of the structure, respectively; $\ddot{u}_g(t)$ is the ground acceleration; [C_d] is an additional damping matrix for the dampers. The equation will be solved by *Newmark-β* method in this paper.

3. Optimization process

3.1 Damper optimization objective functions

The objective function considering the maximal storey drift angle, the maximal acceleration and the maximal displacement of each layer was proposed (Qu and Li 2008) as Eq. (2). Using a damper-controlled structure under seismic load, the structural vibration control target can be satisfied in terms of safety and comfort by the linear combination of the maximum value of the storey drift angle, the maximum value of the absolute acceleration and the maximum value of the absolute displacement with and without control.

$$Z=min(\alpha\frac{\theta_{max}}{\theta_{0,max}}+\beta\frac{a_{max}}{a_{0,max}}+\gamma\frac{u_{max}}{u_{0,max}}) \tag{2}$$

where θ, α, u represent the storey drift angle, absolute acceleration, and absolute displacement of structure under ground motion; subscript *max* indicates the maximum value of controlled structure response; subscript *0*,max indicates the maximum value of uncontrolled structural response. α,

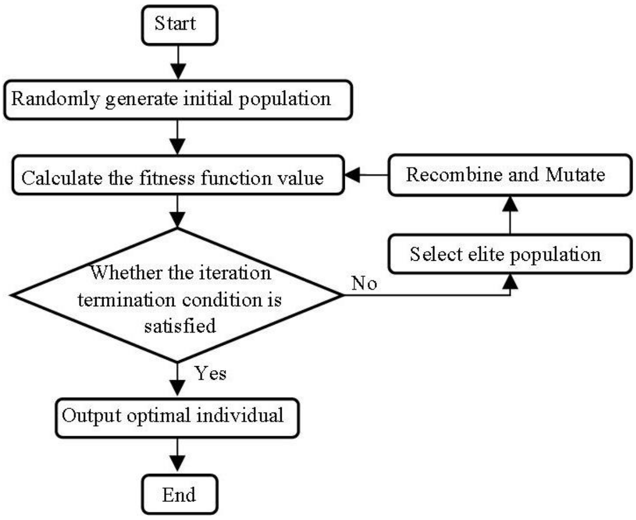


Fig. 1 Flow chart for genetic algorithm

β, γ are weighting factors whose sum is 1. Different combinations of coefficients are given depending on different requirements for safety and comfort in engineering applications.

3.2 Optimization model

The mathematical expression of the damper optimal distribution problem in the structure can be written as Eq. (3)

$$\begin{cases} \min Z(X,t) \\ X=[x_1,x_2,\ldots\ldots x_n]^T \\ \sum x_j \le N \\ \text{s.t.}: x_j \in \{0,1,2,\ldots\ldots N\},\ (j=1,2,\ldots\ldots n) \end{cases} \tag{3}$$

where n is the number of structural layers; N is the upper limit of total dampers number. The upper limit of dampers number per layer is consistent with the upper limit of total dampers number. X is a vector representing damper distribution; x_j is the integer number of dampers on the floor j; Z is the objective function of the optimization problem. Different objective functions can be combined according to different weight coefficients; t is the time of dynamic load.

3.3 Genetic algorithm

Genetic algorithm is a parallel random search optimization method developed by Professor Holland of the University of Michigan in 1962 with simulating the natural genetic mechanism and biological evolution theory. The genetic algorithm first encodes the optimization problem and generates the initial population. For the damper optimization problem, each individual in the population uses integer coding, which represents the damper distribution. Through the calculation of the evaluation function of each generation, the selection, hybridization and variation of the population are carried out. Then the better individuals in the current generation are selected to generate new populations. After multiple generations of calculations, the algorithm finally converges to the optimal solution individual, where the optimal solution represents the damper distribution that optimizes the objective function. The following Fig. 1 shows the flow chart of genetic algorithm:

4. Methodology of machine learning

4.1 Classification and regression

For the damper distribution scheme with low building structures and few optimization schemes, it can be regarded as a classification problem. Through the existing training samples that are the known data and its corresponding outputs, it is trained to obtain a satisfying classifier model which maps all the inputs to the corresponding outputs and makes a simple judgment on the output to achieve the purpose of classification. The classifier has the ability to classify unknown data. However, for the case of more structural layers and more dampers, it is not easy to classify the damper distribution scheme. In this case, such problem is treated as regression problems and each regression model have multiple outputs. After data processing, each multidimensional output represents the number of dampers required for the corresponding number of layers. Unlike classification, regression problems map all inputs to corresponding outputs that are continuous variables. On the other hand, classification is a prediction problem that the output variable is a finite number of discrete variables.

4.2 Support vector machine

Support Vector Machine (SVM) was originally proposed by Vapnik (1963, 1964) and Chervonenkis (1964). SVM can be used for pattern classification and regression, which is considered to be the most successful algorithm before machine learning. The main idea of SVM is to map input vectors to higher dimensional space through non-linear transformation for non-linear separable samples to make them linear separable. A classification hyperplane is created as a decision surface in this high dimensional space making the isolated margin between the positive and negative examples maximized.

A simple two-dimensional vector model is shown in Fig. 2. The separation hyperplane is defined as Eq. (4)

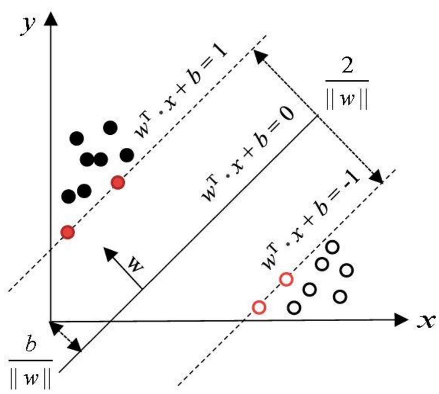


Fig. 2 Optimal hyperplane in the case of linear separability

$$w^T x + b = 0 \tag{4}$$

where x represents data points, w is an adjustable weight vector, b is a bias that represent offset of the hyperplane relative to the origin. T represents the transpose of a vector. The model of SVM is to keep all points to the hyperplane at a certain distance (the function distance=1 in the Fig. 2). Then all the classification points should be on both sides of the support vector of respective category. Positive examples above the hyperplane is defined with labels y=1, while negative examples below the hyperplane is defined with labels y=-1. Obviously there is more than one hyperplane that satisfies this condition. It is designed to find the one with the strongest generalization ability, so that the sum of the distances of all classified points to the hyperplane is the smallest. It has been shown that in this model, the geometric distance is the true distance from the point to the hyperplane. After mathematical derivation, maximizing the separation edge between the two classes is equivalent to minimizing the Euclidean norm of the weight vector w. The optimization model can be expressed as Eq. (5). Because it is a convex function, according to the convex optimization problem theory, the maximum marginal hyperplane can be obtained by using the Karush-Kuhn-Tucher (KKT) condition and the Lagrange multiplier. Supposing there are a total of m points, it can be written as Eq. (5)

$$\begin{cases} \max \dfrac{2}{\|w\|} \Leftrightarrow \min \dfrac{1}{2}\|w\|^2 \\ \text{s.t. } y_i(w^T x_i + b) \geq 1 \quad (\text{i}=1,2,\ldots m) \end{cases} \tag{5}$$

In the previous discussion, it is assumed that the data set is linearly separable. However, for actual application, there may not be the hyperplane that perfectly separates the data sets. In this case, the original space can be mapped to a high-dimensional space. If the data set in the high-dimensional space is linearly separable, then the problem can be solved. In this way, the hyperplane becomes as Eq. (6):

$$w^T \phi(x) + b = 0 \tag{6}$$

where $\phi(x)$ represents the nonlinear mapping of raw data from low-dimensional space to high-dimensional space. This method is called “kernel method” which is realized by kernel functions. Commonly used kernel functions are: linear kernel, polynomial kernel, Gaussian kernel, Laplacian kernel, sigmoid kernel, etc.

4.3 Multilayer perceptron

Multilayer Perceptron (MLP) is inspired by the biological nervous system and usually consists of an input layer, one or more hidden layers and an output layer. MLP can be thought of as a logistic regression classifier in which the input layer is transformed by non-linear transformations that are constantly “learned”. George Cybenko (1989) proposed that a single hidden layer can fit any nonlinear continuous function well as long as the number of nodes is sufficient, making MLP a universal estimator. In 2006, GE Hinton and RR Salakhutdinov (2006) proposed that with the enhancement of computer computing power and the sharp increase of data sets, multiple hidden layers began to be widely used. Multiple hidden layers are composed of multiple single hidden layers, compared with single hidden layers. The multi-hidden layer has strong generalization ability and high prediction accuracy, but the calculation amount is large and the training time is longer. The use of multiple hidden layers proves that multiple hidden layers have a greater advantage, which is called deep learning.

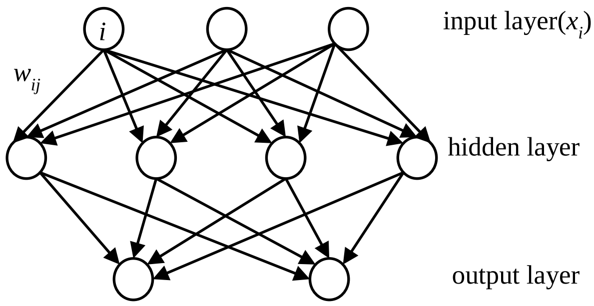


Fig. 3 Sample MLP configuration

Fig. 3 is a single hidden layer MLP structure, in which each layer consists of multiple cells (also called neural nodes) and one circle represents a cell. Each node in the hidden or output layer is fully connected to all nodes in the previous layer, and these connections represent a weighted combination with the nodes in the previous layer. The instance feature vector of the data set starts from being input into the input layer. Then the data are passed to the next layer. The output of each layer through a specific calculation is used as the input of the next layer. The output value of node j shown in Fig. 3 is calculated by the following Eqs. (7)-(8):

$$Z_j = \sum (w_{ij} x_i + b) \tag{7}$$

$$y_j = f(z_j) \tag{8}$$

where x_i represents all inputs; w_{ij} represents the weight between the node i and the node j; b represents the deviation term; Z_j represents the temporary value; y_j represents the node j output value, which is calculated by Z_j through the activation function f. The activation function f represents a nonlinear sigmoid function, enabling the multilayer perceptron to handle research problems in different fields. It has been demonstrated that MLP is capable of approximating any linear or nonlinear function by providing appropriate constraints (Wu and Jahanshahi 2018).

To build an MLP model, the model requires a series of training data to learn the relationship between input and output data. The MLP training usually aims to minimize the error between the predicted value and the true value. The bidirectional iterative update of the weight w_{ij} is performed by the standard Levenberg-Marquardt backpropagation algorithm.

5. Structural model

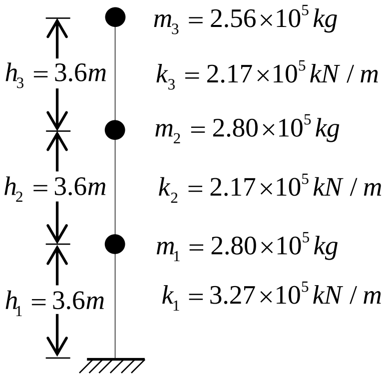


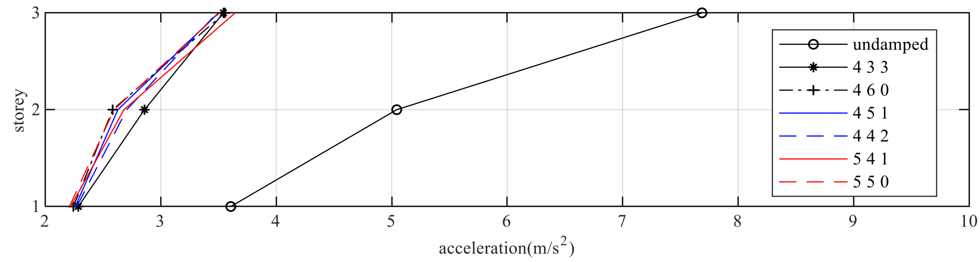
Fig. 4 Three-layer structure feature diagram

5.1 Classification prediction of damper distribution for 3-layer structure

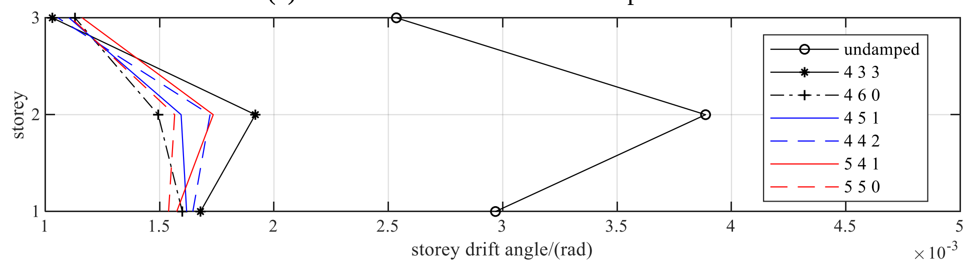
In the first example, a 3-layer reinforced concrete frame structure is selected as the structural model. The structural characteristics are shown in Fig. 4. The damping ratio is 0.05 and the structural period is 0.4479s. The El Centro seismic wave of the Imperial Valley earthquake on May 18, 1940 is selected as the ground excitation. The damping coefficient of the damper is 2.1×10^3 kN·s/m. Using Eq. (2) as the objective function, the optimal distribution will be calculated under the condition that the total number of dampers N is 10.

For low-rise buildings, when the structural properties such as mass, stiffness, and height are determined, the number of dampers distribution is countable. A label can be added to each scheme, which is regarded as a multi-class problem. The feature is also just a three-dimensional vector representing three weight parameters. In this paper SVM and MLP will be used for multi-classification learning. Regarding the number of samples, there are basic quantitative requirements in the training process, because the computer needs to obtain the rules from a large number of samples to establish a function model. Too few samples will result in a poor prediction and meaningless results. Therefore, the step size of three weight parameters is reduced to increase the number of samples, and the following schemes are respectively adopted: (1) if the step size is 0.05, 230 optimized sample data can be obtained; (2) if the step size is 0.02, 1324 optimized sample data can be obtained. Adding an unrepeated portion of the data set with step size of 0.05, a total of 1488 optimized sample data is available; (3) if the step size is 0.01, 5146 optimized sample data can be obtained.

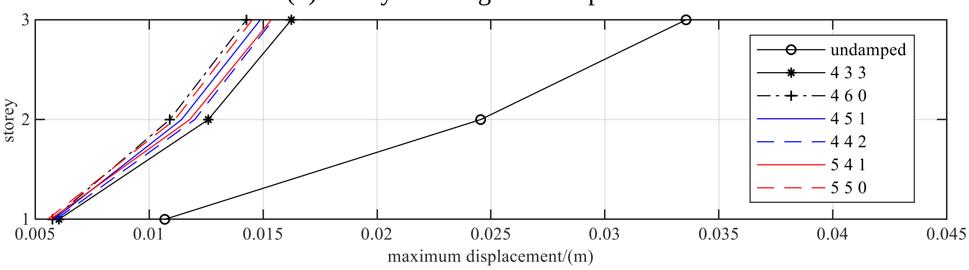
The above cases are optimized using GA. Individuals in the initial population consist of integers. Each individual represents a structural distribution scheme and each integer in individual represents the number of the dampers of corresponding layer. The population size is 100, the number of elites is 10, and the proportion of cross-generations is 0.75. The optimization is completed to obtain the damper distribution scheme, and it is found that there are only five



(a) Maximum acceleration envelope curve



(b) Storey drift angle envelope curve



(c) Maximum displacement envelope curve

Fig. 5 Seismic response of a three-layer structure

Table 1 Damper distribution scheme

Combination	Number of dampers on 1st floor	Number of dampers on 2nd floor	Number of dampers on 3rd floor	Label
1	4	6	0	1
2	4	5	1	2
3	4	4	2	3
4	5	4	1	4
5	5	5	0	5

Table 2 Prediction accuracy of SVM classifier

Sample number	Number of samples	Accuracy of test 1 (%)	Accuracy of test 2 (%)	Accuracy of test 3 (%)	Accuracy of test 4 (%)	Accuracy of test 5 (%)
1	230	94.2029	95.6522	94.2029	97.1014	89.8551
2	1488	98.2103	97.9866	98.434	97.5391	97.9866
3	5146	98.7047	98.5751	98.7694	98.899	98.9637

distribution schemes under the combination of all parameters, as shown in Table 1. Fig. 5 shows the seismic response of the above three-layer structure obtained in the uncontrolled case, uniform distribution case which is 4-3-3 and optimized cases.

5.1.1 Support vector machine classifier

Regarding the SVM classifier used in this paper, the kernel function type uses a linear kernel function and the allowable termination threshold is 0.001. For the sample schemes of three sizes, after the mixed operation, 70% is taken as the training set, and the other 30% is used as the test set. 5 trainings and tests are performed. The test results after the training are shown in Table 2. Every trained SVM model can perform a good classification. As the number of samples increases, SVM becomes more stable.

5.1.2 Multilayer perceptron classifier

Regarding the MLP classifier used in this paper, two hidden layers are selected and the number of hidden layer nodes is [20 5]. The activation functions of the two hidden layers are linear function and logarithmic sigmoid function. For the three-scale sample sets, in order to accelerate the learning process and obtain good learning effects before training, the feature vector is usually standardized before it is passed into the input layer, so that the value is between [-1,1]. After the mixed operation, 70% is taken as the training set, and the other 30% is used as the test set. 5 trainings and tests are performed. The test results after training are as shown in Table 3.

It can be seen that with the increase of the number of samples, the prediction accuracy of MLP model is more and more stable. Both MLP and the SVM can accurately and efficiently identify the damper distribution.

5.2 Regression prediction of damper distribution for 3/5/10-layer structure

In the following example, except for the weights of the objective function, structural properties are also used as the training feature. These structural properties could include

Table 3 Prediction accuracy of MLP classifier

Sample number	Number of samples	Accuracy of test 1 (%)	Accuracy of test 2 (%)	Accuracy of test 3 (%)	Accuracy of test 4 (%)	Accuracy of test 5 (%)
1	230	94.2029	95.6522	98.5507	94.2029	100.000
2	1488	98.6547	97.5336	97.5336	97.0852	98.2063
3	5146	98.5094	98.9631	98.6390	99.2871	98.5094

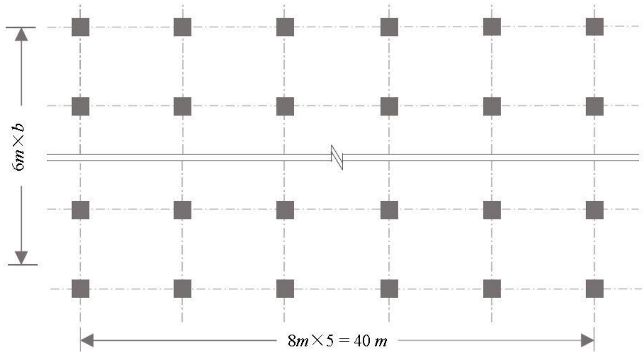


Fig. 6 Layout plan (the number b of spans=[3,4,5])

structural layer numbers, structural mass and structural stiffness. In addition, the number of dampers is taken into account as an indicator of economic factors. The mass and stiffness of each layer can be changed by adjusting the span and height of structure. Considering that the first natural frequency of structure can partly reflect the relationship between the mass matrix [M] and the stiffness matrix [K], three features which are structure span, layer height and the first natural frequency are used to reflect the characteristics of structural properties. Considering the other two characteristics-the number of structural layers and the number of dampers, it can be seen that with the different requirements of the number of structural layers and the number of dampers, the distribution of dampers could have more optimized results. Therefore, the problem must be regarded as a multi-output regression problem instead of a classification problem. The input vector has eight dimensions: the number of structural layers n, the total number of dampers N, the number of spans b, the layer height h, the first weight coefficient of the objective function α, the second weight coefficient of the objective function β, the third weight coefficient of the objective function γ, the first natural frequency of structure ω. In this example, total number of dampers will be 10, 20, and 30, respectively. The number of structure layers will be 3, 5, and 10, respectively. The number of spans will be taken as 3, 4, and 5, respectively. The layer heights will be taken to be 3 m. 3.3 m and 3.6 m, respectively. The output vector sets have ten dimensions and each dimension represents the number of dampers required for the corresponding layer. For the 3-layer structure, the last 7 dimensions of the output vector are set to 0. Similarly, for the 5-layer structure, the last 5 dimensions of the output vector are set to 0. For a 10-layer structure, each dimension may not be 0, which means that layer may be assigned to a damper. In order to eliminate the influence of other factors, the values for other parameters of the structure are shown in Fig. 6 and Fig. 7.

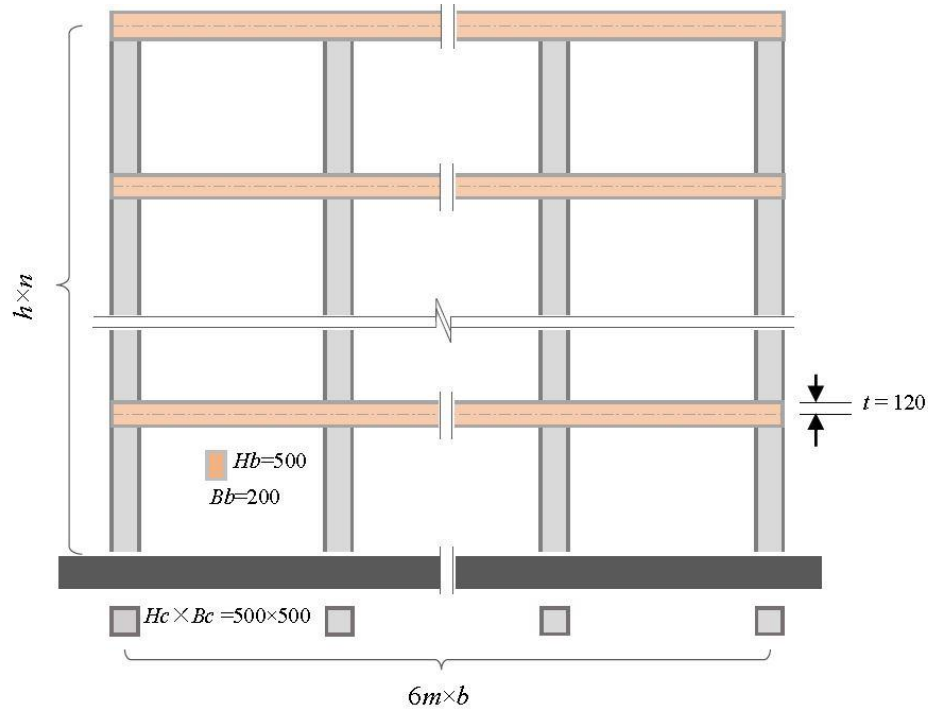


Fig. 7 Vertical Plan (layer height h=[3 m, 3.3 m, 3.6 m]; number of structural layers n=[3,5,10])

Table 4 10 sample inputs from the training set

Sample number	Number of layers	Total number of dampers	number of spans	layer height	α	β	γ	ω
1	5	20	3	3.3	0.3	0.5	0.2	9.562
2	3	10	4	3.3	0	1	0	15.588
3	10	10	5	3	0.4	0.5	0.1	5.408
4	3	10	3	3.3	0.4	0.4	0.2	15.696
5	3	30	4	3.3	0.3	0.7	0	15.588
6	5	10	4	3.3	0	0.3	0.7	9.528
7	3	20	3	3.6	0.4	0.6	0	14.028
8	5	10	4	3	0.6	0	0.4	10.729
9	10	30	5	3.6	0.1	0.3	0.6	4.325
10	10	30	5	3.3	0.1	0.6	0.3	4.814

Table 5 10 sample outputs from the training set

Sample number	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
1	5	8	5	2	0	0	0	0	0	0
2	4	6	0	0	0	0	0	0	0	0
3	0	4	1	0	1	1	1	2	0	0
4	4	6	0	0	0	0	0	0	0	0
5	14	11	5	0	0	0	0	0	0	0
6	1	5	3	1	0	0	0	0	0	0
7	10	8	2	0	0	0	0	0	0	0
8	0	9	1	0	0	0	0	0	0	0
9	4	11	5	2	1	3	4	0	0	0
10	0	7	0	0	4	0	6	7	6	0

Table 6 Sample inputs from the test set

Sample number	Number of layers	Total number of dampers	number of spans	layer height	α	β	γ	ω
1	5	10	4	3.6	0.1	0.4	0.5	8.541
2	10	30	5	3	0.4	0.2	0.4	5.408
3	3	10	4	3.6	0.3	0.5	0.2	13.930
4	3	30	3	3.6	0.1	0.6	0.3	14.028
5	10	30	4	3	0.1	0.8	0.1	5.411
6	5	20	3	3.3	0.3	0.3	0.4	9.562

Concrete material density is 2700 kg/m^3. Column concrete elastic modulus is 3.5×10^{10} Pa. Beam concrete elastic modulus is 3.0×10^{10} Pa and the damping ratio is 0.05. The El Centro seismic wave of the Imperial Valley earthquake on May 18, 1940 is selected as the ground excitation. The damping coefficient of the damper is 2.1×10^3 kN·s/m. Firstly, genetic algorithm is used to obtain the optimized samples. The population size is 200, the number of elites is 20, and the ratio of cross-generation is 0.75. Considering all possibilities and combinations, totally 5346 dampers are selected. After obtaining the optimized sample and performing the mixed operation, 80% is taken as the training set and the other 20% is used as the test set. 10 sample inputs from the training set is selected in Table 4. The sample outputs corresponding to the inputs in training set are shown in Table 5. The underlined zeros “0” indicate that there is no corresponding damper in that floor. Fig. 8 shows the objective function values of different cases.

Before training, in order to speed up the learning process and obtain good learning results, the input vector and output vector are usually normalized first so that the value is between [-1,1]. Then MLP is used to establish the regression model. Before training, some parameters are defined. The number of hidden layers is 2, the number of nodes is [40 30]. 6 sample inputs for the test are selected

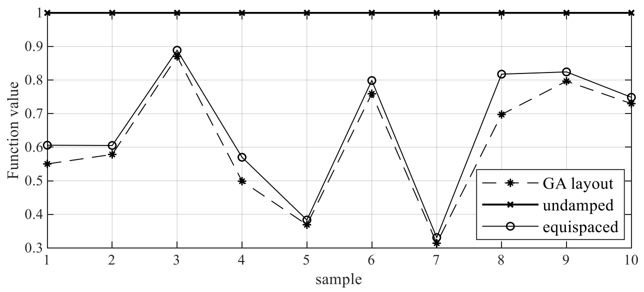


Fig. 8 Objective function value of the ten samples of different cases

Table 7 Outputs predicted by MLP

Sample number	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
1	2.81	3.77	2.21	1.25	0.05	0.02	-0.04	0.01	-0.01	0.00
2	1.21	12.42	6.55	3.84	3.01	1.58	0.78	0.37	-0.17	0.00
3	4.35	5.80	-0.06	0.08	0.02	0.02	0.03	-0.02	0.00	0.00
4	14.56	11.57	3.74	0.15	-0.22	-0.01	-0.05	0.00	0.16	0.00
5	2.34	6.06	4.49	4.82	4.03	2.77	2.24	2.30	0.81	0.00
6	4.86	8.45	5.11	1.57	0.07	0.02	0.03	-0.03	0.00	0.00

from the test set as shown in Table 6. The output of the corresponding prediction is shown in Table 7. The underlined numbers indicate that the corresponding floor does not have dampers.

The values obtained by rounding the prediction values of MLP and the corresponding genetic algorithm optimization values are compared as in Table 8. All damper distributions predicted by MLP in the following are the corresponding results after rounding.

It can be observed that for the results of MLP regression prediction, the total number of dampers in some distribution schemes is not exactly the same as the actual number of optimizations. Some post-processing is performed by reducing the number of dampers in the floors which have excessive dampers and increase the number of dampers in the floors which have less dampers than the required quantity. For schemes where the number of predictions is more than the number of optimizations of genetic algorithm, the number of dampers is reduced, because it will make the optimization objective function larger. Therefore, the principle is to reduce the number of dampers of the floors that have less influence on the objective function. For schemes where the number of predictions is

Table 8 Damper distribution optimized by genetic algorithm and damper distribution by MLP prediction

Sample number	Damper Distribution	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	sum	function value
1	GA	3	4	2	1	0	—	—	—	—	—	10	0.7567
	MLP	3	4	2	1	0	—	—	—	—	—	10	0.7567
2	GA	1	12	6	4	3	2	1	1	0	0	30	0.7440
	MLP	1	12	7	4	3	2	1	0	0	0	30	0.7443
3	GA	4	6	0	—	—	—	—	—	—	—	10	0.4856
	MLP	4	6	0	—	—	—	—	—	—	—	10	0.4856
4	GA	15	11	4	—	—	—	—	—	—	—	30	0.2722
	MLP	15	12	4	—	—	—	—	—	—	—	31	0.2687
5	GA	3	5	4	5	4	3	2	2	2	0	30	0.6828
	MLP	2	6	4	5	4	3	2	2	1	0	29	0.6891
6	GA	5	8	5	2	0	—	—	—	—	—	20	0.5435
	MLP	5	8	5	2	0	—	—	—	—	—	20	0.5435

Table 9 Damper distribution comparison of genetic algorithm, MLP and post-processing

Sample number	Damper Distribution	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	sum	function value
4	GA	15	11	4	—	—	—	—	—	—	—	30	0.2722
	MLP	15	12	4	—	—	—	—	—	—	—	31	0.2687
	Post-process	15	12	3	—	—	—	—	—	—	—	30	0.2727
5	GA	3	5	4	5	4	3	2	2	2	0	30	0.6828
	MLP	2	6	4	5	4	3	2	2	1	0	29	0.6891
	Post-process	2	6	4	5	4	4	2	2	1	0	30	0.6841

less than the number of optimizations of genetic algorithm, the number of dampers is increased. By the similar analysis, the principle is to increase the number of dampers of the floors that have a greater influence on the objective function. In this way, a comparison between the prediction scheme and the genetic algorithm optimization scheme is obtained when the total number of dampers is consistent. For the above six forecasting scheme samples, 1, 2, 3, and 6 are rounded up and the total number of dampers is equal to the quantity requirement, and no post-processing is required. Since the prediction of the 4th sample is more than the optimization requirement, more dampers need to be removed. On the other hand, the prediction of the 5^{th} sample is less than the optimization requirement, therefore the missing damper needs to be added. The post-processing results are shown in Table 9. The post-processing process is

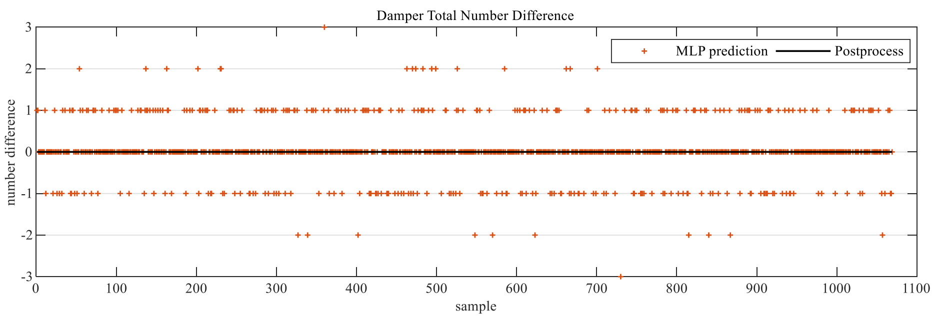


Fig. 9 Damper number difference of MLP and Post-processing relative to GA's optimization

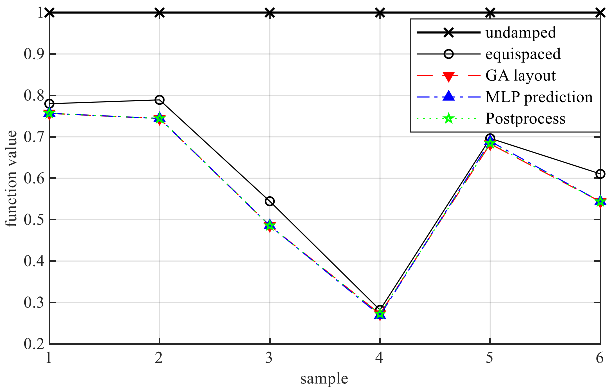


Fig. 10 The objective function value of different cases

different from the exhaustive method of the damper scheme. It is only a small number adjustment based on the MLP prediction, and the required time cost can be neglected. Fig. 9 shows the total number of dampers relative to GA's optimization of each set sample obtained by MLP prediction before and after post- processing. Fig. 10 shows objective function value of the 6 samples in different cases.

Figs. 11-16 show the seismic response (maximum value of the absolute acceleration, maximum value of the storey drift angle, maximum value of the absolute displacement) of the structure in samples 1~6 in the test set under earthquake excitation of different cases. In samples 1, 3 and 6, the MLP-predicted damper distributions are consistent with the genetic algorithm optimization results, and the response curve obtained according to the genetic algorithm optimization scheme represents the multi-layer perceptron and post-processing response. In sample 2, the damper distribution scheme predicted by MLP is inconsistent with the genetic algorithm optimization result. However, the total number of dampers in both cases is the same. Therefore, there is no need to post-process the prediction results of MLP. The response is the same for the MLP prediction and post-processing. In sample 4 and sample 5, the responses are different in five cases.

It can be observed that even under the requirements of different objective function weight values, the damper distributions obtained by genetic algorithm optimization have better control effect than the uniform distribution from the three indicators. After training, the MLP prediction scheme can approximate the results of genetic algorithm in three seismic responses. More importantly, the MLP prediction can consider different weight combinations of the objective function. Fig. 17 shows the objective function of all test set samples in different cases.

Fig. 18 shows the relative errors between the objective

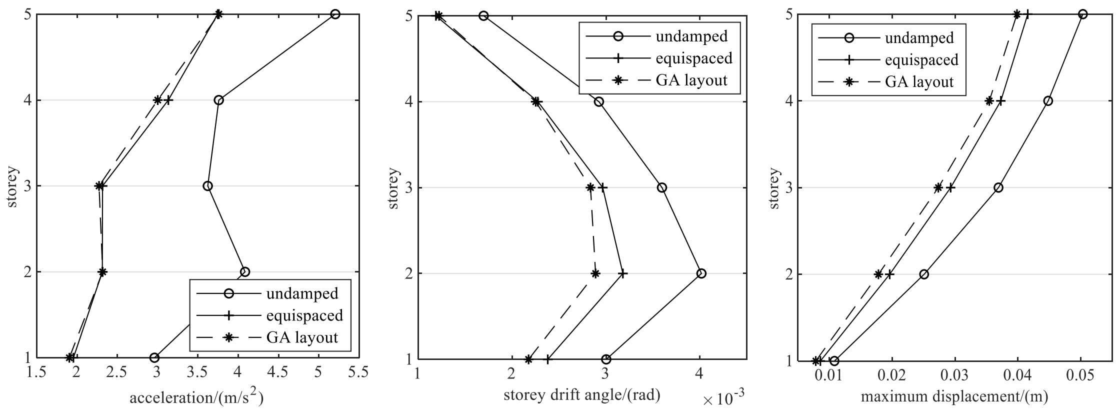


Fig. 11 Structural response of sample 1 selected from test data

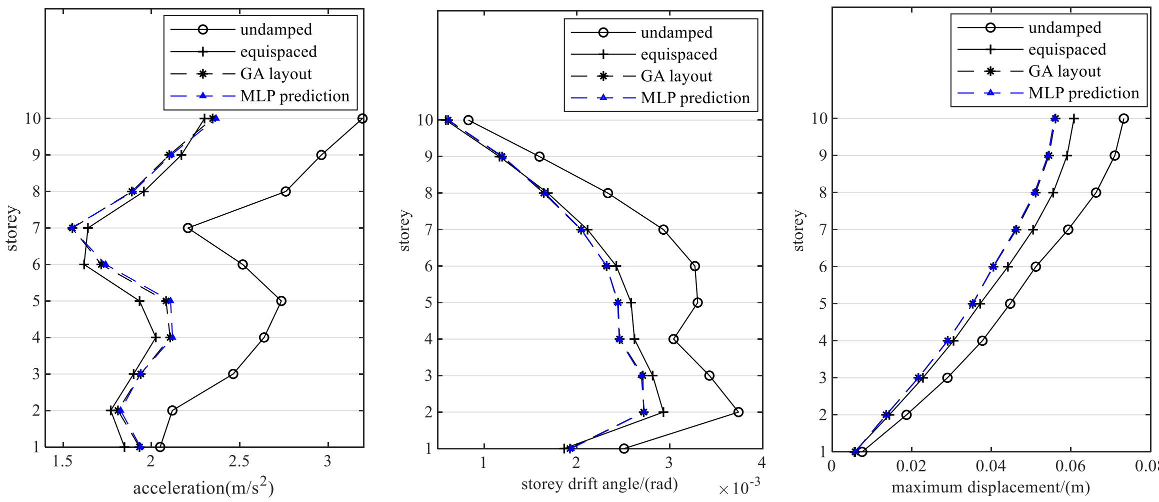


Fig. 12 Structural response of sample 2 selected from test data

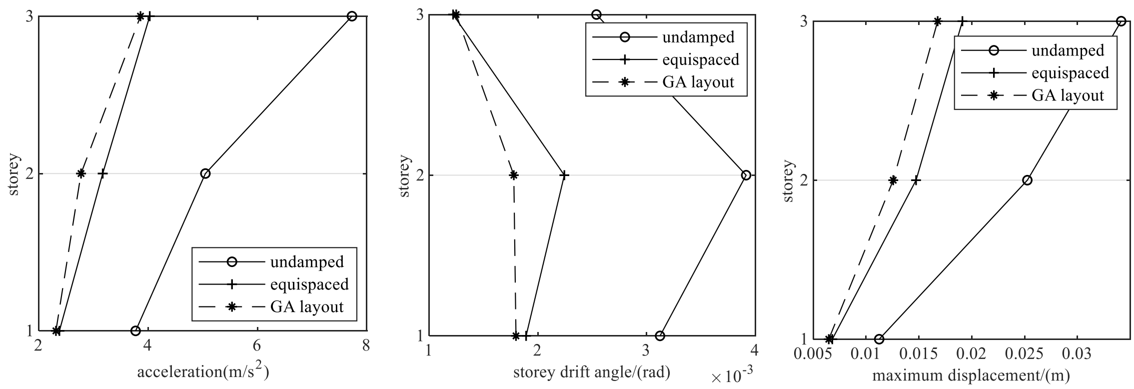


Fig. 13 Structural response of sample 3 selected from test data

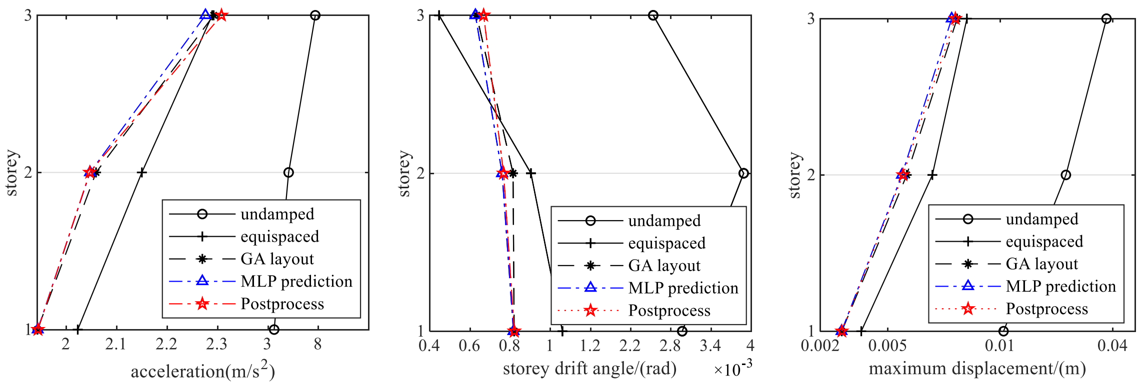


Fig. 14 Structural response of sample 4 selected from test data

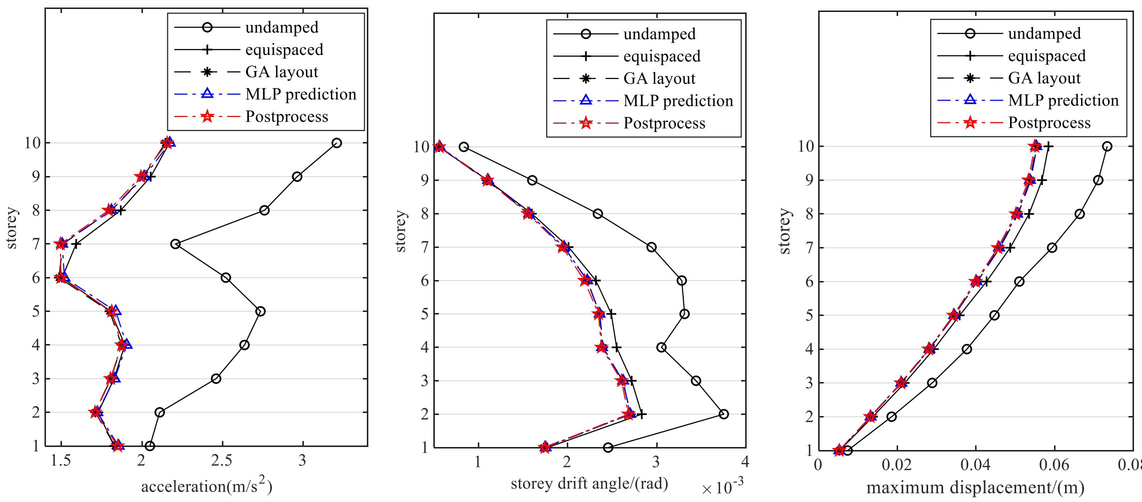


Fig. 15 Structural response of sample 5 selected from test data

function obtained before and after the post-processing of the MLP prediction scheme and the results obtained by GA optimization. It can be seen that the relative error between MLP prediction results and genetic algorithm optimization results are very small, and the relative error of MLP model in this training is less than 5%. After the post-processing of unifying total numbers of dampers got from MLP and GA, the relative errors are generally reduced further and the objective function values are closer to the function value obtained by genetic algorithm optimization. In very few cases, the relative error of post-processing is negative, which should not occur, because it means that a better result is predicted than genetic algorithm and the optimization results of some samples are not global optimal solutions.

6. Conclusions

By extracting the structural properties and the weights of the objective function as training features, the optimal distribution of dampers can be predicted by machine learning method. This paper proves the feasibility of this idea. When there are fewer schemes for optimal distribution of dampers, it can be considered as a multi-class problem to

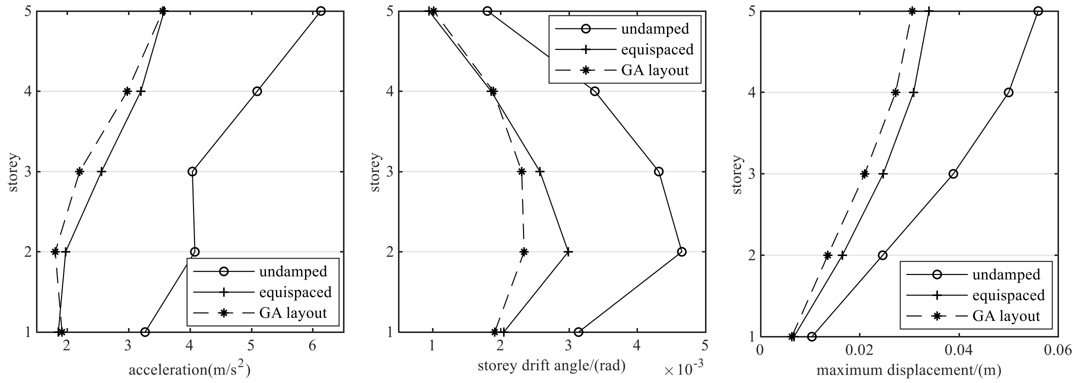


Fig. 16 Structural response of sample 6 selected from test data

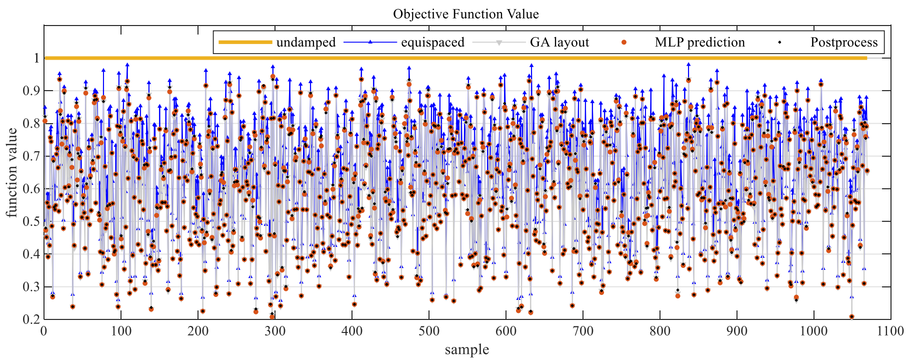


Fig. 17 Objective function of all test set samples of different cases

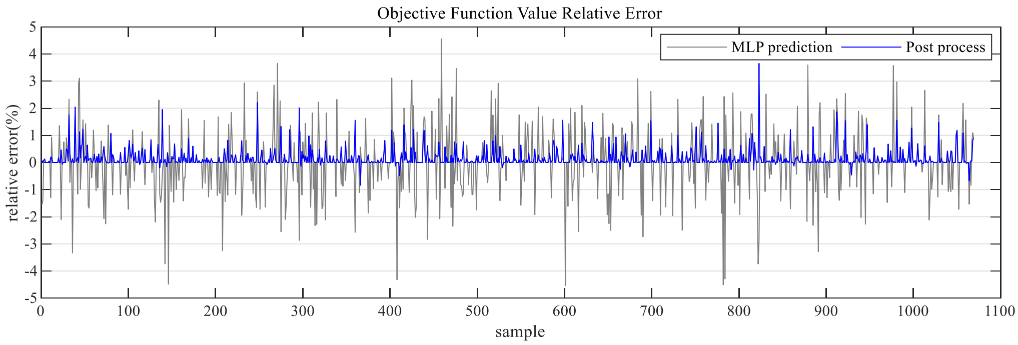


Fig. 18 Relative error relative to GA optimization objective function

solve. Both SVM and MLP classifiers perform well and can predict the optimal results consistent with GA with very high accuracy. As the number of samples increases, the accuracy of the prediction results will become better. With the increase of the number of dampers and structural layers, it is impossible to know all distribution schemes in advance. In this case, it can be regarded as a multi-output regression problem that can be solved by using MLP to predict the number of dampers in each layer. Numerical results show that damper distribution schemes predicted by MLP are very similar to the optimized results obtained from GA. Furthermore, the error of the objective functions can be further reduced by post-processing of unified total damper numbers. From the simulation, it is shown that the relative errors of prediction by machine learning are rather small and most of them are less than 5%. Therefore, the proposed machine learning method can be effectively used to predict the optimized distribution of dampers.

Acknowledgments

This work was supported by the National Science Foundation of China under Award No. 51678116

References

Adachi, F., Yoshitomi, S., Tsuji, M. and Takewaki, I. (2013), "Nonlinear optimal oil damper design in seismically controlled multi-storey building frame", *Soil Dyn. Earthq. Eng.*, **44**(1), 1-13. https://doi.org/10.1016/j.soildyn.2012.08.010.

Akehashi, H. and Takewaki, I. (2019), "Optimal viscous damper placement for elastic-plastic MDOF structures under critical double impulse", *Front. Built Environ.*, **5**, 20. https://doi.org/10.3389/fbuil.2019.00020

Ashour, S.A. (1987), "Elastic seismic response of building with supplemental damping", Univ. of Michigan Ann Arbor, Mi.

Cybenko, G. (1989), "Approximation by superpositions of a sigmoidal function", *Math. Control, Signal. Syst.*, **2**(4), 303-314. https://doi.org/10.1007/BF02551274.

De Domenico, D., Ricciardi, G. and Takewaki, I. (2019), "Design strategies of viscous dampers for seismic protection of building structures: a review", *Soil Dyn. Earthq. Eng.*, **118**, 144-165. https://doi.org/10.1016/j.soildyn.2018.12.024.

Fujita, K. and Takewaki, I. (2012), "Robust passive damper design for building structures under uncertain structural parameter environments", *Earthq. Struct.*, **3**(6), 805-820. http://dx.doi.org/10.12989/eas.2012.3.6.805.

Fujita, K., Moustafa, A. and Takewaki, I. (2010), "Optimal placement of viscoelastic dampers and supporting members under variable critical excitations", *Earthq. Struct.*, **1**(1), 43-67. http://dx.doi.org/10.12989/eas.2010.1.1.043.

Furuya, O., Hamazaki, H. and Fujita, S. (1998), "Study on proper distribution of storey-installation type damper for vibration control of slender structures using genetic algorithm", *ASME-PUBLICATIONS-PVP*, **364**, 297-304.

Guo, Y., Sun, B.N., Ye, Y., Lou, W.J. and Shen, G.H. (2009), "Hybrid genetic algorithm for optimizing damper distribution of transmission towers", *J. Harbin Inst. Technol.*, **10**, 259-264.

Haftka, R.T. and Adelman, H.M. (1985), "Selection of actuator locations for static shape control of large space structures by heuristic integer programing", *Comput. Struct.*, **20**(1-3), 575-582. https://doi.org/10.1016/0045-7949(85)90105-1.

Hinton, G.E. and Salakhutdinov, R.R. (2006), "Reducing the dimensionality of data with neural networks", *Sci.*, **313**(5786), 504-507. https://doi.org/10.1126/science.1127647.

Huang, Z.S., Wu, C. and Hsu, D.S. (2009), "Semi-active fuzzy control of mr damper on structures by genetic algorithm", *J. Mech.*, **25**(1), N1-N6. https://doi.org/10.1017/S172771910000366X.

Kim, J. and Choi, H. (2006), "Displacement-based design of supplemental dampers for seismic retrofit of a framed structure", *J. Struct. Eng.*, **132**(6), 873-883. https://doi.org/10.1061/(ASCE)0733-9445(2006)132:6(873).

Kim, J., Choi, H. and Min, K.W. (2003), "Performance-based design of added viscous dampers using capacity spectrum method", *J. Earthq. Eng.*, **7**(1), 1-24.

Li, H.N. and Qu, J.T. (2010), "Comparison of optimal placement of displacement-based and velocity-based dampers using genetic algorithm", *Chin. J. Comput. Mech.*, **27**(2), 252-257.

Li, L., Song, G. and Ou, J. (2010), "A genetic algorithm-based two-phase design for optimal placement of semi-active dampers for nonlinear benchmark structure", *J. Vib. Control*, **16**(9), 1379-1392. https://doi.org/10.1177/1077546309103277.

Milman, M.H. and Chu, C.C. (1994), "Optimization methods for passive damper placement and tuning", *J. Guid. Control Dyn.*, **17**(4), 848-856. https://doi.org/10.2514/3.21275.

Mohebbi, M. and Joghataie, A. (2012), "Designing optimal tuned mass dampers for nonlinear frames by distributed genetic algorithms", *Struct. Des. Tall Spec. Build.*, **21**(1), 57-76. https://doi.org/10.1002/tal.702.

Natke, H.G. and Soong, T.T. (1993), *Topological Structural Optimization under Dynamic Loads. Computer Aided Optimum Design of Structures III*, Elsevier Science Publishers Ltd.

Qu, J.T. and Li, H.N. (2008), "Research on position optimization of viscoelastic damper in structural damping control", *J. Vib. Shock*, **27**(6), 87-91.

Singh, M.P. and Moreschi, L.M. (2002), "Optimal placement of dampers for passive response control", *Earthq. Eng. Struct. Dyn.*, **31**(4), 955-976. https://doi.org/10.1002/eqe.132.

Takewaki, I. (1997), "Optimal damper placement for minimum transfer functions", *Earthq. Eng. Struct. Dyn.*, **26**(11), 1113-1124. https://doi.org/10.1002/(SICI)1096-9845(199711)26:11.

Takewaki, I. (2009), *Building Control with Passive Dampers: Optimal Performance-based Design for Earthquakes*, John Wiley & Sons Ltd..

Takewaki, I., Yoshitomi, S., Uetani, K. and Tsuji, M. (1999), "Non-monotonic optimal damper placement via steepest direction search", *Earthq. Eng. Struct. Dyn.*, **28**(6), 655-670. https://doi.org/10.1002/(SICI)1096-9845(199906)28:6.

Vapnik, V. (1963), "Pattern recognition using generalized portrait method", *Autom. Remote Control*, **24**, 774-780.

Vapnik, V. and Chervonenkis, A. (1964), "A note on one class of perceptions", *Autom. Remote Control*, **25**, 821-837.

Wu, B., Ou, J.P. and Soong, T.T. (1997), "Optimal placement of energy dissipation devices for three-dimensional structures", *Eng. Struct.*, **19**(2), 113-125. https://doi.org/10.1016/S0141-0296(96)00034-X.

Wu, R.T. and Jahanshahi, M.R. (2018), "Deep convolutional neural network for structural dynamic response estimation and system identification", *J. Eng. Mech.*, **145**(1), 04018125. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001556.

Yan, L.W., Chen, Y.Y., Wang, L. and Su, C. (2010), "Optimum installation of viscous dampers in tall buildings based on relative fitness genetic algorithm", *J. Vib. Shock*, **6**.

Zhang, R. and Soong, T.T. (1992), "Seismic design of viscoelastic dampers for structural applications", *J. Struct. Eng.*, **118**(5), 1375-1392. https://doi.org/10.1061/(ASCE)0733-9445(1992)118:5(1375).

AT