Smart Structures and Systems, *Vol. 36, No. 1 (2025) 51-70*
https://doi.org/10.12989/sss.2025.36.1.051

A ground motion synthesis method based on Kernel Principal Component Analysis and genetic algorithm for accurate seismic analysis

Zhen Liu*[1] and Xingliang Ma [2a]

[1] *School of Management Science and Engineering, Shandong Technology and Business University, No. 191, Binhai Middle Road, Laishan District, Yantai City, Shandong, China*
[2] *School of Civil Engineering and Architecture, Changzhou Institute of Technology, No. 666, Liaohe Road, Xinbei District, Changzhou City, Jiangsu, China*

(*Received February 21, 2025, Revised July 10, 2025, Accepted August 5, 2025*)

Abstract. Efficient and accurate ground motion simulation is key to achieving reliable structural seismic analysis. However, ground motions are significantly influenced by multiple factors, such as the source mechanism, propagation path, and site conditions, resulting in notable nonstationarity and spatial variability. Traditional ground motion synthesis methods based on statistical models are often inadequate to address these complexities. Therefore, this paper proposes a ground motion synthesis method based on Kernel Principal Component Analysis (KPCA) and Genetic Algorithm (GA). By extracting characteristic mother waves from specific earthquake records and using genetic algorithms, artificial ground motions that meet the target response spectrum and peak ground acceleration (PGA) are generated. The method is applied to the Hualien earthquake in Taiwan as an example, and the results show that the synthesized ground motions effectively match the target response spectrum and PGA, with a maximum error within 7%. Further power spectral analysis demonstrates a high similarity between the synthesized and natural ground motions in terms of power spectral characteristics. Moreover, a finite element analysis of a typical frame structure under both artificial and similar natural ground motions from a database shows consistent seismic responses and patterns between the two. To further validate the applicability of the proposed method, ten ground motions were removed from the database, and artificial ground motions were synthesized based on their acceleration response spectra and peak ground velocity (PGV) information. Comparisons of waveform characteristics, PGA, PGV, and acceleration response spectra indicate that the artificial ground motions effectively retain the waveform characteristics of natural ground motions. The proposed method provides an efficient tool for engineering seismic applications, capable of synthesizing artificial ground motions that preserve the waveform and damage characteristics of natural ground motions while meeting specified target response spectrum, PGV, PGA, and other parameters.

Keywords: genetic algorithm; ground motion synthesis; Kernel Principal Component Analysis (KPCA); nonstationarity; response spectrum; seismic simulation

1. Introduction

With the rapid advancement of urbanization, cities have continuously expanded in scale, leading to an increasing concentration of infrastructure-particularly bridges, tunnels, high-rise buildings, and critical energy and transportation facilities-all of which are facing growing seismic threats (Hakamian *et al*. 2023, Chen *et al*. 2024, Ambrosini *et al*. 2021). As urban development becomes more dense and complex, the potential destructive consequences of earthquakes grow more severe, posing significant risks to densely populated urban areas and their interconnected infrastructures (Feng *et al*. 2024, Zhang and Gu 2022, Liu *et al*. 2024b). Therefore, scientifically and systematically assessing the impact of strong seismic events on these large-scale, critical infrastructures has become one of the central challenges in earthquake engineering research. Addressing this issue is essential to enhance urban resilience, minimize the vulnerability of key infrastructure, and protect the safety of both lives and economic activities during major earthquakes.

In seismic analysis, obtaining accurate and reliable ground motion input is fundamental to ensuring the reliability of structural seismic assessments (Wani and Vemuri 2024, Ding *et al*. 2023). Ground motion not only directly influences the safety evaluation of engineering structures but is also closely related to subsequent structural design and the formulation of seismic measures (Fayaz and Galasso 2023). Due to the influence of various factors, such as the source mechanism, propagation path, and site effects, ground motions exhibit significant differences in waveform, frequency characteristics, and energy distribution across different locations (Zhu *et al*. 2022, Chang *et al*. 2022, Hou and Wang 2024). Therefore, when conducting seismic design for structures, one cannot solely rely on standardized ground motion models. Instead, it is essential to incorporate the seismic characteristics of the specific site for a more targeted analysis.

*Corresponding author, Ph.D., Associate Professor,
E-mail: 201913693@sdtbu.edu.cn
[a] Ph.D., Associate Professor

ISSN: 1738-1584 (Print), 1738-1991 (Online)

Ground motions in strong earthquake regions exhibit significant spatiotemporal variability (Takemura *et al*. 2022, Wei *et al*. 2023, Liu 2025), and the scarcity of high-quality samples further complicates ground motion simulation (Zhang *et al*. 2023). Therefore, enhancing sample utilization efficiency during analysis and accurately capturing the characteristics of ground motions are key to precise simulations. Given the complex nature of ground motion signals, traditional approaches often employ mathematical models to fit nonstationary characteristics, while modern methods increasingly leverage intelligent algorithms to extract sample features.

Classical ground motion variability simulation techniques have reached a relatively mature stage. These methods generally rely on existing samples, using specific formulas to describe the nonstationary characteristics of ground motion and incorporating them into stationary models to generate ground motions with spatiotemporal variability (Liu *et al*. 2024a). When extracting and expressing nonstationarity, methodological approaches diverge: some are based on time-domain perspectives, while others focus on frequency-domain analysis. Although different models adopt varying approaches, they share a similar core principle when dealing with nonstationarity, often describing temporal variability using time-varying power spectra and time-varying envelope functions. Many models are based on the Kanai-Tajimi filter model or its modified versions (Kanai 1957), such as those proposed by scholars like Yu *et al*. (2024), Qin and Li (2023), who introduced various improvements by integrating seismic characteristics and site conditions. With the advancement of signal processing technologies, some researchers have applied techniques such as wavelet transform (Wen *et al*. 2023) and Hilbert transform (Li *et al*. 2022) in the construction of time-varying power spectra and envelope functions.

To achieve closer agreement with a target response spectrum, spectrum-compatible techniques are now a mainstay of ground-motion simulation. Pioneering contributions include the frequency-domain iterative algorithm of Gasparini and Vanmarcke (1976), the multi-damping time-domain method of Lilhanand and Tseng (1988), and Abrahamson's (1992) enhancements for non-stationary spectra. These studies established the theoretical and algorithmic foundations on which later optimization-based spectral-matching procedures were built.

Classical stochastic ground-motion models—such as ARMA processes and extended Kanai–Tajimi filters—remain indispensable for capturing the random nature of earthquake excitations (Kanai 1957, Tajimi 1960). When combined with spectral-matching strategies, they offer the dual advantages of statistical consistency and dynamic compatibility, producing motions that satisfy both probabilistic descriptions and engineering design requirements.

In recent years, the rise of multi-objective intelligent-optimization frameworks has accelerated the generation of high-quality artificial ground motions. Within this context, NSGA-II (Deb *et al*. 2002) has become the algorithm of choice owing to its fast, elitist Pareto search, allowing simultaneous optimization of multiple targets such as the response spectrum, waveform similarity, and peak ground acceleration (PGA). Hybrid schemes—most notably the KPCA-GA approach—capitalize on NSGA-II's efficiency to preserve site-specific features while producing motions that match a prescribed design spectrum with high fidelity.

However, with the improved precision of near-fault ground motion recordings (Miller *et al*. 2023, Wilson *et al*. 2024), the temporal variability of ground motions has become increasingly complex, and traditional sample-fitting methods have begun to struggle with these complexities (Aristeidou *et al*. 2024).

Sufficient near-fault ground motion samples are crucial for seismic analysis and artificial ground motion synthesis. In recent years, several high-density, broadband seismic arrays and networks have been established, recording significant seismic variability in strong earthquake regions. However, due to the varying objectives of monitoring network layouts, high-quality ground motion samples remain scarce. Dense arrays have high sampling densities but limited coverage, making it challenging to obtain adequate ground motion samples across broader seismic regions (e.g., the East Bay Seismic Array and Sweetwater Array in the United States). Wide-area networks, though covering large areas, have lower sampling densities, similarly making it difficult to gather a substantial number of high-quality samples (e.g., the NORSAR core array in Northern Europe). Therefore, deploying seismic networks with high density and wide coverage tailored to fault characteristics-such as the seismic network operated by the Central Weather Bureau in Taiwan-becomes particularly important.

To address these challenges, this study proposes an advanced intelligent algorithm combining Kernel Principal Component Analysis (KPCA) and Genetic Algorithm (GA) to synthesize ground motions efficiently based on actual ground motion observational data from the target region. Specifically, KPCA is first employed to extract representative mother waves from the ground motion database of the target region, capturing the primary characteristics of ground motions in that area. Then, key features such as the target response spectrum and peak ground acceleration (PGA) are used as synthesis objectives, and genetic algorithms are utilized to optimize and solve based on the extracted characteristic mother waves, ultimately generating artificial ground motions that reflect the target regional characteristics and meet the synthesis requirements.

To systematically assess the quality of the synthesized ground motions, multiple aspects are analyzed, including response spectrum characteristics, waveform shapes, and time-frequency features. Additionally, finite element analysis is conducted to evaluate the dynamic response of a typical frame structure model, comprehensively assessing the structural response and damage behavior under artificial ground motions. By comparing the synthesized ground motions with natural ground motions, this study verifies the effectiveness and accuracy of the synthesized ground motions, providing a more reliable basis for seismic simulation and analysis in engineering applications. The specific method flowchart for this paper is shown in Fig. 1.

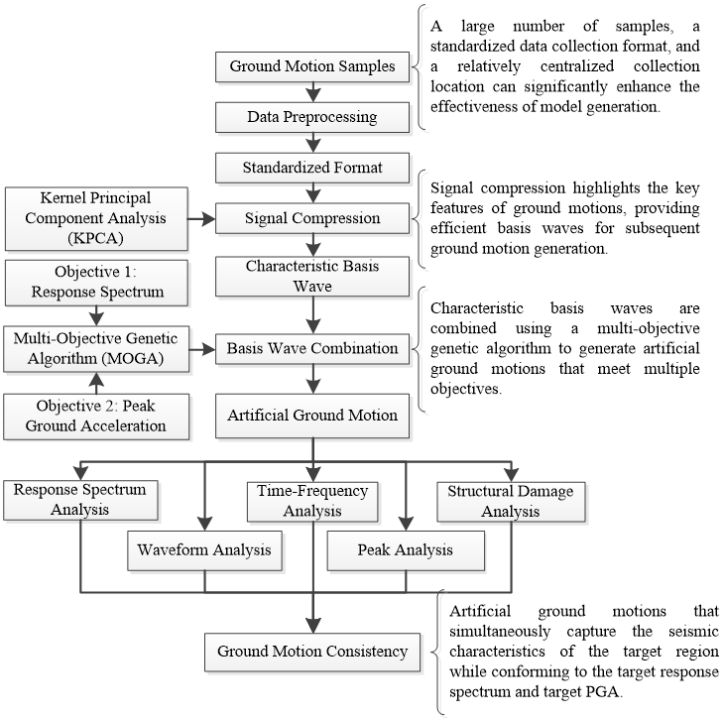


Fig. 1 Research process flowchart

2. Kernel Principal Component Analysis (KPCA)

Kernel Principal Component Analysis (KPCA) is a nonlinear extension of Principal Component Analysis (PCA) designed to handle nonlinearly separable datasets. The core idea of KPCA is to project data from the original space into a high-dimensional feature space using a nonlinear mapping function, making the data more linearly separable in this higher-dimensional space. Once in the high-dimensional space, PCA can be applied to achieve dimensionality reduction (Zhang *et al*. 2020).

Suppose the dataset to be processed is $X = \{x_1, x_2, \ldots, x_n\}$, where n is the number of samples, and each sample x_i is a d dimensional vector. The feature extraction process of KPCA can be roughly divided into the following steps:

(1) High-Dimensional Mapping: The sample data is mapped from the original space to a high-dimensional feature space F using a nonlinear mapping function $\Phi(\bullet)$. In the feature space F, the covariance matrix of the samples can be represented as

$$C_F = \frac{1}{n}\sum_{i=1}^{n}\phi(x_i)\phi(x_i)^T \tag{1}$$

(2) Eigenvalue Decomposition: The covariance matrix C_F is decomposed to obtain its eigenvalues λ and eigenvectors v

$$\lambda v = C_F v \tag{2}$$

Taking the inner product of the mapped sample $\Phi(x_i)$ with Eq. (2) yields

$$\lambda[\phi(x_i)v] = \phi(x_i)C_F v \tag{3}$$

Since $\lambda \neq 0$, in the high-dimensional space, the eigenvector can be expressed as a linear combination of all mapped samples $\Phi(x_i)$

$$v = \sum_{i=1}^{n}\alpha_i\phi(x_i) \tag{4}$$

where α_i are the coefficients.
Substituting Eqs. (4) and (1) into Eq. (3) gives

$$\lambda\sum_{i=1}^{n}\alpha_i[\phi(x_i)\phi(x_k)] = \frac{1}{n}\sum_{i=1}^{n}\alpha_i\sum_{i=1}^{n}[\phi(x_k)\phi(x_j)][\phi(x_k)\phi(x_i)] \tag{5}$$

(3) Definition of Kernel Matrix: Define the matrix K such that $K_{ij} = [\Phi(x_i)\ \Phi(x_j)]$. Then Eq. (5) can be represented as

$$\lambda K\alpha = \frac{1}{n}KK\alpha \tag{6}$$

which can be further simplified to

$$n\lambda\alpha = K\alpha \tag{7}$$

Using the eigenvector α, the eigenvector v of C_F is obtained, which then yields the principal components in the feature space F. Diagonalizing K gives its eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda n$ and the corresponding eigenvectors $\alpha_1, \alpha_2, \ldots, \alpha_n$.

(4) Calculation of Contribution Rate: Generally, the contribution rate of the principal components η_i can be determined using Eq. (8)

Table 1 Main seismic parameters

Earthquake ID	Earthquake name	Occurrence time (Local time)	Magnitude (Richter)	Location	Focal depth (km)
113019	Hualien Earthquake	April 3, 7:58 AM	7.2	25.0 km southeast of Hualien Government	15.5
113021	Aftershock 1	April 3, 8:11 AM	6.5	12.7 km northeast of Hualien Government	5.5
113045	Aftershock 2	April 3, 10:14 AM	6.2	35.8 km northeast of Hualien Government	17.1
113266	Aftershock 3	April 23, 2:26 AM	6.0	29.9 km south of Hualien Government	10.0
113267	Aftershock 4	April 23, 2:32 AM	6.3	17.2 km southwest of Hualien Government	5.5
113351	Aftershock 5	April 27, 2:21 AM	6.1	23.0 km northeast of Hualien Government	24.9

$$\eta_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} \tag{8}$$

KPCA uses a kernel function to map the original low-dimensional data to a high-dimensional feature space, where nonlinear distributions may become linearly separable. Thus, PCA can be effectively applied in this space for dimensionality reduction. KPCA offers flexibility in supporting various kernel functions, such as linear, polynomial, radial basis function (RBF), and sigmoid kernels. Choosing an appropriate kernel function based on the characteristics of the data can improve dimensionality reduction performance and enhance analysis accuracy and efficiency. This advantage makes KPCA highly suitable for feature extraction and dimensionality reduction of complex nonlinear data, presenting significant potential for broad applications.

When applying kernel principal component analysis (KPCA) to extract the complex, nonlinear features of seismic ground motions, it is essential not only to select an appropriate kernel — such as linear, RBF, polynomial, or sigmoid — to enhance dimensionality-reduction performance, but also to rigorously validate both the effectiveness of the reduction and the accuracy of the reconstruction. A widely adopted practice is to compute the cumulative explained variance of the eigenvalues, which gauges the contribution of each principal component and verifies whether the retained dimensions capture the bulk of the original information (Schölkopf *et al*. 1998, Rosipal and Girolami 2001).

3. Strong earthquake samples in the target region

In this study, KPCA is used for dimensionality reduction and compression of ground motion data. Based on ground motion data from the target region, artificial ground motions are synthesized to match a specific ground motion response spectrum using a multi-objective genetic algorithm. Therefore, ensuring that the target region has a sufficient quantity and high quality of ground motion data is crucial for generating efficient base waveforms.

Taiwan is located in the active zone of the circum-Pacific seismic belt, one of the most seismically active areas in the world. In particular, the Hualien-Taitung Rift Valley (Hualien County), situated at the boundary between the Eurasian Plate and the Philippine Sea Plate, experiences frequent seismic activity with high magnitudes, making it a key area for earthquake research and monitoring. This study focuses on several strong earthquakes that occurred in this region in 2024. On April 3, 2024, at 7:58 AM, a magnitude 7.2 mainshock struck the Hualien area, followed by two aftershocks of magnitude 6.5 and 6.2 at 8:11 AM and 10:14 AM on the same day. Additionally, on April 23, 2024, at 2:26 AM and 2:32 AM, two earthquakes of magnitude 6.0 and 6.3, respectively, occurred. On April 27, 2024, at 2:21 AM, another magnitude 6.1 aftershock hit Hualien. A total of six earthquakes with magnitudes of 6.0 or above were recorded, and over 3,000 ground motion data records were collected, forming the ground motion sample database used in this study. The main seismic parameters are shown in Table 1.

From the analysis of the data in Table 1, it can be seen that six significant earthquakes occurred within approximately 30 km of Hualien. The region has a relatively dense distribution of seismic monitoring stations equipped with advanced data collection devices, and the data formats are standardized, providing solid technical support and foundational guarantees for the data compression and simulation analysis conducted in this study.

Studies have shown that ground motion characteristics are influenced by several factors, including the source mechanism, propagation path, and site conditions, and that ground motion waveforms may exhibit unique features in different regions. The ground motion data used in this study are derived from one mainshock and its subsequent five aftershocks. Since the earthquake sources are relatively close, the synthesized artificial ground motions are able to retain many of the typical characteristics of these seismic events.

From the mainshock and five aftershocks, 1,657 sets of ground motion records were carefully selected, each with a length of 30,000 data points and a sampling rate of 100 Hz. After preliminary processing, 1,611 high-quality ground motion records were retained. However, the large volume of data may reduce the efficiency of the subsequent artificial ground motion generation process. Therefore, in this chapter, the KPCA method is further utilized to perform dimensionality reduction and compression of these ground motion data, improving both the efficiency and precision of the ground motion generation process.

In this section, three different kernel functions—Gaussian, Sigmoid, and Laplacian—are employed to reduce

the dimensionality of the ground motion data. Each of these kernel functions has unique characteristics and different abilities to capture the nonlinear structures in the data, allowing for an effective assessment of the performance of different kernel methods for data compression.

(1) Gaussian Kernel:

$$k(x,y) = \exp(-\gamma_1 \|x-y\|^2) \tag{9}$$

Parameter: γ_1 controls the rate of decay, and $\|x - y\|$ represents the Euclidean distance between input vectors x and y.

(2) Sigmoid Kernel:

$$k(x,y) = \tanh(\gamma_2 x^T y + c) \tag{10}$$

Parameters: $\gamma_2$is the scaling parameter, c is the bias term, and $x^T y$ represents the inner product of vectors x and y.

(3) Laplacian Kernel:

$$k(x,y) = \exp(-\gamma_3 \|x-y\|) \tag{11}$$

Parameter: γ_3 controls the rate of decay, and $\|x-y\|$ represents the Euclidean distance between input vectors x and y.

To compare the compression effectiveness of the three kernel functions, the efficiency of compression for the east-west component of the ground motions is used for evaluation. Specifically, Fig. 2 shows the efficiency curves of different kernel functions during the dimensionality reduction process, while Table 2 summarizes the compression efficiency metrics of the three kernels. From Fig. 2 and Table 2, it is evident that the Gaussian, Sigmoid, and Laplacian kernels exhibit differences in the compression performance for ground motion data under various parameter settings.

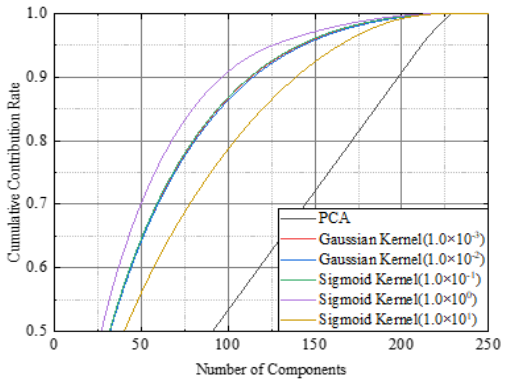


Fig. 2 Compression efficiency of different kernel functions

Analysis of the compression efficiency data for various kernel functions in Table 2 reveals that the Sigmoid kernel demonstrates outstanding performance in dimensionality reduction. Specifically, the Sigmoid kernel requires only 97 characteristic mother waves to effectively capture the core features of the original ground motion database. This indicates that the Sigmoid kernel not only efficiently retains the key information of the ground motion data but also significantly reduces computational complexity.

Moreover, using these carefully extracted characteristic mother waves, artificial ground motions that exhibit specific characteristics can be synthesized by suitable combinations and adjustments. This approach provides a reliable basis for artificial ground motion synthesis, significantly improving computational efficiency while maintaining accuracy.

4. Artificial ground motion generation based on multi-objective genetic algorithm

This section focuses on the acceleration response spectrum and peak ground acceleration (PGA) as key objectives for generating artificial ground motions, utilizing the characteristic ground motion data extracted from Hualien earthquakes in the previous section. To achieve this, an advanced multi-objective genetic algorithm is employed to explore an optimal set of combination coefficients for the characteristic ground motions. This approach not only generates artificial ground motions that meet engineering requirements, such as matching specific target response spectra and PGA values, but also accurately captures the unique characteristics of the Hualien earthquakes.

The multi-objective genetic algorithm simulates natural evolution processes, including selection, crossover, and

Table 2 Cumulative contribution rate of different kernel functions

Gaussian Kernel		Sigmoid Kernel		Laplacian Kernel	
Parameter value	Number of principal components for 90% cumulative contribution rate	Parameter value	Number of principal components for 90% cumulative contribution rate	Parameter value	Number of principal components for 90% cumulative contribution rate
1.00E-04	114	1.00E-04	114	1.00E-04	169
1.00E-03	114	1.00E-03	114	1.00E-03	171
1.00E-02	115	1.00E-02	114	1.00E-02	184
1.00E-01	128	1.00E-01	113	1.00E-01	197
1.00E+00	180	1.00E+00	97	1.00E+00	198
1.00E+01	195	1.00E+01	140	1.00E+01	198
1.00E+02	197	1.00E+02	143	1.00E+02	199

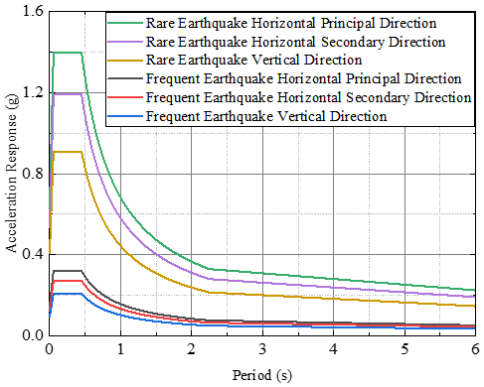


Fig. 3 Target earthquake acceleration response spectrum

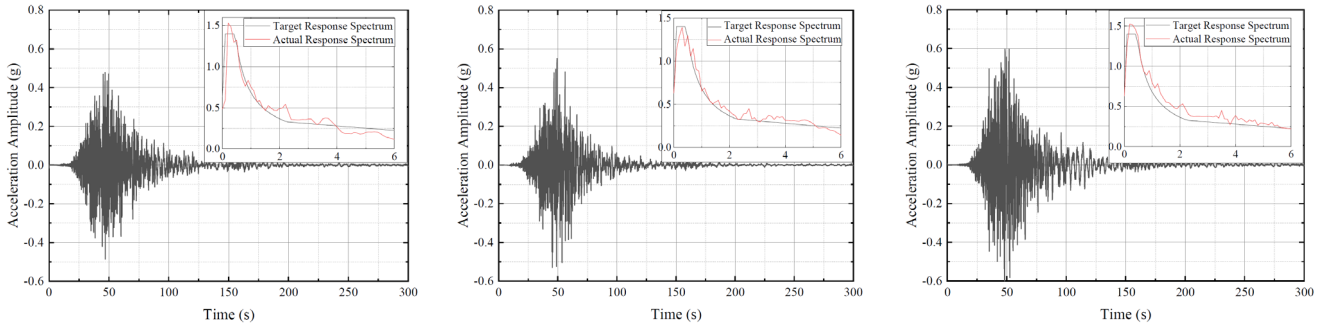
mutation mechanisms, to explore the non-dominated solution set within the solution space. Compared to single-objective genetic algorithms, multi-objective genetic algorithms show significant advantages, as they can simultaneously optimize multiple objective functions, thus identifying multiple optimal solutions that satisfy specific conditions. When dealing with complex multi-objective optimization problems, multi-objective genetic algorithms exhibit excellent robustness and global search capabilities.

The study in this section follows the recommendations outlined in relevant research (Ji *et al*. 2019), fully considering the specific site conditions in the Hualien area of Taiwan. Specifically, the seismic fortification intensity is not less than 9 degrees, the basic design acceleration value is not less than 0.40 g, the site type is medium-hard soil, and the characteristic period is 0.66-1.25 seconds (Kuo *et al*. 2019). The shape of the earthquake acceleration response spectrum, as shown in Fig. 3, serves as an important reference for generating artificial ground motions. The goal of this study is to generate artificial ground motions that not only meet engineering requirements but also accurately reflect the characteristics of the Hualien earthquakes.

Frequent earthquakes refer to seismic events with a relatively high probability of occurrence within the design life of a structure. These earthquakes cause minor effects on the structure but do not lead to significant damage. Frequent earthquakes typically correspond to a 63% probability of exceedance within 50 years. Rare earthquakes, on the other hand, are events with a relatively low probability of occurrence within the design life of a structure. These earthquakes can have a substantial impact on the structure but should not result in overall collapse. Rare earthquakes are associated with longer return periods, corresponding to a 2%-3% probability of exceedance within 50 years. Using the horizontal principal direction of a rare earthquake as an example, this paper describes in detail the ground motion synthesis method based on KPCA and the multi-objective genetic algorithm. First, the earthquake response spectrum of the horizontal principal direction of a rare earthquake is set as the first optimization target, and the PGA is set as the second optimization target to generate artificial ground motions that meet engineering needs. To ensure the reproducibility of the proposed method, this study provides a detailed description of the configuration used for the multi-objective genetic algorithm. Specifically, the following parameter settings were adopted:

a population size of 200, a maximum of 2000 generations, a crossover rate of 0.8, a mutation probability of 0.05, a random mutation strategy, and a tournament selection method. The multi-objective sorting was performed using the NSGA-II (Non-dominated Sorting Genetic Algorithm II), and the stopping criterion was set as reaching the maximum number of generations.

The comparison between the generated ground motion response spectrum and the target response spectrum is



(a) Artificial ground motion 1 (b) Artificial ground motion 2 (c) Artificial ground motion 3

Fig. 4 Comparison of artificial ground motion acceleration time histories and response spectra

Table 3 Relative error of artificial ground motions

Index	Target PGA (g)	Actual PGA (g)	Relative error of PGA	Peak response spectrum (g)	Relative error of response spectrum
1	0.50	0.49	2.78%	1.53	5.24%
2	0.55	0.55	0.09%	1.39	1.73%
3	0.60	0.64	6.70%	1.53	2.35%

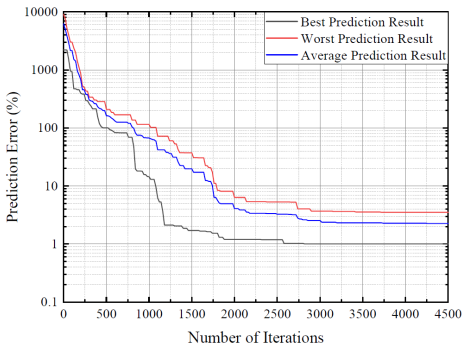


Fig. 5 Variation of relative error of artificial ground motions under different iteration numbers

shown in Fig. 4, and key parameters are listed in Table 3.

By observing the relative errors above, it can be seen that the generated ground motions effectively match the target response spectrum and peak ground acceleration (PGA), with a maximum error controlled within 7%. Additionally, the response spectra of the artificial ground motions maintain good consistency with the target response spectrum at all stages, without significant deviations.

To further explore the effectiveness of the artificial ground motions generated by the proposed method, ten sets of artificial ground motions were generated, and the relative errors under different iteration counts were recorded (as shown in Fig. 5). By analyzing the trend of relative error changes, the accuracy of the proposed method at different iteration stages is illustrated, providing a more intuitive reflection of the stability and precision of the method during the process of generating artificial ground motions.

Fig. 5 shows the variation of the relative error of artificial ground motions under different iteration numbers. The vertical axis represents the relative error (in logarithmic scale), while the horizontal axis represents the number of iterations. The figure includes convergence curves for the best, worst, and average results.

The following observations can be made from Fig. 5:

(1) Overall Trend: All curves show that the relative error gradually decreases as the number of iterations increases, indicating that the genetic algorithm can progressively optimize the generation of artificial ground motions through continuous iterations. The error reduction is particularly significant in the initial stages, while the rate of improvement gradually decreases as iterations progress, ultimately converging.

(2) Difference Between Best and Worst Results: During the early iteration stages, there is a considerable gap between the best and worst errors. As the number of iterations increases, this gap gradually narrows. This suggests that the initial solution of the genetic algorithm may exhibit high uncertainty due to initial conditions, but with multiple iterations, the variability decreases, and the results become stable.

(3) Convergence Characteristics: The curves indicate that the error drops rapidly during the first 1,000 iterations, then gradually slows down, and stabilizes after approximately 3,500 iterations. This pattern demonstrates that the generation method has an initial phase of rapid improvement, followed by a gradual optimization phase. Iteration counts between 2,000 and 3,000 are relatively suitable.

In conclusion, the trend of relative error variation in the figure indicates that the proposed artificial ground motion generation method can progressively optimize the results through multiple iterations, ultimately achieving high precision and stability.

5. KPCA reconstruction process

To further provide a physical interpretation of the effectiveness of KPCA-based dimensionality reduction, this study performs signal reconstruction using the selected principal components after KPCA. By mapping the principal components from the kernel feature space back to the original feature space, the original seismic signals are reconstructed. This inverse mapping procedure serves as a critical validation of the ability of the retained components to preserve the essential characteristics of the signal.

From the original dataset of 1,611 seismic records, 10 records were randomly selected. In the signal generation process, the acceleration time history was defined as the primary optimization objective, while the acceleration response spectrum served as the secondary objective. To quantitatively evaluate the effectiveness of the KPCA-reduced components in signal reconstruction, two widely used indicators are adopted: Root Mean Square Error (RMSE) and the coefficient of determination (R^2). RMSE measures the overall discrepancy between the original and reconstructed signals, while R^2 assesses the goodness of fit.

RMSE can be expressed by Eq. (12) as follows

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(z_i - z_i')^2} \tag{12}$$

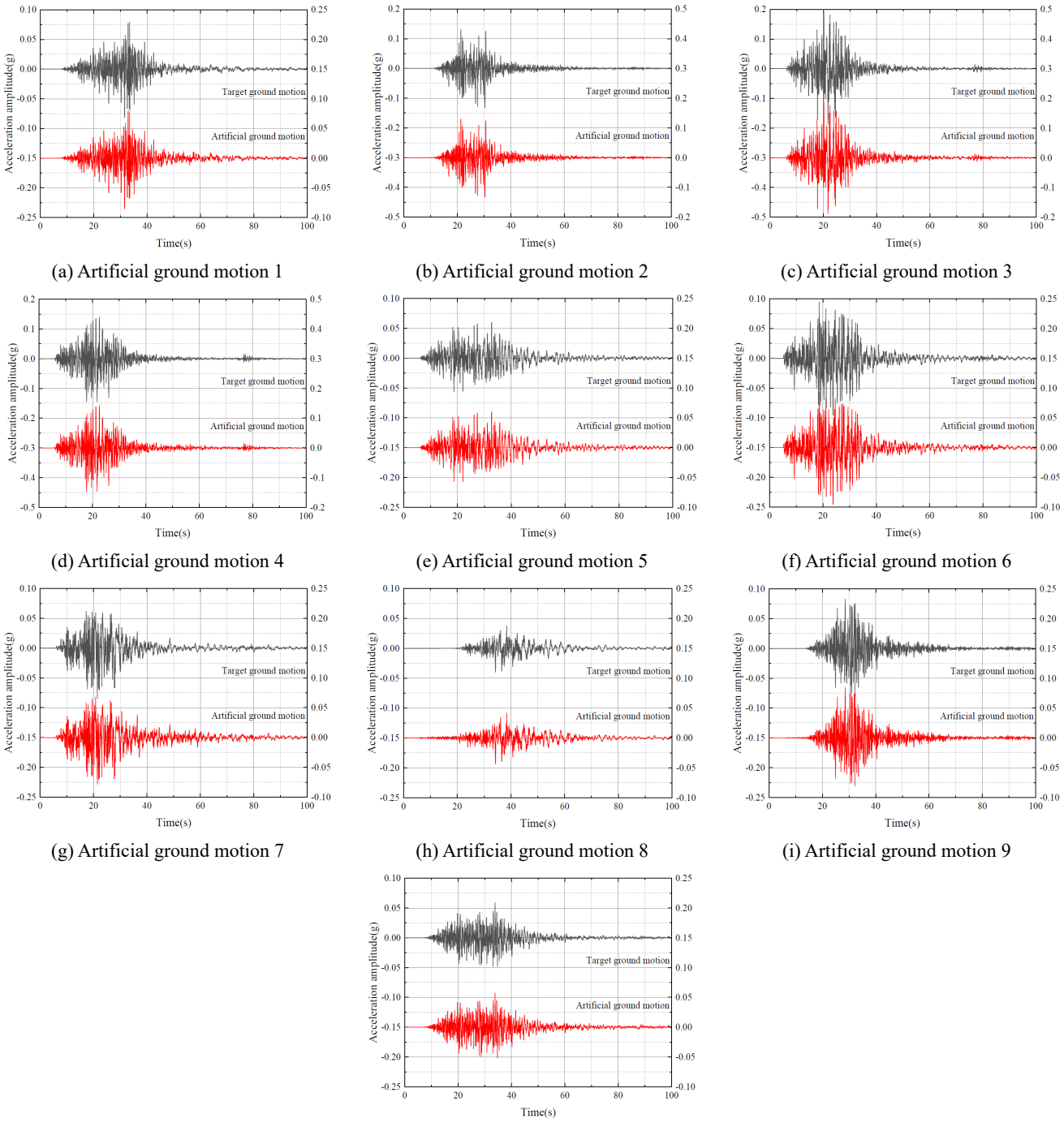
Where z_i denotes the i-th sampling point of the original signal, z_i' represents the i-th sampling point of the reconstructed seismic signal, and N is the total number of sampling points.

R^2 can be expressed by Eq. (13) as follows

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(z_i - z_i')^2}{\sum_{i=1}^{N}(z_i - z_i'')^2} \tag{13}$$

Where z_i'' is the mean value of the original signal.

The primary purpose of this ground motion generation is to demonstrate the effectiveness of KPCA compression. Given that both the acceleration time history and the acceleration response spectrum are set as optimization objectives, the genetic algorithm must undertake a substantial amount of optimization. Accordingly, the parameters were configured as follows: a total population size of 500, a selection size of 200 individuals per generation, a crossover rate of 0.8, a mutation probability of 0.05, and a maximum of 4000 generations. Fig. 6 illustrates



(a) Artificial ground motion 1 (b) Artificial ground motion 2 (c) Artificial ground motion 3

(d) Artificial ground motion 4 (e) Artificial ground motion 5 (f) Artificial ground motion 6

(g) Artificial ground motion 7 (h) Artificial ground motion 8 (i) Artificial ground motion 9

(j) Artificial ground motion 10

Fig. 6 Comparison between artificial and original ground motions

the original seismic records along with the acceleration time histories of the ten generated ground motions.

Fig. 6 compares the waveforms of the original and reconstructed seismic signals when 97 principal components are retained. As shown in the figure, the reconstructed waveform nearly overlaps with the original one under this configuration.

Table 4 presents the quantitative evaluation metrics—RMSE and R^2—for different numbers of principal components. The results indicate that when the number of components increases to 97, the RMSE drops significantly to below 0.001700, while the R^2 exceeds 0.9535, demonstrating that the KPCA method effectively preserves and reconstructs the essential features of the original seismic signals

Further analysis of RMSE and R^2 with respect to the number of principal components reveals that 97 components yield the lowest RMSE and the highest R^2. This suggests that using 97 components achieves high reconstruction efficiency. Including too many components may, in fact, reduce the optimization efficiency of the genetic algorithm and ultimately lower the reconstruction

Table 4 RMSE and R^2 metrics under different numbers of principal components

Number of principal components	RMSE		R2	
	Mean value	Standard deviation	Mean value	Standard deviation
10	2.10E-03	8.89E-04	9.38E-01	1.10E-03
25	2.00E-03	8.93E-04	9.39E-01	4.20E-03
50	2.00E-03	9.26E-04	9.44E-01	1.20E-02
97	1.70E-03	8.57E-04	9.54E-01	2.74E-02
200	2.10E-03	1.15E-04	8.87E-01	1.23E-01
500	2.30E-03	1.44E-04	8.65E-01	1.30E-01
600	2.30E-03	1.83E-04	8.70E-01	1.26E-01
800	2.30E-03	1.53E-04	8.75E-01	1.18E-01

accuracy.

6. Comparison of time-frequency characteristics

To further analyze the similarity between the artificial ground motions and those in the regional ground motion database, three artificially generated ground motions were used as samples to search for similar waveforms in the database. Since this section focuses on waveform similarity analysis, the original ground motions were subjected to appropriate amplitude modulation to facilitate the identification of ground motion segments similar to the artificial ones. Ultimately, similar ground motions to the three generated waveforms were retrieved from the database, as shown in Fig. 7. The similar waveform segments were enlarged for detailed comparison.

By comparing the artificial ground motions with similar natural waveforms from the database, it is evident that certain segments of the artificial ground motions exhibit high similarity to corresponding segments of the natural ground motions. This indicates that the artificial ground motions contain several fragments of natural ground motion waveforms. From this perspective, the artificial ground motions not only resemble the overall trend of natural ground motions but also exhibit distinct natural characteristics in local details. This similarity suggests that artificial ground motions can effectively capture the natural characteristics of ground motions at the target site, providing a more reliable reference for seismic response analysis.

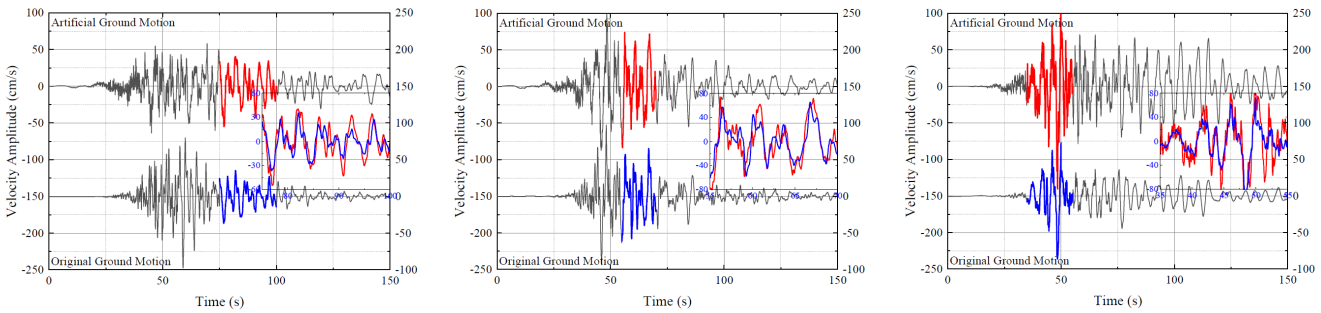
To further explore the differences between artificial and natural ground motions, this section investigates the frequency distribution of seismic energy. Due to differences in generation mechanisms, recording equipment, and site conditions, various types of ground motions often exhibit significant differences in their power spectral density (PSD) functions. Based on this, the performance of artificial ground motions in simulating natural ground motions is analyzed in detail.

First, the response spectrum of a rare ground motion with a horizontal principal direction under 9-degree seismic intensity conditions is set as the primary objective, while different PGA are used as secondary objectives. Eleven sets of ground motions were generated for the east-west direction, and the corresponding PSD plots were drawn (Fig. 8) to further illustrate the frequency domain differences between artificial and natural ground motions

The PSD analysis of the ground motions collected at various sites shows that the artificial ground motions exhibit a high degree of similarity with the corresponding natural ground motions in terms of energy distribution across different frequency bands, peak frequency variation, and attenuation characteristics. This suggests that artificial ground motions can effectively simulate the spectral characteristics of natural ground motions to some extent. However, there are significant differences in PSD between ground motions from different seismic events. For example, in the case of the 1999 Taiwan Chi-Chi Earthquake, although the data were also collected by the Taiwan Central Weather Bureau, the PSD of each ground motion still shows clear differences due to factors such as the source mechanism, energy release, and site conditions.

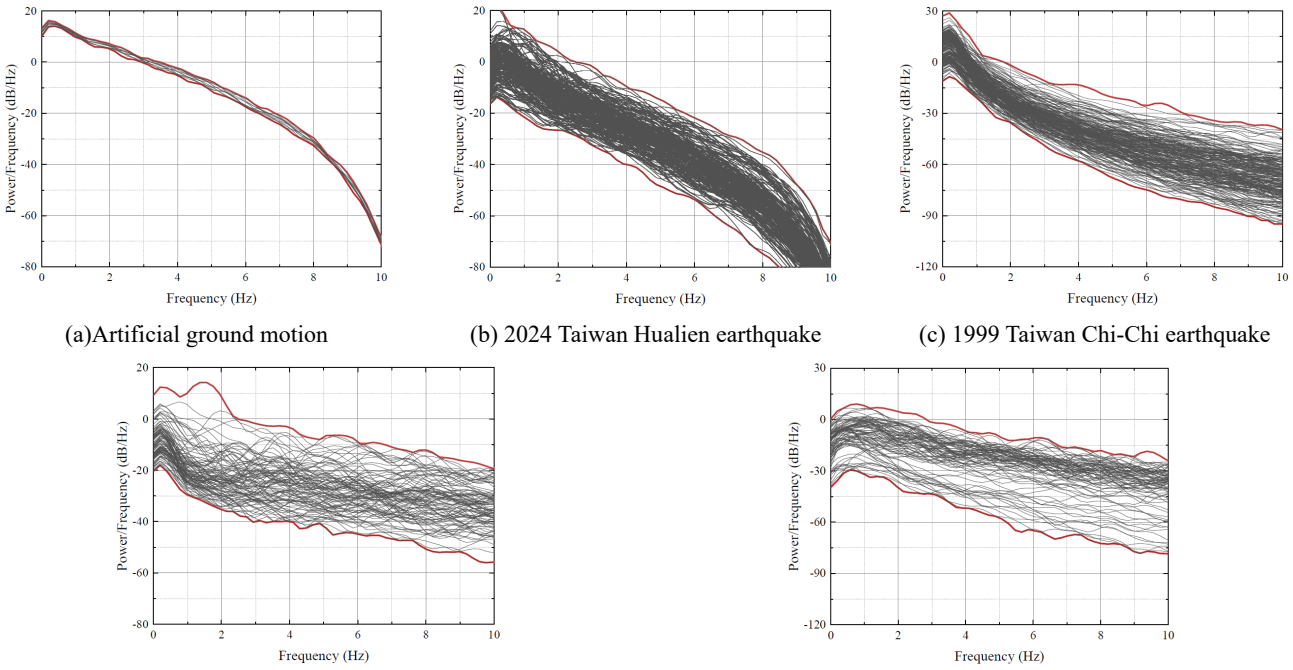
Further analysis of the PSDs of the 2000 Western Tottori Earthquake and the 2004 Parkfield Earthquake reveals even more pronounced differences. These phenomena indicate that ground motion records from different seismic events show significant differences in frequency distribution, energy concentration, and attenuation characteristics due to their unique source mechanisms and geological backgrounds. Moreover, as the variability of ground motion collection conditions increases, the trend of PSD differences becomes more evident.

The consistency in PSD between artificial ground



(a) Artificial ground motion 1 (b) Artificial ground motion 2 (c) Artificial ground motion 3

Fig. 7 Comparison of artificial and natural ground motion waveforms



(a)Artificial ground motion (b) 2024 Taiwan Hualien earthquake (c) 1999 Taiwan Chi-Chi earthquake

(d) 2000 Western Tottori earthquake, Japan (e) 2004 Parkfield earthquake

Fig. 8 Different PSD plots of ground motion

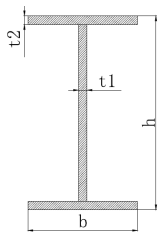
motions and corresponding natural ground motions further demonstrates that artificial ground motions can inherit many spectral characteristics of natural ground motions, providing a solid spectral foundation for simulating realistic site-specific seismic responses.

7. Seismic damage analysis

To investigate the distinctions in seismic damage between artificial synthetic ground motions and natural ground motions, a typical frame structure model was developed using ANSYS software. Comparative analyses were performed by inputting both natural and artificial ground motions, focusing on structural response. To quantify seismic damage, this study adopts the Park–Ang damage index, which integrates peak deformation and cumulative hysteretic energy, and has proven effective for both reinforced-concrete and steel structures (Park and Ang 1985).

The analysis utilized a Benchmark steel structure model, which was originally designed by the SAC Joint Venture in the United States (SAC Joint Venture / FEMA 2000). This model represents a prototypical steel structure designed according to American building standards and is extensively employed in seismic response studies. The model's plan dimensions are 30.5 m × 36.6 m, with a total height of 80.77 m. The structure comprises six spans in the east-west direction and five spans in the north-south direction, with each span measuring 6.10 m. The structure consists of 20 above-ground stories and two underground levels. Except for the first story, each above-ground story has a height of 5.49 m, while the remaining floors are 3.96 m high. The two underground levels each have a height of 3.65 m.

Table 5 Key parameters of beams and columns

Section type	Diagram	Parameters (Unit: mm)			
		h	b	t1	t2
W30×99		754	267	13	17
W30×108		757	267	14	19
W27×84		678	254	12	16
W24×335		699	343	35	63
W24×229		660	333	24	44
W24×192		648	330	21	37
W24×131		622	328	15	24
W24×117		617	325	14	22
W24×84		612	229	12	20
W24×62		602	179	11	15
W21×50		528	166	10	14

The columns of the structure are constructed from 345 MPa steel. Corner columns are box-shaped, whereas other columns are I-shaped. The floor system consists of composite steel-concrete decks, which are assumed to possess infinite in-plane stiffness. The composite steel members have a strength of 248 MPa. Columns are pinned at the foundation, while frame connections are rigidly fixed at ground level. Detailed structural diagrams and material parameters are presented in Fig. 8 and Table 5.

In the structural model, the side length of the steel box columns is 0.38 m, with thickness represented by the symbol t. Column joints are placed every three stories and are capable of resisting bending moments arising from horizontal loads as well as vertical loads. Beam-column

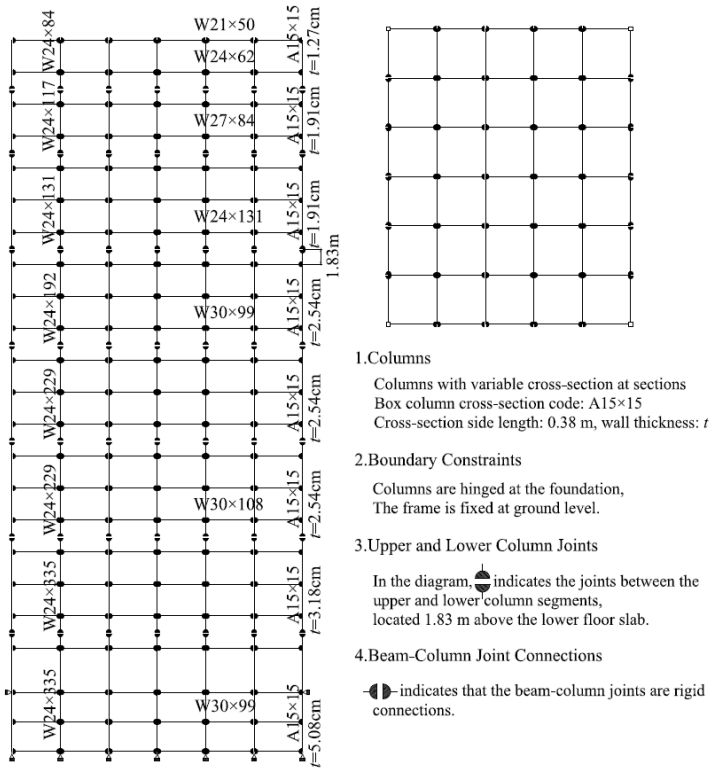


Fig. 9 Steel frame structure diagram

connections are rigid to ensure the overall integrity and stability of the structure under seismic excitation.

The seismic mass of the structure encompasses the self-weight of the horizontal slabs, vertical elements, as well as the weight of ancillary equipment and pipelines. The mass of the first floor is approximately 532 tons, the second floor is 565 tons, each of the third through twentieth floors is 551 tons, the top floor is 583 tons, and the two underground floors each have a mass of 535 tons.

By analyzing the structural model, the differences in seismic damage induced by artificial and natural ground motions can be comprehensively assessed, thereby verifying the efficacy and reliability of artificially synthesized ground motions in replicating realistic seismic responses.

To accurately simulate the dynamic characteristics of the steel structure, the BEAM188 elements were used to model beams and columns, while the SHELL181 elements were used to simulate the slabs. The in-plane stiffness of the slabs was adjusted to approach infinity by appropriately modifying the material parameters, ensuring compliance with the assumption of infinite in-plane stiffness. The additional mass of the structure was modeled using MASS21 elements, and the modal analysis employed the subspace iteration method to determine the natural frequencies and their corresponding mode shapes.

To accurately simulate the nonlinear seismic response of steel structures—particularly the hysteretic energy dissipation and post-yield stiffness behavior of steel—the material model in this study adopts the bilinear kinematic hardening model available in ANSYS. This model incorporates a defined tangent modulus, allowing for a realistic representation of cyclic loading behavior, including strain hardening and hysteretic path tracking. Specifically, the initial elastic modulus of steel is set to 200 GPa, while the post-yield hardening modulus is defined as 2 GPa to reflect the residual stiffness and energy dissipation capacity after yielding.

To capture the nonlinear response of the structure under seismic loading, concentrated plastic hinges are introduced at critical locations. Plastic zones are primarily assigned to beam ends and column bases in each span. The nonlinear behavior of beams and columns is modeled using a concentrated plasticity approach by implementing nonlinear spring elements (COMBIN39) to simulate the moment–rotation behavior across yield, hardening, and ultimate states.

The key parameters for the plastic hinges are determined based on recommendations from FEMA 356 and ASCE 41 (FEMA 2000, ASCE 2017), adjusted for the specific geometric and material properties of the structural

Table 6 First five natural frequencies of the structure

Mode number	Frequency (Hz)	
	This model	Reference (Powell and Allahabadi 1988)
1	0.260	0.261
2	0.840	0.753
3	1.392	1.300
4	1.741	1.830
5	2.489	2.400

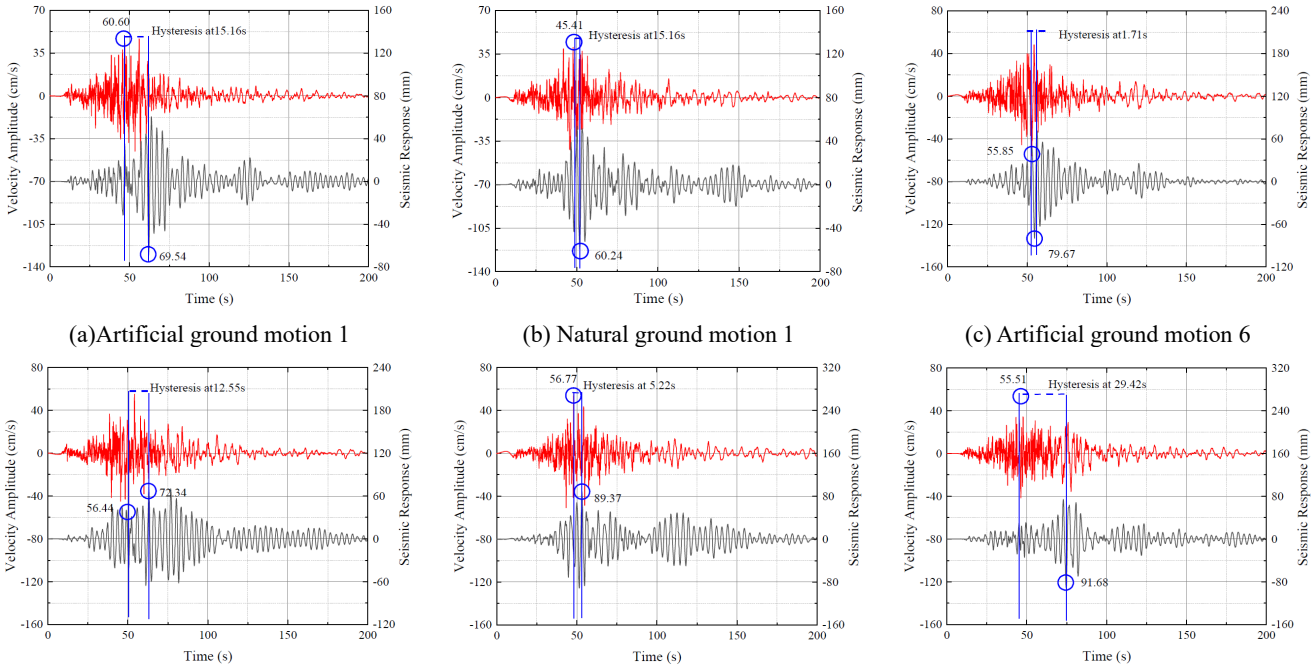
components. The yield rotation is set at 0.01 rad, the ultimate rotation at 0.05 rad, with a post-yield stiffness of 5% of the initial stiffness. The hysteresis model adopts a bilinear hardening rule to reflect the cyclic energy dissipation characteristics.

Using the aforementioned methods, the first five natural frequencies of the structure were successfully determined, with the results shown in Table 6. A comparison of the calculated natural frequencies with data from relevant literature showed good consistency, indicating that the modeling approach employed in this study is both reasonable and accurate. This provides a reliable foundation for subsequent structural dynamic analysis.

Based on the aforementioned parameter settings, eleven sets of generated ground motions were applied in the east-west direction of the structure. ANSYS was used to perform a detailed analysis of the seismic response at key structural nodes. Table 7 presents the input ground motion parameters as well as the response results of key structural nodes. To further explore the differences in structural impact between

Table 7 Key seismic responses of the structure

Ground motion ID	Ground motion information		Structural response	
	PGA (g)	PGV (cm/s)	Top floor displacement (mm)	Maximum inter-story drift angle
1	0.6565	49.39	69.54	1.21E-02
Natural ground motion 1	0.6588	45.41	60.24	1.19E-02
2	0.6736	50.33	69.48	1.21E-02
3	0.6745	59.34	71.15	1.23E-02
4	0.6763	51.60	76.66	1.26E-02
5	0.6903	57.63	78.52	1.29E-02
6	0.6972	55.85	79.67	1.29E-02
Natural ground motion 2	0.7022	56.44	72.34	1.27E-02
7	0.7129	52.37	79.76	1.36E-02
8	0.7190	60.84	80.88	1.40E-02
9	0.7259	43.23	83.81	1.46E-02
10	0.7399	63.97	90.38	1.49E-02
11	0.7475	56.77	91.22	1.62E-02
Natural ground motion 3	0.7476	55.51	91.68	1.61E-02



(a)Artificial ground motion 1 (b) Natural ground motion 1 (c) Artificial ground motion 6

(d) Natural ground motion 2 (e) Artificial ground motion 11 (f) Natural ground motion 3

Fig. 10 Ground motion response time history

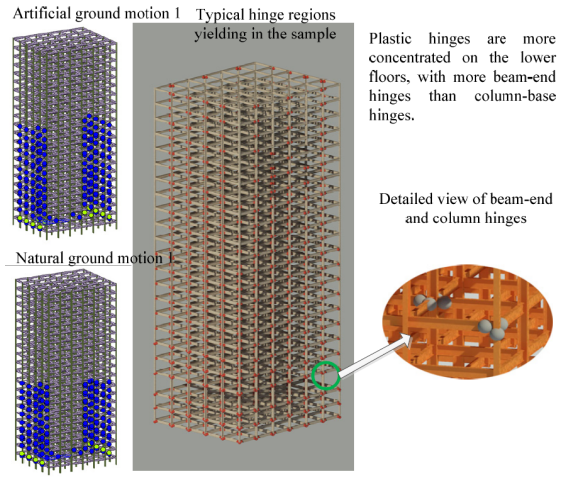


Fig. 11 Distribution of plastic hinges

artificial and natural ground motions, three sets of artificial ground motions were selected, and natural ground motions with similar PGA values were chosen from the ground motion database for comparison.

By analyzing the maximum inter-story drift angles at each floor and the inter-story drift response of the top floor, Fig. 10 illustrates the differences in structural response under various intensities of artificial and natural ground motions.

Fig. 11 illustrates the distribution of plastic regions in the structure under the action of Artificial Ground Motion No. 1 and Natural Ground Motion No. 1. As shown, plastic hinges are primarily concentrated at the beam ends in the lower stories and outer spans, which aligns with the typical stress distribution pattern of structures subjected to seismic excitation. This modeling approach effectively balances overall computational efficiency with the accurate representation of nonlinear deformation and energy dissipation in critical structural components.

Analysis of the plastic hinge development regions under both artificial and natural ground motions reveals that the nonlinear response of the structure is primarily concentrated in certain critical components. In Fig. 11, blue indicates plastic hinges that have reached the yield rotation threshold (0.01 rad), while yellow represents hinges that have reached a rotation of 0.03 rad. Using a concentrated plasticity modeling approach, combined with hysteresis curves and displacement response data, this study identifies and summarizes the locations and development patterns of plastic hinges in the structure: Vertical distribution pattern: Plastic hinges predominantly appear in the lower stories, especially at locations where the first story connects to the foundation. The base columns and lower-story beam ends are typically the earliest to yield, due to significantly amplified seismic shear forces in the lower part of the structure, leading to localized stress concentration. Component type pattern: The structure exhibits a pronounced "strong-column, weak-beam" behavior. This is particularly evident in the outer frame bays, which experience greater bending demands and serve as the primary energy dissipation zones. Yielding among column elements is mostly confined to the base level, while mid-to-upper story columns largely remain elastic, reflecting compliance with seismic design principles. Planar distribution pattern: Plastic hinges tend to form first in the perimeter frames. The corner columns, which are made of box sections and possess higher strength, generally remain elastic, while the plastic hinges are concentrated in the beams and columns along the outer frames. Interior spans and inner columns yield later or remain elastic, demonstrating the combined influence of structural configuration and stress demand on hinge formation. Variation with ground motion intensity: Under Artificial Ground Motion 1 and Natural Ground Motion 1, the east–west components exhibit relatively higher intensities, leading to more pronounced development of plastic hinges at beam ends in the east–west direction. In some cases, the rotation of these hinges reaches up to 0.03 rad. At lower ground motion intensities, plastic hinges are mostly confined to beam ends. However, under higher intensity shaking (e.g., rare earthquake scenarios), additional hinges appear at base columns and joints, indicating a coordinated expansion of yielding across vertical and horizontal structural elements.

In summary, the Benchmark steel frame structure displays a distinct plastic hinge distribution characterized by "interstory concentration, perimeter priority, and beam-end dominance." This pattern is consistent with established findings in the literature regarding the yielding mechanisms of high-rise steel structures under seismic loads (e.g., FEMA 355E, ASCE 41). Moreover, the plastic hinge responses under artificial ground motions closely match those induced by natural earthquakes, confirming that the artificially generated motions can effectively replicate the damage characteristics of real seismic events. These observations provide a valuable foundation for subsequent damage assessment and performance enhancement strategies based on yielding region analysis.

The analysis of structural responses under artificial and natural ground motions reveals the following phenomena: although there are significant differences in waveform characteristics between the artificially generated ground motions (designed to match acceleration response spectra and PGA) and the natural ground motions from the database, the structural damage effects induced by both types of ground motions are highly consistent. Specifically, the average difference in top floor displacement response is only 7.35%, while the difference in maximum inter-story drift angle is just 0.99%. This indicates that, under comparable seismic intensities, artificial ground motions can effectively replicate the structural damage behavior observed under natural ground motions.

The structural damage changes significantly with increasing seismic intensity (PGA). As the PGA increases from 0.65g (corresponding to a maximum inter-story drift angle of approximately 1.20E-02) to 0.75 g (corresponding to a maximum inter-story drift angle of approximately 1.61E-02), the level of structural damage also intensifies. This finding demonstrates that artificial ground motions can accurately simulate the damage trend induced by natural ground motions, exhibiting a high degree of similarity in

structural response.

Further analysis shows that both artificial and natural ground motions induce structural damage that involves an energy accumulation phase lasting several seconds to over twenty seconds. Structural damage triggered by ground motions is primarily controlled by the input energy and its accumulation effects. Since artificial ground motions are comparable to natural ground motions in terms of PGA and energy input, the resulting structural responses are highly consistent.

These findings indicate that artificially generated ground motions, designed based on PGA and acceleration response spectra, can effectively simulate the damage characteristics of natural ground motions. This is especially evident in key structural response indicators – such as top floor displacement and maximum inter-story drift angle – where a strong level of agreement is observed.

8. Seismic damage analysis

To comprehensively evaluate the structural performance under seismic loading, this study adopts the classical Park–Ang damage index to quantify the damage level of structural components during earthquake excitation. This index integrates two critical factors—maximum deformation demand and hysteretic energy dissipation capacity—allowing for a realistic representation of the damage evolution throughout the seismic event. As one of the most widely used empirical damage models in the field of earthquake engineering, the Park–Ang index provides a robust basis for assessing structural integrity and seismic resilience.

The expression for the Park–Ang damage index is as follows

$$D = \frac{\delta_m}{\delta_u} + \beta \cdot \frac{E_h}{F_y \delta_u} \tag{14}$$

Where δ_m is the maximum displacement experienced by the component during the earthquake, δ_u is the ultimate displacement capacity under monotonic loading, E_h is the cumulative hysteretic energy dissipated during the entire seismic process, and F_y is the yield strength of the component. The coefficient β is an empirical calibration factor that reflects the relative weight of energy dissipation in the total damage. Based on established literature (Park and Ang 1985), this study adopts a value of $\beta = 0.1$.

The damage index of the entire steel frame is calculated based on the damage indices of individual components. The weight of each component's damage index is determined according to its cumulative hysteretic energy dissipation. The overall damage index of the steel frame is defined by Eq. (15)

$$\begin{cases} D_w = \sum W_j \times D_j \\ W_j = \dfrac{E_j}{\sum E_j} \end{cases} \tag{15}$$

Where D_w denotes the overall damage index of the steel frame; D_j is the damage index of the j-th component; W_j represents the weight of the j-th component within the entire steel frame; E_j is the cumulative hysteretic energy of the j-th component; and $\sum$ denotes summation.

Compared to using only geometric response parameters such as top-floor displacement or inter-story drift angle, the Park–Ang index offers greater physical significance by accounting for both structural ductility and energy dissipation capacity. It is particularly suitable for assessing

Table 8 Damage index range table

Damage level	Damage index range	Engineering meaning
Slight damage	D < 0.1	Essentially elastic, no repair required
Moderate damage	0.1 ≤ D < 0.4	Localized yielding, inspection and repair recommended
Severe damage	0.4 ≤ D < 0.8	Significant damage, strengthening required
Collapse-level damage	D ≥ 0.8	Component failure; reconstruction or major intervention required

Table 9 Structural maximum inter-story drift angle and structural damage index

Ground motion ID	Ground motion information		Structural response	
	PGA (g)	PGV (cm/s)	Inter-story drift angle	Damage index
1	0.6565	49.39	1.21E-02	0.27
Natural ground motion 1	0.6588	45.41	1.19E-02	0.28
2	0.6736	50.33	1.21E-02	0.28
3	0.6745	59.34	1.23E-02	0.28
4	0.6763	51.6	1.26E-02	0.27
5	0.6903	57.63	1.29E-02	0.28
6	0.6972	55.85	1.29E-02	0.28
Natural ground motion 2	0.7022	56.44	1.27E-02	0.30
7	0.7129	52.37	1.36E-02	0.31

Table 10 Key parameters of sample ground motions

Ground motion ID	Observation point name	Earthquake name	Epicentral distance	PGV(cm/s)			PGA(g)		
				East-west direction	North-south direction	Vertical direction	East-west direction	North-south direction	Vertical direction
N1	F044	Hualien Earthquake	18.91	22.99	17.03	6.14	0.25	0.22	0.08
N2	F045	Hualien Earthquake	60.24	59.51	62.97	17.39	0.36	0.36	0.13
N3	F048	Hualien Earthquake	28.05	53.95	33.94	25.11	0.42	0.17	0.22
N4	F049	Hualien Earthquake	26.95	46.22	36.79	17.72	0.38	0.34	0.15
N5	F067	Aftershock 1	20.58	33.18	18.63	8.43	0.18	0.19	0.06
N6	F068	Aftershock 1	18.6	45.7	40.74	15.02	0.27	0.23	0.1
N7	F076	Aftershock 1	24.6	46	28.6	18.7	0.19	0.21	0.08
N8	E072	Aftershock 2	103.53	25.02	43.52	9.91	0.13	0.12	0.04
N9	E075	Aftershock 2	75.57	36.94	28.5	11.3	0.29	0.22	0.08
N10	EGFH	Aftershock 2	27.15	15.86	15.26	5.66	0.16	0.13	0.06

cumulative damage effects under strong near-fault ground motion input. In this study, the Park–Ang index is applied to evaluate and compare the damage of typical components, further revealing the influence of seismic motion characteristics on structural nonlinear response and damage distribution.

To assess the structural damage levels under different ground motions, this study adopts the widely used Park–Ang damage index in earthquake engineering as the evaluation criterion. This index combines the maximum deformation demand and hysteretic energy dissipation capacity of structural components, effectively reflecting their nonlinear behavior and cumulative damage. Based on literature recommendations and code guidance (Park and Ang 1985), the structural damage is categorized into four levels using the classification criteria listed in the Table 8.

Based on the above calculation results and the classification criteria of the Park–Ang damage index, the structural responses under three sets of natural ground motions and eleven sets of artificial ground motions are converted into corresponding damage indices, as presented in Table 9.

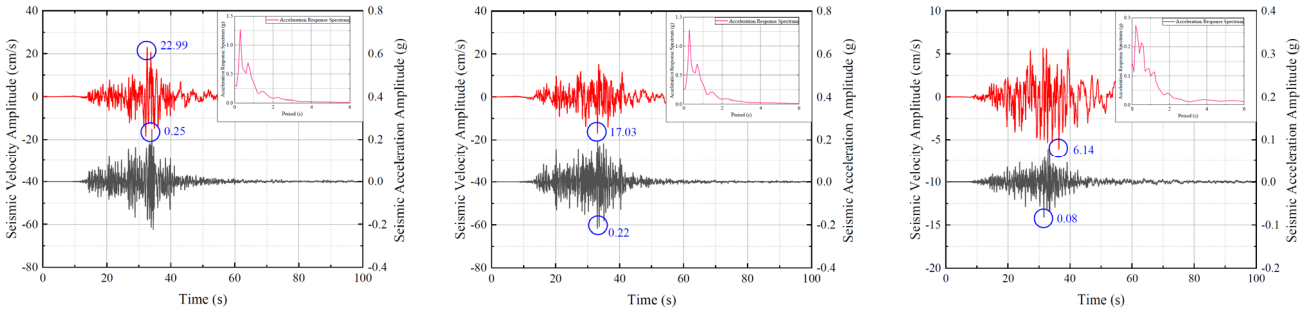
According to the structural response results summarized in Table 9 under both natural and artificial ground motions, the structural damage indices calculated using the Park–Ang model range from 0.27 to 0.37. Overall, this indicates that the structure remains in a moderate damage state under the considered seismic inputs. A comparison between the responses to natural and artificial ground motions reveals a high degree of consistency in both the maximum inter-story drift angle and damage index. For instance, under comparable seismic intensity, the damage indices for Natural Ground Motion 1 and Artificial Ground Motion 1 are nearly identical (0.27–0.28), confirming the effectiveness of the artificial ground motion generation method in maintaining seismic demand consistency. As PGA and PGV increase, both natural and artificial ground motions exhibit a similar trend of increasing inter-story drift and damage indices. This trend is remarkably consistent across both motion types. Notably, Artificial Ground Motion 11 yields the highest damage index of 0.37, approaching the threshold of severe damage, suggesting that strong ground motions induce significant nonlinear response and damage accumulation in the structure.

In conclusion, the artificial ground motions generated using the KPCA-GA method not only closely match natural ground motions in terms of seismic demand metrics, but also show strong agreement in terms of structural damage response. This validates the method's ability to produce artificial ground motions that are not only spectrum-compatible but also damage-consistent with real earthquake records.

9. Simulation and analysis of natural ground motions

To further evaluate the quality and accuracy of the proposed ground motion generation method, natural ground motions from the database were simulated, and a comparative analysis was conducted between the artificial and natural ground motions to verify the effectiveness of the generation method. In this section, ten sets of ground motions were selected from a database that includes one mainshock and five aftershocks as synthesis targets. Remove these ten ground motions from the earthquake training database and retrain the model. Key parameters of the ground motions are summarized in Table 10, and the acceleration time histories, velocity time histories, and acceleration response spectra of four ground motions from the mainshock are shown in Fig. 12. These comparisons allow for a direct analysis of the differences between artificial and natural ground motions in various seismic characteristics. Such analysis not only helps validate the quality of the proposed ground motion generation method but also allows for further evaluation of its applicability and reliability in earthquake engineering.

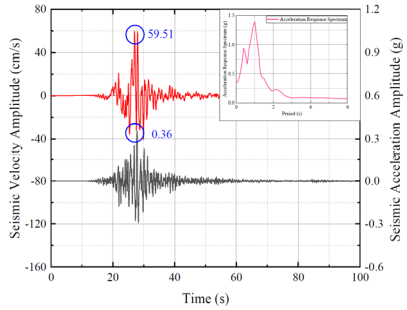
The ground motion data presented in Table 7 cover multiple observation points ranging from areas near the epicenter to those in the far-field region, illustrating variations in ground motion intensity and directionality across different locations. The recorded epicentral distances range from 18.60 km to 103.53 km, encompassing both



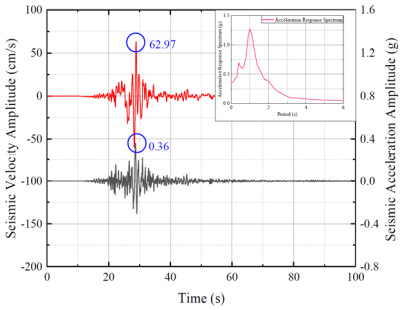
(a) East-west component of natural ground motion N1

(b) North-south component of natural ground motion N1

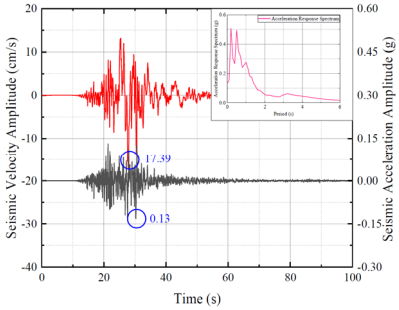
(c) Vertical component of natural ground motion N1



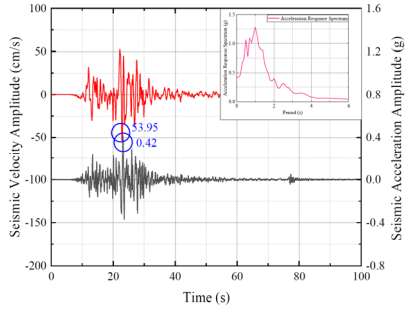
(d) East-west component of natural ground motion N2



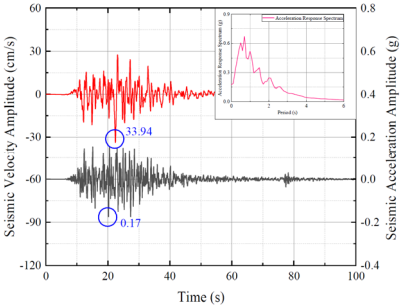
(e) North-south component of natural ground motion N2



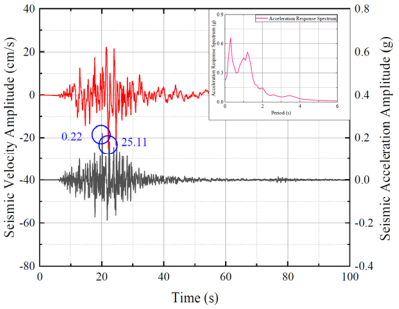
(f) Vertical component of natural ground motion N2



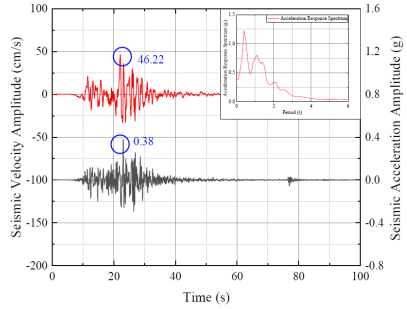
(g) East-west component of natural ground motion N3



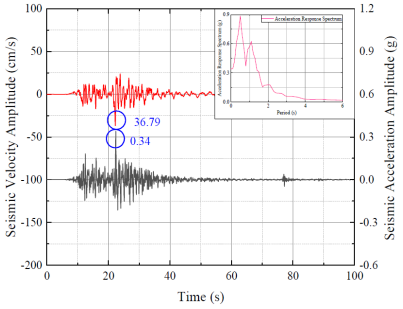
(h) North-south component of natural ground motion N3



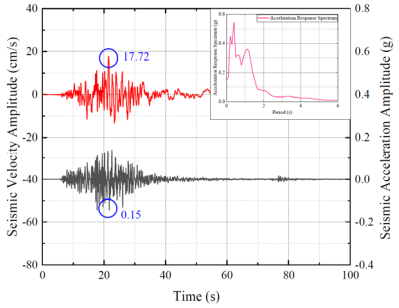
(j) East-west component of natural ground motion N4



(j) East-west component of natural ground motion N4



(k) North-south component of natural ground motion N4

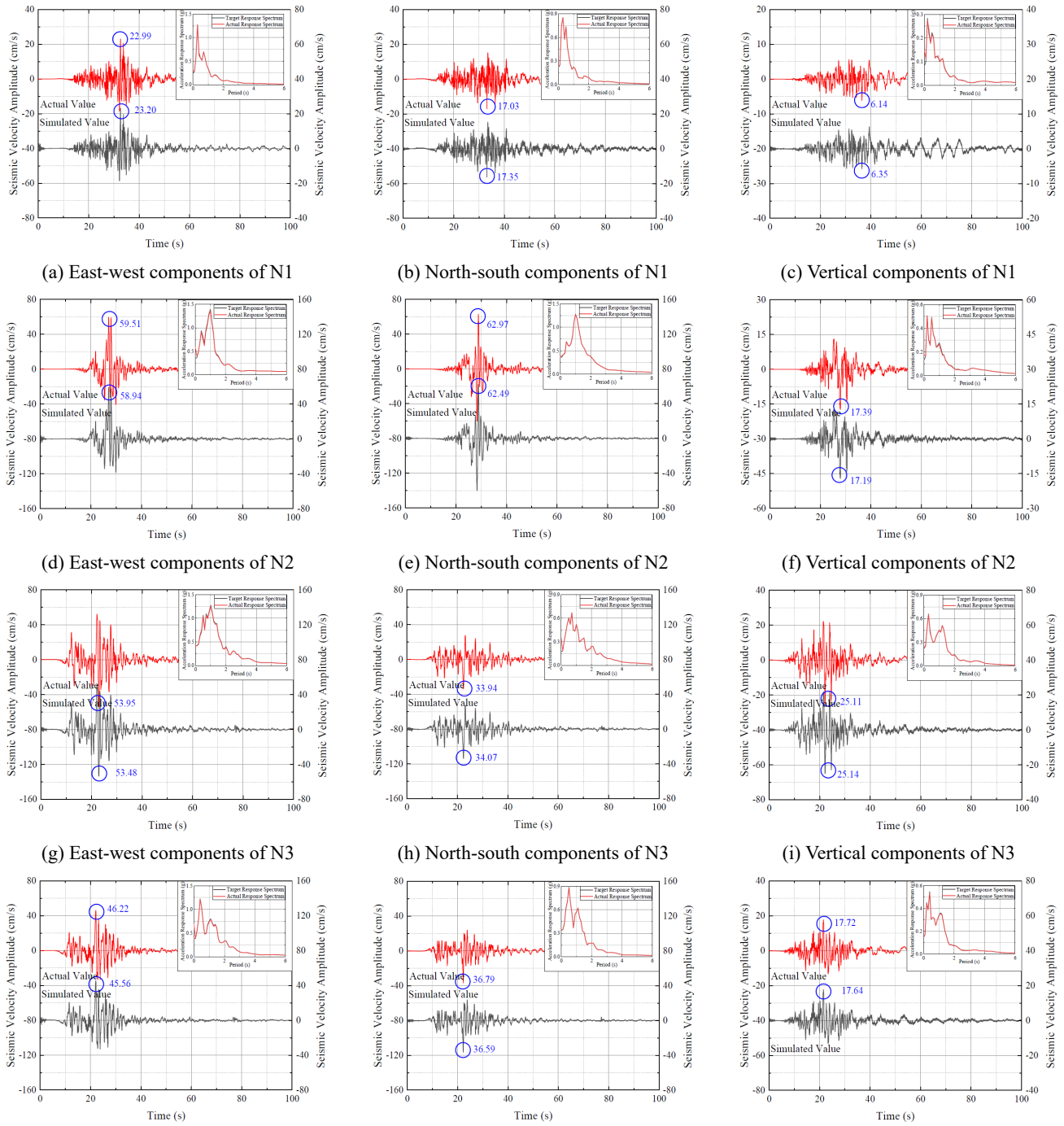


(l) Vertical component of natural ground motion N4

Fig. 12 Time histories of sample natural ground motions

near-field regions (e.g., N1, N6) and far-field regions (e.g., N8, N9), thus reflecting the characteristics of ground motions at different distances. The data include records of one mainshock and two aftershocks, fully capturing the complexity of multi-source ground motions. The spatial distribution of ground motion intensity is extensive, with PGA and PGV covering multiple energy levels, providing a comprehensive data foundation for ground motion generation and analysis.

In this section, the acceleration response spectra of the ground motion samples are set as the primary optimization target, while PGV is set as the secondary target, aiming to generate artificial ground motions that meet engineering requirements. To enhance the accuracy of the generated



(a) East-west components of N1 (b) North-south components of N1 (c) Vertical components of N1

(d) East-west components of N2 (e) North-south components of N2 (f) Vertical components of N2

(g) East-west components of N3 (h) North-south components of N3 (i) Vertical components of N3

(j) East-west components of N4 (k) North-south components of N4 (l) Vertical components of N4

Fig. 13 Comparison of velocity time histories between artificial and natural ground motions

results, each directional component of the ground motion was generated separately. To ensure the validity of the assessment of artificial ground motions, the ten selected ground motions used in this section were removed from the ground motion database to avoid any interference from duplicated data.

The specific parameters of the genetic algorithm used in the generation process were set as follows: a population size of 200, a crossover ratio of 0.08, a mutation probability of 0.05, and a maximum iteration count of 3,000. Using this optimization method, artificial ground motions that closely match the target characteristics were successfully generated, as shown in Fig. 13, with key parameters summarized in Table 11.

Based on the analysis of the above simulation results, it can be concluded that artificial ground motions can accurately simulate the characteristics of natural ground motions. Observations of the velocity time histories for the first four sets of ground motions indicate that the artificial ground motions effectively replicate the waveform features of natural ground motions. This high accuracy in simulation can be attributed to the characteristic waveforms present in

Table 11 Key parameters of artificial ground motion and relative error

Ground motion ID	PGV simulation result (cm/s)			PGV relative error (%)			PGA simulation result (g)			PGA relative error (%)			Acceleration response spectrum relative error (%)		
	East-west direction	North-south direction	Vertical direction	East-west direction	North-south direction	Vertical direction	East-west direction	North-south direction	Vertical direction	East-west direction	North-south direction	Vertical direction	East-west direction	North-south direction	Vertical direction
N1	23.20	17.35	6.35	0.91	1.85	3.35	0.24	0.22	0.08	4.83	1.62	2.89	0.18	0.02	1.41
N2	58.94	62.49	17.19	0.96	0.76	1.15	0.36	0.37	0.13	0.24	3.06	4.11	0.60	0.01	0.21
N3	53.48	34.07	25.14	0.87	0.39	0.12	0.41	0.18	0.22	2.12	3.21	2.76	0.06	0.90	0.04
N4	45.56	36.59	17.64	1.43	0.55	0.46	0.37	0.34	0.15	2.18	0.65	0.26	1.09	0.01	0.03
N5	32.54	18.58	8.16	1.94	0.27	3.22	0.17	0.20	0.06	3.18	4.85	2.61	1.92	0.30	2.13
N6	44.69	40.63	15.26	2.20	0.28	1.56	0.26	0.23	0.09	2.30	1.19	1.26	0.25	0.22	0.30
N7	45.16	28.43	19.26	1.82	0.60	3.01	0.20	0.21	0.08	4.32	2.02	2.07	2.12	1.30	0.38
N8	24.60	43.63	9.67	1.68	0.24	2.42	0.12	0.13	0.04	4.97	7.07	3.28	6.59	1.27	0.18
N9	36.42	28.49	11.12	1.41	0.04	1.60	0.29	0.22	0.08	0.20	1.10	0.50	0.26	0.26	0.59
N10	15.25	15.45	5.76	3.86	1.21	1.74	0.16	0.14	0.06	0.68	7.35	4.08	4.37	1.46	4.31

specific earthquakes, which are compressed using KPCA. During the generation process, the genetic algorithm is capable of combining these compressed characteristic waveforms effectively. Therefore, even though the artificial ground motions are generated with acceleration response spectra as the primary objective, certain waveform segments still exhibit high similarity to natural ground motions.

Further comparisons of relative errors between artificial and natural ground motions in terms of PGA, PGV, and acceleration response spectra reveal the following: the average relative error for the acceleration response spectra is 1.09%, while the relative error for PGV is 1.40%, both of which are controlled within 1.5%. This indicates that the artificial ground motions can accurately replicate the overall characteristics of natural ground motions. However, the average relative error for PGA is 2.46%, which is relatively higher. This is primarily due to the fact that, in the artificial ground motion generation process, the acceleration response spectrum is set as the primary objective and PGV as the secondary objective, while PGA is derived from PGV, leading to an increase in error.

In general, the artificial ground motions demonstrate high precision in waveform, PGV, and acceleration response spectra, effectively replicating the response characteristics of natural ground motions. This method provides a reliable tool for seismic design and structural evaluation in practical engineering, proving its broad potential for application in earthquake engineering.

10. Conclusions

This study proposes a ground motion synthesis method that integrates KPCA with a genetic algorithm to effectively simulate seismic ground motions exhibiting nonstationarity and spatial variability in strong earthquake regions. The results demonstrate that, while introducing a certain degree of uncertainty through the combination of multiple basis waveforms, the method successfully preserves the natural characteristics of the simulated ground motions. Meanwhile, it achieves accurate matching of key seismic features such as the target response spectrum, PGA, and PGV, thereby offering strong support for seismic design applications.

High Target Feature Matching: Through the simulation analysis of the Hualien earthquake in Taiwan, the generated artificial ground motions effectively matched the target response spectra and PGA, with errors controlled within a reasonable range and a maximum error of less than 7%.

Strong Similarity in Power Spectral Density: Further analysis of PSD indicates that the artificial ground motions exhibit high similarity to natural ground motions in terms of frequency domain characteristics, effectively preserving the energy distribution features of natural ground motions across different frequency bands.

High Waveform Feature Reproduction: Targeting the ground motion records from the Hualien earthquake in Taiwan, the artificially synthesized ground motions demonstrated high accuracy in waveform, PGV, and acceleration response spectra, effectively replicating the response characteristics of natural ground motions.

High Consistency in Damage Patterns: Structural damage analysis results indicate that artificial ground motions and natural ground motions show high consistency in the seismic response and damage characteristics of typical frame structures, especially in key indicators such as top floor displacement and maximum inter-story drift angle.

In summary, the proposed method employs KPCA to extract the primary characteristic waveforms of ground motions in the target region and uses a genetic algorithm for optimized generation. By introducing a certain degree of uncertainty while preserving the waveform, time-frequency characteristics, and damage features of natural ground motions, the method synthesizes artificial ground motions based on parameters such as the target response spectrum,

PGA, and PGV. This approach provides a reliable tool for seismic design in practical engineering applications and shows great potential for ground motion simulation in strong earthquake regions. Future research may further explore the applicability of this method under diverse site conditions and varying earthquake magnitudes to enhance its generalizability and robustness in earthquake engineering practices.

Data availability statement

The data that support the findings of this study are openly available in Central Weather Bureau Ground Motion Database at https://gdms.cwa.gov.tw/.

Formatting of funding sources

List funding sources in this standard way to facilitate compliance with funder's requirements:

Funding: This work was supported by the National Natural Science Foundation of China [grant numbers 52208461]; Research on the Collaborative Construction of a Resilient Coastal Protection and Intelligent Emergency Response System in Yantai City [grant numbers: YTSK2025-125].

References

Abrahamson, N.A. (1992), "Non-stationary spectral matching", *Seismol. Res. Lett.*, **63**(1), p. 30.

Ambrosini, D., Mendoza, P.M.B. and Luccioni, B.M. (2021), "Energy dissipation by friction for sliding blocks subjected to near-fault seismic base motion", *Smart Struct. Syst., Int. J.*, **27**(5), 819-836. https://doi.org/10.12989/sss.2021.27.5.819

American Society of Civil Engineers (ASCE) (2017), Seismic Evaluation and Retrofit of Existing Buildings, (ASCE/SEI 41-17), Reston, VA, USA: American Society of Civil Engineers.

Aristeidou, S., Shahnazaryan, D. and O'Reilly, G.J. (2024), "Artificial neural network-based ground motion model for next-generation seismic intensity measures", *Soil Dyn. Earthq. Eng.*, **184**, p. 108851. https://doi.org/10.1016/j.soildyn.2024.108851

Chang, Y.H., Tsai, C.C., Ge, L. and Park, D. (2022), "Influence of horizontally variable soil properties on nonlinear seismic site response and ground motion coherency", *Earthq. Eng. Struct. Dyn.*, **51**(3), 704-722. https://doi.org/10.1002/eqe.3587

Chen, Q., Chen, X. and Zhao, Z. (2024), "Simulation of Near-Fault Pulse-Like Ground Motion and Investigation of Seismic Response Characteristics in Subway Stations", *J. Earthq. Eng.*, **28**(13), 3648-3680. https://doi.org/10.1080/13632469.2024.2356879

Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. (2002), "A fast and elitist multiobjective genetic algorithm: NSGA-II", *IEEE Transact. Evolut. Computat.*, **6**(2), 182-197. https://doi.org/10.1109/4235.996017

Ding, J.Y., Feng, D.C., Brunesi, E., Parisi, F. and Wu, G. (2023), "Efficient seismic fragility analysis method utilizing ground motion clustering and probabilistic machine learning", *Eng. Struct.*, **294**, p. 116739. https://doi.org/10.1016/j.engstruct.2023.116739

Fayaz, J. and Galasso, C. (2023), "A deep neural network framework for real-time on-site estimation of acceleration response spectra of seismic ground motions", *Comput.-Aided Civil Infrastr. Eng.*, **38**(1), 87-103. https://doi.org/10.1111/mice.12830

Federal Emergency Management Agency (FEMA) (2000), Prestandard and Commentary for the Seismic Rehabilitation of Buildings (FEMA 356), Washington, DC, USA: Federal Emergency Management Agency.

Feng, K., Wang, C. and Li, Q. (2024), "Evaluating the Role of Transportation System in Community Seismic Resilience", *Resilient Cities Struct.*, **3**(3), 65-78. https://doi.org/10.1016/j.rcns.2024.05.003

Gasparini, D.A. and Vanmarcke, E.H. (1976), "Simulated earthquake motions compatible with prescribed response spectra. Research Report R76-4", Department of Civil Engineering, MIT, Cambridge, USA.

Hakamian, I., Taghikhani, K., Manouchehri, N. and Memarpour, M.M. (2023), "Soil-structure interaction effects on collapse probability of the RC buildings subjected to far and near-field ground motions", *Earthq. Struct., Int. J.*, **25**(2), 99-112. https://doi.org/10.12989/eas.2023.25.2.099

Hou, R. and Wang, J. (2024), "Uncertainty of Site Period on The Variability of the Site Period-Based Ground Motion Prediction Equations", *Soil Dyn. Earthq. Eng.*, **176**, p. 108345. https://doi.org/10.1016/j.soildyn.2023.108345

Ji, K., Ren, Y., Wen, R. and Kuo, C.H. (2019), "Near-field velocity pulse-like ground motions on February 6, 2018 MW6. 4 Hualien, Taiwan earthquake and structural damage implications", *Soil Dyn. Earthq. Eng.*, **126**, p. 105784. https://doi.org/10.1016/j.soildyn.2019.105784

Kanai, K. (1957), "Semi-empirical formula for the seismic characteristics of the ground", *Bull. Earthq. Res. Inst.*, **35**, 309-25.

Kuo, C.H., Huang, J.Y., Lin, C.M., Hsu, T.Y., Chao, S.H. and Wen, K.L. (2019), "Strong ground motion and pulse-like velocity observations in the near-fault region of the 2018 Mw 6.4 Hualien, Taiwan, earthquake", *Seismol. Res. Lett.*, **90**(1), 40-50. https://doi.org/10.1785/0220180195

Li, B., Cai, Z. and Duan, Z. (2022), "Generating Spectrum-Matched Bidirectional Time Histories using Hilbert-Huang Transform", *J. Build. Eng.*, **58**, p. 105033. https://doi.org/10.1016/j.jobe.2022.105033

Lilhanand, K. and Tseng, W.S. (1988), "Development and application of realistic earthquake time histories compatible with multiple-damping design spectra", *Proceedings of the 9th World Conference on Earthquake Engineering*, **2**, 819-824.

Liu, Z. (2025), "Near-Fault Ground Motion Identification Method Based on Local Mean Decomposition", *J. Earthq. Eng.*, 1-58. https://doi.org/10.1080/13632469.2025.2505976

Liu, C., Macedo, J., Gulerce, Z., Abrahamson, N., Kottke, A., Akbas, B. and Arda Ozacar, A. (2024a), "Estimating systematic source, site, and path effects in nonergodic ground-motion models: Insights from the Turkish ground-motion database", *Bull. Seismol. Soc. Am.*, **114**(6), 3024-3040. https://doi.org/10.1785/0120230288

Liu, Z., Ma, X. and Lv, J. (2024b), "Seismic response of a large-span steel truss arch bridge under nonuniform near-fault ground motions", *Buildings*, **14**(8), p. 2308. https://doi.org/10.3390/buildings14082308

Miller, M.S., Pickle, R., Murdie, R., Yuan, H., Allen, T.I., Gessner, K., Kennett, B.L. and Whitney, J. (2023), "Southwest Australia Seismic Network (SWAN): Recording earthquakes in Australia's most active seismic zone", *Seismol. Soc. Am.*, **94**(2A), 999-1011. https://doi.org/10.1785/0220220323

Park, Y.J. and Ang, A.H.-S. (1985), "Mechanistic seismic damage model for reinforced concrete", *J. Struct. Eng.*, **111**(4), 722-739. https://doi.org/10.1061/(ASCE)0733-9445(1985)111:4(722)

Powell, G.H. and Allahabadi, R. (1988), "Seismic Damage

Prediction by Deterministic Methods: Concepts and Procedures", *Earthq. Eng. Struct. Dyn.*, **16**(5), 719-734. https://doi.org/10.1002/eqe.4290160507

Qin, H. and Li, L. (2023), "Efficient ground motion conditional simulation method for nonstationary multivariate processes based on Kriging method and orthogonal discrete wavelet transform", *Soil Dyn. Earthq. Eng.*, **173**, p. 108149. https://doi.org/10.1016/j.soildyn.2023.108149

Rosipal, R. and Girolami, M. (2001), "An expectation–maximization approach to nonlinear component analysis", *Neural Computat.*, **13**(3), 505-510. https://doi.org/10.1162/089976601300014439

SAC Joint Venture / FEMA (2000), FEMA-355E: State of the Art Report on Past Performance of Steel Moment-Frame Buildings in Earthquakes, Federal Emergency Management Agency, Washington, DC, USA.

Schölkopf, B., Smola, A. and Müller, K.-R. (1998), "Nonlinear component analysis as a kernel eigenvalue problem", *Neural Computat.*, **10**(5), 1299-1319. https://doi.org/10.1162/089976698300017467

Tajimi, H. (1960), "A statistical method of determining the maximum response of a building structure during an earthquake", *Proceedings of the 2nd World Conference on Earthquake Engineering*, Tokyo, Japan, pp. 781-797.

Takemura, S., Obara, K., Shiomi, K. and Baba, S. (2022), "Spatiotemporal Variations of Shallow Very Low Frequency Earthquake Activity Southeast off The Kii Peninsula, along The Nankai Trough, Japan", *J. Geophys. Res.: Solid Earth*, **127**(3), p. e2021JB023073. https://doi.org/10.1029/2021JB023073

Wani, F.M. and Vemuri, J. (2024), "Simulation of pulse-like ground motions with directionality effect for the 2001 Mw 7.6 Bhuj, India earthquake and sensitivity analysis of uncertain model input parameters", *Soil Dyn. Earthq. Eng.*, **183**, p. 108773. https://doi.org/10.1016/j.soildyn.2024.108773

Wei, M., Song, C. and Hu, X. (2023), "A hybrid deep learning model for predicting the residual displacement spectra under near-fault ground motions", *Earthq. Struct., Int. J.*, **25**(1), 15-26. https://doi.org/10.12989/eas.2023.25.1.015

Wen, P., Liu, R. and Wen, R. (2023), "Wavelet Packets-Based Simulation of Non-Stationary Multivariate Ground Motions", *Probabil. Eng. Mech.*, **74**, p. 103495. https://doi.org/10.1016/j.probengmech.2023.103495

Wilson, D.C., Hutt, C.R., Gee, L.S., Ringler, A.T. and Anthony, R.E. (2024), "Global seismic networks operated by the US Geological Survey", *Seismol. Res. Lett.*, **95**(3), 1578-1590. https://doi.org/10.1785/0220230178

Yu, J., Jiang, L., Zhou, W., Liu, X. and Lai, Z. (2024), "Evolutionary Power Spectrum Density of Earthquake-Induced Rail Geometric Irregularities", *Struct. Infrastr. Eng.*, **20**(4), 433-448. https://doi.org/10.1080/15732479.2022.2103155

Zhang, H. and Gu, C. (2022), "Hysteretic model of isolator gap damper system and its equivalent linearization for random earthquake response analysis", *Smart Struct. Syst., Int. J.*, **29**(3), 485-498. https://doi.org/10.12989/sss.2022.29.3.485

Zhang, Q., Li, P., Lang, X. and Miao, A. (2020), "Improved Dynamic Kernel Principal Component Analysis for Fault Detection", *Measurement*, **158**, p. 107738. https://doi.org/10.1016/j.measurement.2020.107738

Zhang, H., Deng, J. and Zhao, Y.G. (2023), "Damping modification factor of pseudo-acceleration spectrum considering influences of magnitude, distance and site conditions", *Earthq. Struct., Int. J.*, **25**(5), 325-342. https://doi.org/10.12989/eas.2023.25.5.325

Zhu, C., Cotton, F., Kwak, D.Y., Ji, K., Kawase, H. and Pilz, M. (2022), "Within-site variability in earthquake site response", *Geophys. J. Int.*, **229**(2), 1268-1281. https://doi.org/10.1093/gji/ggab481