



## Article

# Evaluation of Ground Motion Damage Potential with Consideration of Compound Intensity Measures Using Principal Component Analysis and Canonical Correlation Analysis

Tingting Liu <sup>1</sup>  and Dagang Lu <sup>2,\*</sup> 

<sup>1</sup> School of Civil Engineering and Architecture, Jiangsu University of Science and Technology, Zhenjiang 212100, China; liutingting@just.edu.cn

<sup>2</sup> School of Civil Engineering, Harbin Institute of Technology, Harbin 150090, China

\* Correspondence: ludagang@hit.edu.cn

**Abstract:** The primary motivation of this study is to develop a compound intensity measure (IM) to evaluate ground motion damage potential based on principal component analysis (PCA) and canonical correlation analysis (CCA). To illustrate this, this study examines the correlation among intragroup IMs and intergroup IMs, as well as the correlation between various IMs and response variables. A compound IM, which can be obtained by a linear combination of ten IMs in the log-scale, is utilized to measure the ground motion damage potential. Elastoplastic, bilinear and hysteretic models are utilized to determine peak deformation and hysteretic energy as the response variables of Single-Degree-of-Freedom (SDOF) systems. On the basis of the SDOF systems, the overall structural damage index is obtained by a nonlinear time–history analysis for two reinforced concrete moment frame systems. It is clear that the developed compound IM shows significantly high-level correlation with structural response. The better the correlations, the more one can measure the earthquake damage potential. A single IM alone inadequately characterizes structural damage, highlighting the necessity of multiple IMs to estimate the possibility of structural damage.

**Keywords:** ground motion damage potential; intensity measures; correlation coefficient; principal component analysis; canonical correlation analysis



**Citation:** Liu, T.; Lu, D. Evaluation of Ground Motion Damage Potential with Consideration of Compound Intensity Measures Using Principal Component Analysis and Canonical Correlation Analysis. *Buildings* **2024**, *14*, 1309. <https://doi.org/10.3390/buildings14051309>

Academic Editor: Siau Chen Chian

Received: 16 March 2024

Revised: 16 April 2024

Accepted: 20 April 2024

Published: 6 May 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Performance-based earthquake engineering involves four stages: hazard analysis, structural analysis, damage analysis and loss analysis. The utilization of IMs as a link between seismic hazard, structural and damage analysis has introduced significant uncertainties [1]. The seismic assessment of structures employs various IMs to estimate the demands induced by earthquakes. An evaluation of these adopted IMs is needed due to the uncertainties associated with those estimated by Giovenale et al. [2]. In order to capture the characterization of ground motions, various IMs are often considered to assess the severity of ground motions [3,4] and characterize the ground motion damage potential [5]. IMs are also employed for scaling the ground motion records, as demonstrated by Kurama and Farrow [6] and Akkar and Özen [7]. Consequently, the choice and classification of IMs assume a crucial role in the field of seismic design and analysis.

First-class IMs can be derived from acceleration time histories to characterize ground motion features [8], such as Peak Ground Acceleration (PGA), Peak Ground Velocity (PGV), Peak Ground Displacement (PGD) and so on. PGA and PGV are often employed as IMs for fragility analysis, such as in the study of Akkar et al. [9]. Other IMs describe the time histories obtained by using the structure-specific parameters, such as spectral acceleration (Sa), spectral velocity (Sv), spectral displacement (Sd) and so on. Sa has been widely used in incremental dynamic analysis (IDA) and hazard analysis. Traditional IMs containing PGA or Sa( $T_1$ ) may exhibit substantial record-to-record variability by forcing the histories

of many records to achieve structural response [10]. However,  $Sa(T_1)$  is more effective than PGA, as it better captures differences from record to record, elucidating structural seismic damage. Recent studies have demonstrated that  $Sa(T_1)$  may not be particularly efficient or sufficient for near-source ground motions [11] or certain for some long-period structures [12].  $Sa(T_1)$  does not account for stiffness and strength degradation affected by inelastic lengthening of the period under earthquake ground motions [13].  $S_{a,avg}(T_1, \dots, T_n)$  is identified as the optimal IM to predict inelastic structural response and account for the degrading behavior for Multi-Degree-of-Freedom (MDOF) systems with long periods [14]. Zhou et al. [15] identified optimal IMs for predicting the damage potential of mainshock–aftershock sequences. The optimal IMs are determined based on efficiency, proficiency, sufficiency and robustness [14,16,17]. Various IMs suggested by Hariri-Ardebili and Saouma [18], Kostinakis et al. [19] and Kostinakis and Athanatopoulou [20] are considered. Some IMs that are dependent on structure-specific parameters can reduce the dispersion of the nonlinearity response experienced by buildings under ground motion. Vector-valued IMs based on spectral shape to predict seismic fragility surfaces in reinforced concrete buildings were proposed by Zavala et al. [21]. The modified intensity measure method was used to improve accuracy in a seismic fragility analysis [22].

The Single-Degree-of-Freedom (SDOF) system serves as the simplest model applicable for studying the correlation between IMs and structural response in seismic analysis [23]. The IMs are categorized into acceleration-related, velocity-related and displacement-related groups [24–28], with a majority exhibiting high correlation among them. SDOF systems are employed as study objects to quantitatively elucidate the correlation between intragroup IMs and intergroup IMs, as well as the correlation between various IMs and response variables. The intragroup IM correlations are the correlation of acceleration-related, the correlation of velocity-related, or the correlation of displacement-related IMs. The intergroup IMs correlations are the correlation between acceleration-related and velocity-related, the correlation between velocity-related and displacement-related, or the correlation between displacement-related and acceleration-related IMs. However, an individual IM alone cannot encapsulate sufficient information about ground motion due to the randomness of earthquake ground motion and the complex mechanisms of structural damage. The earthquake damage potential refers to the likelihood of structural damage under ground motion, and the severity of earthquake damage is linked to the IMs. Various IMs are employed to characterize the ground motion damage potential. Zhai et al. [29] proposed a comprehensive method for estimating the damage potential of ground motions, which accounts for the correlation between IMs and response variables to select IMs describing earthquake damage potential based on experience. Building on this study, Zhai et al. [30] introduced a vector-valued IM to quantitatively reflect the damage potential of earthquake ground motion. However, this approach is dependent on ground motion databases and nonlinear analysis results. Subsequently, hybrid IMs were developed by Ozmen [31] to measure the damage potential of ground motion. Chen et al. [32] adopted the composite IM to select the severest input ground motion for underground structures by the rank method. Liu et al. [33] proposed a compound IM by a partial least square regression model, which evaluated the compound IM based on probabilistic seismic demand analysis. Exploratory factor analysis has been proposed for developing a compound IM incorporating the contributions of multiple IMs for predicting the potential damage of ground motion [34]. Chen et al. [35] assessed the damage potential of the ground motion of a tunnel by canonical correlation analysis.

As stated previously, the IMs exhibit correlations among themselves, leading to collinearity issues in multivariate regression models [36]. The diversity in measuring earthquake damage potential is highly dependent on the IMs in order to better capture the strength of the earthquake on a specific structure. The methods of PCA and CCA are employed to mitigate the collinearity problem by reducing the dimensionality of numerous interrelated IMs while retaining as much information as possible from the ground motion database [37]. Many researchers have attempted to focus on artificial neural network and

support vector machine methods to develop a new compound IM. However, a potential drawback of the artificial neural network is that it is time consuming, and the support vector machine is strongly dependent on the types of kernel functions. Unlike the past studies, no optimization algorithm is required in the PCA and CCA, avoiding the dependence on the optimization methods. This paper utilizes 10 IMs to quantify the potential for structural damage. The selected IMs cover a wide range of earthquake characteristics with respect to the peak intensities, mean-square intensities, the spectral intensity and the potential destructiveness, incorporating the frequency content of ground motions. The foundational concept of the PCA has motivated subsequent work, employing the method of CCA to study the relationship between various IMs and structural responses [38].

## 2. Principles of PCA and CCA

### 2.1. Principle Component Analysis

The idea of PCA involves transforming a vector into a set of independent variables within a new coordinate system, aiming to preserve the maximum information from the original variables [37]. The method seeks linear combinations with large variance, and the coefficients of these combinations are determined by the eigenvectors of the covariance matrix.

Suppose  $\mathbf{x} = (X_1, X_2, \dots, X_p)^T$  are the random variables, and the main step is to search a linear function, namely

$$\begin{cases} \mathbf{a}_1^T \mathbf{x} = \alpha_{11}X_1 + \alpha_{12}X_2 + \dots + \alpha_{1p}X_p \\ \mathbf{a}_2^T \mathbf{x} = \alpha_{21}X_1 + \alpha_{22}X_2 + \dots + \alpha_{2p}X_p \\ \vdots \\ \mathbf{a}_k^T \mathbf{x} = \alpha_{k1}X_1 + \alpha_{k2}X_2 + \dots + \alpha_{kp}X_p \end{cases} \quad (1)$$

where  $\mathbf{a}_2^T \mathbf{x}$  is uncorrelated with  $\mathbf{a}_1^T \mathbf{x}$ , and so on, and  $\mathbf{a}_k^T \mathbf{x}$  has a maximum variance subject to being uncorrelated with  $\mathbf{a}_1^T \mathbf{x}, \mathbf{a}_2^T \mathbf{x}, \dots, \mathbf{a}_{k-1}^T \mathbf{x}$ .

The covariance matrix of  $\mathbf{x}$  is  $\text{Cov}(\mathbf{x}) = E[(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^T] \triangleq \mathbf{\Sigma}$ . In a realistic case,  $\mathbf{\Sigma}$  is unknown, and its estimation is followed by a sample covariance matrix. The principal components are solved and correspond to the eigenvalue and eigenvector of  $\mathbf{\Sigma}$  in the multivariate statistical analysis. The eigenvector of  $\mathbf{\Sigma}$ , that is,  $\mathbf{e}_k = (e_{k1}, e_{k2}, \dots, e_{kp})^T$ , has a unit length corresponding to its  $k$ th largest eigenvalue  $\lambda_k$ ; therefore, the  $k$ th PC is given by

$$Y_k = \mathbf{e}_k^T \mathbf{x} = e_{k1}X_1 + e_{k2}X_2 + \dots + e_{kp}X_p \quad \text{for } k = 1, 2, \dots, p \quad (2)$$

with

$$\text{Var}(Y_k) = \mathbf{e}_k^T \mathbf{\Sigma} \mathbf{e}_k = \lambda_k \mathbf{e}_k^T \mathbf{e}_k = \lambda_k \quad (3)$$

$$\text{Cov}(Y_j, Y_k) = \mathbf{e}_j^T \mathbf{\Sigma} \mathbf{e}_k = \lambda_k \mathbf{e}_j^T \mathbf{e}_k = 0, \quad j \neq k \quad (4)$$

It is defined that when  $\mathbf{a}_1 = \mathbf{e}_1$ ,  $\text{Var}(Y_1)$  is the largest, so that

$$\max_{\mathbf{a}_1^T \mathbf{a}_1 = 1} \{\text{Var}(Y_1)\} = \text{Var}(\mathbf{e}_1^T \mathbf{x}) = \mathbf{e}_1^T \mathbf{\Sigma} \mathbf{e}_1 = \lambda_1 \quad (5)$$

Similarly, it is proven that when  $\mathbf{a}_2 = \mathbf{e}_2$ ,  $\text{Var}(Y_2) \leq \lambda_2 \leq \lambda_1$  and  $\text{Cov}(Y_2, Y_1) = 0$ .

In general, most of the variation in  $\mathbf{x}$  will be accounted for by  $m$  principal components, where  $m < p$ , and the criterion for choosing  $m$  PCs is the cumulative percentage of total variation exceeding 85%, defined by

$$\alpha_m = \sum_{k=1}^m \lambda_k / \sum_{k=1}^p \lambda_k \quad (6)$$

where the percentage of variation of the  $k$ th PCs is

$$\omega_k = \lambda_k / \sum_{k=1}^p \lambda_k \quad (7)$$

## 2.2. Canonical Correlation Analysis

Canonical correlation analysis serves as a tool for a multivariate statistical analysis, with its foundational theory originally developed by Hotelling [39]. The method of CCA lends the fundamental concept of dimension reduction with PCA. However, the method of PCA based on the marginal covariance matrix does not account for the correlation between two pairs of variables, and the method of CCA utilizes a joint covariance matrix to depict the relationships between two sets by maximizing the Pearson correlation [40].

Suppose  $\mathbf{x} = (X_1, X_2, \dots, X_p)^T$  and  $\mathbf{y} = (Y_1, Y_2, \dots, Y_q)^T$  are the random variables; therefore, the covariance matrix of  $\mathbf{x}$  and  $\mathbf{y}$  is

$$\text{cov} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{XX} & \boldsymbol{\Sigma}_{XY} \\ \boldsymbol{\Sigma}_{YX} & \boldsymbol{\Sigma}_{YY} \end{bmatrix} \quad (8)$$

In order to study the correlation between  $\mathbf{x}$  and  $\mathbf{y}$ , using the idea of PCA to find the linear combinations, we consider  $U_i$  and  $V_i$  as the comprehensive variables, namely

$$U_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p \equiv \mathbf{a}^T \mathbf{x} \quad (9)$$

$$V_i = b_{i1}Y_1 + b_{i2}Y_2 + \dots + b_{iq}Y_q \equiv \mathbf{b}^T \mathbf{y} \quad (10)$$

under the constraints

$$\text{Var}(\mathbf{a}^T \mathbf{x}) = \text{Var}(\mathbf{b}^T \mathbf{y}) = 1 \quad (11)$$

Given the covariance matrix  $\boldsymbol{\Sigma}$  of random variables  $\mathbf{x}$  and  $\mathbf{y}$ , searching for the vectors  $\mathbf{a}$  and  $\mathbf{b}$ , to achieve the maximum correlation between the comprehensive variables  $U_i$  and  $V_i$ , the correlation coefficient is given by

$$\rho = \frac{\text{cov}(U_i, V_i)}{\sqrt{\text{Var}(U_i)\text{Var}(V_i)}} = \frac{\text{cov}(\mathbf{a}^T \mathbf{x}, \mathbf{b}^T \mathbf{y})}{\sqrt{\text{Var}(\mathbf{a}^T \mathbf{x})\text{Var}(\mathbf{b}^T \mathbf{y})}} \quad (12)$$

Therefore,

$$\rho = \text{cov}(\mathbf{a}^T \mathbf{x}, \mathbf{b}^T \mathbf{y}) = \mathbf{a}^T \text{cov}(\mathbf{x}, \mathbf{y}) \mathbf{b} = \mathbf{a}^T \boldsymbol{\Sigma}_{XY} \mathbf{b} \quad (13)$$

Then, the following is defined:

$$\mathbf{T} = \boldsymbol{\Sigma}_{XX}^{-1/2} \boldsymbol{\Sigma}_{XY} \boldsymbol{\Sigma}_{YY}^{-1/2} \quad (14)$$

where  $\lambda_1^2 \geq \lambda_2^2 \geq \dots \geq \lambda_k^2$  are the nonzero eigenvalues of  $\mathbf{T}\mathbf{T}^T$  and  $\mathbf{T}^T\mathbf{T}$ , and  $\boldsymbol{\gamma}_i$  and  $\boldsymbol{\delta}_i$  are the eigenvectors corresponding to its  $i$ th eigenvalues  $\lambda_i^2$ .

Canonical correlation vectors  $\mathbf{a}_i$  and  $\mathbf{b}_i$  are given by

$$\mathbf{a}_i = \boldsymbol{\Sigma}_{XX}^{-1/2} \boldsymbol{\gamma}_i \quad (15)$$

$$\mathbf{b}_i = \boldsymbol{\Sigma}_{YY}^{-1/2} \boldsymbol{\delta}_i \quad (16)$$

The canonical correlation variables are defined as follows:

$$u_i = \mathbf{a}_i^T \mathbf{x} \quad (17)$$

$$v_i = \mathbf{b}_i^T \mathbf{y} \quad (18)$$

with

$$\text{cov}(u_i, u_j) = \mathbf{a}_i^T \boldsymbol{\Sigma}_{XX} \mathbf{a}_j = \boldsymbol{\gamma}_i^T \boldsymbol{\gamma}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (19)$$

and the same is true for  $\text{cov}(v_i, v_j)$ .

In order to solve the correlation maximization, we fix  $\mathbf{b}$  subject to  $\mathbf{a}^T \boldsymbol{\Sigma}_{XX} \mathbf{a} = 1$ , i.e., the following is solved:

$$\begin{aligned} \max_{\mathbf{a}^T \boldsymbol{\Sigma}_{XX} \mathbf{a} = 1} (\mathbf{a}^T \boldsymbol{\Sigma}_{XY} \mathbf{b})^2 &= \max_{\mathbf{a}^T \boldsymbol{\Sigma}_{XX} \mathbf{a} = 1} (\mathbf{b}^T \boldsymbol{\Sigma}_{YX} \mathbf{a} \mathbf{a}^T \boldsymbol{\Sigma}_{XY} \mathbf{b}) = \max(\mathbf{b}^T \boldsymbol{\Sigma}_{YX} \boldsymbol{\Sigma}_{XX}^{-1} \boldsymbol{\Sigma}_{XY} \mathbf{b}) \\ &= \max(\boldsymbol{\delta}^T \boldsymbol{\Sigma}_{YY}^{-1/2} \boldsymbol{\Sigma}_{YX} \boldsymbol{\Sigma}_{XX}^{-1} \boldsymbol{\Sigma}_{XY} \boldsymbol{\Sigma}_{YY}^{-1/2} \boldsymbol{\delta}) = \max(\boldsymbol{\delta}^T \mathbf{T}^T \mathbf{T} \boldsymbol{\delta}) \end{aligned} \quad (20)$$

where  $\mathbf{a}_1^T \mathbf{x}$  and  $\mathbf{b}_1^T \mathbf{y}$  are the first pairs of canonical correlation variables, and so on,  $\mathbf{a}_i^T \mathbf{x}$  and  $\mathbf{b}_i^T \mathbf{y}$  are the  $i$ th pairs of canonical correlation variables, hence

$$\max_{\mathbf{a}_i^T \boldsymbol{\Sigma}_{XX} \mathbf{a}_i = 1} (\mathbf{a}_i^T \boldsymbol{\Sigma}_{XY} \mathbf{b}_i)^2 = \max(\boldsymbol{\delta}_i^T \mathbf{T}^T \mathbf{T} \boldsymbol{\delta}_i) = \lambda_i^2 \quad (21)$$

Therefore, the canonical correlation coefficients are given by

$$\rho_i = \lambda_i \quad \text{for } i = 1, 2, \dots, k \quad (22)$$

In fact, the joint covariance matrix of  $\mathbf{x}$  and  $\mathbf{y}$  is often unknown; therefore, we apply the sample data to estimate the covariance matrix. In addition, in order to eliminate the influence of dimension, it is important to standardize the original variables.

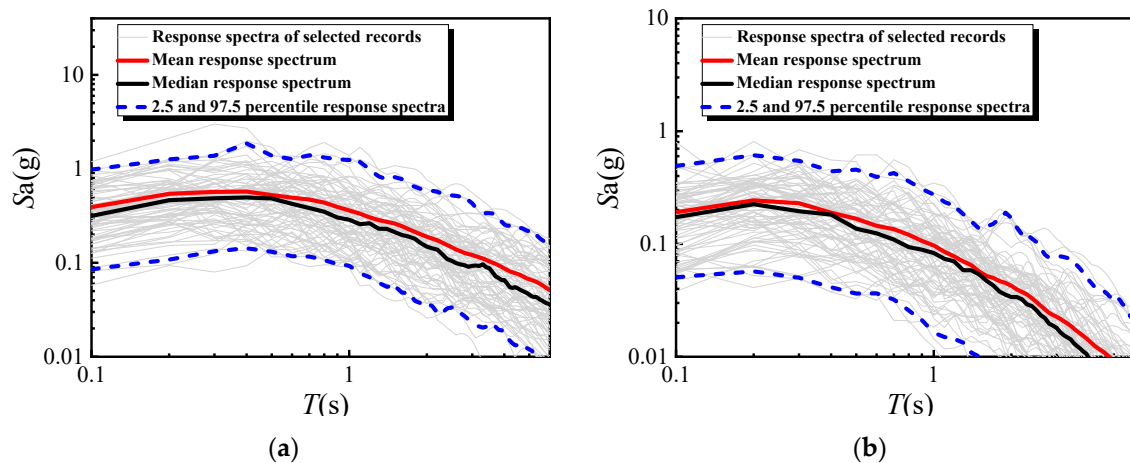
### 3. Selection of Ground Motions and Intensity Measures

#### 3.1. Selection of Ground Motions

Two suites of 40 pairs of horizontal bidirectional ground motions compiled by Baker et al. [41] for the PEER Transportation Research Program were employed as input ground motions. All ground motions within these sets were obtained from the PEER Next Generation Attenuation Project ground motion library [42]. The standardized sets of ground motions presented by Baker et al. [41] were specifically chosen for use in the multi-hazard probabilistic seismic demand model [43], as well as for the fragility analyses [44,45].

The ground motions considered in this study comprise two sets of broad-band ground motions for a soil site experiencing a moderately large earthquake at small distances, as detailed in Baker et al. [41]. These sets are designed to match the median and logarithmic standard deviations of the response spectra of a generic earthquake in California, with one set corresponding to a prediction for a magnitude 7 strike-slip earthquake at a distance of 10 km and the other set corresponding to a prediction for a magnitude 6 strike-slip earthquake at a distance of 25 km. Both sets of ground motions are characterized by an assumed average shear wave velocity in the top 30 m of 250 m/s. Assuming that the selected ground motion contained pulse-like motions with longer periods, more details about the pulse-like ground motions can be found in Hosseini et al. [46].

For the nonlinear dynamic analyses, we employed two sets of 40 pairs of horizontal bidirectional ground motions without scaling. This decision was made to avoid altering the correlation between the IMs and structural damage, which can occur when scaling the ground motions. The majority of correlation studies utilized unscaled accelerograms [8,44,45]. The response spectra for the two unscaled suites of 40 pairs of horizontal bidirectional ground motions are shown in Figure 1.



**Figure 1.** Response spectra of the broad-band ground motions for soil sites with log-log of the axes: (a) set #1A; (b) set #1B.

### 3.2. Selection of Intensity Measures

The chosen IM should effectively capture the ground motion damage potential, and correlation analyses between various IMs and response variables are employed to identify optimal IMs [47–50]. Riddell [27] proposed a comprehensive evaluation to assess the effectiveness of 23 IMs, as detailed in Liu et al. [38]. These IMs are categorized into acceleration-related, velocity-related, and displacement-related indices, following the classification by Chopra [51]. The results indicated that no single IM exhibits a good correlation over the entire frequency range. On the basis of the previous study, the concept of combination IMs was introduced by Ozmen [31] and Morfidis and Kostinakis [52]. Ghotbi and Taciroglu [53] selected the ground motion based on a multi-intensity-measure conditioning approach with an emphasis on diverse earthquake contents.

In light of the aforementioned discussions, the methods of PCA and CCA are developed to identify compound IMs that exhibit a high correlation with structural damage. The CCA utilizes the key concept of dimension reduction from the PCA to select IMs extracted from the 23 IMs. It is noteworthy that not all the IMs are utilized to develop the compound IM due to the possible overlap of information. Increasing the number of IMs in the developed compound IM can naturally improve the predictive capacity of earthquake damage potential. However, the growing complexity of a compound IM may lead to the problem of overfitting.

### 3.3. Analysis of Correlation between IMs

This paper utilizes 23 IMs compiled by Riddell [27] from the existing literature to assess the correlation among them on a logarithmic scale. Figure 2 illustrates the correlation coefficient matrix. Some IMs are excluded, such as  $a_{sq}$ ,  $a_{rs}$ ,  $a_{rms}$ ,  $v_{rs}$ ,  $v_{rms}$ ,  $d_{rs}$  and  $d_{rms}$ , because the correlation coefficients between  $I_A$ ,  $a_{sq}$  and  $a_{rs}$  are nearly 1 in log-scale, and the same for  $P_a$  and  $a_{rms}$ ,  $v_{sq}$  and  $v_{rs}$ ,  $P_v$  and  $v_{rms}$ ,  $d_{sq}$  and  $d_{rs}$  and  $P_d$  and  $d_{rms}$ . The results from Figure 2 show that the intragroup IMs present high correlation among them. In general, the acceleration-related indices include  $PGA$  and  $I_a$ , and  $P_a$  and  $I_c$  have high correlation, as the correlation coefficients are above 0.95 and 0.92, respectively. The velocity-related indices  $PGV$  and  $I_F$ ,  $v_{sq}$  and  $P_v$ ,  $v_{sq}$  and  $I_F$ ,  $v_{sq}$  and  $I_v$  and  $I_F$  and  $I_v$  have high correlation, as the correlation coefficients are above 0.97, 0.93, 0.91, 0.92 and 0.97, respectively; the displacement-related indices are highly correlated among themselves. Finally, the results reveal that the intergroup IMs exhibit low correlation. Specifically, the correlation between the acceleration-related indices and both velocity-related and displacement-related indices is weak. However, the correlation coefficient between the velocity-related indices and displacement-related indices consistently exceeds 0.5. Comparing Figure 2a and Figure 2b, no matter how the ground motion changes, the fact that the intragroup IMs present high



correlation and the intergroup IMs exhibit low correlation will not change. These results align with those presented in the literature by Riddell et al. [27] and Narasimhan and Wang et al. [36].

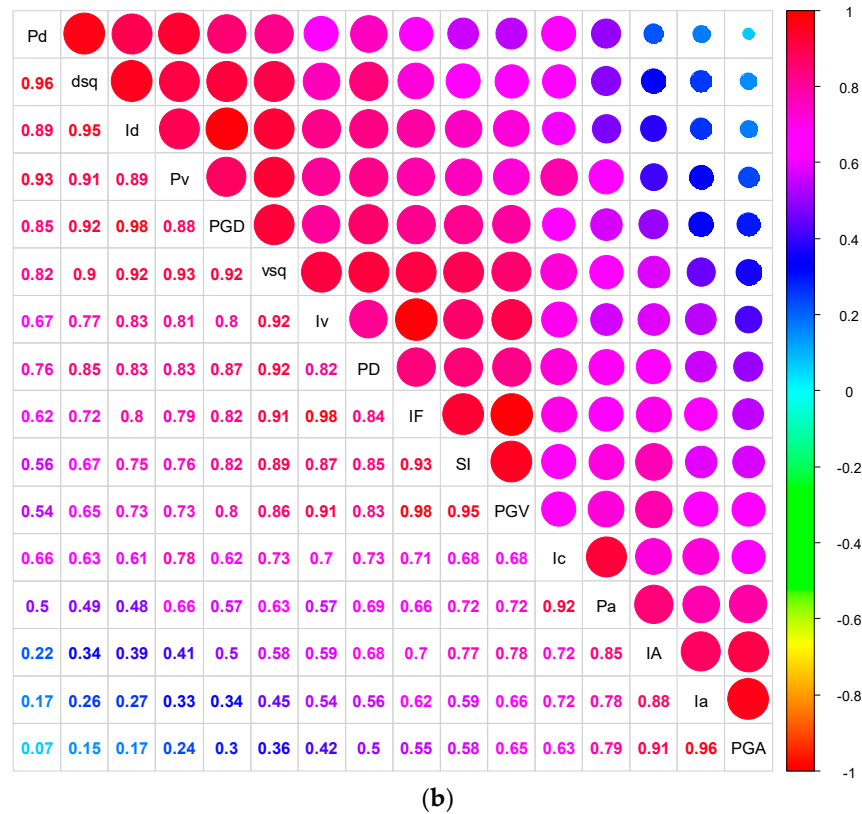
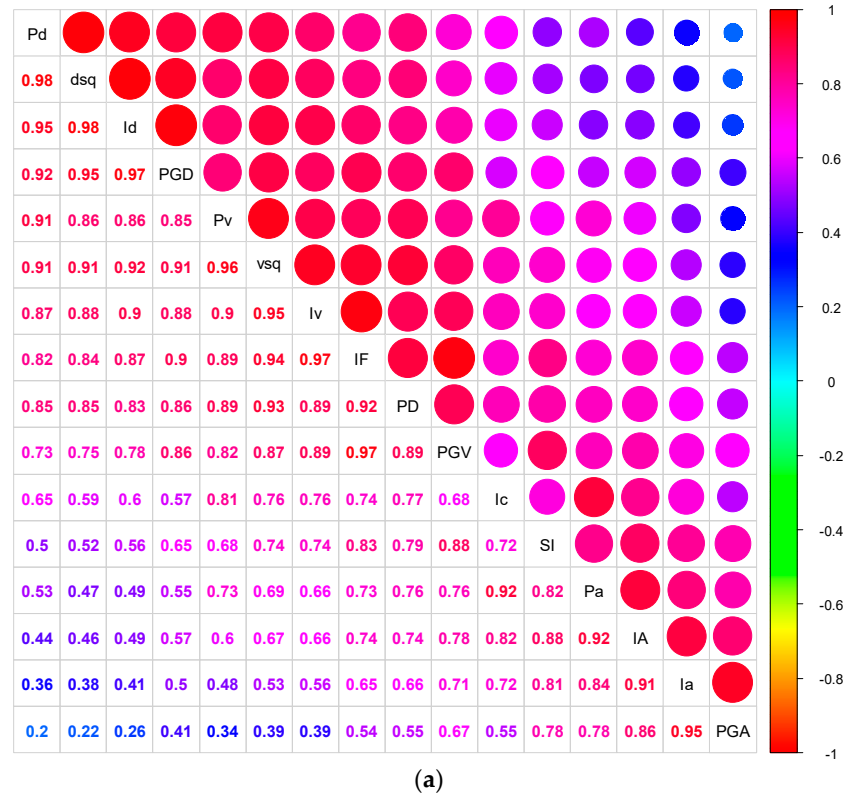
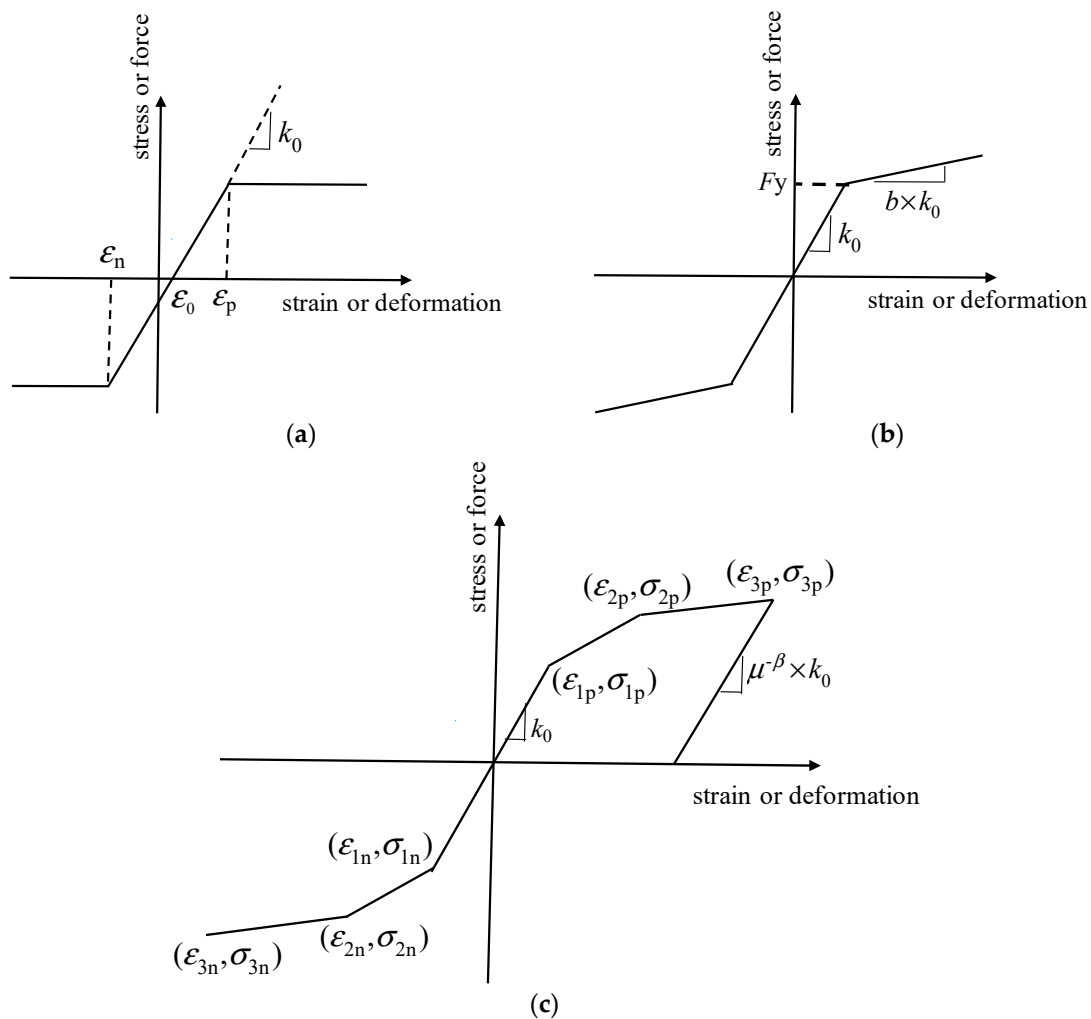


Figure 2. The correlation coefficient matrix of IMs: (a) set #1A; (b) set #1B.

#### 4. Evaluation of Damage Potential of SDOF Systems Using PCA and CCA

##### 4.1. SDOF Systems

The open-source finite-element software OpenSEES has been developed for earthquake engineering. OpenSEES 3.3.0 was employed to construct SDOF systems and conduct nonlinear dynamic analyses. Elastoplastic, bilinear and hysteretic models are utilized to determine peak deformation and hysteretic energy as demand measures (DMs) for the SDOF systems. For the SDOF systems implemented in OpenSEES, the Elastic-Perfectly Plastic material, the Steel01 material and hysteretic materials were adopted, as illustrated in Figure 3.



**Figure 3.** The hysteretic models of SDOF systems based on OpenSEES: (a) Elastic-Perfectly Plastic material; (b) Steel01 material; (c) hysteretic material.

Hysteretic materials incorporating both strength and stiffness deterioration were categorized into two models, one considering the pinching effect and the other not considering it. To enhance practicality, four SDOF models were developed utilizing elastoplastic, bilinear and hysteretic materials, as outlined in Table 1. The material characteristics for these models in OpenSEES are presented in Table 1.



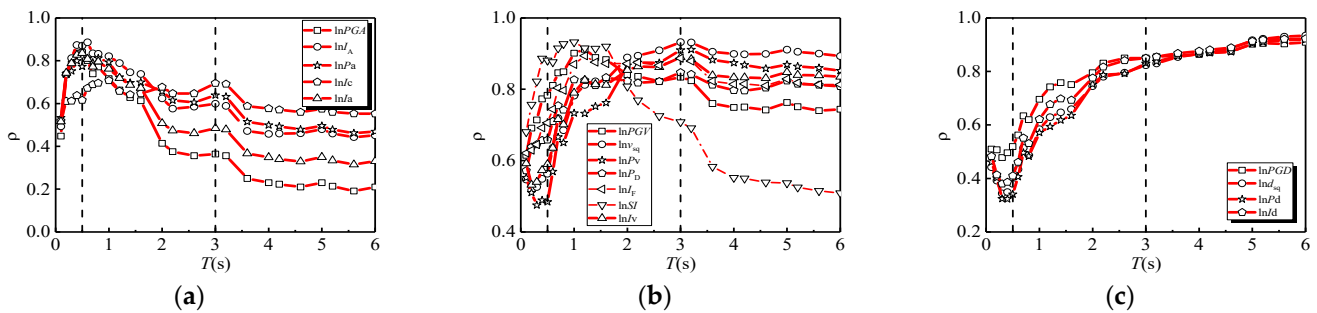
**Table 1.** Four SDOF models and the corresponding material characteristics in OpenSEES.

The Model	Material	The Parameters of Material Characteristics
Model 1	EPP	$k_0$ is the elastic tangential stiffness $\varepsilon_p$ is the yield displacement
Model 2	Steel01	$F_y$ is the yield force $k_0$ is the elastic tangential stiffness $b = 0.05$ is the strain hardening ratio
Model 3	Hysteretic	$\varepsilon_{1p}, \sigma_{1p}$ are the yield displacement and force $\varepsilon_{2p}, \sigma_{2p}$ are the strain hardening displacement and force $\varepsilon_{1n}, \sigma_{1n}$ are the negative of $\varepsilon_{1p}, \sigma_{1p}$ $\varepsilon_{2n}, \sigma_{2n}$ are the negative of $\varepsilon_{2p}, \sigma_{2p}$ does not consider the pinching effect the pinching factor for deformation and force during reloading are set to 1
Model 4	Hysteretic	the parameters are the same as Model 3 considers the pinching effect the pinching factor for deformation and force during reloading are set to 0.8 and 0.2, respectively

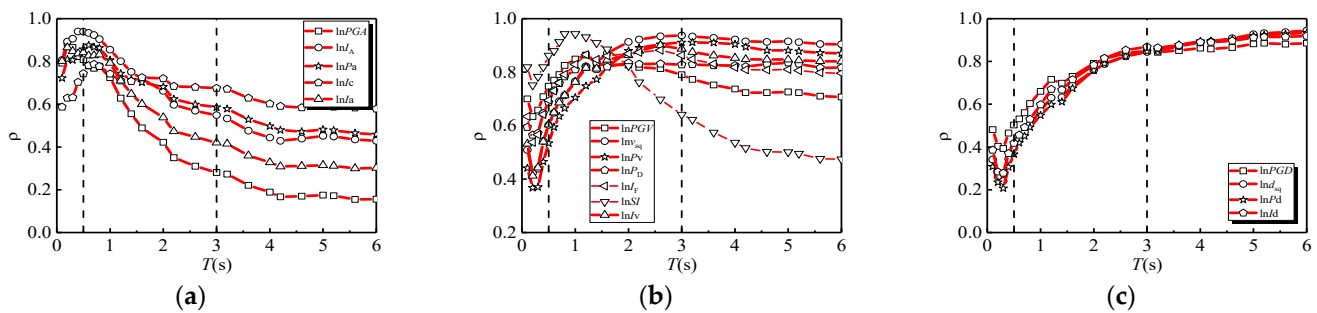
The critical damping is assumed to be 5 percent for the computations in the nonlinear dynamic analysis. The fundamental periods of the 25 SDOF systems are varied from 0.1 s to 6 s at different intervals. The strength reduction factor  $R$  is predetermined at values of 2, 3, 4 and 5 for the elastoplastic systems. For each SDOF system, the yield strength of the hysteretic loop is varied at four strength reduction levels by dividing the elastic spectral strength by each of the yield strength reduction factors. Consequently, a total of 400 nonlinear SDOF systems (4 kinds of hysteretic models  $\times$  25 vibration periods  $\times$  4 strength reduction levels) are subjected to the selected 160 ground motion records, and 64,000 ( $160 \times 400 = 64,000$ ) runs of nonlinear dynamic analyses are conducted by OpenSEES.

#### 4.2. Analysis of Correlation between IMs and DMs

As previously mentioned, two subsets designated as set #1A and set #1B are employed for validating the correlation analysis. In this investigation, set #1A is utilized as the input ground motions. Model 1 is selected as the benchmark model at  $R = 2$ , with peak deformation  $u_{\max}$  and the hysteretic energy  $E_H$  serving as the response variables for the SDOF systems. According to existing research, there is a high linear correlation between seismic intensity and structural damage in logarithmic space. To accurately measure for the damage potential of ground motions, this paper performs a logarithmic transformation on the seismic intensity measures and structural damage. The Pearson correlation coefficients between the IMs of database set #1A and the DMs of the benchmark model in a logarithmic scale are shown in Figures 4 and 5. It is worth noting that  $I_A$ ,  $a_{sq}$  and  $a_{rs}$  have the same correlation as the DMs in the log-scale, and so also the same as  $P_a$  and  $a_{rms}$ ,  $v_{sq}$  and  $v_{rs}$ ,  $P_v$  and  $v_{rms}$ ,  $d_{sq}$  and  $d_{rs}$  and  $P_d$  and  $d_{rms}$ . Therefore,  $a_{sq}$ ,  $a_{rs}$ ,  $a_{rms}$ ,  $v_{rs}$ ,  $v_{rms}$ ,  $d_{rs}$  and  $d_{rms}$  are excluded. Similar results pertaining to set #1B and other models are omitted for brevity and clarity.



**Figure 4.** Relationships between  $\ln IMs$  and  $\ln u_{\max}$ : (a) acceleration-related indices; (b) velocity-related indices; (c) displacement-related indices.



**Figure 5.** Relationships between  $\ln IMs$  and  $\ln E_H$ : (a) acceleration-related indices; (b) velocity-related indices; (c) displacement-related indices.

The  $\rho$ – $T$  relationships presented in Figures 4 and 5 encompass the acceleration-related indices, velocity-related indices and displacement-related indices. Notably, the acceleration-related indices exhibit a high correlation with the DMs within the acceleration region (the short period is  $T = 0$ – $0.5$  s), and the correlation coefficient  $\rho$  generally increases with the structural period  $T$  within the short period, decreasing  $\rho$  for the velocity region (the intermediate period is  $T = 0.5$ – $3$  s) and the displacement region (the long period is  $T > 3$  s). For the velocity-related indices, a robust correlation with the DMs is observed within the intermediate period. However, this correlation diminishes with reductions in the short and long periods. Particularly, the  $SI$  differs from other velocity-related indices, presenting excellent correlation at  $T = 1$  s within the intermediate period but a reduced correlation coefficient  $\rho$  at other period ranges. The correlation of the displacement-related indices with DMs increases as the structural period  $T$  rises, showing high correlation in the long period, which represents the displacement region. These results align with the findings presented in the literature by Ye et al. [54].

Based on the above discussion, it is concluded that no single IM presents a satisfactory correlation with the DMs for all the frequency ranges. Some IMs demonstrate more advantageous relationships with the DMs than others within specific structural period regions. Consequently, multivariate statistical tools, i.e., the methods of PCA and CCA, are adopted to evaluate the compound IMs.

#### 4.3. Evaluation of Damage Potential by Analysis of Correlation between Principal Components of IMs and DMs

Similar expressions among the IMs themselves lead to the issue of collinearity in the method of PCA [33]. In order to avoid missing information, this study cannot select extensively overlapping information. Consequently, some candidate IMs should be screened to determine the optimal compound IMs. Specifically,  $a_{rs}$ ,  $v_{rs}$  and  $d_{rs}$  differ only by a power from  $a_{sq}$ ,  $v_{sq}$  and  $d_{sq}$  in the equations, respectively;  $a_{rms}$ ,  $v_{rms}$  and  $d_{rms}$  also differ only by a power from  $P_a$ ,  $P_v$  and  $P_d$ , respectively;  $I_A$  is excluded since it differs only by a constant from  $a_{sq}$ ; and  $I_c$ ,  $I_a$ ,  $I_f$ ,  $I_v$  and  $I_d$  are combinations of dura-

tion and the other IMs in the logarithmic scale. As a consequence, only ten IMs remain,  $\mathbf{IM} = [PGA, a_{rs}, P_a, I_c, PGV, v_{sq}, P_D, SI, PGD, d_{sq}]$ , which cover a wide range of earthquake characteristics with respect to the peak intensities, mean square intensities, the spectral intensity and the potential destructiveness incorporating the frequency content of ground motions.

Subsequently, in accordance with the fundamental principles of PCA, the adoption of  $\mathbf{x} = \ln \mathbf{IM} - \mu_{\ln \mathbf{IM}}$  as the basic variable is undertaken, and the results of set #1A are presented below. The covariance matrix  $\Sigma$  of  $\mathbf{x}$  and the corresponding eigenvector matrix  $\mathbf{e}$  are evaluated by the following:

$$\Sigma = \begin{bmatrix} 0.348 & 0.238 & 0.462 & 0.240 & 0.250 & 0.294 & 0.397 & 0.274 & 0.205 & 0.239 \\ 0.238 & 0.222 & 0.434 & 0.289 & 0.232 & 0.404 & 0.427 & 0.246 & 0.228 & 0.400 \\ 0.462 & 0.434 & 1.002 & 0.688 & 0.481 & 0.883 & 0.925 & 0.494 & 0.473 & 0.868 \\ 0.240 & 0.289 & 0.688 & 0.558 & 0.319 & 0.731 & 0.698 & 0.321 & 0.366 & 0.811 \\ 0.250 & 0.232 & 0.481 & 0.319 & 0.397 & 0.702 & 0.685 & 0.331 & 0.464 & 0.859 \\ 0.294 & 0.404 & 0.883 & 0.731 & 0.702 & 1.637 & 1.446 & 0.565 & 0.992 & 2.131 \\ 0.397 & 0.427 & 0.925 & 0.698 & 0.685 & 1.446 & 1.488 & 0.574 & 0.896 & 1.893 \\ 0.274 & 0.246 & 0.494 & 0.321 & 0.331 & 0.565 & 0.574 & 0.358 & 0.329 & 0.565 \\ 0.205 & 0.228 & 0.473 & 0.366 & 0.464 & 0.992 & 0.896 & 0.329 & 0.727 & 1.476 \\ 0.239 & 0.400 & 0.868 & 0.811 & 0.859 & 2.131 & 1.893 & 0.565 & 1.476 & 3.341 \end{bmatrix} \quad (23)$$

$$\mathbf{e} = \begin{bmatrix} 0.101 & 0.362 & -0.447 & -0.411 & 0.101 & 0.188 & -0.390 & -0.048 & -0.422 & 0.335 \\ 0.119 & 0.254 & -0.058 & -0.129 & -0.061 & 0.375 & -0.364 & 0.357 & 0.629 & -0.319 \\ 0.259 & 0.557 & 0.274 & -0.271 & -0.028 & -0.401 & 0.109 & -0.199 & -0.141 & -0.492 \\ 0.205 & 0.285 & 0.602 & -0.100 & -0.038 & 0.170 & 0.193 & 0.145 & 0.126 & 0.632 \\ 0.199 & 0.120 & -0.357 & 0.119 & -0.225 & -0.365 & 0.297 & 0.719 & -0.090 & 0.088 \\ 0.439 & -0.039 & 0.147 & 0.519 & -0.435 & -0.044 & -0.523 & -0.098 & -0.198 & 0.025 \\ 0.412 & 0.119 & -0.145 & 0.402 & 0.771 & -0.083 & 0.020 & -0.070 & 0.150 & 0.063 \\ 0.159 & 0.261 & -0.289 & 0.252 & -0.275 & 0.550 & 0.532 & -0.293 & -0.043 & -0.105 \\ 0.282 & -0.146 & -0.311 & -0.253 & -0.260 & -0.362 & 0.047 & -0.420 & 0.528 & 0.288 \\ 0.600 & -0.543 & 0.096 & -0.398 & 0.078 & 0.244 & 0.133 & 0.130 & -0.201 & -0.195 \end{bmatrix} \quad (24)$$

The corresponding eigenvalue of the covariance matrix  $\Sigma$  of  $\mathbf{x}$  is shown in Figure 6. To determine the number of PCs to select, the eigenvalues are examined. The criterion for selection involves identifying the third point at which the graph approximately forms a straight line, as illustrated in Figure 6. The results of the first three PCs are depicted in Figure 7. The first PC accounts for a substantial proportion, amounting to 80.4% of the total variation. The second and third PCs contribute 13.42% and 2.7% of the total variation, respectively. Collectively, the first three PCs account for 96.52% of the variation, offering a comprehensive representation of the total variation and retaining nearly all the information from the original variables in  $\mathbf{x}$ .

As stated above, the new variables resulting from the linear transformation of the first three PCs are presented below:

$$\ln IM_{\text{PCA}} = \sum_{i=1}^3 \omega_i Y_i \quad (25)$$

where  $\omega_i$  is the percentage of variation of the  $i$ th PCs, and  $Y_i$  is the  $i$ th PCs.

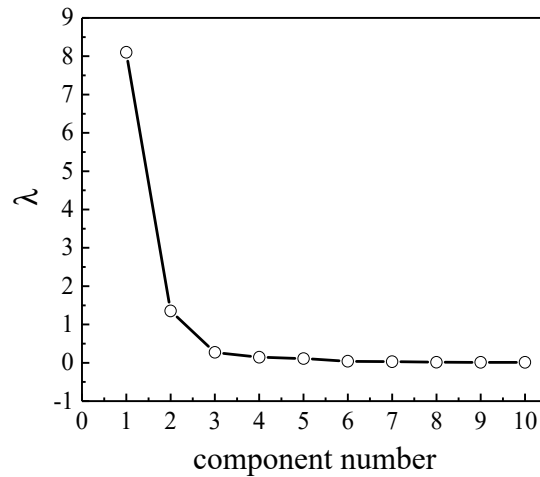


Figure 6. The scree graph of PCA.

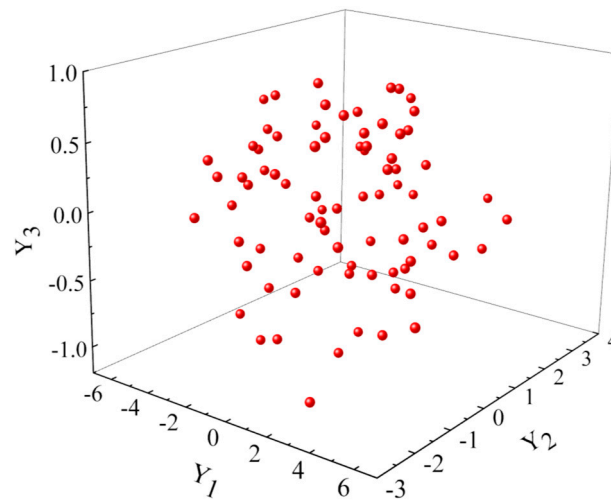


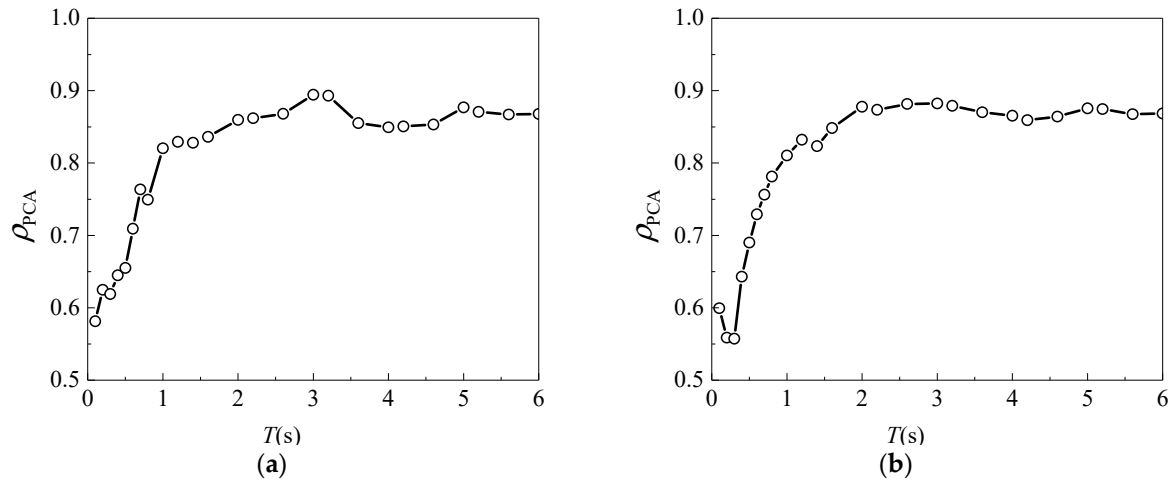
Figure 7. The first three PCs.

To visualize the correlation between the compound IMs and DMs, the correlation coefficient is quantitatively expressed as follows:

$$\rho_{PCA} = \frac{\sum_{i=1}^N (\ln IM_{PCA,i} - \overline{\ln IM_{PCA}})(\ln DM_i - \overline{\ln DM})}{\sqrt{\sum_{i=1}^N (\ln IM_{PCA,i} - \overline{\ln IM_{PCA}})^2 \sum_{i=1}^N (\ln DM_i - \overline{\ln DM})^2}} \quad (26)$$

where  $\overline{\ln IM_{PCA}}$  and  $\overline{\ln DM}$  are the mean values of  $\ln IM_{PCA,i}$  and  $\ln DM_i$  respectively, and  $N$  is the number of pairs of values  $(\ln IM_{PCA,i}, \ln DM_i)$  in the data.

The  $\rho_{PCA-T}$  relationships are presented in Figure 8. The compound IMs exhibit advantageous correlations with DMs compared to other individual IMs, and the compound IMs prove more adept at capturing the characteristics of ground motion. Notably, the correlation between the compound IMs and DMs for all frequency ranges is relatively higher and more stable than those of the single IM. However, it is noteworthy that the correlation coefficient undergoes changes with nonlinear analysis results, indicating that it does not consistently maintain a high correlation. Subsequently, the evaluation of the compound IMs for the optimal correlation with the DMs for all frequency ranges is conducted based on the method of CCA between the IMs and DMs.



**Figure 8.** Relationships between principal components of multivariate IMs and DMs: (a)  $\ln u_{\max}$ ; (b)  $\ln E_H$ .

#### 4.4. Evaluation of Damage Potential by Analysis of Canonical Correlation between IMs and DMs

As previously mentioned, the method of CCA is implemented to determine the correlation between the IMs and structural damage, quantified through the canonical correlation coefficient. The input variables consist of ten candidate IMs. The peak deformation  $u_{\max}$  and hysteretic energy  $E_H$  are the response variables, namely  $\mathbf{DM} = [u_{\max}, E_H]$ . We perform the canonical correlation analysis on the data matrices  $\mathbf{x} = \ln \mathbf{IM} - \mu_{\ln \mathbf{IM}}$  and  $\mathbf{y} = \ln \mathbf{DM} - \mu_{\ln \mathbf{DM}}$  of the benchmark model; when the fundamental periods of the SDOF systems are 1.0 s, the estimated covariance matrices of  $\mathbf{x}$  and  $\mathbf{y}$  are given by

$$\Sigma = \begin{bmatrix} 0.348 & 0.238 & 0.462 & 0.240 & 0.250 & 0.294 & 0.397 & 0.274 & 0.205 & 0.239 & 0.292 & 0.510 \\ 0.238 & 0.222 & 0.434 & 0.289 & 0.232 & 0.404 & 0.427 & 0.246 & 0.228 & 0.400 & 0.267 & 0.481 \\ 0.462 & 0.434 & 1.002 & 0.688 & 0.481 & 0.883 & 0.925 & 0.494 & 0.473 & 0.868 & 0.549 & 0.963 \\ 0.240 & 0.289 & 0.688 & 0.558 & 0.319 & 0.731 & 0.698 & 0.321 & 0.366 & 0.811 & 0.366 & 0.661 \\ 0.250 & 0.232 & 0.481 & 0.319 & 0.397 & 0.702 & 0.685 & 0.331 & 0.464 & 0.859 & 0.392 & 0.638 \\ 0.294 & 0.404 & 0.883 & 0.731 & 0.702 & 1.637 & 1.446 & 0.565 & 0.992 & 2.131 & 0.692 & 1.171 \\ 0.397 & 0.427 & 0.925 & 0.698 & 0.685 & 1.446 & 1.488 & 0.574 & 0.896 & 1.893 & 0.698 & 1.179 \\ 0.274 & 0.246 & 0.494 & 0.321 & 0.331 & 0.565 & 0.574 & 0.358 & 0.329 & 0.565 & 0.386 & 0.674 \\ 0.205 & 0.228 & 0.473 & 0.366 & 0.464 & 0.992 & 0.896 & 0.329 & 0.727 & 1.476 & 0.410 & 0.672 \\ 0.239 & 0.400 & 0.868 & 0.811 & 0.859 & 2.131 & 1.893 & 0.565 & 1.476 & 3.341 & 0.738 & 1.245 \\ 0.292 & 0.267 & 0.549 & 0.366 & 0.392 & 0.692 & 0.698 & 0.386 & 0.410 & 0.738 & 0.478 & 0.777 \\ 0.510 & 0.481 & 0.963 & 0.661 & 0.638 & 1.171 & 1.179 & 0.674 & 0.672 & 1.245 & 0.777 & 1.426 \end{bmatrix} \quad (27)$$

Now, we estimate  $\mathbf{T} = \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2}$  to obtain the nonzero eigenvalues of  $\mathbf{T}\mathbf{T}^T$  and  $\mathbf{T}^T\mathbf{T}$ , and  $\gamma_i$  and  $\delta_i$  are the eigenvectors of  $\mathbf{T}\mathbf{T}^T$  and  $\mathbf{T}^T\mathbf{T}$ , respectively. The first two eigenvalues are

$$\lambda_1^2 = 0.9284, \lambda_2^2 = 0.2693 \quad (28)$$

Then, the first canonical correlation vectors  $\mathbf{a}_1$  and  $\mathbf{b}_1$  are given by

$$\mathbf{a}_1 = \Sigma_{XX}^{-1/2} \gamma_1 = [-0.255, 0.066, 0.240, -0.231, -0.255, -0.166, -0.108, -1.144, 0.161, 0.027]^T \quad (29)$$

$$\mathbf{b}_1 = \Sigma_{YY}^{-1/2} \delta_1 = [-0.711, -0.439]^T \quad (30)$$

Therefore, the first pairs of canonical correlation variables are given by

$$\ln IM_{CCA} = \sum_{i=1}^{10} a_{1i}(T)(\ln IM_i - \mu_{\ln IM_i}) \quad (31)$$

$$\ln DM_{CCA} = \sum_{i=1}^2 b_{1i}(T)(\ln DM_i - \mu_{\ln DM_i}) \quad (32)$$

where  $a_{1i}$  and  $b_{1i}$  are the canonical coefficients making up the first pairs of the canonical correlation variable for  $\ln IM_{CCA}$  and  $\ln DM_{CCA}$ , respectively, and they are the functions of the fundamental period  $T$  of the SDOF systems. Therefore, the first pairs of the canonical correlation coefficient are given by

$$\rho_{CCA}(T) = \frac{\text{cov}(\ln IM_{CCA}, \ln DM_{CCA})}{\sqrt{\text{Var}(\ln IM_{CCA})\text{Var}(\ln DM_{CCA})}} = \lambda(T) \quad (33)$$

The canonical correlation coefficient of the 25 SDOF systems corresponding to the benchmark model under set #1A is illustrated in Figure 9. It is shown that the canonical correlation coefficient  $\rho_{CCA}$  surpasses 0.9 for all frequency ranges, with the exception of  $T = 0.1$  s. Scatter plots illustrating the correlation between the natural logarithms of the IMs and DMs are presented in Figure 10. It illustrates that  $\ln IM_{CCA}$  presents excellent correlation with  $\ln DM_{CCA}$  at the acceleration region ( $T = 0.2$  s), the velocity region ( $T = 1$  s) and the displacement region ( $T = 5$  s), where  $\rho_{CCA}$  is 0.945, 0.964 and 0.963, respectively.

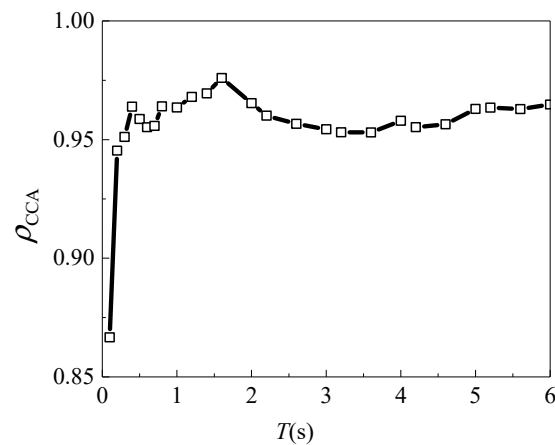


Figure 9. Relationships of benchmark model under set #1A.

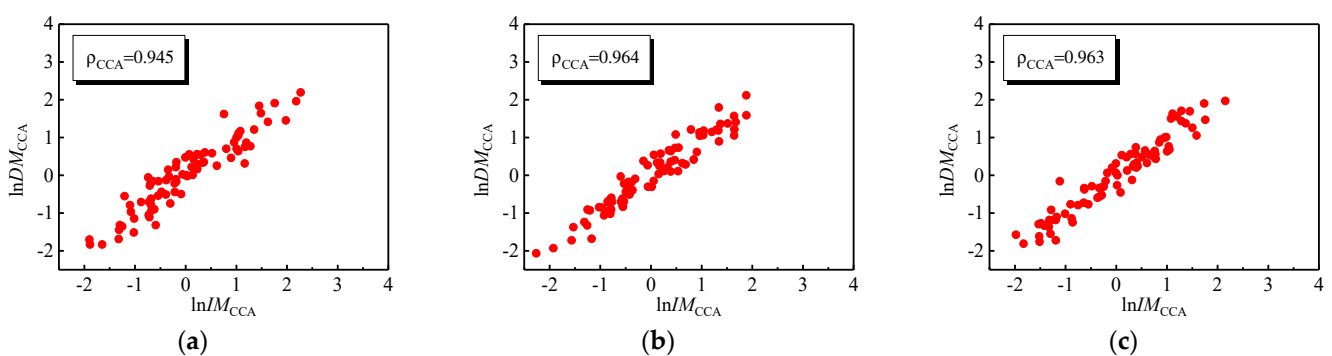


Figure 10. Canonical correlation analysis between  $\ln IMs$  and  $\ln DMs$ : (a)  $T = 0.2$  s; (b)  $T = 1.0$  s; (c)  $T = 5.0$  s.

As previously mentioned, the canonical correlation coefficient  $\rho_{CCA}$  undergoes changes with the canonical coefficients. In order to examine the influence of canonical coefficients, Figure 11 displays the canonical coefficients of the benchmark model derived from the input variables and the response variables. The results show that the canonical coefficients of  $\ln a_{rs}$ ,  $\ln v_{sq}$  and  $\ln SI$  have larger changes over the entire frequency range and are control variables of canonical correlation variables  $\ln IM_{CCA}$ . It can be seen that the canonical coefficients of  $\ln u_{max}$  decrease with  $\ln E_H$  increasing at 0.1–6 s, and vice versa.



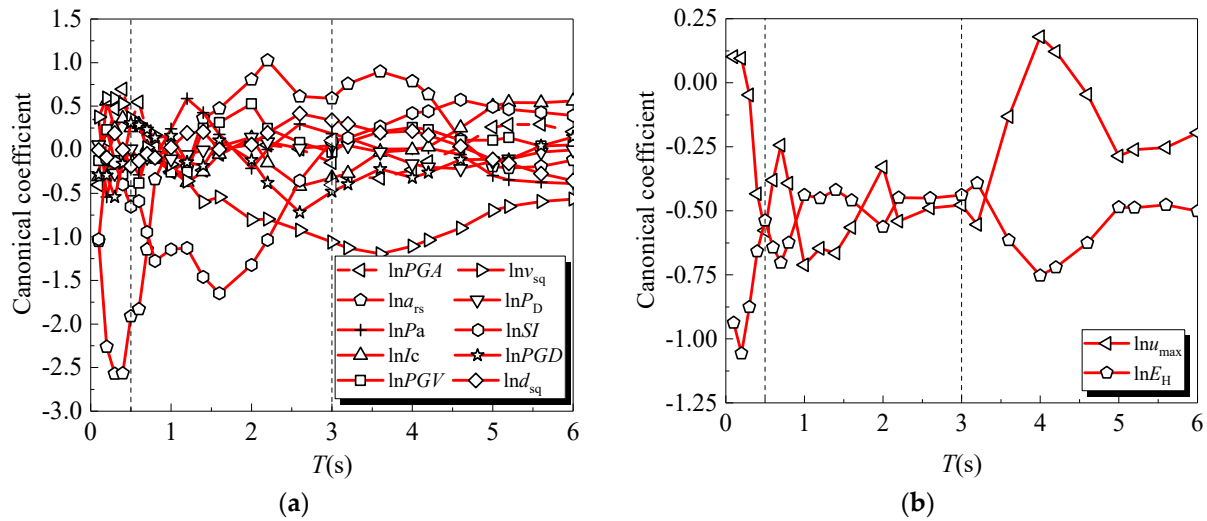


Figure 11. The canonical correlation vectors: (a) lnIMs; (b) lnDMs.

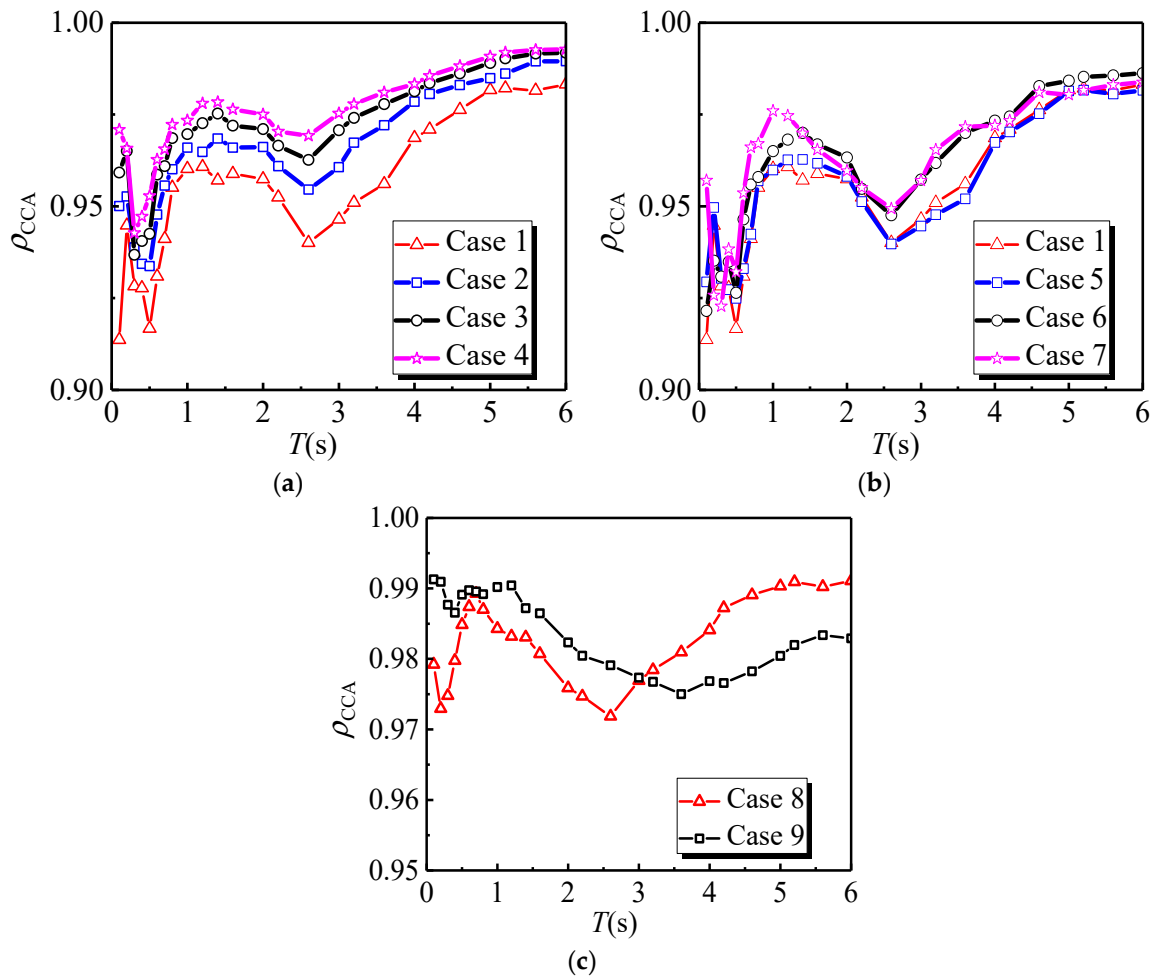
4.5. The Effect of Canonical Correlation Analysis

Nine conditions are established based on different models,  $R$  values and ground motions, as indicated in Table 2. Cases 1–4 represent variations in  $R$  for Model 1 within database set #1B. Cases 5–7 illustrate different models at  $R = 2$  within database set 1#B. Cases 8–9 show different ground motions at  $R = 5$  for Model 4.

Table 2. Number of conditions of different models at different  $R$  under different ground motions.

Conditions	GMs	$R$	The Model Number
Case 1	set #1B	2	Model 1
Case 2	set #1B	3	Model 1
Case 3	set #1B	4	Model 1
Case 4	set #1B	5	Model 1
Case 5	set #1B	2	Model 2
Case 6	set #1B	2	Model 3
Case 7	set #1B	2	Model 4
Case 8	set #1B	5	Model 4
Case 9	set #1A	5	Model 4

The effect of  $\rho_{CCA}-T$  for different models at varying  $R$  values under different ground motions is depicted in Figure 12. It is concluded that  $\rho_{CCA}$  for all cases exceeds 0.9 for the entire frequency region. The correlations generally increase with an increase  $R$  for the same model at the identical period  $T$ , with marginal variations noted for different models at  $R = 2$ . Consequently, the effect of  $\rho_{CCA}$  for different models appears to be less pronounced compared to the different values of  $R$ . Figure 12 also demonstrates significant fluctuations of  $\rho_{CCA}$  arising from nonlinear analysis results under different ground motions. Notably, at  $R = 5$  of Model 4 within database set #1A and set #1B,  $\rho_{CCA}$  surpasses 0.97 for the entire frequency region.



**Figure 12.** The effect of canonical correlation analysis: (a) different  $R$  values; (b) different models; (c) different ground motions.

Above all,  $\rho_{CCA}$  varies concerning different models,  $R$  and ground motions. Nevertheless, the correlations between  $\ln IM_{CCA}$  and  $\ln DM_{CCA}$  are sufficiently high for the entire structural period region, and the compound IM is introduced as a novel intensity measure, rendering it available for a seismic analysis of the SDOF systems. Subsequently, the MDOF systems are evaluated in the subsequent section to portray more intricate nonlinear analysis results than the SDOF systems are incapable of presenting.

## 5. Evaluation of Damage Potential of MDOF Systems

### 5.1. MDOF Systems

The method of CCA was conducted to assess the relationship between the IMs and DMs for SDOF systems. However, such analyses fall short in capturing the complexities introduced by a more intricate nonlinear dynamic history analysis in the MDOF systems. In this investigation, two MDOF systems, featuring four and eight stories, were individually examined. The two-dimensional models, illustrated in Figure 13, consist of three bay frames and were implemented by using OpenSEES. A simplified representation of the RC moment frame system, incorporating key elements, hinges and joints, geometry and dimensions, followed the frameworks proposed by Haselton et al. [55], Liel et al. [56] and Nasrolahzadeh et al. [57]. Lumped plasticity elements were employed to capture the flexural behavior of beams and columns. Plastic hinges at the beam and column ends constrained the finite joint sizes, while beam–column joints were modeled with a joint shear spring. The joint 2D element in OpenSEES effectively captured these behaviors, contributing to the construction of a comprehensive two-dimensional beam–column joint element. Outside

the plastic hinge region, elastic beam–column elements described the material behavior. The building foundations were assumed to have rotational springs, utilizing zero-length elements to simulate typical grade beam designs and soil stiffnesses. The behavior of the rotational spring is generally defined as the trilinear backbone curve and associated hysteretic rules with degrading strength and stiffness developed by Ibarra et al. [58]. RC moment frame systems comprised frame structures and leaning columns interconnected by truss elements. The destabilizing P-Δ effects were accounted for by applying gravity loads on a leaning column in the analysis model. One important design variable is the ratio of tributary areas for gravity and lateral loads, which is primarily affected by whether the building is designed as a space or perimeter frame system. The fundamental periods for the four-story and eight-story RC moment frame systems were determined to be 0.94 s and 1.8 s, respectively.

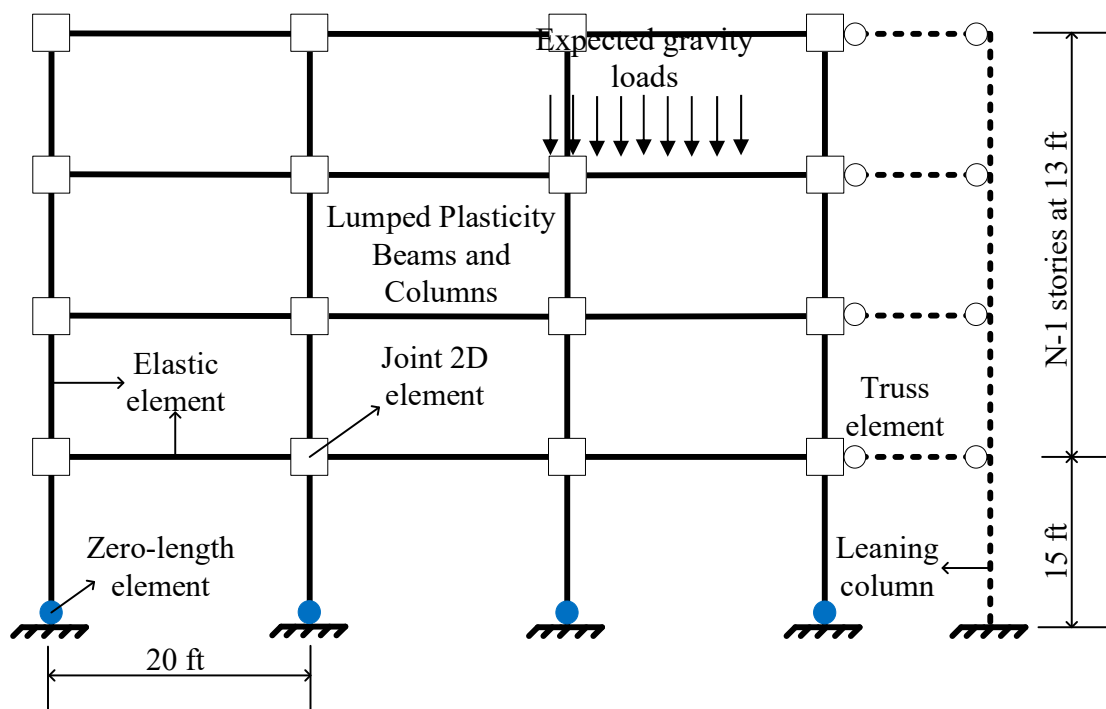


Figure 13. A simplified  $N$ -story model of RC moment frame system.

In order to describe the structural response, a global damage index was proposed by Park and Ang [59]. In this study, the modified Park and Ang damage index, as developed by Kunnath et al. [60], is employed. The general local damage measure is defined for each element as follows:

$$DI = \frac{\theta_m - \theta_r}{\theta_u - \theta_r} + \frac{\beta}{M_y \theta_u} \int E_h \quad (34)$$

where  $\theta_m$  is the maximum rotation;  $\theta_r$  is the recoverable rotation at unloading;  $\theta_u$  is the ultimate rotation capacity;  $\beta = 0.15$  is strength degrading parameter;  $M_y$  is the yield moment; and  $E_h$  is the hysteretic energy.

The overall structural damage index is abbreviated to OSDI, capable of transforming local indices into global indices through the energy weighting factor in the following formula:

$$OSDI = \sum (\lambda_i)_{\text{Story}} (DI_i)_{\text{Story}} \quad (35)$$

$$(\lambda_i)_{\text{Story}} = \left( \frac{E_i}{\sum E_i} \right)_{\text{Story}} \quad (36)$$

where  $\lambda_i$  and  $E_i$  are the energy weighting factor and the total energy absorbed by the  $i$ th story, respectively.

### 5.2. Determination of the Compound IMs for Representing the Ground Motion Damage Potential

Building upon the canonical correlation analysis of the SDOF systems, two sets of 40 pairs of horizontal bidirectional ground motions were adopted as input. The correlations between 10 candidate IMs and the OSDI for two RC moment frame systems are presented in Table 3. It can be seen that  $IM_{CCA}$  exhibits a relatively high correlation with the OSDI, and the canonical correlation coefficients are 0.900 and 0.917 with respect to four-story and eight-story frames, respectively. These coefficients surpass those of individual candidate IMs and  $IM_{PCA}$ . Consequently,  $IM_{CCA}$  is identified as the optimal IM for describing the ground motion damage potential in this study.

**Table 3.** Correlation coefficients between IMs and OSDI in logarithmic scale.

Story	PGA	$a_{rs}$	$P_a$	$I_c$	PGV	$v_{sq}$	$P_D$	SI	PGD	$d_{sq}$	$IM_{PCA}$	$IM_{CCA}$
four	0.716	0.827	0.800	0.784	0.873	0.837	0.833	0.896	0.804	0.756	0.836	0.900
eight	0.590	0.755	0.773	0.806	0.861	0.890	0.860	0.873	0.853	0.827	0.873	0.917

As stated previously, the method of CCA is evaluated here to identify the optimal combinations of 10 IMs for assessing ground motion damage potential. These hybrid parameters can fully capture the characteristic of amplitude, spectrum and duration of the ground motion. Importantly, they exhibit a relatively stable correlation with the OSDI, irrespective of the ground motion database or the number of stories.

Subsequently, the ground motion damage potential is quantitatively determined by the candidate 10 IMs and the corresponding canonical coefficient:

$$P_j = \sum_{i=1}^{10} a_i (\ln IM_{i,j} - \mu_{\ln IM_{i,j}}) \quad (37)$$

where  $P_j$  is the damage potential of the  $j$ th earthquake record;  $\ln IM_{i,j}$  and  $\mu_{\ln IM_{i,j}}$  are the values of the  $i$ th  $\ln IM$  of the  $j$ th ground motion record corresponding to the mean value; and  $a_i$  is the coefficient of potential measures, the value of which follows below:

$$a_i = \begin{cases} [-0.186, 0.385, -0.150, 0.094, 0.360, -0.134, -0.070, 1.068, -0.205, 0.141]^T & \text{for four-story} \\ [-0.233, -0.391, -0.350, 0.642, 0.032, 0.178, 0.105, 0.803, 0.424, -0.247]^T & \text{for eight-story} \end{cases} \quad (38)$$

Therefore, the combined IMs exhibit a robust correlation with the OSDI, providing insight into the ground motion damage potential with respect to four-story and eight-story frames. Employing this approach in future studies may be achieved via improved compound IMs for the evaluation of seismic damage by considering more new models. Compared with the conventional methods, if a new model is considered in a further study, then the coefficients in Equation (38) should be re-calculated. However, the presented approach is an adaptive method, and the main conclusions will not change due to the change in the coefficients.

## 6. Conclusions

This study presents the correlation between intragroup and intergroup IMs, as well as the correlation between various IMs and DMs in an effort to identify the optimal IM for seismic analysis. However, it is observed that no single IM exhibits satisfactory correlation with the DMs for all frequency ranges. Therefore, the compound IMs were determined through the methods of PCA and CCA, both of which demonstrated a high correlation with the DMs. On the basis of the CCA, the compound IMs were utilized as potential measures

to evaluate the damage potential of ground motions for MDOF systems. The key findings are summarized as follows:

1. The intragroup IMs exhibit significant correlation among them in the logarithmic scale. Specifically, the acceleration-related indices demonstrate high correlation, as do those associated with velocity-related indices, while the displacement-related indices are highly correlated among themselves. Conversely, the intergroup IMs reveal low correlation, as the correlation between the acceleration-related indices with both velocity-related indices and displacement-related indices is weak. However, it is observed that the correlation between the velocity-related indices and displacement-related indices demonstrates a relatively higher improvement.
2. For the SDOF systems, the acceleration-related indices ( $PGA$ ,  $I_A$ ,  $P_a$ ,  $I_c$  and  $I_a$ ) demonstrate a notable correlation with the DMs within the acceleration region ( $T = 0\text{--}0.5$  s). Similarly, the velocity-related indices ( $PGV$ ,  $v_{sq}$ ,  $P_v$ ,  $P_D$ ,  $I_F$ ,  $SI$  and  $I_v$ ) exhibit a high correlation with the DMs within the intermediate period ( $T = 0.5\text{--}3$  s). Furthermore, the displacement-related indices ( $PGD$ ,  $d_{sq}$ ,  $P_d$  and  $I_d$ ) correlate significantly with the DMs during the long period ( $T > 3$  s).
3. The compound IM ( $\mathbf{IM} = [PGA, a_{rs}, P_a, I_c, PGV, v_{sq}, P_D, SI, PGD, d_{sq}]$ ) determined through the method of PCA proves to be more adept at capturing the characteristics of ground motions. The correlation between the compound IMs and DMs for the SDOF systems is higher and more stable for all the frequency ranges compared to that of a single IM. This observation serves as a crucial foundation for selecting the optimal IM in earthquake engineering research.
4. The compound IMs determined through the method of CCA exhibit a high correlation with the DMs for the SDOF systems. Notably, the canonical coefficients of  $\ln a_{rs}$ ,  $\ln v_{sq}$  and  $\ln SI$  have larger changes over the entire frequency range, serving as control variables in the canonical correlation analysis. It is observed that to enhance the canonical correlation coefficient  $\rho_{CCA}$ , there is a decrease in the canonical coefficients of  $\ln u_{\max}$  with an increase in those of  $\ln E_H$ , and vice versa.
5. The canonical correlation coefficient  $\rho_{CCA}$  varies concerning different models,  $R$  values and ground motions for the SDOF systems. Generally, it increases with an elevated  $R$  for the same model at a given period  $T$  and shows minimal variation for the different model at the same  $R$ . Notably, the  $\rho_{CCA}$  undergoes significant changes in the nonlinear analysis results under different ground motions. However, it consistently remains high for all frequency regions, irrespective of the changes applied.
6. The compound IMs determined through the method of CCA serve as potential measures for assessing the damage potential of ground motions for the MDOF systems. The correlation of  $IM_{CCA}$  with the OSDI surpasses that of individual candidate IMs and  $IM_{PCA}$ . This approach proves valuable in the selection of unfavorable ground motions and in predicting the response of nonlinear analysis.

**Author Contributions:** Conceptualization, T.L.; methodology, D.L.; software, T.L.; validation, T.L. and D.L.; formal analysis, D.L.; investigation, T.L.; resources, T.L.; data curation, T.L.; writing—original draft preparation, T.L.; writing—review and editing, D.L.; visualization, T.L.; supervision, D.L.; project administration, T.L.; funding acquisition, T.L. and D.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** The financial support received from the National Science Foundation of China (Grant Nos. 52078176 and 51678209) and the Jiangsu Provincial Double Innovation Program (Grant No. JSSCBS20210992) are gratefully appreciated.

**Data Availability Statement:** The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

Abbreviation	Descriptions
PCA	Principal component analysis
CCA	Canonical correlation analysis
IM	Intensity measure
DM	Demand measure
PGA	Peak ground acceleration
$I_A$	Arias intensity
$a_{sq}$	Square acceleration
$a_{rs}$	Root square acceleration
$P_a$	Mean square acceleration
$a_{rms}$	Root-mean square acceleration
$I_c$	Characteristic intensity
$I_a$	Riddell acceleration intensity
PGV	Peak ground velocity
$v_{sq}$	Square velocity
$v_{rs}$	Root square velocity
$P_v$	Mean square velocity
$v_{rms}$	Root mean square velocity
$P_D$	Potential destructiveness
$I_F$	Fajfar intensity
$SI$	Housner spectrum intensity
$I_V$	Riddell velocity intensity
PGD	Peak ground displacement
$d_{sq}$	Square displacement
$d_{rs}$	Root square displacement
$P_d$	Mean square displacement
$d_{rms}$	Root mean square displacement
$I_d$	Riddell displacement intensity

## References

- Porter, K.A. An overview of PEER's performance-based earthquake engineering methodology. In Proceedings of the Ninth International Conference on Applications of Statistics and Probability in Civil Engineering, Francisco, CA, USA, 6–9 July 2003.
- Giovenale, P.; Cornell, C.A.; Esteva, L. Comparing the adequacy of alternative ground motion intensity measures for the estimation of structural responses. *Earthq. Eng. Struct. Dyn.* **2004**, *33*, 951–979. [[CrossRef](#)]
- Housner, G.W. Measures of Severity of Ground Shaking. In Proceedings of the U.S. Conference on Earthquake Engineering, Ann Arbor, MI, USA, 18–20 June 1975; pp. 25–33.
- Housner, G.W.; Jennings, P.C. *Earthquake Design Criteria*; Earthquake Engineering Research Institute: Berkeley, CA, USA, 1982.
- Zhai, C.H.; Xie, L.L. A new approach of selecting real input ground motions for seismic design: The most unfavourable real seismic design ground motions. *Earthq. Eng. Struct. Dyn.* **2007**, *36*, 1009–1027. [[CrossRef](#)]
- Kurama, Y.C.; Farrow, K.T. Ground motion scaling methods for different site conditions and structure characteristics. *Earthq. Eng. Struct. Dyn.* **2003**, *32*, 2425–2450. [[CrossRef](#)]
- Akkar, S.; Özen, Ö. Effect of peak ground velocity on deformation demands for SDOF systems. *Earthq. Eng. Struct. Dyn.* **2005**, *34*, 1551–1571. [[CrossRef](#)]
- Elenas, A.; Meskouris, K. Correlation study between seismic acceleration parameters and damage indices of structures. *Eng. Struct.* **2001**, *23*, 698–704. [[CrossRef](#)]
- Akkar, S.; Sucuoğlu, H.; Yakut, A. Displacement-based fragility functions for low- and midrise ordinary concrete buildings. *Earthq. Spectra* **2005**, *21*, 901–927. [[CrossRef](#)]
- Vamvatsikos, D.; Cornell, C.A. Developing efficient scalar and vector intensity measures for IDA capacity estimation by incorporating elastic spectral shape information. *Earthq. Eng. Struct. Dyn.* **2005**, *34*, 1573–1600. [[CrossRef](#)]
- Luco, N.; Cornell, C.A. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthq. Spectra* **2007**, *23*, 357–392. [[CrossRef](#)]
- Shome, N.; Cornell, C.A. *Probabilistic Seismic Demand Analysis of Non-Linear Structures*; Report No. RMS-35; RMS Program Stanford University: Stanford, CA, USA, 1999.



13. Cordova, P.P.; Deierlein, G.G.; Mehanny, S.S.F.; Cornell, C.A. Development of a two-parameter seismic intensity measure and probabilistic assessment procedure. In *The Second US-Japan Workshop on Performance-Based Earthquake Engineering Methodology for Reinforced Concrete Building Structures*; Pacific Earthquake Engineering Research Center, University of California: Berkeley, CA, USA, 2000; pp. 187–206.
14. Bianchini, M.; Diotallevi, P.P.; Baker, J.W. Prediction of inelastic structural response using an average of spectral accelerations. In *Proceedings of the 10th International Conference on Structural Safety and Reliability (ICOSSAR09)*, Osaka, Japan, 13–17 September 2009; pp. 13–17.
15. Zhou, Z.; Yu, X.H.; Lu, D.G. Identifying Optimal Intensity Measures for Predicting Damage Potential of Mainshock–Aftershock Sequences. *Appl. Sci.* **2020**, *10*, 6795. [[CrossRef](#)]
16. Padgett, J.E.; Nielson, B.G.; DesRoches, R. Selection of optimal intensity measures in probabilistic seismic demand models of highway bridge portfolios. *Earthq. Eng. Struct. Dyn.* **2008**, *37*, 711–725. [[CrossRef](#)]
17. Mehanny, S.S. A broad-range power-law form scalar-based seismic intensity measure. *Eng. Struct.* **2009**, *31*, 1354–1368. [[CrossRef](#)]
18. Hariri-Ardebili, M.; Saouma, V. Probabilistic seismic demand model and optimal intensity measure for concrete dams. *Struct. Saf.* **2016**, *59*, 67–85. [[CrossRef](#)]
19. Kostinakis, K.; Athanatopoulou, A.; Morfidis, K. Correlation between ground motion intensity measures and seismic damage of 3D R/C buildings. *Eng. Struct.* **2015**, *82*, 151–167. [[CrossRef](#)]
20. Kostinakis, K.; Athanatopoulou, A. Incremental dynamic analysis applied to assessment of structure-specific earthquake IMs in 3D R/C buildings. *Eng. Struct.* **2016**, *125*, 300–312. [[CrossRef](#)]
21. Zavala, N.; Bojórquez, E.; Barraza, M.; Bojórquez, J.; Villela, A.; Campos, J.; Torres, J.; Sánchez, R.; Carvajal, J. Vector-valued intensity measures based on spectral shape to predict seismic fragility surfaces in reinforced concrete buildings. *Buildings* **2023**, *13*, 137. [[CrossRef](#)]
22. Ciano, M.; Giofrè, M.; Grigoriu, M. A novel approach to improve accuracy in seismic fragility analysis: The modified intensity measure method. *Probabilistic Eng. Mech.* **2022**, *69*, 103301. [[CrossRef](#)]
23. Yang, D.; Pan, J.; Li, G. Non-structure-specific intensity measure parameters and characteristic period of near-fault ground motions. *Earthq. Eng. Struct. Dyn.* **2009**, *38*, 1257–1280. [[CrossRef](#)]
24. Riddell, R.; Newmark, N.M. *Statistical Analysis of the Response of Nonlinear Systems Subjected to Earthquake*; University of Illinois at Urbana-Champaign: Champaign County, IL, USA, 1979; p. 468.
25. Riddell, R.; Hidalgo, P.; Cruz, E. Response modification factors for earthquake resistant design of short period buildings. *Earthq. Spectra* **1989**, *5*, 571–590. [[CrossRef](#)]
26. Riddell, R.; Garcia, J.E. Hysteretic energy spectrum and damage control. *Earthq. Eng. Struct. Dyn.* **2001**, *30*, 1791–1816. [[CrossRef](#)]
27. Riddell, R.; Garcia, J.E.; Garces, E. Inelastic deformation response of SDOF systems subjected to earthquakes. *Earthq. Eng. Struct. Dyn.* **2002**, *31*, 515–538. [[CrossRef](#)]
28. Riddell, R. On ground motion intensity indices. *Earthq. Spectra* **2007**, *23*, 147–173. [[CrossRef](#)]
29. Zhai, C.H.; Xie, L.L.; Li, S. A new method for estimating strong ground motion damage potential for structures. In *Proceedings of the Ninth International Symposium on Structural Engineering for Young Experts (ISSEYE-9)*, Xiamen, China, 18–21 August 2006; pp. 758–762.
30. Zhai, C.; Chang, Z.; Li, S.; Xie, L. Selection of the most unfavorable real ground motions for low-and mid-rise RC frame structures. *J. Earthq. Eng.* **2013**, *17*, 1233–1251. [[CrossRef](#)]
31. Ozmen, H.B. Developing hybrid parameters for measuring damage potential of earthquake records: Case for RC building stock. *Bull. Earthq. Eng.* **2017**, *15*, 3083–3101. [[CrossRef](#)]
32. Chen, Z.; Yu, W.; Zhu, H.; Xie, L. Ranking method of the severest input ground motion for underground structures based on composite ground motion intensity measures. *Soil Dyn. Earthq. Eng.* **2023**, *168*, 107828. [[CrossRef](#)]
33. Liu, T.-T.; Lu, D.-G.; Yu, X.-H. Development of a compound intensity measure using partial least-squares regression and its statistical evaluation based on probabilistic seismic demand analysis. *Soil Dyn. Earthq. Eng.* **2019**, *125*, 105725. [[CrossRef](#)]
34. Liu, B.; Hu, J.; Xie, L. Exploratory factor analysis-based method to develop compound intensity measures for predicting potential structural damage of ground motion. *Bull. Earthq. Eng.* **2022**, *20*, 7107–7135. [[CrossRef](#)]
35. Chen, C.J.; Geng, P.; Gu, W.Q.; Lu, Z.; Ren, B. Assessment of tunnel damage potential by ground motion using canonical correlation analysis. *Earthq. Struct.* **2022**, *23*, 259–269.
36. Narasimhan, S.; Wang, M.; Pandey, M. Principal component analysis for predicting the response of nonlinear base-isolated buildings. *Earthq. Spectra* **2009**, *25*, 93–115. [[CrossRef](#)]
37. Jolliffe, I.T. *Principal Component Analysis*; Springer: New York, NY, USA, 1986.
38. Liu, T.-T.; Yu, X.-H.; Lu, D.-G. An approach to develop compound intensity measures for prediction of damage potential of earthquake records using canonical correlation analysis. *J. Earthq. Eng.* **2020**, *24*, 1747–1770. [[CrossRef](#)]
39. Hotelling, H. The most predictable criterion. *J. Educ. Psychol.* **1935**, *26*, 139–142. [[CrossRef](#)]
40. Lee, K.; Yoo, J.K. Canonical correlation analysis through linear modeling. *Aust. N. Z. J. Stat.* **2014**, *56*, 59–72. [[CrossRef](#)]
41. Baker, J.W.; Lin, T.; Shahi, S.K.; Jayaram, N. *New Ground Motion Selection Procedures and Selected Motions for the PEER Transportation Research Program*; Peer Report; Pacific Earthquake Engineering Research Center, University of California: Berkeley, CA, USA, 2011.
42. Chiou, B.; Darragh, R.; Gregor, N.; Silva, W. NGA project strong-motion database. *Earthq. Spectra* **2008**, *24*, 23–44. [[CrossRef](#)]

43. Wang, Z.; Padgett, J.E.; Dueñas-Osorio, L. Risk-consistent calibration of load factors for the design of reinforced concrete bridges under the combined effects of earthquake and scour hazards. *Eng. Struct.* **2014**, *79*, 86–95. [[CrossRef](#)]
44. Ramanathan, K.; Padgett, J.E.; DesRoches, R. Temporal evolution of seismic fragility curves for concrete box-girder bridges in California. *Eng. Struct.* **2015**, *97*, 29–46. [[CrossRef](#)]
45. Konstantinidis, D.; Nikfar, F. Seismic response of sliding equipment and contents in base-isolated buildings subjected to broadband ground motions. *Earthq. Eng. Struct. Dyn.* **2015**, *44*, 865–887. [[CrossRef](#)]
46. Hosseini, R.; Rashidi, M.; Bulajić, B.Đ.; Arani, K.K. Multi-objective optimization of three different SMA-LRBs for seismic protection of a benchmark highway bridge against real and synthetic ground motions. *Appl. Sci.* **2020**, *10*, 4076. [[CrossRef](#)]
47. Cantagallo, C.; Camata, G.; Spacone, E.; Corotis, R. The variability of deformation demand with ground motion intensity. *Probabilistic Eng. Mech.* **2012**, *28*, 59–65. [[CrossRef](#)]
48. Kramer, S.L. *Geotechnical Earthquake Engineering*; Prentice-Hall Inc.: Englewood Cliffs, NJ, USA, 1996.
49. Ozmen, H.B.; Inel, M. Damage potential of earthquake records for RC building stock. *Earthq. Struct.* **2016**, *10*, 1315–1330. [[CrossRef](#)]
50. Xu, M.-Y.; Lu, D.-G.; Yu, X.-H.; Jia, M.-M. Selection of optimal seismic intensity measures using fuzzy-probabilistic seismic demand analysis and fuzzy multi-criteria decision approach. *Soil Dyn. Earthq. Eng.* **2023**, *164*, 107615.
51. Chopra, A.K. *Dynamics of Structures*; Prentice Hall Inc.: Englewood Cliffs, NJ, USA, 1995.
52. Morfidis, K.; Kostinakis, K. Seismic parameters' combinations for the optimum prediction of the damage state of R/C buildings using neural networks. *Adv. Eng. Softw.* **2017**, *106*, 1–16. [[CrossRef](#)]
53. Ghotbi, A.R.; Taciroglu, E. Ground motion selection based on a multi-intensity-measure conditioning approach with emphasis on diverse earthquake contents. *Earthq. Eng. Struct. Dyn.* **2021**, *50*, 1378–1394. [[CrossRef](#)]
54. Ye, L.; Ma, Q.; Miao, Z.; Guan, H.; Zhuge, Y. Numerical and comparative study of earthquake intensity indices in seismic analysis. *Struct. Des. Tall Spec. Build.* **2013**, *22*, 362–381. [[CrossRef](#)]
55. Haselton, C.B.; Liel, A.B.; Deierlein, G.G.; Dean, B.S.; Chou, J.H. Seismic collapse safety of reinforced concrete buildings: I. assessment of ductile moment frames. *J. Struct. Eng.* **2011**, *137*, 481–491. [[CrossRef](#)]
56. Liel, A.B.; Haselton, C.B.; Deierlein, G.G. Seismic collapse safety of reinforced concrete buildings: II. comparative assessment of non-ductile and ductile moment frames. *J. Struct. Eng.* **2011**, *137*, 492–502. [[CrossRef](#)]
57. Nasrollahzadeh, K.; Hariri-Ardebili, M.A.; Kiani, H.; Mahdavi, G. An integrated sensitivity and uncertainty quantification of fragility functions in RC frames. *Sustainability* **2022**, *14*, 13082. [[CrossRef](#)]
58. Ibarra, L.F.; Medina, R.A.; Krawinkler, H. Hysteretic models that incorporate strength and stiffness deterioration. *Earthq. Eng. Struct. Dyn.* **2005**, *34*, 1489–1511. [[CrossRef](#)]
59. Park, Y.; Ang, A.H.; Wen, Y.K. Seismic damage analysis of reinforced concrete buildings. *J. Struct. Eng.* **1985**, *111*, 740–757. [[CrossRef](#)]
60. Kunnath, S.K.; Reinhorn, A.M.; Lobo, R.F. *IDARC Version 3.0: A Program for the Inelastic Damage Analysis of Reinforced Concrete Structures*; National Center for Earthquake Engineering Research: Buffalo, NY, USA, 1992.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.