Structural Safety 100 (2023) 102282

Contents lists available at ScienceDirect

Structural Safety

journal homepage: www.elsevier.com/locate/strusafe

A recurrent-neural-network-based generalized ground-motion model for the Chilean subduction seismic environment

Jawad Fayaz [a,b,*], Miguel Medalla [c], Pablo Torres-Rodas [d], Carmine Galasso [a]

[a] Department of Civil, Environmental, and Geomatic Engineering, University College London, London, WC1E 6BT, UK
[b] School of Computing, Engineering, and Digital Technologies, Teesside University, Middlesbrough, TS1 3BX, UK
[c] Facultad de Ingeniería y Ciencias Aplicadas, Universidad de los Andes, Santiago, Chile
[d] Universidad San Francisco de Quito, Colegio de Ciencias e Ingenierías, Campus Cumbayá, Quito, 170901, Ecuador

ARTICLE INFO

Keywords:
Generalized ground motion model
Recurrent neural networks
Deep learning
Subduction ground motions
Long short-term memory

ABSTRACT

This paper proposes a deep learning-based generalized ground motion model (GGMM) for interface and intraslab subduction earthquakes recorded in Chile. A total of ~7000 ground-motion records from ~1700 events are used to train the proposed GGMM. Unlike common ground-motion models (GMMs), which generally consider individual ground-motion intensity measures such as peak ground acceleration and spectral accelerations at given structural periods, the proposed GGMM is based on a data-driven framework that coherently uses recurrent neural networks (RNNs) and hierarchical mixed-effects regression to output a cross-dependent vector of 35 ground-motion intensity measures (denoted as **IM**). The **IM** vector includes geometric mean of Arias intensity, peak ground velocity, peak ground acceleration, and significant duration (denoted as $I_{a_{geom}}$, PGV_{geom}, PGA_{geom}, and $D_{5-95_{geom}}$, respectively), and *RotD*50 spectral accelerations at 31 periods between 0.05 and 5 s for a 5 % damped oscillator (denoted as $S_a(T)$). The inputs to the GGMM include six causal seismic source and site parameters, including fault slab mechanism, moment magnitude, closest rupture distance, Joyne-Boore distance, soil shear-wave velocity, and hypocentral depth. The statistical evaluation of the proposed GGMM shows high prediction power with $R^2 > 0.7$ for most IMs while maintaining the cross-IM dependencies. Furthermore, the GGMM is carefully compared against two state-of-the-art Chilean GMMs, showing that the proposed GGMM leads to better goodness of fit for all periods of $S_a(T)$ compared to the two considered GMMs (on average 0.2 higher R^2). Finally, the GGMM is implemented to select hazard-consistent ground motions for nonlinear time history analysis of a sophisticated finite-element model of a 20-story steel special moment-resisting frame. Results of this analysis are statistically compared against those for hazard-consistent ground motions selected based on the conditional mean spectrum (CMS) approach. In general, it is observed that the drift demands computed using the two approaches cannot be considered statistically similar and the GGMM leads to higher demands.

1. Introduction

Regions such as the Pacific Coast of South and North America (including Chile, Mexico, and the US), and Japan are known to be seismically active with long-rupture subduction zones. Earthquakes arising from subduction are generally very high in magnitude (i.e., megathrust). This means that densely populated cities (e.g., Mexico City, Santiago, Seattle, and Tokyo) can experience high-magnitude and damaging strong ground motions. In the specific case of Chile, near 17.5 million people [1] are exposed to major earthquakes. For instance, three out of the 15 largest earthquakes recorded in the world have occurred next to the Chilean Pacific coast [2]. These earthquakes have led to significant economic and life losses to local communities [2]. The Chilean seismicity is primarily explained by the subduction of the oceanic Nazca plate underneath the continental South American plate. This tectonic setting is considered one of the most active in the world, with annual convergency velocities close to 63 mm/year [3], capable of producing approximately one major event ($M_w > 7.0$) every ten years [4]. Within subduction seismic environments, there are two main types of mechanisms: interface and intraslab earthquakes. The primary notable difference between the two is that intraslab earthquakes typically occur at depths greater than those of the interface events. In Chile,

* Corresponding author at: Department of Civil, Environmental, and Geomatic Engineering, University College London, WC1E 6BT, UK.
E-mail address: j.fayaz@ucl.ac.uk (J. Fayaz).

https://doi.org/10.1016/j.strusafe.2022.102282
Received 30 October 2021; Received in revised form 21 June 2022; Accepted 6 September 2022
Available online 18 October 2022

earthquakes due to the former type occur at the contact between the Nazca and South American plates, while in the latter case, earthquakes occur within the oceanic plate. In the last fifteen years, large magnitude events faced by communities in Chile (e.g., 2010 Maule, $M_w = 8.8$; 2014 Iquique, $M_w = 8.2$; and 2015 Illapel, $M_w = 8.4$, with M_w being their moment magnitude) have motivated the field of seismology and earthquake engineering to undertake practical actions to mitigate seismic risk and particularly improve the seismic hazard assessment through better data, procedures, models, and tools.

Seismic hazard assessment is an essential step for performing probabilistic/deterministic analysis of the impacts of possible earthquakes. Ground-motion models (GMMs) are a critical tool in both probabilistic seismic hazard analysis (PSHA) and deterministic seismic hazard analysis (DSHA). In general, GMMs provide statistical estimations of ground-motion intensity measures (IMs) using the casual seismic source and site information (magnitude, rupture distance, soil shear-wave velocity, etc.). Typical IMs include peak ground acceleration (*PGA*), peak ground velocity (*PGV*), Arias intensity (I_a), strong motion duration (D_s) (in its various definitions), and spectral acceleration at different periods of vibration ($S_a(T)$) for a single-degree-of-freedom (SDOF) system. Several GMMs have been proposed in the past [5] for such IMs; however, $S_a(T)$ has been the primary IM that most GMMs have considered. Due to the difference in the earthquake characteristics between crustal and subduction zones, GMMs are derived separately for the two types of earthquakes, and specific GMMs are developed for the regions near subduction zones. Some of these GMMs are based on worldwide data (e.g., [6,7]) and others on regional data (e.g., Zhaeo *et al.* [8] for Japan; Lin and Lee [9] for Taiwan). Particularly in Chile, two recent studies have provided regression-based GMMs for subduction earthquakes in the region, i.e., Montalva *et al.* [10] and Idini *et al.* [11]. However, similar to the crustal zones, these GMMs are primarily based on $S_a(T)$ and independently developed for each SDOF natural period. Also, in the case of GMMs for other IMs, the models are developed separately, while the databases used by these studies have significant overlap. Hence, as mentioned by Fayaz *et al.* [12], these GMMs neglect the inherent correlations/dependencies between the IMs. Since the IMs of the same ground motion are expected to be cross-correlated/dependent, it is essential that the prediction models accurately maintain such correlations.

Moreover, the cross-IM correlation/dependence is often required for the performance-based seismic design and assessment of structures. For instance, it is extensively used to define target IMs for ground-motion simulation, selection, and modification for engineering applications (e.g., the generalized conditional intensity measure [13]). So far, these correlations are taken into account through *ad-hoc* models estimated in a subsequent analysis stage. For instance, Baker and Jayaram [14] computed the $S_a(T)$ correlations for crustal ground motions through the epsilon parameter/values (ε) and suggested a functional form to approximate these correlations over a range of periods. As an extension to this study, Candia *et al.* [15] developed $S_a(T)$ correlations for Chilean subduction earthquakes. Extending this concept to other IMs, Bradley [13] proposed a generalized conditional intensity measure (GCIM) for hazard consistent ground-motion selection. This model consists of developing the multivariate distribution of any set of ground-motion IMs conditioned on the occurrence of a specific IM obtained from PSHA. Bradley [13] also proposed a ground motion selection algorithm based on a random realization of the conditional multivariate distribution of IMs. One of the drawbacks of this approach is that there is not always a well-defined linear correlation among the different IMs. Moreover, the multivariate lognormal assumption of this approach needs further evaluation. Additionally, the application of the GCIM requires predefined GMMs and correlation structures for the considered IMs; although there have been some initiatives considering Chilean ground motions, they are only focused on the evaluation of $S_a(T)$ correlations and do not consider other IMs, such as I_a, D_s, *PGV*, etc.

For the case of crustal ground motions (using the NGA-West2 database - Enhancement of Next Generation Attenuation Relationships for Western US [16]), Fayaz *et al.* [12] proposed a novel approach in the form of a generalized ground-motion model (GGMM) based on a hybrid recurrent neural network (RNN) that estimates a vector of 29 IMs consisting of geometric means of I_a, *CAV*, D_{5-95}, and RotD50S_a [17] at 26 periods (for a 5 % damped oscillator). In addition, the covariance matrix adaptation evolution strategy (CMA-ES) was further used to propose inter-event and intra-event covariance matrices of the residuals between the recorded and RNN-predicted IMs. The model improves the current GMMs in the literature since it overcomes most of their statistical limitations (such as predetermined functional forms) and incorporates high order dependencies between the vector of the 29 IMs.

Due to the increase in computational resources and high predictive power, more studies are being conducted on utilizing deep-learning-based techniques in the realm of structural and earthquake engineering. Previously, Dhanya and Raghukanth [18] adopted an artificial neural network (ANN) in combination with a genetic algorithm (GA) to train a data-driven model to predict *PGA*, *PGV*, and spectral accelerations at 26 periods between 0.01 and 4 s. However, in that study, the authors utilized conventional feed-forward neural networks, which do not explicitly account for the internal cross-IM dependencies and only attempt to ensure that individual mean IM predictions agree well with observed IM values. Apart from this, studies like Datta et al. [19], Fayaz and Galasso [20], Chiang et al. [21], Hyun-Su [22], Liu and Dai [23], etc., have explored the utilization of convolutional neural networks (CNNs), variational autoencoders (VAE), and RNNs for the real-time estimation of source parameters (e.g earthquake magnitude and location), ground-motion intensity measures (e.g., *PGA*), and structural response parameters (e.g., drift levels). However, such research studies and their related computational tools aim to advance/improve earthquake early warning and are not particularly suitable for ground motion modelling and seismic hazard/risk analysis.

Motivated by the preceding discussion, this study extends the previous research by Fayaz et al. [12] and proposes an RNN- and mixed-effects-regression-based GGMM specifically for the earthquakes recorded in Chile (South America). Unlike previous studies that mainly rely on feed forward networks for modelling amplitude-based IMs and event parameters, this study is a novel attempt to develop an end-to-end framework using long short-term memory (LSTM)-based RNNs and hierarchical mixed-effects regression for generalized cross-dependent predictions of a vector of amplitude-, energy-, frequency-, and duration-based IMs. A total of ~7000 ground motion records from ~1700 events are used to train a data-driven framework of LSTM-based RNN model that uses six source and site parameters (including fault slab mechanism (*F*), magnitude (M_w), closest rupture distance (R_{rup}), Joyner-Boore distance (R_{JB}), soil shear-wave velocity (V_{s30}), and hypocentral depth (Z_{hyp}) to predict a vector of 35 IMs (denoted as **IM** hereafter), including geomean of I_a, *PGV*, *PGA*, D_{5-95} (denoted as $I_{a_{geom}}$, PGV_{geom}, PGA_{geom}, and $D_{5-95_{geom}}$, respectively), and RotD50S_a at 31 periods between 0.05 and 5 s (for a 5 % damped oscillator; denoted as $S_a(T)$). The RNN is particularly trained using an information-theory-based time-series cost function known as index of agreement (IA). Furthermore, the residuals between the true recorded **IM** and the RNN-based estimated **IM** are split into between-event and within-event covariance matrices using hierarchical mixed-effects regression and Pearson correlation analysis. Unlike previous studies, the utilization of advanced LSTM-based RNNs in this study alongwith IA as the cost function retains the internal cross-IM dependencies while showing high prediction power for the amplitude-, energy-, frequency-, and duration- based IMs. Furthermore, the use of mixed-effects regression enables the hierarchical division of the residuals. The proposed framework is then illustrated by selecting ground motions for nonlinear time-history analysis (NLTHA) of a 20-story steel special moment frame (SSMF) developed using the OpenSees [24]. The SSMF model is utilized to perform NLTHA using hazard-consistent ground motions selected based on the proposed

GGMM and conventional conditional mean spectrum (CMS)-based [25] approaches. A total of 180 ground motion components are selected for three hazard levels, *i.e.,* 475, 975, and 2475 years and used for NLTHA.

The primary goals and engineering novelties of this study are threefold: 1) propose a GGMM model to estimate the **IM** vector using a data-driven deep-learning-based model leading to good prediction power while explicitly maintaining the higher-order cross-dependencies between the **IM** components (which include amplitude-, energy-, frequency-, and duration- based IMs); 2) demonstrate the use of GGMM for selecting hazard-consistent ground motions using vector-valued objective function considering the **IM** components; and 3) assessing the applicability and efficiency of the proposed ground-motion selection approach in structural analysis. Therefore, this investigation seeks to provide further insights into the knowledge of subduction ground motions, develop a multi-criteria solution for estimating IMs, and demonstrate a practical way of implementing the proposed GGMM for record selection using a vector-valued objective function.

The proposed GGMM is a general tool that can also be helpful for other possible applications such as 1) the assessment of geotechnical structures (e.g., liquefaction-induced settlements, seismically-induced slope displacements), where the parameters of interest are commonly not only the spectral coordinates $S_a(T)$, but other IMs are required, such as *PGV* [26], and I_a [27]; 2) analysis and validation of simulated ground motions [28–31]; 3) vector-valued PSHA [32], etc. For instance, *CAV* is used in evaluating liquefaction-induced building settlements (e.g., [33,34]). Similarly, *PGV* and I_a are utilized for estimating seismically induced slope displacements and landslides (e.g., [27,35–38]). *CAV*, *PGV* and I_a are also useful assessing liquefaction triggering (e.g., [27,39–41]).

This paper is divided into seven sections, starting with this introductory section. The following section describes the ground-motion database used in this study. Next, the various modelling components of the proposed GGMM are presented in the subsequent sections. Specifically, the mean model using the proposed RNN and the covariance matrices for residuals are described first. For the sake of consistency, the models are explained first, and then their corresponding statistical evaluations are described. After the explanation of the proposed GGMM, its prediction power is statistically compared with other commonly used Chilean GMMs [10,11], and a spectral comparison against the CMS [25] is carried out through likelihood comparisons. Hence, the GGMM is assessed against the currently available models, and its advantages are discussed. Finally, an application of the GGMM for seismic NLTHA of a 20-story building is presented. Conclusions, limitations of this investigation, and recommendations for future work are outlined in the last section of the paper.

2. Ground-motion database

This study uses a set of ~7000 two-component ground-motion acceleration time series from ~1700 seismic events available in the SIBER-RISK strong motion database [42]. The locations of these events are shown in Fig. 1a, and the corresponding $M_w - R_{rup}$ metadata is presented in Fig. 1b. In terms of the two mechanisms, i.e., interface and intraslab, ~5000 ground motion records are from interface events, and ~2000 ground motion records are from intraslab events. As observed from Fig. 1b, the records include events with M_w ranging from 3.7 to 8.8, while R_{rup} ranges from 15 to 300 km. The statistical details of the other event parameters obtained from the metadata include i) $0.9 \leq R_{JB} \leq 298.2$ km; ii) $2 \leq Z_{hyp} \leq 213$ km; and iii) $108 \leq V_{s30} \leq 2127$ m/s.

As discussed in Candia *et al.* [15], SIBER-RISK strong-motion centre applies a third-order bandpass Butterworth filter with corner frequencies at 0.01 and 50 Hz to the acceleration records to remove spurious low-frequency acceleration base shifts. The records are baseline corrected by subtracting a piecewise-linear function to the data using a reversible jump Markov Chain Monte Carlo algorithm [43]. The RotD50S_a spectra ($S_a(T)$) and geomean significant duration $D_{5-95,geom}$ of the ground motion records are presented in Fig. 2a and 2b, respectively. Unlike previous GMMs in Chile, in this study RotD50S_a (current state-of-the-art IM [7,44]) is purposefully selected to represent spectral accelerations. In fact, unlike other definitions (such as geomean, average, etc.), RotD50S_a is independent of the orientation of the recording sensor and leads to the one-to-one causal relation between the source parameters and spectral acceleration. As observed from the figures, the spectra of the selected ground motion represent both low levels of shaking (RotD50S_a < 0.1 g) and high levels of shaking (RotD50S_a > 0.1 g), hence showing the vast range of accelerations used for training the proposed model. For the other IMs, geomean is used for consistency with the state-of-the-art and literature [45]. Furthermore, it can be observed from Fig. 2b the selected ground motions are on average ~40 s long and can go as high as 150 s thereby representing ground motions with different durations.

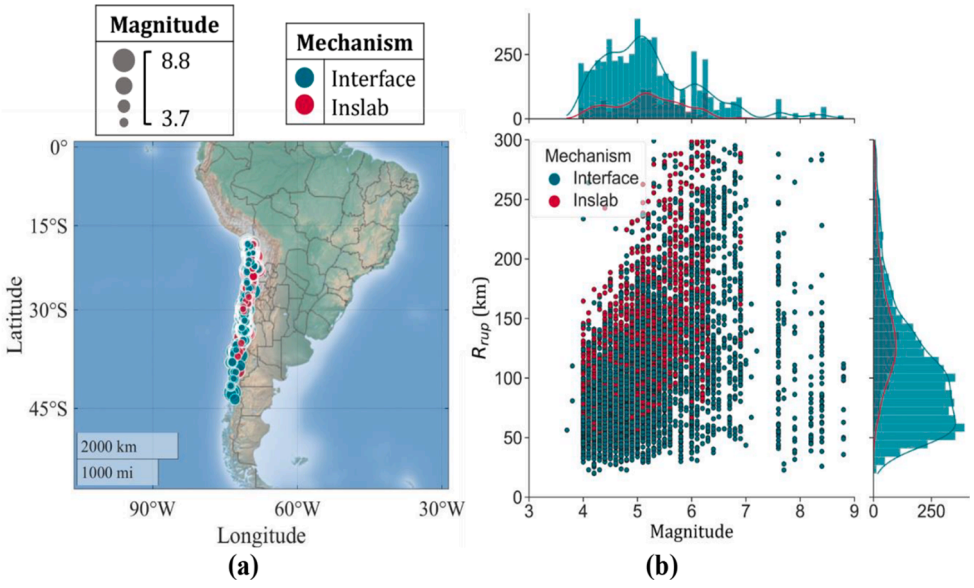

Fig. 1. Earthquake events: (a) epicenters; and (b) M_w-R_{rup} metadata.

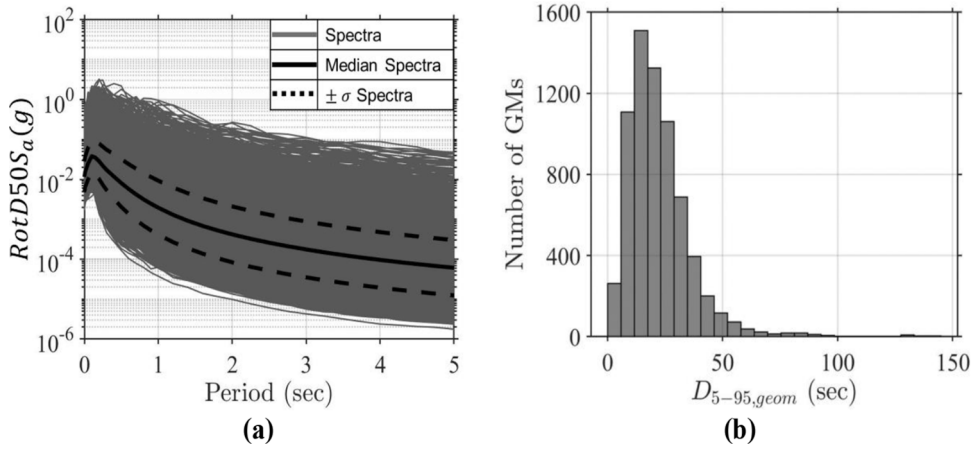

Fig. 2. Subduction database ground motions: a) RotD50S_a spectra; b) $D_{5-95,geom}$.

3. Mean model using recurrent neural network

Similar to Fayaz *et al.* [12], this study employs RNNs to predict the considered **IM** vector and model the IM sequence-dependent behaviour, which means that the values in the output vector are related to each other. The underlying principle of modelling sequences in RNNs is recurrency in artificial neural networks compared to conventional feedforward neural networks. The connections between nodes of an RNN form a directed graph along with sequential data. This is done by feeding back the output of a neural network layer at sequence step t to the input of the same network layer at the sequence step $t + 1$. This allows the RNN to exhibit sequential dynamic behaviour. Since the ground motion **IM** vector represents the IMs of the same ground motion, the values are dependent on each other. Hence, the **IM** vector can be modelled as a sequence of IMs where the individual IMs are dependent on each other. Therefore, an RNN is one of the best candidates to train an unbiased data-driven model that predicts a cross-dependent vector of IMs (**IM**).

Conventional RNNs still suffer from drawbacks, such as short-term memory and vanishing gradients [46], which can be remedied using LSTM units [46]. LSTMs modulate the information flow using internal mechanisms of cell and forget gates. The general details of the LSTM-RNN structures can be obtained from Fayaz *et al.* [12]. LSTM-RNN not only feeds forward but keeps an internal memory to adaptively process the sequences of inputs while maintaining the internal dependencies between the components of the output vector. Furthermore, the recurrent nature of an LSTM-RNN enables performing the same function for each input, copying and sending the data back to the network while producing the output simultaneously. Hence, in this study, LSTM-based RNN cells are utilized for training the neural network.

More specifically, LSTM consists of internal mechanisms called

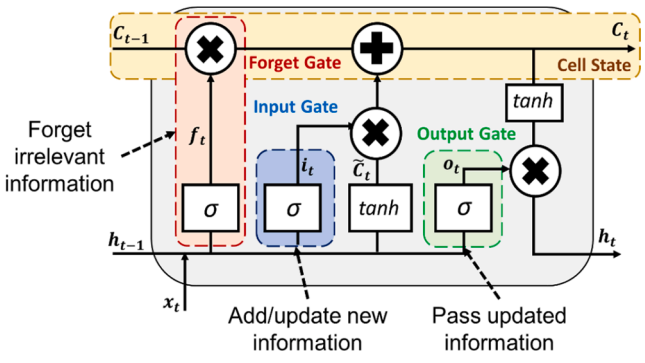

Fig. 3. Long short-term memory (LSTM) cell structure.

'gates' that regulate information flow. A typical LSTM structure is a cell state comprised of three gates is illustrated in Fig. 3 where σ and *tanh* represents sigmoid and hyperbolic- tangent functions, respectively; x_t, h_t, f_t, o_t, i_t, and C_t, represent current input, hidden state, forget vector, sigmoid output, sigmoid input, and cell state, respectively. An LSTM layer consists of recurrently connected blocks, known as memory blocks. Each one contains recurrently connected memory cells and three multiplicative units – the input, output, and forget gates. The forget gate determines what information should be discarded or retained from the previous steps in the sequence based on f_t. The input gate determines how the cell state (C_t) will be updated, and memory will be modified based on the input: x_t, and $h_{t\text{-}1}$. Output gate decides what the next hidden state (h_t) should be and which information will be produced by the state (C_t). For further information, the readers are referred to Fayaz *et al.* [12].

The architecture of the LSTM-based RNN trained in this study is shown in Fig. 4. The neural network structure is trained with cross-validation using randomly selected 80 % data out of the total dataset. The remaining 20 % of the dataset is used as the test set, which is not used in the training of the RNN. The RNN structure is trained using six seismic source and site parameters, including F, M_w, R_{rup}, R_{JB}, V_{s30}, and Z_{hyp}. The parameter F is inputted as a one-hot vector with [1,0] for the interface mechanism and [0,1] for the inslab mechanism. These features describe the physical mechanism of the earthquake source (*i.e.*, F, M_w, Z_{hyp}) and the site characteristics (*i.e.* R_{rup}, R_{JB}, V_{s30}). Although R_{rup} and R_{JB} are correlated with each other, especially for large distances (>50 km), the model proposed in this study performs significantly better when both features are used as inputs. Also, unlike traditional machine learning and statistical models, neural networks do not need explicit feature engineering as the neurons' weights are automatically optimized through gradient descent depending on the sensitivity of the target variables with respect to the input features. While previous studies [55] have considered more causal parameters, based on NGA-West2, this study considers the mentioned six parameters since they are the only ones available and formally published in the Chilean databases and literature. The RNN is trained to output the corresponding 35x1 **IM** vector in the logarithmic space. The training is conducted using IA [74] as the cost function due to its capability to lower the information loss for time series analysis. Since IA value of 1 indicates a perfect match and 0 indicates no agreement between the true and predicted **IM** vectors, the gradient descent is conducted using IA in a negative sense implying a value of −1 for perfect match.

The prediction power of the trained RNN is analyzed by comparing the measured and predicted values of the **IM** vector through the coefficient of determination, R^2. The R^2 for the 32 periods of $S_a(T)$ (also including *PGA*) and the other three IMs are shown in Fig. 5a and 5b, respectively. It can be observed from the figures that in most cases the R^2

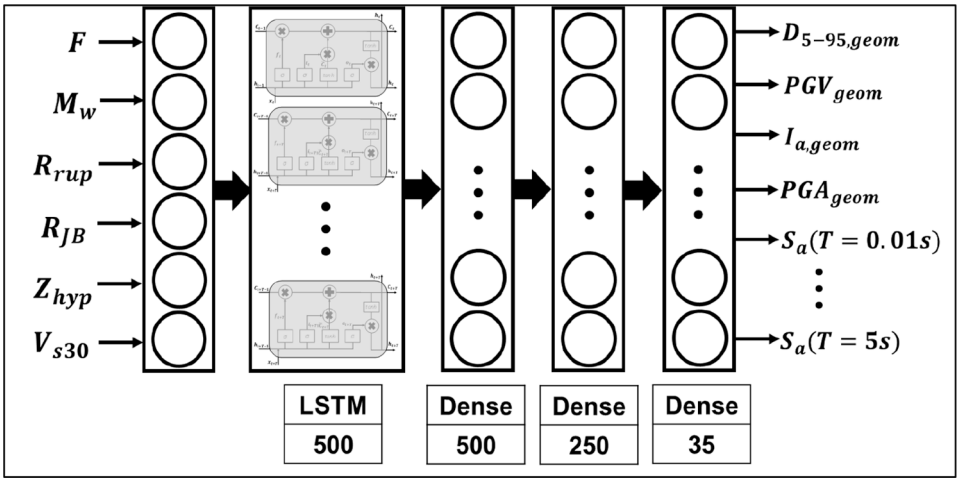

Fig. 4. Proposed RNN architecture.

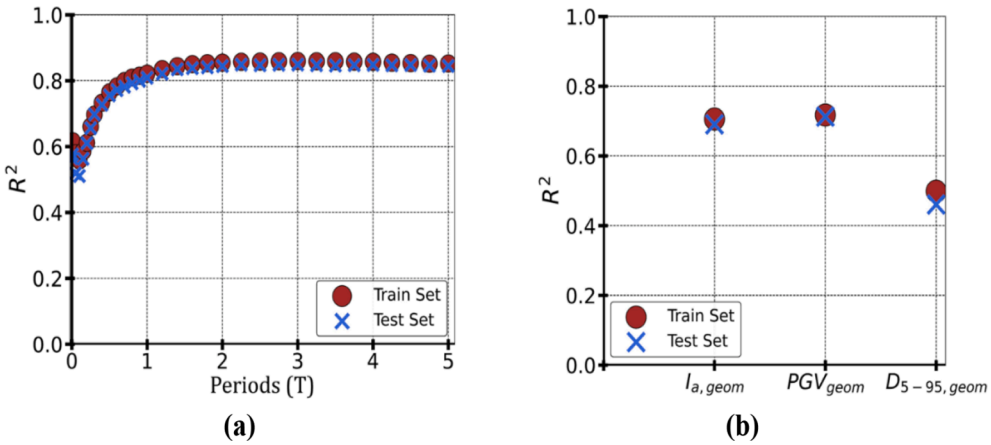

Fig. 5. R^2 for: a) $S_a(T)$; and b) $I_{a,geom}$, PGV_{geom}, and $D_{5-95,geom}$.

is above 0.7 except for $D_{5-95,geom}$ and $S_a(T)$ for very short periods of <0.5 s for both train and test sets. This can be attributed to the fact that the high magnitude subduction ground-motion spectra tend to be less smooth and have high variability for $S_a(T)$ of shorter periods, presumably, because the high-frequency zones in the response spectra of subductive earthquakes in Chile are greatly influenced by the number of asperities that participate in the rupture area, which vary throughout Chile and depend on the rupture areas, and therefore the event magnitudes. As an example, Montalva *et al.* [10] indicates a total standard deviation (in log-space) close to 0.86 for the short period domain (natural periods < 0.5 s) and 0.79 for natural periods between 0.5 and 2.5 s. This can also be observed from Fig. 2a, where for $S_a(T)$ at shorter periods, the median and $\pm\sigma$ have a sharp bent in the trends than the rest of the spectrum. Hence such similar bents are also in the prediction power of the RNN framework due to the high variability involved. Also, engineering structures with such short periods are not commonly designed with advanced engineering tools and may not require sophisticated record selections. The R^2 values are also observed to be very close for both the testing and training sets, showing that the trained RNN does not overfit the data and can model IMs for the ground motions that are not used to train the model (*i.e.,* the test set). It should be noted that the RNN framework is trained to have good prediction power for each IM and maintain the internal cross-dependencies within **IM**. Nevertheless, it will be shown in Fig. 8a (**Section 5**) that the proposed GGMM outperforms the other known GMMs used in Chile.

Furthermore, the prediction power and unbiased nature of the RNN is assessed by checking the normality of the residuals and comparing them against their corresponding M_w and R_{rup}. This is presented in Fig. 6 for four IMs. The plots also show bars demonstrating quantile 1 (Q1), median (50th percentile), and quantile 3 (Q2) of the residuals. In general, these plots show that the residuals tend to be normally distributed with zero medians across all M_w and R_{rup} values. It is further noticed that the residuals for $D_{5\text{-}95,geom}$ are larger than other IMs, which is consistent with their R^2 (*i.e.* larger variability of residuals lead to lower R^2). Finally, the residuals are tested for normality using the Shapiro-Wilk hypothesis test at a 5 % significance level, further validating the normality assumption with no evident bias.

While the internal cross-dependencies within the **IM** vector are not necessarily linear, for the sake of interpretability and brevity, the empirical Pearson correlation matrices of the measured **IM** vector and predicted **IM** vector are compared. The two correlation matrices are presented in Fig. 7 for 9 out of 35 IMs in the **IM** vector. The observations suggest that, in general, the two correlation matrices are visibly similar, and the RNN framework performs well in maintaining the internal correlations in the outputs of the **IM** vector. Furthermore, the correlation values differ mainly within the −0.15 to 0.15 range, which is not expected to alter the IM vector's sequence-dependent prediction significantly. However, it should be noted that an RNN framework does not explicitly intend to preserve the linear correlations within a sequence. Instead, it ensures that the internal dependencies are maintained, which may be highly nonlinear. Hence a mismatch in linear correlations does not necessarily indicate the lower performance of a

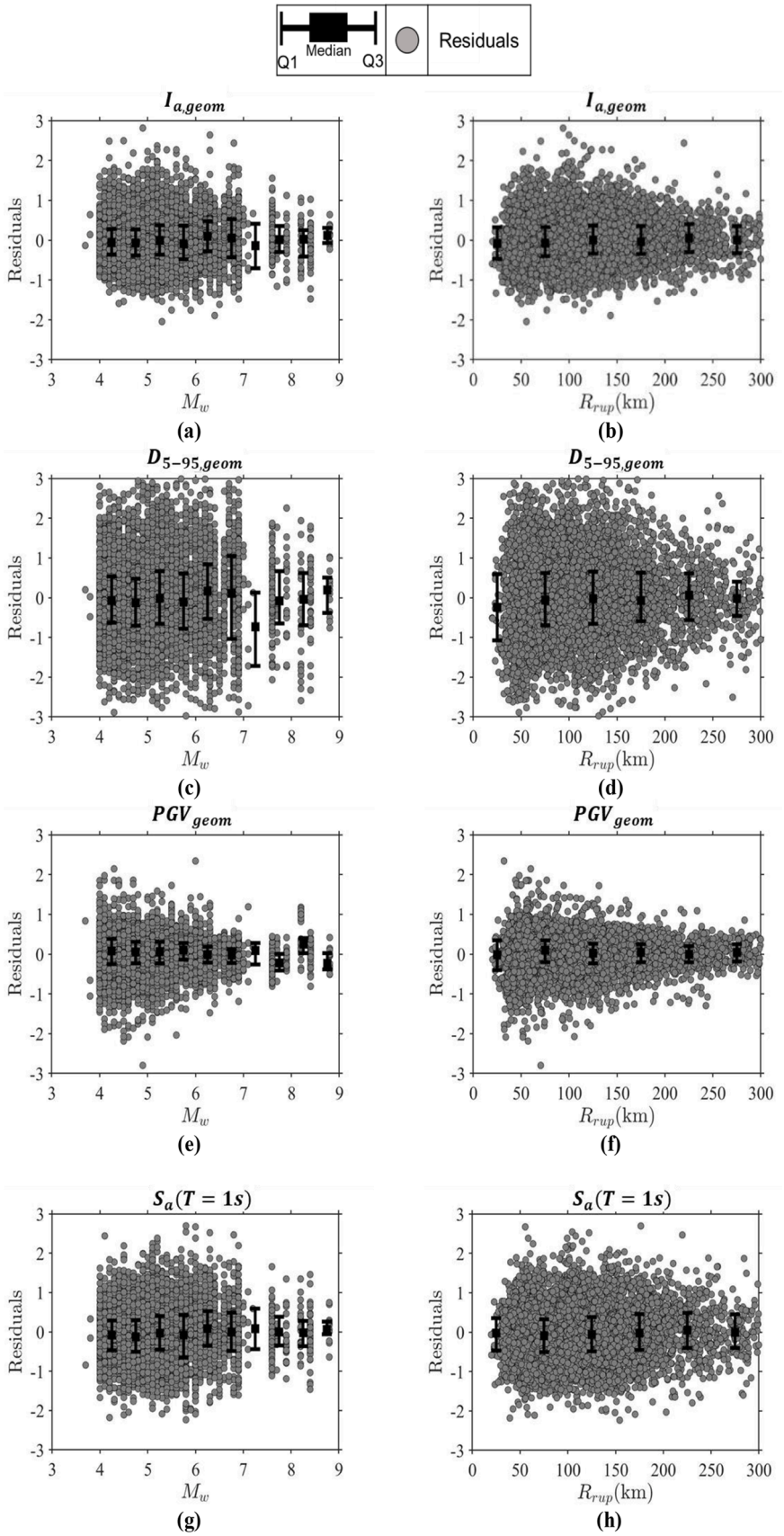

Fig. 6. RNN residuals for: a) $I_{a,geom}$ *vs* M_w; b) $I_{a,geom}$ *vs* R_{rup}; c) $D_{5-95,geom}$ *vs* M_w; d) $D_{5-95,geom}$ *vs* R_{rup}; e) CAV_{geom} *vs* M_w; f) CAV_{geom} *vs* R_{rup}; g)$S_a(T = 1s)$ *vs* M_w, h) $S_a(T = 1s)$ *vs.* R_{rup}.

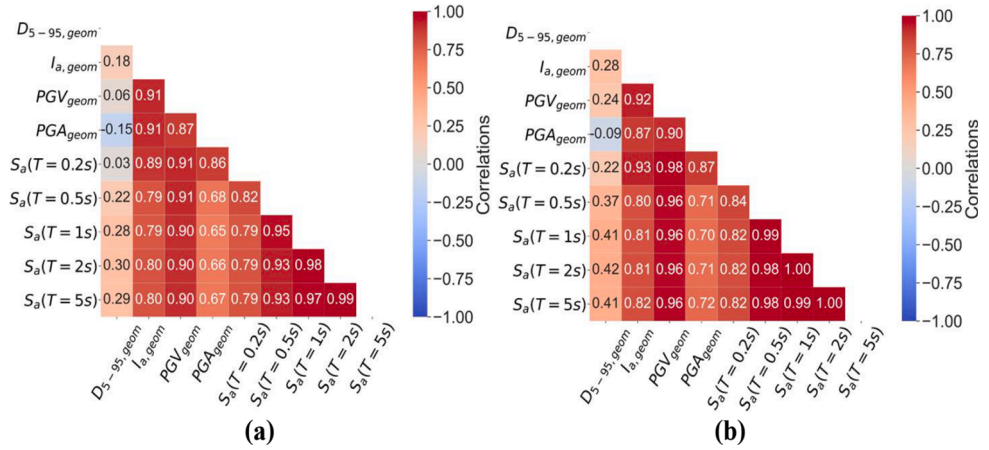

Fig. 7. IM cross-correlations a) measured; and b) predicted.

trained RNN. Nevertheless, the RNN maintains the correlations without any significant evident biases.

4. Covariance matrices for residuals

Due to the hierarchical structure of the ground motions arising from multiple recordings of the same event at different stations and recordings from various events at the same stations, the residuals between the true log-scaled $\mathbf{IM}_{ij}$ vector and RNN-predicted log-scaled $\widehat{\mathbf{IM}}_{ij}$ vector are used to compute 35 values of between-event and within-event variabilities for i^{th} event and j^{th} recording. This is done by fitting a mixed-effects regression [47] model to the residuals as given in Equation (1),

$$\mathbf{IM}_{ij} - \widehat{\mathbf{IM}}_{ij} = \boldsymbol{c}_0 + \eta_i(0, \mathrm{T}^2) + \varepsilon_{ij}(0, \Phi^2) \tag{1}$$

where η_i represents between-event variability with T^2 variance matrix for the 35 IMs (with ${\tau^2}_k$ for k^{th} IM in the **IM** vector), ε_{ij} represents within-event variability with Φ^2 variance matrix for the 35 IMs (with ${\phi^2}_k$ for k^{th} IM in the **IM** vector), and $\boldsymbol{c}_0$ represents any pending bias in the residuals for the 35 IMs. $\boldsymbol{c}_0$ was observed to be very close to zero (failing the regression hypothesis *t*-test at 5 % significance level [48]) and hence dropped in the overall analysis. Also, empirical Pearson correlations are computed for the residuals of **IM** vector, which are then used to convert the between-event and within-event variance matrices into their respective covariance matrices. The independent variances and the correlation structure (only for 31-period $S_a(T)$ spectrum) of the residuals are presented in Fig. 8a and 8b, respectively. It is assumed that the cross-IM correlation structure for ε_{ij} and η_i is the same. In summary, the overall RNN framework is developed for log-scaled mean predictions, and the residuals are used to construct between-event and within-event covariance matrices. The overall model (*i.e.,* mean and covariances) represents a generalized ground motion model (GGMM).

5. Comparison against other Chilean GMMs

In this section, the spectral prediction of the proposed mean GGMM (*i.e.,* RNN predictions) is compared against two popular GMMs utilized in Chile, i.e., (Montalva et al. [10]; Idini et al. [11]), denoted as MBR16 and IRRP17, respectively). However, both the GMMs were originally developed for the estimation of the geometric mean spectrum (*GeomeanS*$_a$). *RotD*50S_a has recently gained popularity in the earthquake engineering community mainly due to its unbiased nature and non-dependence on the orientation of sensors. Hence, to provide a fair comparison between the proposed GGMM and existing GMMs, it is essential to convert the *GeomeanS*$_a$ predictions into *RotD*50S_a without the addition of any variability. Fig. 9a presents the ratios of the *RotD*50S_a and *GeomeanS*$_a$ for all the ~7000 Chilean subduction ground motions for the 32 periods used in this study. It can be observed that while the mean ratios for all periods lie very close to 1, the ratios can go as high as ~3 and as low as ~0.75. Hence, *GeomeanS*$_a$-based GMM can not be directly used for *RotD*50S_a estimates. Several research studies have tried to develop conversion relations between the two types of estimates [49]. Most of these relations are not explicitly developed for Chilean subduction and, based on the authors' implementation, they did not perform very accurately and led to additional sources of variability. Therefore, the predictions of the GMMs (*i.e.*, MBR16 and IRRP17) are

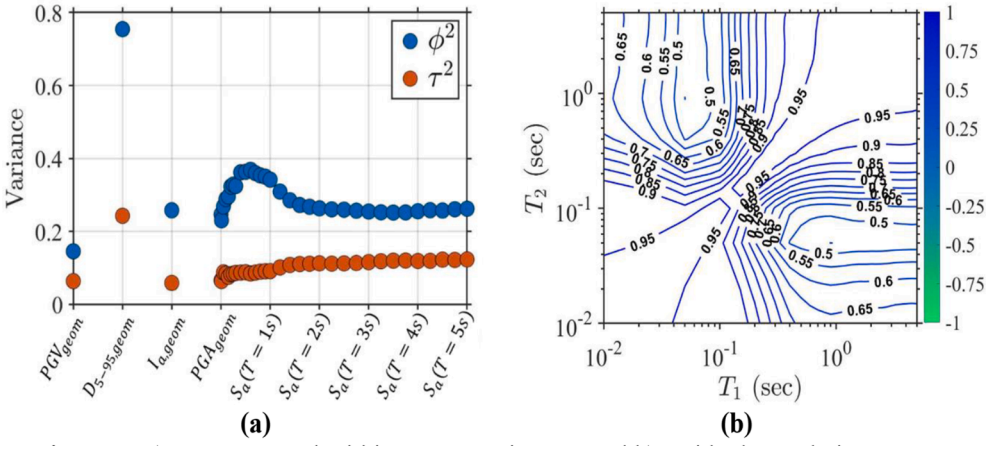

Fig. 8. a) Between- and within- event variances; and b) residual correlation structure.

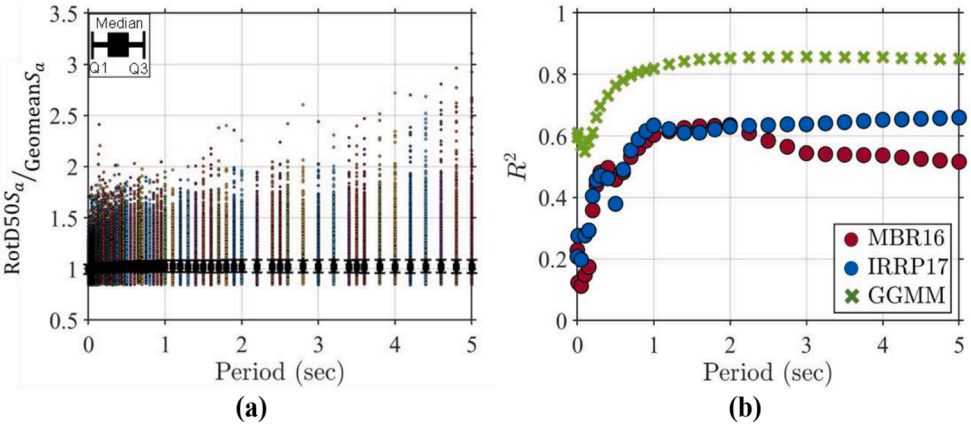

Fig. 9. a) Ratio of *RotD*50S_a and *GeomeanS*$_a$; and b) R^2 GGMM *vs*GMMs.

converted to *RotD*50S_a with no additional variability. This is done using Equation (2).

$$\mu_{\mathrm{GMM},RotD50S_a} = \mu_{\mathrm{GMM},GeomeanS_a} \times \mathrm{True}\,RotD50S_a / \mathrm{True}\,GeomeanS_a \quad (2)$$

First, for all ~7000 ground motions, the two GMMs are utilized to obtain their median estimates of *GeomeanS*$_a$ (this is denoted as $\mu_{\mathrm{GMM},GeomeanS_a}$). Then the obtained $\mu_{\mathrm{GMM},GeomeanS_a}$ estimates are multiplied by the ratio of the true recorded values of *RotD*50S_a and *GeomeanS*$_a$ (computed directly from a time-history analysis of SDOF) of their respective ground motions. This ratio is a scalar value (unique to each ground motion), and each ground motion corresponds to a single factor. Hence by computing the product in Equation (2), no variability is added to the estimates of the GMM. This multiplication process leads to the GMM estimates of *RotD*50S_a (denoted as $\mu_{\mathrm{GMM},RotD50S_a}$). This is equivalent to the GMMs developed for *RotD*50S_a having the same variance and errors that are associated with the GMM *GeomeanS*$_a$. Based on the history of GMMs (such as [6,7,44]), it is generally observed that the variances of the GMMs do not change significantly with the change of spectral acceleration measure. The comparison between the GMMs and proposed GGMM is only done for spectral acceleration and not for other IMs, as at the time of this study, there are no widely accepted GMMs in Chile for the other IMs.

Fig. 9b shows the comparison of the two GMMs and the GGMM estimates using the coefficient of determination R^2. It is observed that for all periods, the proposed GGMM leads to better goodness of fit and performs better than the other two GMMs. The two GMMs are very close in their prediction power, with IRRP17 having a slight edge for longer periods. It can be further observed that the proposed GGMM, on average, leads to 0.2 higher R^2 then the other two GMMs. Hence, it can be concluded that the proposed GGMM performs statistically better than both of the GMMs (*i.e.*, MBR16 and IRRP17) while maintaining the internal cross-dependencies between the spectral accelerations at various periods, which is not present in the GMM estimates.

5.1. Scenario-based comparison against conditional spectrum

To further assess the predictions of the proposed GGMM in comparison to other approaches, two recorded events are randomly selected from the test set (20 % of the dataset) as input scenarios for the GGMM. The outputs of the GGMM are compared against the predictions made using the conditional spectrum (CS) [50] for the same scenarios. The two selected target scenarios involved an event with low magnitude $M_w =$ 5.0 and$R_{rup} =$ 74.1 km and a high-magnitude event $M_w =$ 8.4 and$R_{rup} =$ 62.0 km. The CS was computed based on IRRP17 GMM and the spectral correlation structure proposed by Candia et al. [15] for Chilean subduction ground motions. The IRRP17 GMM outputs are converted to the *RotD*50S_a as described in the previous section using the Equation (2). Both GGMM and CS are used to construct the mean and ± one standard deviation of log($S_a(T)$) for the 32 periods. It should be noted that the proposed GGMM also provides additional IMs apart from log($S_a(T)$). However, such IMs cannot be estimated using CS, so comparison is primarily made using log($S_a(T)$) for 32 periods.

The proposed covariance matrices of the GGMM residuals and the correlation structure of the epsilons (ε) CS are used to construct the ± one standard deviation bands of predictions. This can be done using Equation (3) as exercised by Baker [25]. In Equation (3), ρ represents the correlation between the residuals of k^{th} IM and IM*, σ_{IM_k} represents the standard deviation of IM_k and $\sigma_{IM_k|IM^*}$ represents the standard deviation of the k^{th} IM conditioned on the standard deviation of IM*.

$$\sigma_{IM_k|IM^*} = \sigma_{IM_i}\sqrt{1-\rho^2(IM_k, IM^*)} \quad (3)$$

In the case of CS, Equation (3) is implemented with the concept of ε value of the k^{th} IM rather than IM residuals. For both GGMM and CS, the spectra (and other IMs in the case of GGMM) are constructed using a conditioning period (T) of 1 *sec*. It should be noted that the GGMM framework does not require the conditioning of T to obtain the mean prediction; however, the estimation uncertainty bands (± one standard deviation) for the 35 IMs are conditioned on the variance of the selected IM*. Since the construction of CS also requires an appropriate ε, the true ε of the records are computed directly by comparing the IRRP17 GMM mean estimate and recorded value of log($S_a(T = 1s)$) for the ground motion records. This means that the CS and the true recording match perfectly for log($S_a(T = 1s)$), while the GGMM is not explicitly made to match the recorded log($S_a(T = 1s)$). However, it was observed that the GGMM without any alteration matches the target recorded spectrum at multiple periods.

Fig. 10a and 10c presents the spectrum of the recorded ground motions compared against the estimates of the GGMM and CS (*i.e.*, CMS ±σ) conditioned on $S_a(T = 1s)$ for the two events. In the case of the GGMM, the estimates of the other three IMs are also compared with the recorded values in Fig. 10b and 10d. It can be deduced from Fig. 10a and 10c that the GGMM performs better in describing the short and long period spectral shapes of the selected scenarios as compared to the CS. Moreover, it is observed that the GGMM leads to significantly smaller uncertainty (*i.e.*, standard deviation, which represents 16th and 84th percentiles) in the predictions and still captures the true recorded spectra within the uncertainty prediction bands better than the CS (which has a larger uncertainty band). The efficiency of the proposed model in developing a target spectrum with lower variability can be beneficial in performance-based analysis and engineering applications. The uncertainty reductions in the **IM** estimates can significantly save computational time, enabling practitioners to confidently obtain statistical estimates of structural responses/seismic demands with a reduced number of required ground motions and analyses. It can be noticed that

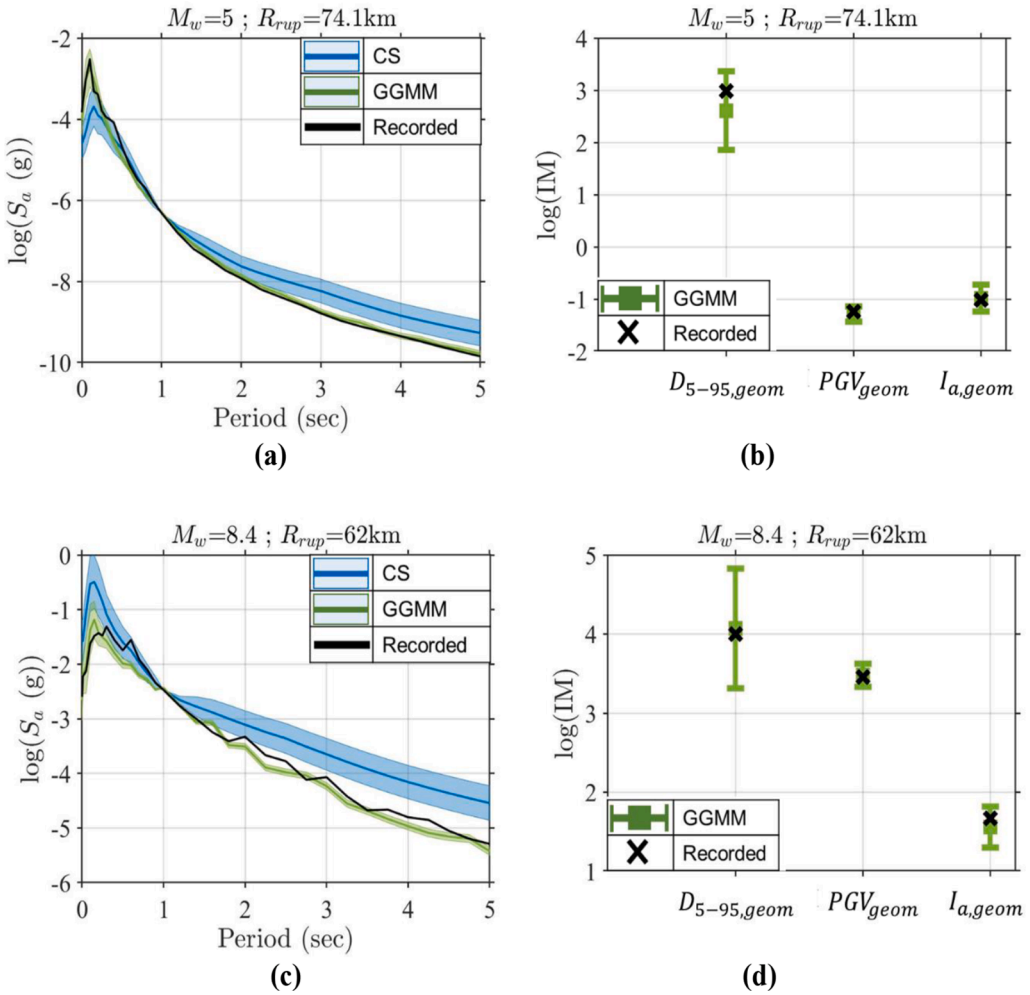

Fig. 10. GGMM *vs* CS for example events: (a) S_a spectra ($M_w = 5$ and $R_{rup} = 74.1$ km); b) other IMs ($M_w = 5$ and $R_{rup} = 74.1$ km); and (c) S_a spectra ($M_w = 8.4$ and $R_{rup} = 62$ km); d) other IMs ($M_w = 8.4$ and $R_{rup} = 62$ km).

the GGMM spectrum (especially for the $M_w = 8.4$ and $R_{rup} = 62.0$ km event) also tries to replicate the spikes (with some bias) in the spectrum observed in the true recordings. This postulates that, with enough data and proper processing techniques, future work on the GGMM framework can be very useful in understanding the peculiarities of the ground motions spectrum and their shapes for a variety of tectonic settings and regions. Furthermore, it is observed that for both scenarios, the other three IMs $D_{5-95_{geom}}$, Ia_{geom}, and PGV_{geom} estimated by the confidence interval of the GGMM encloses the true recorded values. The recorded values are observed to lie very close to the mean log predictions of the GGMM. This suggests that the proposed GGMM framework can accurately capture the values of the 35 IMs and their cross-dependencies.

5.2. Likelihood-based comparison against conditional spectrum

To statistically compare the proposed GGMM (*i.e.*, RNN framework and covariance matrices) against the CS for the predictions of $RotD50S_a$, the GGMM and CS are used to construct the median and ± one standard deviation of the $\log(S_a(T))$ for 32-periods for all ~7000 ground motion records. This is done using the records' seismic source and site parameters as inputs to the GGMM and CS. Similar to the previous section, the spectra are constructed for a conditioning period (T) of 1 *sec* (IM*=$S_a(T = 1s)$). As mentioned in the previous section, since the construction of CS requires an appropriate ε, the true ε of the records were directly computed by comparing the IRRP17 GMM mean estimate and recorded values of $\log(S_a(T = 1s))$ for all the ~7000 ground motion records. This means that the CS and the true recording match perfectly for $\log(S_a(T = 1s))$ and hence the intensity level is appropriately captured by the CS. The GGMM does not require any postprocessing and is not explicitly made to match the recorded $\log(S_a(T = 1s))$. Hence, plots like Fig. 10 are developed for all records in the database. It should be noted here that even though the GGMM also outputs other three IMs in its **IM** vector, only the 32 values of $S_a(T)$ for periods 0 to 5 s are compared against the identical period-range spectra from CS for this section.

The $S_a(T)$ predictions of the GGMM and CS are compared by computing the log-likelihood of the recorded spectrum at each recording station to fall within the uncertainty bands of estimation of the GGMM and CS, using Equation (4). In Equation (4), the log-likelihood for each of the 32 periods is calculated (term inside the brackets) and then summed across the entire spectrum to achieve one estimate per recording. Hence, a higher log-likelihood represents a better predictive power. There are two reasons for using Equation (4) in this form without consideration of covariance instead of variance: i) LSTM-based RNN model does not give correlation values but rather represents the nonlinear dependencies through activation functions and gate mechanisms; hence it would not be possible to use a correlated version of the loglikelihood function in Equation (4); and ii) uncertainty bands for both CS and GGMM are already correlated using Equation (3), hence using correlation in Equation (4) would cause double counting of the correlation values in the likelihood.

$$lnL = \sum_{i=1}^{n=32}\left[-\frac{ln(2\pi)}{2} - \frac{ln(\sigma_i{}^2)}{2} - \frac{1}{2\sigma_i{}^2}(x_i - \mu_i)^2\right] \quad (4)$$

The results of this analysis are presented in Fig. 11a and 11b for all the ground motions against the magnitude and rupture distances of their respective events. For all the cases, it is observed that the estimation of GGMM has a relatively stable log-likelihood indicating no observable bias in the predictions concerning the M_w or R_{rup} of the events. For both scenarios, GGMM performs better than CS, leading to a higher log-likelihood. It is also observed that CS tends to be biased concerning M_w with an increasing trend. In contrast, the GGMM tends to have no observable trends concerning both M_w or R_{rup}. Hence, the brief comparison presented here seems to suggest that the proposed GGMM can represent an improved method for estimating the ground-motion spectra of an earthquake without large variability in predictive power, particularly with respect to the distance to the fault rupture or magnitude of the seismic event.

6. Utilization of the GGMM for structural performance analysis

6.1. Building design and nonlinear model description

To assess the robustness of the GGMM through evaluation of structural response, a twenty-story SSMF building ($T = 2.1$s, fundamental period) is developed in the OpenSees version 2.5.0 simulation platform. The building is designed as per the Chilean seismic code [51] and ANSI/AISC 341–16 [52], considering a seismic zone 3 (*i.e.*, the highest seismic design level), and a soil type C (*i.e.*, very dense soil, with V_{s30} between 350 m/s and 500 m/s). Fig. 12 shows the plan view and elevation of the building. Seismic demands for the designs are calculated by performing response spectrum analysis (RSA). The design process is carried out under the following assumptions:

- Occupancy: office buildings.
- Location: Andalue, Concepcion, Chile.
- Typical Story height: 3.5 m at all stories.
- There is no contribution of the gravity system to the lateral force-resisting system.
- All beam-column connections are detailed as reduced beam section (RBS) connections designed according to ANSI/AISC 358-16 [53] using $a = 0.625\,b_f$, $b = 0.75\,d_b$ and $c = 0.250\,b_f$.
- Embedded base-column connections were considered. Soil-structure interaction was neglected in this study.

SSMFs are typically analyzed considering 2D nonlinear models. Several researchers in the past have used 2D models for nonlinear time history simulations with concentrated plasticity models (e.g., [55,56,58]). P-delta effects are accounted for by including the geometric stiffness in the element stiffness formulation and adding leaning columns due to the gravity loads acting on the gravity frames. All gravity loads (i.e., those acting on the SSMFs and those acting on the leaning column) are consistent with a seismic weight equal to $1.00D + 0.25L$, where D and L are the dead and live design loads, respectively. The structural models represent buildings that comply with current seismic design codes. Thus, inelastic deformations are assumed to be due to flexural plastic hinges at the ends of beams (specifically at RBS locations), flexural or axial-flexural plastic hinges at the ends of columns, and shear deformations at panel zones. Therefore, beams and columns are modelled with elastic elements, and inelastic deformations develop only at the zero-length nonlinear spring elements, which represent the expected plastic hinges locations.

Beams are modelled with three elastic elements and two inelastic rotational springs located at their ends (at the RBS location). The properties of the backbone moment-rotation curve are based on the modelling parameters of Applied Technology Council [59] and Lignos and Krawinkler [60], updated and adjusted after Hartloper [61]. In contrast, columns are modelled with one linear-elastic element and two inelastic rotational springs at their ends. The properties of the backbone curve are based on the recent modelling parameters proposed by Hartloper and Lignos [62]. Consistent with Suzuki and Lignos's [63] findings on axial/flexural interaction modelling, the backbone moment-rotation curve is computed considering an axial force equal to that due only to gravity load. Cyclic deterioration rules of rotational springs in columns and beams follow the originally proposed bilinear hysteretic model [64] and later on modified by Lignos and Krawinkler [60]. Parameter c of the energy-based rule defined by Rahnama and Krawinkler [65] is set equal to unity, which gives an almost constant deterioration rate, and the recommendations of Hartloper and Lignos [62] to model cyclic deterioration are also included.

Shear panel zone behaviour is modelled considering the parallelogram arrange and the moment-rotation relationship of the rotational spring as presented in Gupta and Krawinkler [66], except that the 3 % strain hardening is replaced by a 1 % strain hardening, as recommended in more recent studies (e.g., [63]). On the other hand, column base connections are modelled as linear-elastic elements computing their rotational stiffnesses. These stiffnesses for exposed and deeply embedded base column connections are defined following Kanvinde *et al.* [67] and Torres-Rodas *et al.* [68], respectively. Following the same approach adopted to model plastic hinges in columns, the rotational stiffness of the base plate connections is calculated considering axial forces due only to gravity load. Modelling base connections as linear elastic elements is consistent with the current practice since practitioner engineers prefer to size them using the capacity design criteria (e.g., [69]). Finally, damping is modelled adopting Rayleigh damping

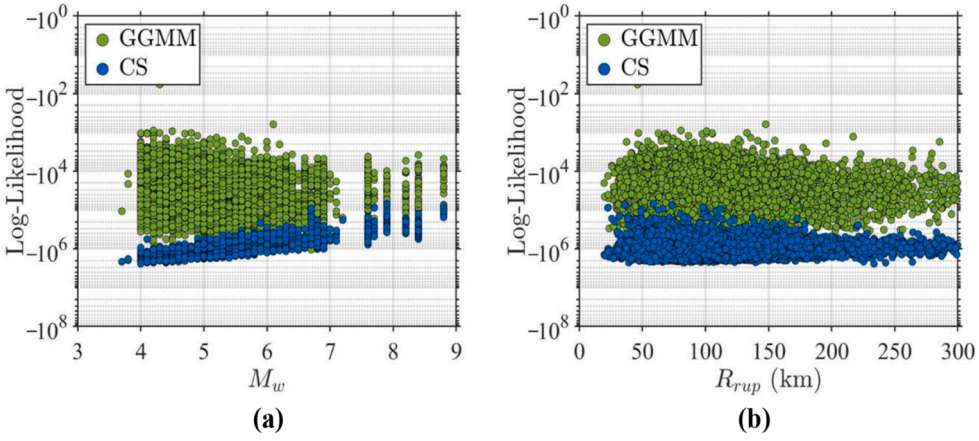

Fig. 11. Likelihood comparison *vs*: a) M_w; and b) R_{rup}.

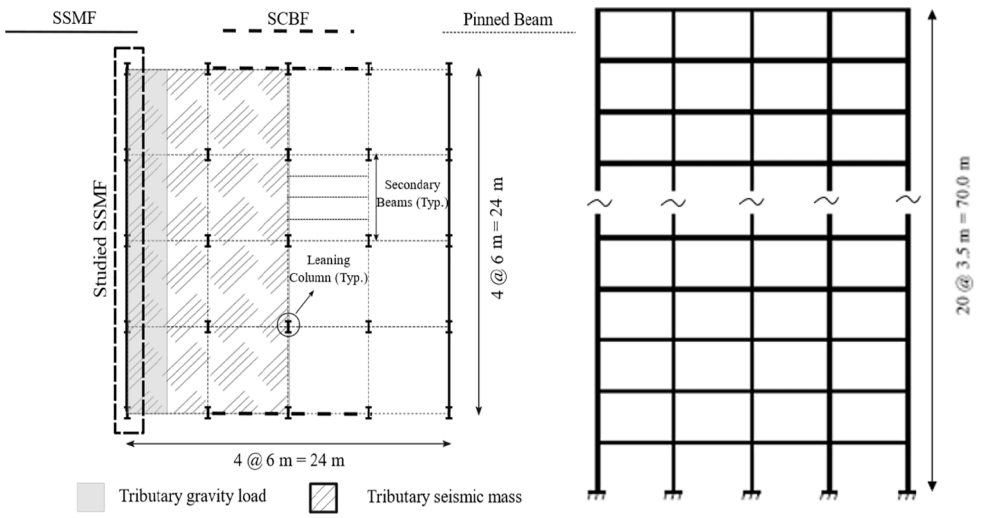

Fig. 12. Building plan and elevation view.

assumptions, and it is assigned only to elastic elements. The model proposed in the Pacific Earthquake Engineering Research Center (PEER) Tall Buildings Initiative (TBI) Guidelines [71] for maximum considered event (MCE) level analysis is considered for defining a fundamental damping ratio of 2.5 %.

6.2. Hazard consistent ground motion selection

As mentioned in the previous subsection, the building is located in Andalue, Concepcion, Chile. Using the hazard assessment platform developed by Candia et al. [15], three different hazard levels of 475, 970, and 2475 years mean return period are considered for structural analysis. The MBR17 GMM and the recently updated recurrence model for Chilean subduction [15] are used for the hazard assessment and hazard disaggregation for the three return periods and two seconds. The disaggregated parameters of the site are provided in Table 1 for the three considered mean return periods.

Table 1 shows that the disaggregation of the hazard does not necessarily lead to an increase in contributing magnitude with an increase in the mean return period (hazard level). As PSHA is based on GMMs such as IRRP17 and MBR16, increasing hazard levels are further characterized by increasing values of ε, representing the number of standard deviations above or below the median IM obtained from the GMM [72]. Therefore, to estimate disaggregated IMs, one needs to use the M_w, R_{rup} distribution as input to a GMM to obtain the log-scale mean prediction (μ) and then add or subtract (depending on positive or negative ε) ε times the standard deviation (σ) of the GMM. Hence, simply using the disaggregated $M_w - R_{rup}$ distribution in the proposed GGMM to obtain IM estimates will be inaccurate. Also, the disaggregated ε is based on current GMMs whose standard deviations (σ) are typically larger than the proposed GGMM; hence ε should not be used directly for the proposed GGMM. For example, for a period of 2 *sec* (~building's fundamental period), MBR16 GMMs have a σ of 0.65 and 0.76 for intraslab and interface mechanisms, respectively, while the proposed GGMM has a σ of 0.61. To develop consistent **IM** using the proposed GGMM, the disaggregated event parameters are first used as input to the CMS-based on the MBR16 GMM. Then the event parameters are used to obtain the median **IM** vector from the GGMM, which is scaled to match the CMS estimate of $S_a(T = 2.1\text{s})$. This is shown in Fig. 13 for all three hazard levels. The ratio between $S_a(T = 2.1\text{s})$ obtained from uniform hazard spectrum (UHS, thereby CMS) and proposed GGMM is denoted as α. To appropriately represent the hazard, the median **IM** vector is scaled by α and utilized as the target for ground motion selection. This ensures that the CMS, UHS, and GGMM times α have the same $S_a(T = 2.1\text{s})$ for the target.

Table 1
Concepción deaggregation event parameters.

Mean Return Period (years)	475	970	2475
M_w	8.02	8.12	8.26
R_{rup}	49.6	43.5	37.7
ε	1.61	1.79	2.04

It should be noted here that the GGMM and CS are conceptually different in developing target spectra. While the GGMM directly (and explicitly) captures the cross-IM dependencies based on the causal source and site parameters, the CMS additionally uses ε to develop the correlation structure. This means that, for a given source and site parameters, the shape of the median GGMM does not vary, while the CMS may lead to differences in spectral shape due to different values of ε. While this is generally true, differences are not particularly significant for typical engineering hazard levels (including 2 %, 5 %, and 10 % probability of exceedance in 50 years; based on the authors' experience) where ε values vary by 1 (and mainly fall between + 1 to + 2) for highly seismic prone zones like Chile and Ecuador. To demonstrate this, an example scenario ($M_w = 8.12$, $R_{rup} = 43.5$ km, $V_{s30} = 360$ m/s) corresponding to 975-year return period in Andalue, Concepcion, Chile is used to construct the CMSs for six values of ε. This is presented in Fig. 14. The figure shows that for the given high-magnitude scenario, ε values do not severely impact the shape of the spectra, and the differences are only observed in their amplitudes. Hence, the use of the α factor to align median-GGMM-based spectra and CMS-based spectra should represent an acceptable method of comparing the ground-motion selection procedure. However, for a more accurate and robust comparison, the hazard analysis utilized for deaggregation should be consistent for the two methods. In this case, the deaggregation is primarily based on GMMs, thereby is more consistent with the considered CMS. Also, a CMS based on the GGMM could be developed and then compared against the conventional CMS approach. However, this is outside the scope of this study, and such scrutiny is not expected to affect the main findings from this study significantly.

GGMM-based ground motions selection is conducted by limiting the scaling such that only scale factors of 0.5 to 2.0 are used. It should be noted here that the scaling process alters all the elements of **IM** vector (*i.e.*, PGV_{geom}, PGA_{geom} and $S_a(T)$ are scaled by α and $I_{a,geom}$ is scaled by α^2) of the ground motions except for the duration $D_{5-95,geom}$. Hence, to select the ground motions that match the **IM** vectors, firstly, the ground motions that have $D_{5-95,geom}$ between 80 % and 120 % of the target $D_{5-95,geom}$ for the given hazard level are used as the ground-motion

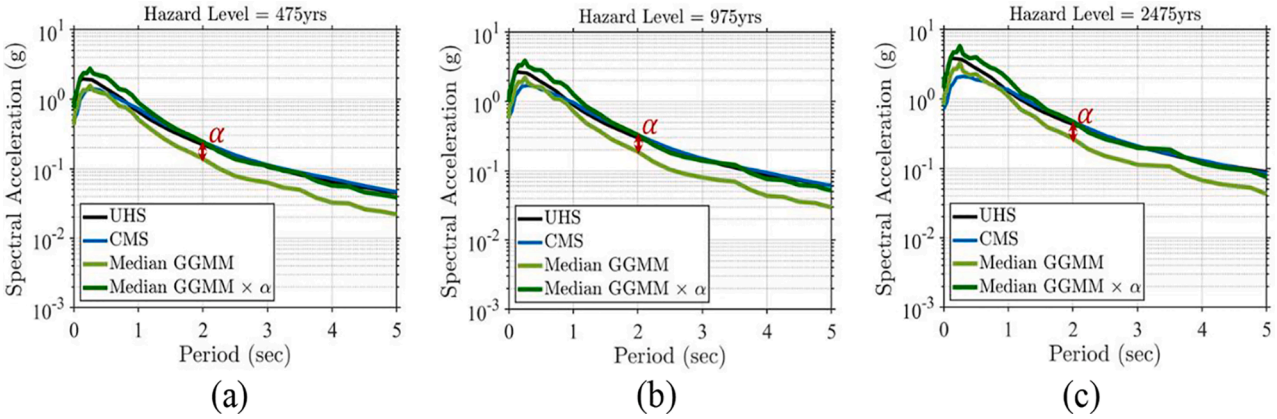

Fig. 13. Scaling of median GGMM to match CMS at $S_a(T = 2.1s)$ for hazard levels: a) 475 years; b) 975 years; and c) 2475 years.

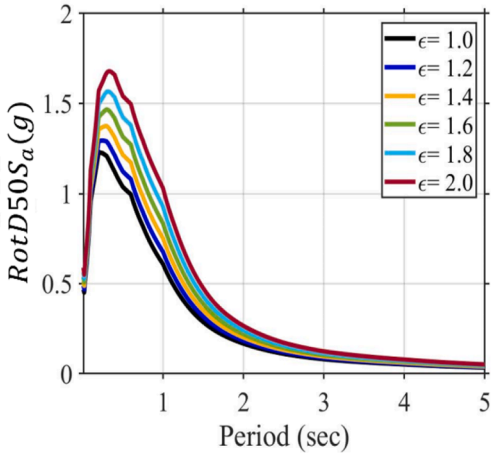

Fig. 14. CMSs for various ε for $M_w = 8.12$, $R_{rup} = 43.5$ km, $V_{s30} = 360$ m/s.

selection pool. Subsequently, the pooled ground motions are scaled from 0.5 to 2.0 at an interval of 0.1 (hence 16 scaling factors), and the resultant 34 × 16 IM vectors (without $D_{5-95,geom}$) are obtained and compared against the target **IM** vector of the given scenario obtained by appropriately scaling the mean GGMM using α. Due to the different scales of the IM values, the comparisons are performed using a weighted average accuracy, where accuracies for $I_{a,geom}$ and PGV_{geom} are computed using the normalized accuracy given in Equation (5), and the accuracy of the $S_a(T)$ spectrum is computed using the IA [74] in Equation (6). Both equations (5) and (6) are bounded between 0 and 1, where 0 is the worst match, and 1 is the best match. Then the accuracy measures are combined using a weighted average accuracy (Acc_{avg}), computed as per Equation (7).

$$Acc_{IM} = 1 - \text{abs}\left(\frac{IM_{Tar} - IM_{GM}}{IM_{Tar} + IM_{GM}}\right) \tag{5}$$

$$IA_{S_a} = 1 - \frac{\sum_{i=1}^{n}\left[S_{a,Tar}(T_i) - S_{a,GM}(T_i)\right]^2}{\sum_{i=1}^{n}\left[\left|S_{a,Tar}(T_i) - \frac{\sum_{i=1}^{n} S_{a,Tar}(T_i)}{n}\right| + \left|S_{a,GM}(T_i) - \frac{\sum_{i=1}^{n} S_{a,Tar}(T_i)}{n}\right|\right]^2} \tag{6}$$

$$Acc_{avg} = \frac{w_1 Acc_{I_{a,geom}} + w_2 Acc_{PGV_{geom}} + w_3 IA_{S_a}}{w_1 + w_2 + w_3} \tag{7}$$

For illustrative purposes, in this case, the weights are chosen as $w_1 = 0.2$; $w_2 = 0.2$; and $w_3 = 0.6$; however, users can freely choose weights based on their own criteria and/or structural system sensitivity. Also, in this study, only $S_a(T)$ values within the period range of 0.5 T_1 (i.e., 1.05 s) and 2 T_1(i.e., 4.2 s) [75] are used in Equation (6) for the record selection. For each hazard level, the ground motions and scaling factor combination leading to the top 30 Acc_{avg} among all ground motions are selected for the building structural assessment. The set of these ground motions is denoted as $GGMM_{IM}$. In addition, the selected ground motions only contain one scaled version. The spectra of the selected ground motions for a return period of 475 years are shown in Fig. 15a. Similarly, 30 ground motions are selected using the developed CMSs as the target spectra and Equation (6) as the objective function. The resulting spectra of the selected ground motions are shown in Fig. 15b for a return period of 475 years. Furthermore, as CMS does not consider other IMs except $S_a(T)$, a separate selection of 30 ground motions is also conducted using only $S_a(T)$ with GGMM as the target spectrum. This is done by using $w_1 = 0$; $w_2 = 0$; and $w_3 = 1$ in Equation (7). Using this criterion, the selected 30 recorded ground motions are shown in Fig. 15c for a return period of 475 years. The set of ground motions selected using this criterion is denoted as $GGMM_{Sa}$. In addition, Fig. 15d illustrates the minimum, median and maximum of the other three IMs of the selected ground motions using the three above-mentioned criteria and the respective target median GGMM × α hazard IMs for the return period of 475 years. This process of ground motion selection is repeated for the other three return periods. It can be observed from Fig. 15 that the selected ground motions lie in a very close range with respect to the target hazard spectrum for Fig. 15b and 15c as compared to Fig. 15a. This is expected as the ground motions are selected based on uni-criteria of $S_a(T)$ match for the former cases while the latter involves other IMs in the selection process, which cause variability in the $S_a(T)$ values. This is further evident in Fig. 15d, where the ground motions selected using GGMM and all IMs (denoted as $GGMM_{IM}$) have the hazard IM very close to the median IM of the selected ground motions. Other cases, where only $S_a(T)$-based selection criteria is used, i.e., CMS and $GGMM_{Sa}$, tend to have a similar level of variabilities in the selected IMs.

6.3. Analysis of building responses

This section discusses the results of the NLTHA and statistically compares the three methods (*i.e.*, CMS, $GGMM_{IM}$, and $GGMM_{Sa}$) of ground-motion selection. The two components of the 30 scaled and selected ground motions are used to conduct NLTHA of the 20-story frame. Thus, for each hazard level, 60 NLTHA are conducted, and the peak interstory drift ratio (PIDR) are recorded. PIDRs are obtained for the three hazard levels and the three methods of ground motion selection. The cases leading to PIDRs > 0.1 are classified as collapse [76]. Fig. 16 shows the histograms with kernel density estimation (KDEs) functions of the non-collapse cases of PIDRs obtained from the three ground motion selection methods (CMS depicted in blue, $GGMM_{IM}$ depicted in green, and $GGMM_{Sa}$ depicted in yellow) for 2475-year return

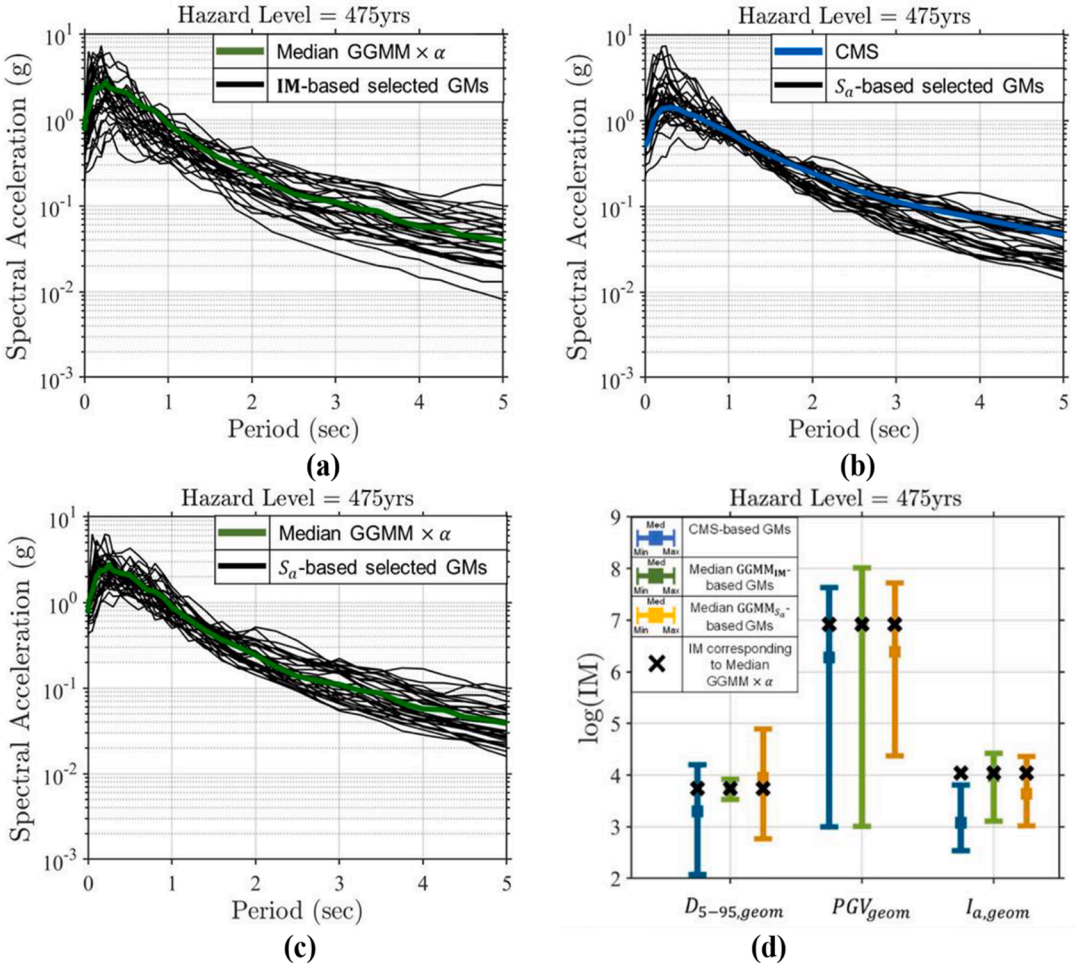

Fig. 15. GMs selected using: a) median GGMM (all IMs considered); b) CMS; c) median GGMM (only S_a considered); and d) comparison of other IMs.

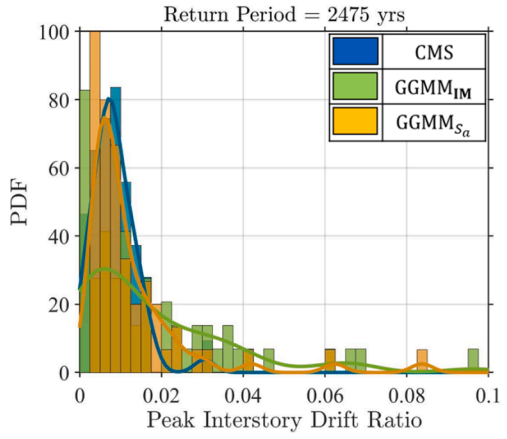

Fig. 16. Peak inter-story drift ratios using the three methods of ground motion selection.

period. The KDEs of the CMS- and GGMM$_{Sa}$- based PIDRs are observed to be quite similar. This implies that the ground motions selected using the GGMM considering only $S_a(T)$ lead to very similar PIDRs as compared to the CMS-based ground motions. In contrast, when the ground motions are selected using the GGMM considering complete **IM** vector, the KDE of the associated PIDRs diverges significantly from the CMS. Furthermore, it is specifically observed that the variability of GGMM$_{IM}$ is higher (*i.e.*, wider KDE), especially for higher levels of PIDRs.

The non-collapse PIDRs are statistically compared using the Kolmogorov–Smirnov (KS) hypothesis test [77]. The hypothesis test is conducted at a 5 % significance level with a null hypothesis that the cumulative distributive functions (CDFs) of the PIDRs from two types of record sequence selection are similar. The similarity is measured using the supremum distance measure between the two CDFs. Hence, *p*-values>0.05 mean that the two CDFs are statistically similar to each other. The comparison is made between the non-collapse PIDRs of ground motions selected through CMS and GGMM$_{IM}$ and CMS and GGMM$_{Sa}$. This comparison is conducted independently for each hazard level (*i.e.,* mean return period of 475, 975, and 2475 years). Fig. 17a presents the results of the KS tests showing that the comparison between CMS and GGMM$_{IM}$ leads to *p*-values<0.05 indicating that the null hypothesis is rejected. Thereby the two distributions cannot be considered statistically similar.

On the other hand, the comparison between CMS and GGMM$_{Sa}$ leads to *p*-values>0.05, denoting their statistical similarity for all hazard levels. Fig. 17b further illustrates the number of collapses obtained in the NLTHA using the three-ground motion selection criterion. In general, the ground motions selected by the mean **IM** vector from the GGMM entail more collapses than the other two criteria for each seismic hazard level. For a return period of 475 years, the ground motions selected by the GGMM$_{Sa}$ did not cause any collapse, while only one of the ground motions selected using the CMS led to structural collapse. For the other two seismic hazard levels (*i.e.*, 975 and 2475 years), the number of collapses from GGMM$_{Sa}$-based selection is relatively higher than the CMS-based selection. Though the median GGMM × α and CMS match at $S_a(T = 2.1s)$, different shapes of the two target spectra for

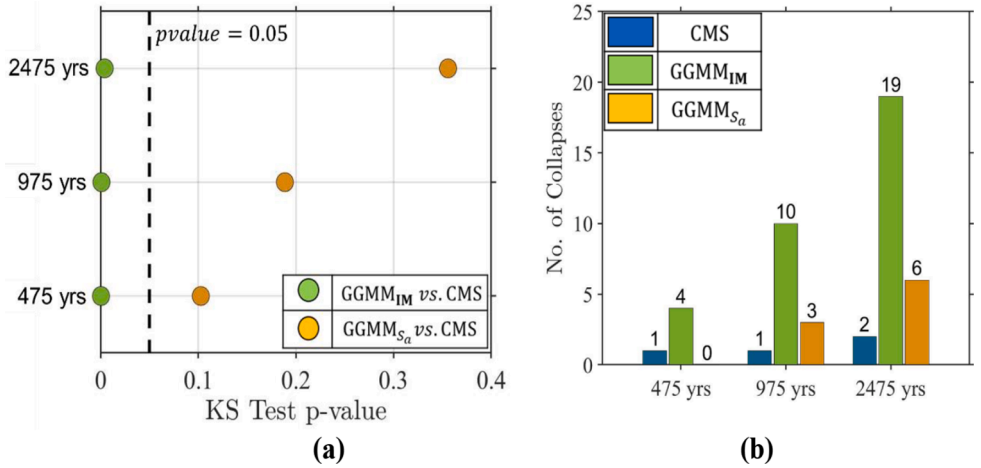

Fig. 17. a) KS test comparison of non-collapse cases; and b) number of collapses for CMs *vs* $GGMM_{IM}$ *vs* $GGMM_{Sa}$.

other $S_a(T)$ lead to the selection of ground motions with different characteristics, thereby leading to variations in the causal PIDRs. This also indicates that a $S_a(T = T_1)$-based ground-motion selection is not sufficient in explaining PIDRs of the buildings. Moreover, results from the analyses of 20-story SSMF indicate that the utilization of $S_a(T)$-based objective function for ground motion selection and NLTHA can lead to underestimating the structural demands (e.g., collapse evaluation), and perhaps objective functions based on a vector of IMs can improve the methods of seismic demand assessment. The results presented herein, particularly for the 975 and 2475 return periods, are similar to those reported in a previous study by Medalla et al. [78]. In that study, it was observed that the median collapse capacity of SSMF buildings with a fundamental period>1.0 s is affected by the variability in other IMs (in particular D_{5-95}) when subjected to high magnitude subduction ground motions (mainly due to the large number of inelastic cycles). In fact, previous studies on the topic (e.g., [79,80]) indicate that collapse capacity is affected (*i.e.*, reduced) by the strong ground motion duration because of the damage accumulated due to the higher number of loading cycles.

7. Conclusions

This paper developed and implemented a recurrent neural network (RNN) based generalized ground-motion model (GGMM) that serves as a sophisticated and improved alternative to the existing ground-motion models (GMMs) for Chile and a novel approach for subduction environments. The proposed GGMM was developed using a database of ~7000 ground motions recorded from ~1700 events selected from the SIBER-RISK strong motion database with magnitudes ranging from 3.8 to 8.8. The GGMM was trained to output a 35x1 cross-dependent vector of intensity measures (**IM**), including $D_{5-95,geom}$, PGV_{geom}, $I_{a,geom}$, and 32-period $S_a(T)$ spectrum. The inputs to the GGMM include six seismic source and site parameters, i.e., the seismic mechanism (F), moment magnitude (M_w), closest rupture distance (R_{rup}), Joyne-Boore distance (R_{JB}), hypocentral depth (Z_{hyp}), and average soil shear-wave velocity at top 30 m (V_{s30}). The parameter F is used as a one-hot vectors for the Interface and Inslab mechanisms. The use of an RNN-based network allows the predictions to be accurate while maintaining the internal nonlinear cross-dependencies within the **IM** vector. In order to overcome drawbacks such as short-term memory and vanishing gradients, Long-Short-Term-Memory (LSTM) units were deployed in the RNN structure. Moreover, the neural network was trained with cross-validation using randomly selected 80 % of the dataset and tested using the remaining 20 % of the dataset.

Traditionally, modelling cross-IM correlations is a two-stepped process, where known GMMs are used to obtain the mean predictions, and conditional spectrum (CS) or generalized conditional intensity measures (GCIM) approaches are employed to implement the cross-correlations within the **IM** vector. However, this process has certain drawbacks such as i) requirement of GMMs for each independent IM in the **IM** vector; ii) assumption of linear correlations between the IMs; iii) general requirement of a functional form for the GMMs; iv) assumption of stationarity of correlations with respect to source and site characteristics; and v) need for a statistically sound procedure to combine the IM predictions from the GMMs and their correlations. The proposed GGMM serves as a more accurate remedy to such limitations using a data-driven deep-learning-based approach.

The predictions of the proposed GGMM were tested using the coefficient of determination, R^2. It was observed that the R^2 is generally above 0.7 for both train and test sets except for $D_{5-95,geom}$ and $S_a(T)$ for very short periods (<0.5 s). Moreover, the $S_a(T)$ predictions of the GGMM were compared against two popular GMMs reported for Chilean subduction, i.e., Montalva *et al.* [10] and Idini *et al.* [11] using R^2. It was observed that the proposed GGMM leads to better goodness of fit for all periods than the other two GMMs and on average, leads to 0.2 higher R^2. Furthermore, the cross-correlation within the GGMM-based predictions of the **IM** vector were compared with the correlations between the true **IM** vector and are found to be identical. The GGMM was further compared with the CS by checking their spectral predictions under two randomly selected recorded scenarios, i.e., $M_w = 5$ and $R_{rup} = 74.4$ km; $M_w = 8.4$ and $R_{rup} = 62$ km. Through a close examination of the two prediction models, it is noticed that the GGMM spectral shape is highly consistent with the spectral shape of the recorded ground motion in both scenarios while having a lower variance band, and at the same time, the GGMM is capable of predicting different cross-correlated **IMs**. This was further seen by comparing the CS against the GGMM through the likelihood measure. It was observed that though the GGMM possesses lower variance, it still leads to higher likelihood as compared to the CS and show no evident bias with respect to M_w or R_{rup}.

Finally, a structural analysis exercise was conducted on a sophisticated nonlinear finite-element model representative of a 20-story steel special moment frame (SSMF) structure. The proposed GGMM and CMS were used to select 30 ground motions with minimal scaling (*i.e.*, scale factors between 0.5 and 2) for three hazard levels (*i.e.*, mean return period of 475, 975, and 2475 years). The peak interstory drift ratios (PIDRs) obtained from the two methods were statistically compared using the KS hypothesis test and deemed to be statistically unsimilar with larger PIDRs recorded using GGMM-based selected ground motions. Thus, the GGMM presented herein offers a robust model with six significant advantages over the conventional approaches: i) higher mean prediction accuracy; ii) inclusion of other different IMs apart from $S_a(T)$; iii) maintenance of cross-IM dependencies (not necessarily linear

correlations); iv) easily re-trainable with additional ground-motion data; v)flexibility for an extension to larger **IM** vectors (i.e., the inclusion of other important IMs); and vi) based on a data-driven approach with minimal practical assumptions involved.

In conclusion, the GGMM proposed herein offers a robust seismic hazard-consistent tool for the Chilean subduction environment that can be used for several purposes, including structural and geotechnical design and analysis (e.g., ground motion selection using multi-objective IM criteria), risk- and reliability-based decision-making (where hazard consistent IMs are required), validation of artificial/simulated ground motion records (by checking if the simulated ground motions comply with a set of possible **IM** vector corresponding to causal parameters), etc. Tools, such as the GGMM model, can provide users (both researchers and professional engineers) with a comprehensive methodology capable of estimating different IMs without the need for more complex routines to link the various models that currently exist for hazard-consistent assessments. The GGMM can allow users to use simple deterministic estimates of a seismic source and site parameters to generate a probabilistic distribution of the corresponding **IM** vector. The proposed GGMM can also be easily re-trained with new ground motion records or extended to a larger **IM** vector and appended to existing seismic hazard analysis toolbox.

8. Data and resource availability

The recorded ground motions used in this study can be obtained from Chilean database (https:// siberrisk.ing.puc.cl/StrongMotionDatabase). The authors have also developed an executable application to obtain estimations from the proposed GGMM. The application can be download easily from the following GitHub repository (https://github.com/jfayaz/GGMMSubd).

Declaration of Competing Interest

This research is funded by the European Union's Horizon 2020 research and innovation programme, specifically grant agreement number 821046: TURNkey "Towards more Earthquake-resilient Urban Societies through a Multi-sensor-based Information System enabling Earthquake Forecasting, Early Warning and Rapid Response actions."

Data availability

Data will be made available on request.

References

[1] Instituto Nacional de Estadisticas (INE). 2021.
[2] Earthquake Lists, Maps and S. U.S. Geological Survey 2020.
[3] Khazaradze G, Klotz J. Short-and long-term effects of GPS measured crustal deformation rates along the south central Andes. J Geophys Res Solid Earth 2003;108.
[4] Centro Sismológico Nacional (CSN). Universidad de Chile. 2021 n.d.
[5] Douglas J. Ground motion prediction equations 1964–2021 2021. http://www.gmpe.org.uk.
[6] Youngs RR, Chiou S-J, Silva WJ, Humphrey JR. Strong ground motion attenuation relationships for subduction zone earthquakes. Seismol Res Lett 1997;68:58–73.
[7] Abrahamson N, Gregor N, Addo K. BC Hydro ground motion prediction equations for subduction earthquakes. Earthq Spectra 2016;32:23–44.
[8] Zhao JX, Jiang F, Shi P, Xing H, Huang H, Hou R, et al. Ground-motion prediction equations for subduction slab earthquakes in Japan using site class and simple geometric attenuation functions. Bull Seismol Soc Am 2016;106:1535–51.
[9] Lin P-S, Lee C-T. Ground-motion attenuation relationships for subduction-zone earthquakes in northeastern Taiwan. Bull Seismol Soc Am 2008;98:220–40.
[10] Montalva GA, Bastías N, Rodriguez-Marek A. Ground-motion prediction equation for the Chilean subduction zone. Bull Seismol Soc Am 2017;107:901–11.
[11] Idini B, Rojas F, Ruiz S, Pastén C. Ground motion prediction equations for the Chilean subduction zone. Bull Earthq Eng 2017;15:1853–80.
[12] Fayaz J, Xiang Y, Zareian F. Generalized ground motion prediction model using hybrid recurrent neural network. Earthq Eng Struct Dyn 2021;50:1539–61.
[13] Bradley BA. A generalized conditional intensity measure approach and holistic ground-motion selection. Earthq Eng Struct Dyn 2010;39:1321–42.
[14] Baker JW, Jayaram N. Correlation of spectral acceleration values from NGA ground motion models. Earthq Spectra 2008;24:299–317.
[15] Candia G, Poulos A, de la Llera JC, Crempien JGF, Macedo J. Correlations of spectral accelerations in the Chilean subduction zone. Earthq Spectra 2020;36:788–805.
[16] Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou B-S-J, et al. NGA-West2 database. Earthq Spectra 2014;30:989–1005.
[17] Boore DM. Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. Bull Seismol Soc Am 2010;100:1830–5.
[18] Dhanya J, Raghukanth STG. Ground motion prediction model using artificial neural network. Pure Appl Geophys 2018;175:1035–64.
[19] Datta A, Wu DJ, Zhu W, Cai M, Ellsworth WL. DeepShake: shaking intensity prediction using deep spatiotemporal rnns for earthquake early warning. Seismol Soc Am 2022;93:1636–49.
[20] Fayaz J, Galasso C. A deep neural network framework for real-time on-site estimation of acceleration response spectra of seismic ground motions. Comput Civ Infrastruct Eng 2022.
[21] Chiang Y-J, Chin T-L, Chen D-Y. Neural network-based strong motion prediction for on-site earthquake early warning. Sensors 2022;22:704.
[22] Kim H-S. Development of seismic response simulation model for building structures with semi-active control devices using recurrent neural network. Appl Sci 2020;10:3915.
[23] Liu T, Dai Z. Real-time prediction of the trend of ground motion intensity based on deep learning. Shock Vib 2021;2021.
[24] McKenna F, Fenves GL, Scott MH. OpenSees: Open system for earthquake engineering simulation. Univ California, Berkeley, CA, Http//Opensees Berkeley Edu., Pacific Earthq Eng Res Center; 2006.
[25] Baker JW. Conditional mean spectrum: Tool for ground-motion selection. J Struct Eng 2011;137:322–31.
[26] Bommer JJ, Alarcon JE. The prediction and use of peak ground velocity. J Earthq Eng 2006;10:1–31.
[27] Macedo J, Abrahamson N, Bray JD. Arias intensity conditional scaling ground-motion models for subduction zones. Bull Seismol Soc Am 2019;109:1343–57.
[28] Petrone F, Abrahamson N, McCallen D, Miah M. Validation of (not-historical) large-event near-fault ground-motion simulations for use in civil engineering applications. Earthq Eng Struct Dyn 2021;50:116–34.
[29] Ming D, Huang C, Peters GW, Galasso C. An advanced estimation algorithm for ground-motion models with spatial correlation. Bull Seismol Soc Am 2019;109:541–66.
[30] Tsioulou A, Taflanidis AA, Galasso C. Validation of stochastic ground motion model modification by comparison to seismic demand of recorded ground motions. Bull Earthq Eng 2019;17:2871–98.
[31] Fayaz J, Azar S, Dabaghi M, Zareian F. Methodology for validation of simulated ground motions for seismic response assessment: Application to CyberShake source-based ground motions. Bull Seismol Soc Am 2021;111:226–41.
[32] Baker JW, Cornell CA. Pacific earthquake engineering research center. Berkeley: Univ California; 2006.
[33] Bray JD, Macedo J. 6th Ishihara lecture: Simplified procedure for estimating liquefaction-induced building settlement. Soil Dyn Earthq Eng 2017;102:215–31.
[34] Bullock Z. Ground motion models for Arias intensity, cumulative absolute velocity, peak incremental ground velocity, and significant duration in New Zealand. Bull New Zeal Soc Earthq Eng 2019;52:193–207.
[35] Rathje EM, Wang Y, Stafford PJ, Antonakos G, Saygili G. Probabilistic assessment of the seismic performance of earth slopes. Bull Earthq Eng 2014;12:1071–90.
[36] Travasarou T. Optimal ground motion intensity measures for probabilistic assessment of seismic slope displacements. Berkeley: University of California; 2003.
[37] Harp EL, Wilson RC. Shaking intensity thresholds for rock falls and slides: evidence from 1987 Whittier Narrows and Superstition Hills earthquake strong-motion records. Bull Seismol Soc Am 1995;85:1739–57.
[38] Keefer DK. Investigating landslides caused by earthquakes–a historical review. Surv Geophys 2002;23:473–510.
[39] Kayen RE, Mitchell JK. Assessment of liquefaction potential during earthquakes by Arias intensity. J Geotech Geoenviron Eng 1997;123:1162–74.
[40] Kostadinov MV, Towhata I. Assessment of liquefaction-inducing peak ground velocity and frequency of horizontal ground shaking at onset of phenomenon. Soil Dyn Earthq Eng 2002;22:309–22.
[41] Orense RP. Assessment of liquefaction potential based on peak ground motion parameters. Soil Dyn Earthq Eng 2005;25:225–40.
[42] Castro, S., Benavente, R., Crempien, J., Candia, G., de la Llera JC. A consistently processed strong motion database for Chilean Earthquakes; 2021.
[43] Gallagher K, Charvin K, Nielsen S, Sambridge M, Stephenson J. Markov chain Monte Carlo (MCMC) sampling methods to determine optimal models, model resolution and model choice for Earth Science problems. Mar Pet Geol 2009;26:525–35.
[44] Campbell KW, Bozorgnia Y. NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. Earthq Spectra 2008;24:139–71.
[45] Campbell K, Bozorgnia Y. Updated Ground Motion Models For AI And CAV Using The NGA-West2 Database 2019.
[46] Schmidhuber J, Hochreiter S. Guessing can outperform many long time lag algorithms 1996.
[47] Demidenko E, Stukel TA. Influence analysis for linear mixed-effects models. Stat Med 2005;24:893–909.

[48] Sheynin O. Helmert's work in the theory of errors. Arch Hist Exact Sci 1995;49:73–104. https://doi.org/10.1007/BF00374700.
[49] Poulos A, Miranda E. Relations between MaxRotD50 and some horizontal components of ground-motion intensity used in practice. Bull Seismol Soc Am 2021.
[50] Lin T, Baker J. Conditional Spectra. In: Beer M, Kougioumtzoglou IA, Patelli E, Au IS-K, editors. Encycl. Earthq. Eng., Berlin, Heidelberg: Springer Berlin Heidelberg; 2014, p. 1–13. 10.1007/978-3-642-36197-5_386-1.
[51] NCh 433. Diseño Sismico de Edificios; 2009.
[52] American Institute of Steel Construction. Seismic Provisions for Structural Steel Buildings. ANSI/AISC 341-16. 2016.
[53] American Institute of Steel Construction. Prequalified Connections for Special and Intermediate Steel Moment Frames for Seismic Applications. ANSI/AISC 358-16. 2016.
[55] Jalayer F, Cornell CA, Chakroborty S, Roy R, Miranda E, Eads L, et al. Assessment of the probability of collapse of structures during earthquakes. Earthq Eng Struct Dyn 2016;128:122. https://doi.org/10.1115/1.4034725.
[56] Torres-Rodas P, Flores F, Pozo S, Astudillo BX. Seismic performance of steel moment frames considering the effects of column-base hysteretic behavior and gravity framing system. Soil Dyn Earthq Eng 2021;144:106654. https://doi.org/10.1016/j.soildyn.2021.106654.
[58] Torres-Rodas P, Flores F, Zareian F. Seismic response of steel moment frame considering gravity system and column base flexibility. 11th Natl Conf Earthq Eng 2018, NCEE 2018 Integr Sci Eng Policy 2018;12:7383–93.
[59] ATC P. ATC 72-1: Modeling and acceptance criteria for seismic design and analysis of tall buildings. Atc 72 2010;1.
[60] Lignos DG, Krawinkler H. Deterioration modeling of steel components in support of collapse prediction of steel moment frames under earthquake loading. J Struct Eng 2011;137:1291–302. https://doi.org/10.1061/(asce)st.1943-541x.0000376.
[61] Hartloper AR. Updates to the Asce-41-13 Nonlinear modelling provisions for performance-based seismic assessment of new existing steel moment resisting frames. McGill University (Canada); 2017.
[62] Hartloper A, Lignos D. 11.29: Updates to the ASCE-41-13 provisions for the nonlinear modeling of steel wide-flange columns for performance-based earthquake engineering. Ce/Papers 2017;1:3072–81.
[63] Suzuki Y, Lignos DG. Large scale collapse experiments of wide flange steel beam-columns. 8th Int. Conf Behav Steel Struct Seism Areas 2015.
[64] Ibarra LF, Medina RA, Krawinkler H. Hysteretic models that incorporate strength and stiffness deterioration. Earthq Eng Struct Dyn 2005;34:1489–511. https://doi.org/10.1002/eqe.495.
[65] Rahnama M, Krawinkler H. Effects of soft soil and hysteresis model on seismic demands, vol. 108. John A Blume Earthquake Engineering Center Standford; 1993.
[66] Gupta A, Krawinkler H. Estimation of seismic drift demands for frame structures. Earthq Eng Struct Dyn 2000;29:1287–305. 10.1002/1096-9845(200009)29:9<1287::AID-EQE971>3.0.CO;2-B.
[67] Kanvinde AM, Grilli DA, Zareian F. Rotational stiffness of exposed column base connections: experiments and analytical models. J Struct Eng 2012;138:549–60. https://doi.org/10.1061/(asce)st.1943-541x.0000495.
[68] Rodas PT, Zareian F, Kanvinde A. Rotational stiffness of deeply embedded column-base connections. J Struct Eng 2017;143:04017064. https://doi.org/10.1061/(asce)st.1943-541x.0001789.
[69] Torres-Rodas P, Fayaz J, Zareian F. Strength resistance factors for seismic design of exposed based plate connections in special steel moment resisting frames. Earthq Spectra 2020;36:537–53. https://doi.org/10.1177/8755293019891714.
[71] Initiative PTB. Guidelines for performance-based seismic design of tall buildings. Technical Report PEER 2017; 2017.
[72] Baker JW, Allin CC. Spectral shape, epsilon and record selection. Earthq Eng Struct Dyn 2006;35:1077–95.
[74] Willmott CJ, Ackleson SG, Davis RE, Feddema JJ, Klink KM, Legates DR, et al. Statistics for the evaluation and comparison of models. J Geophys Res Ocean 1985;90:8995–9005.
[75] Fayaz J, Rezaeian S, Zareian F. Evaluation of simulated ground motions using probabilistic seismic demand analysis: CyberShake (ver. 15.12) simulations for Ordinary Standard Bridges. Soil Dyn Earthquake Eng 2021;141:106533. https://doi.org/10.1016/j.soildyn.2020.106533.
[76] Applied Technology Council. Quantification of building seismic performance factors. US Department of Homeland Security, FEMA P695; 2009.
[77] Massey Jr FJ. The Kolmogorov-Smirnov test for goodness of fit. J Am Stat Assoc 1951;46:68–78.
[78] Medalla M, Lopez-Garcia D, Zareian F. Seismic characterization of steel special moment frames subjected to megathrust earthquakes. Earthq Spectra 2020;36:2033–57.
[79] Chandramohan R, Baker JW, Deierlein GG. Quantifying the influence of ground motion duration on structural collapse capacity using spectrally equivalent records. Earthq Spectra 2016;32:927–50.
[80] Bravo-Haro MA, Elghazouli AY. Influence of earthquake duration on the response of steel moment frames. Soil Dyn Earthq Eng 2018;115:634–51.