17th World Conference on Earthquake Engineering, 17WCEE
Sendai, Japan - September 13th to 18th 2020
Paper N° C001603
Registration Code: A00555

APPLICATION OF GENETIC ALGORITHMS TO SELECT A GROUND MOTION SET FOR CONDUCTING HC-IDA

S. G. Shrestha [(1)], R. Chandramohan [(2)], R. P. Dhakal [(3)]

[(1)] PhD Candidate, University of Canterbury, Christchurch, New Zealand, srijana.gurungshrestha@pg.canterbury.ac.nz
[(2)] Lecturer, University of Canterbury, Christchurch, New Zealand, reagan.c@canterbury.ac.nz
[(3)] Professor, University of Canterbury, Christchurch, New Zealand, rajesh.dhakal@canterbury.ac.nz

Abstract

This paper describes the application of genetic algorithms to select a generic set of ground motions that can be used to conduct hazard-consistent incremental dynamic analysis (HC-IDA) on a wide range of structures located at any site. HC-IDA is a recently developed procedure that overcomes the primary drawback of traditional incremental dynamic analysis (IDA) by enabling the computation of a hazard-consistent collapse fragility curve. Hence, it offers an alternative to the commonly employed hazard-consistent multiple stripe analysis (MSA) procedure, but without the need for site-specific ground motion selection. The response spectral shapes and durations of the ground motions used to conduct HC-IDA should ideally be uniformly distributed over the range of response spectral shapes and durations likely to be expected at a wide range of representative sites. This uniform distribution of response spectral shapes and durations enables the structural failure surface to be estimated with the least amount of uncertainty. In this study, response spectral shape is quantified using the scalar metric S_aRatio, while duration is quantified using 5-75% significant duration ($Ds_{5\text{-}75}$). The ranges of S_aRatio and $Ds_{5\text{-}75}$ values anticipated at Wellington, New Zealand are computed using the generalized conditional intensity measure (GCIM) framework. A genetic algorithm is employed to select a suitable record set from a database of 2467 ground motions recorded from both shallow crustal and subduction earthquakes. Genetic algorithms employ operations such as mutation, crossover, and selection, inspired by the process of natural selection in evolution, to optimize highly nonlinear functions. The Latin hypercube sampling technique is used in this study to select the sets of ground motions constituting the first generation of chromosomes that have approximately uniform marginal distributions of S_aRatio and $Ds_{5\text{-}75}$. The fitness of the ground motion sets, quantified using the Kolmogorov-Smirnov test, is then optimized over successive generations by crossover and mutation operations. The selected ground motions are demonstrated to be able to predict the failure surface of a steel moment frame building more precisely compared to the FEMA far-field set. Hence, they can be used to compute the hazard consistent fragility curve of a wide range of structures located at a wide range of sites using HC-IDA.

Keywords: ground motion selection; genetic algorithms; incremental dynamic analysis; hazard-consistent; collapse fragility

1. Introduction

This paper describes the application of genetic algorithms to select a generic set of ground motions for conducting hazard-consistent incremental dynamic analysis (HC-IDA) [1]. HC-IDA is a recently developed procedure that imparts incremental dynamic analysis (IDA) the ability to compute hazard-consistent collapse risk estimates, provided the ground motions used in the procedure satisfy certain broad criteria. Such a generic ground motion set, once selected, can be employed to analyze different types of structures located at a wide range of sites. This makes HC-IDA easier to conduct compared to multiple stripe analysis (MSA) [2], which requires the selection of hazard-consistent site and structure specific record sets at different intensity levels.

The ground motions for HC-IDA are required to broadly cover the range of response spectral shapes and durations expected at a particular site of interest. Such ground motions should ideally be uniformly distributed to minimize the uncertainty in predicting ground motion collapse intensity. Selecting a single set of ground motions to fulfill these criteria for a set of targets is a complex optimization problem. Therefore, this study utilizes genetic algorithms which imitate the process of natural selection to select an optimal set of ground motions. The algorithm starts with an initial population of candidate record sets generated using the Latin hypercube sampling technique. The Kolmogorov-Smirnov (K-S) goodness of fit test result and Pearson's correlation coefficient are used by the algorithm to evaluate the fitness of record sets produced in each successive generation. The K-S test quantifies how closely the distribution of response spectral shapes and durations of a candidate set of records matches the expected marginal uniform distributions, while the Pearson's correlation coefficient quantifies the orthogonality of the response spectral shapes and durations of the records. The algorithm terminates after a certain number of generations have elapsed, producing an optimal set of ground motions.

This procedure is employed to select a generic set of ground motions covering the range of ground motion response spectral shapes and durations anticipated at Wellington, New Zealand. Wellington was chosen since it is a densely populated metropolitan region with high seismic risk and is exposed to both crustal and subduction earthquakes. The distributions of anticipated ground motion response spectral shapes and durations at other sites in New Zealand are expected to be contained within the corresponding distributions for Wellingon, permitting the selected record set to be used to analyse buildings at a wide range of sites in New Zealand.

2. Target distributions of response spectral shapes and durations

The spectral acceleration (S_a) at the fundamental modal period of vibration (T) at 5% damping ($S_a(T)$) is employed as the primary intensity measure (IM) used to characterize the ground motion intensity. It has been shown to be an effective predictor of structural response for a wide range of structures [3] and widely used in current structural design and assessment practice. Two secondary IMs—response spectral shape and duration—are used in conjunction with $S_a(T)$ since they have both been demonstrated by previous studies ([1], [4]–[6]) to be good predictors of structural collapse capacity. 5-75% significant duration ($Ds_{5\text{-}75}$) [7] is used in this study to quantify ground motion duration since it has been shown to be well suited to selecting ground motions for collapse risk assessment [1]. It is defined as the time interval over which 5 to 75% of the cumulative integral of the square of the ground acceleration is accumulated. Response spectral shape is quantified using a dimensionless parameter S_aRatio proposed by Eads et al. [4], which has also been shown to be a good predictor of structural collapse capacity. S_aRatio is defined as the ratio of $S_a(T)$ to the geometric mean of the portion of the response spectrum that lies between the periods 0.2T and 3.0T, as shown in Eq. (1). The response spectra of two ground motions with high and low S_aRatio values scaled to a common $S_a(1.2s)$ value are plotted in Fig. 1. The response spectrum of the ground motion with a low S_aRatio value, recorded from the 1990 Manjil, Iran earthquake, contains relatively high spectral ordinates at periods above and below 1.2s, while the ground motion with a high S_aRatio value, recorded from the 2011 Tohoku, Japan earthquake, contains relatively low spectral ordinates at periods above and below 1.2s.

$$S_a Ratio(T, 0.2T, 3.0T) = \frac{S_a(T)}{S_{a,avg}(0.2T,\ 3.0T)} \tag{1}$$

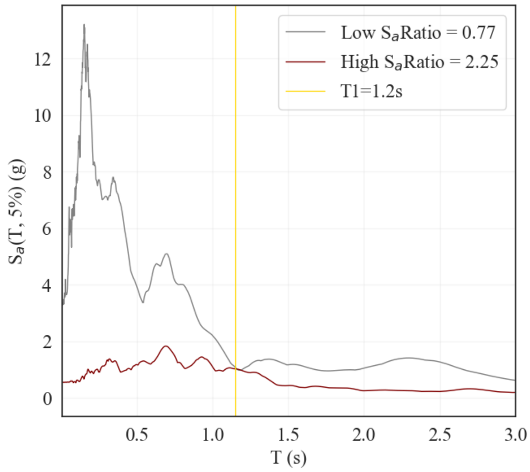
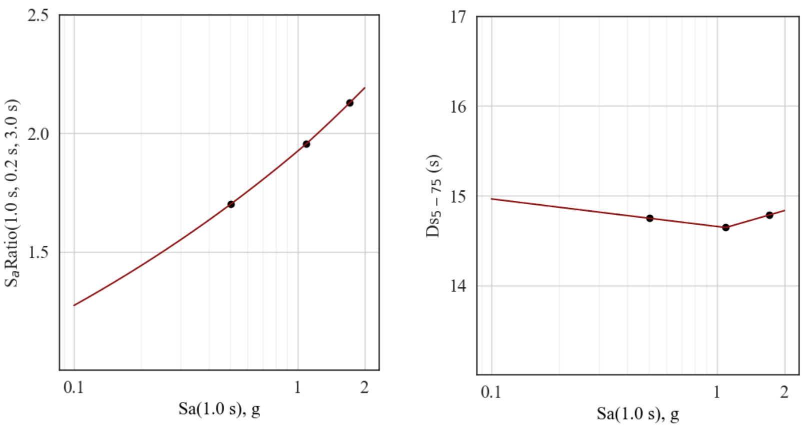


Fig. 1 – Response spectra of two ground motions scaled to a common value of S_a(1.2s) = 1g with low and high S_aRatio(1.2s, 0.2s, 3.6s) values.

The hazard-consistent target distributions of S_aRatio and $Ds_{5\text{-}75}$ anticipated in Wellington are computed using the generalized conditional intensity measure (GCIM) framework [8]. The GCIM computations require the use of prediction models for response spectra and duration, as well as models for the correlation between the prediction residuals of response spectral ordinates and duration. Additional details regarding the GCIM computations can be found in [1]. The GCIM computations were carried out using the open-source seismic hazard analysis platform, OpenQuake [9]. S_aRatio and $Ds_{5\text{-}75}$ targets were computed at Wellington, New Zealand, conditional on the exceedance of S_a at different periods (0.1 s, 0.2 s, 0.5 s, 1.0 s, 2.0 s, and 5.0 s) at three different hazard levels (10%, 2%, and 0.5% in 50 years). Wellington is chosen since it is a densely populated metropolitan region with a high seismic risk, that is susceptible to both short duration ground motions from crustal earthquakes and long duration ground motions from subduction earthquakes. Hence, the range of anticipated ground motion response spectral shapes and durations in Wellington is expected to encompass the anticipated ranges at a number of other sites in New Zealand. The median S_aRatio and $Ds_{5\text{-}75}$ targets, conditional on different exceedance probabilities of S_a(1.0 s) in Wellington are plotted in Fig. 2.



(a) (b)

Fig. 2 – Median (a) S_aRatio(1.0 s, 0.2 s, 3.0 s) and (b) $Ds_{5\text{-}75}$ targets, conditional on different levels of S_a(1.0 s).

3. Database of candidate ground motions

A database of 17,148 ground motions recorded from 334 earthquakes ranging in magnitude from M_W3.4 to M_W7.9 from the PEER NGA-West2 database [10], and 3973 ground motions recorded from large magnitude earthquakes such as the 1985 Michoacan, Mexico; 2010 Maule, Chile; and 2011 Tohoku, Japan earthquakes, was first assembled. Low intensity records with peak ground acceleration (PGA) lesser than 0.1g or peak ground velocity (PGV) lesser than 10 cm/s were first screened out of this database. Among the remaining records, only those with S_aRatio and $Ds_{5\text{-}75}$ values that lie within the union of the 5th to 95th percentile marginal ranges of anticipated S_aRatio and $Ds_{5\text{-}75}$ values in Wellington, conditional on the 2% in 50 year exceedance probability of S_a at 0.2 s, 0.5 s, 1.0 s, and 2.0 s, are selected as candidate ground motions. The application of these criteria produced a databse of 2467 candidate ground motions. The S_aRatio and $Ds_{5\text{-}75}$ values of these candidate ground motions are plotted in Fig. 3, along with the 5th to 95th percentile marginal target ranges.

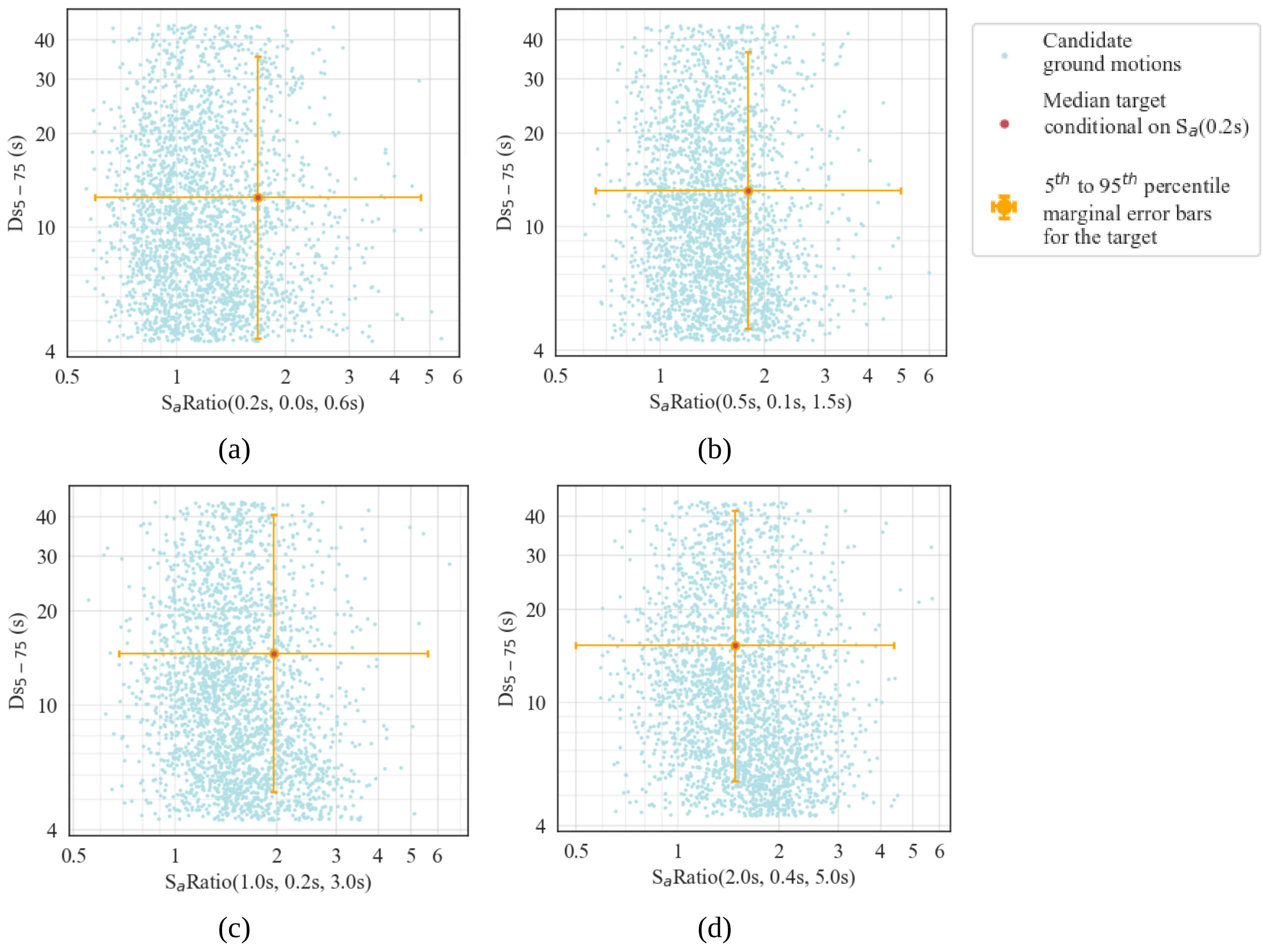


Fig. 3 – Candidate ground motions and the distributions of the S_aRatio and $Ds_{5\text{-}75}$ targets conditional on the 2% exceedance probability of S_a at (a) 0.2 s, (b) 0.5 s, (c) 1.0 s, and (d) 2.0 s

A subset of records is now selected from this database of candidate ground motions to conduct HC-IDA. This record set should ideally satisfy the following criteria to enable the estimation of the structural failure surface with the least amount of uncertainty, as discussed in Chandramohan [1].

i. The records must have uniform marginal distributions of S_aRatio (computed at different periods) and Ds_{5-75} values over the corresponding ranges of values anticipated at the site.

ii. The S_aRatio and Ds_{5-75} values of the records must exhibit orthogonality.

iii. It must contain a relatively large number of ground motions.

Selecting a single set of ground motions that satisfy these criteria is a complex optimization problem. A set of 100 ground motions can be selected from the 2467 candidate ground motions in more than 2.3 * 10^{180} combinations. Genetic algorithms (GAs) [11] are employed in this study to select an optimal set of ground motions. An overview of the employed ground motion selection procedure is presented in Fig. 4.

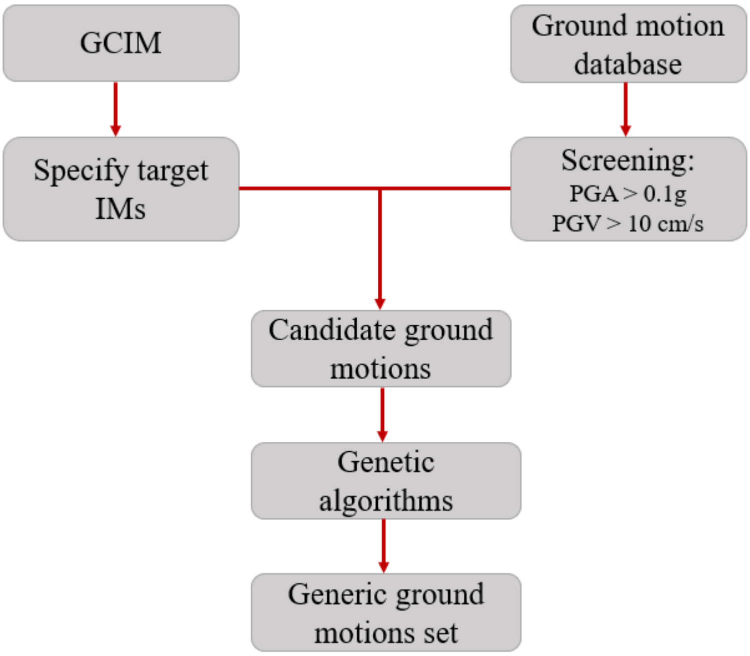


Fig. 4 – Ground motion selection procedure

4. Genetic algorithms for ground motion selection

Genetic algorithms (GAs) are used in this study to select a broad generic set of 100 ground motions from the database of 2467 candidate ground motions. GAs are based on Darwin's evolutionary theory, i.e., survival of the fittest. The basic components of the GA used in this study are described below:

i. **Initial Population**: This is the group of candidate record sets used to initialise the first generation of the optimisation algorithm. Each candidate record set is expressed as a binary string of 2467 0s and 1s [11], where 0 indicates a record from the database is not included in the set, and 1 indicates a record is. Since a set of 100 records is to be selected, the total number of 1s in a string should be 100. An initial population of 20 record sets with approximately uniformly distributed S_aRatio and Ds_{5-75} values was generated using the Latin hypercube sampling (LHS) method [12].

ii. **Fitness function**: This is a function characterising how well a record set matches the selection objective. "Fitter" record sets that better match the objective stand a higher chance of being picked as parents to produce offspring in the next generation [13]. The fitness function we used is a linear combination of *p*-values from a series of Kolmogorov-Smirnov (K-S) goodness of fit tests and a set of Pearson's correlation coefficients. The *p*-values of the K-S tests quantify how closely the distributions of the S_aRatio values computed at periods 0.2 s, 0.5 s, 1.0 s, and 2.0 s, and the Ds_{5-75} values of the records in a set follow uniform distributions, while the Pearson's correlation coefficients between the S_aRatio at each period and Ds_{5-75} quantify the degree of orthogonality between them.

iii. **Crossover**: This is the procedure used to produce offspring for the next generation from a randomly selected pair of parent candidate record sets from the current generation. We used a single-point crossover pattern, with the mid-point of each parent binary string serving as the crossover point. Hence, an offspring is produced by splicing the first half of one parent's binary string and the second half of the other [11]. This process is repeated to develop 10000 generations of offspring and the fittest offspring from the final generation is chosen as the optimal record set.

iv. **Mutation**: This is the process of randomly flipping some bits from 0 to 1 or vice versa during crossover, to introduce diversity in successive generations [11] . In this study, 4.

The response spectra of the 100 optimal ground motions selected using the GA, along with the mean, 15th, and 85th percentile spectra are plotted in Fig. 5. The distributions of the S_aRatio and $Ds_{5\text{-}75}$ values of the selected ground motions are shown in Fig. 6, overlaid on the 5th to 95th percentile marginal error bars of the targets conditional at different periods. While the $Ds_{5\text{-}75}$ values of the ground motions appear to be approximately uniformly distributed, the S_aRatio values, especially at lower periods, appear to deviate slightly from uniformity. These deviations at lower periods can be attributed to the scarcity of ground motions with large S_aRatio values in the database.

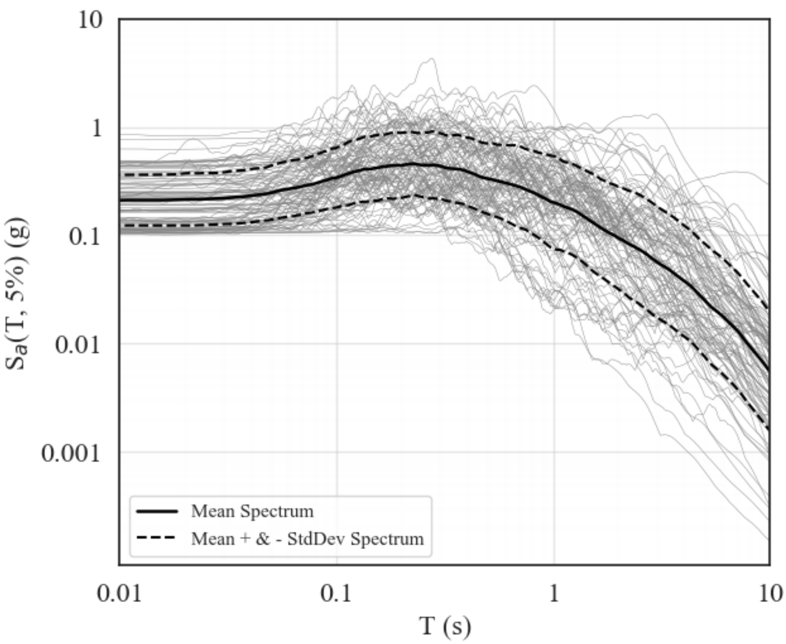


Fig. 5 – Response spectra of the 100 selected records.

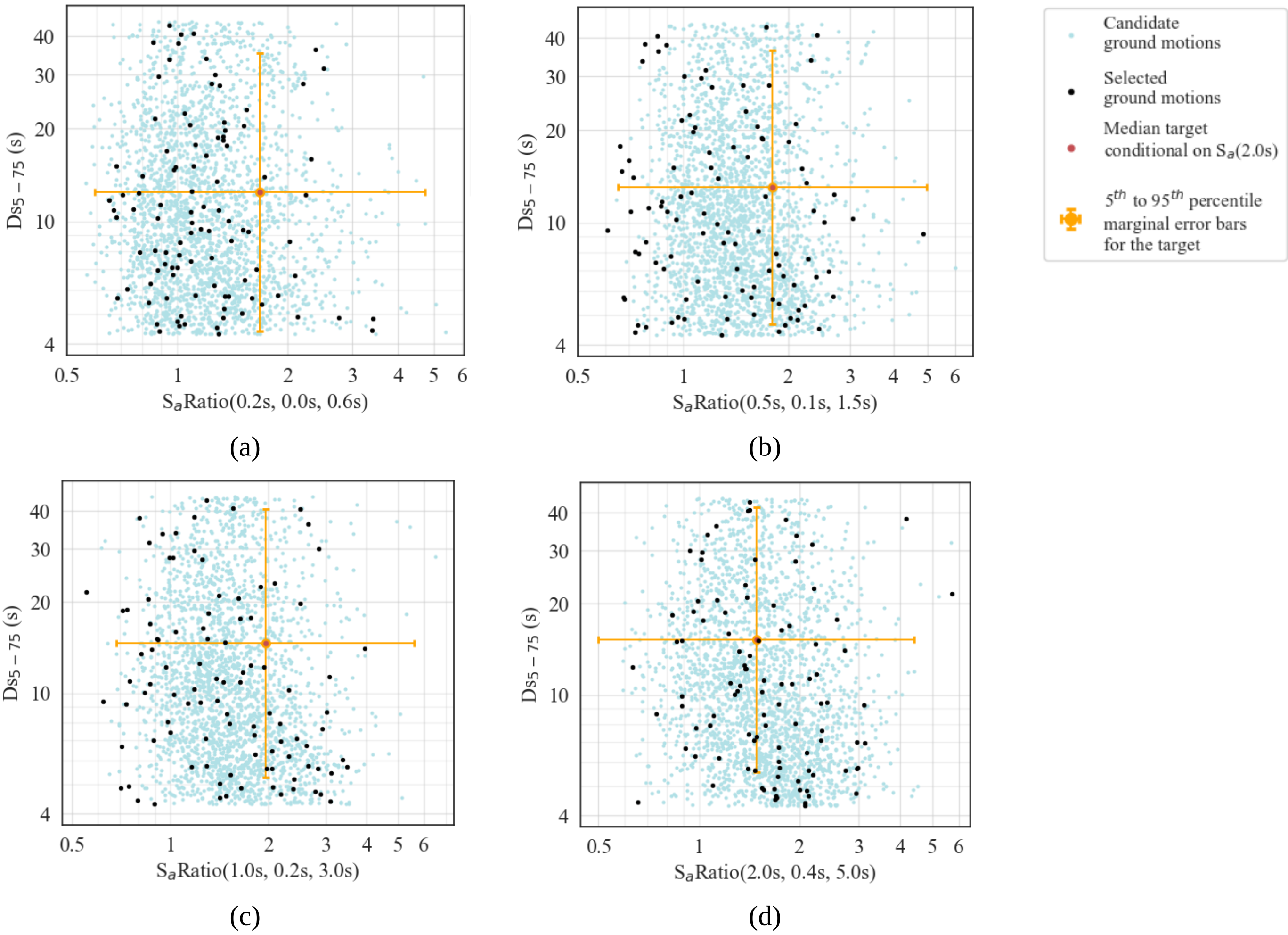


Fig. 6 – S_aRatio and $Ds_{5\text{-}75}$ values of the selected and candidate ground motions overlaid on the targets conditional on the 2% in 50 year exceedance probability of S_a at (a) 0.2 s, (b) 0.5 s, (c) 1.0 s, and (d) 2.0 s

5. Structural model and incremental dynamic analysis

A four-storey ductile steel moment resisting frame designed according to NZS 1170.5 [14], as a part of loss assessment study conducted by Sullivan et al. [15], is analysed using the selected ground motion set. This moment frame represents a typical mid-rise office building in the Wellington central business district (CBD) and is designed to be located on a site with soil type C. The first storey of the frame is 4.5 m tall and all other storeys are 3.6 m tall. It has three bays that are each 8 m wide. The fundamental modal period of the frame is 1.2 s.

The 2D model of the frame was created in OpenSees (Mazzoni et al. 2006) using a lumped plasticity approach, as illustrated in Fig. 7. The beams and columns are modelled using elastic beam-column elements. Zero-length rotational plastic hinges are placed at the reduced beam sections (RBS) of the beams and the ends of the columns. The hysteretic behaviour of the plastic hinges is modelled using the modified Ibarra-Medina-Krawinkler bilinear model [16], [17]. This hysteretic model is capable of capturing the in-cycle and cyclic degradation in strength and stiffness of the structural components, required to capture the effect of duration on structural response. The hysteretic shear behaviour of the finite panel zones is modelled using a quadrilinear backbone curve [18]. A pin-connected leaning column is modelled to capture the destabilising P-Δ effect of the internal gravity frames.

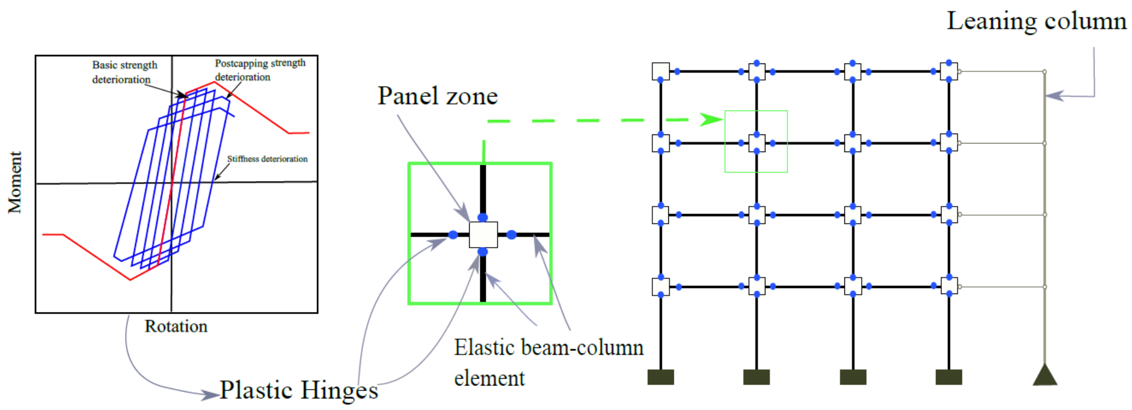


Fig. 7 – Schematic of the four-storey steel moment frame building model, and the modified IMK bilinear hysteretic model [17] used for plastic hinges.

Incremental dynamic analysis (IDA) [19] is conducted by progressively scaling each ground motion until it causes structural collapse, which is identified by the unbounded increase in simulated storey drift ratio (SDR) above a threshold of 0.10. The lowest intensity a ground motion needs to be scaled to, to cause structural collapse, is termed its collapse intensity [1]. IDA curves showing the variation in peak SDR with increasing Sa(T) are plotted for the 100 ground motions in Fig. 8. The explicit central difference was used to conduct all nonlinear dynamic analyses since it has been shown to be more robust (against non-convergence) and efficient compared to implicit time integration schemes [20]. All analyses were also conducted using efficient parallel algorithms on supercomputers accessed via DesignSafe [21].

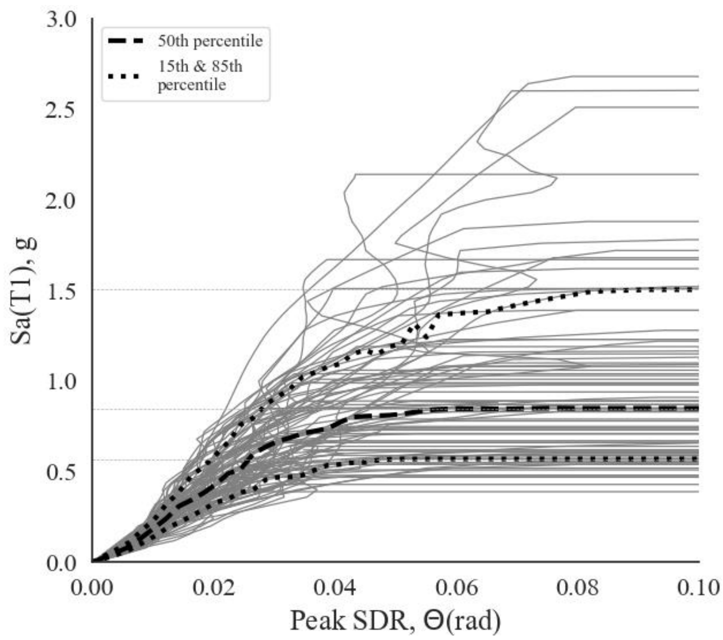


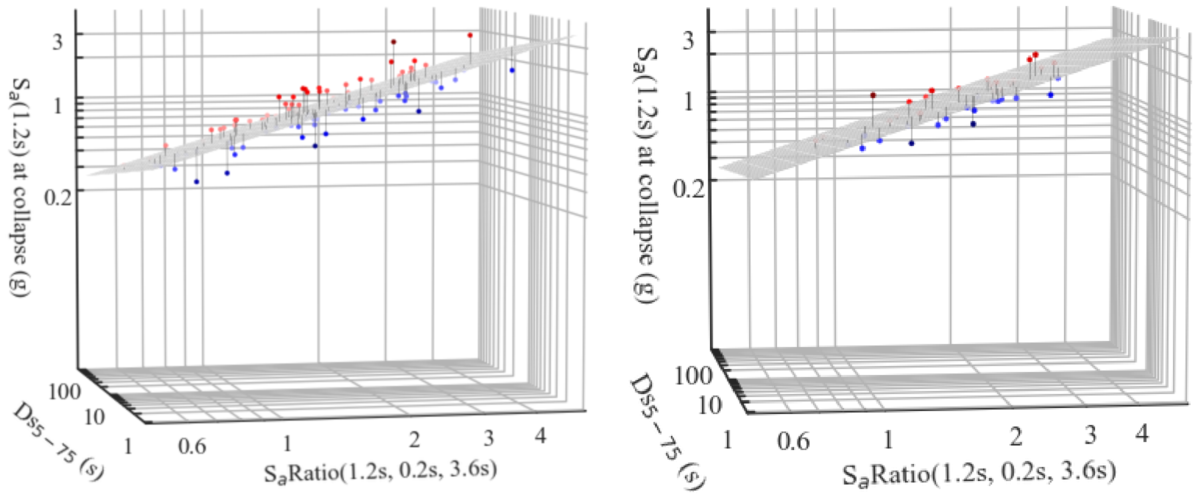
Fig. 8– Incremental dynamic analysis (IDA) curves

6. Structural failure surface

The structural failure surface is computed by fitting the multiple linear regression model in Eq. (2) to the IDA results. This regression model predicts the natural logarithms of the ground motion collapse intensities using the natural logarithms of the S_aRatio and $Ds_{5\text{-}75}$ values of the ground motions as predictors. The mean prediction, which represents a plane in 3-dimensional space, is plotted in Fig. 9(a). The coefficients c_{ss} and c_{dur} in the equation characterise the partial derivatives of the plane and quantify the sensitivity of ground motion collapse intensity to the effects of response spectral shape and duration respectively. ε represents the error term of the regression model and characterises the observed scatter in the data points about the plane.

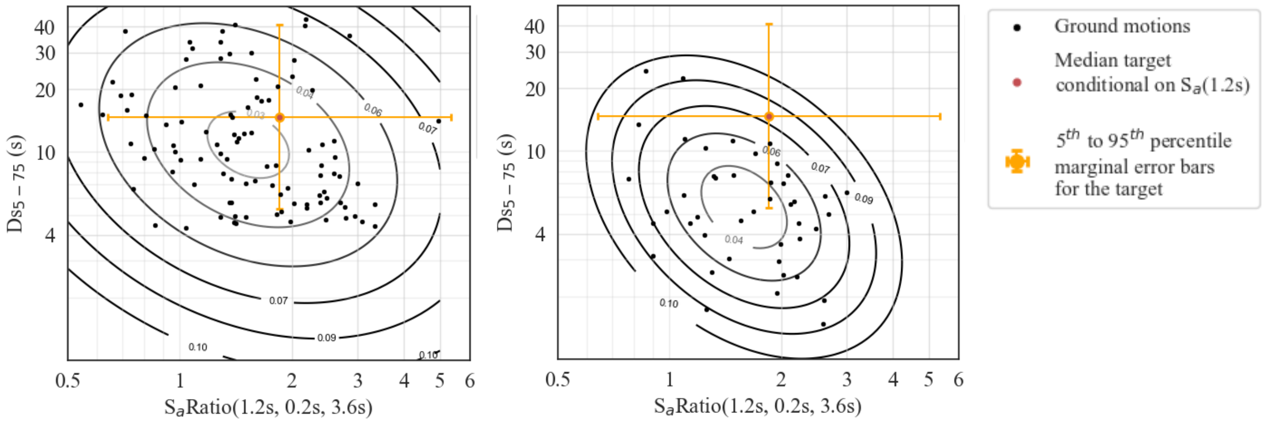
$$\ln S_a\left(T_1\right) \text{ at collapse} = c_0 + c_{ss}\ln S_a\text{Ratio} + c_{dur}\ln Ds + \varepsilon \tag{2}$$

The performance of the selected record set is compared to the FEMA P695 far field set [22], which is a commonly used record set consisting of 44 ground motions recorded from moderate magnitude crustal earthquakes. The failure surface computed using the FEMA P695 far field set is shown in Figure 9(b). The coefficients of determination (R^2) are 0.75 and 0.75 using the selected records set and the FEMA P695 far field set respectively. The standard error in predicting the median collapse intensity given the S_aRatio and $Ds_{5\text{-}75}$ of a ground motion is depicted using contours in Fig. 10.



(a) (b)

Fig. 9 – Regression models fit to the collapse intensities of the four-storey steel moment frame using (a) the selected set of 100 ground motions and (b) FEMA P695 far field set.



(a) (b)

Fig. 10 – Contour plots of the standard error in predicting the median collapse intensity, overlaid on the S_aRatio and Ds_{5-75} values of the (a) selected record set and (b) the FEMA P695 far field set, and the 5th to 95th percentile marginal error bars of the S_aRatio and Ds_{5-75} targets conditional on the 2% in 50 year exceedance probability of S_a(1.2s).

7. Discussion

The inclination of the failure surface in Fig. 9 indicates that the structure collapses at lower intensities under ground motions of longer durations and lower S_aRatio values. This implies that longer duration ground motions and those with lower S_aRatio are inherently more damaging. The R^2 values obtained from the regression analyses indicate that the S_aRatio and Ds_{5-75} of ground motions are able to explain approximately 75% and 74% of the variance in the ground motion collapse intensities obtained from the selected records set and FEMA P695 far field set respectively. Hence, response spectral shape and duration are seen to be good predictors of ground motion collapse intensity. As observed from the contour lines in Fig.10, the standard error in predicting the median collapse intensity in the region bounded by the marginal error bars of the S_aRatio and Ds_{5-75} targets is higher for the FEMA far-field records set compared to the selected ground motions. This is a consequence of the non-optimal distribution of S_aRatio and Ds_{5-75} values of the records in the FEMA P695 far field set. Hence, the selected record set is expected to produce a more precise estimate of the hazard-consistent collapse fragility curve using the HC-IDA procedure, than the FEMA P695 far field set.

8. Conclusion

A genetic algorithm based procedure to select a generic set of 100 ground motions for conducting HC-IDA is proposed in this study. The records are selected so as to have uniform and orthogonal marginal distributions of response spectral shape (quantified by S_aRatio) and duration (quantified by Ds_{5-75}). The records are selected to broadly cover the range of ground motion S_aRatio and Ds_{5-75} values anticipated in Wellington, New Zealand. A four-storey modern steel moment frame building is analysed using the selected record set, and its performance is compared against the FEMA P695 far field set. The structural failure surface is estimated using both record sets, and the standard error in predicting the median height of the failure surface is quantified. It is observed that the selected record set enables the estimation of the median height of the failure surface with lesser uncertainty within the region defined by the 5th and 95th percentile marginal error bars of the S_aRatio and Ds_{5-75} targets. Hence, the selected records set is expected to produce a more precise estimate of the hazard-consistent collapse fragility curve computed using the HC-IDA procedure, compared to the FEMA P695 far field set. The superior performance of the selected record set indicates its suitability for conducting HC-IDA on a wide range of structures located at a number of different sites in New Zealand.

9. Acknowledgements

This project was (partially) supported by QuakeCoRE, a New Zealand Tertiary Education Commission-funded Centre. This is QuakeCoRE publication number 0550. The authors would like to thank Timothy J. Sullivan, Amirhossein Orumiyehei, and Trevor Z. Yeow for sharing their ductile steel moment resisting frame model. The authors also acknowledge the computational resources provided by DesignSafe.

10. References

[1] R. Chandramohan, "Duration of earthquake ground motion: Influence on structural collapse risk and integration in design and assessment practice," Stanford University, 2016.

[2] F. Jalayer, "Direct probabilistic seismic anaysis: implementing non-linear dynamic assessments," Stanford University, 2003.

[3] N. Buratti, "Earthquake Intensity and Related Ground Motion," in *15th World Conference in Earthquake Engineering*, 2012.

[4] L. Eads, E. Miranda, and L. Dimitrios, "Spectral shape metrics and structural collapse potential," *Earthq. Eng. Struct. Dyn.*, vol. 45, no. 10, pp. 1643–1659, 2016.

[5] C. B. Haselton, J. W. Baker, A. B. Liel, and G. G. Deierlein, "Accounting for Ground-Motion Spectral Shape Characteristics in Structural Collapse Assessment through an Adjustment for Epsilon," *J. Struct. Eng.*, vol. 137, no. 3, pp. 332–344, 2011.

[6] M. Raghunandan and A. B. Liel, "Effect of ground motion duration on earthquake-induced structural collapse," *Struct. Saf.*, vol. 41, pp. 119–133, 2013.

[7] M. D. Trifunac and A. G. Brady, "A study on the duration of strong duration earthquake ground motion," *Bull. Seismol. Soc. Am.*, vol. 65, no. 3, pp. 585–626, 1975.

[8] B. A. Bradley, "A generalized conditional intensity measure approach and holistic ground-motion selection," *Earthq. Eng. Struct. Dyn.*, vol. 39, no. 12, pp. 1321–1342, Oct. 2010.

[9] M. Pagani *et al.*, "OpenQuake engine: An open hazard (and risk) software for the global earthquake model," *Seismol. Res. Lett.*, vol. 85, no. 3, pp. 692–702, 2014.

[10] T. D. Ancheta *et al.*, "PEER NGA-West2 Database : A Database of Ground Motions Recorded in Shallow Crustal Earthquakes in Active Tectonic," *15th World Conf. Earthq. Eng.*, p. 6, 2012.

[11] D. A. Coley, *An Introduction to Genetic Algorithms for Scientists and Engineers*, vol. 37, no. 2. World Scientific Publishing Co. Pte. Ltd., 1995.

[12] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 42, no. 1, pp. 55–61, 2000.

[13] F. Naeim, A. Alimoradi, and S. Pezeshk, "Selection and scaling of ground motion time histories for structural design using genetic algorithms," *Earthq. Spectra*, vol. 20, no. 2, pp. 413–426, 2004.

[14] NZS 1170.5, *Structural design actions Part 5: Earthquake actions - New Zealand*, no. 1. New Zealand, 2004.

[15] T. J. Sullivan, A. Orumiyehei, and T. Z. Yeow, "Options for simplified loss assessment," *(SubmittedUnder Rev.*, 2017.

[16] L. F. Ibarra, R. A. Medina, and H. Krawinkler, “Hysteretic models that incorporate strength and stiffness deterioration,” *Earthq. Eng. Struct. Dyn.*, vol. 34, no. 12, pp. 1489–1511, 2005.

[17] D. G. Lignos and H. Krawinkler, “Deterioration Modeling of Steel Components in Support of Collapse Prediction of Steel Moment Frames under Earthquake Loading,” *J. Struct. Eng.*, vol. 137, no. 11, pp. 1291–1302, 2011.

[18] K. D. Kim and M. D. Engelhardt, “Monotonic and cyclic loading models for panel zones in steel moment frames,” *J. Constr. Steel Res.*, vol. 58, no. 5–8, pp. 605–635, 2002.

[19] D. Vamvatsikos and C. Allin Cornell, “Incremental dynamic analysis,” *Earthq. Eng. Struct. Dyn.*, vol. 31, no. 3, pp. 491–514, 2002.

[20] R. Chandramohan, J. W. Baker, and G. G. Deierlein, “Robust and efficient nonlinear structural analysis using the central difference time integration scheme,” in *1st European Conference on OpenSees*, 2017.

[21] E. M. Rathje *et al.*, “DesignSafe: New Cyberinfrastructure for Natural Hazards Engineering,” *Nat. Hazards Rev.*, vol. 18, no. 3, pp. 1–7, 2017.

[22] FEMA P695, “Quantification of Building Seismic Performance Factors,” Federal Emergency Management Agency, Washington, 2009.