COMPDYN 2013
4th ECCOMAS Thematic Conference on
Computational Methods in Structural Dynamics and Earthquake Engineering
M. Papadrakakis, V. Papadopoulos, V. Plevris (eds.)
Kos Island, Greece, 12–14 June 2013

MEASURING BIAS IN INCREMENTAL DYNAMIC ANALYSIS USING BOOTSTRAP

Athanasia E. Zacharenaki[1*], Michalis Fragiadakis[2], and Manolis Papadrakakis[1]

[1] Institute of Structural Analysis and Antiseismic Research, National Technical University of Athens,
Iroon Polytecheiou Str., Zografou Campus, 15780, Athens, Greece
azachare@mail.ntua.gr, mpapadra@central.ntua.gr

[2] Department of Civil and Environmental Engineering, University of Cyprus
P.O. Box 20537, 1678 Nicosia, Cyprus
mfrag@ucy.ac.cy

Keywords: Incremental Dynamic Analysis, Cloud Analysis, Record Scaling, Bias, Bootstrap.

Abstract. *We present a methodology for the evaluation of the effect of scaling when Incremental Dynamic Analysis (IDA) is performed. The median capacity curve of IDA is compared to the capacity curve obtained using cloud analysis. Cloud analysis data contain results obtained using unscaled natural and synthetic ground motion records. Synthetic records were used due to the lack of a statistically significant number of natural records for large intensities. Nonlinear regression is performed with the aid of the Local Regression Smoothing Algorithm (LOESS) in order to post-process the results of cloud analysis. The primary difference between the two methods is that cloud analysis allows obtaining capacity curves without scaling the ground motion records, as opposed to the IDA algorithm. To investigate the statistical significance of this comparison, the bootstrap method is used. The bootstrap method is a powerful and easy-to-implement tool that allows calculating confidence intervals. Using bootstrap we are able to measure the bias introduced by record scaling when IDA is adopted. Thus, the bias is examined quantitatively and qualitatively for the full range of limit-states, yielding useful conclusions regarding scaling and its legitimacy in the context of IDA. A three-storey and a nine-storey steel moment resisting frames are used as testbeds for our investigations.*

1 INTRODUCTION

Incremental Dynamic Analysis (IDA) [1] is a method where the mathematical model of the structure is subjected to a number of ground motion records. Every record is scaled to several levels of intensity, producing the structure's capacity curve in terms of an Engineering Demand Parameter (EDP) versus an Intensity Measure (IM). IDA provides a powerful performance estimation framework, which, however, is often questioned due to the scaling of records with factors that considerably differ from one. This practice leads to ground motions that may not represent a realistic physical process and may under or over estimate the demand, or in other words, may introduce bias in the capacity estimation. This study investigates in a systematic manner the effect of record scaling providing an approach for measuring the bias introduced when IDA analysis is performed.

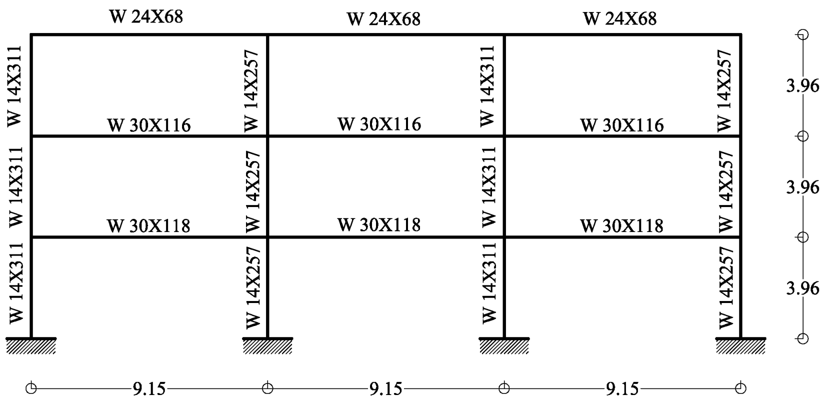
In Shome et al. [2] it is shown that small-to-moderate scaling factors do not introduce bias in the response estimation. It was also shown that there are structures for which scaling does not introduce bias, e.g. moderate period buildings in sites with no directivity. Vamvatsikos and Cornell [1] discuss how accurate the scaling practice is within the frameworks of IDA. They say that the problem depends on the EDP, the IM, the structure and the record population. They conclude that scaling is legitimate when the choice of the IM is such that the IM values, conditional on the EDP, are effectively independent of the magnitude and the distance scenario. Moreover, Iervolino and Cornell [3] observed that scaling arbitrarily selected records to match the strength of stronger records does not introduce bias in the seismic demand estimations. Luco and Bazzuro [4] suggest biased responses when the mean scale factor of a bin was larger than one. Furthermore, according to Baker [5] when the number of records that are scaled up is approximately equal to the number of records that are scaled down unbiased median interstory drift ratios are obtained.

The issue of selecting and scaling records without biasing the response has been investigated by several other researches. Katsanos et al. [6] discuss the different record selection procedures for seismic design. Watson-Lamprey and Abrahamson [7] show that a proper selection of records does not lead to biased results. Baker and Cornell [8] proposed selecting seismic records using the epsilon 'ε'-method in order to reduce the bias. Iervolino et al. [9] compared different procedures in terms of inelastic seismic response that led to obtain sets of spectral matching accelerograms for nonlinear dynamic analysis of structures. The outcome of the analysis leads to the fact that artificial or adjusted accelerograms may underestimate the displacement response when compared to original natural records which are considered as a benchmark. Kayhan et al. [10] propose a model to obtain input ground motion datasets compatible with given design spectra using a meta-heuristic harmony search algorithm.

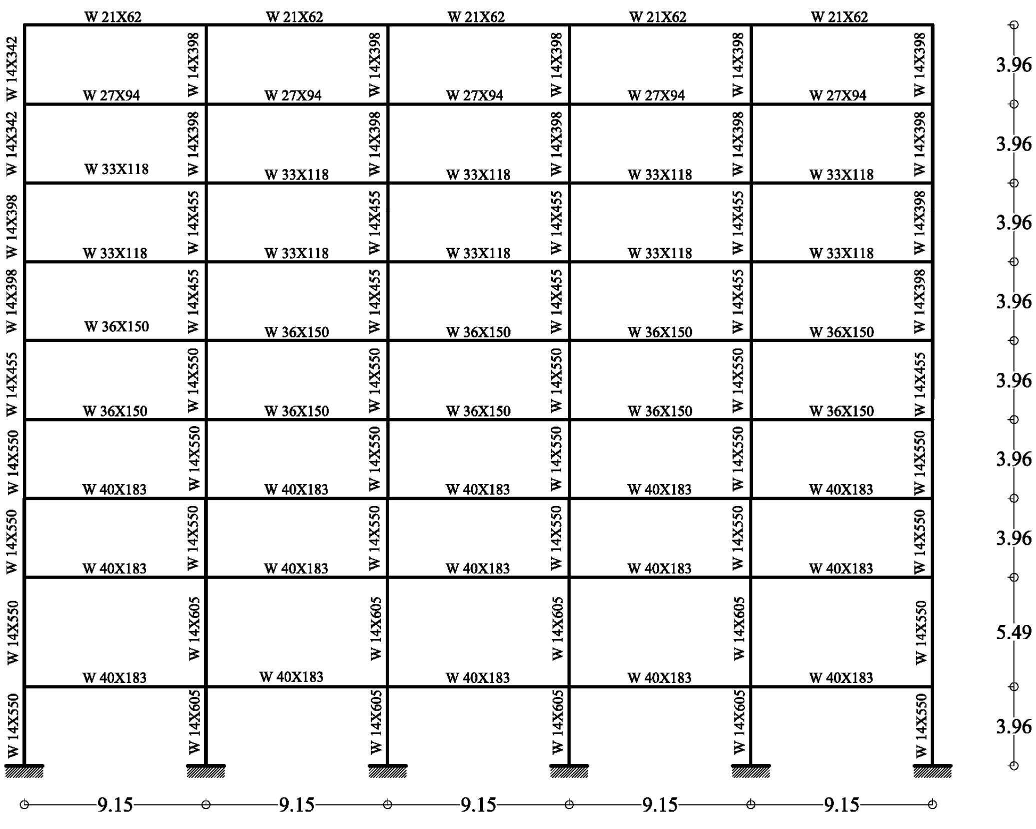
The above studies discuss the selection and scaling of ground motion records. They, also discuss the bias introduced in the response estimation when nonlinear response history analysis is performed. In this study the issue of scaling records within Incremental Dynamic Analysis (IDA) is investigated. To assess IDA, the response statistics of an Engineering Demand Parameter (EDP) are obtained using a large number of natural and synthetic records that are left unscaled. The structures considered, are two steel moment-resisting frame buildings. The bias introduced due to record scaling is assessed both qualitatively and quantitatively for the full range of limit-states.

2 BUILDINGS CONSIDERED

The multi-degree of freedom (MDOF) structures used in this study are two steel moment-resisting frames that have been designed for a Los Angeles site according to the 1997 NEHRP (National Earthquake Hazard Reduction Program) provisions. More specifically we study a three-storey and a nine-storey steel moment resisting frame, denoted as LA3 and LA9, respectively. Both buildings have been designed according to the strong-column, weak-beam design philosophy [11], while all analyses were conducted using the Opensees program [12]. Centerline models are used to model the two-dimensional exterior moment-resisting frame of each building. The fundamental periods of the frames are T_1=0.93sec and T_1=2.34sec, respectively. Therefore both moment-resisting frames are in essence first-mode dominated, while the LA9 building has some sensitivity also to higher modes.



(a)



(b)

Figure 1: Geometry and cross-sections of steel moment resisting frames: (a) LA3 building, (b) LA9 building.

3 GROUND MOTION RECORDS

For IDA analysis a set of 30 records was used. These records correspond to relatively large magnitudes of 6.5-6.9 that have been recorded on dense soil and bear no marks of directivity. Details about these records are given elsewhere [13].

For cloud analysis 1480 natural and synthetic ground motions are used. 1015 natural records were chosen randomly from the PEER NGA database [14] ensuring uniform processing. These records were selected giving no regard to a certain Intensity Measure (IM) and assuming dense soil. A large number of these records correspond to low magnitudes and when left unscaled give low $S_a(T_1,5\%)$ intensities and thus are not capable of causing yielding or collapsing of our structures.

For cloud analysis we have also used 465 synthetic records. The synthetic records were obtained following the procedure discussed in [15]. Assimaki et al. [15] combine regional velocity and attenuation structures that have been initially compiled using near-surface geotechnical data with the crustal velocity structure at three downhole arrays in Southern California. Broadband ground motion simulations were next conducted for rupture scenaria of weak, medium and large magnitude events (M_w=5÷7.5) and three component seismograms were computed on a surface station grid at distances 2-75km from the surface projection of the fault [16]. In this study we used horizontal components with magnitudes 6, 6.5, 7.5 each within a PGA range of 0.1~2.0g. Approximately 465 out of 3150 ground motions fulfilled the latter restriction.

4 METHODOLOGY

4.1 Incremental Dynamic Analysis and Cloud Analysis

In performance-based earthquake engineering (PBEE), the capacity curve of a building can be formed in the plane of an Engineering Demand Parameter (EDP) versus an Intensity Measure (IM). The Intensity Measure may be any of the record's characteristic parameters, such as peak ground acceleration (PGA), Arias Intensity or the five-percent damped, first-mode spectral acceleration, $S_a(T_1,5\%)$. The latter parameter takes into account the structure's first mode and is considered as an efficient, sufficient and practical IM which reflects the nonlinear response of the structure. The EDP considered here is the maximum interstorey drift over all the stories and is denoted as θ_{max}.

Incremental Dynamic Analysis (IDA) and Cloud Analysis (CA) are two popular seismic performance estimation methods. In IDA the mathematical model of the structure is subjected to a suite of ground motion records with increasing scaling factors offering thorough seismic demand and capacity prediction capability. Figure 2 shows, as an example, a 30-record IDA curves and the corresponding median curve of a nine-storey steel frame. To perform IDA several algorithms are available with the hunt-and-fill algorithm to be the most efficient [1]. According to Figure 2, IDA has to be repeated using different ground motion records in order to obtain meaningful statistical averages of the response, usually expressed as the median capacity curve and the 16^{th} and 84^{th} fractiles of the response.

In CA the mathematical model of the structure is subjected to a suite of ground motion records with a common scale factor, usually equal to one. The EDP-IM dataset which is formed from the nonlinear response history analysis is called "cloud".Once the cloud is formed a capacity curve can be obtained by nonlinear regression. When the scale factor is one the CA is called "single CA" since a single cloud in the EDP-IM plane has been formed. When several increasing scale factors are considered it is called "multiple CA".

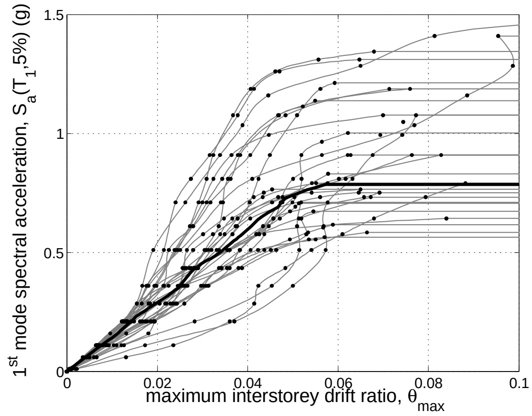


Figure 2: 30-record IDA curves and the corresponding median curve of a nine-story steel frame.

The LOESS (Locally Weighted Scatterplot Smoothing) algorithm [17] is used to perform nonlinear regression on the cloud of EDP-IM data when cloud analysis is adopted. This is a nonlinear regression algorithm that uses second degree polynomials and least square fitting to fit a polynomial curve on the data. A span of the moving average is needed for this algorithm in order to define a window of neighboring points that will be included in our estimations. The effect of the span value chosen can be seen in Figure 3. A large span of the moving average leads to increased smoothness, while a small span decreases the smoothness giving a curve that is more sensitive to the data (Figure 3a).

Nonlinear regression and the span value of the moving average are sources of additional bias on the seismic performance estimation. To decrease this outcome we chose the *k*-fold cross validation algorithm to obtain an optimal span value [17]. This algorithm has the following steps: i) the cloud is randomly partitioned to *k* subsamples, ii) a single subsample is retained as the validation cloud set, and iii) the remaining *k*-1 subsamples are used as training set to generate the LOESS curve. The mean squared error is the square of the distance between the LOESS curve of the training set and the curve produced by the testing set. The mean squared error with the span value is presented in Figure 3b. In this way the goodness-of-fit as function of the span value is evaluated. As optimal span value we take the one that minimizes the sum of the squared errors, i.e. the minimum value of the curve of Figure 3b.

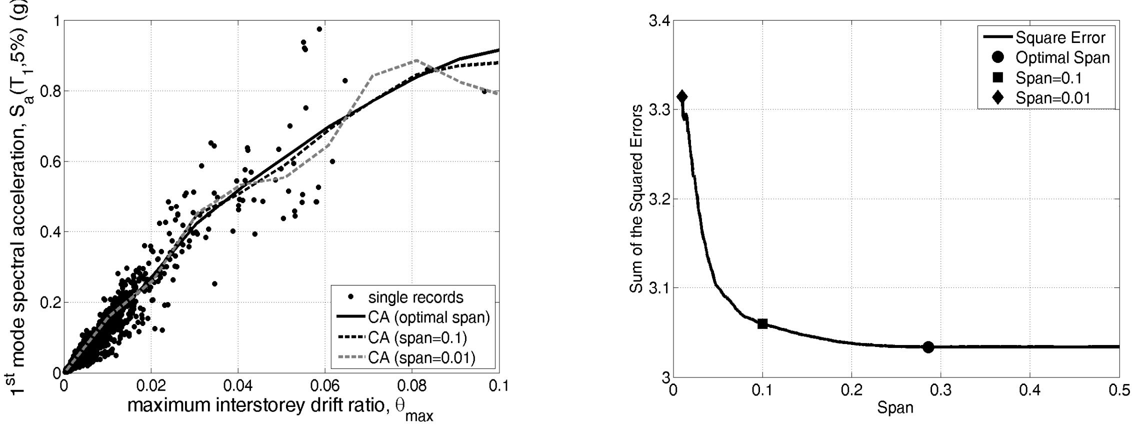


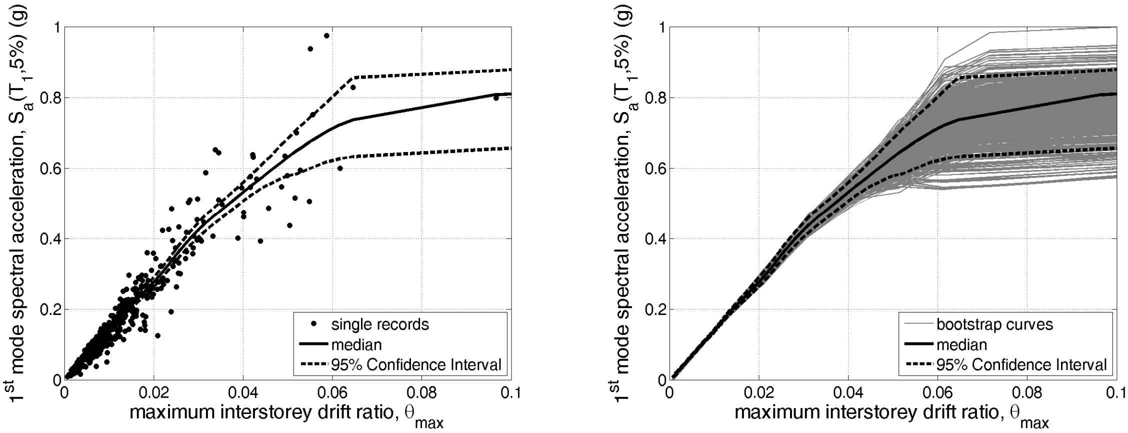
Figure 3: (a) Capacity curves for different span values of the LOESS fit and (b) square error of the LOESS fitting as function of the span.

4.2 Bootstrap analysis

Bootstrap [18] is used in order to investigate the statistical significance of our response statistics. The confidence interval of a quantity, here the IM values can be obtained with the aid of bootstrap. Bootstrap is an easy-to-implement tool which entails sampling with replacement from the initial population in order to generate a new population which we use to calculate the confidence intervals.

Bootstrap confidence intervals can be calculated for both IDA and CA. The process followed for both methods is the same. In both methods, the data is scattered on the EDP-IM plane. For both cases the bootstrap method resamples with replacement to obtain 1000 samples. Sampling with replacement means that after we have randomly drawn an observation from the original sample, we put it back before drawing the next observation. In IDA analysis this is repeated for the whole range of EDP values, since as samples we consider the θ_{max} and $S_a(T_1,5\%)$ values of every single record IDA curve. When cloud analysis is performed, every sample produces a new population of EDP-IM pairs on which nonlinear regression is performed, with the optimal span value as discussed in paragraph 4.1.

In Figure 4 the derivation of the bootstrap confidence intervals of the nine-storey steel moment-resisting frame is shown. In Figure 4a cloud analysis is performed on the scattered data, while in Figure 4b the bootstrap curves are shown. The dashed bold lines show the 95% bootstrap confidence interval while the solid bold lines the median curve obtained with LOESS. According to Figure 4a for $\theta_{\max}$ over 0.06 the original data are becoming scarce. However, this occurs for large drift values not affecting the early limit-states that usually are of interest. Moreover, due to limitation of cloud analysis to provide the dispersion, we are limited to study only the bias of the median $S_{\mathrm{a}}(T_1,5\%)$ capacities.



(a) (b)

Figure 4: Estimate of the median and 95% Confidence Intervals, versus: (a) the initial scattered data, and (b) 1000 capacity curves generated after bootstrapping the results of cloud analysis.

5 NUMERICAL RESULTS

Figure 5 presents the capacity curves and their 95% confidence intervals for the three-storey and the nine-storey building. For the LA3 frame the median IDA and CA curves coincide until $\theta_{\max}$=0.03. Beyond this value there is a slowly increasing deviation until $\theta_{\max}$=0.15. IDA underestimates the demand until $\theta_{\max}$=0.12 and beyond this value it overestimates it. The maximum difference between both curves is 0.4g in $\theta_{\max}$=0.15. For the nine-storey building, the IDA curve is entirely included within the confidence intervals of the CA curve as can be seen in Figure 5b. Therefore, IDA and CA analysis produce estimates of the capacity very close thus indicating that the bias in IDA is not significant.

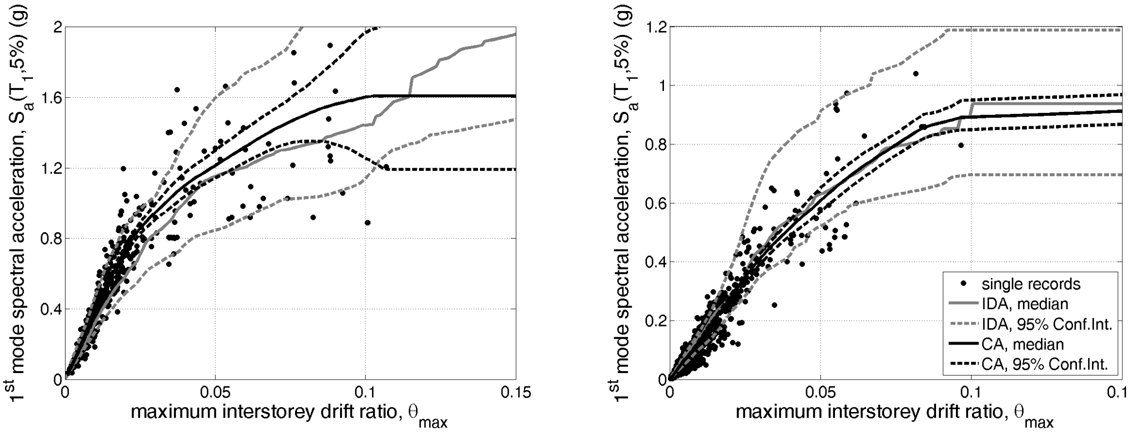


Figure 5: Median capacity curves and their 95% confidence intervals: (a) three-storey building (LA3), and (b) nine-storey building (LA9).

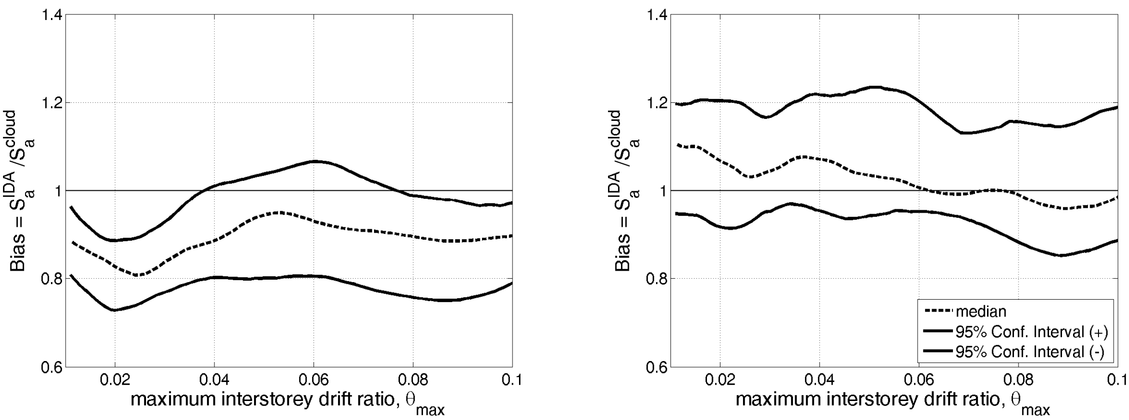
Bias is the systematic over or under-estimation of the $S_a(T_1,5\%)$ capacity conditional on the limit state θ_{max} demand. The measure of the bias here adopted is based on cloud analysis which is considered as unbiased. Therefore, the bias is quantified as:

$$bias = \frac{\left(\bar{S}_a\left(T_1,5\%\right)|\,\theta_{\max}\right)_{IDA}}{\left(\bar{S}_a\left(T_1,5\%\right)|\,\theta_{\max}\right)_{cloud}} \tag{1}$$

where $\bar{S}_a\left(T_1,5\%\right)|\,\theta_{\max}$ is the bootstrap $S_a(T_1,5\%)$ capacities conditional on maximum interstorey drift, θ_{max}. To quantify the bias and calculate the corresponding confidence intervals we perform bootstrap on the bias values as estimated using Eq.(1). Therefore, the confidence intervals of the bias show the effect of scaling the ground motion records within IDA.

The confidence intervals of the bias give an estimate whether the scaling practice biases the response. The bias quantification can be realized by observation of the width of the confidence intervals, the position of the unity line and their symmetry. If the unity line is not entailed in the confidence interval, then we assume that, at the 95% level, scaling is not the cause of the difference. Still, if the interval contains the unity line, we have reasonable evidence reinforcing the previous assumption that when the unity line is close to one bound there is some bias induced by scaling. Little evidence exists if the unity line is in the middle of the interval. In this case, scaling has doubtful effect in the estimation of bias. The width of the confidence interval, also provide an indication of the variability of the parameter studied.

Figure 6 presents the results of the bias for the three-storey and the nine-storey steel moment resisting frames as function of θ_{max}. As Figure 6 shows, the bias in the overall nonlinear structural response is more prominent in the three-storey than in the nine-storey building as the state of collapse is approached. According to Figure 6a for the three-storey building a slight bias is observed, since both bounds of the interval are below the unit line, while the distance of the bounds starts to increase from θ_{max}=0.02 until 0.05 and decreases from θ_{max}=0.05 until θ_{max}=0.1. For the nine-storey building where the IDA and CA curves were close (Figure 5b), the confidence intervals are symmetric with respect to the unity line, as shown in Figure 6b, thus indicating that the effect of bias is small for this building.



(a) (b)

Figure 6: Bootstrap 95% confidence intervals on the ratio of the median $S_a(T_1,5\%)$-capacities given θ_{max} of the IDA case over the cloud case (Eq. 2) for: (a) three storey building, (b) the nine storey building.

6 CONCLUSIONS

This study presents an approach for approximately assessing the effect of scaling in Incremental Dynamic Analysis. We perform nonlinear regression on cloud analysis containing building responses from unscaled natural and synthetic ground motion records. On this basis the influence of the scaling is enabled with the aid of the bootstrap method, which allows to quantify the bias. For the three-storey steel moment resisting frame the IDA curve does not induce significant bias in the seismic capacity assessment, where a small effect is observed from θ_{max}= 0.02 until 0.1. For the nine-storey building the bias was not significant for the whole range of limit states. In conclusion, for both buildings the bias introduced by IDA is small and within limits acceptable in the engineering practice.

REFERENCES

[1] D. Vamvatsikos, C.A. Cornell, Incremental dynamic analysis. *Earthquake Engineering & Structural Dynamics*, **31**, 491–514, 2002.

[2] N. Shome, C.A. Cornell, P. Bazzurro, J.E. Carballo, Earthquakes, records, and nonlinear MDOF responses. *Earthquake Spectra*, **14**, 469–500, 1998.

[3] I. Iervolino, C.A Cornell, Record selection for nonlinear seismic analysis of structures. *Earthquake Spectra*, **21,** 685–713, 2005.

[4] N. Luco, P. Bazzurro, Does amplitude scaling of ground motion records result in biased nonlinear structural drift responses? *Earthquake Engineering & Structural Dynamics*, **36**, 1813-1835, 2007.

[5] J.W. Baker, Measuring bias in Structural Response caused by ground motion scaling. *8th Pacific Conference on Earthquake Engineering*, December 5-7, 2007.

[6] E.I. Katsanos, A.G. Sextos, G.D. Manolis, Selection of earthquake ground motion records: A state-of-the-art review from a structural engineering perspective. *Soil Dynamics and Earthquake Engineering*, **30**, 157-169, 2010.

[7] J. Watson-Lamprey, N. Abrahamson, Selection of ground motion time series and limits on scaling. *Soil Dynamics and Earthquake Engineering,* **26,** 477-482, 2006.

[8] J.W. Baker, C.A. Cornell, Vector-Valued Ground Motion Intensity Measures for Probabilistic Seismic Demand Analysis. *John A. Blume Earthquake Engineering Center. Report No. 150, Stanford, CA*, 2006.

[9] I. Iervolino, F. De Luca, E. Cosenza, Spectral shape-based assessment of SDOF nonlinear response to real, adjusted and artificial accelerograms. *Engineering Structures*, **32**, 2776-2792 , 2010.

[10] A.H. Kayhan, K.A. Korkmaz, A. Irfanoglu, Selecting and scaling real ground motion records using harmony search algorithm. *Soil Dynamics and Earthquake Engineering*, **31,** 941-953, 2011.

[11] FEMA-350, Recommended seismic design criteria for new steel moment-frame buildings. Report No. FEMA-350, *SAC Joint Venture*, *Federal Emergency Management Agency*, Washington DC, 2000.

[12] F. McKenna, G.L Fenves, The OpenSees Command Language Manual -Version 1.2. *Pacific Earthquake Engineering Research Centre*, *University of California*, *Berkeley*, 2001.

[13] D. Vamvatsikos, C.A. Cornell, Developing efficient scalar and vector intensity measures for IDA capacity estimation by incorporating elastic spectral shape information, *Earthquake Engineering & Structural Dynamics*, **34**, 1573-1600, 2005.

[14] PEER. PEER NGA Database. Pacific Earthquake Engineering Research Center. Available from http://peer.berkeley.edu/smcat/, last accessed on 2008.

[15] D. Assimaki, J. H Steidl, P.-C Liu, Attenuation and Velocity Structure for Site Response Analyses via Downhole Seismogram Inversion. *Pure and applied geophysics*, **163**, 81-118, 2006.

[16] D. Assimaki, W. Li, M. Fragiadakis, Site Effects in Structural Performance Estimation, *Earthquake Spectra*, **28**, 859-883, 2012.

[17] W.S. Cleveland, S.J. Delvin, Locally-weighted Regression: An approach to Regression Analysis by local fitting. *Journal of American Statistical Association*, **83**, 596-610, 1988.

[18] B. Efron, R. Tibshirani, *An introduction to the Bootstrap*. New York, Chapman & Hall, 1993.