SECED 2015 Conference: Earthquake Risk and Engineering towards a Resilient World
9-10 July 2015, Cambridge UK

A CORRELATION STUDY OF BUILDING MONETARY LOSS WITH SEISMIC DEMANDS THROUGH INSURANCE CLAIMS ANALYSES OF 2011 TOHOKU EARTHQUAKE

Tao LAI[1], Arash NASSERI[2], Jayadipta GHOSH[3], Alvaro FARIAS[4] and Mesut TUREL[5]

Abstract: Correlation of building damage with ground motion parameters has been studied by many researchers using analytical modelling. This paper presents a correlation study using observational data consisting of ground motion recording and insurance claim data from the 2011 Tohoku earthquake. Considering both ground motion parameters and structural demand parameters, through statistical analysis, this study evaluates the capability of various parameters in predicting building damage. It is shown that while no single parameter consistently stands out, Arias Intensity (IA) and spectral acceleration at periods close to the fundamental period of structures correlate well with building damage for all regression methods and goodness-of-fit measures used.

Introduction

Over the past several decades, researchers have attempted to quantify seismic damage to buildings as a function of ground motion intensity parameter. Relationship between building damage and the ground motion intensity parameter is manifested in fragility curves for different damage states, or in "damage functions" which blends fragility curves for each damage state and their economic consequence (e.g. repair cost) into a single curve [Rota et al., 2010; Kircher et al., 1997; Ramamoorthy et al., 2006; and Hwang and Huo, 1994]. However, since most of these functions are developed using purely analytical methods, the uncertainty of damage prediction for a given ground motion intensity parameter is underestimated when compared to actual observations. Few studies exist on assessing the correlation of different ground motion intensity parameters with observed damage from real earthquakes. The study presented here attempts to address this issue using real insurance-claims data and actual ground motion recordings from the 2011 Tohoku earthquake.

The magnitude 9.0 Tohoku earthquake on March 11, 2011 near the northeast coast of Japan is one of the greatest earthquakes in history [Takewaki et al., 2011]. The earthquake spawned a massive tsunami that devastated coastal areas. In addition to tsunami damage in the inundated areas, shake damage was widespread especially in the inland areas [AIR-Worldwide, 2011]. In light of extensive instrumentation in Japan, this earthquake has produced a wealth of seismic recordings that allows researchers to perform studies which are often impaired by lack of data. In addition to strong motion records, this study benefits from a large suit of insurance claims data filled and evaluated after the earthquake. Combination of the ground motion and claims data provides a unique opportunity to reassess the correlation of different ground motion intensity parameters with observed building damage.

Recorded ground motion time histories are used to determine 15 different intensity parameters. These can be broadly grouped into structure-independent and structure-dependent parameters. After aggregating the claims loss data within a 1km radius of the seismograph stations, the predictive capability of each ground motion intensity measure is

[1] Assistant Vice President, AIR Worldwide, Boston, MA; tlai@air-worldwide.com
[2] Senior Research Engineer, AIR Worldwide, Boston, MA; anasseri@air-worldwide.com
[3] Assistant Professor, Indian Institute of Technology, Bombay India; jghosh@civil.iitb.ac.in
[4] Research Engineer, AIR Worldwide, Boston, MA; afariasvelasco@air-worldwide.com
[5] Research Engineer, AIR Worldwide, Boston, MA; mturel@air-worldwide.com

assessed using different regression functions and a multitude of goodness-of-fit estimates. The following sections of this paper describe the ground motion records, insurance-claims data and the different ground motion intensity parameters considered in this study. The ensuing sections present a discussion on the adopted nonlinear model fitting techniques and error measures used in this study. Results are presented in the subsequent section which offers different perspectives on the error measures, in absolute or scaled terms independent of the model fit and measurement technique.

Ground motion records and the insurance claim data

Following the 2011 Tohoku earthquake, over 3 million claims data were made available to the research team at AIR Worldwide. The data comprised structural characteristics, replacement value, damage and repair cost (payment). Damage ratio, defined as the ratio of repair cost to replacement value shows the economic consequence of the earthquake and is an indirect indication of the level of damage experienced. In this study damage ratios are taken as the primary measure of damage.

Using the claims data, damage ratios are calculated for three common residential construction types, namely, Wood, Steel and Reinforced Concrete buildings of varying number of stories grouped as low-rise (1- 3 stories), mid-rise (4-7stories) and high-rise (8 stories and higher). The percentage of data available for each category is shown in Table1. As evident from the table, the majority of the claims belongs to the category of low-rise wood-frame buildings, and are hence chosen to perform this study. The histogram of the damage ratios for low-rise wood buildings available from the claims data as shown in Fig.1 demonstrates the wide variability in observed damage estimates ranging from very low to significantly high damage.

Table 1. Percentage of claims data for different categories of building type and height

Building Type	Building Height		
	Low-rise	Mid-Rise	High-Rise
Wood	82.30%	0.0078%	0.0075%
Steel	7.85%	0.18%	0.009%
Concrete	1.18%	0.74%	0.23%

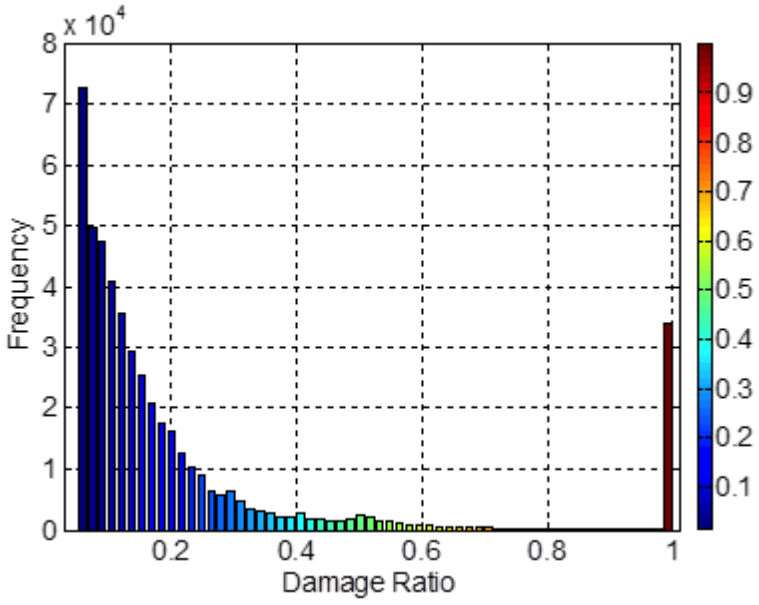

Figure 1. Histogram of damage ratios for low-rise wood-frame structures in Tohoku earthquake

The ground motion time history records used in this study are obtained from 1226 recording stations (K-Net and Kik-net), made available by the National Research Institute for Earth Science and Disaster Prevention in Japan. Geographic distribution of the available claims data for low-rise wood-frame buildings and seismic recording stations are shown in Fig.2 along with the damage ratios inferred from the claims data. Out of the 1226 recording stations, 327 stations were found to be within a 1km radius of the claims locations. The damage ratios within the 1km radii were grouped and averaged to reduce the impact of ground motion uncertainty in this study.

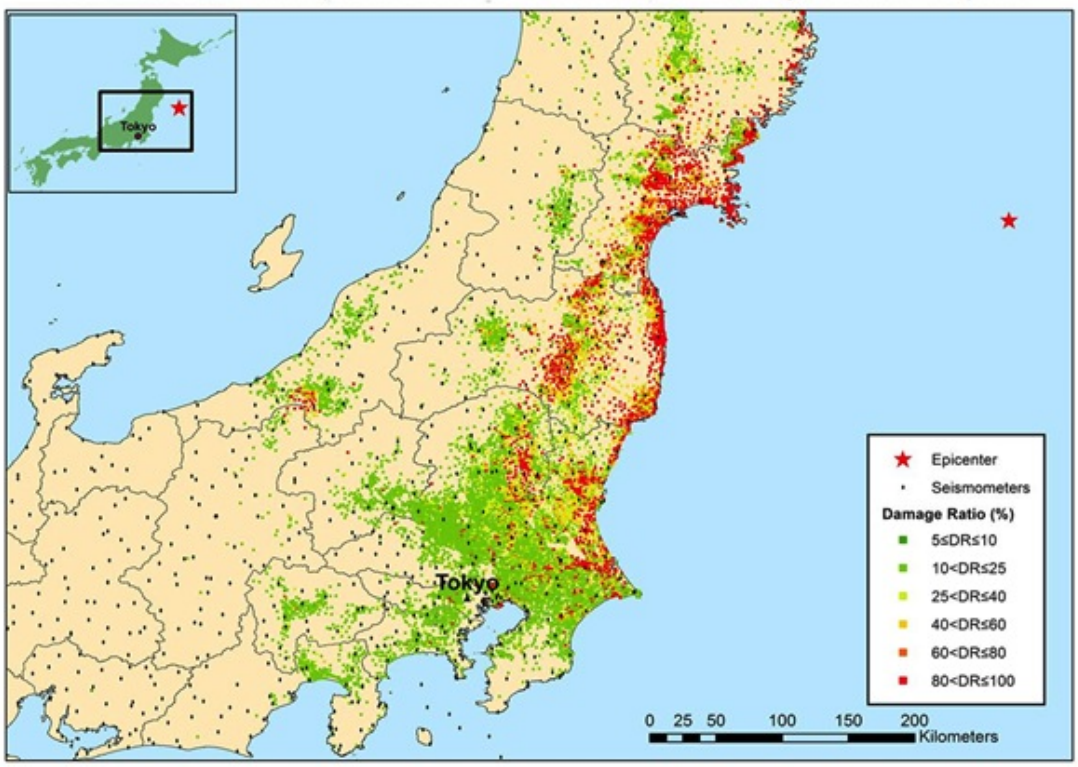

Figure 2. Distribution of seismic recording stations, and the low-rise wood frame damage ratio from the available claims data of the Tohoku earthquake

Ground motion intensity parameters studied

A total of 15 different ground motion intensity parameters, commonly used to correlate with building damage, are considered in this study. Among these parameters 9 are structure-independent intensities which are determined solely from the ground motion records and the remaining 6 are parameters that depend on structure's period. Table2 briefly describes these intensity parameters and the acronyms used throughout the remainder of this paper.

Table 2. Description of the intensity parameters studied

Intensity Measure	Description
PGA (g)	The maximum acceleration (PGA), velocity (PGV) and displacement (PGD) experienced by the ground at a given location during an earthquake. These measures are widely used in engineering practice, partially because they are easily measured
PGV (m/s)	
PGD (m)	
PGA/PGV (1/s)	PGA is not proportional to PGV and the peak of each parameter happens at different frequencies [Goodno, 2009]. PGA/PGV is therefore more descriptive of the ground motion than either measure alone

VSI (m)	Velocity Spectrum Intensity (VSI) is the integral of the pseudo-velocity spectrum (PSV) over a wide period range. For the purpose of this study the integral is taken over a 0 to 4.5s period range with 5% viscous damping
IA (m/s)	The Arias Intensity (IA) measures ground motion intensity using Eq.1 [Wilson, 1993]: $IA = \frac{\pi}{2g}\int_0^{T_{gm}} A(t)^2 dt$ (1) where T_{gm} is the duration of the ground motion and A(t) is the acceleration time history
RMSA (g)	The Root Mean Square of the Acceleration time history given in Eq.2 [McCann, 1980] $RMSA = \sqrt{\frac{1}{T_{gm}}\int_0^{T_{gm}}[A(t)]^2\, dt}$ (2)
CAV (m/s)	The Cumulative Absolute Velocity measures the absolute value of the integral of the acceleration time history given by Eq.3. CAV accounts for the duration effects of earthquake ground motion [Campbell and Bozorgnia, 2010] $RMSA = \sqrt{\frac{1}{T_{gm}}\int_0^{T_{gm}}[A(t)]^2\, dt}$ (3)
Number of Cycles	Several definitions of the number of cycles of a ground motion exist [Hancock, 2006]. For the purpose of this research, the number of cycles is simply the number of full cycles completed by the ground during earthquake ground motion
S_{aT} (g)	Spectral acceleration at period T. In this study 5% damped spectral acceleration at 0.2, 0.3, 1.0, 2.0, 3.0 and 4.0 seconds are considered

Fitting methods and error measures

Relationship between building damage and intensity parameters usually follows a nonlinear monotonically increasing shape. Three monotonically increasing functions are examined in this study to fit a curve to the observed damage and intensity parameters. Performance of each intensity parameter as a predictive measure depends on the function used in the fit and on the criteria set for evaluating errors. Thus, in this study several error measures are utilized to reduce the bias in the performance evaluation. The steps involved in this procedure are outlined below

- For a particular ground motion intensity measure *IM*, divide the intensity range into *n* bins as [0, IM_1, IM_2, …, IM_n]
- For each intensity bin [IM_i , IM_{i+1}] calculate the mean of claims damage ratio
- Fit a nonlinear function *F* to damage ratios in step 2 as a function of the ground motion intensity *IM*
- Compute the goodness of fit of the function *F* to the damage data points using an error measure *E*. These error measures are examined to identify which ground motion intensity correlates better with observed damage

The nonlinear fits used in this study are briefly described below:

Lognormal Cumulative Density Function (LCDF):

The lognormal cumulative distribution function described by Eq.4 is characterized by two distribution parameters μ and σ which are, respectively, the mean and standard deviation of

the associated normal distribution. The parameters μ and σ are calculated after minimizing the residual error between the function prediction and the claims damage ratio.

$$LCDF = \frac{1}{2\pi\sigma}\int_0^{im}\frac{e^{-\frac{[\ln(t)-\mu]^2}{2\sigma^2}}}{t}\,dt \tag{4}$$

Adaptive Cubic Spline Function (ACSF):
Intensity range is divided into 10 equally spaced intervals or "knots" and within each knot a piecewise function of the form depicted in Eq.5 is fitted. The spline is fitted using the Shape Language Modelling tool in MATLAB [D'Errico, 2009].

$$ACSF = a\left(im - IM_{knot_i}\right)^3 + b\left(im - IM_{knot_i}\right)^2 + c\left(im - IM_{knot_i}\right) + d \tag{5}$$

Piecewise Linear Functions (PLF):
PLF applies linear interpolation to each interval of the intensity parameter and can be formulated as shown in Eq.6.

$$PLF = DR_i + (DR_{i+1} - DR_i)\frac{(im - IM_i)}{(IM_{i+1} - IM_i)} \tag{6}$$

Examples of the nonlinear model fits are presented in Fig.3 for a structure independent ground motion intensity PGA and a structural frequency dependent intensity $Sa_{1.0}$.

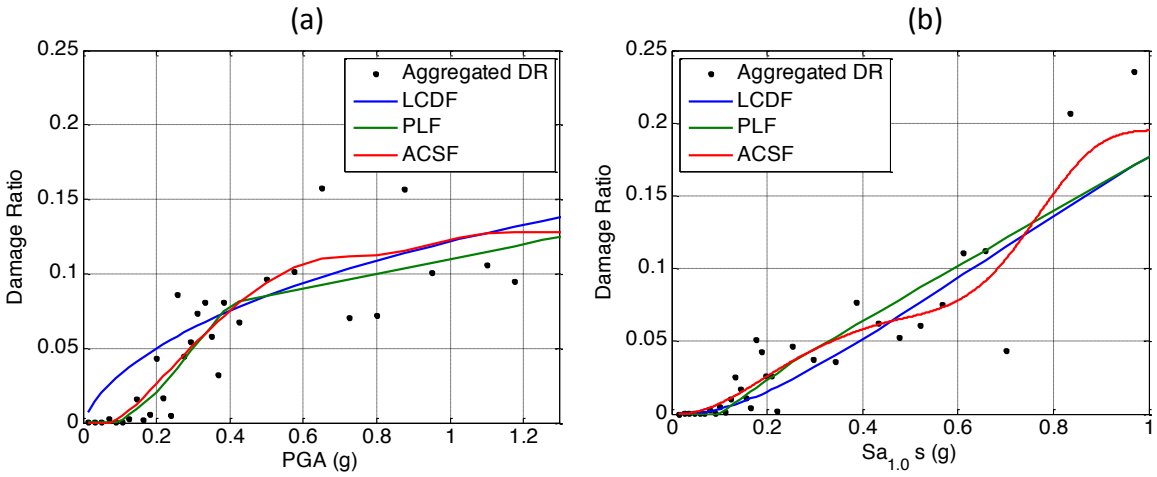

Figure 3. Adopted nonlinear model fits to aggregated claims damage ratio as a function of ground motion intensities: a) PGA and b) Sa1.0

In order to evaluate the goodness-of-fit and ultimately the predictive capability of each intensity parameter six error measures, as described below, are utilized in this study.

Mean Square Error (MSE):
The mean square error is a measure of the average of the square of differences between the actual observations and the model prediction and is expressed as in Eq.7

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{DR}_i - DR_i\right)^2 \tag{7}$$

where *n* is the number of damage ratio data points, $\widehat{\mathrm{DR}}_\mathrm{i}$ is the predicted damage ratio from the nonlinear model fit and DR_i is the actual damage ratio data point.

Mean Absolute Error (MAE):
Mean absolute error [Willmott and Matsuura, 2005] quantifies the average error of the model fit regardless of the sign of the error as expressed in Eq. 8

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\widehat{DR}_i - DR_i| \tag{8}$$

Mean Absolute Percentage Error (MAPE):
MAPE, as defined by Eq.9, is a normalized error is a widely used in statistics, especially in relation to trend estimation.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\widehat{DR}_i - DR_i}{DR_i}\right| \tag{9}$$

Systematic Mean Absolute Percentage Error (SMAPE):
The *SMAPE*, defined in Eq.10, is a variation on the *MAPE* that takes the average of the absolute value of the observations and predictions in the denominator.

$$SMAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\widehat{DR}_i - DR_i}{(\widehat{DR}_i + DR_i)/2}\right| \tag{10}$$

Mean Percentage Error (MPE):
MPE, defined in Eq.11 is similar to the *MAPE* except for it uses the actual, rather than the absolute error between observations and predictions. Since positive and negative errors can offset each other, *MPE* can be used as a measure of the bias in the predictions.

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{DR}_i - DR_i}{DR_i} \tag{11}$$

1-R^2 Error (RE):
Generally in regression analysis, the coefficient of determination, R^2, is a popular measure of model fit wherein higher the value of R^2, the better the fit. Since the other goodness of fit measures adopted in this study are error estimates, the authors have used a unique variation of coefficient of determination, namely, 1- R^2, such that the direction of discrepancy measurement is same as the other error measures.

Results and discussion
The error measures adopted in this study, as explained above, do not have consistent units and therefore cannot be directly compared. In order to unify the error measures for the purpose of evaluating the intensity measures, the errors are scaled to a range of [0, 1] across different model fits using Eq.12. The scaling ensures that for a particular model fit, the ground motion intensity with the poorest predictive capability has a scaled value of 1 (highest error) while the ground motion intensity with the best damage prediction has a scaled value of 0 (least error).

$$SE_{i,j,k} = \frac{AE_{i,j,k} - \min\left(AE_{1,j,k}, AE_{2,j,k}, \ldots, AE_{6,j,k}\right)}{\max\left(AE_{1,j,k}, AE_{2,j,k}, \ldots, AE_{6,j,k}\right) - \min\left(AE_{1,j,k}, AE_{2,j,k}, \ldots, AE_{6,j,k}\right)} \tag{12}$$

In the above equation, $SE_{(i,j,k)}$ is the scaled error for the i^{th} ground motion intensity corresponding to the j^{th} error measure and the k^{th} model fitting technique. Fig. 4 shows examples of the scaled estimates of two error measures 1-R^2 and *MAPE*.

Although scaling ensures that the error measures are independent of the measuring procedure, it does not unify the impact of fitting techniques. An additional step is required to

render these error measures independent of the model fitting procedure. This is achieved by averaging the errors and repeating the scaling procedure. Table 3 summarizes this process and presents the scaled average error for all the intensity parameters. Note that the columns labelled *LCDF*, *PLF* and *ACSF* show the error measures scaled using Eq.12 for the respective fitting techniques. Column Scaled Average shows the rescaled measures after averaging for various fitting techniques.

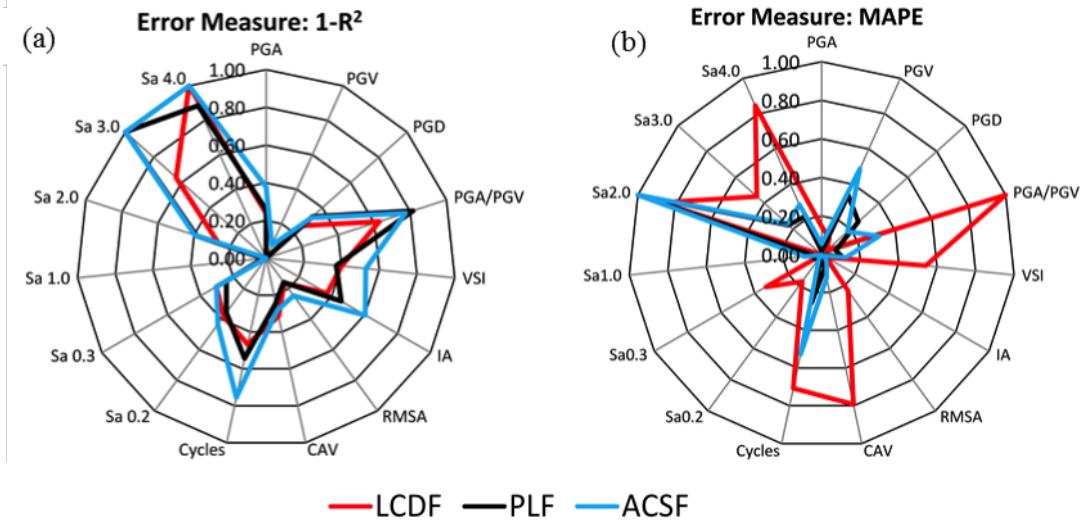

Figure 4. : Scaled damage prediction error estimates for two error measures a) 1-R2, and b) MAPE corresponding to the three adopted nonlinear model fits to claims damage data

Table 3. Scaled error independent of the error measures (columns LCDF, PLF, ACSF) and independent of fitting function (column Scaled Average).

Intensity Parameter	LCDF	PLF	ACSF	Average	Scaled Average
PGA	0.229	0.085	0.333	0.216	0.237
PGV	0.071	0.170	0.417	0.219	0.241
PGD	0.186	0.404	0.271	0.287	0.318
PGA/PGV	0.829	0.553	0.708	0.697	0.783
VSI	0.629	0.468	0.583	0.560	0.628
IA	0.100	0.085	0.208	0.131	0.141
RMSA	0.300	0.106	0.167	0.191	0.209
CAV	0.686	0.191	0.313	0.397	0.442
Cycles	0.629	0.511	0.813	0.651	0.731
Sa 0.2	0.271	0.106	0.208	0.195	0.214
Sa 0.3	0.286	0.000	0.042	0.109	0.116
Sa 1.0	0.000	0.021	0.000	0.007	0.000
Sa 2.0	0.614	1.000	1.000	0.871	0.982
Sa 3.0	0.500	0.660	0.708	0.623	0.699
Sa 4.0	1.000	0.766	0.896	0.887	1.000

Fig.5 graphically illustrates the results in a spider-web plot similar to Fig.4. Panel *a* shows scaled errors (independent of error measures) for the three fitting techniques while panel *b* shows the scaled average values which are independent of both error measures and fitting techniques. It can be seen from Fig.5b that among all the ground motion intensities under consideration, $Sa_{1.0}$ emerges as the best predictor of damage. Among all the structural period dependent ground motion intensity measures, Sa0.2, Sa0.3 and Sa1.0 are found to

predict damage fairly well compared to other higher-period intensities. This observation is consistent with the fact that the fundamental structural period of typical low-rise wood-frame buildings in Japan ranges between 0.1-0.3 seconds with potential elongation of the structural period following yielding under seismic forces. Among the structural independent intensity measures, the Arias Intensity (*IA*) is found to be a good estimator of damage. This is consistent with past research (Travasarou et al., 2003) in which Arias intensity was found to perform particularly well for the prediction of the response of short-period structures.

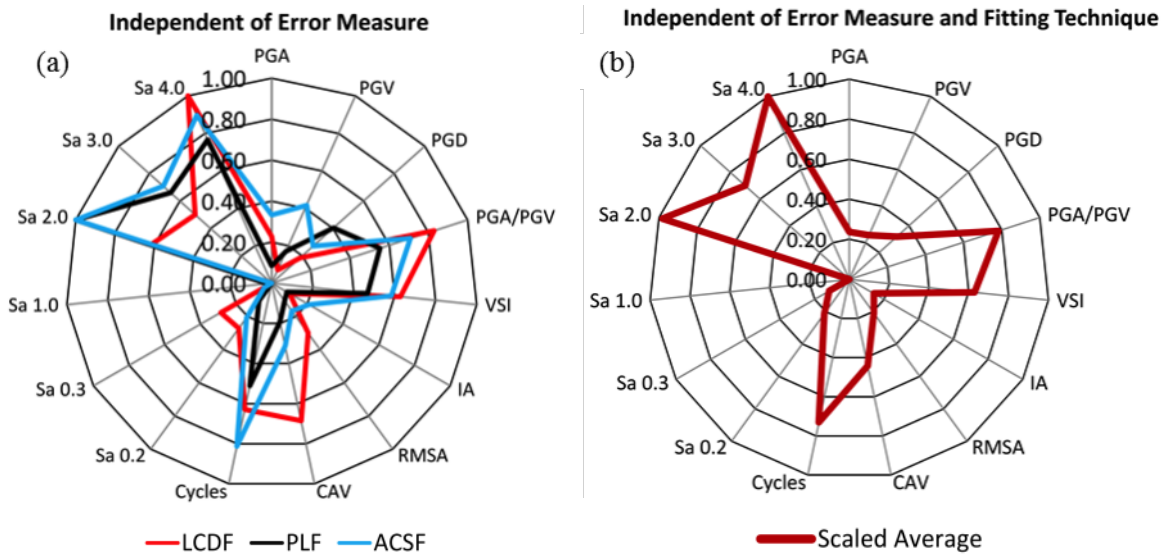

Figure 5. : Scaled damage prediction error estimates a) independent of error measuring technique, and b) independent of both the error measuring technique and nonlinear fitting

Conclusion

This study attempts to identify different ground motion intensity measures that correlate well with observed damage following earthquake events. To this end, 15 different ground motion intensity parameters derived from actual ground motion records of the 2011 Tohoku earthquake are investigated. Damage ratios calculated from insurance claims data averaged within a 1km radius of recording stations are taken as sample observations. Three different nonlinear fitting methods are used to fit the damage ratios as a function of the ground motion intensity. Goodness of fit of each of these fitting techniques is then evaluated using six different error measures.

Examination of the scaled error (unitless range of [0,1] where 1 denotes the least correlation) indicates that no single intensity parameter emerges as the best predictor for a particular fitting technique. Averaging and rescaling the errors across all the fitting methods to eliminate the dependency on fitting techniques, one observes that among the structure-independent ground motion intensity measures, Arias Intensity is reasonably well correlated with damage. Among the structure-dependent intensity measures, spectral acceleration at 1.0 seconds is found to perform best, closely followed by spectral accelerations at periods of 0.2 seconds and 0.3 seconds which are consistent with the natural period of typical low-rise wood-frame buildings in Japan.

REFERENCES

Takewaki, I., Murakami, S., Fujita, K.; Yoshitomi, S., Tsuji, M., (2011). *The 2011 off the Pacific coast of Tohoku earthquake and response of high-rise buildings under long-period ground motions*. Soil Dyn. Earthq. Eng. 2011, 31, 1511–1528

AIR-Worldwide, 2011. *Understanding Earthquake Risk in Japan Following the Tohoku-Oki Earthquake of March 11, 2011* http://www.air-worldwide.com/Publications/White-Papers/Understanding-Earthquake-Risk-in-Japan-Following-the-Tohoku-Oki-Earthquake-of-March-11,-2011/

Rota, M., Penna, A.; Magenes, G., (2010). *A methodology for deriving analytical fragility curves for masonry buildings based on stochastic nonlinear analyses*. Eng. Struct. 2010, 32, 1312–1323.

Kircher, C. A., Nassar, A. A. Kustu, O., Holmes, W. T., (1997). *Development of Building Damage Functions for Earthquake Loss Estimation*. Earthq. Spectra 1997, 13, 663–682.

Ramamoorthy, S. K., Gardoni, P., Bracci, J. M., (2006). *Probabilistic demand models and fragility curves for reinforced concrete frames*. J. Struct. Eng. 2006, 132, 1563–1572.

Hwang, H. H., Huo, J.R., (1994). *Generation of hazard-consistent fragility curves*. Soil Dyn. Earthq. Eng. 1994, 13, 345–354.

Goodno, B., (2009). *Improving the Seismic Performance of Existing Buildings and Other Structures.* American Society of Civil Engineers, 2009.

Wilson, R. C., (1993). *Relation of Arias intensity to magnitude and distance in California*. US Department of the Interior, US Geological Survey, 1993.

McCann, M. W., (1993). *RMS acceleration and duration of strong ground motion*. Stanford Univ., 1980.

Campbell, K. W., Bozorgnia, Y., (2010). *A Ground Motion Prediction Equation for the Horizontal Component of Cumulative Absolute Velocity (CAV) Based on the PEER-NGA Strong Motion Database*. Earthq. Spectra 2010, 26, 635–650.

Hancock, J., (2006). *The influence of duration and the selection and scaling of accelerograms in engineering design and assessment,* Imperial College London (University of London), 2006.

D'Errico, J., (2009). *SLM - Shape Language Modeling*. File Exchange - MATLAB Central http://www.mathworks.com/matlabcentral/fileexchange/24443-slm-shape-language-modeling

Willmott, C. J., Matsuura, K., (2005). *Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance*. Clim. Res. 2005, 30, 79–82.

Travasarou, T., Bray, J. D., Abrahamson, N., (2003). *A. Empirical attenuation relationship for Arias Intensity*. Earthq. Eng. Struct. Dyn. 2003, 32, 1133–1155.