Proceedings of the 9th U.S. National and 10th Canadian Conference on Earthquake Engineering
Compte Rendu de la 9ième Conférence Nationale Américaine et
10ième Conférence Canadienne de Génie Parasismique
July 25-29, 2010, Toronto, Ontario, Canada • Paper No 219

ASSESSING THE EPISTEMIC UNCERTAINTY OF GROUND-MOTION PREDICTIONS

J. Douglas[1]

ABSTRACT

Due to the limited observational datasets that are available for the derivation of ground-motion prediction equations (GMPEs) for the estimation of ground shaking from future earthquakes there is always epistemic uncertainty in the estimated median ground motion. Because of the increasing quality and quantity of strong-motion datasets it would be expected that the epistemic uncertainty in ground-motion prediction (related to lack of knowledge and data) is decreasing. In this study the predicted median ground motions from over 200 GMPEs for two scenarios (M_w 6 at r_{jb}=20km and M_w 7.5 at r_{jb}=10km) are plotted against date of publication to examine whether the scatter in the predictions (a measure of epistemic uncertainty) is decreasing with time for two regions. In addition, these predictions are compared to the median ground motion (and its confidence limits) for these scenarios based on averaging observations from a large strong-motion dataset at different dates. The epistemic uncertainty in predicting the median ground motions for a given scenario using an empirical GMPE are related to the confidence limits of the regression results, which is in turn dependent on the quantity and distribution of the underlying dataset. Confidence limits for three generations of two series of GMPEs are derived and examined for evidence that these confidence limits are narrowing.

It is found that there are still considerable differences in predicted ground motions from the various GMPEs and that estimates do not seem to be converging over time, although for western North America (WNA) predictions for moderate earthquakes have show a high level of consistency since the 1980s. A good match is observed between the predictions from GMPEs and the median ground motions and their confidence limits based on observations from similar scenarios, although the confidence limits from the observations are tighter than the variation between predictions from GMPEs would suggest they should be. The confidence limits derived for the series of GMPEs from WNA are, as expected, getting narrower with time but even for moderate earthquakes there is still roughly a 20% uncertainty in the median predicted ground motions, which is consistent with the scatter in predictions from various recent GMPEs for WNA.

[1] Visiting Professor, Earthquake Engineering Research Centre, University of Iceland, 800 Selfoss, Iceland. On teaching leave from BRGM, Orléans, France.

Introduction

Ground-motion prediction equations (GMPEs) have been developed since the 1960s and to date over 250 models for peak ground acceleration (PGA) and over 150 models for linear elastic response spectral ordinates have been published (Douglas, 2008). During this same period the quality and quantity of strong-motion data available to constrain these models has increased dramatically, particularly in the past two decades with the occurrence of various moderate and large earthquakes in well-instrumented areas (e.g. Loma Prieta, 1989; Northridge, 1994; Kocaeli, 1999; Düzce, 1999; Chi-Chi, 1999; Parkfield, 2004). These data have significantly improved the understanding of strong ground motion. Epistemic uncertainty in ground-motion prediction is related to a lack of data and knowledge and can be large. Within probabilistic seismic hazard assessments it has become standard practice to account for epistemic uncertainty by using logic trees, modeling the degrees of belief in various inputs (e.g. Bommer & Scherbaum, 2008). This uncertainty can have a large impact on the mean probabilities of exceedence. Thanks to the collection and analysis of considerable data from recent earthquakes the epistemic uncertainty in ground-motion prediction should have been coming down since the 1960s. This article investigates whether this is true using a variety of techniques. It should be noted that the procedures tested here probably only lead to lower bound estimates of epistemic uncertainty – true assessments would require more sophisticated methods and expert judgment.

In the next section, median predicted ground motions from hundreds of GMPEs are plotted against publication date to see whether these predictions are tending to a single value, which in the absence of epistemic uncertainty they should. In addition, on these plots are added the median ground motion (and its confidence limits) for the considered scenarios based on strong-motion records from a large databank. Since the results of hundreds of studies are being considered here, the analysis undertaken here is an example of 'meta-analysis', a common approach in, for example, medical research but not common in engineering seismology. Following this meta-analysis, confidence limits on the median predicted ground motions for three generations of two series of commonly-used GMPEs (one for western North America and one for Europe, the Mediterranean and the Middle East) are assessed to see how they are affected by the accumulation of new data and knowledge.

The history of ground-motion predictions

The empirical ground-motion models up to the end of 2008 summarized in the reports of Douglas (2004, 2006, 2008) have been programmed. Some authors do not report the coefficients of their models or the original reference of the model could not be retrieved in time for the submission of this article and, therefore, a few of the listed GMPEs were not programmed. In total 254 empirical GMPEs are considered here. In addition, 21 models (generally for stable continental regions, SCRs) based on ground-motion simulations, often the stochastic method, were also included in this analysis (some of these models provide multiple sets of coefficients to account for epistemic uncertainty, which have been considered in the analysis for SCRs). When sufficient information exists the differences in independent variables (e.g. magnitudes, style-of-faulting, horizontal component definition and source-to-site distance) between models have been minimized using the methods and equations discussed in Bommer et al. (2005) and Beyer &

Bommer (2006). Many (particularly early) studies do not provide sufficient information on, for example, magnitude scales used or definition of horizontal component so these adjustments could not be made. Therefore, the predictions shown for some models could be in error by up to 20%. For this article, models have been adjusted to: moment magnitude (M_w), distance to the surface projection of the rupture (Joyner-Boore distance) (r_{jb}), vertical-dipping strike-slip faulting and the geometric mean of the two horizontal components. Due to lack of space, only two strong-motion intensity parameters have been considered: PGA and (pseudo)-spectral acceleration (SA) at 1s natural period and 5% damping. Since some datasets (e.g. Joyner & Boore, 1981) have become standards for the testing of new regression techniques this could lead to an apparent underestimation of the true uncertainty in ground-motion predictions in the following graphs. For easy comparison, in following figures the maxima of the y-axes are always 25 times the minima.

Figure 1 presents the predicted PGAs and SAs at 1s from all programmed GMPEs against publication date for a M_w 6 strike-slip earthquake recorded at r_{jb}=20km on a stiff soil/soft rock site (V_{s30}=620m/s). This scenario was considered since it is roughly the best-represented scenario in global strong-motion databases and hence if predictions from GMPEs are becoming more similar then it should be noticeable for such a scenario. A few predicted ground motions are off the graphs shown on Figure 1 often due to extrapolating a model far outside its range of applicability. Figure 1 shows that the dispersion in predicted earthquake ground motions from different GMPEs is large (the ratio between the smallest and largest predictions is greater than ten) and that this scatter is not obviously reducing with time (even when considering only models passing basic quality control criteria). Assuming that predictions from GMPEs from all regions will tend to a single value over time makes the assumption that regional dependence of ground motions (e.g. Douglas, 2007) is minimal. However, this assumption is the focus of strong debate in the literature (e.g. see references in Douglas, 2007). In addition, Figure 1 plots many models that do not pass basic quality assurance criteria concerning peer-review, publication of basic information on the underlying dataset, independent and dependent parameters and extrapolation far outside the range of applicability of the model (e.g. Cotton et al., 2006). These issues complicate the use of Figure 1 when looking for a reduction in epistemic uncertainty. Therefore, it was decided to conduct the analysis for two smaller geographical regions for which many models exist: western and eastern North America, WNA and ENA, and also to indicate those models that were published in peer-reviewed international journals and give basic details of the datasets used for their derivation (criteria 2 and 3 of Cotton et al., 2006), which act as basic quality control.

Epistemic uncertainty in ground-motion prediction is linked to the quality and quantity of available strong-motion data. Therefore, given an abundant data set, predictions for a considered scenario should match the average ground motion observed for that scenario. Hence, in addition to plotting the predictions from various GMPEs the median ground motion for the considered scenario obtained from a large (over 13 000 records from over 2 500 events) strong-motion database (the strong-motion archive of Imperial College London from 2004 with the addition of many more accelerograms) are shown up to a given date. The median should track the predictions and the variability in the median should also show a reduction. Note that the variability in the median is not the same as the aleatory variability in an individual estimated ground motion (modelled by the standard deviation, σ) since given enough data the median can be well predicted but the variability in an individual estimate will remain high due to the complexity of earthquake ground motion generation and propagation (e.g. Douglas, 2007). The variabilities of the medians are computed here by dividing the standard deviation by √n, where n is the number of records used to compute the standard deviation. When constructing the strong-motion database employed here particular attention was given to the collection of data from shallow crustal earthquakes in active tectonic regions. Therefore, medians and variabilities are only plotted on the graphs for WNA and not on the graphs for ENA (an SCR). The medians and their variabilities were computed by considering the available data within 0.5-Mw units and 10km of the scenario of interest and excluding a consideration of local site conditions and style of faulting. Tests were made using narrower bins and considering only records from stiff soil sites and similar results were obtained.

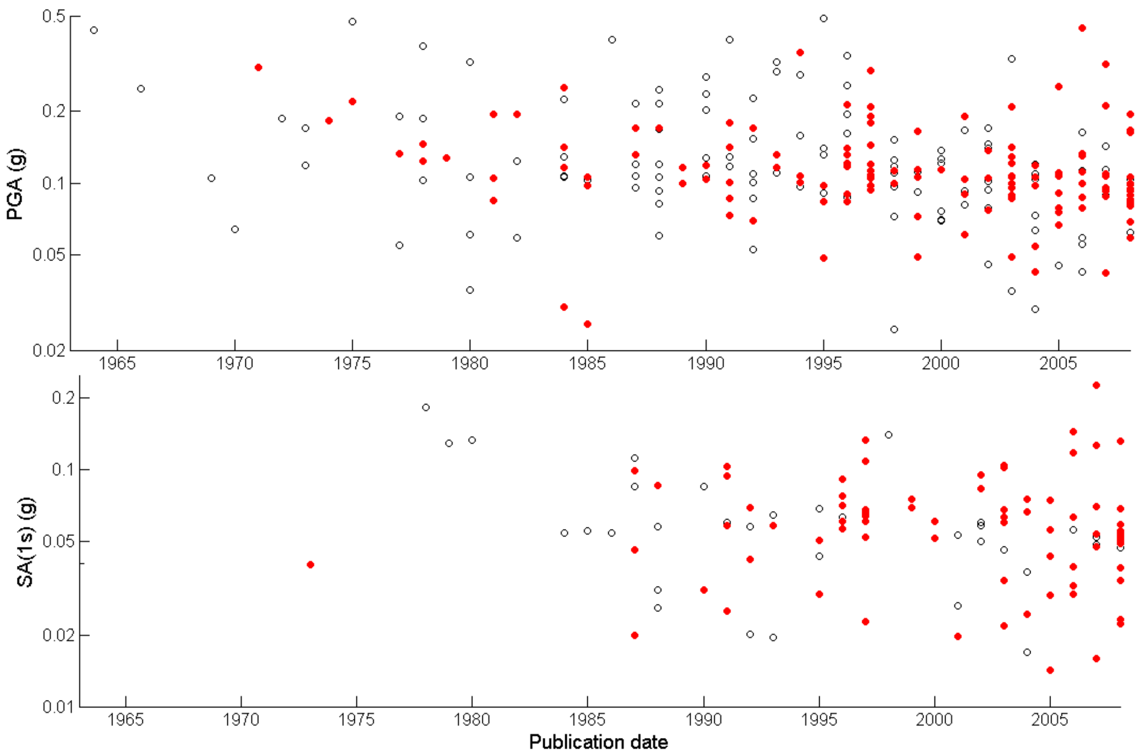

Figure 1. Predicted PGA and SA at 1s against publication date for over 250 models published in the literature. Filled red circles indicate models published in peer-reviewed international journals and for which basic information on the used dataset is available.

Figure 2 shows the results of the analysis for WNA for the scenario: M_w 6 at r_{jb}=20km. This figure does not clearly show a decrease in the scatter of predicted ground motions after about 1980 (the narrowing of the limits at that time could be related to the occurrence of the Imperial Valley earthquake in 1979, which significantly increased the amount of near-source strong-motion data available). Scatter in predictions for this scenario in models for WNA is small for PGA (the highest and lowest predictions are only about 25% different) but larger for SA at 1s (predictions differ by up to a factor two), probably related to the variety of techniques used to model site effects in the considered GMPEs. The median ground motions and their confidence limits obtained from averaging observations match the predictions from the GMPEs quite well although they are slightly higher. The confidence limits from the observations are narrower than the scatter in the predictions from the GMPEs, which suggests that the variation in predictions from GMPEs could be slightly over-estimating the actual uncertainty in the prediction of the median ground motion for this scenario.

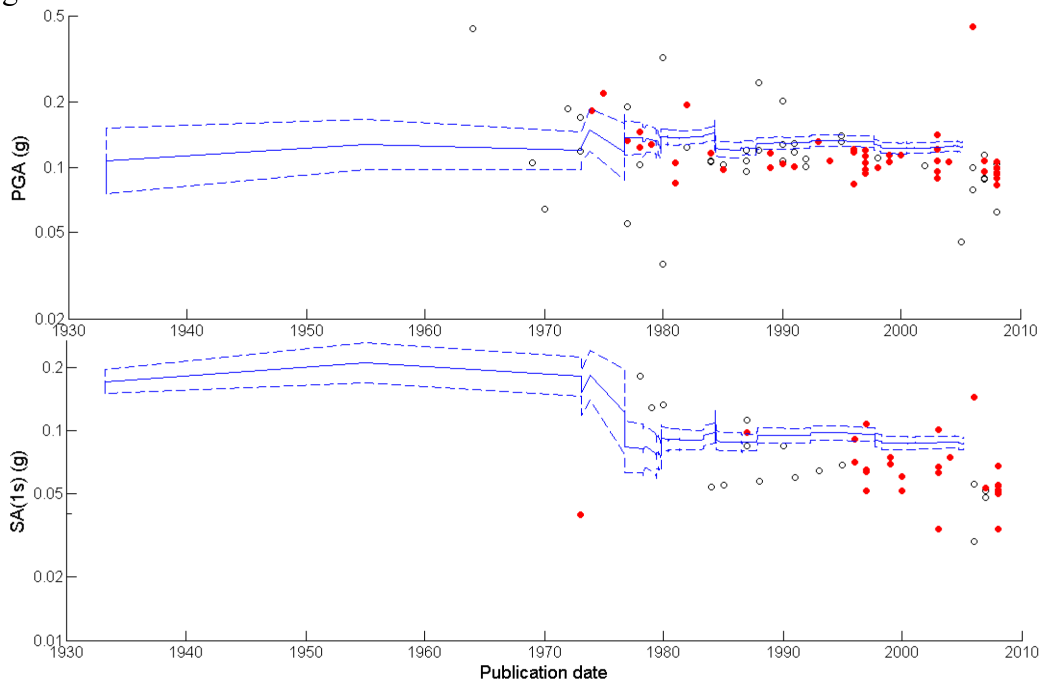

Figure 2. Like Figure 1 but only for WNA models. Also indicated are the median ground motions (solid blue line) and their 16th and 84th confidence limits (dashed blue line) based on averaging of records up until that date (see text for details). Up until the end of 2005, 253 records were used to compute these averages.

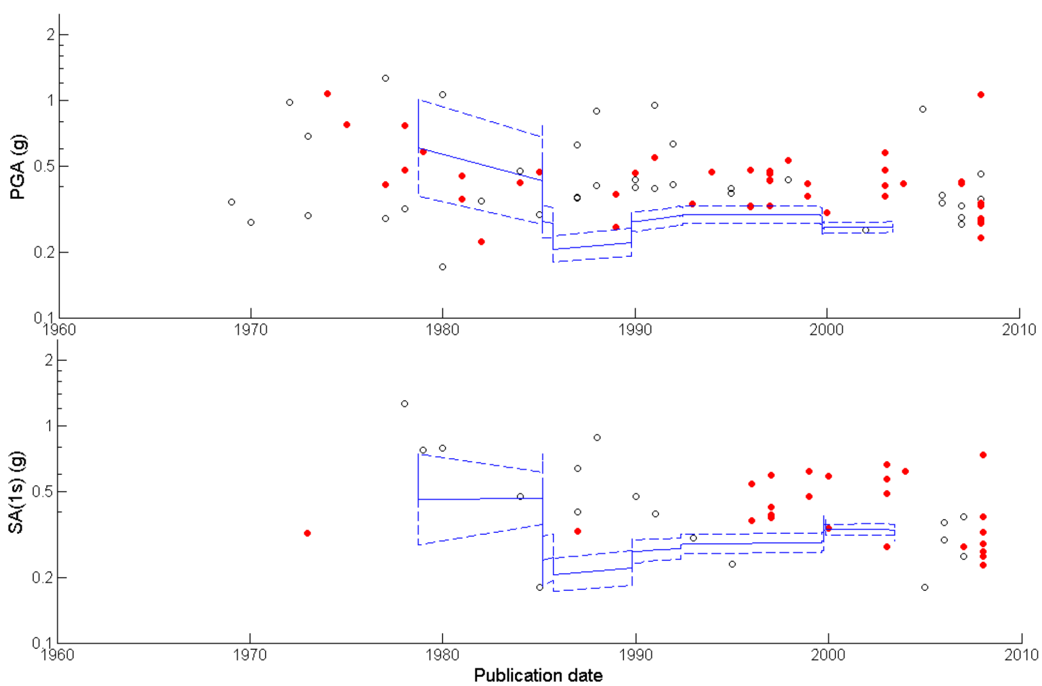

Figure 3. Like Figure 2 but for M_w 7.5 at r_{jb}=10km. Up until the end of 2003, 129 records were used to compute these averages.

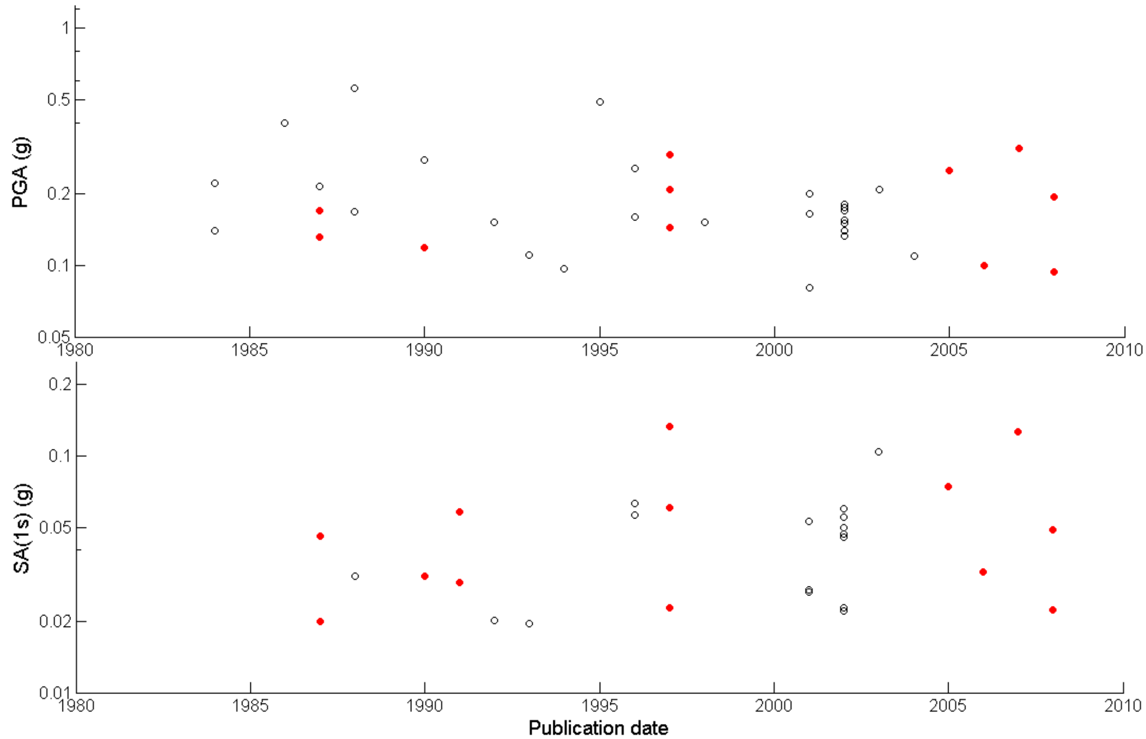

Figure 4. Like Figure 1 but only for SCR models and for very hard rock (V_{s30}=2.8km/s).

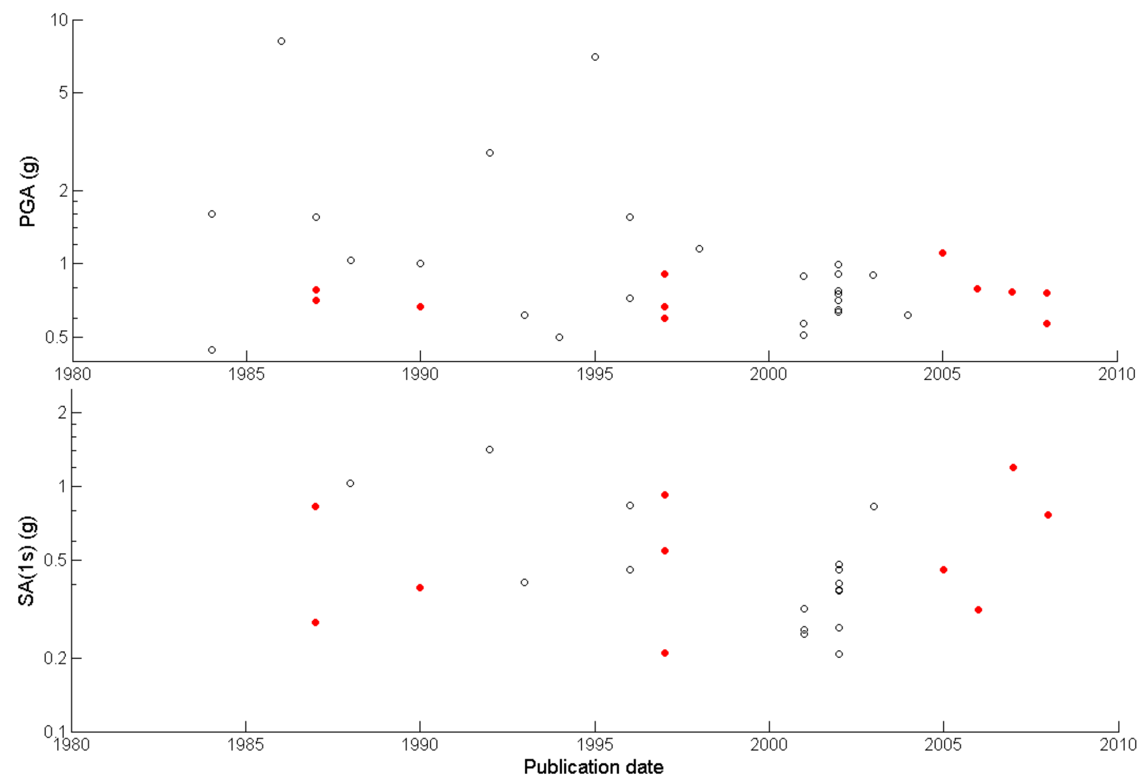

Figure 5. Like Figure 4 but for M_w 7.5 at r_{jb}=10km.

Figure 3 repeats the exercise for the scenario M_w 7.5 at r_{jb}=10km, for which observational strong-motion databases are still limited. As expected the scatter in predictions for this scenario shows a slight decrease with time, particularly in the past decade when more data to constrain predictions for larger magnitudes became available. In addition, the predictions from the GMPEs show a greater scatter for this scenario than for M_w 6 at r_{jb}=20km since there are still few records to help modellers constrain their GMPEs in the near-source region of large magnitudes. The median ground motions based on the averaging of observations match the predictions reasonably well although they are below the predictions from the GMPEs. Like for the previous scenario the confidence limits on the medians from the averaging are narrower than the scatter in the predictions from the GMPEs. This is partly due to multiple records from the same earthquake being used to compute these averages, which could lead to an underestimation of the true variability in the medians.

Figure 4 and 5 show the results of a similar analysis for stable continental regions (SCRs, mainly ENA). Note that unlike the previous figures, the vast majority of predictions plotted come from simulation-based GMPEs rather than empirical GMPEs due to the lack of strong-motion data from this region. This means that the variation amongst predictions is more under the control of the GMPE developers since they define the parameters and ranges used to develop the simulations that are a basis of these relations. Both these figures show that there has not been much, if any, decrease in the variation amongst predictions from the GMPEs for SCRs over the past 25 years. In addition, the difference between predictions for SCRs is higher than for WNA. These findings are in line with expectations since there is still sparse strong-motion data from SCRs from earthquakes with M_w>6 that can be used to constrain predictions. It is surprising that the scatter in PGA predictions for the M_w 7.5-r_{jb}=10km scenario is not greater than the scatter for the smaller event at greater distances since there are no records from SCRs available for such a scenario. This strongly suggests that the variation in predictions from GMPEs for SCRs is underestimating the true epistemic uncertainty in the prediction of ground motions for large events in such regions.

Confidence limits on predictions

As another measure of epistemic uncertainty in empirical ground-motion predictions Douglas (2007) discusses the use of the 5 and 95% confidence limits of GMPEs. These limits show how robustly the median ground motion is predicted. Since confidence limits of GMPEs have never been published to our knowledge (except for seven simple GMPEs in Douglas, 2007), in this study the datasets used to derive some GMPEs have been employed to compute the confidence limits on predictions. To investigate the reduction in epistemic uncertainty with time, two series of widely-used GMPEs have been considered: Joyner & Boore (1981), Boore et al. (1993) and Boore & Atkinson (2008) for WNA; and Ambraseys & Bommer (1991), Ambraseys et al. (1996) and Ambraseys et al. (2005) for Europe, the Mediterranean and the Middle East (EMME). Only PGA has been considered here. The same functional forms adopted by the authors of these studies have been used except for the model of Boore & Atkinson (2008), which uses a complex functional form. For this dataset the functional form of Akkar & Bommer (2007) is used instead. Standard nonlinear least-squares regression is used since this allows confidence limits to be easily approximated through the covariance matrix. Note that the purpose of this analysis is not to derive GMPEs for seismic hazard analysis but to roughly measure the improvement in empirical ground-motion predictions over the past couple of decades. To more easily examine the width of the confidence limits of the various models the ratios between the 95% and 50% confidence limits of the median PGA are plotted for three magnitudes (M_w 4, 6 and 8) against Joyner-Boore distance (r_{jb}) for the two sequences in Figure 6.

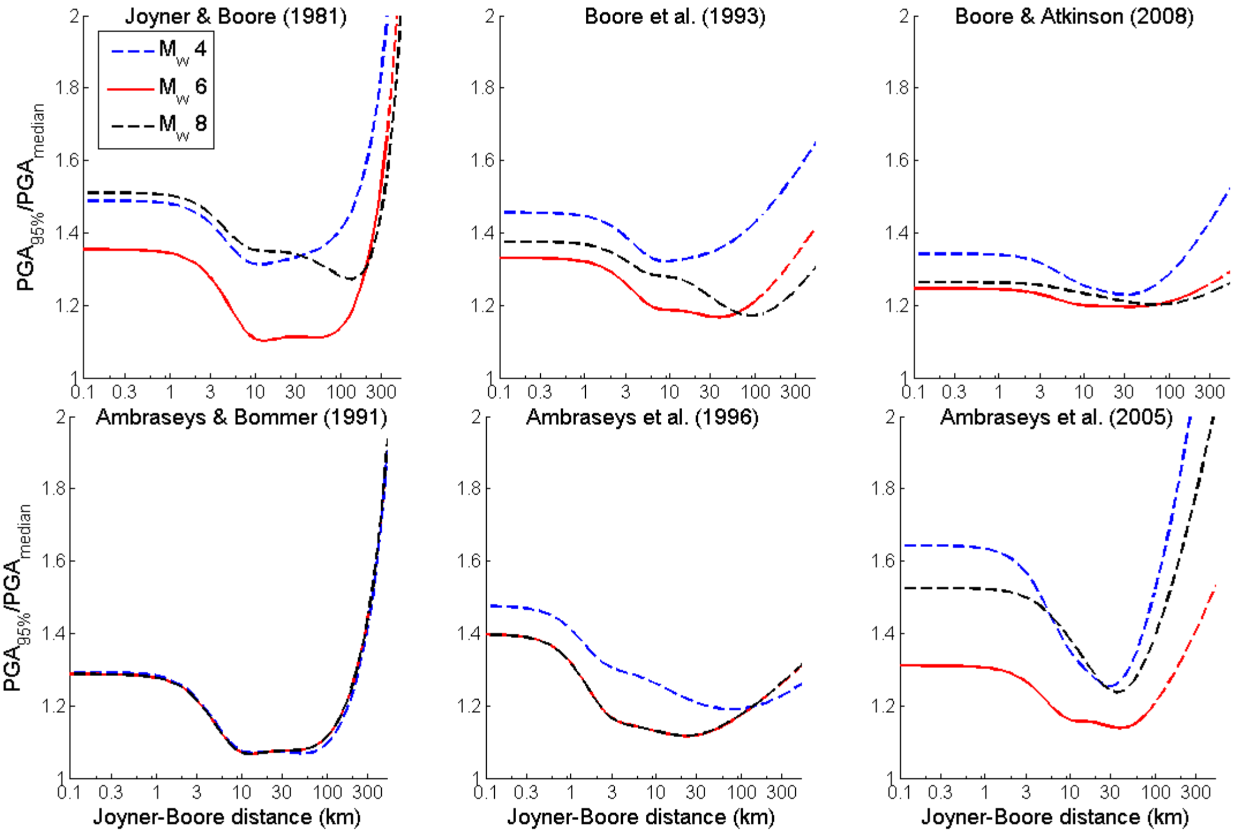

Figure 6. Ratios between the 95% confidence limits and the 50% confidence limits of the median rock PGA from strike-slip earthquakes for GMPEs re-derived using various datasets for WNA (top graphs) and for EMME (bottom graphs). For the datasets of Ambraseys & Bommer (1991) and Ambraseys et al. (1996), which use M_s, the GMPEs have been derived using this scale and M_w has been converted to M_s to draw these graphs. Dashed lines signify where the models have been extrapolated outside their strict ranges of applicability.

The graphs for the sequence of GMPEs for WNA show the expected behaviour with a narrowing of the confidence limits (reducing ratio between the 95 and 50% confidence limits) between successive models (from 1981, 1993 and 2008). Note, however, that these ratios are still high (1.2-1.5) and consequently there is still uncertainty in the prediction of the median PGA, even for moderate events far from the source. The graphs for the GMPEs from EMME suggest that the epistemic uncertainty in prediction of the median PGA is *increasing* with time (from 1991, 1996 and 2005) since the ratios are getting larger. Unlike for the GMPEs from WNA for which the quantity of data used has been steadily increasing (182, 271 and 1574) the quantity of points used by Ambraseys & Bommer (1991) is only slightly less than that used by Ambraseys et al. (2005) (529 compared to 595 points). In addition, Ambraseys et al. (2005) use a more complex model with ten free coefficients rather than only four for the model of Ambraseys & Bommer

(1991). The data used by Ambraseys et al. (2005) seems to have been insufficient to enable all these ten coefficients to be robustly estimated. However, Ambraseys et al. (2005) had more records available from M_w>7 events than did Ambraseys & Bommer (1991) (45 compared to 28) and, therefore, it seems that these computed confidence limits are not providing the entire picture concerning the reliability of the derived GMPEs. This warrants further investigation.

Conclusions

In this article the expected reduction in epistemic uncertainty in earthquake ground-motion prediction thanks to the accumulation of data and knowledge has been indirectly assessed using two methods: one based on the dispersion in predicted ground motions from published GMPEs and the other based on the confidence limits on predicted median PGAs from three generations of two series of GMPEs. This analysis has shown that although epistemic uncertainty seems to be reducing slightly for moderate events, there is still large uncertainty in the estimation of ground motions even for scenarios that are well represented in strong-motion databanks (e.g. M_w 6 at r_{jb}=20km). For areas such as WNA the epistemic uncertainty in the median PGA seems to be between 20-50% and for longer periods this uncertainty could be higher (perhaps up to 100%). For areas with limited datasets, e.g. SCRs, the uncertainty in the median ground motions appears to be higher (factors of three between predictions from different models are not uncommon). The results presented in this article clearly demonstrate that even for the best-investigated regions (e.g. WNA) that epistemic uncertainty in ground-motion prediction is real and must be accounted for in seismic hazard assessments through the use of multiple models capturing the range of possible median ground motions and their associated aleatory variabilities.

Acknowledgments

I thank: Landsvirkjun and the University of Iceland for funding my one-year visiting professorship at the Earthquake Engineering Research Centre, David Boore for sending me the Boore & Atkinson (2008) PGA dataset and Julian Bommer for sending me the dataset of Ambraseys & Bommer (1991) and for his detailed comments on a previous draft of this paper.

References

Akkar, S. and J. J. Bommer, 2007. Prediction of elastic displacement response spectra in Europe and the Middle East. *Earthquake Engineering and Structural Dynamics*, 36(10), 1275-1301.

Ambraseys, N. N. and J. J. Bommer, 1991. The attenuation of ground accelerations in Europe. *Earthquake Engineering and Structural Dynamics*, 20(12), 1179–1202.

Ambraseys, N. N., K. A. Simpson, and J. J. Bommer, 1996. Prediction of horizontal response spectra in Europe. *Earthquake Engineering and Structural Dynamics*, 25(4), 371–400.

Ambraseys, N. N., J. Douglas, S. K. Sarma and P. M. Smit, 2005. Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: Horizontal peak ground acceleration and spectral acceleration, *Bulletin of Earthquake Engineering*, 3(1), 1-53.

Beyer, K., and J. J. Bommer, 2006. Relationships between median values and between aleatory variabilities for different definitions of the horizontal component of motion. *Bulletin of the Seismological Society of America*, 96(4A), 1512–1522. doi: 10.1785/0120050210.

Bommer, J J. and F. Scherbaum, 2008. The use and misuse of logic trees in probabilistic seismic hazard analysis, *Earthquake Spectra*, 24, 997-1009.

Bommer, J. J., F. Scherbaum, H. Bungum, F. Cotton, F. Sabetta, and N. A. Abrahamson, 2005. On the use of logic-trees for ground-motion prediction equations in seismic-hazard analysis, *Bulletin of the Seismological Society of America,* 95(2), 377–389.

Boore, D. M. and G. M. Atkinson, 2008. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s. *Earthquake Spectra*, 24(1), 99-138.

Boore, D.M., W.B. Joyner, and T.E. Fumal, 1993. Estimation of response spectra and peak accelerations from western North American earthquakes: An Interim Report", Open-File Report 93-509, United States Geological Survey, Denver, U.S.A.

Cotton, F., F. Scherbaum, J. J. Bommer, and H. Bungum, 2006. Criteria for selecting and adjusting ground-motion models for specific target regions: Application to central Europe and rock sites. *Journal of Seismology*, 10(2), 137–156. DOI: 10.1007/s10950-005-9006-7.

Douglas, J., 2004, Ground motion estimation equations 1964–2003: Reissue of ESEE Report No. 01-1: 'A comprehensive worldwide summary of strong-motion attenuation relationships for peak ground acceleration and spectral ordinates (1969 to 2000)' with corrections and additions. Technical Report 04-001-SM, Department of Civil and Environmental Engineering; Imperial College of Science, Technology and Medicine; London; U.K.

Douglas, J., 2006, Errata of and additions to 'Ground motion estimation equations 1964–2003'. Intermediary report BRGM/RP-54603-FR, BRGM, Orléans, France.

Douglas, J., 2007. On the regional dependence of earthquake response spectra. *ISET Journal of Earthquake Technology*, 44(1), 71-99.

Douglas, J., 2008, Further errata of and additions to 'Ground motion estimation equations 1964-2003', Final report, BRGM/RP-56187-FR, BRGM, Orléans, France.

Joyner, W. B. and D. M. Boore, 1981. Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. *Bulletin of the Seismological Society of America*, 71(6), 2011–2038.