Chapter 3
Analysis of Regional Ground Motion Variations for Engineering Application

Jonathan P. Stewart

Abstract An important question for many ground motion hazard analyses is the degree to which ground motion prediction equations (GMPEs) developed for one region may have bias for a different region. A closely related problem is the applicability of multi-regional GMPEs to a particular region, even if that region contributed some fraction of the database. It is well known that ground motions show distinct characteristics for stable continental regions, subduction zones, and active tectonic regions with shallow crustal earthquakes. Here I consider variations among active regions with shallow crustal earthquakes. For such regions having sufficient data that meaningful comparisons are possible, I review four approaches for evaluating regional variations: (1) direct comparisons of medians from GMPEs; (2) analysis of variance; (3) overall goodness of fit metrics; and (4) verification of specific GMPE attributes relative to regional data. For engineering application, the objective of the comparison should be to evaluate whether median predictions show statistically similar trends with respect to magnitude-scaling, distance-scaling, and site effects across the range of magnitudes and distances controlling the seismic hazard, as well as consistent standard deviation terms.

3.1 Introduction

The attributes of earthquake ground motion intensity measures (IMs) are predicted using ground motion prediction equations (GMPEs), which describe the variation of the median and standard deviation of an IM with respect to source, path, and site parameters. A review of the vast literature on GMPEs is beyond the scope of this article (see Douglas, 2003, 2006 for reviews).

Earthquakes from subduction zones, shallow sources in active regions, and stable continental regions produce ground motions with distinct attributes, and

J.P. Stewart (✉)
University of California, Los Angeles, CA, USA
e-mail: jstewart@seas.ucla.edu

M. Garevski, A. Ansal (eds.), *Earthquake Engineering in Europe*, Geotechnical, Geological, and Earthquake Engineering 17, DOI 10.1007/978-90-481-9544-2_3,

hence different GMPEs are needed. However, within active regions having crustal earthquakes, there remains a lack of consensus on the manner by which to manage regionalization. As described by Douglas (2007), many investigators in Italy, Turkey, France, Spain and elsewhere assume a high degree of regionalization and develop GMPEs from small datasets derived from local regions (sometimes only a few thousand km^2 in size or arbitrarily defined by political boundaries). Due to the small size of these datasets, the GMPEs are not useful for prediction of the effects of relatively large earthquakes that form the basis for engineering design.

Another approach is to develop GMPEs from a large database derived from multiple regions. This was the approach of the Next Generation Attenuation (NGA) project that resulted in 2008 GMPEs by Abrahamson and Silva (AS), Boore and Atkinson (BA), Campbell and Bozorgnia (CB), and Chiou and Youngs (CY). The database used in that project (Chiou et al., 2008) included world-wide shallow crustal earthquakes from active regions including California, Taiwan, Japan, Turkey, Greece, Italy, New Zealand, and elsewhere.

The arguments for and against regionalization are well described elsewhere (e.g., Bommer, 2006; Douglas, 2007). The objective of this paper is to review and critique four methods for comparing ground motions from different regions. The emphasis here is on engineering application; hence, I assume that the intent of regionalization studies is the development or verification of GMPEs that can be used in probabilistic analyses that often find the hazard to be controlled by earthquakes of moderate to large magnitude at modest to close distance.

The four methods I will describe are comparisons of GMPE attributes; analysis of variance (Douglas, 2004a, b, 2007); overall goodness-of-fit (Scherbaum et al., 2004; Stafford et al., 2008); and verification of specific GMPE attributes relative to regional data (Scasserra et al., 2009a).

3.2 Methods for Analyzing Regional IM Variations

3.2.1 Comparison of GMPEs

Ground motion prediction equations (GMPEs) provide estimates of the median and log-normal standard deviation of ground motion. The attribute of GMPEs that is most often compared is the median and its variation with distance and magnitude for a reference site condition. As described by Douglas (2007), such comparisons are often problematic when one or more of the GMPEs is derived from small datasets because the standard error of the median (i.e., the uncertainty in the location of the median) is high and is not considered in the comparison. The standard deviation of the GMPEs is also critical for ground motion hazard analysis, but is seldom compared.

Figure 3.1 compares medians from GMPEs derived from large data sets (hence relatively small uncertainty in medians). Median peak horizontal ground accelerations (PGA) and 5%-damped pseudo spectral acceleration from two European

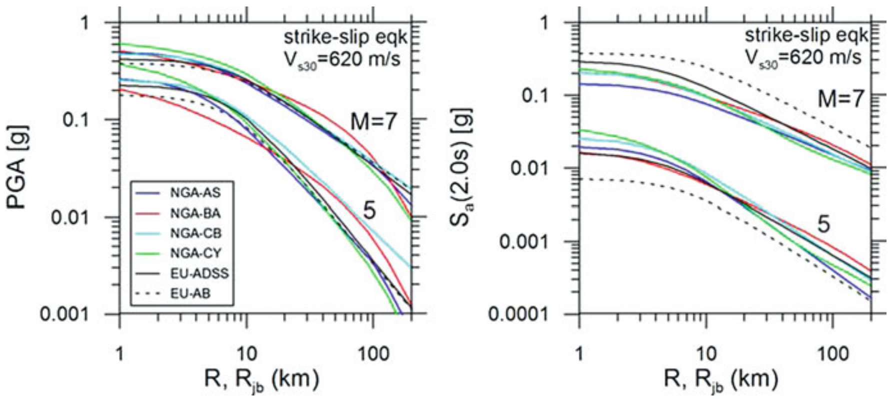

Fig. 3.1 Comparison of median predictions of PGA and 2.0 s pseudo spectral acceleration for strike slip earthquakes and soft rock site conditions from NGA and European GMPEs. AS = Abrahamson and Silva (2008); BA = Boore and Atkinson (2008); CB = Campbell and Bozorgnia (2008); CY = Chiou and Youngs (2008); ADSS = Ambraseys et al. (2005); AB = Akkar and Bommer (2007). Adapted from Scasserra et al. (2009a)

models (Akkar and Bommer, 2007; Ambraseys et al., 2005) are compared to those from NGA models. The European and NGA predicted medians generally compare well over the range of distances and magnitudes well constrained by the data. The bands of results for the two magnitudes generally show reasonably consistent vertical offsets from model-to-model (e.g., the difference between M7 and M5 PGA at $R_{jb} = 30$ km is reasonably consistent across models). This suggests generally consistent levels of magnitude scaling. The slopes of the median curves for a given magnitude are generally steeper for the European relations than the NGA relations for PGA, suggesting faster distance attenuation of this parameter.

In Fig. 3.2, I compare standard deviations from the AB European GMPE to a representative NGA GMPE (CY). Note that two standard deviation terms are shown. Standard deviation σ represents intra-event dispersion, which can be interpreted as the average level of dispersion from individual, well-recorded earthquakes. Term τ represents inter-event dispersion, or the standard deviation of event terms. Since event terms represent the average misfit of a GMPE to the data for a given event, τ represents event-to-event variability of the IM.

The comparison in Fig. 3.2 indicates relatively consistent τ terms and σ terms at large magnitude, but larger low-magnitude σ terms in the AB model relative to the CY model. This represents a potential example of regional variability, as both data sets apply for shallow crustal earthquakes in active regions.

3.2.2 Analysis of Variance

This approach was applied by Douglas (2004a) to compare ground motions for five local regions within Europe, Douglas (2004b) to compare ground motions from

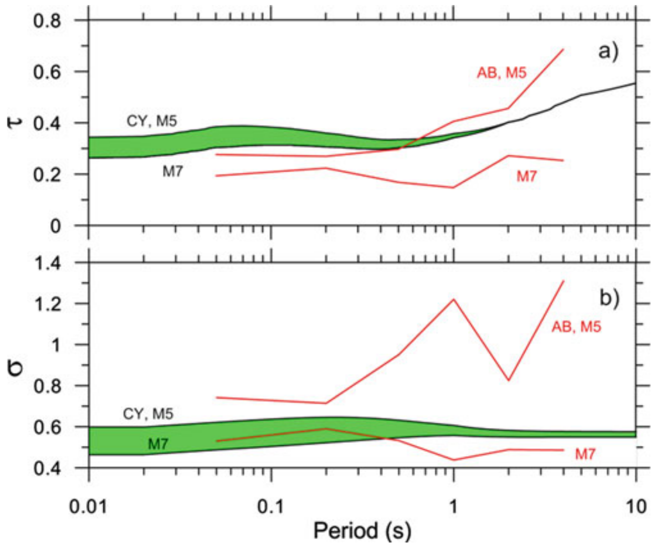

Fig. 3.2 Standard deviation terms σ and τ from European and NGA relation

Europe, New Zealand, and California, and Douglas (2007) for two local regions within Italy. The procedure involves calculating the mean (μ) and total variance (σ_T^2) of the log of data inside particular magnitude and distance bins (*M*-*R* bins) for two different regions (e.g., Europe and California) and combined data for those regions. The distance metric used by Douglas is the closest distance to the surface projection of the fault for $M > 6$ and epicentral distance otherwise. Individual data points are adjusted for a linear site factor before the calculation of mean and variance. These results are then used in two ways. First, for a given *M*-*R* bin and pair of regions, the variance of the combined data for both regions [termed $(\sigma_T^2)_{inter-region}$] is compared to the within-region variance [termed $(\sigma_T^2)_{intra-region}$] using statistical tests that evaluate whether the data sets are significantly distinct. If $(\sigma_T^2)_{inter-region} > (\sigma_T^2)_{intra-region}$ in a statistically significant way, there is likely to be significantly different medians between regions. The second use of the binned results is to plot medians for each *M*-*R* bin together for pairs of regions as shown for example in Fig. 3.3.

Using the above approach, Douglas (2004a) found similar variances for the various regions in Europe, indicating a lack of regional variations. Accordingly, Douglas (2004b) combined all of the European data into a single category for comparison to New Zealand and California data. The Europe-California comparisons indicate that approximately half of the *M*-*R* bins demonstrate significantly different inter- and intra-region variances. The distinction was towards larger ground motions in California (Douglas, 2004b). Figure 3.3 shows an example comparison of California and European medians from Douglas (2004b). The results indicate that the California and European medians for most *M*-*R* bins are similar at short

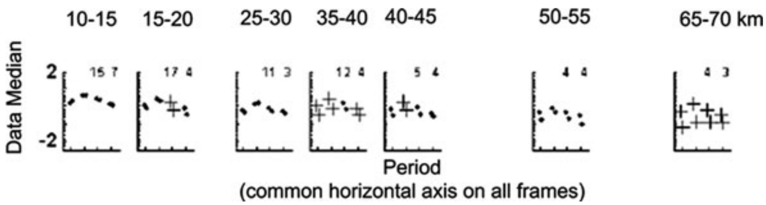

Fig. 3.3 Medians of data from $M = 6$–6.25 earthquakes within distance bins from California and Europe. Data from the two regions are shown side-by-side (California on left, Europe on right) with dots when the differences are not statistically significant and with crosses when significant at the 95% confidence level. Frames not shown for poorly populated bins. Modified from Douglas (2004b)

distance (<20 km), whereas California amplitudes are larger at larger distances (>30 km). Thus, Douglas' (2004b) finding of larger California ground motions could be expressed as more rapid distance attenuation in Europe.

3.2.3 Overall Goodness of Fit of GMPE to Data

This approach, developed by Scherbaum et al. (2004), provides an evaluation of overall goodness-of-fit of a GMPE to a dataset. A normalized residual is calculated for recording j from event i in a dataset as:

$$Z_{T,ij} = \frac{\ln\left(IM_{\mathrm{obs},ij}\right) - \ln\left(IM_{\mathrm{mod},ij}\right)}{\sigma_T} \tag{1}$$

where ln $(IM_{\mathrm{obs},ij})$ represents the IM value from the record, ln $(IM_{\mathrm{mod},ij})$ represents the median model prediction for the same magnitude, site-source distance, and site conditions of the record, and σ_T represents the total standard deviation of the model ($\sigma_T^2 = \sigma^2 + \tau^2$). If the data is unbiased with respect to the model and has the same dispersion, the normalized residuals (Z_T) should have zero mean and standard deviation of one – i.e., the properties of the standard normal variate. Accordingly, in simple terms, the procedure of Scherbaum et al. (2004) consists of comparing the actual Z_T distribution to that of the standard normal variate. Note that this procedure tests both misfit of the median and standard deviation.

Figure 3.4 shows an example application of this approach by Stafford et al. (2008), who also extended the method to consider both inter- and intra-event variability. They compared European data to the BA NGA relation and several European GMPEs. The specific example shown in Fig. 3.4 is intra-event normalized residuals for PGA relative to the BA relation. The BA relation was shown to match the median of the European data nearly as well as European GMPEs. The BA standard deviation, however, is lower than implied by the European data, resulting in the misfit of the histogram relative to the standard normal variate shown in Fig. 3.4.

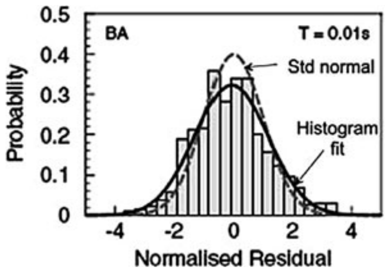

Fig. 3.4 Histogram of intra-event residuals of European data for PGA relative to BA GMPE. Note lack of bias indicated by median near zero. The histogram fit has a standard deviation larger than unity, indicating larger σ in the dataset than in the model from Stafford et al. (2008)

3.2.4 Verification of Specific GMPE Attributes Relative to Regional Data

This approach involves comparing regional data to GMPEs developed for a different region or for a diverse set of regions (e.g., shallow crustal earthquakes in active regions world-wide, as in NGA). As noted by Douglas (2007), this approach can provide misleading results if the available data lies outside the magnitude and distance range for which the GMPE is valid. However, when sufficient data is available to enable a valid comparison, it is possible to verify specific attributes of the GMPE relative to the data such as magnitude-scaling, distance-scaling, site effects, and standard deviation terms. This approach has been applied by Scasserra et al. (2009a) using the NGA GMPEs and data from pre-2009 Italian earthquakes.

To begin, residuals are evaluated between the data and a particular GMPE referred to with index k. Residuals are calculated as:

$$\left(R_{i,j}\right)_k = \ln\left(IM_{i,j}\right)_{\text{data}} - \ln\left(IM_{i,j}\right)_k \tag{2}$$

Index i refers to the earthquake event and index j refers to the recording within event i. Hence, $(R_{i,j})_k$ is the residual of data from recording j in event i as calculated using GMPE k. Term ln $(IM_{i,j})_{\text{data}}$ represents an IM computed from recording j. Term ln $(IM_{i,j})_k$ represents the median calculated using GMPE k in natural log units.

The analysis of residuals with respect to magnitude-, distance, and site-scaling requires that event-to-event variations be separated from variations of residuals within events. This is accomplished by performing a mixed effects regression (Abrahamson and Youngs, 1992) of residuals according to the following function:

$$\left(R_{i,j}\right)_k = c_k + (\eta_i)_k + \left(\varepsilon_{i,j}\right)_k \tag{3}$$

where c_k represents a mean offset (or bias) of the data relative to GMPE k, η_i represents the event term for event i (explained below), and $\varepsilon_{i,j}$ represents the intra-event residual for recording j in event i. Event term η_i represents approximately the mean offset of the data for event i from the predictions provided by the GMPE median (after adjusting for mean offset c_k, which is based on all events). Event terms provide

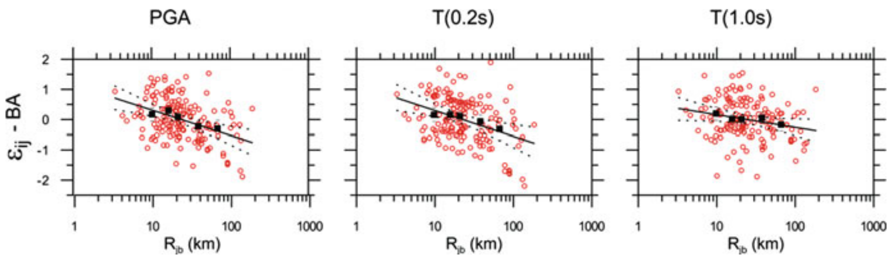

Fig. 3.5 Trend of intra-event residuals of Italian data for PGA relative to BA GMPE. Modified from Scasserra et al. (2009a)

a convenient mechanism for testing the ability of a GMPE to track the magnitude scaling of recordings in a dataset. Event terms are assumed to be normally distributed, and have zero mean and standard deviation $= \tau$ (in natural log units). Intra-event error ε is also assumed to be normally distributed with zero mean and standard deviation $= \sigma$.

Scasserra et al. (2009a) applied the above methodology to an Italian data set that had been carefully screened for record quality and for which the necessary metadata on site conditions, magnitudes, and distance parameters was available (Scasserra et al., 2009b). Careful and consistent screening of the data in that manner is necessary to obtain meaningful results from the analyses.

The analysis of Scasserra et al. (2009a) showed a general lack of significantly non-zero values of c, suggesting a lack of overall bias in the NGA GMPEs relative to the Italian data. As shown in Fig. 3.5, intra-event residuals (ε) demonstrated a negative trend with distance, indicating faster attenuation of the Italian data than in the NGA GMPEs for high-frequency spectral accelerations. That bias was not present for low-frequency spectral accelerations. After removing the distance attenuation bias from the GMPEs, no statistically significant bias of the event terms with respect to magnitude was found, indicating consistent magnitude-scaling between the NGA GMPEs and Italian data. As found by Stafford et al. (2008), intra-event standard deviation term σ from the Italian data was larger than in the NGA GMPEs, but inter-event standard deviation term τ was similar.

3.3 Discussion and Conclusions

The multi-regional database used to develop the NGA GMPEs is large (3,551 recordings from 173 earthquakes; subsets used for particular GMPEs). As noted by Stafford et al. (2008), because of the large size and high quality of the NGA database, certain effects are captured that could not be evaluated using only data from a single region. Examples include depth to top-of-rupture, magnitude- and/or site-dependent standard deviations, and nonlinear site response. The NGA data also provides the opportunity to constrain relatively complex functional forms for magnitude and distance scaling as compared to regional models.

Because of the relative sophistication of the NGA GMPEs, it is of interest to evaluate whether they can be applied to specific geographic regions. In this article, I have briefly described four procedures by which this evaluation can be performed. I assume for the sake of this discussion that a strong motion database is available for the region under consideration, and that this database has an appropriate level of processing, screening, and available meta-data.

For such situations, the goodness-of-fit approach of Scherbaum et al. (2004), later modified by Stafford et al. (2008), assesses GMPE performance in an overall sense – i.e., all aspects of the model (magnitude-scaling, distance-scaling, site effects) are evaluated together. If one or more of these model components is in error, that effect could be obscured through compensating errors in the analysis of normalized residuals. Accordingly, while the results of Stafford et al. (2008) are certainly promising with respect to the application of NGA relations in Europe, they do not specifically address whether individual components of the NGA models are adequate with respect to European data.

When the required data can be assembled, a more complete picture of regional variations emerges when specific attributes of a GMPE are tested against the regional dataset in the manner described by Scasserra et al. (2009a). These comparisons require selection of data that lies within the range of applicability of the GMPE. For the Italian dataset, this approach enabled the finding of faster distance attenuation and higher intra-event standard deviation described above. Those specific misfits, in turn, can be corrected within the multi-region GMPE for application in a local region. This allows advantages of the relatively sophisticated GMPEs to be leveraged in the local region without the introduction of obvious, first-order bias.

The above approaches rely on the availability of a GMPE that can be used for comparison to regional data. If such a GMPE is unavailable or judged to be problematic, an alternative approach is needed. The analysis of variance approach of Douglas (2004a, b, 2007) avoids reliance on GMPEs, focusing instead on comparisons of binned data from two regions. This approach provides valuable insights into fundamental features of datasets from two regions, but does not directly result in a usable GMPE, which is needed for hazard analyses.

All of the approaches described in this article have advantages and limitations. My goal here is not to recommend a single approach, but to highlight available methodologies and the types of insights that can be gained from them. Ultimately, the goal of regionalization studies should be to support the development and use of reliable GMPEs in engineering practice. Those GMPEs should be constrained to the maximum extent possible by recordings/events that span the range of magnitudes and distances controlling seismic hazard at the return periods of engineering interest. GMPEs based solely on small, local databases are unlikely to meet this standard for the foreseeable future.

References

Abrahamson NA, Silva WJ (2008) Summary of the Abrahamson and Silva NGA ground motion relations. Earthquake Spectra 24(S1):67–97

Abrahamson NA, Youngs RR (1992) A stable algorithm for regression analyses using the random effects model. Bull Seism Soc Am 82:505–510

Akkar S, Bommer JJ (2007) Prediction of elastic displacement response spectra in Europe and the Middle East. Earthquake Eng Struct Dyn 36:1275–1301

Ambraseys NN, Douglas J, Smit P, Sarma SK (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. Bull Earthquake Eng 3(1):1–53

Bommer JJ (2006) Empirical estimation of ground motion: advances and issues. In: Proceedings of the 3rd international symposium on the effects of surface geology on seismic motion, vol 1. Grenoble, France, pp 115–135

Boore DM, Atkinson GM (2008) Ground motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 and 10.0 s. Earthquake Spectra 24(S1):99–138

Campbell KW, Bozorgnia Y (2008) NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD, and 5%-damped linear elastic response spectra for periods ranging from 0.01 to 10 s. Earthquake Spectra 24(S1):139–171

Chiou BS-J, Darragh R, Dregor D, Silva WJ (2008) NGA project strong-motion database. Earthquake Spectra 24(S1):23–44

Chiou BS-J, Youngs RR (2008) An NGA model for the average horizontal component of peak ground motion and response spectra. Earthquake Spectra 24(S1):173–215

Douglas J (2003) Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectra ordinates. Earth Sci Rev 61:43–104

Douglas J (2004a) An investigation of analysis of variance as a tool for exploring regional differences in strong ground motions. J Seism 8:485–496

Douglas J (2004b) Use of analysis of variance for the investigation of regional dependence of strong ground motion. In: Proceedings of the 13th world conference on earthquake engineering, Vancouver, BC, Paper 29 (electronic file)

Douglas J (2006) Errata of and additions to "Ground motion estimation equations 1964–2003". Intermediary Report BRGM/RP-54603-FR, Bureau de recherches géologiques et minières

Douglas J (2007) On the regional dependence of earthquake response spectra. ISET J Earthquake Technol 44(1):71–99

Scasserra G, Stewart JP, Bazzurro P, Lanzo G, Mollaioli F (2009a) A comparison of NGA ground-motion prediction equations to Italian data. Bull Seism Soc Am 99(5):2961–2978

Scasserra G, Stewart JP, Kayen RE, Lanzo G (2009b) Database for earthquake strong motion studies in Italy. J Earthquake Eng 13(6):852–881

Scherbaum F, Cotton F, Smit P (2004) On the use of response spectral reference data for the selection and ranking of ground motion models for seismic hazard analysis in regions of moderate seismicity: the case of rock motion. Bull Seism Soc Am 94(6):2164–2185

Stafford PJ, Strasser FO, Bommer JJ (2008) An evaluation of the applicability of the NGA models to ground motion prediction in the Euro-Mediterranean region. Bull Earthquake Eng 6:149–177