Comparison of GOF_MO to Other Proposed GOF Measures

Anderson (2004) proposed a goodness-of-fit measure for the purpose of validation and verification of broadband synthetic seismograms. In the following, we compare our GOF measure to that proposed by Anderson.

Anderson's GOF measure includes 10 separate metric comparisons. These comparisons are generally applied to 10 individual band-passed frequency ranges, to provide a GOF for a given pair of time series and to quantify the contributions from different bandwidths. Anderson's GOF value for seven of these intensity metrics is calculated using the equation:

$$\exp(-RAT^2) \text{ , where } \quad RAT = \frac{(y-x)}{\min(x,y)} \quad . \qquad \text{(S8)}$$

This measure is used for the metrics PGA, PGV, PGD, Fourier spectrum, Arias Intensity, energy integral, and the response spectra. The remaining three intensity metrics included in Anderson's GOF measure are the cross-correlation, the Arias duration and the energy duration. The cross correlation GOF is 10 times the cross correlation if the value is positive, or else the GOF is set to zero (eq. S1). The energy and Arias duration GOF values are computed as

$$10 * \left(1 - \max\left(\left|x_i - y_i\right|\right)\right) \text{ ,} \qquad \text{(S9)}$$

where x and y are the normalized Arias integrals or the normalized energy integrals. This comparison tracks the maximum differences in the timing of the cumulative energy of the time series. Anderson's GOF measure is suited to evaluate both low-frequency and broadband seismograms.

To facilitate a comparison between our and Anderson's GOF measures, let y = α*x for the discrete scalar values x and y. If we assign y to be greater than or equal to x, the value of α must always be greater than or equal to 1. We can then write the argument (NR) of the erfc function in eq. 1 as:

$$\left[2*\frac{|x-y|}{x+y}\right]=\left[2*\frac{|x*(1-\alpha)|}{x*(1+\alpha)}\right]. \qquad \text{(S10)}$$

All values computed in eq. S10 are greater than or equal to zero, thus we can re-write this expression as:

$$\left[2*\frac{|x*(1-\alpha)|}{x*(1+\alpha)}\right]=\left[2*\frac{(\alpha-1)}{(\alpha+1)}\right] \qquad \text{(S11)}$$

Thus, our GOF measure can be expressed in terms of α as

$$\mathrm{erfc}\left(2*\frac{(\alpha-1)}{(\alpha+1)}\right). \qquad \text{(S12)}$$

Similarly, Anderson's GOF measure (eq. S8) can be expressed as:

$$\exp\left(-(\alpha-1)^2\right) . \qquad \text{(S13)}$$

These simple transformations allow us to analyze the GOF measures using a single independent variable.

For convenience, we have removed the scaling factors (100 for our and 10 for Anderson's GOF measures, respectively) in eqs. S12 and S13 to limit the range of each measure between the values of 0 and 1, with 1 being a perfect match. The two measures are compared graphically in Figure 1 of the journal article. Our GOF measure has very good resolution in the higher values (100 - 90) and a lower resolution in the lower values (<35). While the choice of the two GOF measures depends on the specific application, we expect that the high resolution for the smallest misfits using our GOF measure is generally desirable, while the resolution for comparisons of time series with larger misfits is of less significance.

The two GOF measures are similar in many ways. They are monotonically decreasing, and they cross at approximately the same value (~ 0.35) for $\alpha \approx 2.0$. This level of misfit may be an appropriate cut-off level in that values below this threshold may indicate unacceptable discrepancies between the time series (Anderson, 2004). We calculate many of the same metrics: PGA, PGV, PGD, Fourier Spectrum, response spectrum, cumulative energy and cross-correlation. However, the GOF estimates are computed differently for the cross-correlation and Fourier spectrum in the two methods. Our GOF calculation uses a smoothed Fourier spectrum (see eq. S6) in order to minimize undesirable effects from erratic spectral variation, while Anderson's GOF estimate uses an unsmoothed spectrum. Moreover, the cross-correlation and the Fourier spectrum for our GOF estimate are calculated using the format of the input time series (i.e. velocity), while Anderson specifically calculates the Fourier spectrum and the cross-correlation on the acceleration time series. For this study, we applied Anderson's cross-correlation to the velocity time series to enable a direct comparison with our GOF measure.

Anderson's GOF measure includes Arias and energy durations. These durations are based on discrepancies in the arrival time and magnitude of the Arias and energy integrals. In comparison, our GOF calculates the energy duration for the input time series based on the model given by Trifunac (1977) and Jarpe and Kasameyer (1996). Additionally, we calculate a GOF value for the ratios between the IE ratios at multiple periods.

Finally, Anderson's GOF measure is very flexible in the fact that the measure produces values for multiple pre-determined bandwidths. Such detailed frequency-specific measures may be useful for structural response, when sufficient information is available for the built environment. Such detailed measures are not included in the current version of our GOF, but would be a straight-forward extension. Our GOF generates a comprehensive measure based on a single user-defined bandwidth for simplicity. This allows the user to specify the bandwidth of interest providing a more specific GOF value for a given application. In addition, our method exports the IE and SA16 values along with their corresponding GOF values explicitly for each period within the user-defined bandwidth.

As a verification exercise within SCEC, Bielak et al. (2010) carried out a detailed comparison of 0-0.5Hz synthetics at 10 sites from three different numerical codes (two regular finite-difference meshes, SDSU, and URS, and one variable-mesh finite elements, CMU) for the ShakeOut scenario in southern California. The ShakeOut scenario represents a hypothetical M7.8 earthquake on the southern San Andreas fault, with the finite-fault source described in Graves et al. (2008). Bielak et al. used both the GOF measures by Anderson and Kristekova et al. (2006) to compare the three sets of LF synthetic generators. Here, we compare the GOF results for the three sets of synthetics used in Bielak et al.'s with Anderson's measure to those using the method by Kristekova et al. (see Table S1). The values from Anderson's method are higher than those derived from our measure on average for all involved metrics except FS and Xcor, as expected for time series with generally low misfits (see Figure 1), although some of the GOF values at the individual bandwidths show a minor degree of variation. These differences are likely caused by the different computations of duration and Fourier spectrum. For example, the normalized mean residuals between Anderson's and our GOF values on the overall comparison for the duration and Fourier spectrum metrics are -0.014 and 0.035, respectively, with standard deviations of 0.126 and 0.074 respectively (see Table S1, d) indicating that Anderson's method tends to generate higher values for the duration calculation while the our GOF calculation using the smoothed Fourier spectrum provides a better fit. However, both GOF measures agree on the general conclusions for the verification study by Bielak et al.

References

Anderson J. (2004). Quantitative measure of the goodness-of-fit of synthetic seismograms, Proceedings of the 13th World Conference on Earthquake Engineering, Vancouver, paper 243.

Bielak, J., R. Graves, K. Olsen, R. Taborda, L. Ramírez-Guzmán, S. Day, G. Ely, D. Roten, T. Jordan, P. Maechling, J. Urbanic, Y. Cui, G. Juve (2009). The ShakeOut earthquake scenario: verification of three simulation sets, *Geophysical Journal International*, **180**, 375-404, doi: 10.1111/j.1365-246X.2009.04417.x

Graves, R. and A. Pitarka (2004). Broadband time history simulation using a hybrid approach. *Thirteenth World Conference on Earthquake Engineering Conference Proceedings*, Vancouver, BC. Paper 1098.

Jarpe, S. P., and P. W. Kasameyer (1996). Validation of a procedure for calculating broadband strong-motion time histories with empirical green's functions, *Bull. Seis. Soc. Am*. **86**, 1116-1129.

Kristekova, M., J. Kristek, P. Moczo, and S.M. Day (2006). Misfit criteria for quantitative comparison of seismograms, *Bull. Seism. Soc. Am., Vol 96,* 1836-1850.

Trifunac, M.D., and B. Westermo (1977). A note on the correlation of frequency-dependent duration of strong earthquake ground motion with the Modified Mercalli Intensity and the geologic conditions at the recording stations, *Bull. Seis. Soc. Am*. **67**, 917-927.