EqHaz - An open-source probabilistic seismic hazard code based on the Monte Carlo simulation approach

Karen Assatourians and Gail M. Atkinson

For submission to Seism. Res. L.

July 2012

Introduction

We have developed a suite of open-source FORTRAN programs to perform probabilistic seismic hazard analysis (PSHA) by the Monte Carlo simulation method (Musson 1999, 2000; Hong and Goda, 2006). Our program suite, EqHaz, is designed to optimize the speed, flexibility and ease of application for typical PSHA problems in moderate-seismicity environments, such as eastern North America (ENA), for which detailed information on the source characteristics and wave propagation attributes of specific active fault sources are unavailable. The program can handle areal and fault sources with magnitude recurrence statistics as described by Gutenberg Richter (untruncated, truncated, or tapered), or a user-specified discretized cumulative or incremental magnitude-recurrence distribution. Ground-motion amplitudes are modeled as user-specified functions of magnitude and distance (in a table format). The ground-motion formulation is flexible so that future modifications may also allow direct use of a catalogue of simulated records in place of ground-motion prediction equations (GMPEs). Both epistemic and aleatory uncertainty in the key input parameters can be modeled. The user may treat these uncertainty sources as equivalent (in terms of our ability to predict future ground motions), enabling the treatment of ground motions realized over a long simulated catalogue as an extreme value statistical problem. Alternatively, the user may treat epistemic uncertainty separately using confidence fractiles on the input parameters, as has been traditional practice (e.g. McGuire, 2004).

Outputs of EqHaz include simulated earthquake catalogues generated from the user-specified seismicity parameters, ground-motion catalogues at a site, mean-hazard probability curves, and mean-hazard motions at specified return periods calculated for a grid of points (to generate hazard maps). Hazard de-aggregation occurs naturally, as the program tracks which simulated catalogue events cause exceedance of specified target amplitude levels. The outputs include a list of the magnitude-distance-sigma combinations that contribute the peak motions within specified duration windows (where sigma is the standard deviation of the GMPE).

The outputs of EqHaz have been validated against those of other industry-standard programs such as EZFRISK and FRISK88, for selected test problems. The EqHaz programs provide a flexible alternative to other available programs such as OpenSHA and EZFRISK that is suitable for many applications, particularly those for which areal sources dominate. The EqHaz program suite fills what we have perceived as a void in PSHA software, in that it is simple to use, runs quickly (in minutes) on a standard PC, is open-source and freely available, and provides the capability to analyze both aleatory and epistemic uncertainty for many typical PSHA problems in moderate-seismicity environments in a flexible way. It also enables research into a number of topics in PSHA, which are facilitated by the Monte Carlo simulation approach to performing the computations.

Basic Methodology

In the following description of the EqHaz methodology, we focus mainly on the case of a single set of input parameters, for ease of understanding. This is the classic 'best estimate' case, in which the hazard is described by a single set of source zones, each of which follows a prescribed magnitude recurrence relationship truncated at a maximum magnitude (M_x). Each source zone may be either an areal source (polygon) or a fault source (line segments, with a specified dip). Three GMPEs, with prescribed aleatory variability (sigma) and epistemic uncertainties, are used to characterize the expected amplitudes (e.g. such as the response spectral amplitude at a specific period) as a function of magnitude and distance, for some reference site condition. The problem is analyzed using three sequential programs: EqHaz1 simulates an earthquake catalogue using the source zone geometries and magnitude recurrence statistics; EqHaz2 calculates ground motions received at a site and produces a mean-hazard curve; and EqHaz3 compiles statistics on maximum amplitude values for a given set of return periods, along with information on motions that exceed the mean-hazard amplitude for the given return periods. Hazard fractiles are also part of the EqHaz3 outputs, for the case in which epistemic and aleatory uncertainties have been treated separately. The steps in these programs are described below. Although we focus on the ‘best estimate’ case for simplicity, we also mention how parameters are treated in the more general case in which there is modeled epistemic uncertainty in input parameters.

EqHaz1 - Catalogue generation

EqHaz1 generates a synthetic earthquake catalogue for given seismic source geometries and seismicity parameters (magnitude recurrence and hypocentral depth), for events greater than a minimum magnitude of interest. EqHaz1 generates a user-specified number of subcatalogues, each of which has a specified duration (e.g. 2000 subcatalogues of 2475 years each). The output is a simulated earthquake catalogue. Each record of the synthetic catalog contains the following fields: subcatalogue number, time (in decimal years), coordinates (latitude, longitude, depth) and magnitude of earthquake, source zone number, and the values of seismicity parameters (Gutenberg Richter parameters No and β, and maximum magnitude M_x) implemented for generating the record. Figure 1A shows the flowchart of the EqHaz1 program (with Figure 1B referring to EqHaz2).

As shown in Figure 1A, EqHaz1 reads the input file, then uses Monte Carlo simulation to generate the events of a subcatalogue, following the probability distributions given by those parameters. EqHaz1 sorts and writes the output, and then repeats the process for all subcatalogues. EqHaz1 generates the time, magnitude, and coordinates of each record as follows. For assigning the time of a record, EqHaz1 picks a time randomly between zero and the subcatalogue duration, thus assuming a stationary Poisson process. (Note: future enhancements will allow consideration of time-dependent distributions.) For assigning a magnitude, EqHaz1 works with M_x and the Gutenberg-Richter parameters N_0 and β for the source zone (either a single set of magnitude-recurrence parameters, or random draws from a range of possible sets for the entire subcatalogue, if epistemic uncertainty is being considered), using a Monte Carlo draw that follows the prescribed probability distribution. The cumulative magnitude-frequency distribution equation of the magnitude recurrence parameters is constructed and used to obtain a random realization of a magnitude value that follows this distribution. For assigning focal depth, EqHaz picks randomly a value from a specified depth distribution in the input file, for the entire subcatalogue (or for each record, if epistemic uncertainty is not considered). The depth distribution is given as a list of alternative depths and weights; the user may also apply a specified perturbation to randomize the depths.

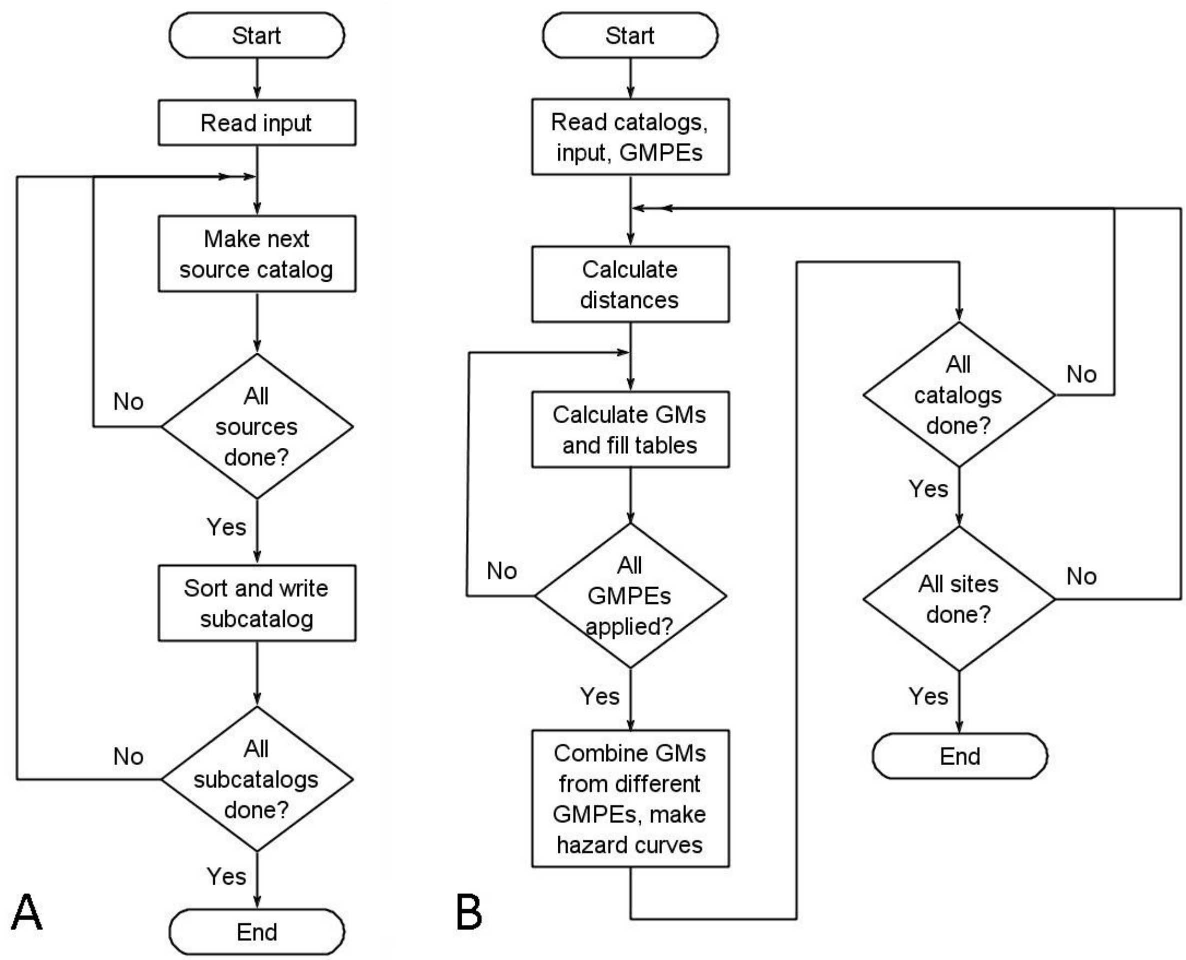

Figure 1. A) Flowchart of program EqHaz1; B) Flowchart of program EqHaz2; input to EqHaz2 is the output from EqHaz1.

To assign epicentral coordinates to the simulated event, the process is different for areal and fault sources. If the source is areal, a random point inside the polygon specifying the seismic source is selected. The derived epicenter may be further perturbed by a user-specified value to simulate "fuzzy" source boundaries. If the source is a fault, a random point along the fault strike is selected. The location of the epicenter is assumed to be on a line perpendicular to the fault strike, passing through this point; the event depth and fault dip determine the offset distance from the line of fault strike. (Note: EqHaz1 accepts a multi-segment fault source.)

EqHaz1 generates an arbitrarily-large number of records for each source for each subcatalogue (e.g. 100000). It discards all those events that are smaller than the specified minimum magnitude of the source. Then it repeats the same procedure for all other seismic sources and saves all results. Next it sorts all these records chronologically, and writes this subcatalogue to an output file. The program repeats this process for the specified number of subcatalogues.

Figure 2 shows an example of a generated 2475-year subcatalogue for an area of eastern Canada along the St. Lawrence River (Montreal-Ottawa region). The seismic source zone boundaries utilized for the catalogue generation are shown in the figure. They are modeled after those presented in Atkinson and Goda (2011), with some minor revisions. In essence, the model provides a "historical earthquake source model" that hosts earthquakes of **M**≤6.5 (where **M** is moment magnitude). It is overlain by a characteristic "Iapetan Rifted Margin (IRM)" source that hosts random large earthquakes of **M**6.5 to 7.5. The magnitude recurrence parameters are obtained from historical seismicity and paleoseismic information as described in Atkinson and Goda (2011). Details of the model are not provided here as they are not important for our purpose; we wish merely to demonstrate the program operations. The input parameters for this model are provided in the example application files that are distributed with EqHaz. The files demonstrate an example hazard calculation with this model for Montreal, Quebec.

EqHaz2 - Calculation of mean-hazard curve

EqHaz2 reads the earthquake catalogues generated by EqHaz1, and generates ground-motion catalogues for a specified site, along with a mean-hazard curve, using specified ground-motion prediction equations (GMPEs), provided in tabular format. The tabular format allows for maximum flexibility in specifying the GMPEs. If multiple GMPEs are desired, there will be several alternative tables, with associated relative weights; the alternative GMPEs are used to model epistemic uncertainty in ground motion generation and propagation.

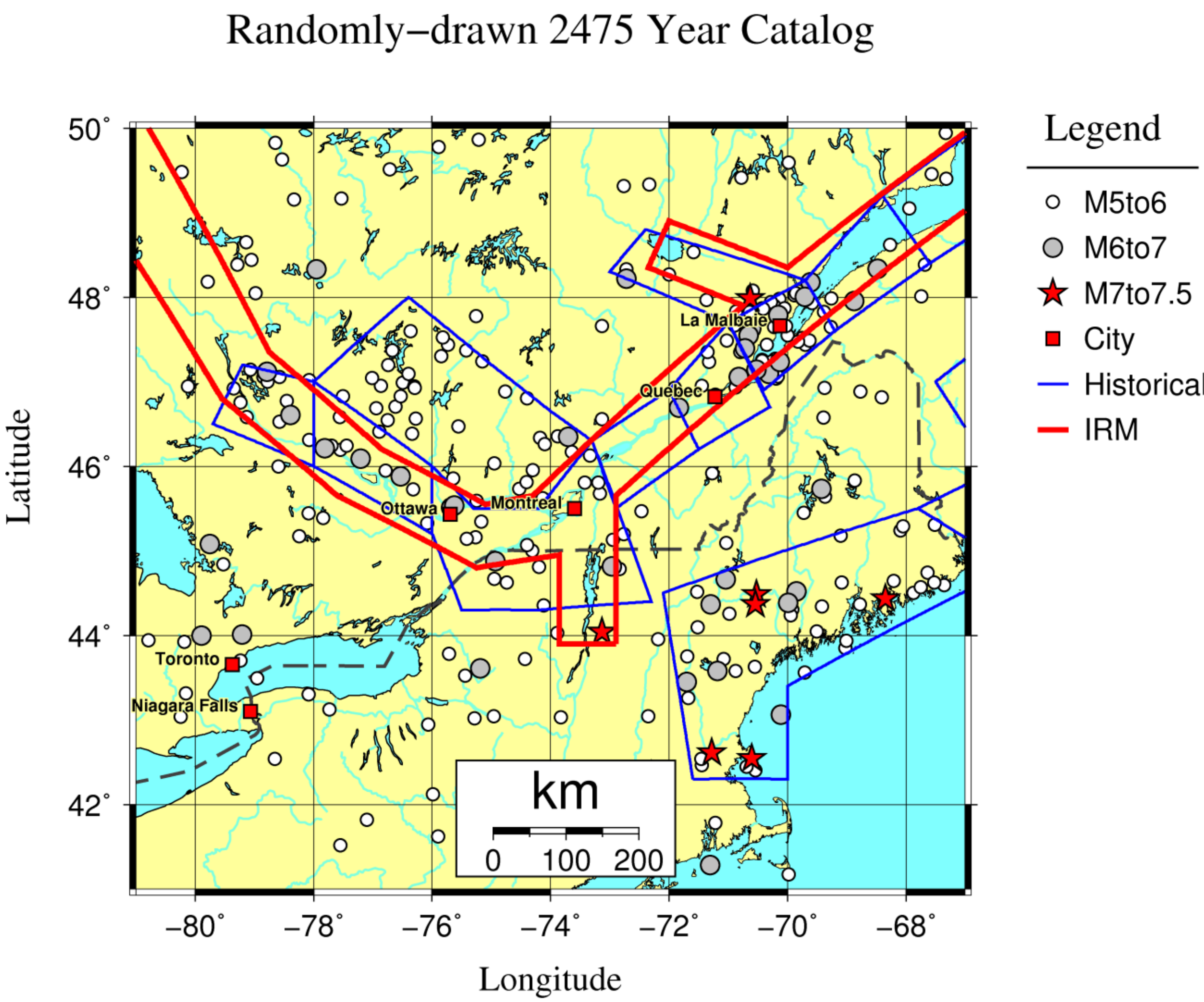

*Figure 2 – Source zone boundaries of historical seismicity zones (blue lines; **M**≤6.5) and overlying characteristic IRM zone (red lines: **M**6.5 to 7.5), along with a 2475 year subcatalogue of events of **M**≥5 generated by EqHaz1. Events outside the source zones come from a prescribed "background" seismicity rate. Example hazard calculations will be performed at Montreal.*

The EqHaz2 program reads three sets of input files: i) earthquake catalogue file(s) generated by EqHaz1; ii) GMPE table(s), which include the median predicted ground motions as a function of magnitude, distance, and frequency, and their associated aleatory variability (sigma); and iii) an input file specifying the location of sites, EqHaz1 catalogue weights (in the case of multiple catalogues), and GMPE weights (in the case of multiple GMPEs). The program generates three sets of outputs: i) a table of ground motions for each combination of site, catalogue, and GMPE;

ii) an expected ground motion table for each site, using the combined effects of all considered GMPEs (for multiple, weighted GMPEs); and iii) the mean-hazard curve for each site (considering all specified uncertainties). EqHaz2 also generates a table of mean-hazard motions for a user-specified return period, at a large number of sites, for hazard-map preparation purposes. Figure 3 is an example of a map made from such a table.

Figure 1B shows the flowchart of EqHaz2. After reading the site coordinates, GMPEs with their corresponding weights, and synthetic catalogue(s), EqHaz2 calculates R_{epi} (epicentral) and R_{hyp} (hypocentral) distances for each record in the catalogue with respect to the site. It also calculates equivalent closest distances to the fault plane (R_{cd}) and its surface projection (Joyner-Boore distance, R_{jb}). The equivalent R_{cd} and R_{jb} values for areal sources are specified by statistical models, with choices including the model of Goda et al. (2010), or the simplified model from Atkinson (2012), as developed for seismic hazard mapping applications in Canada. (Note that some of these choices include options for specifying smaller fault sizes to model high stress drop, suitable for applications in ENA.) For fault sources, distance metrics are calculated in greater detail, based on the fault geometry, location of hypocenter on the fault area, and rupture area size, as estimated based on the empirical relations of Wells and Coppersmith (1994). The next step is to calculate the received ground motions at the site. This calculation is carried out by linear interpolation of the median motions, as prescribed by log (ground-motion) in log (frequency), log (distance), and magnitude space provided by the GMPE tables (i.e. the GMPEs given in tabular format as specified in the input file). A random number drawn from the standard normal distribution is multiplied by the given sigma value (variability) of the GMPE model, and added to the median log ground motions, to model the aleatory variability in ground motions. After calculating ground motions for all records of a catalogue, corresponding to a site and GMPE, EqHaz2 writes the generated ground-motion catalogue. Ground-motion catalogues are generated for each combination of site, earthquake catalogue, and GMPE; this is the first set of EqHaz2 outputs.

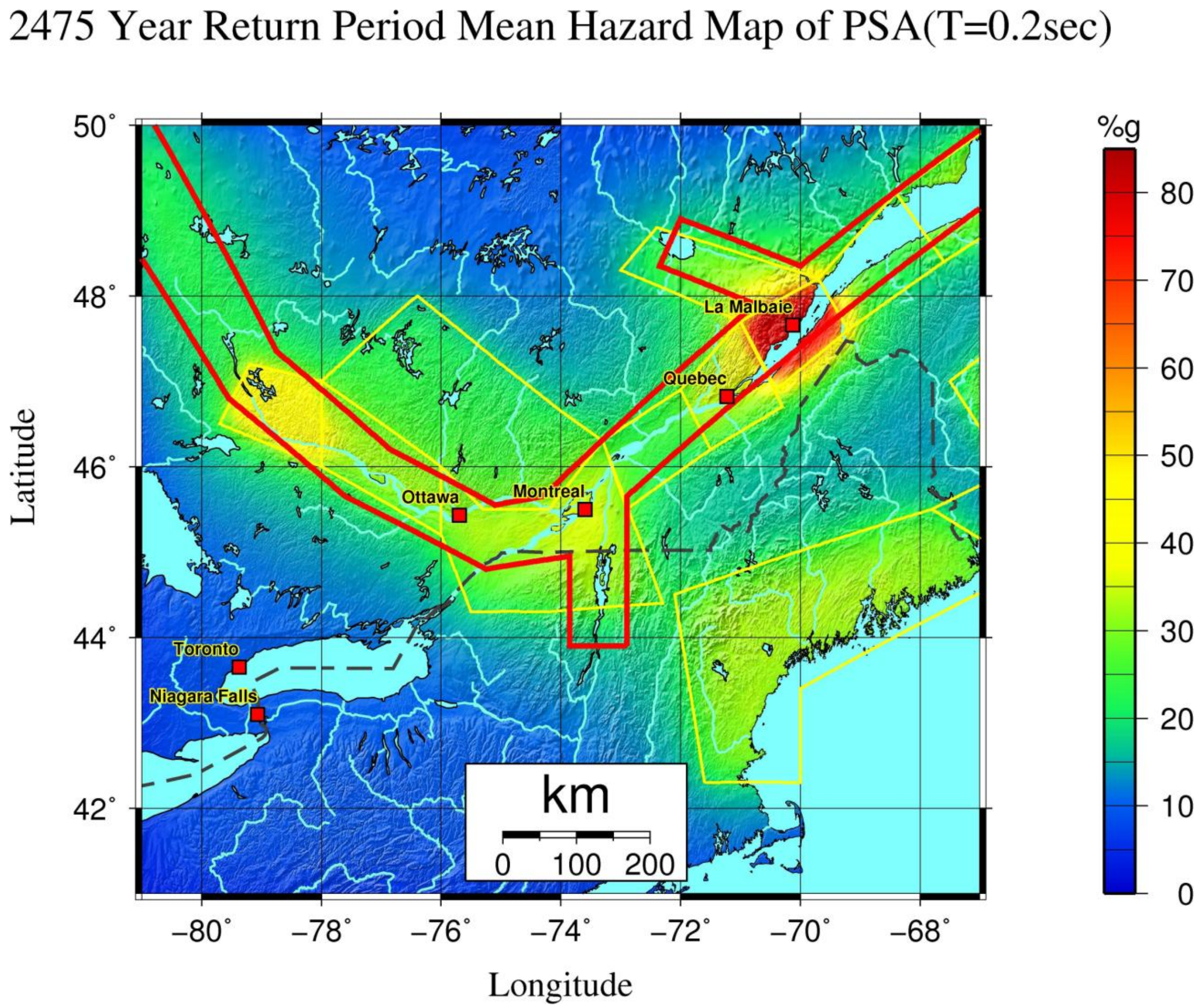

Figure 3 – A sample 1/2475 per annum mean-hazard map of PSA(T=0.2s) based on the Atkinson and Goda 2011 source model for eastern Canada(as shown in Figure 1).

In the next step, EqHaz2 makes a combined ground-motion output table by drawing from the catalogues generated for each GMPE (considering their weights); for each record, one of the GMPE results is drawn. The resulting ground-motion catalogue, combining all specified GMPEs, is the second set of EqHaz2 outputs; this ground-motion catalogue is used for the preparation of mean-hazard curves in EqHaz2. (Mean-hazard curves derived from the individual subcatalogues are used to calculate ground-motion fractiles in EqHaz3.)

To obtain the mean-hazard curve, EqHaz2 counts the number of exceedances of specified ground-motion levels in the generated ground-motion catalogue, and divides these numbers by the equivalent total duration of the catalogue (=number of subcatalogues × duration of a subcatalogue) to find the rate of exceedance of each ground-motion amplitude level. Rates of exceedance are used to calculated probabilities, assuming a Poisson process. Tabulated values of ground-motion amplitude levels versus their corresponding rates and probabilities of exceedance constitute the hazard curve, and are the third set of EqHaz2 outputs. This is a mean-hazard curve. It is important to note that the mean-hazard curve is not sensitive to the relative treatment of epistemic and aleatory uncertainty; the mean is the same regardless of how they are partitioned (see McGuire, 2004). The program repeats the process for all combinations of input synthetic catalogues and sites.

Figure 4 provides an example mean-hazard curve for Montreal, calculated from a 4.95Myr synthetic catalogue, based on the source model illustrated in Figure 2, using a weighted selection of ENA GMPEs proposed by Atkinson (2012) to characterize epistemic uncertainty for hazard mapping in eastern Canada. The actual GMPEs (provided in tabular format in the EqHaz example application files) are not important here, as this is just an illustration of the output.

EqHaz3 - Statistics of amplitude exceedance

EqHaz3 sorts the ground-motion catalogue generated by EqHaz2 in decreasing-amplitude order. This sorted list is used to make a table of fractiles of ground-motion amplitudes for given return periods (where return period is the inverse of annual probability of exceedance). It also uses the ground-motion catalogue to tabulate the contributions to hazard in magnitude-distance space (de-aggregation). The approach of EqHaz to characterizing uncertainty in seismic hazard inputs is flexible. It allows for the separate treatment of epistemic and aleatory uncertainties ("traditional" approach to uncertainty) or an alternative approach that "mixes them" (treats them as being equivalent). The latter approach, which we prefer, models amplitudes in the context of an extreme-value distribution of motions received at a site; we discuss this approach in more detail in the next section.

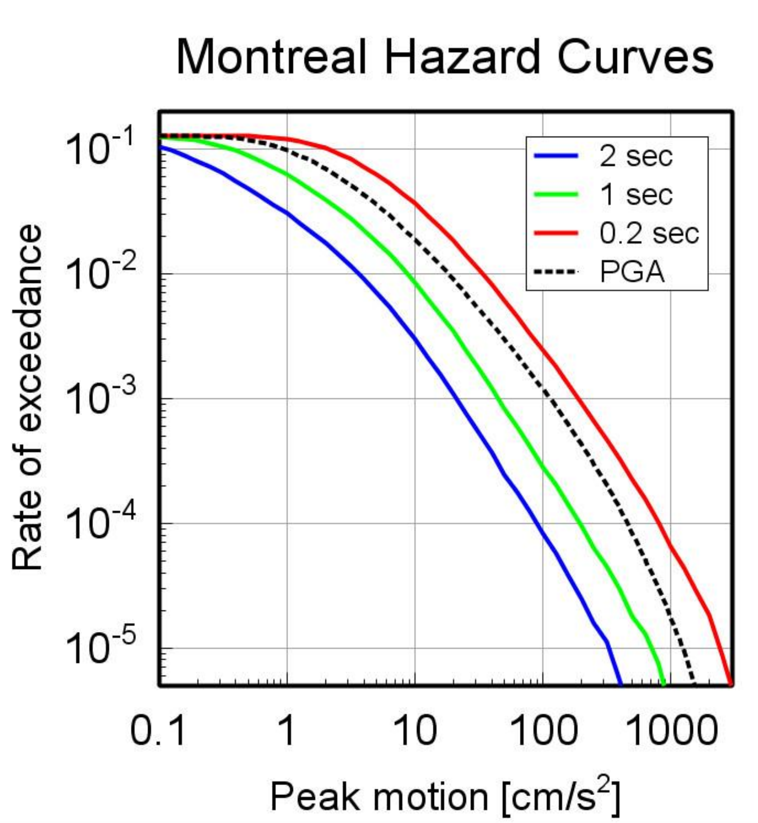

Figure 4 – Plot of mean-hazard curves constructed by EqHaz2 for the example application in Montreal. Ground-motion amplitudes are plotted for horizontal-component pseudo-acceleration, for a reference site condition, for periods of 0.2, 1 and 2 sec, and peak ground acceleration (PGA).

We note that if a user wishes to implement the traditional approach to hazard analysis, in which epistemic uncertainty is specified as being distinct from aleatory uncertainty, this should be done consistently through all three modules of EqHaz. This may be accomplished by setting the appropriate flag in the EqHaz1 input, to force it to apply a fixed set of input parameters to the simulated earthquakes of each subcatalogue. Likewise, EqHaz2 inputs should be set to use an identical GMPE for each record within a subcatalogue. Finally, a flag in EqHaz3 is set to calculate mean–hazard motions separately for each subcatalogue. In this way, each subcatalogue represents a distinct combination of input variables that sample the prescribed epistemic uncertainties. An example of the traditionally-calculated 2475-year return period motions for the example problem (hazard at Montreal) is shown in Figure 5.

In EqHaz3, we seek to characterize the distribution of amplitudes received at a site, over multiple realizations of a specified return period. The inputs to EqHaz3 are the ground-motion catalogues generated by EqHaz2, and a control file specifying return periods of interest. Four types of output are generated. The first output is just the sorted list of ground-motion values; note that by dividing the order of each record in this catalogue by the total duration of the catalogue (which is the product of the subcatalogue duration and the number of subcatalogues), we can assign a rate of exceedance to each ground-motion amplitude value. The second type of output is a table of fractiles of peak motions with specified return periods. To generate these tables, the program takes the whole catalogue (comprised of all subcatalogues) and subdivides it into continuous, non-overlapping, equal-duration windows (for example, windows of 2500 years, if we are interested in a 2500 year return period). Then it finds the maximum motions within each such window, and sorts them in decreasing order. If there are N windows, then the percentile associated with the i^{th} motion in the sorted list of N values will be $100\times(N-i)/N$. In this way we find the fraction of times that we would expect the maximum amplitude at a site, say in a 2500 year period, to exceed a certain value. To illustrate, if we have 100 realizations of a 2500-year time window, we may find that the maximum PGA exceeds 10%g in 10 of the realizations, or 10% of the time; in this case 10%g would represent the 90^{th} percentile motion for a 2500-year return period. One table of extreme-value statistics is created for every specified return period. Figure 5 shows an example of these peak motion percentiles for a 2475 year return period, as derived from a 4.95 Myr synthetic catalogue. EqHaz3 tabulates the contributions to the extremes in magnitude-distance bins, which provides basic deaggregation information.

The third type of output from EqHaz3 provides more detailed and formatted deaggregation information. After sorting the ground-motion catalogue, N_e of the records having the largest amplitudes are selected, such that the ratio of catalogue duration to N_e provides the required return period. Then the distribution of these motions and their relative frequencies are tabulated against a list of magnitude-distance bins. Figure 6 shows an example of the deaggregation calculation for Montreal, from a 4.95Myr synthetic catalogue generated using the seismic source model of Figure 2, and ground motions as shown in the hazard curves of Figure 4.

The fourth type of output is the fractile table. EqHaz3 calculates the mean–hazard amplitudes of each subcatalogue, sorts the values in decreasing order, assigns percentiles to each value and writes them in the fractile table. We have found that if the length of the subcatalogue is of the order of ~10 times the required return period, the calculated mean–hazard motions of the subcatalogues and the resulting fractiles will be stable and reliable. Figure 5 shows fractiles of 2475-year return period motions, derived from 100 synthetic subcatalogues of 49500 years (4.95Myr catalogue in total), and compares them with the peak motion percentiles for a 2475 year return period derived from the same catalogue. We discuss the differences between these approaches in the next section.

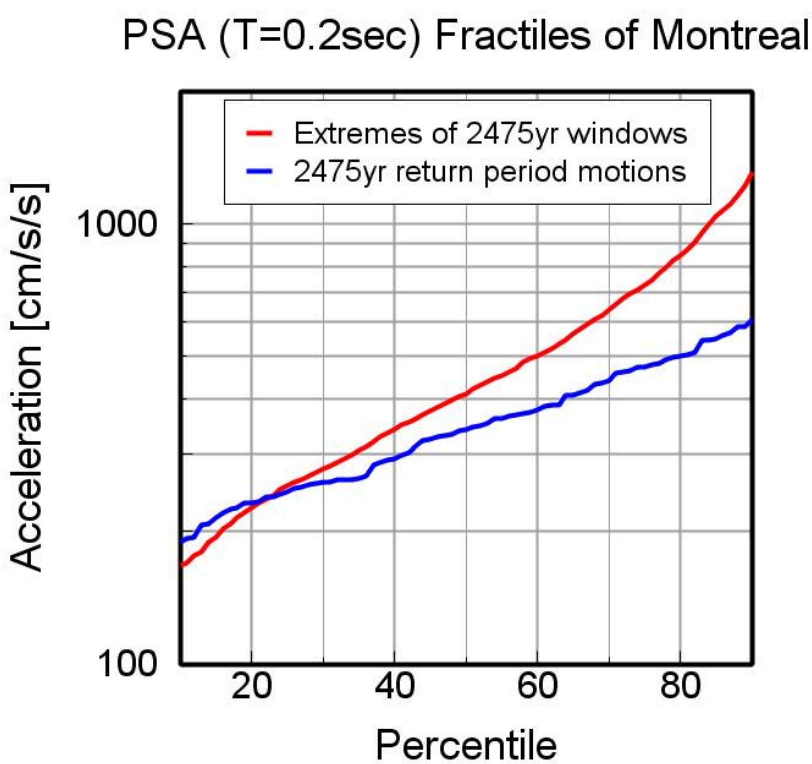

Figure 5 – Fractiles for the 2475-year return period motions for PSA(T=0.2s) at Montreal, comparing the amplitudes calculated using the traditional separation of aleatory and epistemic uncertainty (blue line), and the approach in which fractiles of extreme values in 2475-year windows are calculated (mixing epistemic and aleatory variability; red line). These results are derived from a 4.95Myr synthetic catalog based on the Atkinson and Goda (2011) model for

eastern Canada; note the mean-hazard amplitudes are the same for both approaches, though the fractiles differ.

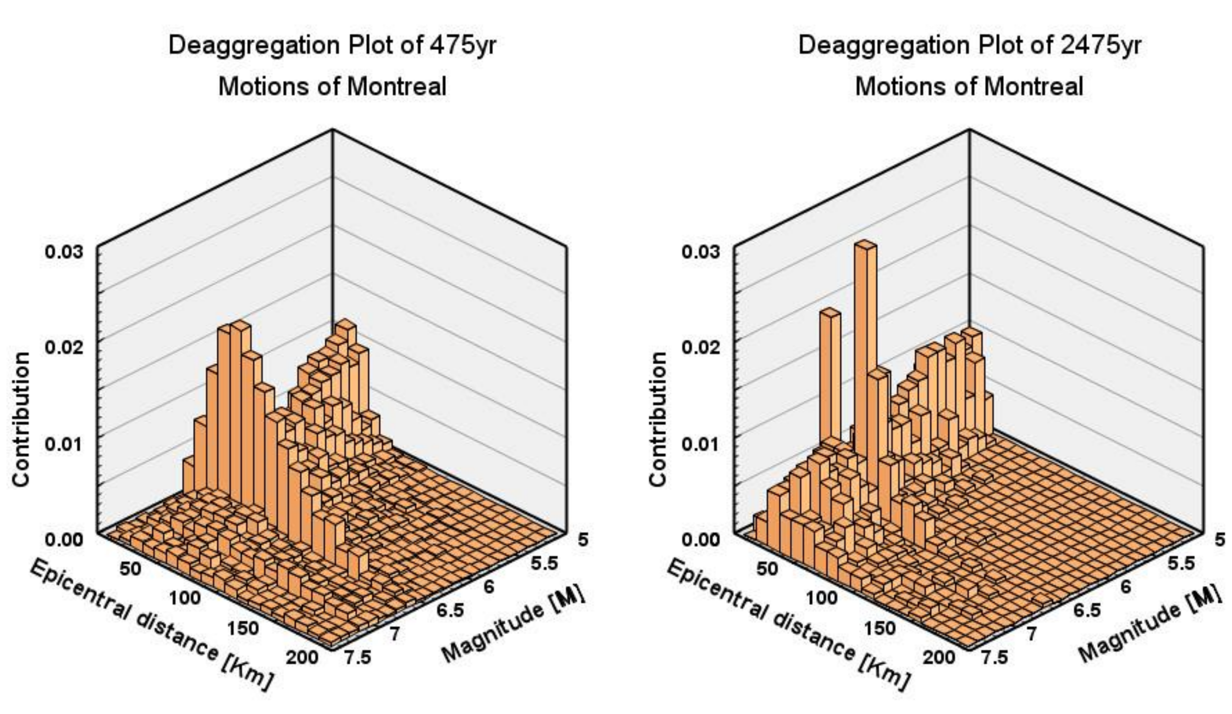

Figure 6. Contributions of different magnitude-distance bins to the T=0.2sec PSA hazard in Montreal, calculated from a 4.95Myr synthetic catalogue.

Treatment of uncertainty

Traditional Approach

Various alternatives may be formulated for each of the major inputs to a PSHA, and we are uncertain as to which of these inputs may be the most correct representation of the seismic hazard environment. This uncertainty in the input parameters has long been recognized in PSHA as epistemic uncertainty (e.g. McGuire, 2004; Abrahamson and Bommer, 2005; Bommer and Abrahamson, 2006; Bommer and Scherbaum, 2008). The traditional approach to handling epistemic uncertainty evolved gradually from the 'best estimate' case. First, sensitivity analyses were used to consider the range of estimates that might be obtained for the amplitude associated with a specific return period, as the key input parameters were varied. Typically, judgment was then exercised to pick a 'design spectrum' (assuming amplitudes were specified at multiple vibration periods), based on the sensitivity cases. This evolved to the use of logic trees (Kulkarni et al, 1984; Bommer et al., 2005), in which all possible combinations of uncertain inputs (including alternative models of source zonation, magnitude recurrence and GMPEs, for example) were used to produce a family of output probability curves, the weighted combination of which could be used to calculate a mean-hazard curve, and fractile curves expressing our confidence in the estimate, given the identified epistemic uncertainties (e.g. McGuire, 2004; Bommer and Scherbaum, 2008).

To calculate fractile hazard curves based on this traditional concept, one must make a clear distinction between aleatory uncertainty and epistemic uncertainty. The aleatory uncertainty, such as the variability about a median GMPE, is integrated over in the hazard computation. The epistemic uncertainty, which is ascribed to alternative choices for the best median values for specific parameter choices, is treated through producing and weighting multiple answers to the question: "what is the probability of exceeding a given amplitude?" In the context of EqHaz, this requires running EqHaz1 so as to generate subcatalogues that each have a fixed combination of input parameters (a distinct set of magnitude recurrence parameters and maximum magnitude) for each zone; the combination of parameters that applies to the subcatalogue are drawn randomly for each of them, based on the given parameter weights. The subcatalogues generated in this way reflect epistemic uncertainty in the input parameter values in the traditional manner, with each subcatalogue representing a unique combination of input parameters.

For each catalogue generated by EqHaz1, there are alternative median GMPEs that may apply. EqHaz2 reads the EqHaz1 output catalogues and uses a GMPE to calculate received ground motions at a site. In the traditional approach to epistemic uncertainty, we randomly select a GMPE that will apply to all records of each subcatalogue, based on a weighted random draw from the suite of alternative GMPEs. This maintains the uniqueness of the parameter combination within each subcatalogue. The motions at the site are calculated accordingly for all records of a catalogue. Then EqHaz 3 reads the table of motions generated by EqHaz2, and calculates the hazard curves for each subcatalogue separately. These curves are used to calculate the classical fractiles (e.g. median, 84th percentile) for the Uniform Hazard Spectrum (McGuire, 2004) at any given probability, from the EqHaz2 outputs.

The extreme value method

In the approach to uncertainty described above, a clear distinction is made between aleatory and epistemic uncertainty. It is widely recognized that this distinction is largely artificial (Bommer et al., 2005; Bommer and Scherbaum, 2008). In fact, one might argue that all uncertainty is epistemic, in that if we knew the earth's processes and parameters with sufficient accuracy and precision, then the entire problem would be essentially deterministic. In light of these considerations, we have formulated an alternative to the traditional approach to uncertainty that we prefer, in which epistemic and aleatory uncertainties are treated as being equivalent. The basic idea is that we consider all uncertainties in the same way, through Monte Carlo simulation, throughout the PSHA. This allows efficient sampling of the parameter space, and lends itself to the treatment of the resulting amplitudes using extreme-value statistics in EqHaz3. In EqHaz1, we create a long-duration catalogue (e.g. 5,000,000 years, say, subdivided into a number of subcatalogues for convenience) for each alternative source zonation model. Within each source-zonation catalogue, we use the Monte Carlo approach to make a weighted random draw of all parameters that will apply in generating each event – including its location, depth, recurrence model and M_x. Thus each generated catalogue contains a long sequence that implicitly contains both the epistemic and aleatory variability in these parameters (appropriately weighted), for that particular source zonation model. In EqHaz2 we associate each event in each catalogue with a GMPE and a sigma level, again using a weighted random draw from both the available median GMPEs and their variability. This allows a mean-hazard curve to be constructed for each source model, with the weighted sum of these curves being used to construct the final (composite) mean-hazard curve (considering all of the alternative source zonations). Note that this mean-hazard curve will equal that produced by the traditional approach, as the mean is insensitive to the subdivision of aleatory versus epistemic uncertainty (McGuire, 2004).

Unlike the mean, the fractiles expressing uncertainty in the PSHA results are sensitive to how epistemic uncertainty is modeled. In our preferred approach, we use EqHaz3 to subdivide the EqHaz2 output files into blocks corresponding to a specific return period. For example, a 5,000,000-year catalogue of motions received at a site can be subdivided into 2000 sub-catalogues, each of 2500 years. For each of these sub-catalogues, we find the maximum amplitude. This is one realization of the 1/2500 year motion. We sort these motions to provide fractiles that tell us how often the 1/2500 year motion is below a certain value. Because of the sampling technique used, these are smooth curves. These amplitudes follow a general extreme-value distribution, which may also be used to fit the statistical parameters of amplitudes. Because we can fit the extreme-value parameters robustly, this approach facilitates investigation of the occurrence of extremes (i.e. motions with very low probabilities of exceedance).

This blended approach to uncertainty differs from the current-practice treatment of uncertainty. What we are doing is defining the expected future distribution of amplitudes, based on uncertain input parameters. No distinction between epistemic and aleatory uncertainty is made in this context. We note that this approach is the same in concept as the extreme-value method described decades ago by Esteva (1976) and Milne and Davenport (1969) in early PSHA developments. The difference is that because we are using a catalogue simulation approach, we eliminate the problems associated with sparse statistics that plagued these early applications. In concept, we see no fundamental reason why epistemic and aleatory uncertainty should be treated separately, especially since in practice such a distinction is not actually feasible. Therefore we prefer our alternative approach as based on the extreme value concept. It is interesting that treating the uncertainty in this way facilitates the occasional realization of extreme amplitudes over long periods of time from unlikely combinations, and results in greater predicted extremes for low-probability fractiles (see Figure 5).

EqHaz Validation

The EqHaz1program is straightforward to validate for a given set of magnitude recurrence parameters. When a sufficiently long catalogue is generated for a given source zone, of the order of 1,000,000 to 10,000,000 years for a low- to moderate-seismicity environment, we can obtain the input magnitude recurrence parameters back exactly (N_0 and β of the Gutenberg Richter relation, for example), simply by using the statistics of the synthetic catalogue to determine the Gutenberg Richter parameters. We performed such tests for several examples to ensure that the EqHaz1 program does indeed produce synthetic catalogues that follow the input distributions. This process also allows us to see how an apparent Gutenberg Richter line can change between sub-catalogues; for example, we can obtain these relationships for 500 year windows within a long catalogue to obtain an appreciation for uncertainty in parameter estimation from the short historical record. Over a long period of time, these fluctuations average out to allow the underlying relationship to be accurately recovered.

We have validated the EqHaz2 and EqHaz3 programs to obtain ground motion probabilities, for both the traditional and the extreme value approach to uncertainty, as follows.

Traditional approach

To check that our program can reproduce the results of a traditional PSHA, we repeated the calculations that were done by an engineering consulting company to obtain seismic hazard estimates for a number of dam sites across Ontario, Canada. The company (Klohn Crippen Berger Ltd.; KCBL) developed a relatively-simple 81-branch logic tree to characterize the main uncertainties in seismic source zonation and recurrence rates, maximum magnitude and GMPEs, for sites in Ontario. They used EZFRISK, a well-known industry standard developed by Risk Engineering Inc. (now available through Lettis Inc.), to produce a hazard curve for each of the 81 branches, for each site. They developed post-processing software (a script in visual basic) to combine the 81 hazard curves to produce the mean-hazard curve, as well as median and 84th percentile curves. We ran their 81 input models using EqHaz1 and EqHaz2, combining the output curves using our software (traditional method as described above). We verified that we obtain the same answers (mean, median 84th %), for three different sites and three different ground-motion periods, as was obtained independently by KCBL using the same input parameters.

As an additional check, we verified our mean-hazard curves against the corresponding mean-hazard results of the Geological Survey of Canada (GSC) as obtained using FRISK88 for the 2010 national seismic hazard maps of Canada (Adams and Halchuk, 1993; Halchuk, pers. Comm., 2012) for several selected sites (both their H and R models).

Extreme-Value approach

We reiterate that our preferred alternative approach, as based on the blending of aleatory and epistemic uncertainty in the context of an extreme-value problem, produces the same mean-hazard curve as that obtained using the traditional approach. It thus remains to verify and understand the output of EqHaz3, which provides fractile amplitude distributions for specified return periods, under this approach. We did this by inputting the vector of maximum amplitudes in 2500-year sub-catalogues (for a 5,000,000 year catalogue), as obtained by EqHaz2, into MATLAB (Mathworks, Inc.). We used MATLAB to fit a general extreme-value distribution to the maxima in the 2500 year windows, the statistics of which were used to calculate the amplitude fractiles for a 2500 year period. These matched very closely with the fractiles obtained by EqHaz3; note that theEqHaz3 fractiles are based on sorting the peak amplitudes of non-overlapping equal-duration windows outputs from EqHaz2, rather than by fitting extreme-value statistics. The fitting of extreme-value statistics to the EqHaz2 outputs can be used to investigate the statistics of extreme motions at very low probabilities under this blended approach to aleatory and epistemic uncertainty.

Future work

An advantage of the Monte Carlo formulation as used by EqHaz is that it lends itself well to future refinements, due to the flexibility of the formulation. Future developments that we plan to implement include the following:

1- incorporation of a smoothed seismicity model as an alternative to source zones.

2- flexible options for fuzzy boundaries for source zones.

3- incorporation of time-dependent models of earthquake occurrence.

4- an option to draw time histories from a suite of simulations or a ground-motion catalogue, rather than using GMPEs.

The EqHaz suite and documentation is freely available at www.seismotoolbox.ca. Please be aware that although the program is provided, we do not have the resources to provide user support. Nor can we warrant that the program is free of errors. Thus it is recommended that users be those that are very familiar with PSHA, and that have an understanding of its strengths, limitations, and proper usage.

References

Abrahamson, N. A., and J. J. Bommer (2005). Probability and Uncertainty in Seismic Hazard Analysis, *Earthquake Spectra* **21,** 603-607.

Adams, J., and S. Halchuk (2003). Fourth generation seismic hazard maps of Canada; values for over 650 Canadian localities intended for the 2005 National Building Code of Canada, *Open-File Report - Geological Survey of Canada* 155.

Atkinson (2012). "White Paper" on ground-motion prediction equations for use in the 2012 National Seismic Hazard Maps. Draft Manuscript. March 2012.

Atkinson, G. M., and K. Goda (2011). Effects of seismicity models and newground motion prediction equations on seismic hazard assessment for four Canadian cities, *Bull. Seismol. Soc. Am.* **101**, 176–189.

Bommer, J. J. , and F. Scherbaum (2008). The use and misuse of logic-trees in Probabilistic Seismic Hazard Analysis, *Earthquake Spectra*, ***24***(4), 997-1009.

Bommer, J. J., and N. A. Abrahamson (2006). Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates? *Bulletin of the Seismological Society of America* **96,** 1967-1977.

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,** 377-389.

Esteva, L. (1976). Seismicity, Elsevier Sci. Publ. Co., New York, N.Y., United States (USA), United States (USA).

Hong, H. P., and K. Goda (2006). A comparison of seismic-hazard and risk deaggregation, *Bulletin of the Seismological Society of America* **96,** 2021-2039.

McGuire, R. K. (2004). Seismic hazard and risk analysis, Vol. MNO-10, Earthquake Engineering Research Institute, Oakland, Calif.

Milne, W. G., and A. G. Davenport (1969). Distribution of earthquake risk in Canada, *Bulletin of the Seismological Society of America* **59,** 729-754.

Musson, R. M. W. (1999). Determination of design earthquakes in seismic hazard analysis through Monte Carlo simulation, *J.Earthquake Eng.* **3,** 463-474.

Musson, R. M. W. (2000). The use of Monte Carlo simulations for seismic hazard assessment in the U.K. *Ann.Geofis.* **43,** 1-9.

Wells, D. L., and K. J. Coppersmith (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bulletin of the Seismological Society of America* **84,** 974-1002.