FATALITY MODELS FOR THE U.S. GEOLOGICAL SURVEY'S PROMPT ASSESSMENT OF GLOBAL EARTHQUAKES FOR RESPONSE (PAGER) SYSTEM

K.A. Porter[1], K.S. Jaiswal[2], D.J. Wald[2], P.S. Earle[2], and M. Hearne[2]

[1] *Professor, Dept. of Civil, Environmental, & Architectural Engineering, University of Colorado, Boulder, USA*
[2] *United States Geological Survey, Golden CO, USA*
Email: keith.porter@colorado.edu, kjaiswal@usgs.gov, wald@usgs.gov, pearle@usgs.gov, mhearne@usgs.gov

ABSTRACT

The US Geological Survey is adding post-earthquake fatality estimation capability to its Prompt Assessment of Global Earthquakes for Response (PAGER) system. PAGER's goal is to inform early and rapid post-earthquake decisions about humanitarian assistance before ground-truth and news information can be acquired, and to examine hypothetical scenarios for risk-management purposes. In its post-earthquake mode, PAGER monitors the USGS's near real-time global earthquake solutions, automatically identifies the important events, and estimates the population exposed to various levels of shaking intensity. To enhance those capabilities, we develop several vulnerability models for estimating shaking-related deaths and other impacts from earthquakes anywhere in the world. Three candidates are discussed here.

KEYWORDS: loss estimation, seismic vulnerability, PAGER

1. POST-EARTHQUAKE LOSS MODELING

Computerized post-earthquake loss estimation has existed since predictive catastrophe models were developed. Commercial catastrophe models such as those developed by RMS, AIR, EQECAT, and others have been used at least since the 1989 Loma Prieta earthquake for rapid post-earthquake loss evaluation of insured loss or societal loss. The California-centric EPEDAT software (Eguchi et al. 1997) was designed explicitly for post-earthquake estimation of deaths, injuries, economic loss, and lifeline damage. FEMA's public software HAZUS-MH (NIBS and FEMA 2003) and related software developed for Australia (EQRM), Norway (Selena), and Turkey (HAZTURK) likewise quantify societal risk before earthquakes or estimate losses just afterwards. More recently, the Russian software EXTREMUM (Shakramanian et al. 2000) and the related QUAKELOSS software estimate deaths and injuries for earthquakes anywhere in the world, and the alert system GDACS provides qualitative warning levels shortly after earthquakes and other natural disasters. Aside from the commercial catastrophe software, most of these systems are open to some extent, with clear documentation, in some cases freely available underlying databases (e.g., HAZUS-MH), and in a few cases open source software code so that users can understand and develop a degree of trust in model outcomes (e.g., EQRM, Selena, and HAZTURK).

The USGS PAGER program currently estimates population exposed to various levels of shaking shortly after earthquakes, and rapidly and automatically publishes these estimates. New developments in PAGER seek to estimate fatalities for humanitarian decision-making purposes. Its goal by December 2008 is to include a probabilistic fatality estimate for any earthquake anywhere in the world, and to be entirely open (documentation, databases, software code) and freely publically available. No existing models have all of these attributes. An important challenge has been the vulnerability functions: relating shaking intensity to fatality rate, which here means shaking-related deaths as a fraction of population exposed to a given shaking intensity.

2. OBJECTIVES FOR PAGER VULNERABILITY MODELS

Empirical model: The goal for this model is to estimate country-specific mean fatality rate as a function solely of MMI, without reference to parameters other than earthquake source parameters (e.g., magnitude, location, time of day). The model should employ only total recorded shaking-related deaths (no deaths from secondary hazards) in a large catalog of past earthquakes and the estimated population exposed to various MMI levels in each event. Using this model, PAGER should estimate total event-level fatalities in future earthquakes within an average of ½ to 1 order of magnitude, with higher accuracy in highly fatal events. A procedure is required to quantify uncertainty in future earthquake fatality estimates.

Semi-empirical model: The goals for this model are the same as in the empirical case, except that mean fatality rates are to be estimated by structure type, which also requires (a) estimating the population by structure type at the time of the earthquake; (b) estimating the collapse rate by structure type as a function of intensity, and (c) estimating the total fatalities given the collapse of a structure. The model can employ expert opinion or regression analysis of historic collapse or fatality rates by country and structure type.

Analytical model: The goals for this model are the same as in the semi-empirical case, except that collapse rates are to be calculated by structural engineering principles. Intensity measures could be scalar or vector measures of spectral response, could include magnitude, distance, duration, etc.—any parameter readily available from ShakeMap procedures. "Structural engineering principles" here includes knowledge of material properties, mechanics of materials, construction practices, and structural analysis to estimate structural behavior.

3. BACKGROUND ON LOSS MODELS

Before addressing the development of the three PAGER vulnerability models, it is worthwhile to briefly summarize loss-estimation methodologies in general, and PAGER's methodology in particular. Quantitative models all work generally the same way:

(1) Determine earthquake magnitude (denoted here by M) and location. In its post-earthquake mode, PAGER monitors the USGS's near real-time global earthquake solutions, automatically identifying possibly important events. For large events, point sources are later replaced with a finite fault model. The interested reader can learn more from http://earthquake.usgs.gov/regional/neic/who_we_are.php.
(2) Apply a ground motion prediction equation (GMPE) to estimate shaking intensity either on a gridded basis or at points (e.g., city centroids, at which the entire city population is assumed to be located for modeling purposes). PAGER uses a roughly 1-km grid. Depending on magnitude, depth, and seismic domain, PAGER uses Atkinson and Boore (2006), Youngs (1997) intraslab relationship, Youngs (1997) interface relationship, or Boore et al. (1997). Where intensity data are available, such as from Did You Feel It? or recorded ground motion, local intensity corrections are made to match the observed shaking. Shaking intensity h can be measured in terms of MMI (calculated, e.g., using Wald et al. [1999] from PGA and PGV estimated by the GMPE) or using spectral measures such as S_a(0.3 sec, 5%) and S_a(1.0 sec, 5%). Wald and Allen (2007) is used for site soil classification.
(3) Determine the population in each gridcell or at each point, V_i. PAGER employs the LandScan 2006 gridded global population database (Bhaduri et al. 2002).
(4) Many models assign the population in each gridcell or point to various building types, accounting for time of day. PAGER's semi-empirical and analytical models both do so; see Jaiswal and Wald (2008a) for details.
(5) Where building types are used, apply vulnerability functions to each combination of value exposed, building type, and intensity level, to estimate loss at each location. A vulnerability function provides loss as a function of input excitation, which here means fatality rate versus either MMI or other intensity measure. In subsequent

equations, N is the number of locations of interest, i denotes an index for these locations, T is the number of possible structure types in the country or region where the earthquake occurs, and j is an index to structure type. Let V_i denote population at location i, and p_{ij} denotes the fraction of the population at location i in structure type j at the time of the earthquake. Let $y_j(h_i)$ denote the mean fatality rate in structure type j given shaking intensity h_i, which is the best estimate of shaking intensity at site i (usually from the GMPE alone). That is, $h_i = f(M, R_i, S_i, \&)$, where f denotes the GMPE, M is magnitude, R_i is the distance from the source to location i, S_i is the site classification (e.g., NEHRP A, B, C, etc.), and "&" denotes parameters of the GMPE in addition to M, R, and S. On a first-order basis, total societal loss (e.g., number of fatalities), denoted by L, is given by Equation 1, where E denotes expected value. The present paper focuses on PAGER's development of $y_j(h_i)$.

$$E\left[L\right] \approx \sum_{i=1}^{N}\sum_{j=1}^{T} V_i p_{ij} y_j\left(h_i\right) \tag{1}$$

(6) There are numerous sources of uncertainty and various ways to propagate it. The PAGER models currently employ a country-specific error term determined from hindcasting of past losses. Let Ψ denote a normalized error term for L, i.e., a variable with unit mean and some distribution to be determined, which can vary by country. Let l denote a particular value of L in a future event in that country. Let P denote probability. Then in a future event, the probability distribution of fatalities is given by:

$$P\left[L \le l\right] = P\left[\Psi \le \frac{l}{E\left[L\right]}\right] \tag{2}$$

The distribution of Ψ is derived from hindcasting losses in the country of interest. In the following, ψ denotes a particular value of Ψ, Q denotes the number of events in the catalog of historic earthquakes in the country, for which the number of fatalities is recorded, and k denotes an index to those events. Let l_k denote the number of fatalities in event k, and let $E[L_k]$ denote the model-estimated fatalities. I denotes the indicator function: 1 if the expression inside the brackets is true, 0 otherwise. Then the distribution of Ψ is estimated by:

$$P\left[\Psi \le \psi\right] = \frac{1}{Q}\sum_{k=1}^{Q} I\left[\frac{l_k}{E\left[L_k\right]} \le \psi\right] \tag{3}$$

A lognormal probability distribution fit to Ψ commonly passes a Lilliefors goodness-of-fit test.

4. DEVELOPMENT OF THE PAGER EMPIRICAL MODEL

See Jaiswal et al. (2008) for details of PAGER's empirical loss model. In it, the fatality rate $y(h)$ is expressed as a lognormal cumulative distribution function:

$$y\left(h\right) = \Phi\left(\frac{\ln\left(h/\theta\right)}{\beta}\right), \tag{4}$$

where Φ is the standard normal cumulative distribution function, h is shaking intensity (instrumental MMI) and θ and β are distribution parameters. The model does not use structure types, so $T = 1$, $p_{i1} = 1$, and Eq 1 simplifies to

$$E\left[L\right] \approx \sum_{i}^{N} y\left(h_i\right) V_i \tag{5}$$

Here, V_i is estimated as discussed above. The two free parameters, θ and β, are derived by finding the values that minimize an objective function η calculated from a catalog of past events (Allen et al. 2008). The catalog includes approximately 1,000 events with estimated values of total population by intensity level and a recorded number of shake-related deaths (zero or greater). We considered four objective functions:

$$\eta_1 = \sum_{k=1}^{Q} \left| E[L_k] - l_k \right| \tag{4}$$

$$\eta_2 = \sum_{k=1}^{Q} \left(E[L_k] - l_k \right)^2 \tag{5}$$

$$\eta_3 = \sum_{k=1}^{Q} \left(\ln\left(E[L_k]\right) - \ln\left(l_k\right) \right)^2 \tag{6}$$

$$\eta_4 = \ln\left(\sqrt{\frac{1}{Q}\sum_{k=1}^{Q} \left(E[L_k] - l_k \right)^2} \right) + \sqrt{\frac{1}{Q}\sum_{k=1}^{Q} \left(\ln\left(E[L_k]\right) - \ln\left(l_k\right) \right)^2} \tag{7}$$

Using conventional norms such as Equations (4) and (5) tends to fit the model best for the more-fatal earthquakes, as opposed to smaller events, because the large events dominate total historical losses. Equation (6) tends to cause the model to fit the smaller, more frequent losses better because of the larger number of these events in the catalog and because it is sensitive only to the *ratio* of observed to estimated losses, rather than the absolute number of deaths. The objective function shown in Equation (7) was therefore developed to balance the two extremes: the first summand works to fit the model to the larger events, the second, to the smaller ones. Figure 1 shows the result of using Equation (7) for Turkish earthquakes between 1973 and 2007. Earthquakes with zero recorded deaths have been taken as 0.1 deaths for calculation purposes, and earthquakes without a recorded number of deaths (zero or otherwise) are ignored. Except for a few outliers, the model estimates the fatalities for most of the events within 1 order of magnitude, with approximately equal accuracy at low and high recorded deaths.

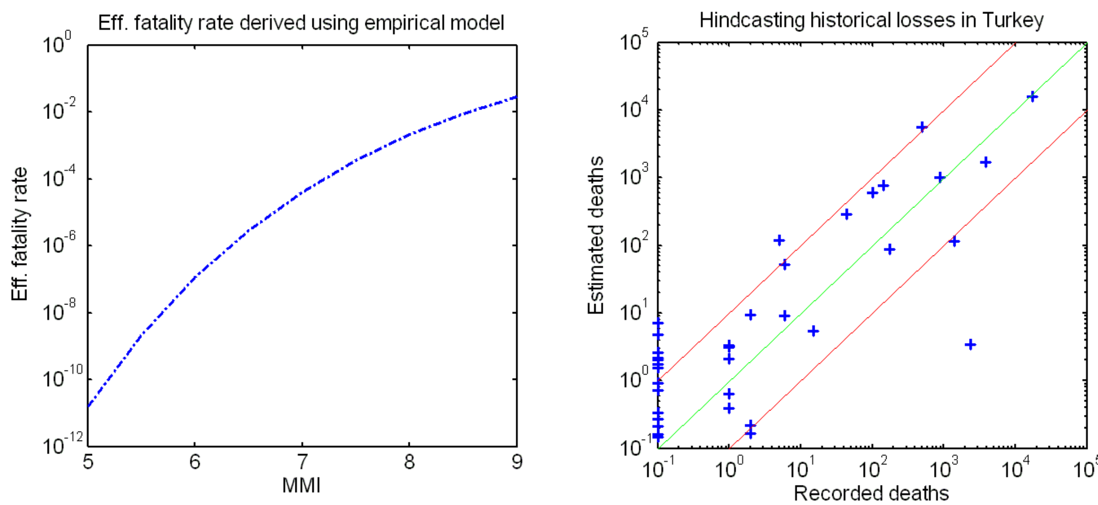


Figure 1. Fatality estimation using empirical loss modeling for Turkey.

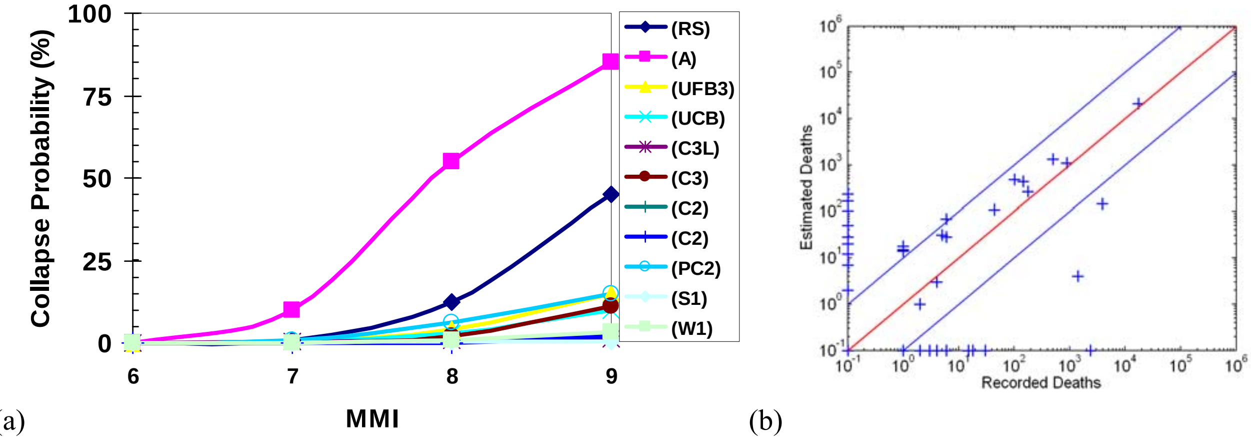
Only 35 countries have 4 or more fatal earthquakes in the catalog. To estimate fatality rates for countries that have inadequate data, we developed a regional vulnerability model, combining the events from several neighboring countries. The regionalization scheme (see Figure 2) is based on groups of countries with similar building stocks and other socio-economic characteristics, as estimated by Jaiswal and Wald (2008a). We considered past earthquake data, broad knowledge of building inventory, socio-economic characteristics in the form of Human Development Indicators (HDI) and climatic condition etc. For example, Canada and the United States outside of California are grouped together in a region, as are west-central Africa, central Asia and Northern Europe. Again, regions are only used for countries without sufficient earthquake experience in the catalog. For example Peru, Chile, and Colombia have sufficient data to create country-specific fatality models, and events in these countries have been grouped together for use in earthquakes in Ecuador, Bolivia and Argentina. The proposed scheme is broad, qualitative, based on limited data and subjective judgment, and hence should be revisited periodically as data become available.

Figure 2. Regionalization scheme used for empirical loss modeling (Jaiswal et al., 2008).

5. DEVELOPMENT OF PAGER SEMI-EMPIRICAL VULNERABILITY MODEL

Detail on the semi-empirical model is presented in Jaiswal and Wald (2008a). In it, building inventories and vulnerability are country-specific. Population and shaking intensity for each earthquake are estimated at ~1 km grid cells. Indoor population by occupancy and structure type is estimated as a function of time of day: 10 AM – 5 PM (*day*), 10 PM – 5 AM (*night*), and other times are termed *transit*. Population is then distributed among three broad occupancy categories, i.e., residential, non-residential and outdoor population by identifying local time of day of the earthquake, population density in terms of urban or rural as defined in the CIESEN (2004) GRUMP database, and country-level workforce distribution by sector of employment. Fatality estimation in this model consists of (a) estimating the fraction of indoor occupants by structure type j and occupancy category at the grid-cell level; (b) summing the fractions over occupancies by intensity level i, to produce p_{ij} of Equation (1); (c) multiplying by the collapse probability of each structure type at each intensity level; (d) multiplying by the estimated fraction of occupants in collapsed buildings who die as consequence of collapse, by structure type; multiplying the collapse probability and the indoor fatality rate given collapse produces $y_j(h)$ of Equation (1). (e) Finally, one sums fatalities over structure types and intensity levels. The building inventory database is detailed in Jaiswal and Wald (2008b); WHE experts estimated collapse probability by structure type and intensity level, as detailed in Porter et al. (2008). The fatality rate given collapse is adopted from HAZUS-MH (NIBS and FEMA 2003), as well as from in-progress work by So and Spence (2007).

Our initial analysis using WHE collapse fragility functions tends to overestimate fatalities in historical earthquakes. The overestimation is more significant for smaller earthquakes than for larger earthquakes, suggesting that collapse fragility functions have a greater positive bias relative to reality at lower intensity levels. Figure 3 is an example: Figure 3a shows collapse fragility functions used for Turkey, and Figure 3b shows estimated versus recorded deaths in Turkish earthquakes, using the semi-empirical model; it is comparable to Figure 1. (See Porter et al. [2008] for discussion of collapse fragility functions, and Jaiswal and Wald [2008b] for inventory data.) Using workforce data for Turkey, we estimate the population distribution for Turkey as follows. Daytime: 22% residential, 45% nonresidential, and the balance outdoors. At night the figures are 97.9% and 2%, respectively, and during transit, 54% and 8%. The indoor fatality rates given the collapse of Turkish buildings are: small woodframe: 0.13%; various concrete structure types: 15%; adobe: 6%; and unreinforced brick or concrete block masonry: 8%.



(a) (b)

Figure 3. (a) Fragility functions used for Turkey and (b) hindcasting accuracy for Turkish earthquakes

6. ANALYTICAL VULNERABILITY MODEL

A third approach to fatality-rate vulnerability functions is similar to HAZUS-MH (NIBS and FEMA 2003). In it, a building type is idealized by a single-degree-of-freedom nonlinear damped harmonic oscillator with given yield and ultimate points in the space of (S_d,S_a), plus elastic damping ratio B_E, and three κ terms to relate hysteretic energy dissipation to earthquake duration. (Which of the three κ values is used in any given analysis depends on earthquake magnitude.) HAZUS-MH uses a pushover curve that is linear up to the yield point (D_y,A_y), perfectly plastic after the ultimate point (D_u,A_u), with an elliptical spline between them. This requirement was relaxed for PAGER, and pushover curves were allowed to reflect strength degradation. Effective damping B_{eff} is taken as B_E where $S_d \le D_y$, and increases where $S_d > D_y$ to account for hysteretic energy dissipation in a procedure that is not detailed here (see NIBS and FEMA 2003). Using the capacity spectrum method of structural analysis, one calculates at a building's location an idealized response spectrum parameterized by the site-corrected 5%-damped elastic spectral acceleration response at 0.3 and 1.0-sec periods—denoted here respectively by $S_a(0.3,5\%)$ and $S_a(1.0,5\%)$—and depicted in the space of (S_d,S_a), as shown in Equation (8). In Equation (8), T_{AVD} denotes the period at the intersection between the constant-acceleration and constant-velocity portions of the response spectrum with damping ratio B_{eff}. The structural response of the building is the point where the pushover curve intersects the damped response spectrum with the same damping ratio. We developed a non-iterative procedure to determine the intersection that is not detailed here.

$$\begin{aligned} S_a &= S(0.3,5\%)/R_A \qquad && 0 < T \le T_{AVD} \\ &= S(1.0,5\%)/(R_V T) && T_{AVD} \le T \\ T &= 0.32\sqrt{S_d/S_a} \end{aligned} \tag{8}$$

where
$$R_A = 2.12\Big/\left(3.21 - 0.68\ln\left(100B_{eff}\right)\right) \tag{9}$$

$$R_V = 1.65\Big/\left(2.31 - 0.41\ln\left(100B_{eff}\right)\right) \tag{10}$$

$$T_{AVD} = \left[S_a(1.0,5\%)R_A\right]\Big/\left[S_a(0.3,5\%)R_V\right] \tag{11}$$

Once the performance point is determined, the probabilities of 5 structural damage states are calculated using lognormal fragility functions, and the mean fatality rate E[L] is calculated using the theorem of total probability:

$$E[L] = \sum_{ds=1}^{4}\left(P_{ds} - P_{ds+1}\right)f_{ds} + P_5 f_5 \tag{12}$$

where P_{ds} denotes the probability of reaching or exceeding structural damage state *ds*, and f_{ds} denotes the mean fatality rate given structural damage state ds. For damage states 1, 2, 3, and 4, P_{ds} is given by

$$P_{ds} = \Phi\left(\ln\left(S_d/\theta_{ds}\right)/\beta_{ds}\right) \tag{13}$$

where Φ denotes the cumulative standard normal distribution, S_d is the spectral displacement at the performance point, θ_{ds} is the median value of S_d associated with entering damage state *ds*, and β_{ds} denotes its logarithmic standard deviation. P_5 is taken as a P_c x P_4, where P_c represents a fixed fraction of building area collapsed given that the structure is in damage state 4. The fatality vulnerability of a structure type is thus defined by 22 parameters: (D_y,A_y), (D_u,A_u), B_E, the three κ values, four θ_{ds} and four β_{ds} values, P_c, and five f_{ds} values. These 22 parameters are fixed for a given structure type. Given them, one can calculate $E[L]$ for any combination of magnitude M, $S_a(0.3,5\%)$, and $S_a(1.0,5\%)$. In a transformation not detailed here, we replace one of the spectral acceleration values with knowledge of fault distance *R,* NEHRP site classification *S*, and seismic domain *X* (plate boundary or continental interior).

For US construction, NIBS and FEMA (2003) offer values of all 22 parameters. For non-US construction, we reduce to 11 the number of required parameters by ignoring all damage states but collapse, and use experimental observations to derive these in collaboration with experts from the World Housing Encyclopedia (see Porter et al. 2008). We are using these parameters to create tabular vulnerability functions for each US and non-US structure type, relating mean indoor fatality rate to either $S_a(0.3,5\%)$, or $S_a(1.0,5\%)$, for any combination of M = 5, 6, 7, or 8; R = 10, 20, 40, or 80 km, S = A, B, C, D, or E, and X = plate boundary or continental interior.

7. CONCLUSIONS

The US Geological Survey's PAGER project has developed three methods to estimate earthquake fatality vulnerability functions, i.e., relationships between shaking intensity and fraction of people killed from shaking-related damage. One is entirely empirical, producing a country or region-specific vulnerability function that depends solely on the total population exposed to various levels of MMI. It was developed by finding the fatality-rate vulnerability function for each country that best hindcasts historic losses in a catalog of events since 1973. Another, termed semi-empirical, is also country-specific, but relates fatality rate to MMI by building type. It employs the judgment or analysis of experts convened by the World Housing Encyclopedia (WHE). A third is analytical, employing HAZUS-style analysis of buildings idealized as single-degree-of-freedom nonlinear oscillators, with 5 possible damage states determined using fragility functions and a fatality rate of each state. It uses HAZUS-MH parameters for US construction, and for non-US structure types draws largely upon laboratory experimental investigation by another set of WHE experts. PAGER will add these vulnerability functions to its current capabilities to estimate the shaking-related deaths shortly after the occurrence of future earthquakes to inform humanitarian aid decisions in the hours or days after the earthquake and before first-hand observations are available to give a clearer picture of need.

REFERENCES

Allen, T.I., D.J. Wald, A.J. Hotovec, P.S. Earle, and K.D. Marano (2008). An Atlas of ShakeMaps for Selected Global Earthquakes. USGS Open File Report 2008-1236, US Geological Survey, Golden CO.

Atkinson, G.M., and D.M. Boore (2006). Earthquake ground-motion prediction equations for eastern North America. *Bulletin of the Seismological Society of America;* December 2006; 96 (6), 2181-2205

Badal, J., M. Vazquez-Prada, and A. Gonzalez (2005). Preliminary quantitative assessment of earthquake casualties and damages. *Natural Hazards* 34:353-474

Bhaduri, B., E. Bright, P. Coleman, and J. Dobson (2002). LandScan: locating people is what matters. ***Geoinformatics*** **5:2** (April 2002) 34-37

(CIESIN) Center for International Earth Science Information Network (2004). Global Rural-Urban Mapping Project (GRUMP), Alpha Version: Urban Extents. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw (Jan 2007).

Eguchi, R.T., J.D. Goltz, H.A. Seligson, P.J. Flores, N.C. Blais, T.H. Heaton, and E. Bortugno (1997). Real-time loss estimation as an emergency response decision support system: the Early Post-Earthquake Damage Assessment Tool (EPEDAT). *Earthquake Spectra*, 13 (4), 815-833

(NIBS and FEMA) National Institute of Building Sciences and Federal Emergency Management Agency, 2003. ***Multi-hazard Loss Estimation Methodology, Earthquake Model, HAZUS®MH Technical Manual***, FEMA, Washington DC, 690 pp

Jaiswal, K. S., and D. J. Wald (2008a). A semi-empirical approach for rapid earthquake loss estimation. *U.S.G.S. Open File Report* (in preparation).

Jaiswal, K. S., and D. J. Wald (2008b). Creating a global building inventory for earthquake loss assessment and risk management. *U.S.G.S. Open File Report* 2008-1160, pp 109.

Jaiswal, K., D.J. Wald, and M.G. Hearne (2008). Estimating casualties for large worldwide earthquakes using an empirical approach, *U.S.G.S. Open File Report* (in preparation).

Porter, K.A. (2008). Cracking an open safe: HAZUS vulnerability functions in terms of structure-independent spectral acceleration. *Earthquake Spectra* (submitted).

Porter, K.A., Jaiswal, K.S., Wald, D.J., Greene, M., and C. Comartin (2008). WHE-PAGER Project: A new initiative in estimation global building inventory and its seismic vulnerability. *Proc. 14th WCEE Beijing*, China.

Shakhramanian, M.A., V.I. Larionov, G.M. Nigmetov, and S.P. Sutschev, (2000). *Assessment of the Seismic Risk and Forecasting Consequences of Earthquakes While Solving Problems on Population Rescue (Theory and Practice)*. Ministry of Russian Civil Defense and Disaster Management Research Inst., Moscow, Russia, 180 pp.

So, E., and R. Spence (2007). Personal communication.

Wald, D.J., V. Quitoriano, T.H. Heaton, and H. Kanamori (1999). Relationships between peak ground acceleration, peak ground velocity and Modified Mercalli Intensity in California. ***Earthquake Spectra***, **15:3**, 557-564

Wald, D.J. and T.I. Allen, 2007. Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, **97**, 1379-1395

Wald, D.J, P.S. Earle, K. Porter, K. Jaiswal, and T.I. Allen (2008). Development of the U.S. Geological Survey's Prompt Assessment of Global Earthquakes for Response (PAGER) System, *14WCEE*, Beijing.

Youngs, R.R. (1997). Strong ground motion attenuation relationships for subduction zone earthquakes. *Seismological Research Letters*. **68:1**, 58-73. Jan.-Feb. 1997