Bull Earthquake Eng (2013) 11:347–363
DOI 10.1007/s10518-012-9373-8

ORIGINAL RESEARCH PAPER

Estimating shaking-induced casualties and building damage for global earthquake events: a proposed modelling approach

Emily So · Robin Spence

Received: 26 May 2011 / Accepted: 6 August 2012 / Published online: 30 August 2012

Abstract Recent earthquakes such as the Haiti earthquake of 12 January 2010 and the Qinghai earthquake on 14 April 2010 have highlighted the importance of rapid estimation of casualties after the event for humanitarian response. Both of these events resulted in surprisingly high death tolls, casualties and survivors made homeless. In the Mw = 7.0 Haiti earthquake, over 200,000 people perished with more than 300,000 reported injuries and 2 million made homeless. The Mw = 6.9 earthquake in Qinghai resulted in over 2,000 deaths with a further 11,000 people with serious or moderate injuries and 100,000 people have been left homeless in this mountainous region of China. In such events relief efforts can be significantly benefitted by the availability of rapid estimation and mapping of expected casualties. This paper contributes to ongoing global efforts to estimate probable earthquake casualties very rapidly after an earthquake has taken place. The analysis uses the assembled empirical damage and casualty data in the Cambridge Earthquake Impacts Database (CEQID) and explores data by event and across events to test the relationships of building and fatality distributions to the main explanatory variables of building type, building damage level and earthquake intensity. The prototype global casualty estimation model described here uses a semi-empirical approach that estimates damage rates for different classes of buildings present in the local building stock, and then relates fatality rates to the damage rates of each class of buildings. This approach accounts for the effect of the very different types of buildings (by climatic zone, urban or rural location, culture, income level etc), on casualties. The resulting casualty parameters were tested against the overall casualty data from several historical earthquakes in CEQID; a reasonable fit was found.

Keywords Building damage · Casualties · Earthquakes · Loss estimation · Modelling

E. So (✉)
Department of Architecture, University of Cambridge, 1-5 Scroope Terrace, Cambridge CB2 1PX, UK
e-mail: ekms2@cam.ac.uk

R. Spence
Cambridge Architectural Research Ltd., 25 Gwydir Street #6, Cambridge CB1 2LG, UK

1 Introduction

Rapid urbanisation and urgent housing needs combined with a lack of financial resources and institutional development in the form of earthquake code advances and implementation means that at present, the inhabitants of many cities in the developing world are exposed to greater than average earthquake risk. This ever-changing exposure and vulnerability landscape poses significant challenges for earthquake loss estimation. In many regions, though traditional forms of construction for dwellings are being replaced with new buildings that incorporate steel and reinforced concrete, the chances of survival of their occupants during strong earthquakes do not seem to have improved. Traditional building techniques are gradually diminishing in the rapidly depopulating rural parts of the developing world, while “modern” construction practices are now common even in regions that are least developed. The housing stock is what can be described as ‘second generation’ housing where traditional techniques and materials are mixed with new materials, such as reinforced concrete.

It is now widely recognised that the number of earthquake shaking deaths is closely related to the numbers of buildings that fully or partially collapse (So 2009). This number is also agreed to depend on the ground shaking intensity, the numbers of buildings affected at a particular intensity, their occupancy and the fatality rate among occupants, as suggested by Coburn and Spence (2002). Thus it is expected that better estimation of seismic intensity, combined with knowledge of the vulnerability of the building stock and population density will provide a better way to estimate earthquake shaking deaths. Moreover the advent, in recent years, of methods for immediate post-earthquake mapping of the expected ground shaking intensity, has provided the opportunity to provide estimates of likely casualties within a few minutes of the occurrence of an event, with potentially huge benefits for the mobilisation of search and rescue and relief efforts. Two separate approaches to the definition of such estimates have been proposed: by the U.S. Geological Survey Prompt Assessment for Global Earthquake Response (PAGER) group (Wald et al. 2010) and the World Agency of Planetary Monitoring and Earthquake Risk Reduction (WAPMERR) group (Trendafiloski et al. 2011).

PAGER developers propose three separate approaches, an empirical, a semi-empirical and an analytical approach (Jaiswal et al. 2011). In the empirical approach, a fatality rate is proposed as a proportion of the population exposed at each intensity level, and depends on the shaking intensity according to a lognormal function, with values of the two separate parameters defining the function, and one defining an uncertainty factor, each given for different countries or regions of the world (Jaiswal et al., OFR 2009). Population affected at any intensity is determined by overlaying the USGS ShakeMap (created within 30 min of the earthquake’s occurrence) with the LandScan global population maps (Dobson et al. 2003). The semi-empirical approach aims to develop a better casualty estimate by using, for the area affected at each intensity level, the number of buildings, and their vulnerability to collapse at the estimated ground shaking, combined with an estimate of the fatality (or lethality) rate among occupants, given collapse (Jaiswal and Wald 2010). To apply this approach to PAGER for worldwide earthquakes, the collapse fragilities were assembled using an expert judgment approach using the experts from 26 countries who have contributed to the World Housing Encyclopedia/ Earthquake Engineering Research Institute (WHE-EERI) survey; HAZUS fatality rates (NIBS-FEMA 2006) were used for inside the US and for outside the US, rates were derived from the results of the EU LessLoss project (Spence 2007). The analytical approach is similar in concept, but uses the HAZUS capacity-spectrum approach to estimate collapse rates for the building types affected (Porter 2009).

Table 1 Building vulnerability classes

Class	Description
A	Weak masonry
B	Load-bearing masonry, unreinforced
C	Structural masonry; pre-code reinforced concrete (RC) frame
D	Moderate code reinforced concrete (RC) frame; concrete shear wall; timber frame
E	Steel frame; high code RC

The casualty estimation approach adopted by WAPMERR in their QLARM loss estimation method (Trendafiloski et al. 2011), is similar in concept to the semi-empirical approach of PAGER. However, a more elaborate world wide population and building stock exposure database has been developed, in which the building stock and population for each city and rural area in the world is estimated, and divided between five vulnerability classes, A to E as defined in EMS-98 (Grünthal 1998). A separate collapse rate for each vulnerability class is determined for all of nine global regions using the World Housing Encyclopedia data (WHE). However, for fatality and injury rates, data proposed by HAZUS (NIBS-FEMA 2006) are used as default, with updates using observed data.

Estimates of casualties are routinely published by WAPMERR for the benefit of post-disaster rapid response agencies; such estimates are also compiled (and, since October 2010 routinely published), by PAGER. The estimates given by WAPMERR are often within a factor of two of the final death toll, but some significant discrepancies occur. Two inherent weaknesses in all methods so far used are:

(1) lack of detailed knowledge of the building stocks in many areas of the world (particularly in areas where the most vulnerable populations live), and
(2) inadequate knowledge of the fatality rate given building collapse, and understanding of the factors affecting this rate.

In this paper an alternative semi-empirical method is proposed. This method adopts a simplified single-parameter approach to defining the distribution of the building stock into vulnerability classes, using empirical collapse rate data for each vulnerability class, and adopting occupancy and fatality rates based on recent research. The results of this approach are then tested against actual casualty data that are assembled in the Cambridge Earthquake Impacts Database (CEQID, www.ceqid.org).

There are six components to the approach.

- First, the building stock in any location is defined in terms of its distribution among five vulnerability classes (Classes A to E, Table 1). A binomial distribution is assumed as this allows a single parameter, p, to be used to define the entire distribution, with relatively small implications for the resulting loss estimates (Sect. 2.5.1).
- The population distribution by building class is then derived using data (or assumptions) on relative occupancy levels for different vulnerability classes.
- For each vulnerability class, empirical data have been used to estimate the expected proportion of the buildings in that class to collapse (damage state d5) or to sustain very heavy damage (d4), at increasing levels of ground motion intensity, measured by the Modified Mercalli intensity scale (MMI).
- For each building class, a value of the lethality rate (fatality rate) for buildings which are at damage levels d5 (collapsed) or d4 (very heavily damaged) is determined.

- Using an assumed instantaneous occupancy rate for the area (depending on the time of day of the event and other factors), and assumed lethality rates for the building classes, an estimated fatality rate for the combined building stock is calculated at each given intensity.
- The number of deaths in any zone affected by a given intensity is then determined based on the total estimated population of that zone. The total number of deaths caused by the event is calculated by summing over all settlements with significant population and potentially destructive ground shaking, which can range upwards from MMI = 5 depending on the vulnerability of the affected buildings.

The approach is designed to be incorporated within systems such as the USGS PAGER, in which the distribution of ground shaking in MMI and the estimated population at a given intensity are already calculated. The additional information needed to produce a casualty estimate is just the p value (or values) attributable to the building stock of the area, and the occupancy and lethality parameters to be used. This approach is much simpler in application to many parts of the world than the currently proposed PAGER approach, which depends on distributing the building stock in any affected region into up to 80 vulnerability classes (Wald et al. 2010).

The following section discusses how a prototype approach for rapid global casualty estimation has been developed; Sect. 3 describes its testing against geographically distributed casualty data available for eight earthquakes in CEQID.

2 A prototype global casualty estimation model

2.1 Casualty model

In the model, a location, l, is a town, village or district within which the total population is known, or may be estimated, and over which the ground shaking intensity level is assumed constant within an estimated variation.

The estimated number of people at location l in building class i who are killed by a given earthquake event (K_{il}) is:

$$K_{il} = O_{il} * P_{il} * (d5_{il} * L5_i + d4_{il} * L4_i) \tag{1}$$

O_{il} is the average occupancy rate (i.e., the proportion of the normally resident population who are actually inside the building at the time of the event) in building class i and location l.

P_{il} is the total number of people normally resident in building class i in location l.

$d5_{il}$ and $d4_{il}$ are the proportions of the buildings of class i that collapse (damage level d5) and are heavily damaged (d4) respectively, in location l, given the assumed ground shaking intensity[1].

$L5_i$ and $L4_i$ are the lethality rates, i.e., the proportion of occupants killed, in buildings class i which collapse (d5) or are heavily damaged (d4) respectively.

Following Eq. 1, the total number of casualties, K, in the event across all building types and affected locations is:

$$K = \sum_{l} \sum_{i} [O_{il} * P_{il} * (d5_{il} * L5_i + d4_{il} * L4_i)] \tag{2}$$

[1] The damage levels d4, d5 used here have the same definitions as the equivalent damage grades (grade 4, grade 5, defined in EMS-98.

Note that the model does not include fatalities in buildings damaged at levels lower than d4, or fatalities occurring outside buildings.[2] It is also limited to shaking-induced casualties, so makes no attempt to estimate fatalities from earthquake-triggered events such as landslides, tsunamis and fire following.

The following sections describe how we estimate at location l: (1) the distribution of the normally resident population among the building classes P_{il}; (2) the occupancy rate at the time of the event O_{il}; (3) the ground shaking intensity; (4) the collapses and heavily damaged rates, $d5_{il}$ and $d4_{il}$; (5) the lethality rates by damage states, $L5_{il}$ and $L4_{il}$ and (6) the cumulative uncertainties of these variables.

2.2 Definition of building classes and distribution of the population among building classes

Given the aggregated uncertainties inherent in the many stages of this modelling process, the model adopts a simplified way to estimate the subdivision of buildings in an area among vulnerability classes. A total of five vulnerability classes (A to E) are assumed to be sufficient to capture the principal variations in collapse rates and lethalities. The classes and the building typologies generally assumed to belong to these classes are shown in Table 1. It is assumed that each of these classes is characterised by a unique vulnerability function (in terms of damage for a given shaking intensity) with its associated uncertainty, and the grouping of building typologies in each class is organised on the basis of observed performance. The classes in Table 1 are somewhat similar but not identical to those proposed in EMS-98[3] (Grünthal 1998). It is important to note that the proposed methodology requires the user to allocate classes defined by particular building typologies into the right vulnerability classes.

A further assumption is that the distribution of the building stock among these five classes in any location can be modelled using a binomial distribution, defined by a single binomial parameter p, which adequately characterises the distribution of vulnerabilities at that location. Thus the proportion of buildings in class i, given p, is defined by

$$\mathrm{PrB_i} = (4!/((i!)\,(4-i)!)) * p^i * (1-p)^{4-i} \quad \text{for } 0 \le i \le 4 \tag{3}$$

where i ranges from 0 to 4, representing vulnerability classes A–E. These are ordered from most vulnerable to least vulnerable.

For trial applications of the model described in Sect. 3, best-fit values of the binomial parameter p have been derived from building data recorded in CEQID, using a least squares procedure. The procedure adopted makes use of the building classification system of post-event damage surveys in the area concerned, and to reclassify these buildings according to vulnerability classes A to E as shown in Table 1. The value of p used was that which minimised the estimation error across all five classes, and these values are shown in Table 2. The table also shows the distribution of the building stock between the five vulnerability classes from the actual data and that derived from the best-fit value of p.

In reality, building stock distributions do not always closely follow a binomial distribution, but it is shown later in Table 4 that the additional error in the calculated collapse rates from making this assumption is small compared with the overall uncertainty in the model. For use in Eq. 1, however we need the distribution of the population by building class, which further depends on the numbers of occupants of the different building classes.

[2] Evidence from recent studies (So 2009; Petal 2004; Koyama et al. 2011), indicates that fatalities from buildings damaged at damage level d3 or below form a very small proportion of the total.

[3] The description "Structural masonry" (vulnerability class C) is taken to include both "load-bearing masonry with RC floors", and reinforced or confined masonry" as defined in EMS-98. EMS-98 includes a sixth class, F representing buildings of very high earthquake resistance; these buildings are very rare.

Table 2 Modelled and observed building stock distributions associated with the events analysed

Year	Event		Derived p value	Proportion of buildings in different vulnerability classes				
				A	B	C	D	E
1995	Kobe	Data		0.00	0.12	0.00	0.58	0.31
		Modelled	0.78	0.00	0.03	0.18	0.42	0.37
2006	Yogyakarta	Data		0.59	0.15	0.26	0.00	0.00
		Modelled	0.11	0.63	0.31	0.06	0.00	0.00
1999	Kocaeli	Data		0.00	0.04	0.93	0.03	0.00
		Modelled	0.5	0.06	0.25	0.38	0.25	0.06
1980	Irpinia	Data		0.55	0.14	0.31	0.00	0.00
		Modelled	0.08	0.72	0.25	0.03	0.00	0.00
2005	Kashmir	Data		0.65	0.34	0.01	0.00	0.00
		Modelled	0.1	0.66	0.29	0.05	0.00	0.00
1999	ChiChi	Data		0.00	0.28	0.63	0.09	0.00
		Modelled	0.46	0.09	0.29	0.37	0.21	0.04
2001	Bhuj	Data		0.44	0.51	0.03	0.03	0.00
		Modelled	0.18	0.45	0.40	0.13	0.02	0.00
1993	Latur	Data		0.90	0.10	0.00	0.00	0.00
		Modelled	0.01	0.96	0.04	0.00	0.00	0.00

If P is the total population and P_i is the total population in building class i in the area considered, then

$$\begin{aligned} P_i &= NB_i * PpB_i \\ &= NB * PrB_i * PpB_i \\ &= NB * PrB_i * PpBR_i * PpB_0 \end{aligned}$$

and

$$P = NB * PpB_0 \left(\sum_i \left(PrB_i * PpBR_i \right) \right) \tag{4}$$

where NB is the total number of buildings, and NB_i is the number of buildings in building class i;

PrB_i is the proportion of buildings in class i (i.e., NB_i/NB);

PpB_i is the average population per building in class i, and PpB_0 is the average population per building in a single family dwelling in the area;

and $PpBR_i$ is the average occupants per building in class i as a ratio of the average number of residents per building in a single family dwelling in the area.

Thus the proportion of the population in building class i is:

$$P_i/P = PrB_i * PpBR_i / \sum_i \left(PrB_i * PpBR_i \right) \tag{5}$$

Equation 5 is independent of both the total number of buildings and the average population per building in a single-family dwelling. To determine the vector $PpBR_i$ across the building classes it is necessary to assume or have some knowledge of the proportions of multi-family housing (or non-residential buildings) among the buildings of that class in the area affected.

2.3 Occupancy rates at the time of the event

It is well-established that instantaneous occupancy rates vary according to time of day and whether a building is residential or commercial (i.e., primarily a place of work) (e.g., Coburn and Spence 2002). The occupancy rate can also be affected by the pattern of ground shaking; in some cases initial low-intensity shaking provides the opportunity for residents to leave before the strong shaking precipitating collapse occurs (So 2009).

2.4 Intensity of ground shaking at a location

The model described is intended for use in conjunction with estimates of shaking intensity derived from USGS ShakeMaps. The ShakeMap provides, using ground motion attenuation relationships, an estimate of MMI intensity for any point location, with decimal subdivision to 0.01 intensity units (Allen et al. 2008), although the presented method does not justify this level of precision. In the case of this application, if the location is a town or district, the intensity used can be taken at the middle of the principal population centre, or if the population is dispersed, an average of intensity values across the local populated area can be used.

2.5 Collapse and heavy damage rates

2.5.1 Defining damage levels

To determine the relationships between ground shaking intensity and the probabilities of different damage states, damage data assembled in CEQID have been analysed. CEQID contains damage data from surveys in more than 600 locations in 53 earthquakes and includes more than 1.5 million buildings (Spence and So 2010). For each of these locations, a ShakeMap MMI level has been determined using the ShakeMap archive, and further analyses of the damage data have been performed, grouping each building class into one of the five classes shown in Table 1.

For each building class the best-fit relationship to a cumulative normal distribution has been determined, with estimated uncertainty parameters, $\varepsilon_1, \varepsilon_2$. The probability P that the damage level d exceeding damage level d4 or d5 at location $\mathrm{I_l}$ is given by:

$$P(\mathrm{d} = \mathrm{d5} \mid \mathrm{I_l}) = \boldsymbol{\Phi}(\alpha_\mathrm{i} * \mathrm{I_l} + \beta_\mathrm{i}) + \varepsilon_1 \tag{6}$$

$$P(\mathrm{d} \geq \mathrm{d4} \mid \mathrm{I_l}) = \boldsymbol{\Phi}(\alpha_\mathrm{i} * \mathrm{I_l} + \delta_\mathrm{i}) + \varepsilon_2 \tag{7}$$

$\boldsymbol{\Phi}$ is the standard normal cumulative distribution function. $\mathrm{I_l}$ is the ground shaking intensity (in MMI units from the USGS ShakeMap) at location l. α_i, β_i, and δ_i are parameters for the given building class determined by regression analysis.

Since vulnerability class D contains buildings with widely different lethality potential (RC frame and timber frame), two separate subclasses, D1 and D2, have been defined in this case. In each location, or earthquake, the vulnerability parameters chosen for Class D will be those appropriate for the type of buildings of class D (i.e., D1 or D2) commonly found in the area. Thus fixed values of the three parameters α_i, β_i, and δ_i are applied to buildings of each class A, B, C, D1, D2 and E and Table 3 shows the values of the constants determined

Table 3 Values of parameters α, β, and δ defining proportions of buildings at collapse (d5) and heavy damage (d4) levels, and lethality rates (L4, L5) at each damage level, for all building vulnerability classes

Class	d5			d4	
	α	β	L5	δ	L4
A	0.178	−3.087	0.200	−2.110	0.0500
B	0.496	−5.976	0.078	−5.108	0.0195
C	0.297	−4.483	0.250	−3.932	0.0625
D1	0.419	−5.400	0.250	−5.000	0.0625
D2	0.260	−5.211	0.013	−3.759	0.0034
E	0.505	−6.256	0.278	−5.859	0.0695

Table 4 The effect of categorisation of entire building stocks with a single p value and binomial distribution assumption on the estimation of collapse rates; calculated for both intensity MMI = 8 and MMI = 9

Event	Error in collapse rate at I = 8 (%)	Error in collapse rate at I = 9 (%)
Kobe	6	10
Yogyakarta	−5	−9
Kocaeli	−15	−25
Irpinia	−14	−13
Kashmir	0	2
ChiChi	−13	−11
Bhuj	1	5
Latur	−3	0
Average (ratio)	−5	−5
RMS error	9	12

for each building class. ε assumed in Eqs. 6 and 7 are based on the standard errors in the best-fit regression relationships derived from the underlying data.

The effect of categorisation of entire building stocks with a single p value and binomial distribution assumption on the estimation of collapse rates has been calculated for both intensity MMI = 8 and MMI = 9. The average root mean square (RMS) error in estimation was found to be 9 % at MMI = 8 and 12 % at MMI = 9, as shown in Table 4. Approximate data on the breakdown of the same building stock data by building height have been used to estimate the PpBR values appropriate to these locations, and these are later shown in Table 5.

2.5.2 Comparison with collapse rates from other studies

Figure 1 compares the mean values of collapse rates (d5) proposed by Lagomarsino and Giovinazzi (2006) for their Classes A, B and C, (here labeled A*, B* and C*) and those derived in the analysis described above from Eqs. 6 and 7. There are inevitably some differences in definition of the classes, and possibly also of interpretation of the damage levels,

Table 5 Assumed and calculated standard deviations, variance and summation to estimate variance of the estimated total number killed

Parameter	Source of data	Typical values	Assumed and calculated standard deviation of lognormal
Population	LandScan	1,000–100,000	0.20
Instantaneous occupancy, PpBR	Judgement	0.3–0.9	0.25
Building stock distribution (p)	CEQID analysis	0.1–0.8	0.20
Population per building ratio	CEQID analysis	1–15	0.20
Collapse rate (including uncertainty in MMI from PGA/PGV)	CEQID analysis	1–7 %	1.04
Lethality given collapse	CEQID/Judgement	0.1–0.3	0.95
Total			1.47

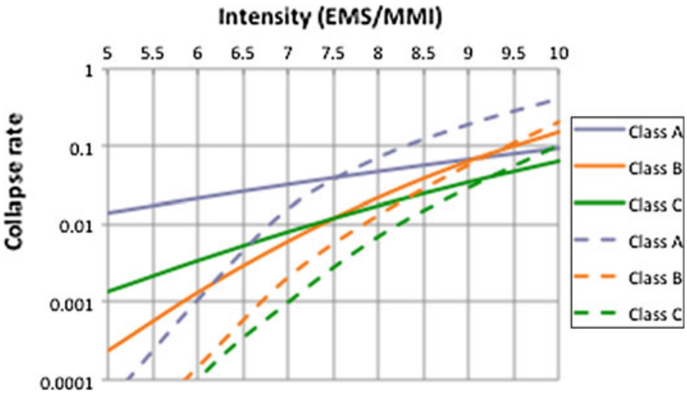

Fig. 1 Comparison between collapse rates as a function of ground shaking intensity derived in this study (Classes A, B and C) and the nearest equivalent classes derived by Lagomarsino and Giovinazzi (2006) (labeled A*, B* and C*)

since the RiskUE[4] study was for European building types only and the curves shown are derived from the EMS-98 Intensity Scale definitions.

The estimates of collapse rates determined using the parameters in Table 3, while similar in the range 7 < MMI < 8.5, are considerably lower at higher intensities than those proposed in the RiskUE project and are higher at lower intensities. Reasons for these differences need further investigation, but it is speculated that this may, in part, stem from the use of USGS estimated values for MMI, rather than intensities derived solely from local damage observations where assessments will tend to define an area of higher intensity where there

[4] The European RISK-UE project was launched in 1999, at the end of the International Decade for Natural Disaster Reduction (IDNDR). The project started in January 2001 and ended in September 2004. The project itself involved the assessment of earthquake scenarios based on the analysis of the global impact of one or more plausible earthquakes at city scale, within a European context.

is a concentration of local damage. However, since in this model the intensities used are to be derived from the USGS ShakeMap, the damage functions derived from the ShakeMap estimated intensities are the most appropriate for this empirical analysis.

Collapse rates derived from the WHE/EERI expert-elicitation study are even higher than those proposed by the Risk-UE study, suggesting a tendency among the experts to overestimate collapse potential of a building class at a given ground shaking (Pomonis et al. 2011).

2.6 Lethality rates for each building class

Lethality rates used in this model are empirically derived. A new set of general lethality rates relating to building collapse was presented by Spence (2007) based on recent studies. The LessLoss study reviewed fatality data for typical building types in three European cities and presented injury distributions for collapsed timber, steel, reinforced concrete and masonry buildings. These data have been applied to this model for use with the six vulnerability classes adopted (Table 3), where again separate classes, D1 and D2, have been defined to deal with different lethality potential for two types of structure. In each location, or earthquake, the lethality rate chosen for class D will be that appropriate for the type of buildings of class D (i.e., D1 or D2) most commonly found in the area. Since no lethality data are available for heavily damaged buildings, lethality rates at damage level d4 are taken as 25 % of those for d5 (Spence 2007). Uncertainties in lethality rates are generally large and are estimated based on the empirical data available.

2.7 Cumulative uncertainties

Based on Eqs. 1–7, it is clear that the estimated error in numbers of people killed in a location, e_K will be some complex function of the errors in the estimates of the total population e_P; the occupancy rate e_O; the parameter p defining the building stock distribution, e_p; the collapse or heavy damage rate given ground shaking intensity e_{coll}; the ground shaking intensity e_I; and the lethality rate e_L. Thus

$$e_K = f\left(e_P, e_O, e_p, e_{coll}, e_I, e_L\right) \tag{8}$$

A simplified way to estimate the cumulative error is to assume all these components are lognormal in form and that all variables are independent, in which case we can estimate:

$$\sigma_k = \sqrt{\left(\sum \sigma_x^2\right)} \tag{9}$$

where σ_k is the standard deviation of the error in the estimation of the number killed, and σ_x is the standard deviation of the individual components of Eq. 2. Table 5 shows the estimated standard deviations of the individual components from which an estimated cumulative standard deviation is derived. A series of numerical simulations were used to assess the cumulative uncertainty in the collapse rate, based on known uncertainties in ground motion estimation using Boore et al. (1997), its conversion to MMI using Wald et al. (1999), and collapse rate estimations from MMI intensity using Eqs. 6 and 7. The final cumulative uncertainty is dominated by the uncertainties in the collapse and lethality rates for the building classes.

This is an approximate approach to addressing uncertainty (and ignores correlation between the separate variables); further work will entail a more rigorous assessment of the resulting combined uncertainty. However, this approximation fits the spread of empirical data fairly well (see Figs. 2, 3, 4 and 5).

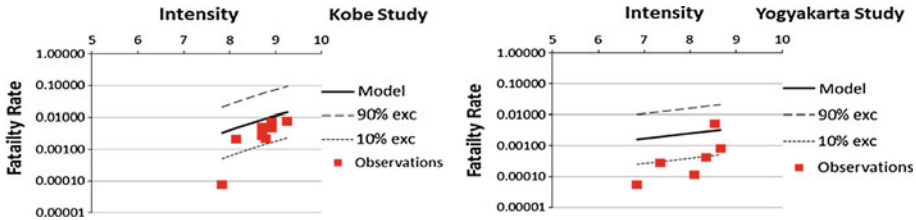

Fig. 2 Reported and estimated death rates and expected uncertainties (10 and 90 % exceedence) for 9 districts in the Kobe 1995 earthquake (*left*) and 6 districts in the Yogyakarta 2006 earthquake (*right*)

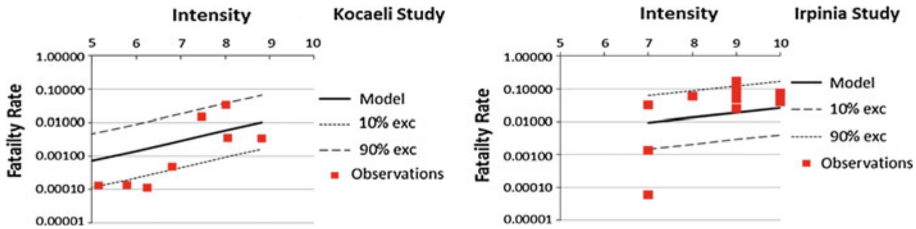

Fig. 3 Reported and modeled fatality rates and expected uncertainties for 10 municipalities in the Kocaeli 1999 earthquake (*left*), and 14 municipalities in the Irpinia 1980 earthquake (*right*)

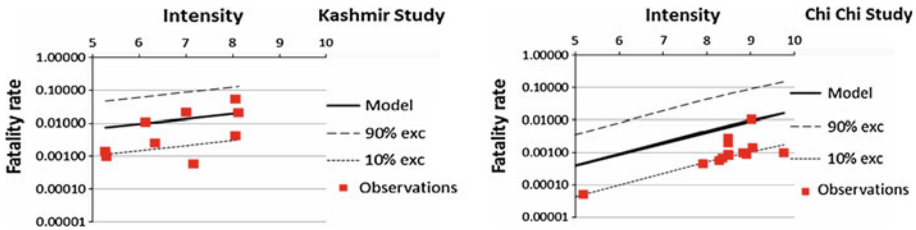

Fig. 4 Reported and estimated deaths and expected uncertainties (10 and 90 % exceedence) in 9 districts in the Kashmir 2005 earthquake (*left*) and in 15 provinces in the Chi Chi 1999 earthquake (*right*)

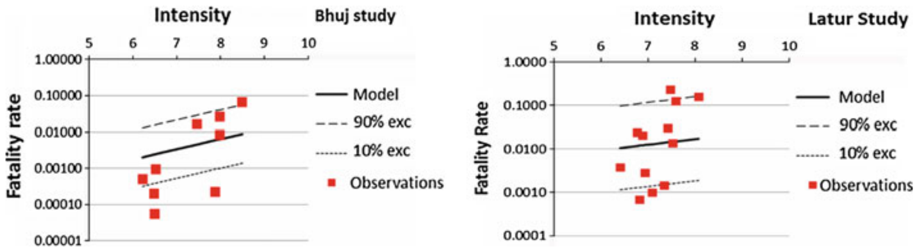

Fig. 5 Reported and estimated deaths and expected uncertainties (10 and 90 % exceedence) in 10 districts in the Bhuj 2001 earthquake (*left*) and in 12 zones in the Latur 1993 earthquake (*right*)

3 Testing the model against reported casualty data

The approach described above was tested against data on reported deaths from the eight earthquakes shown in Table 6. The building distributions from these surveys were used to derive p values for the analyses but the fatality data had not been used. For all of these events, fatality data were available in CEQID on a geographically distributed basis, by village or district.

Table 6 Parameters used in the application of the model for the eight earthquakes (for relative occupancy NR indicates classes assumed to have insignificant numbers)

Event	Day and time of day	Number of districts	ShakeMap intensity range	Derived p value	Assumed instant. Occ'cy	PpBR Classes A, B	PpBR Class C	PpBR Class D	PpBR Class E
Kobe, 1995	Tuesday, 05:46	9	8.1–9.3	0.78	0.9	1	1	1	5
Yogyakarta, 2006	Saturday, 05:54	6	6.8–8.5	0.11	0.3	1	1	1	NR
Kocaeli, 1999	Tuesday, 03:02	10	4.3–8.8	0.5	0.9	1	5	10	5
Irpinia, 1980	Sunday, 19:35	13	7.0–10.0	0.08	0.9	1	2	2	2
Kashmir, 2005	Saturday, 08:52	9	5.3–8.1	0.1	0.9	1	1	1	NR
Chichi, 1999	Tuesday, 01:47	15	4.7–9.7	0.46	0.9	1	5	10	10
Bhuj, 2001	India's 52nd Republic Day, 08:46	10	6.2–8.5	0.18	0.9	1	1	5	5
Latur, 1993	Thursday, 03:56	12	6.4–8.1	0.01	0.9	1	1	NR	NR

In most cases, data were collected and published by national studies, though the accuracy of these data cannot be ascertained. For each event a USGS ShakeMap was available, providing the ground shaking intensity estimates for each location. In addition, for each location an estimated or official total population exposed was available, enabling a fatality rate to be determined for each location, and aggregated across all locations.

For each event in Table 6, Figures 2 through 5 plot the reported fatality rate against the USGS ShakeMap intensities at several locations. On each graph, the expected fatality rate has also been plotted using Eq. 2 as a continuous function over the total population at each intensity level. Values of p, instantaneous occupancy rates used and assumed values of occupancy distribution between vulnerability classes, PpBR, are shown in Table 6. The lethality rates at damage states d4 and d5 used are shown in Table 3. Estimated 90 and 10 % exceedence probability values of fatality rates are also plotted using the estimated cumulative uncertainties shown in Table 5.

Instantaneous occupancy is one of the most difficult parameters to determine. For this application a default value of 0.9 was used in all cases, except for the Yogyakarta event for which a better estimate of the daytime occupancy of 0.3 was known from field studies (So 2009).

The comparative plots shown in the figures above indicate that there is a wide scatter of fatality rates, but nevertheless for all events there is a tendency to higher fatality rates at higher intensities, as would be expected. Most of the observed fatality rates from the different locations are captured between the exceedence bounds. In total, 17 out of 84, i.e., 20 % of the data points fell outside the 10 and 90 % exceedence levels suggesting that the uncertainty bands are adequate, although only two of these were above the 10 % exceedence level, and 15 were below the 90 % exceedence level.

For each event, the total numbers of deaths estimated and reported across all the study locations have been calculated, and the ratios of reported deaths to deaths expected from the model for the event have been determined. This ratio is a measure of the overall error in the modelling accuracy. Table 7 shows these ratios.

In spite of the wide variations in fatality numbers in district-by-district location, for six of the eight events the overall number of deaths calculated by the model was within a factor of 2.5 of the reported values, and in the worst case, for Chi Chi earthquake was about five times the reported deaths. The correlation coefficients for fatality rates versus intensity are positive and greater than 48 % in all events except Chi Chi (−5 %). It is worth noting that in the Chi Chi case alone there was no correlation between the observed death rate and ShakeMap intensity (i.e., death rates do not increase with the increase in intensity levels), which suggests that there may be some systematic error in the reported deaths or exposed populations for that event.

For some events additional explanations for the differences can be identified. In the Yogyakarta earthquake it was found from the surveys that many of the occupants had moved outside their homes at the moment of the main shock, leading to the lower than expected fatality rate (reported/modelled = 0.46). While in the Kashmir, Pakistan event, the higher than expected fatality rate (reported/modelled = 1.12) can possibly be attributed to the extensive ground failures, which enhanced building collapse rates in many of the locations worst affected. In the Bhuj event the overall casualty estimate shown above was revised downwards by the Gujarat State Disaster Management Agency to 12,221 in 2002 (but without giving revised district breakdowns). This would have reduced the reported/estimated ratio to 1.3. What has not been modeled in this study are influence of distance and intensity ranges. In future work, the authors intend to carry out additional sensitivity studies to test the effects of these parameters in the model.

Table 7 Comparison of observed and expected/ modelled deaths in the total population for the 8 earthquakes studied

Event	Number of districts	ShakeMap intensity range	Total population	Reported deaths	Correlation reported fatality rate versus intensity	Modelled deaths	Ratio reported/modelled
Kobe	9	8.1–9.3	1,578,718	6,030	0.87	14,007	0.43
Yogyakarta	6	6.8–8.5	4,700,254	5,760	0.48	12,608	0.46
Kocaeli	10	4.3–8.8	15,125,288	15,227	0.51	33,561	0.45
Irpinia	13	7.0–10.0	92,396	2,176	0.51	1,239	1.76
Kashmir	9	5.3–8.1	4,796,208	72,695	0.57	64,814	1.12
Chichi	15	4.7–9.7	740,178	990	−0.05	4,725	0.21
Bhuj	10	6.2–8.5	1,284,507	18,403	0.60	9,359	1.97
Latur	12	6.4–8.1	139,811	6,255	0.63	1,853	3.37

4 Comments on the proposed methodology and further work

The method proposed is simple and approximate. It is an attractive alternative as the approach uses a single parameter to define the building distribution in the area of interest. It is intended to provide an estimate of deaths from ground shaking to within about a factor of 3 or 4 in most situations, using only crude information about the quality of the building stock affected. The predictions from this model were compared with reported data from eight earthquake events. A wide scatter of fatality rates was observed when plotted for separate locations against local intensity, but there is in nearly all cases a tendency for fatality rates to increase with increasing estimated ground shaking. In most cases the modelled fatalities summed across the affected area are within a factor of 2.5 of those reported, and in the worst case, it was within a factor of five. In cases where there are the largest discrepancies, these can be accounted for by particular factors affecting estimated collapse rates or occupancy at the time of the event.

There are obvious limitations to this approach. Deaths and injuries outside a building or collateral hazards such as landslides, fires and tsunamis are not included. As shown in recent events in Christchurch in New Zealand and Sendai in Japan, these numbers can be significant. The estimates of ground shaking intensity are subject to all the limitations of the ShakeMap intensity estimates as discussed by Wald et al. (2008), with additional errors from the assignment of a single intensity value to some district sized areas. Likewise the LandScan-based population estimates have limitations inherent in that dataset. Assuming the building stock can be classified with five classes and follows a binomial distribution introduces some further uncertainty, which has been shown to be comparatively small. Collapse rates for the given classes at each intensity level are based on extensive damage data analysis using ShakeMap intensities; however they do not take account of impending revisions to the ShakeMap Atlas (Worden et al. 2010). Lethality rates have used a single rate for each building class across all locations: a more detailed assessment of lethality rates is in progress. Occupancy rates are assumed, again, based on limited data. Moreover, the model has been tested so far against data from only a few earthquakes with a limited range of building types, and with ground shaking intensities reaching MMI = 10 in only 2 cases. A systematic study of the uncertainties in this procedure has yet to be carried out. This will be a key aspect of future planned improvements.

Much remains to be done to build an operationally robust casualty estimation method from the approach adopted, but this initial study suggests that the approach has potential value, in enabling the quality of the local building stock, occupancy factors and lethality rates to be specifically included in the model. Mapping of the parameter p for national and local building stocks will be an essential element of the process. This can be done using available building survey data, or based on the USGS inventory data (Jaiswal and Wald 2008). Alternatively, it could be done based on damage survey data derived from the CEQID for different localities, taking into account actual earthquake performance.

A future version of this model will also differentiate classes of masonry based on their roof type: light roof or heavy roof. The weight of roofing materials makes a significant difference in casualty potential but this is not yet captured in any of the existing models. This was shown in field data collected from Tabas, Bam and, Muzaffarabad where collapsed class A buildings with heavy roofs were responsible for the deaths of up to 25 % of a village's population. Much lower levels were evident where buildings had light roofs (So 2009).

4.1 In closing

With an ever changing building environment, particularly in developing countries, it is of prime importance to capture and model the impact of the damage to buildings on their occupants in rapid loss estimation models. Currently, the available operational loss models such as PAGER compute losses directly from relationships between ground motion and fatalities. Furthermore, the instantaneous occupancy rates using working/ school hours, transit and non-work hours are being explored within the PAGER semi-empirical model (Jaiswal and Wald 2010).

The model described here is preliminary and needs further development in many respects, but it offers a simple calculator for losses based on direct empirical data. The approach is an attractive alternative for future casualty modelling.

Acknowledgments Our thanks go to National Earthquake Hazards Reduction Program (NEHRP) for funding the initial part of this research and to the USGS PAGER team for the continual collaboration and support.

References

Allen TI, Wald DJ, Hotovec AJ, Lin K, Earle PS, Marano KD (2008) An Atlas of ShakeMaps for selected global earthquakes. U.S. Geological Survey Open-File Report 2008-1236, 47 p
Boore DM, Joyner WB, Fumal TE (1997) Equations for estimating horizontal response spectra and peak acceleration from Western North American earthquakes: a summary of recent work. Seismol Res Lett 68/1:128–153
Coburn A, Spence R (2002) Earthquake protection. Wiley, London
Dobson JE, Bright EA, Coleman PR, Bhaduri BL (2003) LandScan: a global population database for estimating populations at risk. In: Mesev V (ed) Remotely Sensed Cities. Taylor & Francis, London, pp 267–281
GEM, Global Earthquake Model. http://www.globalquakemodel.org/. Accessed 21 Feb 2011
Grünthal G (ed) (1998) European macroseismic scale 1998. Council of Europe, Luxembourg
Jaiswal KS, Wald DJ (2008) Creating a global building inventory for earthquake loss assessment and risk management. U.S.G.S. Open File Report 2008-1160
Jaiswal KS, Wald DJ (2010) Development of a semi-empirical loss model within the USGS prompt assessment of global earthquakes for response (PAGER) system. In: Proceedings of the 9th US and 10th Canadian conference on earthquake engineering: reaching beyond borders, 25–29 July 2010, Toronto, Canada
Jaiswal KS, Wald DJ, Hearne M (2009) Estimating casualties for large worldwide earthquakes using an empirical approach. U.S. Geological Survey Open-File Report 2009-1136
Jaiswal K, Wald D, Earle P, Porter K, Hearne M (2011) Earthquake casualty models within the USGS prompt assessment of global earthquakes for response (PAGER) system, chapter 6. In: Spence R, So E, Scawthorn C (eds) Human casualties in earthquakes: progress in modelling and mitigation. Springer, Berlin
Koyama M, Okada S, Ohta Y (2011) Major factors controlling earthquake casualties as revealed via a diversified questionnaire survey in Ojiya City for the 2004 mid-Niigata earthquake, chapter 14. In: Spence R, So E, Scawthorn C (eds) Human casualties in earthquakes. Springer, Berlin
Lagomarsino S, Giovinazzi S (2006) Macroseismic and mechanical models for the vulnerability assessment of current buildings. Bull Earthq Eng, Special Issue “Risk-Ue Project” 4(4)
National Institute of Building Sciences-Federal Emergency Management Agency (NIBS-FEMA) (2006) HAZUS-MH MR2 technical manual
Petal MA (2004) Urban disaster mitigation and preparedness: the 1999 Kocaeli earthquake. PhD dissertation, University of California, Los Angeles
Pomonis A, Kappos A, Panagopoulos G, Karababa F (2011) Seismic vulnerability and collapse probability assessment for Greek buildings, chapter 11. In: Spence R, So E, Scawthorn C (eds) Human casualties in earthquakes: progress in modelling and mitigation. Springer, Berlin
Porter KA (2009) Cracking an open safe: HAZUS vulnerability functions in terms of structure—independent spectral acceleration. Earthq Spectra 25(2):361–378
Spence R (ed) (2007) Human losses, chapter 5 in earthquake disaster scenario prediction and loss modeling for urban areas. LessLoss report no. 2007/07, pp 51–58

Spence RJS, So E (2010) A global earthquake building damage and casualty model, chapter 5. In: Spence R, So E, Scawthorn C (eds) Human casualties in earthquakes: progress in modelling and mitigation. Springer, Berlin

So E (2009) The assessment of casualties for earthquake loss estimation. PhD thesis, University of Cambridge, Cambridge

Trendafiloski G, Wyss M, Rosset P (2011) Loss estimation module in the second generation software QLARM, chapter 7. In: Spence R, So E, Scawthorn C (eds) Human casualties in earthquakes: progress in modelling and mitigation. Springer, Berlin

Wald DJ, Quitoriano V, Heaton TH, Kanamori H (1999) Relationships between peak ground acceleration, peak ground velocity and modified Mercalli intensity in California. Earthq Spectra 15/3:557–565

Wald DJ, Lin K, Quitoriano V (2008) Quantifying and qualifying ShakeMap uncertainty. U.S. Geological Survey Open-File Report 2008-1238, 27 p

Wald DJ, Jaiswal KS, Marano KD, Bausch, DB, Hearne MG (2010) PAGER—Rapid assessment of an earthquake's impact. U.S. Geological Survey Fact Sheet 2010–3036, 4p

WHE, World Housing Encyclopedia. www.world-housing.net. Accessed 10 Oct 2010

Worden CB, Wald DJ, Allen TI, Lin K, Garcia D, Cua G (2010) A revised ground-motion and intensity interpolation scheme for ShakeMap. Bull Seismol Soc Am 100(6):3083–3096