Australian Earthquake Engineering Society 2014 Conference, Nov 21-23, Lorne, Vic

The production of a robust worldwide rapid socio-economic loss model for earthquake economic loss and fatality estimation: success from 2009-2014

James Daniell[1,2] & Friedemann Wenzel[2]
[1] John Monash Scholar, The General Sir John Monash Foundation, Level 5,
30 Collins Street, Melbourne, Victoria, Australia 3000
[2] Geophysical Institute & Center for Disaster Management and Risk Reduction Technology (CEDIM), Karlsruhe Institute of Technology (KIT), Hertzstrasse 16, 76187, Karlsruhe, Germany
.

E-Mail: j.e.daniell@gmail.com

ABSTRACT:

An empirical global earthquake loss model is presented based on empirical data from 1900-2008. Over 7300 damaging earthquakes have occurred since 1900 via the CATDAT Damaging Earthquakes Database, occurring in different socioeconomic situations. Disaggregated fatalities and economic losses and costs from earthquakes as well as secondary effects (landslide, liquefaction, tsunami) have been collected for each event.

A virtual earth was created from 1900-2014 for 255 countries, taking into account border changes, population, building stock, economic production, earthquake vulnerability and socioeconomic indices such as development. Collected intensity maps from 3000+ damaging earthquakes were then superimposed on this exposure to recreate each historical earthquake. 450 seismic resistant codes were also reviewed in order to provide a global seismic code influence index.

The result of this geophysical, engineering, and socio-economic analysis was the production of a set of socioeconomic fragility functions to accurately determine the economic loss and cost and number of fatalities from a damaging earthquake anywhere in the world, within minutes and hours after an event (J. E. Daniell's Doctoral Thesis). This was tested for the period of 2009-2014 with 100% success for economic losses and 98% for fatalities and used on www.earthquake-report.com and can also be used for future risk assessment.

Keywords: Economic losses, earthquake loss assessment, global risk, fatality estimation, insurance.

1. Introduction

As soon as possible after an earthquake, it is useful to predict an approximate number of deaths and fatalities. Over the past few years many efforts have been made to create a loss estimation procedure that works globally to produce results. In addition, quite a few other rapid regional and local scale loss estimation procedures exist.

Table 1: A synopsis of the components of open source worldwide ELE software packages with respect to rapid loss estimation and use in this study.

Type of Software Package	**Acc.**	**Open/ Closed**	**Lang.**	**Tested Locations**	**Rapid Loss**	**Exposure**	**Vuln.**	**Hazard**	**Social**	**Economic**
ELER	**YES**	**O**	Matlab	Turkey, Europe	**Yes**	B,P,Es	Both	MM, Sp	Ss, Auto	None
EQRM	**YES**	**O**	Python	Australia adapted	**No**	B,P,Es	Both	Sp	Ss	DC
ELER has 3 versions – Level 0, Level 1 and Level 2										
CATDAT-EQLIPSE (this study)	**YES**	**C**	Matlab, GIS	Global	**Yes**	P, Es	Emp	MM	Ss	DC
EQLIPSE has 2 versions – Q and R										
Extremum	**NO**	**C**	Win, GIS	Turkey etc.	**Yes**	B,P,Es	Emp	MM	Ss Auto	RE
HAZ-Taiwan (TELES)	**NO**	**C**	C++ and MapInfo	Taiwan	**Yes**	B,I,P, Es	Anl.	Sp, SE	Ss Auto Sc	DC, IO
HAZUS-MH	**YES**	**O**	VB6, C++, Arc	USA	**Yes ***	B,I,P, Ec	Anl.	Sp, SE	Ss, Sc	DC, IO
InaSAFE Earthquake	**YES**	**O**	Java, QGIS	Jakarta	**Yes ***	B,I,P, Es	Both	MM, Sp, SE	Sc	DC
PAGER	**NO**	**C**	Matlab, unknown.	Worldwide	**Yes**	B, P, Es	Both	Sp, MM, SE	Ss Auto	Many DC, RE
PAGER has 3 different versions – Empirical, Semi-Empirical and Analytical (Jaiswal et al. (2011))										
QLARM	**NO**	**C**	Java	Worldwide	**Yes**	B, P, Es	Both	MM	Ss Auto	No
REDARS	**YES**	**C**	GUI Windows, Basic	California	**Yes**	I,Es	Emp	Sp, SE	No	DC, IO
REDAS	**NO**	**C**	GUI Windows, Basic	Philippines	**Yes**	B,P,I, Es	Emp	MM, SE	Ss	DC
SAFER	**NO**	**C**	Same as SELENA	European Settings	**Yes**	B,P,I, Es	Both	Sp	Ss	DC
SELENA	**YES**	**O**	Matlab, C++	Oslo	**Yes**	B,P,Es	Anl.	Sp	Ss	DC
SES2002 and ESCENARIS	**NO**	**C**	VB	Spain	**Yes**	B,I,P, Es	Emp	MM	Ss, Sc	DC, RE
SIGE/ESPAS	**NO**	**C**	VB	Italy	**Yes**	B,I,P, Es	Emp	MM	Ss, Sc	DC, RE
PLINIVS	**NO**	**C**	Unk. – DPMs	Naples	**Yes**	B,P	Emp	MM	Ss	No
QuakeIST	**NO**	**C**	C++	Lorca, Faial, Iceland	**Yes**	B,I,P	Both	Sp, MM	Ss, DI	DC, RE

IERRS (Istanbul), SUPREME (Tokyo), READY (Yokohama) and AFAD-RED (Turkey) are additional rapid loss models that have not been reviewed, as there are not many details on the methodology.

O/C – open or closed software package; VB = Visual Basic, Arc=ESRI ArcGIS or similar, *= as shown below.

Exp =Exposure, B=Buildings, I=Infrastructure, P=Population, Es=economic values for infrastructure, Ec=complex economic values using regional or location assessment values
Haz =Hazard, MM=intensity, Sp=spectra, SE=secondary effects
Vuln =Vulnerability Type, Anl.=Analytical, Emp.=Empirical
Soc. =Social, Ss=simple social analysis, Sc=complex social analysis, US=user inputted curves or assessment, Auto=automatic analysis for rapid loss estimation, DI=Disruption Index
Econ. = Economic, DC= direct conversion to replacement cost, RE= rapid loss estimation possible, IO= indirect and additional analysis.

From these different software packages, a variety of results have been obtained with them and these were examined in Daniell et al. (2011). The key global methods are the PAGER system and the QLARM System, as explained in the aforementioned table for global fatality and economic loss quantification.

The overall goal of this study was to develop a dynamic approach to rapidly and accurately calculate a range of fatalities and economic losses from earthquakes anywhere in the world using the input of an intensity based map and historical earthquakes as a proxy over multiple temporal and spatial scales. This required the development of socio-economic indicators through time to scale historical earthquakes, with these indicators being able to change as more earthquakes occur.

The focus that the approach should be dynamic and use available data worldwide was a key concept of this goal. Temporal relationships of socio-economic losses were explored in order to calibrate loss functions not just based on a relationship of intensity to population. Similarly, exploring post-earthquake building damage, casualty and economic loss functions was required. The method created is a fast approach for calculating losses, given an event on the spatial scale of the population centroids, or any inputted population grid. It works on the premise of functions that have been derived from historical losses and then adjusted objectively and subjectively to account for changes worldwide in the socio-economic climate.

This study focusses on the vulnerability conversion of economic losses and fatalities, and does not address the problem of hazard in rapid loss estimation, except for the use of intensity maps in historical earthquakes (although other methods have been implemented in other studies by the author). The methodology is data intensive, requiring much global data for analysis; on the other hand, it requires much less data than other global systems, relying on the premise that only having a few components will increase robustness and accuracy compared to a very complex system.

As described in the paper of Daniell et al. (2011) at the AEES, the methodology uses MMI from historical earthquakes as a proxy for creating a loss estimate. There have been over 200 casualty functions globally used over the last 60 years. Some range from a study for single events, to full global loss methodologies. Historically, there have been many key fatality-intensity methodologies used (Figure 1), with the first key function being that of Haihua (1987) for various Chinese earthquake events. Tiedemann (1989) looked a single relation, given different vulnerabilities. QLARM/EXTREMUM looked at global fatalities using intensity, vulnerability and fatality ratio matrices in the 1990s and 2000s. Similarly, the work of Christoskov, Badal and Samardzhieva (1992) in these two decades also looks at a similar methodology. Currently, 2 main methodologies, ELER (2007) and USGS-PAGER through Jaiswal et al. (2009) and Jaiswal and Wald (2010), are shown using global historical fatalities from 1973-2008 vs. shakemaps and hindcasted national level population data, and are currently globally available. In addition, So and Spence (2009) use building typology superclass fatality ratios and intensity to provide good estimates of fatalities.

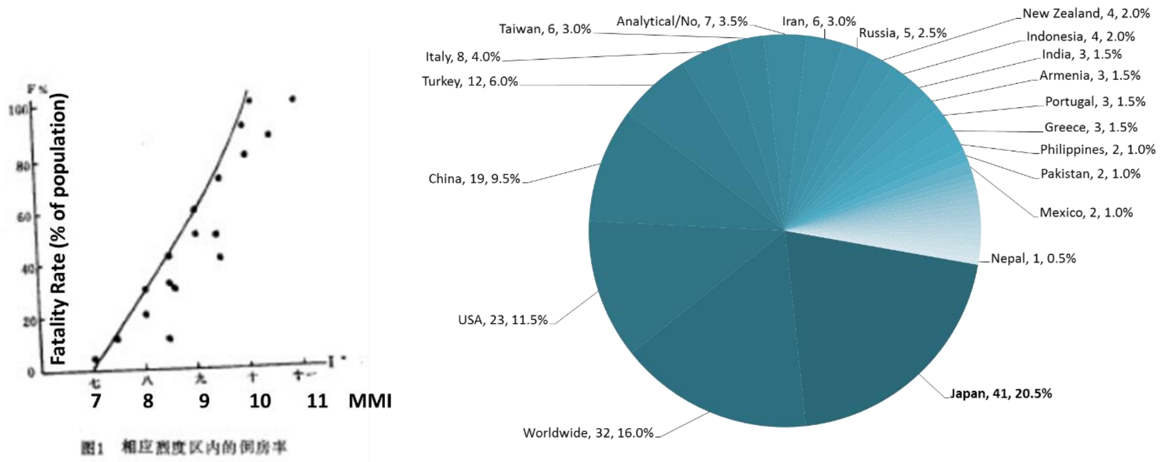


Figure 1: Left: The intensity-fatality ratio method of Haihua (1987); Right: The location of the 200+ casualty functions researched globally for different locations.

For economic functions, over 400 vulnerability studies and functions have been researched and collected; however, the authors are sure that many more exist, particularly in the insurance industry. Of these vulnerability studies, very few relate directly to intensity-economic loss relationships, as they are mainly fragility functions converted to loss, or are for very specific locations globally. For economic losses, there have been 3 main studies historically in this field. Tiedemann and SwissRe (1989) calculated simple equations connecting intensity and economic loss globally. Three curves based on income (low, moderate, high) were created by Chen et al. (2001c), and were based on vulnerability functions from around the world. It can be seen, however, that the results are high compared with the data recorded. Badal (2005) similarly used a MunichRe type loss function (cubic) vs. exposure and intensity from historical events to calculate economic losses with an upper and lower function. During the course of this study, Jaiswal and Wald (2011, 2013) also extended the fatality methodology to calculating reconstruction costs using a similar methodology of fitting functions against MunichRe data versus tangible and intangible wealth, using the same fatality function form of HAZUS (1999), Porter (2007), Jaiswal et al. (2009).

In full risk assessment for buildings and infrastructure, spectral ordinates should be used relating to detailed fragility and vulnerability functions to create losses. However, for rapid loss estimation, due to the amount of variability, the intensity is a reasonable, fast approach.

2. Methodology for the socioeconomic vulnerability functions

The methodology attempts to fit a function to historical earthquake losses using the macroseismic intensity and the exposure and socioeconomic climate of the location at the time of the event. This is done using a process of comparing the observed parameters to the calculated parameters in order to optimise the function.

Observed Parameters: CATDAT Damaging Earthquakes Database value of fatalities (disaggregated) and economic loss & cost as per the papers Daniell (2010), Daniell et al. (2011) from past AEES conferences. Information on 7208 damaging earthquakes from 1900-2012 were collected from over 24,000 sources in the database over the past years.

Calculated Parameters:
A **socioeconomic fragility function** to recreate the historical loss value **at the time of event,** calculating either:
1) Fatalities per intensity bound vs. population exposure
2) Economic losses and costs per intensity bound vs. economic exposure

The process of the **social fragility function** for a single historical event is as follows:-

1. Have a metric of the ground shaking based damage for an individual earthquake. This can either be point based (macroseismic data intensity points) or area based (distribution of intensity over a certain area). The preferable value is MMI, as this is the metric that is used in this study. Other intensity scales can be changed via the damage scales reviewed. If the MDP version does not have enough information to derive a population value, a kriging method is used.
2. For each of the point or area intensity values (vector/polygon format), recreate the population of the time using census and other forms of data. This will be in raster format. This is then joined as a value to the polygon intensity.
3. These are then split into smaller administrative units or geocells in order to assign socio-economic parameters from the databases.
4. The death toll for the particular earthquake is attached to the particular earthquake from the CATDAT Damaging Earthquakes Database. It is important that this is disaggregated into only ground motion shaking related fatalities, rather than including secondary effects.
5. The social parameters are joined to the area based on the databases.

Similarly to HAZUS (1999) and Porter et al. (2007), the normal cumulative distribution function of a lognormal function is used for fatalities, and is optimised for the 2 parameters, theta and beta. The incomplete beta distribution cdf can also be defined over the same range but does not give a lot of improvement on the Level 2 data. Similarly, Weibull distributions, power laws and other polynomial and distribution functions were tested like in Figure 2 without achieving the accuracy necessary for a better regression than that of the cdf, given the better fit in the intensity less-than-8 region. Whether 2 functions will better fit the data is still to be examined in the future.

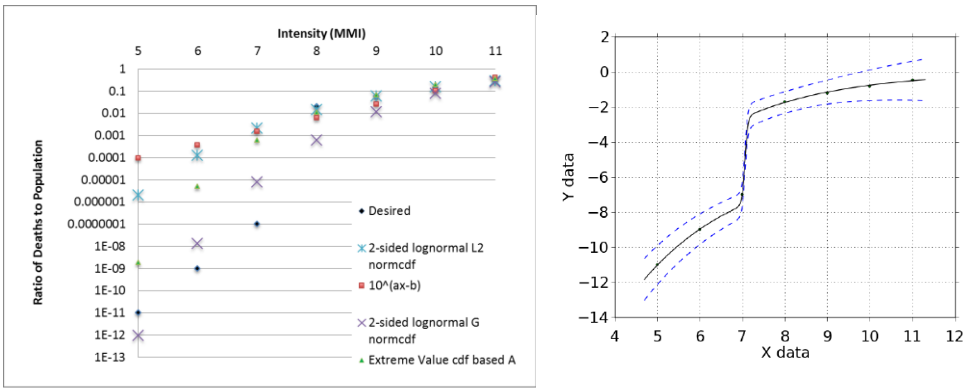


Figure 2: Left: A step in the fatality function due to similar building collapse is difficult to pick up with less than 2 coefficients.; Right: A step in the fatality function due to similar building collapse is modelled with a double exponential function, with 5 coefficients, but is unrealistic for a robust global methodology.

$$FatRatio_i = \Phi\left(\frac{\ln\left(\frac{I}{\theta}\right)}{\beta} \times \left(\frac{HDI_{max}}{HDI_{event}}\right)^n\right) \quad \text{Eq. 1}$$

$$Population_i = Nighttime_i \times Occupancy\ at\ Time\ of\ Day_i \quad \text{Eq. 2}$$

$$C_i = Population_i \times FatRatio_i \quad \text{Eq. 3}$$

I=Intensity, θ = shape parameter 1, β = shape parameter 2, Φ = normal cdf, HDI = human development index

This process is undertaken over each of the intensity bound-population couplets that could contribute to the overall loss and is summed. An example is shown of the optimisation of the three Italian events in Figures 3 and 4.

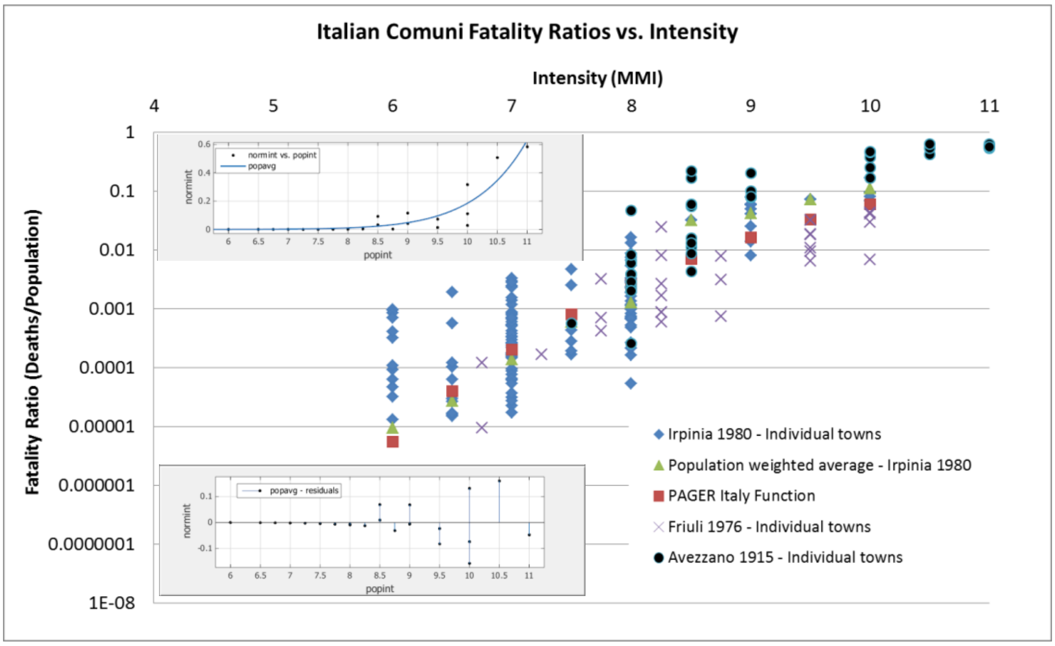


Figure 3: The fatality ratios of all towns affected by the 3 earthquakes – 1915 Avezzano, 1976 Friuli and 1980 Irpinia, including over 300 individual points to fit the fatality ratio; population weighted average for the Irpinia earthquake and the fatality function of Jaiswal and Wald (2010) for Italy is shown. The fit and residuals of the fatality ratio for 3 Italian events against the total fatality ratio per intensity class is also shown.

The state of Abruzzo at the time of the Avezzano quake in 1915 was below the Italian average in terms of many key socio-economic metrics. Education (proxy of mean years of schooling) at the time had a value of 0.733 of the national average, and income was at 0.67 of the national average (Felice, 2005; 2011). This was the same for the 1980 Irpinia earthquake with Campania's income being around 69% and education around 91% of the national average. Friuli at the time of the 1976 earthquake had a higher income (103%) and education (113%) standard than the Italian average at the time of their event. Education can often be used as a proxy for building practice, and income for building quality. The effect of this parameter can be seen on the following diagram. It can also be seen that the PAGER Italy empirical fatality function of Jaiswal and Wald (2010) fits the data quite well.

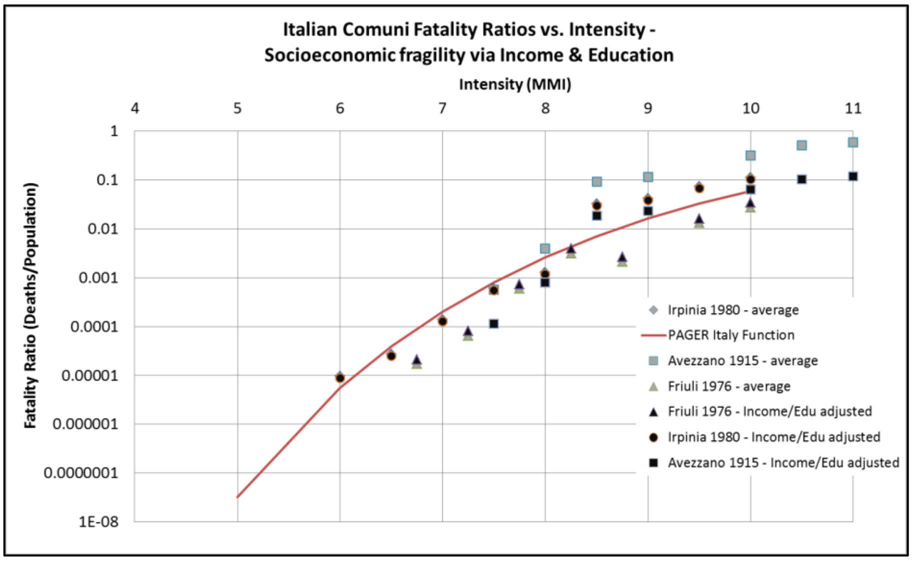


Figure 4: The effect of improving the fatality function via an income-education index adjustment (black) as compared to the grey values. The residual error reduces and the values can be then disaggregated for the individual events, showing the effect of creating a "socio-economic fatality function" vs. a standard regression.

The regression for each of these groupings a-c is then undertaken using the regression approach (shown above) of a tanhyp or normal cumulative density function distribution optimised for the earthquakes, as has been tested through the Level 2/3 casualty data. An assignment of the distribution translates a single value of deaths over the population intensity bounds, and then this is modified depending on which socio-economic parameters correspond to the earthquake in question. The expected deaths are then optimised against the fatalities observed in events.

Each of these events is then grouped together in one of three ways:-

a) As single countries with no additional socioeconomic parameters (traditional method);
b) As single countries adapted by the joined socioeconomic parameters; and
c) As socioeconomic parameters joining together countries.

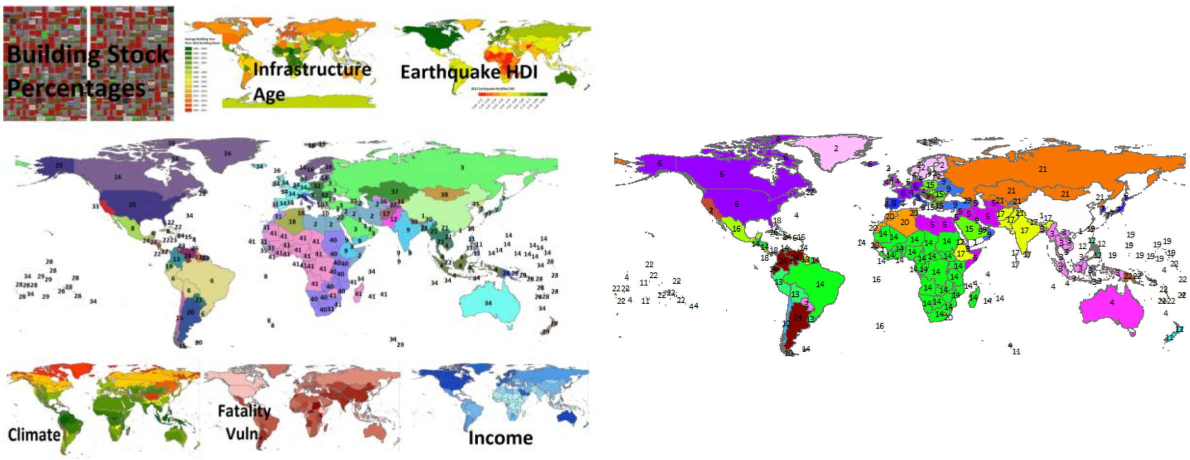


Figure 5: The regionalisation strategy for fatalities (Left) and economic losses (Right). Some of the various components which were looked at in the country classification system are shown in the fatality side of the figure.

Should enough fatal earthquakes not be available from 1900-2008, then a regionalisation approach for the grouping is undertaken shown in Figure 5, such as looking at building typologies, economic situation and practices. This is minimised for the final global method by

using similar HDI and vulnerability events to populate locations where not many damaging earthquakes have occurred.
The variability of the function can be quantified by looking at the residual error. The functions can then be optimised as more data becomes available. The functions are consequently used for the [intensity, population, socioeconomic parameter] triplets for each new event. The combination of an L2 (least squares) norm and G (log) norm into a L2+G norm by Jaiswal et al. (2009b) is a good solution to the problem of upper and lower residuals; however, for the global method, a dual function is tested, with the L2 (least squares) norm being used for the top, and G (log) norm for the bottom, thus retaining the best possible regression.

For Italy, θ is equal to 12 and β is equal to 0.142.

$$\zeta_{v2} = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}\left(\left[\ln(O_i + 0.1) - \mu_{lnO_i|lnC_i})\right]^2\right)} \qquad \text{Eq. 4}$$

Oi=observed event losses, Ci=calculated event losses, N = number of events

where, in the same way as the PAGER model (Jaiswal et al. (2011)), the boundaries of loss can be calculated using the log residual error to give a range of losses based on the uncertainty of the model (Figure 6). The variation allows for users to understand that fatalities often do not fit a perfect trend, given the amount of uncertainties. The functions are also checked over 100 events with lower level fatality data (i.e. town, city, province). This ensures that the function has a reasonable shape.

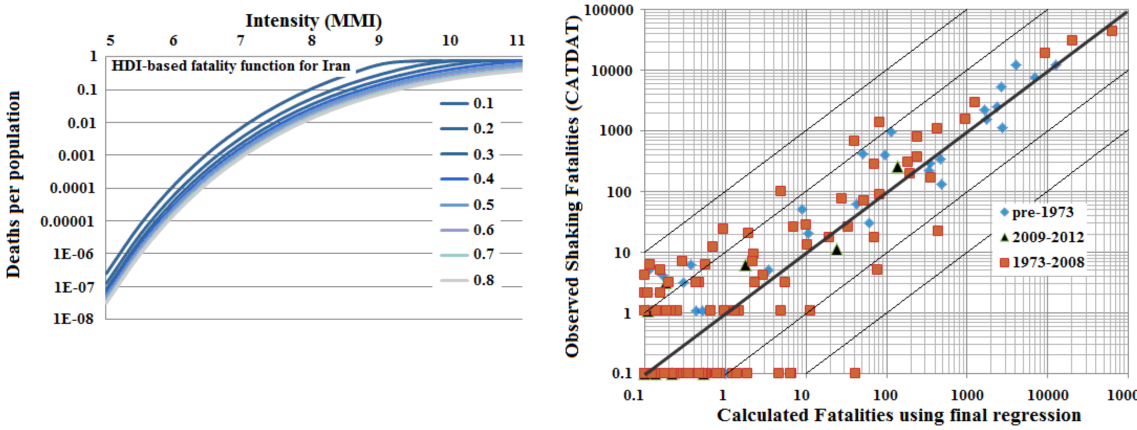


Figure 6: Left: The fatality functions for Iran based on HDI through time; Right: The final fit for Iran based on a time of day and an HDI based fatality function from 1900-2012, calculated on 1900-2008 data (Daniell, 2014).

In order to evaluate the economic losses from any earthquake event worldwide, the historical earthquake event data must be aggregated. The process of the economic fragility function for a single historical event is as follows:-
Steps 1-3 are as for the fatality fragility function above.

4. For each of the point or area intensity values (vector/polygon format), recreate the capital stock and GDP of the time using census and other forms of data. This will be in raster format. This is then joined as a value to the polygon intensity.
5. These are then split into smaller administrative units or geocells in order to assign additional socio-economic parameters from the databases.

6. The direct economic loss for the particular earthquake is attached to the particular earthquake and the gross capital stock. It is important that this is disaggregated into only ground motion shaking related economic losses, rather than including secondary effects, given differences in the methodologies.
7. The economic parameters are joined to the area, based on the databases.

The grouping of the countries for creating functions using economic loss data is then also, as in the fatality model, done with similar economies, historical backgrounds and other disaster data where inadequate earthquake data is available in a country. A regionalisation approach for the grouping is also undertaken, such as looking at building typologies, economic situation, and building practices. The reassessment of all economic loss estimates to date, into a net capital loss, gross capital stock cost form and GDP loss value creates a new insight into losses, but also indicates a major research gap to be filled in the future. It is important to note that a major focus of the methodology is the collection of economic exposure, as the loss estimates are better constrained by having accurate economic exposure assessment.

The regression for each of these groupings, as in the fatality model, are undertaken also for the capital stock and GDP portions versus economic losses over the intensity brackets, using a tanhyp or lognormal cumulative distribution function optimised for the earthquakes. Level 2/3 economic data is also sourced to create the shape, and the tanhyp function provides a reasonable fit, given that this test is required to ensure that a completely different shape has not been seen in past events but fits a distribution of some singular economic loss values distributed over intensities. A tanhyp function has been used with success in Giovinazzi (2005) for European typologies. The form is then modified to allow for fitting of economic loss vulnerability functions.

$$LossRatio(Int) = 0 < \left\{\left[1 + tanh\left(\frac{Int - a}{b}\right)\right] - \left[1 + tanh\left(\frac{4.9 - a}{b}\right)\right]\right\} \div (HDI_{event}\ /HDI_{max}\ )^{n} < 0.9 \qquad \text{Eq. 5}$$

This has also been optimised by getting the best fit between the observed and calculated events for each regionalisation.

3. Exposure Modelling

A key output of this study has been the establishment of the exposure and vulnerability component globally. The production of a virtual earth of infrastructure typologies, population estimation on multiple levels (country, province, sub-province and city) using census information as well as proxies enables the next earthquake to be quantified accurately, allowing for picking up of regional differences such as the difference between an Istanbul (western Turkey) and Van (eastern Turkey) earthquake in terms of economic exposure as well as population and vulnerability.

For the current Earth (circa 2013), extensive work into establishing the economic exposure through net capital stock and gross capital stock has been undertaken using historical depreciation rate methods and existing studies to establish a value of assets that could be impacted by an earthquake currently. These differed by a factor of 3 to 4 from some existing datasets, but were validated against national databases and many existing studies, thus proving the major improvement. Globally, cities with over 750,000 population, as well as capital cities, have been studied specifically with respect to their overall assets, GDP, HDI, population and dynamics. A combination of Level 1 (country), Level 2 (province) and Level 3 (subprovince) modelling has been undertaken over time in order to provide a view of population, infrastructure and contents replacement, infrastructure and contents worth, and GDP since 1900 for each country around the world.

The development of new country-based social and economic indices provides an input to change the vulnerability functions into socio-economic vulnerability functions to better convert to economic and social losses, taking into account the change in vulnerability through time, as well as allowing for comparison between the setting of historical earthquakes. These have been produced in two ways: 1) temporal-spatial – containing data from 1900 to before 2013 for each country; and 2) current socio-economic indices. The impact of these indices was explored with regard to historical losses and the loss output. Temporal-spatial indices produced included the first country-based HDI (Human development index) value consistent from 1900 onwards for all countries as well as unskilled wage, GDP, Capital Stock, CPI, construction cost, and exchange rate data.

Population

This was needed in order to relate the value of fatalities or economic losses historically to the population or exposed assets at the time of the event. In order to build this system, various population datasets have been aggregated and hindcasted to historical values. This often causes slight discrepancies in individual locations where the urban/rural makeup has changed markedly or the population centres have changed within a province or subprovince. An example is shown below for France (Figure 7). There are also some locations for the earth where only full country information is available. Where possible, this has been collected on lower spatial levels as far back as possible and then adjusted using the country rate before that time. Census data has been used for locations with good census data, as in the case of ABS Australia.

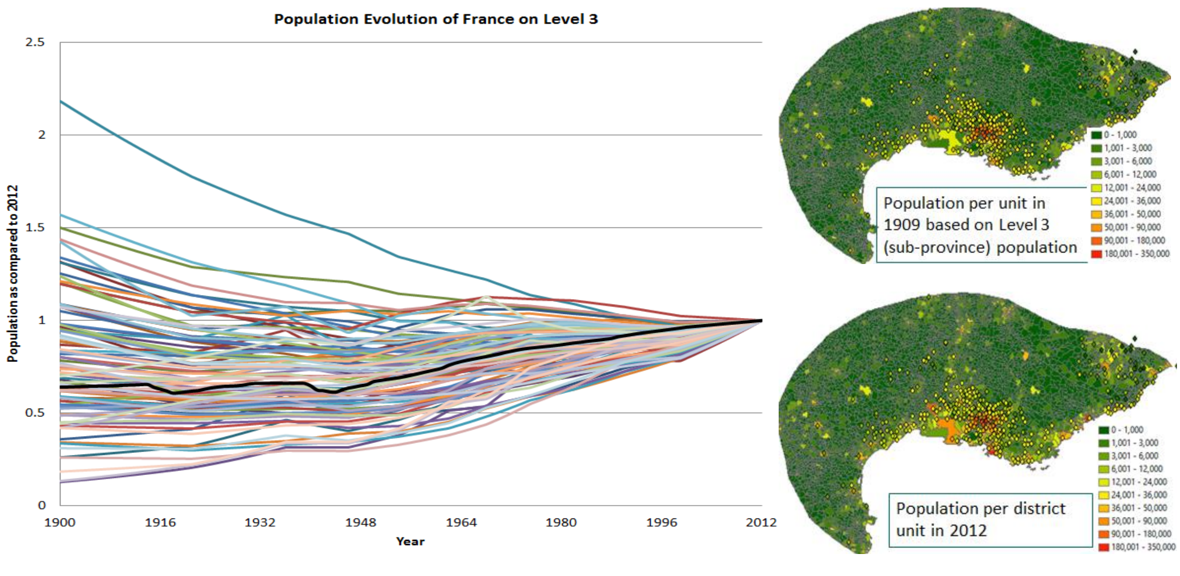


Figure 7: Population modelling for the 1909 Provence earthquake, showing Level 3 population variability for France (Level 3 was the highest level of hindcasted population data, with most on Level 2 (province))

Capital Stock and GDP

In a similar way to the population, the capital stock and GDP of the globe has been collected from 1900 onwards in order to determine the approximate value of the earth at each point in order to compare the potential exposure to earthquakes compared to recorded economic losses. Both the net (depreciated) and gross (replacement) stock has been calculated of equipment, contents, buildings and infrastructure; as well as a potential GDP portion exposed. To do this, depreciation rates were estimated as well as capital stock series sourced from census data, statistical agencies and various authors. This forms the economic exposure for the loss analysis as partly shown in Figure 8.

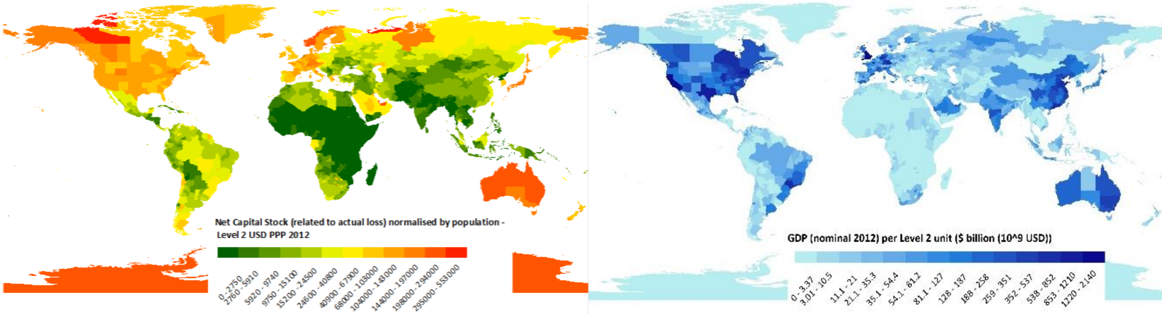


Figure 8: Left: The global net capital stock per capita on Level 2 (infrastructure and buildings' actual cost standing on the earth), Right: GDP on Level 2.

Human Development Index

The human development index forms a key adjusting process for the socioeconomic fragility function in order to standardise events from lower socio-economic climates and earlier years into a single function for use in today's terms. The UNDP definition of combining life expectancy, mean and expected years of schooling, literacy rate and GDP per capita has been calculated from 1900-2013. National and provincial level information was combined in order to create this index over time (Figure 9). An example for France 1909 is shown in Figure 10.

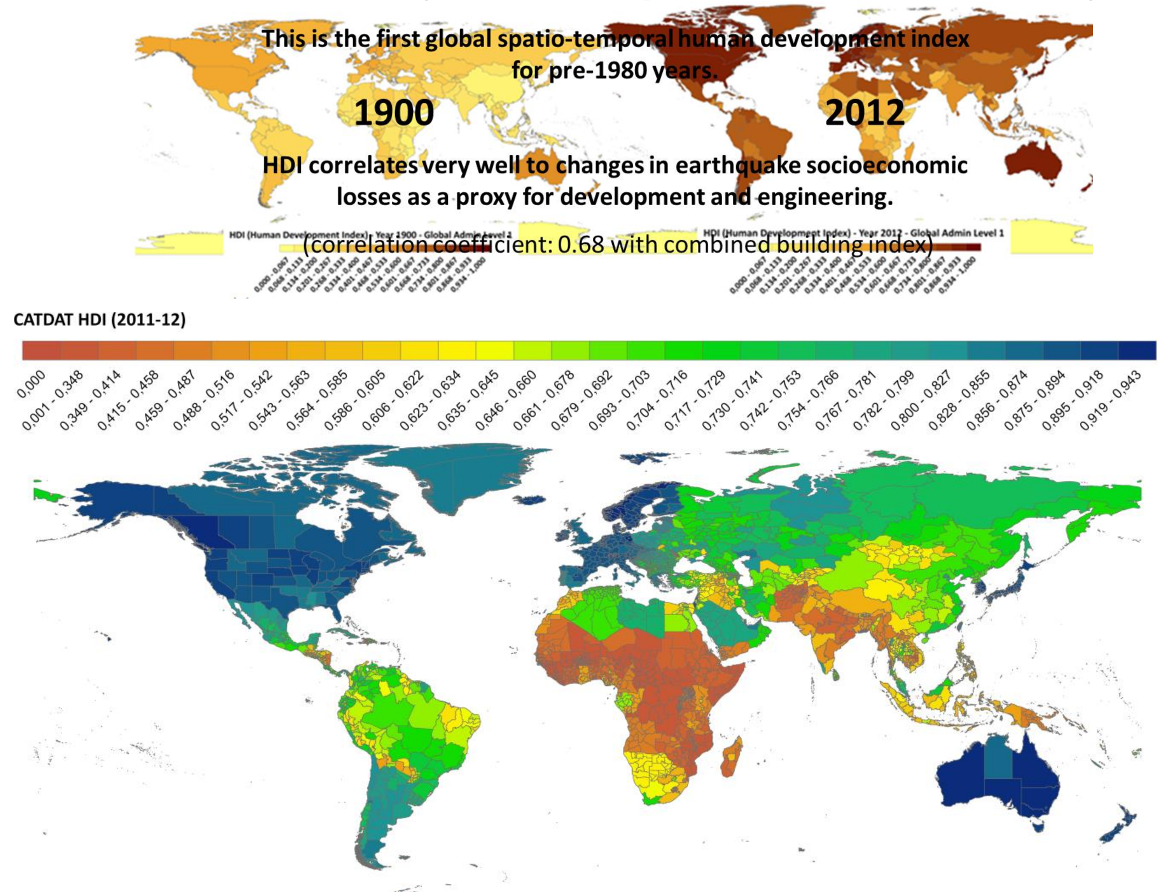


Figure 9: Upper: The 1900-2012 Human Development Index has been created for each country. Lower: Level 2 Human Development Index used for calibration of functions globally. It can be seen that there is much difference in HDI from state to state.

Additional exposure analysis is shown in the work of Daniell et al. (2014) – Paper No. 1400, and Daniell and Wenzel – Paper No. 1505 from ECEES.

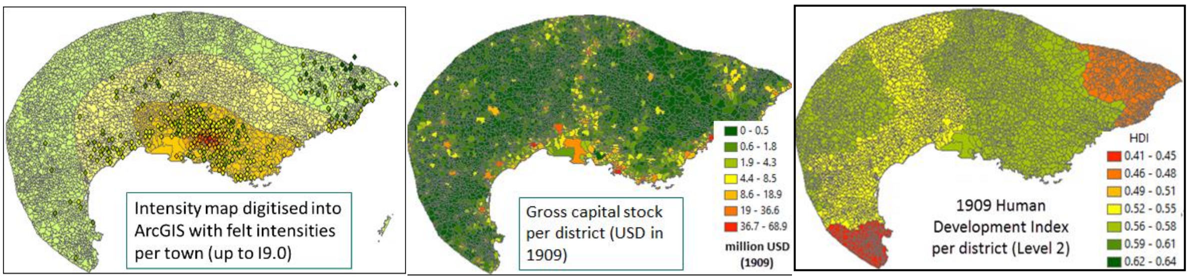


Figure 10: The intensity, gross capital stock and human development for the 1909 Provence earthquake example shown earlier.

4. Historical Hazard Modelling

Many studies have been made to quantify the effects of historical earthquake events. These have been integrated into a global database, combining the spatial, temporal, socio-economic and seismological aspects of the particular earthquakes. Each individual earthquake was examined, and the key earthquakes were searched for in multiple languages in order to attempt to improve the dataset. Isoseismals from many different atlases and sources were collected over the time period and these have been integrated into the database to create a larger sample space through time and to create a more accurate basis for analysing historic earthquakes. An example is shown for the Buyin-Zara event in 1962 in Iran (Figure 11).

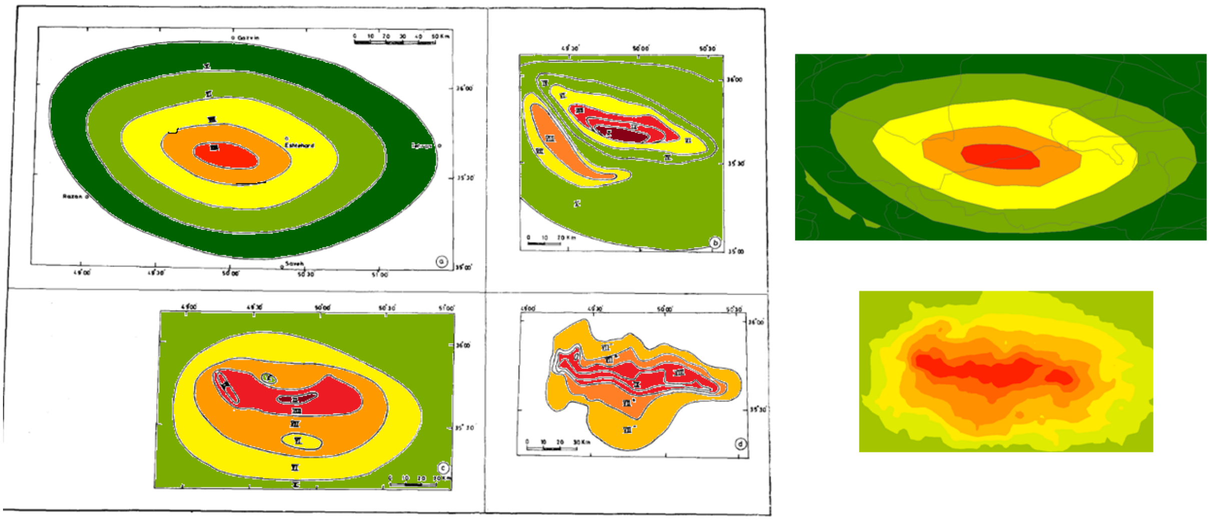


Figure 11: Various isoseismals used for the Buyin Zahra earthquake of 1962 - from clockwise top, Mohajer and Pierce (1963); Abdalian (1963); digitised intensity map using Mohajer and Pierce (1963) and Omote et al. (1965); digitised intensity map using Ambraseys (1963) combined with USGS Shakemap; Ambraseys (1963); Omote et al. (1965).

Large events have been digitized, allowing for reanalysis of these events using current parameters or integration into the functions established (Figure 12). The earthquakes were all disaggregated with respect to shaking and secondary losses; it has also been an extensive effort during this study creating disaggregation of social and economic losses from earthquakes through time from 1900-2014.

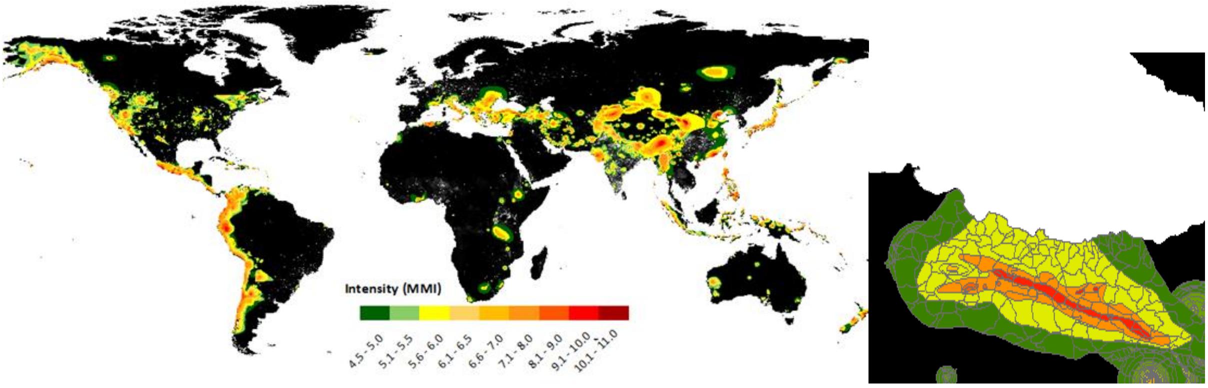


Figure 12: Left: Isoseismal maps created and collected worldwide correlated to damaging earthquakes in the database with an intensity greater than 5. In some cases, damaging earthquakes only had bounds to a minimum of intensity 6, and thus were not extrapolated. Right: An example of the splits of administrative units over an intensity map to form intensity, administrative unit couplets for the Erzincan 1939 event.

5. Risk Modelling

Combining the methodology with the exposure and hazard modelling allowed for the testing of the 1900-2008 model for each earthquake from 2009-2012 (and 2013/14 not presented in this paper). Over the past 5 years, in conjunction with earthquake-report.com, different versions of the socio-economic fragility functions have been produced for each major event and tested in real-time, with changing intensity maps as information came in. When testing the results in the days after an event, case studies have been undertaken using more traditional methods of vulnerability assessment and also forensic analysis to validate or disprove the findings of the economic and fatality modelling. The events from 2009-2012 have been tested with reasonable results, showing the usefulness of such a methodology (Figure 13). It is expected that the functions can be optimised in the future by inclusion of the 2009-2012 data. For fatalities, the methodology predicted 45% of fatal events to within 50% of the death toll reported and 59% for all events. 69% of fatal events were calculated to within 100% of the fatalities reported. 90% of the fatal events tested and 89% of all events were within one order of magnitude below or above the mean.

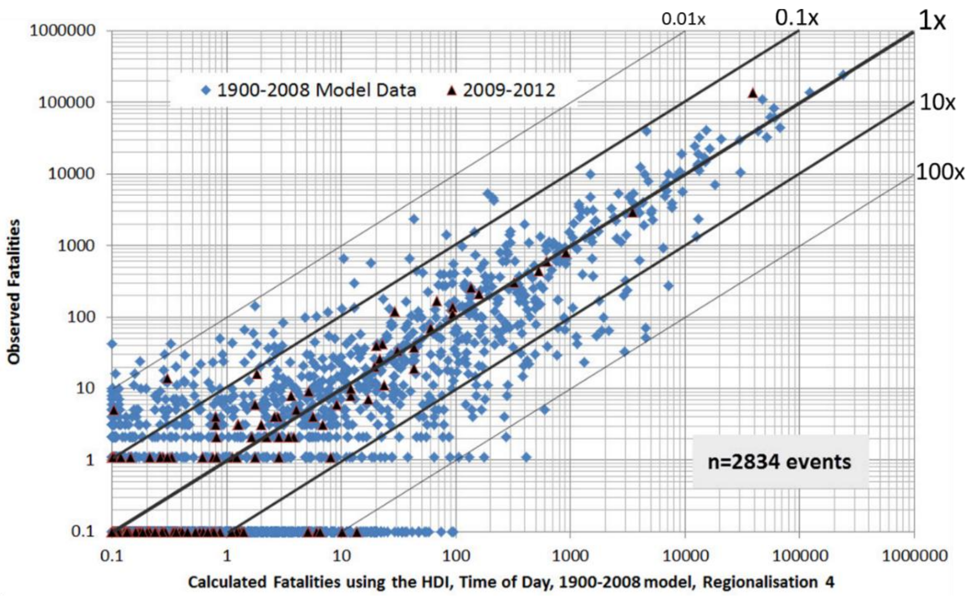


Figure 13: The tested 2009-2012 shaking fatality data using the regionalisation 4 country grouping built from 1900-2008. It can be seen that there are 2 fatal quakes underestimated by the model of 5 and 13 deaths respectively. In addition, 5 misses with overestimates over 2 fatalities result from the model.

The direct comparison of a loss potential is difficult against other models, given the additional historical analysis of isoseismals, shakemaps and population in this study. The functions do not match with those of Jaiswal et al. (2009), with the final functions being sometimes more than an order of magnitude different.

For economic losses, the methodology predicted 42% of events to within 50% of the economic costs reported and 65% of events to within 100% of the economic costs reported. 98% of the events tested were within one order of magnitude in either direction as seen in Figure 14. Three key types of functions have been established; economic cost; economic loss; and fatality for rapid loss estimation. These can also be output in various forms, be it as a socioeconomic parameter or as a country socioeconomic fragility function.

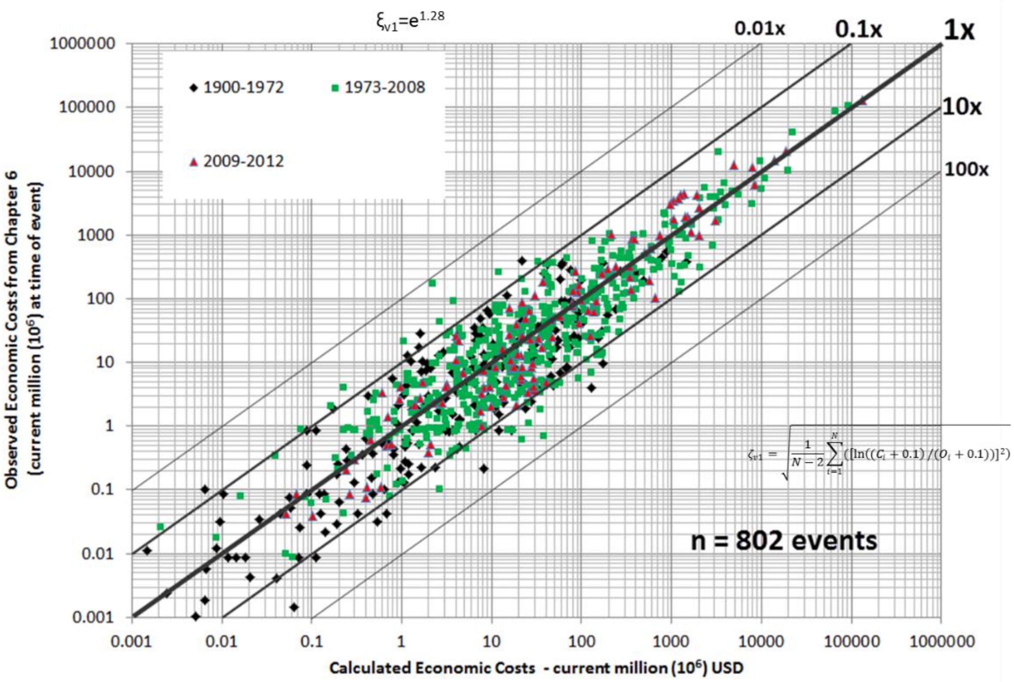


Figure 14: The finalised economic cost model (regionalisation 6, HDI) in 2012 adjusted dollars, testing the period from 2009-2012. (log residual error, ξv1=1.28)

Large events have been seen to be more accurate than smaller events, due to the ability of the empirical function to pick up a higher number of smaller losses from historic events, as shown in Figure 15. Intensity assignments in larger events are also often more widely reported.

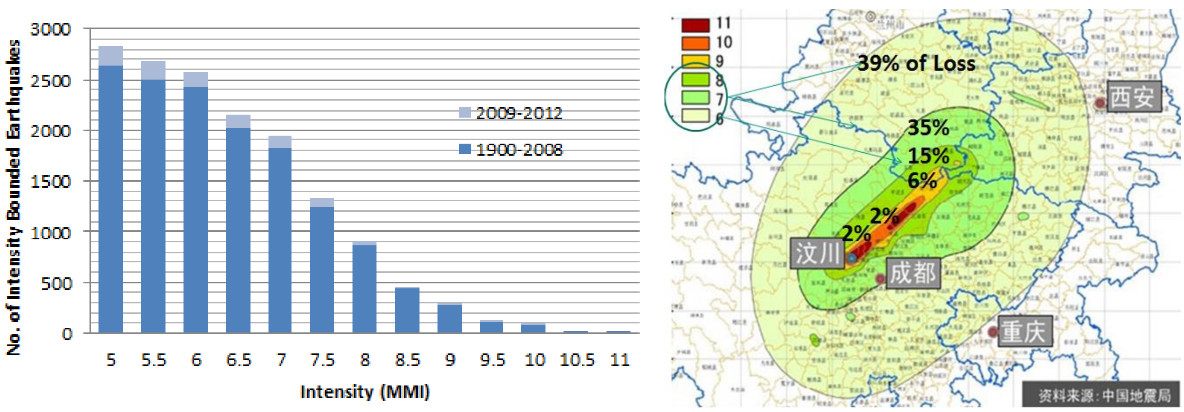


Figure 15: Left: The intensities within the events in the constructed database from 1900-2008 vs. 2009-2012; Right: The economic loss in each intensity bound for the 2008 Sichuan earthquake.

The ability of the model to output globally with the confidence that the socioeconomic parameters are reasonable in any country is a big advantage of such a methodology. It has been shown to give reasonable results when tested against earthquakes over the last 5 years, and has been integrated with www.earthquake-report.com into worldwide reporting of earthquakes, which also allows for social input of data into the functions in the future.

6 Limitations of the study

The methodologies provided are intended for the rapid estimation of economic costs and losses. Although there are many overlaps with traditional engineering based methods, detailed studies are required of cities, locations and building typologies, as there are many components that are not integrated into these simplistic functions of a few parameters. Although every best effort has been made to fill the world globally in terms of population based on census data, this data is often not correct, as shown by the problem in Germany of overcounting before the census of 2011; thus variability could exist. Modelling of building damage and fatalities through time and spatially have been attempted via these global and regional trend analysis for the time period of 1900-2012; however, unexpected damage can often result, leading to an inherent variability in all fatality and economic modelling.

The functions currently have the limitation that the isoseismals and intensity maps created have not been completely reviewed from 1973 onwards. Although a major effort was made to check the shakemaps from USGS against recorded intensities and to collect additional isoseismals, this could only be carried out for a few events and will be the subject of future study, as there are still events in the database where shaking and mapped intensities do not match. The concept of defining intensities is difficult, given that there is no scale which correctly characterises all of the building types of the globe accurately, thus making consistent intensity output impossible. In addition, the use of a continuous scale is not what was intended technically, but is used globally in loss estimation and analysis procedures.

Uncertainties exist along the chain of events from the earthquake source through to the economic loss function. The vulnerability datasets used in this study are based on the country level; hence, regional differences are not accounted for if different from HDI and the socioeconomic exposure quantified on these levels. Therefore, a future effort would be the collection of vulnerability changes over nations, in order to keep filling in the virtual earth. In locations where very few events have occurred, unknowns can always occur with regionalisations of countries, perhaps creating errors in these cases.

This study is presumed to include the uncertainties of intensity maps; however, unexpected building damage patterns can often lead to higher death counts. In addition, fatality rates are often never known, given the extent of some disasters, and therefore there will always be this uncertainty. In addition, censoring and loss of historical records, and manipulation or simply incapability of loss number creation from governments, has occurred through the period of 1900-2012.

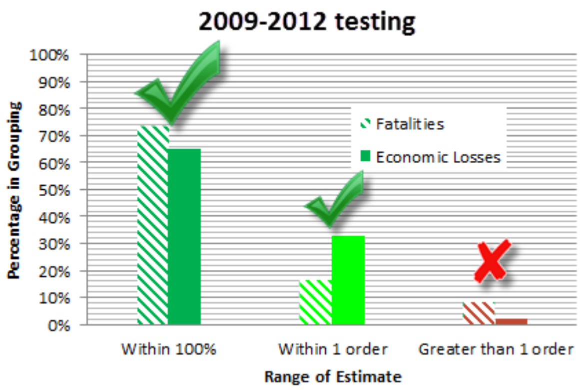
Within the measurement of the socioeconomic parameters, proxies and assumptions are made through time, given historical gaps in data, lack of collection and sporadic data points. As further studies are undertaken by the authors, these trends will be subject to constant change.. There will always be additional values that can be collected.

The quantification of economic exposure is one that has difficulties, given that capital stock is reasonably difficult to quantify accurately because of the number of different asset types, service lives and data points. In this study a simplified methodology has been used, including the results of more complex methodologies to trend these in places through time; however, large differences in capital stock across a nation may occur and may not be related to the output proxy used in this study. The methodology detailed in this study also only refers to shaking losses, and does not include secondary effects losses (although other methods have been established externally).

7 Conclusion

The methodology provides a rapid set of vulnerability functions for fatalities and economic losses which can be used quickly after an event. The methodology has been shown to create more consistent estimates than existing global methodologies and has shown great success globally from 2009-2013 in predicting the scale of disasters, which has been useful for many insurance, government and aid agencies. It should be noted, however, that the inherent variability in estimates once the intensity is defined cannot likely be improved much past this point, even given new information.

Consistent worldwide multi-level socioeconomic databases have been built in the course of this study and have been checked against existing small-scale studies. These databases have been used in various fields of study and not just earthquake engineering. The highest number of damaging earthquakes ever studied in this way has provided a valid sample space, with over 2835 damaging earthquake events looked at for the fatality functions. The economic loss functions have also fitted very well with existing engineering fragility functions globally, providing a good validation for the methodology. The use of province and sub-province data has enabled much more accurate results than existing methodologies.



This methodology is not a substitute for in-depth analytical and empirical analyses using spectral ordinates and traditional building and infrastructure fragility functions which need to be undertaken in any major risk assessment, but is intended for rapid loss analysis in the hours and days after the event, before detailed estimates are available.

Acknowledgements

I would like to thank the General Sir John Monash Foundation for funding my doctoral studies. In addition, the authors would like to thank Kay Mitusch, Stefan Hinz, Mustafa Erdik, Bijan Khazai, Andreas Schaefer and Jonathan Santiago for the useful discussions and comments.

References

Abdalian, S. (1963). *Seismotectonics of Iran* (Vol. 16): Impr. de L'Université de Tehran.

Allen, T. I., Wald, D. J., Hotovec, A. J., Lin, K. W., Earle, P. S. & Marano, K. D. (2008). *An atlas of shakemaps for selected global earthquakes*. Retrieved from http://purl.access.gpo.gov/GPO/LPS100315

Ambraseys, N. N. (1963). The Buyin-Zara (Iran) earthquake of September, 1962; a field report. *Bulletin of the Seismological Society of America*, *53*(4), 705–740.

Australian Bureau of Statistics (ABS). (1889-2012). *Year book Australia*. Canberra: Australian Bureau of Statistics.

Badal, J., Vázquez-Prada, M. & González, A. (2005). Preliminary quantitative assessment of earthquake casualties and damages. *Natural Hazards*, *34*(3), 353–374.

Chen, Q.-F., Chen, Y. & Chen, L. (1997a). Earthquake loss estimation with GDP and population data. *Acta Seismological Sinica*, *10*(4), 791–800.

Chen, Y., Chen, Q.-F. & Chen, L. (2001c). Vulnerability analysis in earthquake loss estimate. *Natural Hazards*, *23*(2-3), 349–364.

Daniell J.E., Wenzel F., Khazai B., Santiago J.G. and Schäfer A. (2014) "A worldwide seismic code index, country-by-country global building practice factor and socioeconomic vulnerability indices for use in earthquake loss estimation," Paper No. 1400, 15th ECEE, Istanbul, Turkey.

Daniell, J. E. (2010f). The CATDAT Damaging Earthquakes Database, Paper No. 06. In *Australian Earthquake Engineering Society 2010 Conference, Perth, Australia* .

Daniell, J. E., Vervaeck, A. & Wenzel, F. (2011h). A timeline of the socio-economic effects of the 2011 Tohoku Earthquake with emphasis on the development of a new worldwide rapid earthquake loss estimation procedure. In *Australian Earthquake Engineering Society 2011 Conference, Barossa Valley, Australia* .

Daniell, J.E. (2014) "Development of socio-economic fragility functions for use in worldwide rapid earthquake loss estimation procedures", Doctoral Thesis, Karlsruhe Institute of Technology, Karlsruhe, Germany.

Daniell, J.E. (2014) "Socioeconomic impact of earthquake disasters.", in "Earthquake Hazard, Risk, and Disasters", ed: Prof. Max Wyss, Earthquake and Seismic Hazards and Disasters. Elsevier.

Daniell, J.E. & Wenzel F. (2014a) "The production and implementation of socioeconomic fragility functions for use in rapid worldwide earthquake loss estimation", Paper No. 490, 15th ECEE (European Conference of Earthquake Engineering), Istanbul, Turkey.

Daniell, J.E. & Wenzel F. (2014b) "The Economics of Earthquakes: A reanalysis of 1900-2013 historical losses and a new concept of capital loss vs. cost using the CATDAT Damaging Earthquakes Database," Paper No. 1505, 15th ECEE (European Conference of Earthquake Engineering), Istanbul, Turkey.

Erdik, M. & Safak, E. (2008). Earthquake Early Warning and Rapid Response System (ELER) Istanbul. In *International Earthquake Conference, Los Angeles, USA*.

Felice, E. (2005). The regional added value. An estimation for the years 1891 and 1911 as well as some long term elaborations (1831-1971) (in Italian). *Rivista di storia economica*, *21*(3), 273–314.

Felice, E. (2011). *The determinants of Italy's regional imbalances over the long run: exploring the contributions of human and social capital* (Oxford University Economic and Social History Series No. _088). Retrieved from http://ideas.repec.org/p/nuf/esohwp/_088.html

Giovinazzi, S. (2005). *The Vulnerability Assessment and the Damage Scenario in Seismic Risk Analysis* (Ph.D. Thesis). Technical University Carolo-Wilhelmina, Braunschweig, Germany. Retrieved from http://rzbl04.biblio.etc.tu-bs.de:8080/docportal/servlets/MCRFileNodeServlet/DocPortal_derivate_00001757/Document.pdf

Haihua, L. (1987) "Preliminary Analysis on the Highest Rate of Casualty in Earthquake Disasters," Journal of Catastrophology, 1987-02, Seismological Institute of Lanzhou, SSB, China.

Jaiswal, K. S. & Wald, D. J. (2011). *Rapid Estimation of the Economic Consequences of Global Earthquakes* (Open-File Report No. 2011-1116). Retrieved from http://pubs.usgs.gov/of/2011/1116/

Jaiswal, K. S. & Wald, D. J. (2010). An empirical model for global earthquake fatality estimation. *Earthquake Spectra*, *26*(4), 1017–1037.

Jaiswal, K. S. & Wald, D. J. (2013). Estimating Economic Losses from Earthquakes Using an Empirical Approach. *Earthquake Spectra*, *29*(1), 309–324.

Jaiswal, K. S., Wald, D. J. & Hearne, M. (2009). *Estimating Casualties for Large Earthquakes Worldwide Using an Empirical Approach: U.S. Geological Survey Open-File Report OF 2009–1136,*. Retrieved from http://pubs.usgs.gov/of/2009/1136/pdf/OF09-1136.pdf

Mohajer, G. A. & Pierce, G. R. (1963). Qazvin, Iran, earthquake. *AAPG Bulletin*, *47*(10), 1878–1883.

Omote, S., Nakagawa, K., Kobayashi, H., Kawabata, S. & Nakaoka, E. (1965). A report on the Buyin earthquake (Iran) of September 1, 1962. In *Proceedings of the 3rd World Conference for Earthquake Engineering, New Zealand*.

Samardjieva, E. & Badal, J. (2002). Estimation of the expected number of casualties caused by strong earthquakes. *Bulletin of the Seismological Society of America*, *92*(6), 2310–2322.

Samardjieva, E. & Oike, K. (1992). Modelling the number of casualties from earthquakes. *Journal of Natural disaster science*, *14*(1), 17–28.

So, E. & Spence, R. J. S. (2009). *Estimating shaking-induced casualties and building damage for global earthquake events* (Final Technical Report, NEHRP Grant. No. 08HQGR0102).

Wyss, M. (2004a). Earthquake loss estimates in real-time begin to assist rescue teams worldwide. *Eos, Transactions American Geophysical Union*, *85*(52), 565–570.