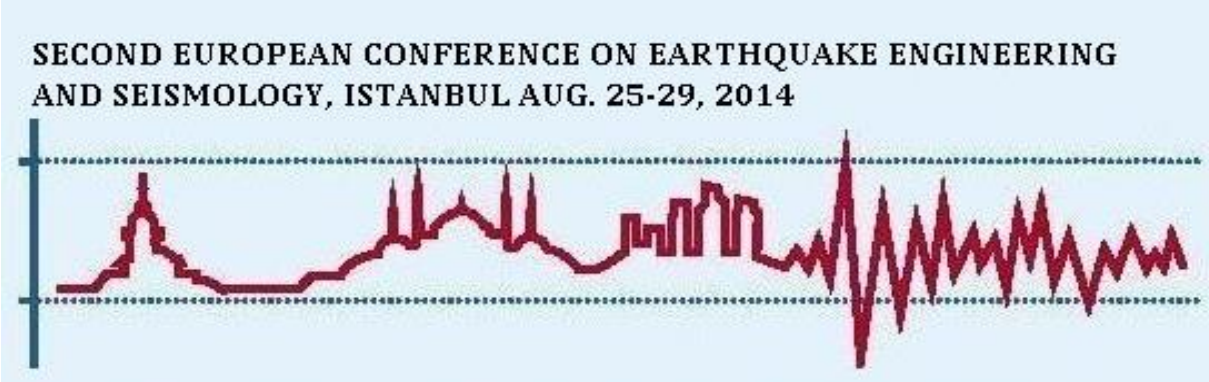


THE GEM EARTHQUAKE CONSEQUENCES DATABASE

Emily SO[1,2] and Antonios POMONIS[3]

ABSTRACT

To improve the performance of buildings and other structures in earthquakes, it is essential, in addition to analytical and experimental studies, to observe and quantify the performance of structures after damaging earthquakes. Only by such observations can we gain confidence that structural design is achieving its aim of providing life safety and by field observations, the causes of the failure of aspects of design can be identified, and steps to rectify them can be found. Earthquake damage observations and surveys are also of crucial importance for assessing the vulnerability of existing buildings, in applications such as rapid impact assessments, mitigation planning and risk modelling. For best use to be made of these data, it is essential to bring them together and make it accessible to the wider research community, to enable cross-event analysis to take place, and to determine to what extent lessons learnt in one country or region can be applied elsewhere.

The Global Earthquake Consequence Database (GEMECD) has been developed under the auspices of GEM, the Global Earthquake Model over the past three years, with the participation of partners in Latin America (ERN-AL), Oceania (GNS Science), the Far East (Kyoto University), the Middle East (Kandilli Observatory) and North America (SPA Risk) to make possible for the first time, an easy and open access to data that have been painstakingly collected by various authorities and researchers around the world after important earthquakes. What sets GEMECD apart is its global outreach and its adherence to inventory and structural taxonomies developed in parallel within the GEM project. This ensures that GEMECD can be directly utilised for vulnerability assessments and loss model validations for many regions around the globe.

The database has been designed in such a way as to be able to capture the full spectrum of earthquake consequences which can be visualised as a matrix of the interaction between the various inventory assets and the earthquake-related damage agents. The database assembles consequence data of five different categories and consists of data sets on ground shaking damage to standard buildings; human casualty studies and statistics; ground shaking consequences on non-standard buildings, critical facilities, important infrastructure and lifelines; consequences due to secondary, induced hazards (landslides, liquefaction, tsunami and fire following) to all types of inventory classes and lastly, socio-economic consequence and recovery data (provided by CRED, Louvain University). Underlying each of the 71 events in the database is a set of upgraded shake maps developed by the US Geological Survey which serves as a common hazard denominator for analyses.

Central to GEMECD is the ability of carrying our analyses across events through a set of tools housed within GEM's OpenQuake. For earthquake engineers, this would enable cross-event analyses

[1] Director, Cambridge Architectural Research Ltd., Cambridge, UK, emily.so@carltd.com
[2] Director, University of Cambridge Centre for Risk in the Built Environment CURBE, Cambridge, UK, ekms2@cam.ac.uk
[3] Director, Cambridge Architectural Research Ltd., Cambridge, UK, antonios.pomonis@carltd.com

to be derived for given building inventory classes, and levels of ground motion, leading to more robust empirical vulnerability relationships, but also enable layman users to learn about earthquakes of the past.

This paper presents the GEMECD, its contents and capabilities. Housed within a public GIS platform, it is hoped that GEMECD will serve to inform users on consequences from past events, as a benchmarking tool for analytical loss models and to support the development of tools to create vulnerability data appropriate to specific countries, structures, or building classes. By seeking out relationships linking past events and the environments they affect, it is hoped that our understanding of the potential regional consequences from earthquakes will improve, especially in parts of the world where there have been fewer documented events in the recent past.

INTRODUCTION

Despite a plethora of losses in past earthquakes and post-earthquake reconnaissance missions to investigate the causes of these losses, the quality and uniformity of the loss data collected for some of these events can be described as poor, due to inconsistencies in methods and reporting. Though limited, these data provide important constraints for loss model development, validation and calibration. The utility of existing global data can also be substantially improved with dedicated focus on proper data taxonomy and uniformity, and aggregation of widely dispersed and variable format data sets into a formal database. Given the state of past data, then, it is equally important to develop a proper framework for collecting perishable loss data for earthquakes in the future. For GEMECD, we envisioned a database that will inform data collection and integration and enable a clearer understanding of the true consequences of earthquakes around the World, thus contributing to the eventual risk mitigation actions that bring those at risk closer to a desirable level of earthquake protection.

The aim of the Global Earthquake Consequence Database (GEMECD) project under the auspices of GEM, the Global Earthquake Model is to make possible, for the first time, the easy and open access to data that have been painstakingly collected by various authorities and researchers after important recent and historical earthquake disasters around the globe. By seeking out relationships linking past events and the environments they affect, it is hoped that our understanding of the potential regional consequences from earthquakes will improve, especially in parts of the world where there have been fewer documented events in the recent past. This has been accomplished by a thorough investigation of the key consequences of past earthquakes around the World, compiled into a database of the best available data. The data collated, learning from local knowledge and existing regional initiatives, draws on regional capabilities and international reconnaissance missions and are linked to a GIS mapping of ground shaking and other induced secondary hazards. What sets this database apart from previous efforts is a strict adherence to the developing inventory and structure taxonomies being developed in parallel to this effort, ensuring that GEMECD can be directly utilised for vulnerability development and loss model validation.

The nature of this project is collaborative. A consortium of regional experts have assembled a structured and web-accessible geospatial database of consequence data including building damage, human casualties, social disruption, and other secondary impacts from events around the World in the past four decades. The database has also been designed for users to contribute their own data from past events and document future consequences, and will be housed in the GEM OpenQuake platform for users to take advantage of the other tools and databases in GEM. This platform will contain features to enable post-processing across events, structure types and ground motion severities for specific consequences globally or within geographic regions.

THE GEMECD CONSORTIUM

The GEMECD consortium has been formed to ensure the best available knowledge is employed to fulfil each of these tasks while including partners with expertise in all the important earthquake-prone regions of the World. The existing working relationships and geographical awareness of this international group of partners was crucial in this global initiative; the project team is shown in Figure 1.

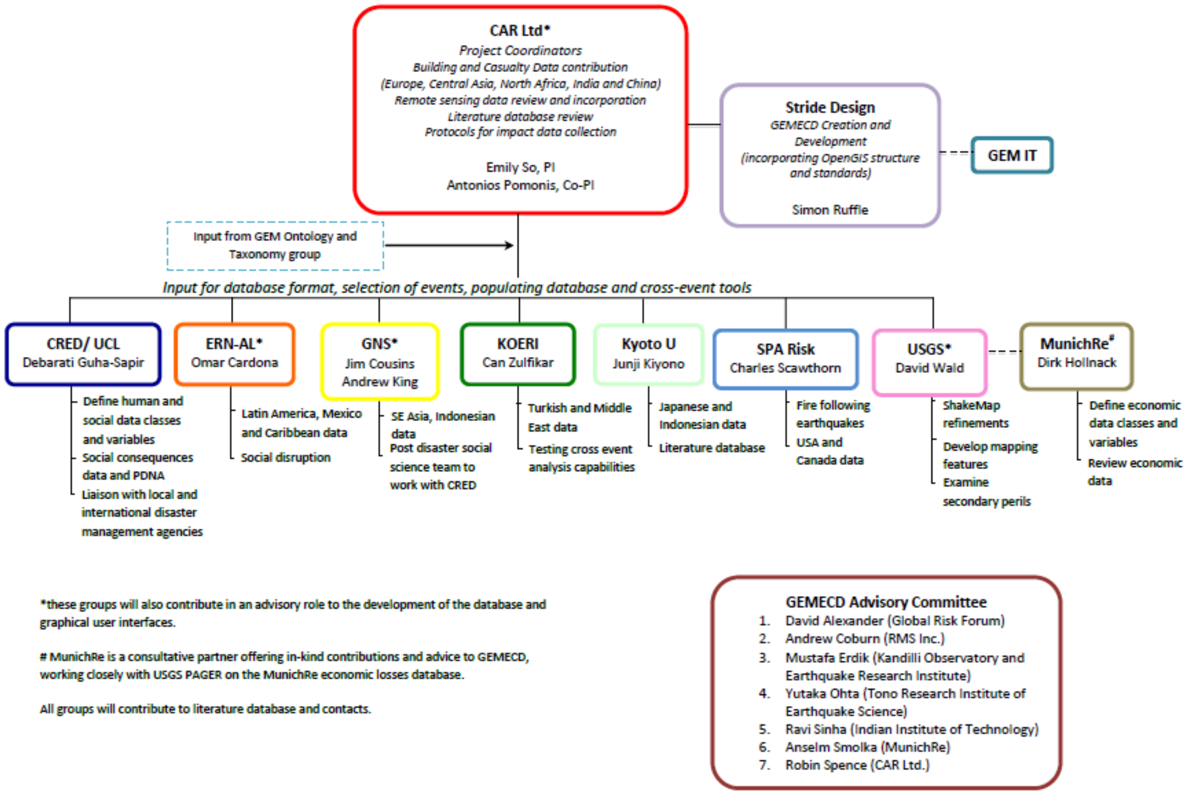


Figure 1. Organogram of the GEMECD consortium

Within the GEM development agenda it was envisaged that GEMECD will provide data that are useful for seismic risk modelling purposes and particularly to furnish other GEM Risk projects with the necessary information to carry-out inventory assessments, vulnerability assessments, model calibrations and model validations. The main aim is to assist with modelling of risk to standard buildings due to ground shaking, as this interaction of hazard and exposure is the greatest contributor to earthquake risk worldwide. GEMECD therefore has a heavy focus on studies on building damage due to ground shaking, although the database also cover other aspects of earthquake consequences. There are five different categories of consequences data in GEMECD:

a) Ground shaking damage to standard buildings.
b) Human casualty studies and statistics.
c) Ground shaking consequences on non-standard buildings, critical facilities, important infrastructure and lifelines.
d) Consequences due to secondary, induced hazards (landslides, liquefaction, tsunami and fire following) to all types of inventory classes, i.e. standard and non-standard buildings, lifelines and critical facilities.
e) Socio-economic consequence and recovery data.

GEMECD has been designed to capture the full spectrum of earthquake consequences and can be visualised as a matrix of the interaction between the various inventory assets and the earthquake-related damage agents, as shown in Figure 2.

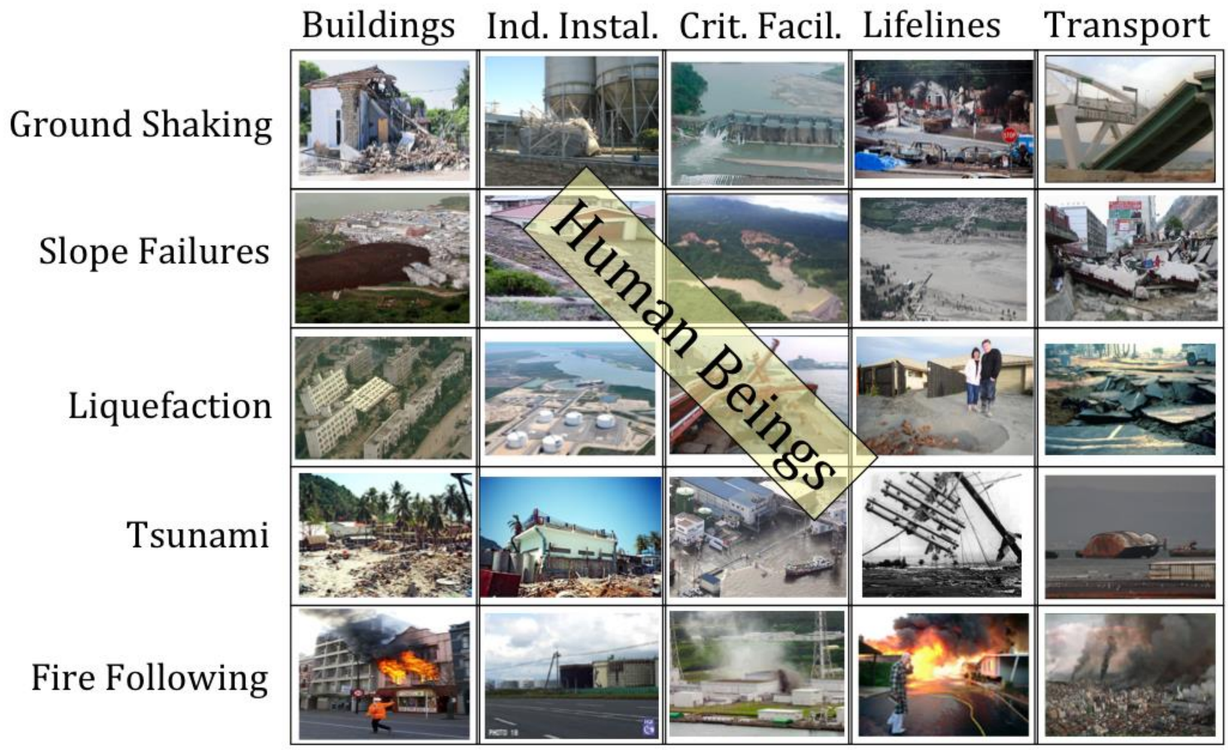


Figure 2. Matrix of consequences considered in GEMECD

GEMECD currently contains 71 event overviews with 98 accompanying U.S. Geological Survey ShakeMaps including high resolution zoomable maps and related products. ShakeMaps provide the site hazard basis or "denominator" for GEMECD consequences at their location of occurrence. In order to improve the maps in the ShakeMap Atlas v1.0 (Allen et al., 2008) for the GEMECD events the USGS has carried out the following during the project:

- Manual determination of event-specific sampling resolution and map dimensions. A subset of 30 data-rich events which have been carefully examined at larger scales.
- Improvement of the selection scheme to choose the appropriate suite of equations for a given event. W; USGA employ NGA and other state-of-the art Ground Motion Prediction Equations (GMPEs), Intensity Prediction Equations (IPEs) and ground-motion to conversion equations -GMICEs- (see Johnson et al., 2012; García et al., 2014) and apply a new regionalisation algorithm that substantially improves the way prediction and conversion equations are chosen for a given event, combining source and seismotectonic information from different sources (García et al., 2012).
- Manual analysis, quality assurance, and selection of the best equations for a given event. Within a given suite of equations (e.g., stable continental region, interface, etc.), there are usually multiple options available. The choice of the specific equation to be used for a specific given earthquake is determined by testing different IPE/GMICE/GMPE combinations and examining their data fitting in terms of the intensity, PGA, and PGV grids and the regression plots (allowing for inter-event term bias corrections). Final maps were reviewed in terms of the intensity grid quality and reliability of the product, including. These improvements were done both using both the maps (not only MMI and, but also for the GM parameters maps, as well as) and the regression plots, was assessed. A careful review has been conducted for the uncertainty map, a crucial addendum of the intensity grid. As a consequence, uncertainty on the GEMECD events has systematically reduced from ShakeMap Atlas v1.0.

At the time of writing, 99 building damage studies are included in GEMECD. Each study contains detailed study description, building class definition, taxonomy string for each examined building class, damage scale definition, study accuracy and reliability assessment and accompanying shape files of the examined locations. Locations can be of the following three types: point geo-coordinates, administrative boundary (using the GADM database) or user defined boundary. The same requirements were set for other consequences and GEMECD houses 43 casualty studies, 43 critical building and infrastructure studies and 23 socio economic reports (the latter related to 18 events). This equates to 3,070 study locations worldwide comprising of 33,763 individual survey value records. In total, partners have uploaded more than 580 geo-referenced photographs into the database.

THE DATABASE STRUCTURE AND LIST OF EVENTS

The list of events and types of consequence studies included in GEMECD is available in separate reports in GEM NEXUS (http://www.nexus.globalquakemodel.org/gemecd/files). The core (cross-event) database structure of GEMECD is shown in Figure 3. As shown, the database is study based and as such, it is highly flexible for searching by asset types. The associated tools and all documentation on the database creation can also be found on NEXUS.

Input to and output from the database has been carried out via a web based interface (http://gemecd.org/), which is a tool specially developed for the task of supporting the data input process, with an XML-based import format. Partners inputting studies to GEMECD were generally choosing to upload their contributions via the XML format using the four Excel macro templates developed and use the web based interface as a maintenance and checking tool. The gemecd.org site also has a facility for uploading shapefiles which can be used to show geographical extents of studies and partners have also been making use of this. At present, this is password protected and designed for internal use only but in the coming year, it is hoped that these uploading features will be integrated into GEM's OpenQuake.

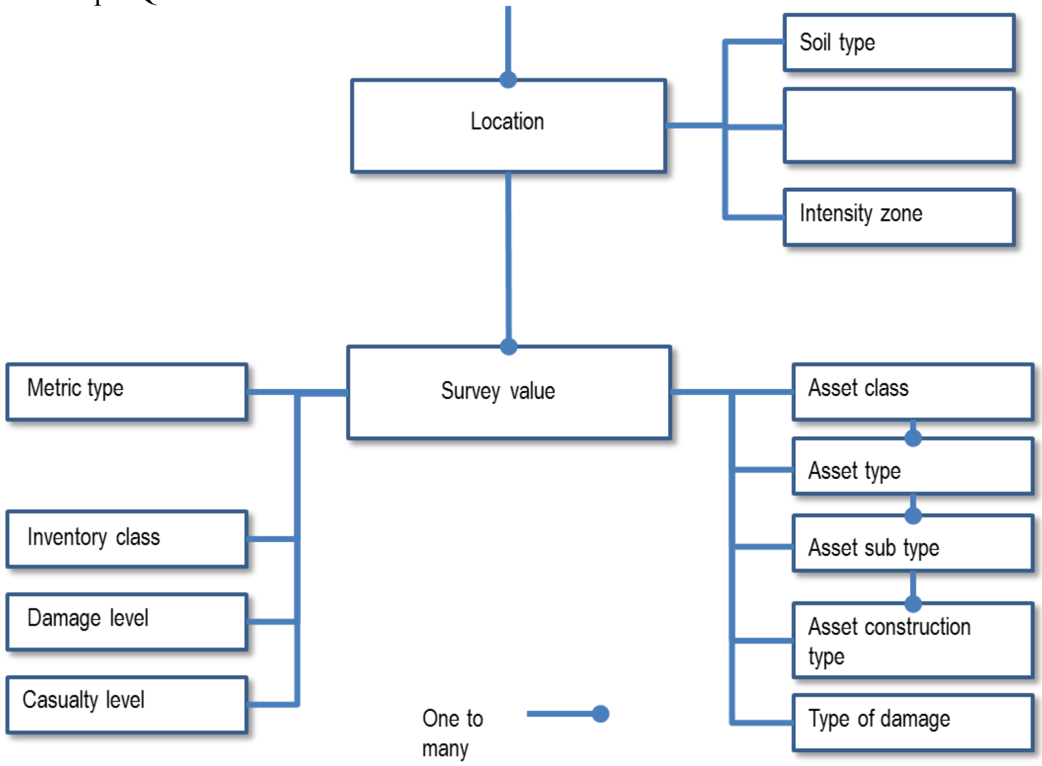


Figure 3. The core (cross-event) database structure of GEMECD

In the final year of the three year GEM project which started in 2010, test pages of how GEMECD could be directly implemented into OpenQuake platform were created and these are shown in the figures below. The figures also show how the database has been organised and the types of information that are housed at each level of the database.

OpenQuake maps have ability to add overlays – e.g. in Figure 4, the GRUMP[4] Population layer has been added to the map. On the top of the webpage, there are also quick filter buttons on the map which allow the user to select events that contain data they may be interested in. At present there are three main categories in the database, namely buildings, casualty and critical buildings with infrastructure, with secondary-induced hazards to be added.

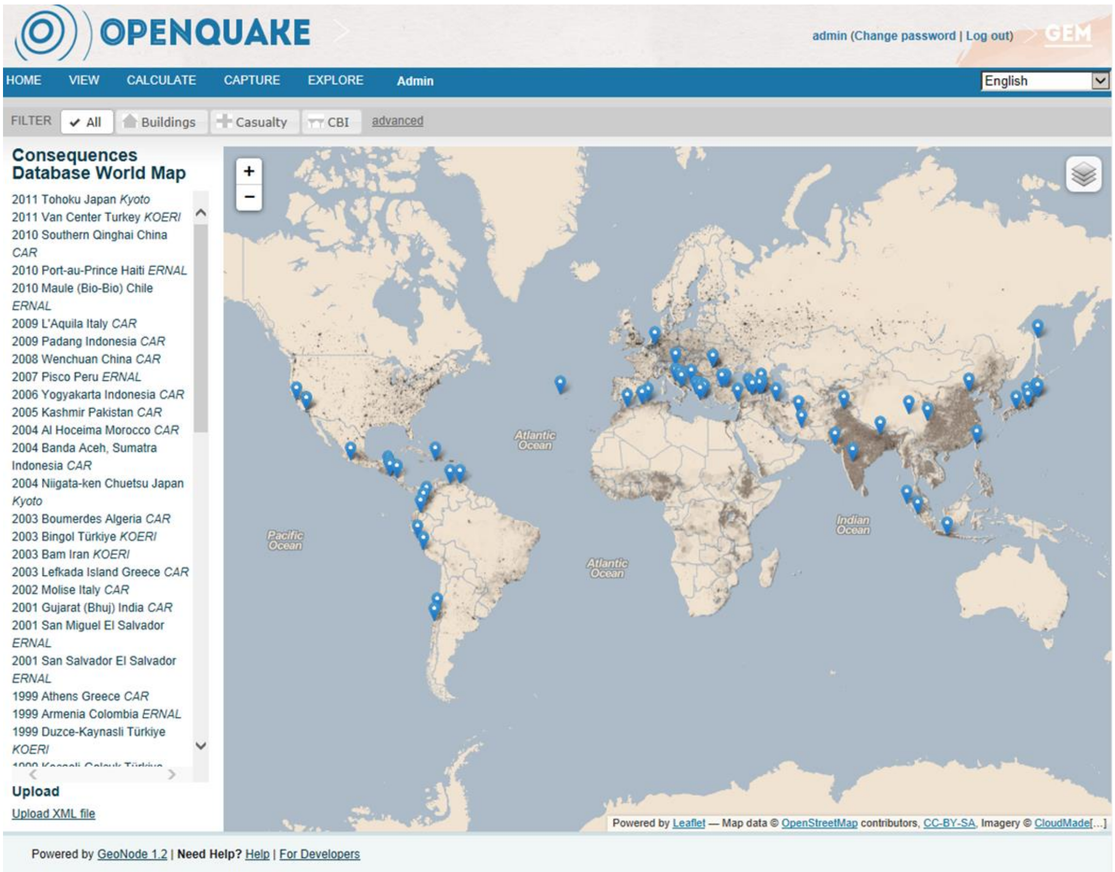


Figure 4. World events map with a map marker for each event in the Consequences Database, with full list down on the left hand side

From the World map one can then select an individual event and obtain an overview. This would take the user into the event area which shows the epicentre, 5km radial rings and the surveyed locations included in the database, as shown in Figure 5. Some locations are points, some been aggregated into GADM[5] administrative polygons as is shown in the events overview page below, while others are user defined boundaries. Once in the event, the user can choose to overlay the USGS ShakeMap for the event, also shown in Figure 5.

[4] GRUMP- Global Rural-Urban Mapping Project

[5] GADM- Database of Global Administrative Areas

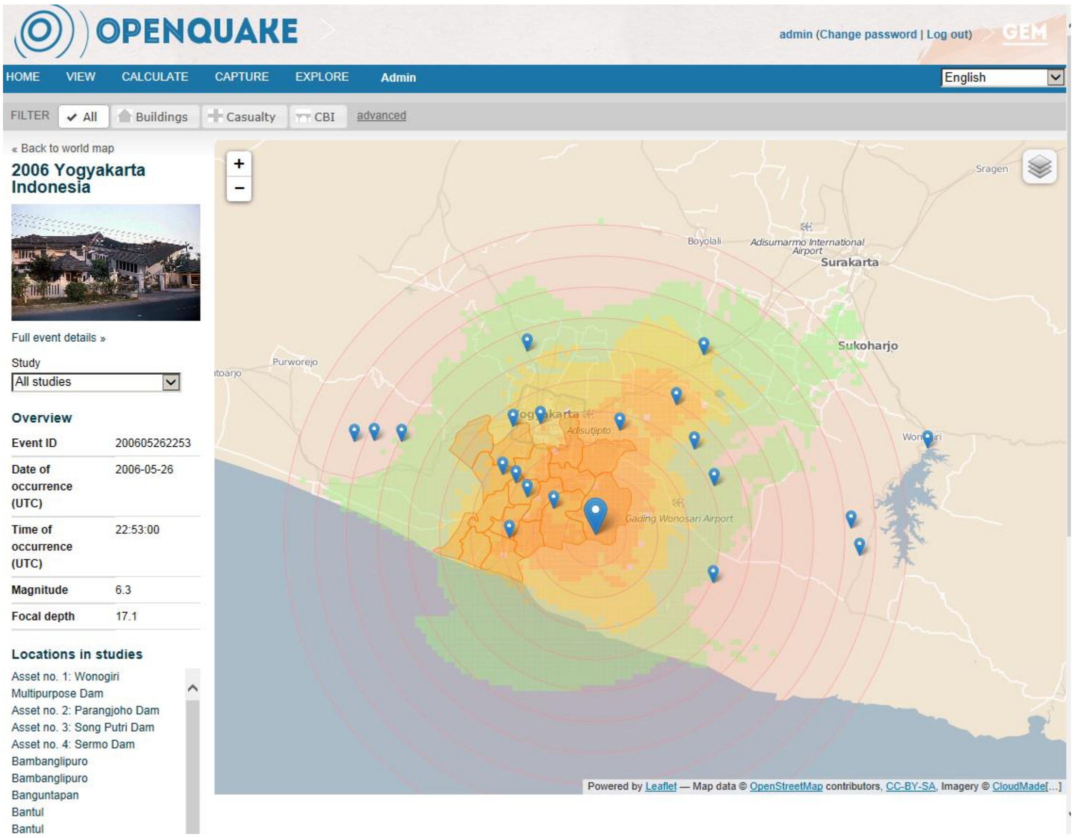


Figure 5. Event overview page with a USGS ShakeMap overlay

Once in the event details page, as shown in Figure 6, the key facts will be displayed as well as the following items:

1. Event introduction
2. Impact summary
3. Demographics
4. Human impact
5. Building damage
6. Socioeconomic loss
7. Key event facts
8. Main Data Source studies
9. Epicentre location map

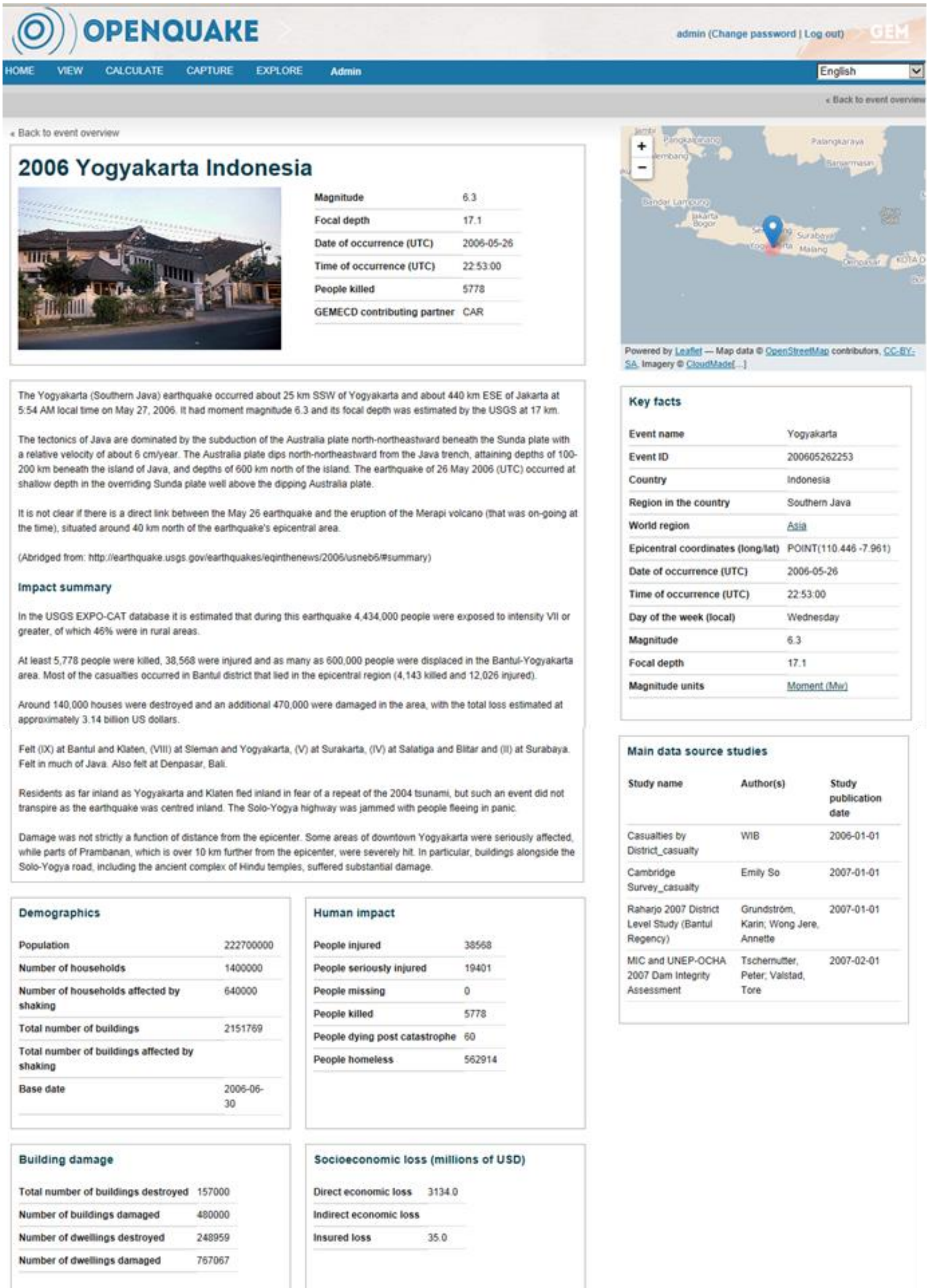
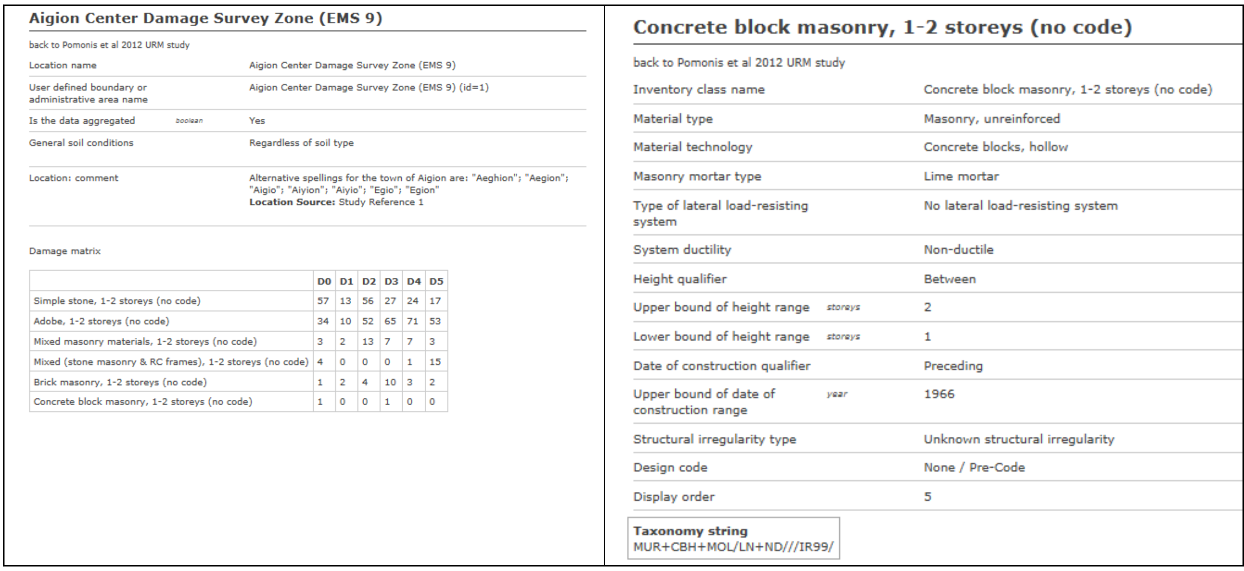


Figure 6. Event Key Facts for 2006 Yogyakarta, Indonesia

Alternatively, by clicking on the name of the study one wishes to view or the locations on the map from the events overview page shown in Figure 5, the user would be directed to the specified location. This can be in the form of a building damage matrix, a casualty matrix, critical facilities asset damage study. Figure 7 shows one of a buildings damage study, carried out after the 1995

Aigion earthquake in Greece by types of structure, which can be mapped to the standardised GEM taxonomy used in other GEM projects.



Aigion Center Damage Survey Zone (EMS 9)

back to Pomonis et al 2012 URM study

Location name		Aigion Center Damage Survey Zone (EMS 9)
User defined boundary or administrative area name		Aigion Center Damage Survey Zone (EMS 9) (id=1)
Is the data aggregated	boolean	Yes
General soil conditions		Regardless of soil type
Location: comment		Alternative spellings for the town of Aigion are: "Aeghion"; "Aegion"; "Aigio"; "Aiyion"; "Aiyio"; "Egio"; "Egion" **Location Source:** Study Reference 1

Damage matrix

	D0	D1	D2	D3	D4	D5
Simple stone, 1-2 storeys (no code)	57	13	56	27	24	17
Adobe, 1-2 storeys (no code)	34	10	52	65	71	53
Mixed masonry materials, 1-2 storeys (no code)	3	2	13	7	7	3
Mixed (stone masonry & RC frames), 1-2 storeys (no code)	4	0	0	0	1	15
Brick masonry, 1-2 storeys (no code)	1	2	4	10	3	2
Concrete block masonry, 1-2 storeys (no code)	1	0	0	1	0	0

Concrete block masonry, 1-2 storeys (no code)

back to Pomonis et al 2012 URM study

Inventory class name		Concrete block masonry, 1-2 storeys (no code)
Material type		Masonry, unreinforced
Material technology		Concrete blocks, hollow
Masonry mortar type		Lime mortar
Type of lateral load-resisting system		No lateral load-resisting system
System ductility		Non-ductile
Height qualifier		Between
Upper bound of height range	storeys	2
Lower bound of height range	storeys	1
Date of construction qualifier		Preceding
Upper bound of date of construction range	year	1966
Structural irregularity type		Unknown structural irregularity
Design code		None / Pre-Code
Display order		5

Taxonomy string
MUR+CBH+MOL/LN+ND///IR99/

Figure 7. Location details showing buildings damage matrix and the details of a specific building type, mapped to the GEM taxonomy from the Pomonis et al. (2012) study for the Aigion earthquake in 1995.

Secondary-Induced Hazards

It was the ambition of the GEMECD consortium to collect observational data on ground failures, i.e. landslides, fault ruptures, liquefaction, and induced effects such as tsunamis and fire following, with descriptions, photographs and locations to enable their display on maps and linkage to levels of ground shaking and to more detailed reports for 25+ earthquakes (taken from the original proposal). To date there are 14 events where secondary-induced hazards information and studies have been uploaded, either in form of a map of induced landslides from the event, e.g. landslide data set from Keefer, 2002 for the 1989 Loma Prieta, California, M6.9 earthquake shown in Figure 8, or as studies of consequences (building damage or casualties) due to secondary- induced hazards.

Overall, the collection of secondary-induced hazards was difficult, especially in the location of good quality data and specific consequences studies that could be geocoded and accompanied by photos of each affected asset. First, there are relatively few useful datasets available, either mapped or (better) in digital form. Second, most of these datasets are not comprehensive, nor complete. We refer to comprehensive as mapping all, for example, landslide occurrences over a specified dimension (e.g., 1 m^2) and complete as covering the entire domain over which landslides might have occurred for a given earthquake. These terms are essential for accurately modelling landslide occurrence from a statistical standpoint. Often these conditions must be inferred for a given study, since standards for the collection of these data do not exist. All USGS studies (Harp and Jibson, 1996) are publicly available and can be freely redistributed, but few other studies provide secondary hazard data with comparable liberal use permissions. For example, despite successfully gathering datasets for some key earthquakes, USGS were unable to release some of the data due to the significant limitations and challenges in collecting and redistributing landslide geospatial databases. Therefore in terms of secondary-induced hazard geospatial databases, the GEM community may need to acquire data for some key older events in a piecemeal fashion. As found in some cases, this requires obtaining permission from the authors in order to allow redistribution. This is being investigated at present.

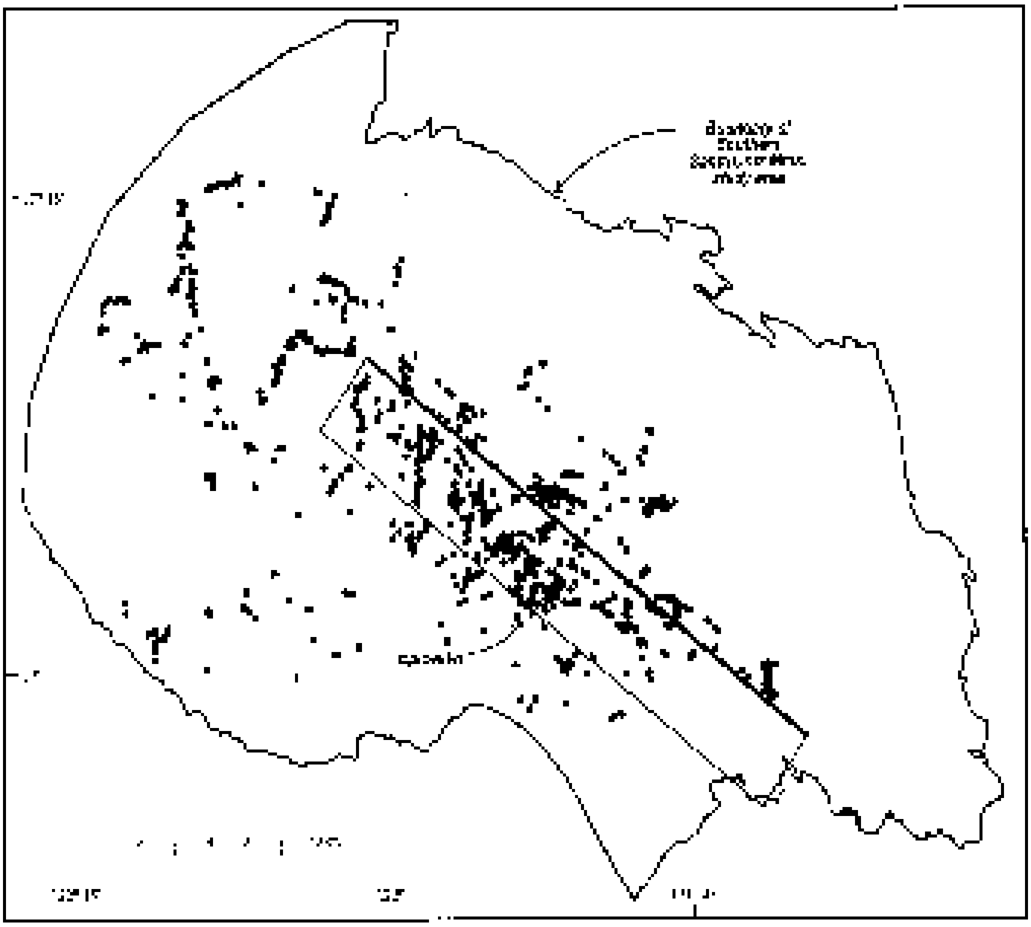


Figure 8. Example of a landslide data set (from Keefer, 2002) for the 1989 Loma Prieta, California, M6.9 earthquake

DATA UPLOAD

To aid data upload into the central database, four templates for data upload have been created:

1) aggregated building damage study due to ground shaking"
2) asset by asset damage study for critical buildings and infrastructure (CBI)
3) human casualty studies
4) secondary hazard damage studies (landslides, fire following earthquake, tsunami and liquefaction)

These templates constitute an effort to streamline the workflow of populating the database. The templates, which have gone through multiple iterations, are designed to consolidate and systematise all data required for damage studies. In addition to presenting all required data fields in an organised fashion, the templates also include detailed formatting instructions that allow the users to prepare and collate their information in compliance with the specification of the GEMECD database and the GEMECD data collection guidelines. All of the technical documentation and templates are part of GEMECD and were deliverables to the Global Earthquake Model.

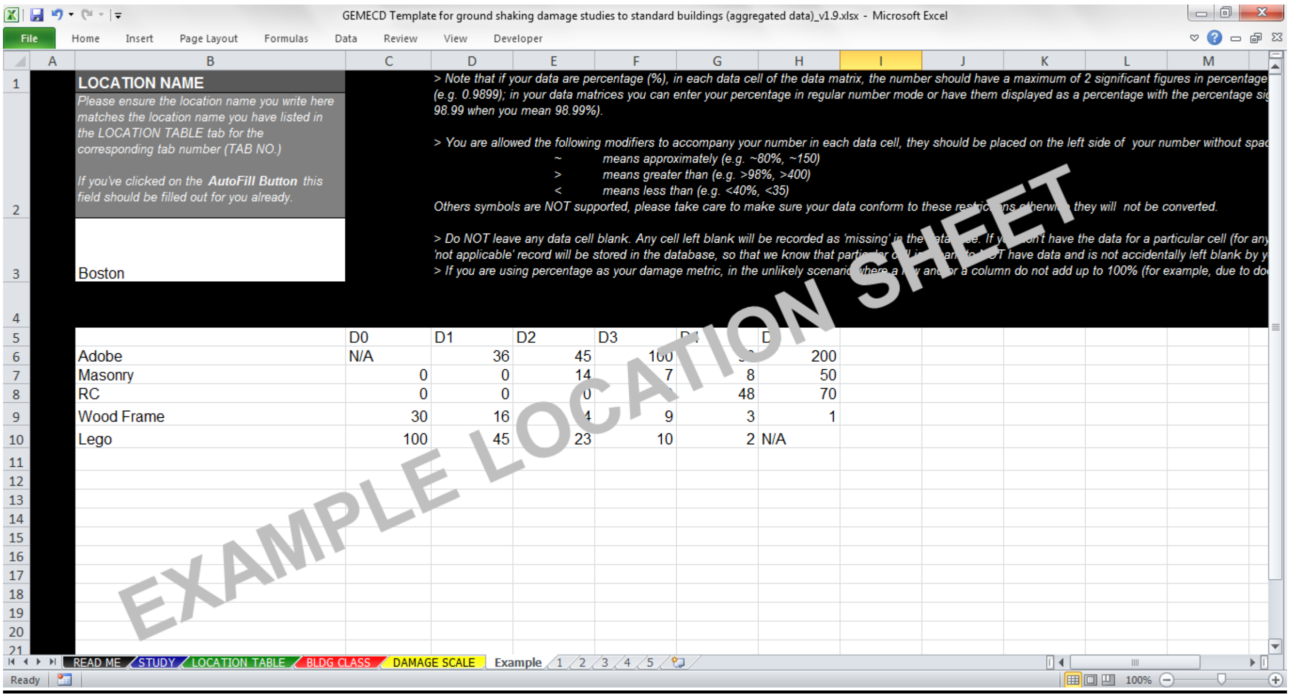


Figure 9. Building damage page for a specific location in the study (Excel Macro template)

OTHER FEATURES IN GEMECD

Apart from being able to view spatial and tabular information by event and individual studies, the GEMECD also allows for cross event analyses. Three groups of users were identified for the development of cross event analyses tools in GEMECD, they are:

1. non-expert users (or those identified by GeoHazards International in their user needs assessment) or someone who just wants to know about earthquakes in a particular country or period perhaps filtered by their consequences,
2. professionals interested in drawing out vulnerability data for a particular building type, and
3. the insurance industry.

The cross event analyses tools in GEMECD are still being developed in conjunction with OpenQuake, which will be used to generate analyses and aggregation of data in tabular and graphic forms. At the event levels, parameters from the Event Overview and Socio-Economic parameters can be filtered. These can take the form of events by magnitude ranges, events by total economic loss, events by insured loss, events by fatalities, events by number of buildings collapsed or destroyed etc.

At a more detailed survey level, the CEA tools have analysis capabilities to search and plot for specified events, studies or by aggregated building classes/ damage levels. One issue that has not been resolved and will require further work is a unified international building damage scale. One of the key requirements of GEMECD is that all data are inputted as depicted in the original studies but where possible, partners mapped these to the European Macroseismic Scale, EMS-98. Though there are over 80 damage surveys to standard buildings due to ground shaking in the database, only 13 such studies were carried out using EMS-98, as the scale came to use only in recent years. The examples therefore developed for the CEA tools will be based on the studies across circa 10 events. It should however be possible to add some more studies and events after the unified mapping of the damage grades has been completed, as these tools will be integrated within OpenQuake.

Accompanying the GEMECD is also a fully searchable literature database of all the documents from which the data in the database have been extracted from, either in English or in its original language.

IN CLOSING

The purpose of this paper is to provide an introduction to the GEM Earthquakes Consequences Database (GEMECD). GEMECD has been created to provide a standardised repository for future users to view past consequences of significant earthquakes in the past 40 years and also to record and submit their own studies, using a common platform**.** Housed within a public GIS platform, it is hoped that GEMECD will serve as a benchmarking tool for analytical loss models and to support the development of tools to create vulnerability data appropriate to specific countries, structures, or building classes in the future as part of the Global Earthquake Model, GEM.

REFERENCES

Allen TI, Wald DJ, Hotovec AJ, Lin K, Earle PS and Marano KD (2008) "An Atlas of ShakeMaps for selected global earthquakes*". Open-File Report 2008-1236*, U.S. Geological Survey, Colorado, 41 pp.

García, D., D.J. Wald, and M.G. Hearne (2012). A global earthquake discrimination scheme to optimize ground-motion prediction equation selection. *Bull. Seism. Soc. Am.* 102, 1, 185-203.

Garcia D, Wald DJ, Marano K, Mah R, Johnson K and Hearne MG (2014). ShakeMap Atlas 2.0., U.S. Geological Survey, USGS *Open-File Report*, in progress.

Johnson, K.L., D. García, C.B. Worden, K.W. Lin, R. Mah, K.D. Marano, M. Hearne, and D.J. Wald (2012). ShakeMap best practices: historic and modern events. *Seism. Res. Lett.* 83, 2, 407.

Harp EL, and Jibson RW (1996). Landslides triggered by the 1994 Northridge, California, earthquake. Bull *Seism. Soc. Am*, 86, 1B, S319–S332.

Keefer DK (2002) Statistical analysis of an earthquake-induced landslide distribution – the 1989 Loma Prieta, California event. Eng. Geology 58, 213–249.

Pomonis A, Gaspari M. and Karababa FS (2012) "Seismic vulnerability assessment for buildings in Greece based on observed damage data sets" *Bollettino di Geofisica Teorica ed Applicata*, 54.

Website reference:

GEM NEXUS http://www.nexus.globalquakemodel.org/