A Global Building Inventory for Earthquake Loss Estimation and Risk Management

Kishor Jaiswal,[a] M.EERI, **David Wald,**[b] M.EERI, **and Keith Porter,**[c] M.EERI

We develop a global database of building inventories using taxonomy of global building types for use in near-real-time post-earthquake loss estimation and pre-earthquake risk analysis, for the U.S. Geological Survey's Prompt Assessment of Global Earthquakes for Response (PAGER) program. The database is available for public use, subject to peer review, scrutiny, and open enhancement. On a country-by-country level, it contains estimates of the distribution of building types categorized by material, lateral force resisting system, and occupancy type (residential or nonresidential, urban or rural). The database draws on and harmonizes numerous sources: (1) UN statistics, (2) UN Habitat's demographic and health survey (DHS) database, (3) national housing censuses, (4) the World Housing Encyclopedia and (5) other literature. [DOI: 10.1193/1.3450316]

INTRODUCTION

In order to assess the impact of an earthquake on the built environment, it is essential to know the structural systems of buildings and their performance in past earthquakes, engineering standards adopted during construction, and the location and distribution of vulnerable building stock in the shaken area. It is evident from past fatal earthquakes around the world that the existence of vulnerable buildings in high intensity areas has in most cases controlled the total human losses. For example, collapsed adobe and masonry buildings caused more than 90% of the approximately 26,000 deaths in the M6.6 Bam, Iran, earthquake of 2003 (Kuwata et al. 2005). Among the 25,000 fatalities in the 1988 Spitak, Armenia, earthquake, collapse of 72 precast concrete-framed buildings in Leninakan (52% of the total), and of 43 of composite precast concrete-framed buildings with stone masonry infill walls (73% of all buildings in Spitak) dominated the casualties (Krimgold 1989). Collapse of weak masonry and reinforced concrete-framed construction in the Bhuj, India earthquake of 2001 resulted in 80% of total fatalities (Madabhushi and Haigh 2005). Similarly, among the 2,360 deaths caused by the 1999 M7.8 Chi-Chi Taiwan earthquake, 94% of the victims died because of building collapse, and 43.5% of those victims were residing in mud-brick residences (Tien et al. 2002). Earthquakes with similar population exposure killed far fewer people in countries with less-vulnerable construction; see Table 1 for comparisons. Many variables other than structure types contribute to the contrast in the death toll. Although the construction

[a] U.S. Geological Survey, Golden CO (contracted through Synergetics Incorporated) kjaiswal@usgs.gov
[b] U.S. Geological Survey, Golden CO
[c] University of Colorado at Boulder

Earthquake Spectra, Volume 26, No. 3, pages 731–748, August 2010; © 2010, Earthquake Engineering Research Institute

Table 1. Samples of comparable events with very different fatality rates. Columns labeled "Pop MMI 6+" and "MMI 8+" indicate the estimated population (in 1,000 s) exposed to shaking intensities of 6 or greater and 8 or greater at the time of the earthquake. "Dead" indicates reported earthquake fatalities, also in 1,000 s

Event	Pop MMI 6+	Pop MMI 8+	Dead		Compare With	Pop MMI 6+	Pop MMI 8+	Dead
M6.6 Bam 2001	136	81	26.3	↔	M7.7 Guam 1993	203	142	0
M6.7 Spitak 1988	548	47	25	↔	M6.5 Imperial Valley 1979	720	100	0
M7.6 Bhuj 2001	2,800	664	20	↔	M7.1 Philippines 1994	1,600	570	0.08

practice in the worst cases described are not indicative of modern design and construction, construction does matter, probably more than any other variable in most earthquakes. And yet little is known on a consistent basis about building stocks worldwide; there is no public, open, global database of construction practice, use, and occupancy.

Various inventory data do exist. Shakhramanian (2000) discusses an effort by Emercom of the Russian Federation to develop the global loss modeling software Extremum, which includes a global inventory that characterizes building stocks according to seismic resistance, but the software and its underlying data are unavailable for public use or scrutiny. HAZUS-MH (FEMA 2006) contains estimated building stocks in the United States by 128 categories, and while the data are not easy to extract from HAZUS, they are freely available and well documented. The HAZUS methodology for creating its building-stock database is complex, but in overview, it involves using housing, population, and economic census and other economic data to estimate population by census area and each of 28 occupancy types. Population is then multiplied by estimates of average square footage per person by occupancy class to arrive at square footage by census area by occupancy class, which is then distributed among 128 structure types using engineering judgment. The process is time-consuming and country-specific, and would be difficult to replicate and validate worldwide.

A building-stock database exists for Istanbul, Turkey, created at least in part by contractors examining individual buildings using a modified form of the FEMA (ATC 2002) rapid visual screening instrument (Yakut 2004). Geoscience Australia (GA) is developing a national building-exposure database, using in part a national, geo-coded address file of residences, along with building-specific data (e.g., address, number of occupants, type of dwelling and its valuation) from the Australian Bureau of Statistics (ABS) and the Australian Housing Survey (AHS; Nadimpalli et al. 2002). For commercial and industrial buildings, it will rely on GA survey studies, the CityScope database, and a generic approach to estimate the structural properties, business information, and values.

Models developed by and for the insurance industry are also said to contain estimates of portions of the building stock in various countries, but these are likewise publicly unavailable. Various sources listed in Table 2 and Table 3 and others discussed later provide some relevant information, but were generally not intended to inform engineer-

Table 2. Data sources. High quality refers to data compiled from engineering or telephonic surveys, field visits for ground-truth, or data from local engineering experts; Medium refers to data from general field surveys and assignments not based on engineering standards; Low refers to data from non-engineering agencies that are not specifically meant for engineering risk analysis. We evaluated each data source discussed below using these criteria and carried out quality assignment. Readers are referred to Jaiswal and Wald (2008) for further details on procedure adopted for quality ratings of each source.

Sr. No.	Source of Data	Building Stock	Quality	Global Coverage
1.	**World Housing Encyclopedia** (developed by EERI, USA)	**1. Residential:** a. Construction type b. Occupancy	 High Medium	110 residential construction types in 37 countries. Exact fraction of each housing type per country is unknown. Day and night occupancy by construction type is available.
2.	**UN Database** (UN 1993, UN-HABITAT 2007)	**1. Residential:** a. Construction type b. Occupancy	 Medium Low	Data for 44 countries give fraction of housing units in that country by exterior wall material. Data for 110 countries give average number of people per housing unit.
3.	**Census of Housing** (data compiled from housing census statistics)	**1. Residential:** a. Construction type b. Occupancy **2. Non-residential:** a. Construction type b. Occupancy	 Medium Low Medium Low	197 countries conducted housing census in 1990. Several do not publish housing statistics online. Most of the Census surveys do not include information about non-residential building inventory.
4.	**Published Literature** (e.g., research articles; reports; non-proprietary information)	**1. Residential:** a. Construction type b. Occupancy **2. Non-residential:** a. Construction type b. Occupancy	 High Low Medium Medium	~15 countries contained high quality information based on survey and verification of other published information (e.g., census/tax assessor's data). The day and night time occupancy by construction type is not available.
5.	**PAGER-WHE Project** (http://pager.world-housing.net)	**1. Residential:** a. Construction type b. Occupancy **2. Non-residential:** a. Construction type b. Occupancy	 High Medium High Medium	To date, EERI has helped to compile a database of more than 25 countries through a survey and internal review.

ing risk analysis. There appears to be no comprehensive public-domain building inventory appropriate for earthquake loss estimation. Under the auspices of the Prompt Assessment of Global Earthquakes for Response (PAGER) project, the U.S. Geological Survey (USGS) set out to create such a database.

Table 3. Other publications providing country-specific building-stock data

Sr. No.	Country	Vintage	Data Source
1.	Albania	2001	2001 Albania Housing Census data
2.	Algeria	1983	Petrovski (1983)
3.	Argentina	2002	Rodriguez et al. (2002)
4.	California (USA)	2002	HAZUS-MH (FEMA 2006)
5.	Iran	2005	Tobita et al. (2007)
6.	Iraq	1983	Petrovski (1983)
7.	Italy	2006	Dolce et al. (2006)
8.	Jordan	1983	Petrovski (1983)
9.	Pakistan	2008	Maqsood and Schwarz (2008)
10.	Russian Federation	2000	Shakhramanian et al. (2000)
11.	Saudi Arabia	1983	Petrovski (1983)
12.	Sudan	1983	Petrovski (1983)
13.	Syrian Republic	1983	Petrovski (1983)
14.	Taiwan	2002	Tien et al. (2002)
15.	Turkey	2002	Bommer et al. (2002)
16.	United States of America	2002	HAZUS-MH (FEMA 2006)

OBJECTIVES

The PAGER system already produces ShakeMaps shortly after all global earthquakes ($M \geq 5.5$), and estimates the population exposed to various levels of MMI (Wald et al. 2005, 2008). It is developing the capability to estimate fatalities as well, for use by planners and emergency-response decision-makers. PAGER's objective is currently to achieve an order-of-magnitude accuracy in the fatality estimate, which can be highly valuable in the days or weeks after an earthquake until ground truth is available. One route to creating those fatality estimates is to estimate building stocks exposed to shaking and the population within each type at the time of the earthquake, along with the shaking intensity to which the buildings and people were exposed. By developing and applying seismic vulnerability functions (e.g., Porter 2009; Jaiswal and Wald 2010) that relate shaking to damage and damage to fatalities, the PAGER program will produce fatality estimates within minutes after the occurrence of earthquakes, worldwide, for emergency-management purposes. Every stage of the estimate will be based on peer-reviewed methods and open data.

The present work addresses creation of a regional building inventory and its spatial, structural, and occupancy characteristics. Our objective is to create and publicly disseminate a database in the format shown in Table 5. The database contains four tables, each containing the country-specific distribution of the population among different structure types. One table is required for each combination of urban or rural, residential or non-residential construction.

The objectives for the inventory database include a few key noteworthy features. It is to be global, open, public, and collaboratively developed, with country-level geographic resolution, which would seem to make it the first such database with these characteristics. Second, its structure type category system must be nearly exhaustive, i.e., capable of including structure types common anywhere in the world. Its taxonomy must be collapsible, in that it should allow for categories with detailed characteristics such as height (where such data are available), and also include aggregate categories for structure types whose details are unavailable (especially when obtained through housing census or similar surveys). This will allow for somewhat greater versatility in loss estimates than if the taxonomy were defined only at detailed or only at an aggregate level. We set out to include a quality assignment (low, medium, high) and data vintage, reflecting varying confidence in the available data and our interpretation of it. Finally, we sought to make the database modular, in that it can be incrementally improved as better data become available.

METHODOLOGY

With the objectives and available literature and data in mind, we now present the PAGER inventory development, in three phases:

1. Data acquisition, preparation and confidence rating
2. Data aggregation and quality ranking
3. Data assignment for missing entries

The reader is referred to Jaiswal and Wald (2008) for more detail on the process of inventory distribution, along with sample calculations performed for one of the countries.

PHASE I: DATA ACQUISITION, PREPARATION AND CONFIDENCE RATING

This task began with the acquisition of the source data listed in Tables 2 and 3. Each source in the tables provides, for one or more countries, an estimate of the fraction of urban, rural, residential, or nonresidential construction represented by each of several structure types. To adapt these disparate data sources to populate the database outlined in Table 3 required converting the various building category systems and quantities of construction from their native format to a common, globally applicable taxonomy.

Before continuing, it is necessary to establish such a building taxonomy. Several building typologies exist. Prominent among these are: ATC-13 (ATC 1985); EMS-98; ATC-14 (ATC 1987), which was used with modifications in *FEMA 154* (ATC 2002), HAZUS-MH (FEMA 2006), and other FEMA-funded efforts; the World Housing Encylopedia (WHE, http://www.world-housing.net/); and RISK-UE 2004 (Mouroux et al. 2004).

Any taxonomy is a compromise between simplicity and thoroughness. For PAGER's purposes, PAGER-STR builds upon several of these existing typologies, harmonizes construction types that appear in several category systems, and adds a few that are absent from existing schemes, a prerequisite for inventory development at a global scale. In particular, it merges the FEMA taxonomy (e.g., *ATC-14, FEMA 154,* etc.) with those of WHE (for unreinforced masonry types; refer to http://www.world-housing.net/) and EMS-98 (for some European types that do not appear in the first two systems; European

Seismic Commission Working Group on Macroseismic Scales 1998). A few types were added that appear in none of these taxonomies. The resulting list is long: 54 types, 11 of which have the option for an additional, three-category height suffix—low-rise, midrise, and high-rise as shown in note a, and one type has the option for 2-category height suffix as shown in note b of Table 4—for a total of 89 types. This greatly exceeds that of WHE (10 types), EMS-98 (15 types), *FEMA 154* (15 types times three height categories), and ATC-13 (40 types), but on the other hand is smaller than HAZUS-MH (36 types each with up to seven design levels). Some distinctions might seem minor (e.g., those based on the type of mortar), but we have generally included those that either were common to the existing taxonomies or because post-earthquake reconnaissance workers or others noted that the distinctions seemed important to seismic performance. Furthermore, as noted in our objectives, the taxonomy is collapsible, with nine aggregate types W, S, C, etc., that comprise two or more detailed types also shown in the taxonomy. It seems likely that the PAGER STR taxonomy will evolve, as did several of its predecessors. In the next version currently in development, for example, it will include confined masonry.

1. ***Map building category system.*** The building category system of the source data was mapped to PAGER-STR (Table 4). Several datasets such as the UN Database (1993) describe buildings by predominant wall material (wood, concrete, brick, stone etc.). For example, housing units with brick and mud as a construction material of external walls have been mapped with “Unreinforced fire brick masonry with mud mortar” or UFB1 structure type. Even though there are large numbers of building types in the PAGER-STR list, most countries’ building stocks can be adequately represented for PAGER’s limited purposes by a few types, selected via one-to-one mapping from types named in the source data to appropriate PAGER-STR types. To create the mapping we used our judgment, informed by various sources such as Web-accessible photos of construction, World Housing Encyclopedia housing prototypes, and other miscellaneous text describing common country-specific building practices. Whenever the available data seemed to indicate details such as structural system, number of stories, etc., we mapped to a detailed PAGER-STR category rather than its aggregate type. Figure 1 illustrates the mapping of building description found in housing census data compiled by statistical center of Iran to equivalent PAGER-STR categories. For example, the steel skeleton structure type in Iran is mapped to aggregate steel construction (S) category due to lack of additional information about structural system. Similarly the brick-steel and stone-steel are mapped to UFB4 type, which represents brick masonry with rigid diaphragms. Brick-wood is mapped to UFB3, a gravity load bearing masonry wall structure with a timber-frame system. The cement block construction is mapped to UCB, and reinforced concrete structure is mapped to C3 as shown in Table 4. The mapping scheme is developed for each data source and is detailed in Jaiswal and Wald (2008) along with their country-specific applications.
2. ***Map quantities of construction.*** The source data quantify the building stock in a variety of measures: housing units, buildings, building volume, etc. These disparate quantification systems were mapped to PAGER’s, which because it is concerned first with human impacts quantifies the fraction of population by

Table 4. PAGER-STR, structure types used for developing the PAGER global inventory database

LABEL	DESCRIPTION
W	**Wood**
W1	Wood frame, wood stud, wood, stucco, or brick veneer
W2	Wood frame, heavy members, diagonals or bamboo lattice, mud infill
W3	Wood frame, prefabricated steel stud panels, wood or stucco exterior walls
W4	Log building
S	**Steel**
S1	Steel moment frame[a]
S2	Steel braced frame[a]
S3	Steel light frame
S4	Steel frame with cast-in-place concrete shear walls[a]
S5	Steel frame with unreinforced masonry infill walls[a]
C	**Reinforced Concrete**
C1	Ductile reinforced concrete moment frame[a]
C2	Reinforced concrete shear walls[a]
C3	Nonductile reinforced concrete frame with masonry infill walls[a]
C4	Nonductile reinforced concrete frame without masonry infill walls[a]
C5	Steel reinforced concrete (steel members encased in reinforced concrete)[a]
PC1	Precast concrete tilt-up walls (low rise)
PC2	Precast concrete frames with concrete shear walls[a]
TU	Precast wall panel construction (mid to high rise, former Soviet Union style)
RM	**Reinforced Masonry**
RM1	Reinforced masonry bearing walls with wood or metal deck diaphragms[b]
RM2	Reinforced masonry bearing walls with concrete diaphragms[a]
MH	**Mobile Homes**
M	**Mud Walls**
M1	Mud walls without horizontal wood elements
M2	Mud walls with horizontal wood elements
A	**Adobe Block (Unbaked Dried Mud Block) Walls**
A1	Adobe block, mud mortar, wood roof and floors
A2	Same as A1, bamboo, straw, and thatch roof
A3	Same as A1, cement-sand mortar
A4	Same as A1, reinforced concrete bond beam, cane and mud roof
A5	Same as A1, with bamboo or rope reinforcement
RE	**Rammed Earth/Pneumatically Impacted Stabilized Earth**
RS	**Rubble Stone (Field Stone) Masonry**
RS1	Local field stones dry stacked (no mortar). Timber floors. Timber, earth, or metal roof.
RS2	Same as RS1 with mud mortar.
RS3	Same as RS1 with lime mortar.
RS4	Same as RS1 with cement mortar, vaulted brick roof and floors
RS5	Same as RS1 with cement mortar and reinforced concrete bond beam.
DS	**Rectangular Cut Stone Masonry Block**

Table 4. (cont.)

LABEL	DESCRIPTION
DS1	Rectangular cut stone masonry block with mud mortar, timber roof and floors
DS2	Same as DS1 with lime mortar
DS3	Same as DS1 with cement mortar
DS4	Same as DS2 with reinforced concrete floors and roof
UFB	**Unreinforced Fired (baked)Brick Masonry**
UFB1	Unreinforced brick masonry in mud mortar without timber posts
UFB2	Unreinforced brick masonry in mud mortar with timber posts
UFB3	Unreinforced fired brick masonry, cement mortar, timber or timber-and-steel gravity system
UFB4	Same as UFB3, but with reinforced concrete floor and roof slabs
UCB	**Unreinforced Concrete Block Masonry, Lime/Cement Mortar**
MS	**Massive Stone Masonry in Lime/Cement Mortar**
INF	**Informal Construction (makeshift dwellings, made from plastic/GI sheets or other material)**
UNK	**Unknown (Not specified)**

[a] Suffix for height: L=1–3 stories, M=4–7 stories, H=8+stories.
[b] Suffix for height: L=1–3 stories, M=4–7 stories.

structure type, density (urban or rural), and use (residential or nonresidential). For example, PAGER quantifies the fraction of the urban population that dwells in the given structure type in a given country, recognizing that significant fractions of the population do not have workplaces or work outdoors, and assuming that most of the population do dwell in buildings. We generally equated fraction of housing units, or of volume, etc., with fraction of population. For example, where a source said 10% of housing units are adobe, we equated that with 10% of the population living in adobe dwellings. Where a source made no distinction between urban and rural construction, we imposed none. Where a source purported to reflect only a portion of a country (e.g., field surveys such as Faccioli et al. 1999; Ozmen 2000; Tobita et al. 2007), and no other sources were avail-

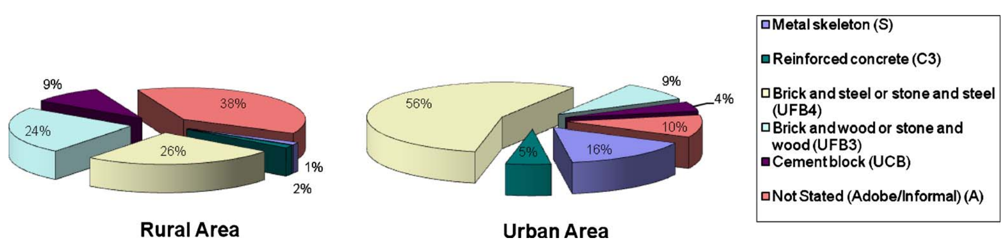

Figure 1. Housing stock distribution and attribute mapping (shown in brackets) for rural and urban areas of Iran (Source: Statistical Center, Iran).

Table 5. Layout of PAGER building inventory database

FIELDS	TYPE	COMMENT
1. Sr. No.	Integer	Table index
2 to 6. Country	Text & Integer	ISO 2-character & 3 character country code (ISO 3166-1), ISO numeric country code and 2 fields for country name
7. PAGER vulnerability code	Integer	Low, medium or high
8. Rating	Text	Low, medium or high
9 to 14. Source & country-pairing information	Text & Integer	Original = data available directly for this country Neighbor = taken from neighboring country WHE-Survey = from PAGER-WHE Project; By Judgment = from PAGER-WHE neighbor
15. Source file	Text	Name of the appropriate file used for development of particular dataset
16. Year	YYYY format	Year in which the data were published
17 to 105. Structure type population	Float	Fraction of population (0.0 to 1.0; urban, rural, residential, or nonresidential) in the given structure type. Each column (17 to 105) refers to a different structure type.

able for that country, we applied the data to the entire country and applied a lower quality rating (discussed later).

3. ***Label data quality***. We assigned a high, medium, or low data-quality rating to each data source. The rating is based on (a) the procedure used to collect the field data in the first place (ordinary or local survey versus engineering survey), (b) the objective of the original data collection (gathering census or other demographic information versus informing an engineering loss estimate), and (c) the degree to which engineering experts were involved in the raw data compilation. High quality refers to data compiled from engineering surveys, possibly by telephone or field observations; or other data from local engineering experts. Medium quality data are those that are generally indicative of structure type and compiled by general field survey, but not by or for engineers. Low quality refers to everything else.

At this stage, many countries have some information available to populate fields 17 through 105 of Table 5. Where data are missing, e.g., information is lacking about the fraction of the population working in the given structure types, the relevant fields in the database are assigned a “null” value. In some cases, there is more than one source available to populate a given field, so the data sources are not yet actually merged into the final database. At the end of this phase, each data source and the relevant information contained therein has been identified, rated, and compiled for further processing in Phase II.

PHASE II: DATA PRIORITIZATION, MERGING, AND COUNTRY ASSIGNMENTS

In the previous step we compiled available data sources and mapped their attributes to PAGER structure types, quantities, and use classes, and assigned to each a quality level. In Phase II these sources are merged. For each country, if only a single database provides information for a certain attribute, it is assigned directly to the PAGER inventory database along with its quality rating. If none of the databases provide information for certain attribute of a country, we assign the missing information based on an *a priori* country-pairing described in the next phase.

If multiple data sources address a particular country and attribute, then the data with the higher quality rating are employed and the other data are not used. If two data sources have the same quality rating, then the newer data are employed and the older data ignored. If the data have the same quality and vintage, then we preferred peer-reviewed data (e.g., published in archival journals) to other data. Attributes that still have no relevant data (i.e., still have a null or missing assignment) are flagged for further processing in Phase III, described next.

PHASE III: DATA DEVELOPMENT FOR MISSING ENTRIES AND SYNTHESIS

More than half of all countries have no direct information about building inventory distribution in the above-named resources. On a first-order basis, it seems reasonable to assign their inventory characteristics from neighboring countries. It is quite common to have similar construction practice and building characteristics among neighboring countries. How then to select the most appropriate neighbor with inventory data to copy (which we term country pairing)?

We compiled a first-order vulnerability ranking scheme by which to group countries into 5 regions of relatively uniform seismic vulnerability. Each country was labeled with a region number 1 through 5, with 1 being least vulnerable (e.g., California, New Zealand) and 5 the most vulnerable (e.g., Afghanistan, Iran). By "vulnerability" we mean the damage susceptibility of a structure (or building stock in general within a given region or country) at various levels of ground motion intensity, as opposed to other social or economic effects resulting from natural disasters.

Principal among many factors considered in making the assignment were building codes and enforcement and past seismic performance of buildings in the country. Countries known to have few highly vulnerable structures were assigned as a part of region 1. Countries known to have large quantities of questionable engineered construction were assigned to region 3. Countries with large fractions of the building stock constructed of adobe or rubble masonry were assigned to region 5. Regions 2 and 4 were used as in-

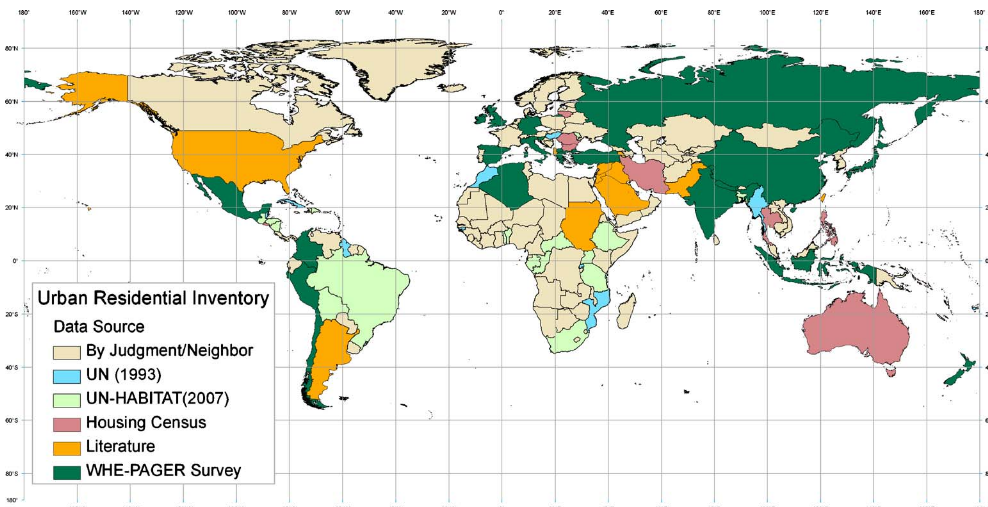

Figure 2. Global coverage of PAGER building inventory database showing data compiled from different sources.

termediates, when neither 1 & 3, nor 5 seemed to apply well. Although the scheme is broad, non-quantitative and subjective, it was used only for country pairing for building inventory assignments in cases where the data are missing.

To populate the building stock distribution of countries with missing data, we selected the neighbor in the same vulnerability region, with the most recent data and with highest quality rating. For more detail, see Jaiswal and Wald (2008). The completion of phase-III produced a complete global building inventory database with the format shown in Table 5. Figure 2 shows the primary data source for each country. Even though 44 countries initially had data relevant to construction material/type in the UN (1993) database, only a handful actually appear in the final compilation due to the quality and vintage of the data source as shown in Figure 2. The confidence rating map shown in Figure 3 illustrates 34 countries with high rating, 11 with medium, and the remaining countries with low rating for the urban residential building inventory. The building inventories that have high rating are generally from the countries where most of the earthquake related fatalities occurred during the last century and those are important for PAGER fatality estimation purposes.

INVENTORY UPDATING

The database described here requires routine updating as better data become available. We developed a tool for viewing and updating country-specific data, including its source, vintage, and quality (Figure 4). Ongoing efforts will improve the data quality and also fill missing data identified here. One such effort is the collaboration with the World Housing Encyclopedia, mentioned earlier, in which experts organized by the World Housing Encyclopedia are providing estimates of the distribution of building

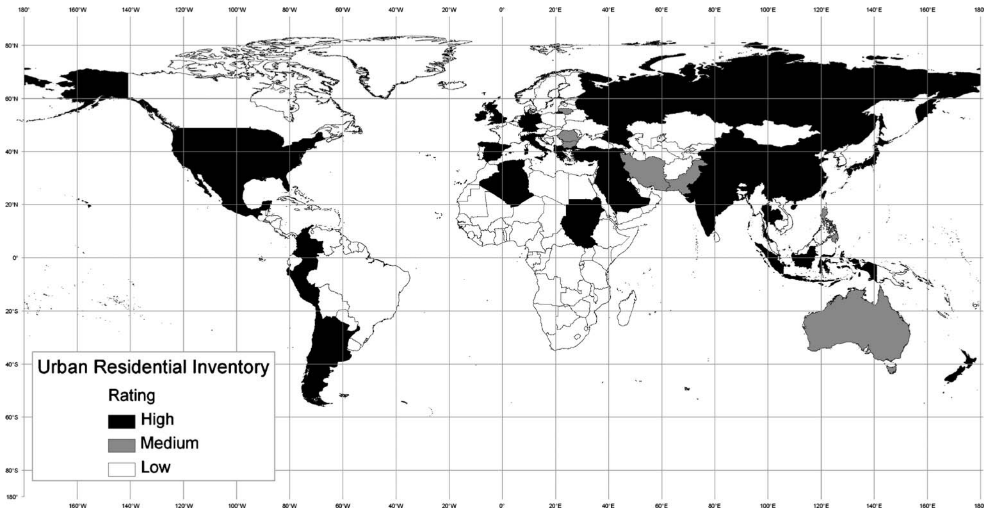

Figure 3. Map showing quality rating of the PAGER urban residential building inventory data for different countries.

stock and their seismic vulnerability on a country-by-country basis. The emphasis in this effort is on countries that are both high risk and data-poor. To date we have acquired through this effort estimated building-type distributions for more than 25 countries; that number is expected to exceed 35 by mid-2009, and could continue to increase. The growing responses of professionals especially from earthquake prone countries have clearly demonstrated the need and practicality of such an open exchange.

WEB DELIVERY OF GLOBAL BUILDING INVENTORY

The database developed during this investigation is being made available online at http://pubs.usgs.gov/of/2008/1160/. It allows engineers and professionals anywhere in the world to download the data and to contribute or suggest modifications for their respective countries. The feedback received from respective country experts will be utilized to upgrade the current, default PAGER inventory database.

IMPLEMENTATION CHALLENGES

That such a database did not yet exist in the public domain hints at the challenges to creating one. Principal among the challenges we faced were: disparate media, nomenclature, and geographic resolution, and the paucity of information about occupancy patterns by time of day. Some of the data we employed were only available in bound paper documents, some of which required help from United Nations officials to acquire. Digital documents included HTML, spreadsheets, and word-processing text documents including scanned PDF. Significant though straightforward effort was required to bring these into uniform format. Some data categorized building types in unfamiliar nomenclature idiosyncratic to the relevant country or local region. We addressed this problem

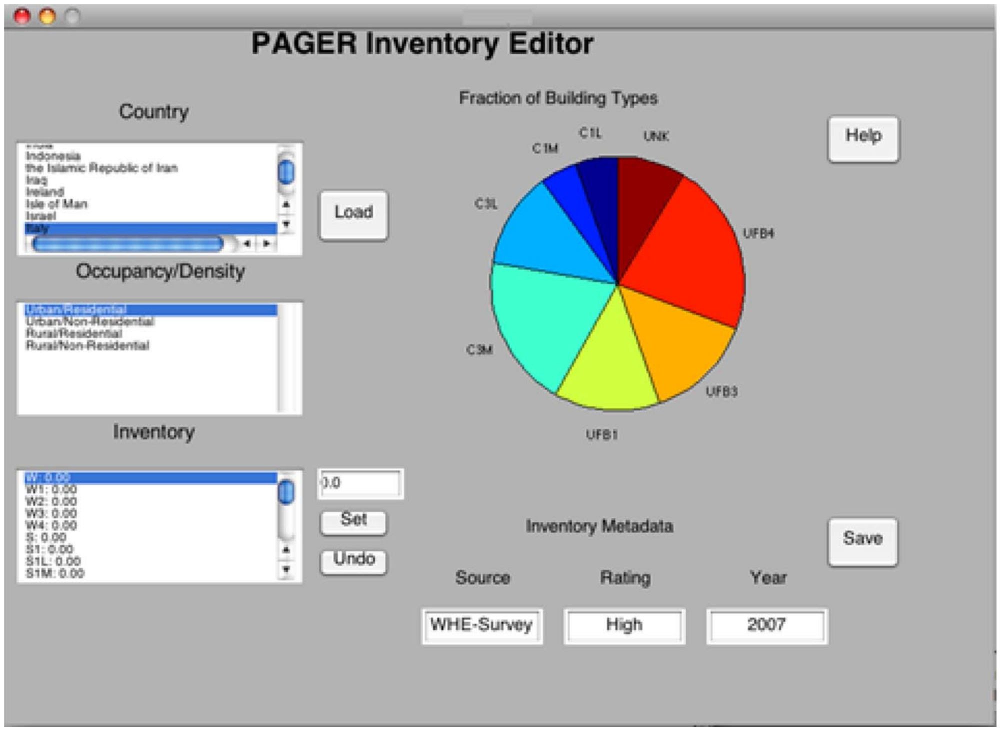

Figure 4. Screen shot of PAGER building inventory editor, showing inventory data for Italy along with source, rating and vintage of the data.

by Internet searches of construction-related websites with representative images of the construction, or by interpreting the literal English-language translation of the term in question. Problems of geographic resolution mostly had to do with interpreting earthquake-engineering literature that touches on construction characteristics in sub-country regions that had been affected by a particular earthquake. This kind of local information is problematic when it comes from large countries where construction likely varies significantly between climatic and economic regions.

Time-of-day occupancy patterns were particularly problematic because of the paucity of relevant data. PAGER requires occupancy information by construction type and time of day to estimate casualties. Average occupancy varies by time of day, day of week, and by type of occupancy (e.g., single—or multi-family residential, commercial, or industrial). Of the public data sources known to us, only the World Housing Encyclopedia provides the average occupancy of a construction type by time of day, and that too is limited to a few countries. We deduced missing information in part from data about average number of persons per housing unit in the UN database, and filled in the rest from engineering judgment.

Future efforts to enhance global building-stock models for catastrophe risk assessment might be aided by a global standard building taxonomy, participation by local engineering experts, and development of incentives and efficient means to contribute to and validate the database. It would be valuable to revisit the country-level building-type assignments for large countries. Finally, it would be valuable to revisit the data-quality-rating scheme discussed above with an approach that somehow quantifies error rates in building-type assignment for each kind of data source.

LIMITATIONS

This is an initial attempt to create a global, open inventory of building stocks for purposes of catastrophe risk modeling. It contains no hazard information, and therefore cannot by itself be used to estimate risk. It is far from being definitive, and is limited first by the narrow objectives for which it was created, second by the limited resources currently devoted to its creation, and third by the paucity of relevant data. As a consequence of our narrow objectives of estimating fatalities in future earthquakes, the database does not contain building attributes that would be of interest in other perils (e.g., window protection, elevation above grade, or flammability, relevant to wind, flood, or fire risk, respectively). It does not contain estimates of building area, value, or quantification of uncertainty, which would all be required to begin to approach the actuarial quality required for insurance risk modeling. However, we see no fundamental reason why these limitations cannot be overcome by others with comparable effort.

Likewise, we put only limited efforts into refining our mapping scheme for interpreting structure types in the UN or housing census databases. Better building data exist for many countries, but much of it is proprietary and is either cost-prohibitive to acquire, overly constrained in terms of making the relevant information publically available, or is simply unavailable. We made no attempt to purchase commercial building data. In any event such data are often focused on insured (and likely engineered) properties and thus may not be useful for overall country-wide inventory and impact assessment. Because of the limited resources available to us we made no attempt to employ satellite or aerial imagery to infer building characteristics.

More serious however are the limitations in the geographic scope and descriptive detail of the data we examined. For many countries, available construction information was only partially informative of important engineering details such as structural materials and lateral force resisting system. For many countries, building inventory and vulnerability data were either patchy or completely missing. The use of a country-pairing scheme to identify the "best neighbor" depending on data quality and vintage allowed the development of default inventory distributions for such countries. Such mapping of inventory distribution from one country to a neighboring country through an assumption of country-pairing was carried out as a last resort and these countries were assigned low quality ratings in our database. Our country-pairing was based purely on our judgment and must suffice until better quality data become available. Finally, there are undoubtedly more detailed local building surveys that we have not examined. Through current efforts of the World Housing Encyclopedia's WHE-PAGER project (e.g., Porter et al. 2008) and ongoing research we will attempt to collect and incorporate more of these; the methodology is designed to do so. These limitations undoubtedly contribute to error in loss estimates. The quantification of such error will be the subject of other PAGER publications.

SUMMARY AND CONCLUSIONS

The methodology described in this paper provides a relatively simple, low-cost framework for creating an open, global, country-level building inventory database that aggregates data sources with variable quality and vintage. The main purpose of the inventory is for earthquake fatality assessment for the USGS PAGER system with order-of-magnitude accuracy, which can be highly valuable for humanitarian aid decisions in the days or weeks after an earthquake until ground truth is available (Wald et al. 2005). The methodology consists of the identification of data sources; attribute mapping; quality assessment and rating; synthesis of data, and where necessary, the assignment of inventory distributions to the countries lacking data from neighboring countries.

The framework described here is not entirely new and is in many ways analogous to inventories developed using techniques from ATC-13 (1985) and HAZUS-MH (FEMA 2006). It is not the first global building stock database; the Russian program Extremum seems to hold that distinction. However, PAGER's inventory database appears to be the first to be open, publically available, transparently developed, and created in part through collaboration with independent earthquake engineering experts from around the world.

The database is limited in several important ways: it focuses on features relevant to fatality risk in earthquakes; it does not take advantage of some available information such as commercial databases or remote sensing data; and it is seriously limited by the paucity of available data, with more than half the world's countries wholly lacking in publically available building data. We are working to overcome some of these limitations, such as through development of sub-country inventories for priority countries where data is readily available and also through a collaborative effort with experts organized through the World Housing Encyclopedia to gather data on building stock distributions, especially for high-vulnerability, high-seismicity countries. The inventory framework has been explicitly developed as a starting model that allows for its constitutive data to be replaced with higher quality and higher resolution data as they become available. We are aware that for many regions, obtaining more detailed inventories will be simply a matter of time and effort; in other areas, given the difficulty of ascertaining data, it will not be tractable in the near future. Hence, the reliance on our default, country-wide datasets may be applicable for certain regions of the world for some time to come. Given the limitations noted above, three central conclusions from this effort are that:

1. The PAGER inventory framework provides a globally consistent approach for treating diverse data with varying uncertainty, quantity, and quality, to produce a building inventory useful for estimating social and economic impacts of global earthquakes.
2. It defines somewhat more clearly the development needs for improving the quality of existing building data, especially in certain hazard-prone countries.

3. It demonstrates the practicality of open data development and data-sharing mechanisms at a global scale. It allows for the continual improvement of global building inventory data, especially for non-residential buildings, through a global collaboration. The data can be used for PAGER and other loss estimation needs, such as societal risk mitigation decision-making. The Global Earthquake Model (GEM), for example, aims to inform such decisions, and GEM's leadership is interested in using the database to that end.

ACKNOWLEDGMENTS

We thank Dr. Mboup Gora of the United Nation's UN-HABITAT for providing demographic and health survey data. The valuable contribution made by participants to WHE-PAGER survey (http://www.pager.world-housing.net/) is gratefully acknowledged. Mike Hearne developed the PAGER inventory editor and together with Paul Earle and Trevor Allen, provided additional advice and assistance.

REFERENCES

Applied Technology Council (ATC), 1985. *Earthquake Damage Evaluation Data for California*, Applied Technology Council, Redwood City, California.

——, (ATC), 1987. *ATC-14, Evaluation of the Seismic Resistance of Existing Buildings*, Redwood City, CA, 370 pp.

——, (ATC), 2002. *FEMA 154: Rapid Visual Screening of Buildings for Potential Seismic Hazards: A Handbook*, Second Edition, Federal Emergency Management Agency, Washington, D.C.

Bommer, J., Spence, R., Erdik, M., Tabuchi, S., Aydinoglu, N., Booth, E., Del Re, D., and Peterken, O., 2002. Development of an earthquake loss model for Turkish catastrophe insurance, *J. Seismol.* **6**, 431–446.

Dolce, M., Kappos, A., Masi, A., Penelis, G., and Vona, M., 2006. Vulnerability assessment and earthquake damage scenarios of the building stock of Potenza (Southern Italy) using Italian and Greek methodologies, *Eng. Struct.* **28**, 357–371.

Faccioli, E., Pessina, V., Calvi, G., and Borzi, B., 1999. A study on damage scenarios for residential buildings in Catania City, *J. Seismol.* **3**, 327–343.

Federal Emergency Management Agency (FEMA), 2006. *HAZUS-MH MR2 Technical Manual*. Washington, D.C., http://www.fema. gov/plan/prevent/hazus/hz_manuals.shtm [22 Aug 2008].

Jaiswal, K., and Wald, D. J., 2008. *Creating a Global Building Inventory for Earthquake Loss Assessment and Risk Management*, USGS Open File Report, OF *2008-1160*, 103 pp., http://pubs.usgs.gov/of/2008/1160/.

Jaiswal K., and Wald, D. J., 2010. Development of a semi-empirical loss model within the USGS Prompt Assessment of Global Earthquake for Response (PAGER) system in *Proceedings of the 9th U.S. and 10th Canadian Conference on Earthquake Engineering: Reaching Beyond Borders*, July 25–29, 2010, Toronto, Canada, Paper No. 1095.

Krimgold, F., 1989. Search and rescue, *Earthquake Spectra* **5** (S1), 136–149.

Kuwata, Y., Takada, S., and Bastami, M., 2005. Building damage and human casualties during the Bam-Iran earthquake, *Asian Journal of Civil Engineering (Building and Housing)* **6**, 1–19.

Madabhushi, S., and Haigh, S., 2005. *The Bhuj India Earthquake of 26th Jan 2001: a Field Report by EEFIT, Earthquake Engineering Field Investigation Team*, Institution of Structural Engineers, London.

Maqsood, S., and Schwarz, J., 2008. Analysis of building damage during the 8 October 2005 earthquake in Pakistan, *Seismological Research Letters* **79**, 163–177.

Mouroux, P., Bertrand, E., Bour, M., Le Brun, B., Depinois, S., and Masure, P., 2004. The European RISK-UE project: an advanced approach to earthquake risk scenarios, *Proc. of the 13th World Conference on Earthquake Engineering*, Vancouver, BC, Canada, paper 3329 (CD-Rom).

Nadimpalli, K., Cornish, L., and Kazemi, S., 2002. National Exposure Information System (NEXIS) at Geoscience Australia, 2006, (http://www.ga.gov.au/image_cache/GA10003.pdf) 1–2.

Ozmen, B., 2000. Isoseismic map, human casualty and building damage statistics of the Izmit earthquake of August 17, 1999, *Third Japan-Turkey Workshop on Earthquake Engineering*, February 21–25, 2000, Istanbul, Turkey.

Petrovski, J., 1983. Engineering measures for earthquake risk reduction in the Arab countries, in *Assessment and Mitigation of Earthquake Risk in the Arab Region* (Cidlinsky K., and B. M. Rouhban, eds.), Prepared by UNESCO for the Arab Fund for Economic and Social Development, 173–218.

Porter, K. A., Jaiswal, K. S., Wald, D. J., Greene, M., and Comartin, C., 2008. WHE-PAGER project: A new initiative in estimating global building inventory and its seismic vulnerability, *Proc. 14th World Conference on Earthquake Engineering*, October 2008, Beijing, China, Paper No. S23–016.

Porter, K., 2009. Cracking an open safe: HAZUS vulnerability functions in terms of structure-independent spectral acceleration, *Earthquake Spectra* **25**, 361–378.

Rodriguez, V. I., Yacante, M. I., and Reiloba, S., 2002. *World Housing Encyclopedia Reports 1, 2, 70, & 89*; Argentina.

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

Tien, Y. M., Juang, D. S., Pai, C. H., Hisao, C. P., and Chen, C. J., 2002. Statistical analyses of relation between mortality and building type in the 1999 Chi-Chi earthquake, *J. Chin. Inst. Eng.* **25**, 577–590.

Tobita, T., Miyajima, M., Fallahi, A., Alaghebandian, R., and Ghayamghamian, M. R., 2007. Seismic intensity estimation through questionnaire survey and collapse rates of various building types in the 2003 Bam, Iran earthquake, *Earthquake Spectra* **23**, 841–865.

United Nations (UN), 1993. *Housing in the World-Graphical Presentation of Statistical Data*, New York.

UN-HABITAT. 2007. Housing in the World-Demographic and Health Survey Database (personal communication).

Wald, D. J., Worden, B. C., Quitoriano, V., and Pankow, K. L., 2005. ShakeMap Manual: Technical Manual, User's Guide, and Software Guide, U. S. Geological Survey, Techniques and Methods 12-A1, 128 pp.

Wald, D. J., Earle, P. S., Allen, T. I., Jaiswal, K. S., Porter, K. A., and Hearne, M., 2008. Development of the U. S. Geological Survey's PAGER system (Prompt Assessment of Global Earthquakes for Response). *Proc. 14th World Conference on Earthquake Engineering*, October 2008, Beijing, China, Paper No. 10-0008.

Yakut, A. 2004. Personal communication.

(Received 25 November 2008; accepted 22 November 2009)