Originally published as:

Pittore, M., Wieland, M. (2013): Towards a rapid probabilistic seismic vulnerability assessment using satellite and ground-based remote sensing. - Natural Hazards, 68, 1, 115-145

DOI: [10.1007/s11069-012-0475-z](https://doi.org/10.1007/s11069-012-0475-z)

Toward a rapid probabilistic seismic vulnerability assessment using satellite and ground-based remote sensing

Massimiliano Pittore · Marc Wieland

Received: date / Accepted: date

Abstract Natural hazards such as earthquakes threaten millions of people all around the world. In a few decades most of these people will live in fast-growing, inter-connected urban environments. Assessing risk will, therefore, be an increasingly difficult task that will require new, multidisciplinary approaches to be tackled properly. We propose a novel approach based on different imaging technologies and a Bayesian information integration scheme to characterize exposure and vulnerability models, which are among the key components of seismic risk assessment.

Keywords earthquake · vulnerability · risk management · urban planning · remote sensing · omnidirectional imaging · bayesian networks

PACS PACS 89.60Gg · PACS 89.65.Lm2 · PACS 89.20.Kk

1 Introduction

Seismic risk is usually defined by three fundamental components, each of which should be separately quantified [1], [2], namely: *seismic hazard occurrence probability* (in the following referred to as *hazard*), *exposure* and *vulnerability*. *Hazard* represents the likelihood of a given location experiencing a certain level of ground-shaking. The *exposure* model identifies the inventory of people, buildings or other assets exposed to the hazard. The *vulnerability* model describes how the exposed assets will be affected by the hazard, for instance quantifying the amount of expected damage to different types of buildings under a given seismic load. In this paper we are focusing on exposure and vulnerability models, considered to be the main sources of spatio-temporal change in seismic risk. Seismic hazard usually exhibits a low rate of change over short time spans (from a few months to a few years,

This work was supported by PROGRESS (Georisiken im Globalen Wandel), Helmholtz-EOS (Earth Observation System), EMCA (Earthquake Model Central Asia) and CEDIM (Center for Disaster Management and Risk Reduction).

M. Pittore, M. Wieland
GFZ German Research Centre for Geosciences
Telegrafenberg, 14473 Potsdam, Germany
E-mail: pittore@gfz-potsdam.de

not considering aftershocks), and is almost independent of anthropic activities [3]. On the contrary, the rapid sprawling of built-up areas, unplanned settlements and the general rapid changes modern cities and mega-cities undergo heavily affect the space- and time-dependency of exposure and vulnerability models. A thorough, up-to-date inventory of all exposed assets (for instance residential buildings) is often unfeasible or very time- and resource-intensive. A reliable model of seismic vulnerability is even more difficult to obtain, involving usually an on-site complete analysis of the structures and materials by experienced engineers. As a result, seismic risk assessment is often based on incomplete, uncertain or strongly aggregate data [4], [5]. A reliable, spatially-defined distribution of potential losses can greatly improve the organization of rapid response actions, but where heterogeneous information is being integrated, a clear understanding of the involved uncertainties should be possible within the model. This paper is ideally the continuation of a former publication [6], where a stratification scheme based on the analysis of multi-temporal medium-resolution satellite images were first introduced. In the same publication an in-situ data collection based on omni-directional imaging were also suggested, but the integration of collected information into a vulnerability model was not addressed. The current contribution thus further extends towards a complete assessment of inventory and seismic vulnerability, by proposing an original information integration scheme based on Bayesian Networks. Moreover, a building spatial density model based on selective building footprint digitization has been introduced in order to complement the building inventory model and to further improve the overall vulnerability assessment. For sake of clarity the main components of the overall scheme are briefly reviewed, In Section 2, a general outline of the proposed approach is provided, while Section 3 briefly introduces the study area. In Section 4, the stratification and sampling scheme, based on the processing of remote sensing data is briefly recalled. In section 5, an *in-situ* data collection strategy, taking advantage of omni-directional imaging is described, along with a manual satellite image interpretation stage. In section 6, a fully probabilistic information integration approach for vulnerability assessment, based on Bayesian Networks, is presented. Finally, in sections 7 and 8, an exemplification of a building inventory and structural seismic vulnerability model assessment is presented and discussed along with a preliminary validation. In the last section, conclusions and future work are outlined.

2 Outline

The proposed approach is based on three main logic modules: *stratification*, *information collection*, and *information integration* (see Figure 1a). The purpose of *stratification* is to subdivide the area of interest into sub-areas which can be considered homogeneous with respect to the type of data to be collected. The more homogeneous a sub-area is, the less intensive will be the data collection required to have a reliable description of the sub-area itself. This module will be briefly outlined and interested reader can refer to the previous publication [6] for further details. Once the area of interest has been divided in to sub-areas, the *information collection* phase can start. In this phase, several strategies can be followed, including direct on-site surveys and the processing or manual interpretation of high-resolution satellite data. The last module, *information integration*, is de-

voted to properly integrate all collected, extracted or computed information into a unified description taking into account, whenever possible, the uncertainties involved. Among different approaches to information integration, we suggest the use of Bayesian networks (or *belief* networks). The flowchart of the proposed approach, depicted in Figure 1b, is simple but effective; once the collection phase has been completed, a check of the resulting homogeneity can lead to a new, improved stratification. Once the information has been integrated with already available data, an evaluation of the resulting model in terms of uncertainty can lead to a new collection phase. Following this approach, assessing exposure and vulnerability models for risk assessment is no longer a punctual process, but rather a dynamical process that can be scaled according to the stakeholders' resources and constrains. Furthermore, the *Update cycle* (see Figure 1b) can be repeated over time as the environmental conditions change or need for less uncertain models arises.

Fig. 1: Outline of the data collection and integration scheme. (a) The logical scheme. (b) The flowchart of the whole process.

3 Study area

As a case study, the capital of Kyrgyzstan, Bishkek, is considered and used throughout the paper for exemplifying the different components of the scheme. Bishkek is located just north of the Kyrgyz Ala-Too mountain-range bordering the Chu basin (see Fig. 2). The city lies in a plain formed by mainly by fluvial deposits of the Ala-Archa and Alamedin rivers. The region of Bishkek in the Chu basin refers to the North Tien-Shan seismic zone, which belongs to one of the most seismically

hazardous areas in Central Asia. The GSHAP seismic hazard map shows a peak ground acceleration (PGA) of 4.5 m/s^2 with a probability of 10% to be exceeded in 50 years for the area [7]. The city itself was founded in 1825 and counted 859'800 inhabitants in 2010 [8]. Since the 1970's the city almost doubled its population (436'459 inhabitants in 1970). It is therefore not only the largest city in Kyrgyzstan at present, but moreover is also rapidly expanding both in terms of area and inhabitants. Rapid urban growth coupled with mainly non-engineered construction practices which show high structural vulnerability [7] lead to an increasingly high level of seismic risk in the area [9]. The actual area under observation within this study covers an area of 665 km^2 and is outlined in Fig. 2.

Fig. 2: Location and extent of the study area Bishkek, Kyrgyzstan. On the left: the GSHAP map (1999), on the right: Landsat TM (2009)

4 Stratification

In order to achieve a reasonable stratification of the considered town, we suggest a multi-stage approach (Figure 3); the spatial extent of the town is first subdivided into small geographical units, obtained by unsupervised low-level segmentations of a time-series of medium-resolution multi-spectral satellite images (see Sec. 4.1). Then the geographical units are labeled and grouped into bigger clusters depending on their predominant landuse/landcover and approximate age using partly supervised classifications and change-detection analysis (see Sec. 4.2). The technical details can be found in [6]. In this section we provide a brief discussion on the approach and a preliminary validation.

Stratification is a well known sampling technique [10], particularly useful if the population from which the data are to be sampled can be clustered into different *strata*, each one exhibiting a relative homogeneity with respect to the population as a whole. Sampling from a suitable stratification ensures an optimal trade-off between time and resource-effectiveness of the sampling, and allows for controlling the uncertainty of the estimates. Since we are interested in describing the exposed inventory and its seismic vulnerability, the population we focus our attention on is represented by the inventory of the assets physically exposed to the earthquake threat. Herein we will mostly focus on the residential buildings of the city, hence information about land-use/land-cover will be used to filter out areas mainly characterized by industrial or commercial buildings. The reason is two-fold: on the one hand, residential buildings are strongly related to social vulnerability, and as such are of particular interest, while on the other, assessing the vulnerability of industrial or commercial buildings is very difficult, and is beyond the scope of the present paper. The natural stratification of such buildings population would take into account the building type, according to a specific taxonomy. Unfortunately this is not the most viable approach, since in most cases information about the composition of the exposed inventory is not available, nor the location of the different building types within the considered urban environment. A more useful stratification is defined on a spatial base, based on the hypothesis that the building inventory composition, and therefore the vulnerability model, follow Tobler's law of spatial autocorrelation [11]. This hypothesis proposes that often neighboring buildings share several features related to vulnerability, such as age, typology and materials of construction, occupation, and can therefore be clustered together. While this basic hypothesis is used in many studies dealing with urban structure types [12, 13] the scale at which homogeneity should be addressed is still a matter of debate. It is clear, for instance, that a city block should be considered homogeneous, whilst larger areas are more likely to show an uneven distribution of the features of interest. We decided to focus on a spatial scale ranging from one to several blocks, also considering that other interesting risk-related features such as soil-amplification effects are usually investigated at a comparable scale in urban environments.

4.1 Unsupervised segmentation

The main purpose of segmentation is to provide an object-based, higher level description of a satellite image capturing the urban environment of interest. The

resulting segments need to be considered sufficiently homogeneous to be treated as atomic entities in every subsequent processing stage, while still retaining significant local information in terms of spectral content and image texture. The resulting segmentation therefore reduces the complexity of the image, providing an alternative, representationally efficient description [14]. Several algorithms for image segmentation have been proposed in the literature in the last decade. As proposed in [6] we have opted for the approach proposed by Felzenszwalb and Huttenlocher [15], which is computationally efficient and has the property of capturing perceptually important grouping in the processed image. In order to provide a proper base for the stratification, the segmentation algorithm is applied to three medium-resolution Landsat images of the region of interest, acquired in 1977, 1994 and 2009 (Fig. 3 A). A scale parameter and a merge parameter must then be chosen, which is done on a trial-and-error basis.

Fig. 3: Stratification workflow exemplified for Bishkek, Kyrgyzstan.

4.2 Supervised classification and multi-temporal change-detection

The segmentation stage described in Sec. 4.1 is the input of the main stratification procedure. In this stage each of the computed segments is assigned a label describing its land-use/land-cover and approximate age. For the land-use/land-cover classification we use a supervised approach based on statistical learning inference. Support Vector Machines (SVM) have been used as the reference learning engine [16]. The feature vectors describing the segments to be classified are composed of 26 features, which include the mean and standard deviation of values in the six spectral bands of Landsat, the mean and standard deviation of the Normalized Difference Vegetation Index (NDVI) and two texture descriptors derived from the

Gray-Level Co-occurrence Matrix (GLCM). Training samples for the SVM learning model have been selected based on local expert knowledge, manual image interpretation and GPS-photos. A total of 10 different class labels, listed in Table 1, have been used to provide a comprehensive classification (Fig. 3 B). The classes have been selected following the approach outlined in [7]. A preliminary evaluation of the performance of this processing stage is provided in Table 2. The confusion matrix has been built using 10 randomly selected and manually labeled test samples *per* class. Also, the test samples were labeled based on a mixture of manual image interpretation, GPS-Photos and local expert knowledge.

Table 1: Landuse/landcover (LULC) class labels used in the classification of the extracted segments

class	description
1	1-2 storey masonry, brick (individual apartment houses) - type 1
2	1-2 storey masonry, brick (individual apartment houses) - type 2
3	1-2 storey masonry, brick (individual apartment houses) - type 3
4	3-6 storey masonry, brick, concrete, panel buildings (apartment blocks)
5	7-9 storey concrete panel, frame + monolithic (apartment blocks)
6	industrial, commercial
7	mixed built-up area
8	vegetation
9	water
10	other (rocks, bare soil)

Table 2: Confusion matrix relative to the classification described by Table 1. A total of 10 test samples *per* class have been used.

	class	Reference										TP (%)
		1	2	3	4	5	6	7	8	9	10	
Classification	1	9	0	0	0	0	0	1	0	0	0	90
	2	0	8	0	1	0	0	1	0	0	0	80
	3	0	0	8	0	0	0	2	0	0	0	80
	4	0	0	0	9	1	0	0	0	0	0	90
	5	0	0	0	0	5	1	2	0	0	2	50
	6	0	0	0	0	0	9	0	0	0	1	90
	7	0	0	1	2	0	0	7	0	0	0	70
	8	0	0	0	0	0	0	0	10	0	0	100
	9	0	0	0	0	0	0	0	2	7	1	70
	10	0	0	0	0	0	0	0	1	0	9	90
TN (%)		100	100	88.9	75	83.3	90	53.8	76.9	100	69.2	81

In order to take into account the temporal changes in the built-up environment over the considered time-span, that is adding temporal information to the final stratification, the extent of built-up areas for each time-stamped image has been delineated using a simple two-class (built-up/not built-up) learning engine (Fig. 3 B). An evaluation of the performance of this processing stage in terms of receiver operating characteristic curves (ROC) is provided in Fig. 4. Differences

Fig. 4: ROC for three binary urban extent classifications from Landsat images.

in the performance of the built-up areas' delineation are dependent upon both the different quality of imagery sources, and on the training set used to train the learning engine (availability of reliable ground-truth data for 1977 built-up mask, for instance, is limited with respect to more recent imagery). In 1977 a total area of 113 km^2 was classified as being built-up in the study area. In 1994 the built environment accounted for 145 km^2 , while in 2009 the city reached a total built-up area of 207 km^2 inside the study area.

Considering both the 7 labels referring to the built-up environment (classes 1-7 in Table 1) and the results of a post-classification change-detection analysis of the three time-stamped built-up masks (Fig. 3 C), a total of 21 possible combinations (defining the final stratification) can be delineated. To describe the *strata* also in terms of their spatial extent, the individual segments are merged together according to their combination of attributes. The final stratification for Bishkek is shown in Fig. 3 D. Further explanations and details can be found in [6].

5 Data Collection

Given a stratification of the region of interest, the next logical phase is data collection. For the case of Bishkek, a preliminary sampling has been performed to collect visual information about the buildings inventory and, based on this, to estimate its seismic vulnerability.

Since we aim at illustrating a solution for rapid screening and assessment, in this paper we will only cover visual, indirect assessments. It is nevertheless clear that the proposed approach can (and should, when possible) be complemented by more in-depth surveys [17]. For the sake of exemplification, we considered different types of data collection; on-site rapid image capturing based on omni-directional imaging, and a remote, off-line manual interpretation of satellite imagery (building footprints digitizing). In the latter case, some information was already available from OpenStreetMap which has been integrated into the obtained dataset.

5.1 Building footprints collection

Buildings footprints provide useful information about the considered buildings inventory stock. Each footprint provides the geographical location (and orientation) of the building, its outer shape, and its area. Such information is increasingly available worldwide, even if on a very discontinuous base, thanks to open, collaborative projects such as OpenStreetMap [18]. A complete manual digitization of a large town is often very time and resource-intensive [19]. Unsupervised approaches for automatic extraction of footprints from high-resolution satellite images are available or under study, but often, when applied extensively and in complex urban environments, still do not provide a sufficient level of accuracy, particularly when the total number of buildings should be estimated [20–22]. On the other hand, a focused manual digitization, possibly taking advantage of already available, free of charge data, proves to be a viable and efficient alternative. A total of 8750 building footprints have been collected within a *sampling frame* composed of 21 sample areas (one sample area for each stratum, as depicted in Fig. 5, of which 8255 buildings were manually digitized during a four-week project, and 455 were imported from OpenStreetMap. The collected footprints' area amounts to about 1.5 km^2 , equal roughly 0.75% of the total 2009 built-up area (estimated to be 207 km^2) in Bishkek. Details on the selection and sizing of the sample areas can be found in [6].

Another 13484 building footprints have been collected outside the above-described sampling frame, of which 3305 were imported from OpenStreetMap. These footprints are not used in the subsequent density estimation, because they do not provide a comprehensive coverage of the areas, but they are used in the inventory and vulnerability composition assessment described in Sec. 6. All collected buildings footprints are stored in a PostgreSQL/PostGIS database, where they can easily be accessed for processing or GIS-based visualization.

An exploratory analysis of the collected footprints in terms of the area distribution among the different strata already shows that stratification is indeed successful in separating different populations of buildings. As shown in Fig. 6, buildings belonging to segments labeled with classes 1,2,3 actually exhibit very similar footprint area distributions. We further observe that buildings belonging to segments

Fig. 5: Bishkek: manually digitized footprints, along with the stratification. Respectively in blue and red the considered sample and validation areas are shown.

labeled by class 4, 5, 6 and 7 exhibit clearly separate area distributions. The mean values of area are compatible with the type of buildings expected in the respective segments. For instance classes 1,2 and 3 refer to smaller buildings, mostly belonging to single-family compounds often composed by several structures, while class 4 refers to bigger multi-family houses. In segments classified as 6 and 7, labeled ‘industrial/commercial’ and ‘mixed built-up’, the footprint areas are more widely spread. Thus, these segments are less homogeneous in terms of inventory composition. In order to further assess the stratification based on the collected footprints, the distribution of areas with respect to estimated age is shown in Fig. 7. It is possible to note that buildings belonging to segments classified as 1,2,3 (1-2 storey houses) and 4 (3-6 storey buildings) do not show a significant change in area distribution with built-up age. In contrast, buildings belonging to segments classified as 5 (7-9 storey buildings) exhibit a decrease in their mean footprint area over the 1994-2009 period. This is likely to depend on the recent increase in the number of monolithic high-rise buildings, whose footprints are usually more compact. Also, the segments classified as 7 (mixed built-up) show a clear decreasing trend. This could be partially due to the reduced frequency of buildings with bigger footprints included in the most recent “mixed built-up” segments and hence are classified as outliers, although this should be confirmed by further analysis. The latter preliminary evaluations confirm that the proposed stratification correctly subdivided the urban areas into clusters that are homogeneous with respect to the predominant land-use/land-cover.

Fig. 6: Distribution of area of digitized building footprints in the different strata, grouped by landuse/landcover. The LULC classes are clearly separated also in terms of area distribution, therefore indirectly confirming the quality of stratification.

Fig. 7: Distribution of area of digitized building footprints in the different strata, grouped by age.

5.2 Building density estimation

Buildings' footprints are a valuable resource for assessing the composition of the buildings' inventory, as they provide information on the shape, area and location of a sample of buildings. This information can be used in turn to infer useful characteristics of the buildings inventory such as building density. The latter is important, for example in order to provide quantitative loss estimates, which is of particular interest for decision makers or insurance companies. However, information about the number of residential or commercial buildings is seldom available, and often uncertainties are not explicitly considered. In this section we describe a simple approach for estimating building density on a *per-stratum* base, in order to obtain the expected number of buildings for each spatial segment.

The development of a realistic model of the spatial distribution of buildings in an urban environment can prove to be a challenging task, given the high number of interacting phenomena which are contributing to it. For the sake of simplicity, we suppose that the spatial placing of buildings can be modeled as a stationary Poisson point process [23]. A Poisson process is a stochastic process which counts the number of events inside each of a number of non-overlapping finite subregions of some vector space V (in our case $\mathbb{R}^2$), assuming that each count represents an independent random variable following the Poisson distribution. More specifically, let S be a subset of $\mathbb{R}^2$, and A a family of subsets of S . If $\{N(A)\}_{A,\lambda}$ is a homogeneous Poisson point process with intensity λ , then the following holds:

$$P(N(A) = k) = \frac{e^{-\lambda|A|}(\lambda|A|)^k}{k!} \quad k = 0, 1, \dots, n \quad (1)$$

Where $N(A)$ is the number of events in A , $P(N(A) = k)$ is the probability for this number to be exactly equal to k , and $|A|$ is a measure of the area of A . This means that the probability of the number of events $N(A)$ depends on the set A only through its size $|A|$.

We furthermore hypothesize that the stochastic process is stationary within each considered segment. This is a strong simplification, since clustering and inhibition of buildings placement happen naturally in urban environments, following the topology of road networks, the availability of services, etc. Missing further detailed information, stationarity is a common-sense choice, but we decided to apply a boot-strap approach to the Poisson intensity estimation in order to encode the missing information in terms of uncertainty. For each sample area, $N_{ssa} = 100$ possibly overlapping sub-areas have been randomly selected. For each selected sub-sample area, a stationary Poisson point process has been fitted. The mean value of the fitted intensities is considered as the intensity reference value for the sample area (and thus for the relative stratum), and the standard deviation σ_{sa} of the fitted intensities describes the estimated uncertainty through a standard 95% confidence interval *c.i.* = $\pm \frac{1.96 * \sigma_{sa}}{\sqrt{N_{ssa}}}$.

A preliminary validation of the proposed approach has been conducted. A set of 10 validation areas (six of which refer to *stratum* type 2, two to *stratum* 10 and two to *stratum* 12), where complete building footprints delineation is available, have been considered. For each validation area, a confidence bound on the expected number of buildings has been computed, based on the intensity and uncertainty of the Poisson process estimated for the underlying stratum. The confidence bound is compared in Fig. 8 with the actual number of delineated footprints in each

validation area. Despite the preliminary nature of the assessment, a reasonable agreement is found between the estimated and measured values. We note, though, the greater variability of the areas within stratum 2. At a first glance, this appears to be related to the less ordered arrangement of buildings in this stratum, where empty spaces alternate with high-density patches. On the contrary, the other two strata show a more regular pattern, therefore determining a better performance of the fitted model.

Fig. 8: Comparison between expected number of buildings (represented by confidence bounds) and the counted building footprints for the considered test areas (depicted in Fig. 5).

The number of buildings expected for each stratum, described in term of its dominant land-use/land-cover (LULC) is shown in Fig. 9. The total number of buildings resulting from this estimate for the study area in Bishkek is $N_b = 182000 \pm 14000$.

Fig. 9: Expected number of buildings, by land-use/land-cover, with standard errors for Bishkek.

5.3 Omnidirectional Mobile Mapping

Mobile mapping systems are increasingly used for collecting geo-referenced, multi-sensorial data for a variety of purposes, ranging from urban planning to road maintenance to tourism applications [24]. Usually a mobile mapping system is composed of one or more image capturing devices and a Global Positioning System (GPS) fixed on a car and driven around. Other sensors like laser scanners, inertial measurements units (IMUs) and microphones can be added, depending on the specific application. In this context, omni-directional imaging [25,26], recently brought to the public attention by Google StreetViewTM, is a very efficient way to collect and store geo-referenced visual information. It involves very simple deployment and operation, with no need for skilled operators, nor for sophisticated aiming devices, therefore shifting the focus of the applications from the collection to the (usually remote and off-line) interpretation of the collected data. During two weeks of field-work in Bishkek, approximately 1.000.000 images (equal to 3 TB of data) were acquired. The field-data-collection has been guided along pre-calculated routes using GPS-based navigation system. A total area of approximately 30 km^2 and an overall street-length of 170 km has been covered by the survey. For better management, all meta-data describing the collected images, their geographical locations, their azimuths and speeds are stored in a PostgreSQL/PostGIS database. The database allows for several spatial reasoning applications involving the collected images. For instance, since building footprints are available, a visual sample of the buildings referred to by the collected footprints is straightforward. In the next section, preliminary automatic processing of the acquired omni-directional images is presented, aiming at inferring 3D information about the imaged buildings by exploiting the particular omni-directional image geometry. The proposed approach, already outlined in [6], is briefly recalled in the following section for the sake of clarity. Some new results are provided along with a preliminary validation.

5.4 Building height from omni-directional images

A preliminary assessment of the capability of the automatic processing is described in this section. The goal is to select a set of processing routines able to extract information from the image database, which could be used both in a subsequent unsupervised analysis of inventory composition and vulnerability, and as an augmentation of the experience for a user manually interpreting the images. For example, an engineer analyzing the images for rapid vulnerability screening could find it useful to have displayed ancillary information extracted off-line, such as the building height, the estimated number of floors, the number or size of visible openings, etc. in case such information would not be already available from other sources. We focused on the building height extraction, also because it is a parameter instrumental to infer the building type, which is directly affecting the vulnerability and is often difficult to estimate from remotely-sensed imagery (without resorting to expensive and not always available LIDAR or high-resolution stereo data). The approach exploits optical flow computation to obtain a dense set of visual correspondences which in turn are used to infer geometrical translation and rotation between the images. From each pair of correspondences a 3D point is computed using a simple triangulation technique [27]. Metric calibration

is then applied based on GPS data embedded in the captured images. An example of the resulting 3D reconstruction is shown in Fig. 10. Let us note that such a reconstruction is covering a very extended field of view, since the two buildings depicted in the Figure are each nine storeys high, and not more than 25 meters away from the observer location. Covering such a field of view with a conventional camera would imply capturing several images with different aims.

Fig. 10: Example, for Bishkek, of a 3D reconstruction based on the processing of omni-directional image stream. The map in the lower left corner shows the locations of the optical center of the camera when the images were shot. In orange, several digitized footprints are also shown. In the upper part one of the images of the processed sequence is displayed. On the right, the 3D reconstruction based on omni-directional image processing is displayed as a point cloud. All coordinates are in meters. It is possible to recognize the outlines of the two buildings at the sides of the street.

The above described processing has been applied to an extended sample of collected images, namely a 3D reconstruction, employing 200 images over the entire set of sequences acquired in Bishkek, Kyrgyzstan. All computed 3D measurements have been used to assign a height value to a two-dimensional sparse grid. Grid cells are 10 meters wide squares, whose origin resides for the sake of simplicity in the first processed location. The size of the grid is coarse enough to partly filter out the 3D noise of reconstruction (see Fig. 11). In Fig. 12 the distribution of the estimated building heights is provided and subdivided according to different land-use/land-cover labels. The frequency distribution of estimated building heights are in general agreement with the land-use/land-cover labels. We observe, though, that the histograms related to LULC: *3-6 storeys masonry brick/concrete/panel* and LULC: *7-9 storeys concrete panel frames/monolithic* are unexpectedly similar. This could be due to a partial misclassification of some segments, or to a bias of the height estimation due to occluding trees (generating *ghost* measurements).

Fig. 11: Results of extracting 3D measurements from omni-directional image processing. Coloured cells represent elements of the sparse measurements grid. Four different sub-areas of Bishkek are shown.

Fig. 12: Distribution of estimated building height in spatial segments classified with different land-use/land-cover labels. The upper histograms are in accordance with the expected height values, while the bottom histograms show a larger spread of the values.

An intensive validation of the proposed approach for height estimation will be conducted as soon as a suitable ground-truth dataset is available. As a preliminary assessment, we provide in the following a few examples of the results so far. Four cases have been selected for which a ground-truth height is available, and are depicted in Fig. 13. In order to discuss the usability and possible sources of uncertainty, two of the cases (cases a,b of Fig. 13) refer to optimal conditions, and two (cases c,d of Fig. 13) refer to sub-optimal conditions. Optimality herein is defined by a few basic requirements: the building should not be too far from the camera (10 to 40 m, also depending on its height), there should not be major visual occlusions (vegetation, poles, roadsigns) and the camera should not be subject to strong erratic accelerations. In our experience within the town of Bishkek, such simple requirements are very often not complied with: many buildings are surrounded by tall trees, for instance, and small ones can be occluded by high fences and walls. Furthermore, the uneven surface of several streets (some of which are not sealed) exposes the acquisition system to sudden accelerations due to the rolling motion of the car. Where the basic requirements are complied with, satisfactory results are to be expected. A further source of uncertainty is the GPS signal. Since the metric calibration is based on the measurement of the displacement of the camera across three consecutive frames, an error in the position would affect the 3D measurements. This uncertainty is currently deemed of minor importance and is not accounted for.

Fig. 13: Examples of building height estimation, compared with available ground truth information. The upper row (cases *a*, *b*) shows results obtained under optimal conditions, the lower row (cases *c*, *d*) shows the degraded quality of the assessment where non-optimal conditions are encountered.

6 Integration

6.1 Bayes networks

Bayesian networks (BN) were developed in the early 1980s to facilitate the tasks of inference in artificial intelligence (AI) systems. In these tasks, it is necessary to find a coherent interpretation of incoming observations that is consistent with both the observations and the prior information at hand [28],[29]. Bayesian networks have been recently proposed, within the context of earthquake engineering, for the spatial modeling of seismic vulnerability [30] for insurance pricing applications or to tackle the broader problem of managing uncertainties in seismic risk [31]. Throughout this section we will refer to Bayesian networks (also known as *belief* networks) as directed acyclic graphs (DAG) composed of a set of vertexes representing discrete random variables $X_1, \dots, X_n$. Conditional independence relationships among the random variables are visually encoded by arcs connecting vertexes (see Fig. 14). Usually, the joint probability of the considered set of vari-

Fig. 14: Section of a directed acyclic graph (DAG) showing a random variable X_i and its parents PA_i .

ables can be expressed as:

$$p(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \prod_i p(X_i = x_i | X_1 = x_1, \dots, X_{i-1} = x_{i-1}) \quad (2)$$

The most relevant feature of Bayesian networks is that the joint probability of the whole set of variables can be defined as:

$$p(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \prod_{i \in PA_i} p(X_i = x_i | PA_i) \quad (3)$$

where PA_i refers to the parents of the vertex describing the random variable X_i , and $P(X_i = x_i | PA_i)$ indicates the conditional probability of variable X_i assuming value x_i *given* the particular state of its direct ancestors PA_i . Therefore, Bayesian networks allow to drastically reduce the complexity of representation and computation of joint probabilities by exploiting the conditional independence relationships

among the involved variables. Those relationships can be stated *a priori*, on the basis of established knowledge, or can be statistically inferred from experimental data (or both, see for instance [32]). As suggested by Eq. 3, the complexity of the resulting joint distribution is driven by the node with the highest number of ancestors. The local conditional probability distributions encode the *prior* knowledge about the modeled system. The interesting property of Bayesian network is that, whenever new information is provided, therefore forcing one or more nodes representing the random variables into a state (hard evidence) or into a distribution (soft evidence), the state of all other nodes can be computed and a new set of *posterior* distributions obtained. This process, also called *belief updating*, conforms to the well known Bayes Rule:

$$p(X_i = x_i | X_e = x_e) = \frac{p(X_i = x_i, X_e = x_e)}{p(X_e = x_e)} \quad (4)$$

where X_e is the random variable which forcibly assumes the evidence value x_e and is tied to variable X_i by a causal interrelation. Since belief updating is a NP-complete problem, several sampling-based approaches have been proposed for solving it in the general case [33]. We used in this work an open source C++ implementation of Bayesian networks included in the DLib library. Building simple Bayesian networks as those described in the next section is usually performed through the following steps:

- Select the nodes considering the random variables of interest. For the sake of simplicity, each variable is described by a finite discrete number of states.
- Connect the nodes with directed arcs encoding the causal interrelation among the considered variables.
- Define the conditional probability tables of the nodes with at least a parent, and the prior distributions for the others.

6.2 Exposure and vulnerability probabilistic estimation

In this subsection a Bayesian network is used to estimate, for each of the buildings whose footprint has been collected (see Sec. 5), the most likely building type and its most likely structural vulnerability class according to EMS-98 [34]. According to the EMS-98 classification, vulnerability is described by 6 mutually exclusive states (A, B, C, D, E, F) with A referring to the highest structural vulnerability and F the lowest one. This inference process actually integrates available relevant information, regardless of its source, propagating the uncertainties encoded by the probability distributions.

The proposed BN, sketched in Fig. 15, is composed of the nodes summarized in Table 3. A more detailed description of the nodes is provided in Tables 4 to 7b. For each node without parents, the prior distribution is also provided. The prior for nodes A (age) and S (LULC) are based on the considered stratification. The prior for H (height) is a flat distribution, because we do not want to make any particular hypothesis on the distribution. The conditional probability tables for the nodes with parents are also provided. Since most of the conditional probability tables have been compiled manually, the structure of the Bayesian network and the number of states for each node have been kept intentionally small. The conditional

probability table describing nodes F (no. of storeys), T (type) and V (vulnerability) have been inferred for Bishkek from the World Housing Encyclopedia (WHE) reports [35]. These reports, compiled by local engineers, cover many predominant building types of several countries along with an evaluation of their structural vulnerability. In particular for Bishkek, a total of 8 reports have been used. An example of a conditional probability table constructed for node V (Vulnerability) is given in Table 8.

Its important to stress the fact that the only information used to infer the building vulnerability is its type according to World Housing Encyclopedia (WHE). Moreover, the only structural parameter that relates to the building type is the number of storeys, the others being the land-use/land-cover and the approximate age of the *stratum* where the building is laying. This representation of the conditional relationships of building type and seismic vulnerability on structural parameters is overly simplified, but the aim of this paper is to propose a method to efficiently and automatically integrate already available information rather than re-assessing the vulnerability of the considered inventory stock. In this case, all information about seismic vulnerability is extracted from the WHE reports.

Fig. 15: Bayesian network encoding the dependence of a buildings vulnerability on the state of the node encoding the building type and its ancestors.

Table 3: Bayes network for vulnerability assessment. Summary of nodes.

node	description	no. of states	parents	table
V	vulnerability	6	T	4
T	building type	8	A,F	5
A	stratum-age	3	-	6a
F	no. of storeys	7	S,H	7b
S	stratum-LULC	3	-	6b
H	building height	7	S	7a

Table 4: Node relative to building seismic vulnerability.

state	vulnerability (EMS-98)
1	<i>A</i>
2	<i>B</i>
3	<i>C</i>
4	<i>D</i>
5	<i>E</i>
6	<i>F</i>

Table 5: Node: “building type”.

state	building type	WHE report
1	houses with mud walls and thatch roofs	42
2	two-story unreinforced brick masonry building with wooden floor	41
3	buildings with hollow clay tile Load-bearing walls	34
4	reinforced concrete frame buildings Without beams (serie KUB)	39
5	single family brick masonry house	36
6	pre-cast reinforced concrete frame building with cruciform and linear-beam elements	33
7	prefabricated concrete panel buildings with monolithic panel joints	38
8	buildings with cast in-situ load-bearing reinforced concrete walls	40

Table 6: Nodes relative to stratification.

(a) Node: “age”

state	age	prior
1	until 1977	0.4095
2	between 1977 and 1994	0.2889
3	between 1994 and 2009	0.3016

(b) Node: “stratum”

state	stratum	prior
1	1-2 stories, masonry, brick	0.829
2	3-6 stories, masonry, reinforced concrete, panel	0.143
3	7-9 stories, reinforced concrete, panel, monolithic	0.028

Table 7: Nodes relative to building height.

(a) Node: “height”			(b) Node: “num. storeys”	
state	height (meters)	prior	state	num. storeys
1	0-5	0.14286	1	1-2
2	5-10	0.14286	2	3-4
3	10-15	0.14286	3	5-8
4	15-20	0.14286	4	9
5	20-25	0.14286	5	10-11
6	25-30	0.14286	6	12
7	30-50	0.14286	7	13-18

Table 8: Conditional probability table, node: “building vulnerability”.

type	P(A)	P(B)	P(C)	P(D)	P(E)	P(F)
0	0.9334	0.04757	0.01903	0	0	0
1	0.175	0.56706	0.19842	0.01984	0.01984	0.01984
2	0.175	0.56706	0.19842	0.01984	0.01984	0.01984
3	0.175	0.56706	0.19842	0.01984	0.01984	0.01984
4	0.0181	0.215	0.427	0.2964	0.0435	0
5	0.0181	0.215	0.427	0.2964	0.0435	0
6	0	0.0303	0.06061	0.22	0.43911	0.24998
7	0	0.0303	0.06061	0.22	0.43911	0.24998

In order to provide an example of the proposed integration approach, as well as a preliminary validation on real-world cases, the posterior distributions relative to the buildings depicted in Fig. 13 (see Section 5.4) are provided in Fig. 16. The values in light-green refer to the “measured” data, and as such are considered *evidences*, with a 100% belief (belief quantifies the certainty of the collected data). Had information on the uncertainty of the measurements been available beforehand, the probability distribution of the evidence could be modified to take it into account. The ground-truth values for the considered buildings, are: height=30m, storeys=9, EMS-98 vulnerability=E (detailed information on the WHE building type is not available). As we can observe from Fig. 16, when the height is correctly estimated, both the number of storeys and the vulnerability are correctly inferred. Small variations of measured height, as for instance between cases *a* and *b*, still allow for a correct final estimation. As expected, when the height measure is wrong, as in case *c* for instance, the system is not able to provide sound results. Building 179 has not been considered because it would have produced the same output as building 269 (see Fig. 13, Section 5.4).

Fig. 16: Posterior probability distributions computed for the buildings depicted in Fig. 13. In light green are displayed the values inserted as evidence; the strongly peaked probabilities indicate a strong belief in the values.

7 Exposure inventory

Using the techniques described in the previous section, a preliminary assessment of the inventory stock of Bishkek has been computed. For each digitized building either inside or outside the sampling areas mentioned in Sec. 5, the BN described in Sec. 6 has been initialized with the available data, namely the state of the nodes *stratum-age*, *stratum-LULC* and, if available, the state of node *building height* (in case it is not available, its prior distribution is automatically used). The recomputed posterior probability distributions of node *WHE building-type* are then averaged for each building, based on the underlying stratum, and describe the relative frequencies of the building types listed in Table 5. It is important to note that the average probability distributions represent both the available information and the current uncertainties. As a matter of fact, for instance, a flat probability distribution ($p(state_i) = \frac{1}{n_{states}}$) could represent either complete ignorance of the real inventory composition, or a full inventory description where all building types are equally represented. In Fig. 17 the resulting distributions for all strata not classified as “commercial, industrial” is provided. A meaningful

Fig. 17: Probability distribution of the considered building types according to the WHE taxonomy, for each stratum not classified as “commercial, industrial”.

aggregation of the building types can be useful considering that, based on the WHE reports, some of the building types can be grouped together based on their expected vulnerability in terms of EMS-98 classification. Namely, type 1 should mostly refer to vulnerability *A*, types 2-4 should refer to vulnerability *B* (also extending from *A* to *C*), types 5-6 refer to vulnerability *C* (ranging from *B* to *D*) and types 7-8 are usually associated with vulnerability *E* (ranging from *D* to *F*). The average distributions depicted in Fig. 17 could thus be aggregated following these trends, and the result is shown in Fig. 18. In order to better analyze the inventory composition and distribution, the aggregated probability distributions

Fig. 18: Probability distributions of the aggregated building types, averaged based on three aggregated land-use classes.

are also re-aggregated based on three main land-use attributes of strata, namely classes 1-3, class 4 and class 5, as described in Table 1. This classification takes into account a coarser, but intuitive classification of the Bishkek inventory. Examining Fig. 18 it is possible to note that buildings of type 1 and 2 are more likely to be found in areas classified as low-rise, whereas on the other hand, areas with high-rise buildings also show a relative abundance of buildings of type 7 and 8. The strata with medium-rise, mixed buildings also reflect this heterogeneity in the flatness of the probability distribution. The estimates computed in Sec. 5.2 of the density, and therefore the number of buildings in the considered strata can now be applied to give to the inventory model a more quantitative assessment. In Table 9 a summary of the estimated number of aggregated building types is provided, considering the aggregated land-use classification as a spatial reference. The uncertainty in the building numbers is significant, and depends on the size of the sample used for the density estimation, and on the simplistic model used in the modeling of the density distribution itself. In Fig. 19, the spatial distribution

Table 9: Estimated number of buildings by WHE type, with uncertainty, relative to three examples of land-use classes.

aggregated land-use	type 1 $\pm\sigma$	type 2-4 $\pm\sigma$	type 5-6 $\pm\sigma$	type 7-8 $\pm\sigma$	total $\pm\sigma$
1-2 storey masonry, brick	26130 $\pm$ 9645	68773 $\pm$ 25385	37222 $\pm$ 13739	26261 $\pm$ 9693	158386 $\pm$ 58464
3-6 storey masonry, brick, concrete, panel buildings	407 $\pm$ 132	1816 $\pm$ 591	1316 $\pm$ 429	1450 $\pm$ 472	4989 $\pm$ 1625
7-9 storey concrete panel, frame + monolithic	65 $\pm$ 27	413 $\pm$ 174	228 $\pm$ 96	471 $\pm$ 198	1177 $\pm$ 496

of the areas with aggregated land-use attributes considered in Fig. 17 and Table 9 is shown.

Fig. 19: Spatial distributions of areas with aggregated land-use attributes considered in Fig. 17 and Table 9.

8 Vulnerability assessment

A preliminary estimation of the vulnerability model has been performed, following the same procedure described in the Sect. 7. For each stratum the posterior probability distributions of node *vulnerability*, described in Table 4, has been computed by averaging over each considered building footprint. In Fig. 20 an overview of the probability distribution of the EMS-98 vulnerability states is provided for each stratum not classified as “commercial, industrial”.

Fig. 20: Probability distribution of the vulnerability in EMS-98 scale for each stratum not classified as “commercial, industrial”.

Considering the aggregated inventory composition described in Sec. 7 and displayed in Fig. 18, we can compute the relative vulnerability probability distribution. Fig. 21 shows the resulting distributions. We observe that these distributions loosely follow the trend highlighted in Fig. 18. For instance, as expected, the vulnerability in the strata with land-use attribute 1 (“1-2 storeys brick, masonry”) is higher than in the other strata, and more dominated by vulnerability classes *A* and *B*. Likewise, the average vulnerability in strata with land-use 3 (“7-9 storeys concrete panel, frame + monolithic”) is lower and with higher relative frequency of low-vulnerability classes such as *D*, *E* and *F*.

As an alternative and synthetic scalar indicator for vulnerability, we can also define the *mean vulnerability index* (MVI) indicator as:

$$MVI = \frac{1}{(n-1)} \left(\sum_{i=0 \dots n-1} p(V_i)(n-i) - 1 \right) \quad (5)$$

Fig. 21: Probability distribution of vulnerability in the EMS-98 scale for three aggregated land-use attributes (see also Fig. 18).

where n is the number of possible vulnerability states (in our case $n = 6$), and $\{V_{i=0\dots5} = A, B, C, D, E, F\}$ represent the vulnerability states, expressed in the EMS-98 scale. Vulnerability states are mutually exclusive (each exposed asset is described by only one vulnerability state) and are given a probability of occurrence $p(V_i)$. The MVI index varies linearly from 0 (indicating the lowest degree of structural vulnerability) to 1 (indicating the highest degree of structural vulnerability) and provides an intuitive and easy way to describe the structural seismic vulnerability of buildings. We can take advantage of the scalar nature of MVI to depict the spatial distribution of the estimated vulnerability for Bishkek, as shown in Fig. 22. It is interesting to note that most of the newest urban areas, located for instance in the south-western corner of Bishkek where the town is expanding, show a higher structural vulnerability relative to the central, older part of the city. This of course is a preliminary observation that needs to be validated and possibly complemented by further analysis.

Fig. 22: Spatial distribution of mean vulnerability index (MVI) in Bishkek, Kyrgyzstan. A 2009 Landsat TM image is displayed in the background.

8.1 Preliminary validation

In order to provide a preliminary validation of the obtained vulnerability model in Bishkek, several buildings whose seismic vulnerability has been assessed through a detailed inspection by local engineers have been considered. This set of analyzed buildings represents a particular sampling of the distribution of seismic vulnerability in Bishkek, result of a previous systematic field activity. The nature of the sampling and the heterogeneous spatial distribution of the assessed buildings is not optimal for a thorough validation but can provide some useful preliminary information. For each building the corresponding mvi has been computed. Figure 23 shows the spatial distribution of the (around 900) sampled buildings superimposed to the extracted *strata*. Both the *strata* and the sampled buildings are colored according to their mvi values. In the figure it is possible to note as in *strata* where a lower vulnerability value is expected, the sampled buildings exhibit lower mvi values.

Fig. 23: Spatial distribution of sampled vulnerability and comparison with inferred values.

Figure 24, A), B) shows the distribution of the seismic vulnerability of sampled buildings, expressed in EMS-98 scale. The two *strata* shown in Fig. 23 are considered. Even though the empirical distributions differ from those in Fig. 20, the model correctly predicted a higher frequency of low-vulnerability buildings for *stratum* 10 with respect to *stratum* 2, where higher vulnerability buildings are present. Figure 24, C) shows the difference in mvi between the sampled buildings and the (average) value associated with the underlying *stratum*. It is possible to remark that the computed vulnerability model apparently underestimates the vulnerability with respect to the considered sample. Nevertheless, many of the sampled buildings are in accordance with the expected mvi values associated to the

stratification. Since the considered sample is covering only a subset of the stratification, only a preliminary and qualitative validation can be performed.

Fig. 24: A),B) Empirical distributions of seismic vulnerability of sampled buildings in the two *strata* considered in Fig. 23. The distributions should be compared with those in Fig. 20. C) difference between average vulnerability of strata and vulnerability of sampled buildings, expressed as *mvi*.

9 Conclusions and future work

In this paper we described a novel approach to rapid inventory and vulnerability model assessment based on the combined use of remotely sensed and ground based imaging to select optimal sampling and collection strategies, and the use of a fully probabilistic Bayesian networks-based information integration. The current contribution extends and completes a previous work, described in [6]. We showed how, based on a sound stratification, heterogeneous information can be used to quickly infer a model of exposition and vulnerability for a urban environment. A limited amount of manually digitized buildings footprints, for instance, is used to estimate local spatial density and total number of buildings. Omni-directional imaging, suggested as an in-situ visual sampling tool, has proven to be easy to deploy and to operate. In this paper we started exploring the use of automatic off-line processing of the acquired omni-directional images to infer seismic vulnerability-related information which could later enrich the operators visual experience. The obtained results show that a completely automated off-line extraction of relevant parameters is possible, but its inherently complex nature requires more structured and sophisticated techniques in order to be fully effective. Finally, we showed that a flexible information integration process based on Bayesian networks allows for a sound treatment of uncertainties and for a seamless merging of very different data sources, including legacy data, expert judgment, automatic data-mining based inferences, etc. The proposed approach can be easily scaled based on the target location features, and on the specific needs or constraints of the users (civil protection agencies or decision-makers, for instance). Moreover, our goal is to show that seismic risk assessment ought to be considered as a continuous process rather than a punctual assessment, in order to take into account the temporal dynamics of the exposed assets. From this perspective, a comprehensive management of uncertainties is instrumental for the iterative sampling/modeling of the inventory stock and its seismic vulnerability, leading therefore to a more thorough and robust risk assessment.

Results from the proposed processing stages have been applied to the city of Bishkek, the capital of Kyrgyzstan. A first plausibility check has been performed by analyzing the vulnerability composition and distribution of a set of buildings previously assessed by local engineers. The preliminary assessments provided new, interesting insights about the seismic vulnerability distribution and composition of Bishkek, which will be validated and completed in the near future with the help of local experts. The promising results obtained so far open up a wide range of possible future research activities. From the remote-sensing point of view, we will focus on:

- increasing the temporal resolution of time-change analysis for better built-up and age estimation;
- exploring the automated analysis of high-resolution satellite images as a complementary information source;
- further exploring the information collection, acquisition and use of manually digitized footprints and improving its statistical characterization.

Regarding omni-directional imaging:

- the manual interpretation of omni-directional images by experts will be extensively tested;

- height computation from omni-directional images will be improved upon and complemented by pattern recognition tools.

Finally, the use of Bayes networks will be further explored. In the considered application, the Bayes networks have been used to encode the conditional relationships between some properties of the buildings and their seismic vulnerability, as implicitly described in the World Housing Encyclopedia reports. In the future more complex network topologies will be considered, including unsupervised learning of conditional probability tables from available and collected data. We should also remark that all described activities have been carried out using free and open-source software tools and libraries and, where possible, freely available data sources.

Acknowledgements The authors wish to thank the unknown reviewers whose constructive comments and suggestions considerably improved the manuscript. The authors are also grateful to Dr. Kevin Fleming for stimulating many inspiring discussions. Particular thanks to CAIAG (Central Asian Institute for Applied Geoscience, Bishkek) and IntUIT (International University for Innovation Technologies, Bishkek) for their support during the ground-based data-capturing.

References

1. A. W. Coburn, R. Spence, and A. Pomonis, *Vulnerability and Risk Assessment*. UNDP Disaster Management Training Programme, 1994.
2. G. M. Calvi, R. Pinho, G. Magenes, J. J. Bommer, L. F. Restrepo-Velez, and H. Crowley, "Development of seismic vulnerability assessment methodologies over the past 30 years," *ISET journal of Earthquake Technology*, vol. 43, no. 472, pp. 75–104, 2006.
3. A. J. Michael, "Random variability explains apparent global clustering of large earthquakes," *Geophysical Research Letters*, vol. 38, Nov. 2011.
4. S. Steimen, D. Fh, D. Giardini, M. Bertogg, and S. Tschudi, "Reliability of building inventories in seismic prone regions," *Bulletin of Earthquake Engineering*, vol. 2, no. 3, pp. 361–388, 2004.
5. P. Grossi, P. Kleindorfer, and H. Kunreuther, "The impact of uncertainty in managing seismic risk: the case of earthquake frequency and structural vulnerability," *Wharton Financial Institution Center*, vol. 99, no. 23, 1999.
6. M. Wieland, M. Pittore, S. Parolai, J. Zschau, B. Moldobekov, and U. Begaliev, "Estimating building inventory for rapid seismic vulnerability assessment: Towards an integrated approach based on multi-source imaging," *Soil Dynamics and Earthquake Engineering*, 2012.
7. M. Erdik, T. Rashidov, E. Safak, and A. Turdukulov, "Assessment of seismic risk in tashkent, uzbekistan and bishkek, kyrgyz republic," *Soil Dynamics and Earthquake Engineering*, vol. 25, pp. 473–486, Aug. 2005.
8. "Kyrgyzstan, brief statistical handbook 2007-2010," 2011.
9. D. Bindi, M. Mayfield, S. Parolai, S. Tyagunov, U. Begaliev, K. Abdrakhmatov, B. Moldobekov, and J. Zschau, "Towards an improved seismic risk scenario for bishkek, kyrgyz republic," *Soil Dynamics and Earthquake Engineering*, vol. 31, pp. 521–525, Mar. 2011.
10. W. Cochran, *Sampling techniques*. New York: Wiley, 3rd ed., 1977.
11. W. R. Tobler, "A computer movie simulating urban growth in the detroit region," *Economic Geography*, vol. 46, no. 2, pp. 234–240, 1970.
12. A. Rapoport, "Neighborhood heterogeneity or homogeneity. the field of man-environment studies," vol. 1, pp. 65–77, 1980.
13. S. Steiniger, T. Lange, D. Burghardt, and R. Weibel, "An approach for the classification of urban building structures based on discriminant analysis techniques," *Transactions in GIS*, vol. 12, pp. 31–59, Feb. 2008.
14. J. Shi and J. Malik, "Normalized cuts and image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 8, pp. 888–905, 2000.

15. P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, 2004.
16. V. Vapnik, *The nature of statistical learning theory*. New York: Springer, 2nd ed., 2000.
17. "Rapid visual screening of buildings for potential seismic hazards - a handbook - FEMA 154," Mar. 2002.
18. F. Ramm, J. Topf, and S. Chilton, *OpenstreetMap : using and enhancing the free map of the world*. Cambridge, England: UIT Cambridge, 2011.
19. D. Ehrlich, G. Zeug, J. Gallego, A. Gerhardinger, I. Caravaggi, and M. Pesaresi, "Quantifying the building stock from optical high-resolution satellite imagery for assessing disaster risk," *Geocarto International*, vol. 25, pp. 281–293, July 2010.
20. O. Aytakin, I. Ulusoy, A. Erener, and H. S. B. Duzgun, "Automatic and unsupervised building extraction in complex urban environments from multi spectral satellite imagery," in *Recent Advances in Space Technologies, 2009. RAST'09. 4th International Conference on*, pp. 287–291, 2009.
21. L. Durieux, E. Lagabrielle, and A. Nelson, "A method for monitoring building construction in urban sprawl areas using object-based analysis of spot 5 images and existing GIS data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 63, no. 4, pp. 399–408, 2008.
22. M. Wieland, M. Pittore, S. Parolai, and J. Zschau, "Exposure estimation from Multi-Resolution optical satellite imagery for seismic risk assessment," *ISPRS International Journal of Geo-Information*, vol. 1, pp. 69–88, May 2012.
23. D. R. Cox and V. Isham, *Point processes*. London; New York: Chapman and Hall, 1980.
24. C. V. Tao and J. Li, *Advances in mobile mapping technology*. London; New York: Taylor & Francis, 2007.
25. H. Haggrén, H. Hyypää, O. Jokinen, A. Kukko, M. Nuikka, T. Pitkäänen, P. Pöntinen, and P. Rönnholm, "Photogrammetric application of spherical imaging," in *Panoramic Photogrammetry Workshop, Dresden, 19/22 February*, 2004.
26. S. Teller, "Toward urban model acquisition from geo-located images," *Proceedings of Pacific Graphics, Singapore*, pp. 45–51, 1998.
27. A. Torii, A. Imiya, and N. Ohnishi, "Two-and three-view geometry for spherical cameras," in *Proceedings of the Sixth Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras*, 2005.
28. J. Pearl, "bayesian networks: a model of self-activated memory for evidential reasoning," in *Proceedings of Seventh Annual Conference of the Cognitive Science Society*, Aug. 1985.
29. J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Mateo, Calif.: Morgan Kaufmann Publishers, 1991.
30. L. Li, J. Wang, and H. Leung, "Using spatial analysis and bayesian network to model the vulnerability and make insurance pricing of catastrophic risk," *International Journal of Geographical Information Science*, vol. 24, pp. 1759–1784, Nov. 2010.
31. Y. Y. Bayraktarli and M. H. Faber, "Bayesian probabilistic network approach for managing earthquake risks of cities," *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, vol. 5, pp. 2–24, Mar. 2011.
32. D. Heckerman, D. Geiger, and D. Chickering, "Learning bayesian networks: The combination of knowledge and statistical data," *Machine learning*, vol. 20, no. 3, pp. 197–243, 1995.
33. A. Gelfand and A. Smith, "Sampling-based approaches to calculating marginal densities," *Journal of the American statistical association*, pp. 398–409, 1990.
34. G. Grunthal and E. S. Commission., *European macroseismic scale 1998 : EMS-98*. Luxembourg: European Seismological Commission Subcommission on Engineering Seismology Working Group Macroseismic scales, 2nd ed., 1998.
35. "World housing encyclopedia: an EERI and IAEE project." <http://www.world-housing.net/>.