

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

Massimiliano Pittore & Marc Wieland

Natural Hazards

Journal of the International Society
for the Prevention and Mitigation of
Natural Hazards

ISSN 0921-030X

Nat Hazards
DOI 10.1007/s11069-012-0475-z



Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

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

Massimiliano Pittore · Marc Wieland

Received: 31 March 2012 / Accepted: 25 October 2012
© Springer Science+Business Media Dordrecht 2012

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

1 Introduction

Seismic risk is usually defined by three fundamental components, each of which should be separately quantified (Coburn et al. 1994; Calvi et al. 2006), 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, not considering aftershocks) and is almost independent of anthropic activities (Michael 2011). 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

M. Pittore (✉) · M. Wieland
GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany
e-mail: pittore@gfz-potsdam.de

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 (Steimen et al. 2004; Grossi et al. 1999). 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 (Wieland et al. 2012), 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 was also suggested, but the integration of collected information into a vulnerability model was not addressed. The current contribution thus further extends toward 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. The proposed information collection and integration scheme can be used in complex, rapidly changing urban environments to obtain, with a limited amount of resources, a sound model of exposure and seismic vulnerability. In Sect. 2, a general outline of the proposed approach is provided, while Sect. 3 briefly introduces the study area. In Sect. 4, the stratification and sampling scheme, based on the processing of remote sensing data, is briefly recalled. In Sect. 5, an in situ data collection strategy, taking advantage of omni-directional imaging, is described, along with a manual satellite image interpretation stage. In Sect. 6, a fully probabilistic information integration approach for vulnerability assessment, based on Bayesian Networks is presented. Finally, in Sects. 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 Fig. 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 (Wieland et al. 2012) for further details. Once the area of interest has been divided into 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 devoted 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 Fig. 1b,

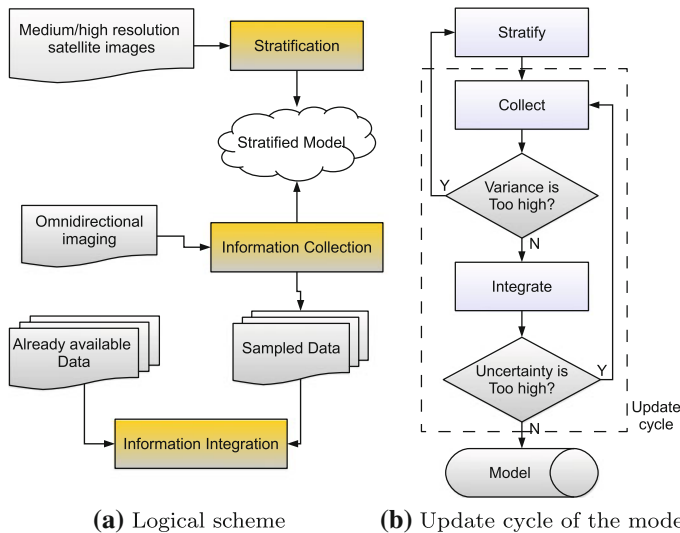


Fig. 1 Outline of the data collection and integration scheme. **a** The logical scheme. **b** The flowchart of the whole process

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 Fig. 1b) can be repeated over time as the environmental conditions change or need for less uncertain models arises.

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 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 (Erdik et al. 2005). The city itself was founded in 1825 and counted 859,800 inhabitants in 2010. Since the 1970s 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 (Erdik et al. 2005) lead to an increasingly high level of seismic risk in the area (Bindi et al. 2011). The actual area under observation within this study covers an area of 665 km^2 and is outlined in Fig. 2.

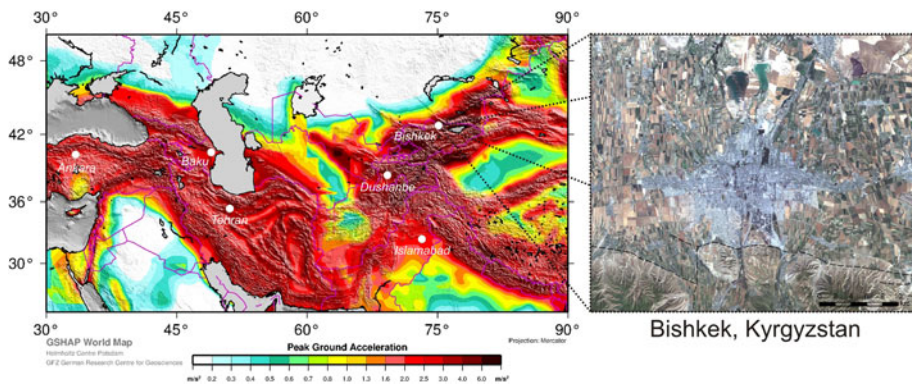


Fig. 2 Location and extent of the study area Bishkek, Kyrgyzstan. *On the left:* the GSHAP map (1999), *on the right:* Landsat TM (2009)

4 Stratification

In order to achieve a reasonable stratification of the considered town, we suggest a multi-stage approach (Fig. 3); the spatial extent of the town is first subdivided into small geographic units, obtained by unsupervised low-level segmentations of a time-series of medium-resolution multi-spectral satellite images (see Sect. 4.1). Then, the geographical units are labeled and grouped into bigger clusters depending on their predominant land use/land cover and approximate age using partly supervised classifications and change-detection analysis (see Sect. 4.2). The technical details can be found in (Wieland et al. 2012). In this section we provide a brief discussion on the approach and a preliminary validation.

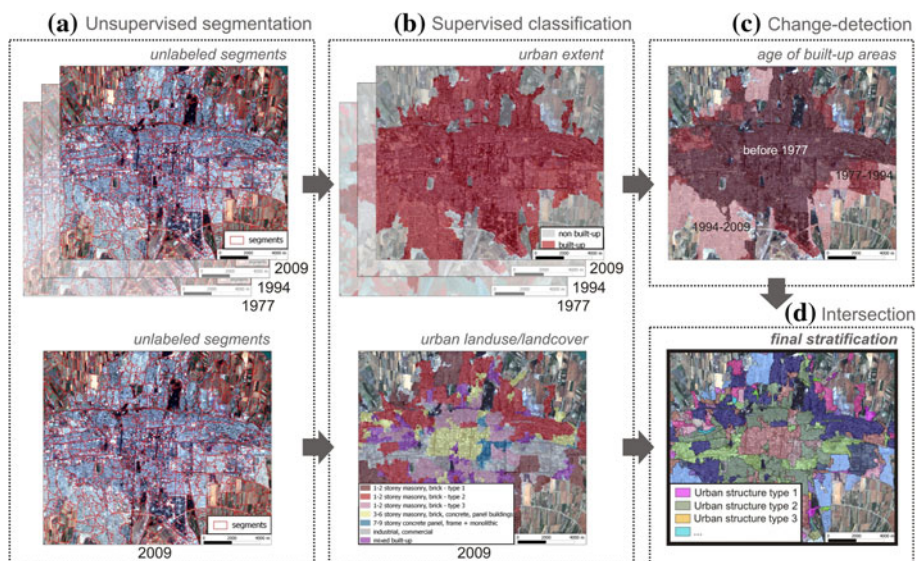


Fig. 3 Stratification workflow exemplified for Bishkek, Kyrgyzstan

Stratification is a well-known sampling technique (Cochran 1977), 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 (Tobler 1970). 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 (Rapoport 1980; Steiniger et al. 2008), 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, while 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 (Shi and Malik 2000). Several algorithms for image segmentation have been proposed in the literature in the last decade. As proposed in (Wieland et al. 2012) we have opted for the approach proposed by Felzenszwalb and Huttenlocher (2004), 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. 3a). A scale parameter and a merge parameter must then be chosen, which is done on a trial-and-error basis.

4.2 Supervised classification and multi-temporal change detection

The segmentation stage described in Sect. 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 (Vapnik 2000). 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 (Erdik et al. 2005). 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.

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. 3b). An evaluation of the performance of this processing stage in terms of receiver operating characteristic curves (ROC) is provided in Fig. 4. Differences in the performance of the built-up areas' delineation are dependent upon both the different quality of imagery sources and 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 km² was classified as being built-up in the study area. In 1994 the built environment accounted for 145 km², while in 2009, the city reached a total built-up area of 207 km² 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. 3c), a total of 21 possible combinations (defining the final

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

Reference											
Class	1	2	3	4	5	6	7	8	9	10	TP (%)
<i>Classification</i>											
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

A total of 10 test samples *per* class have been used

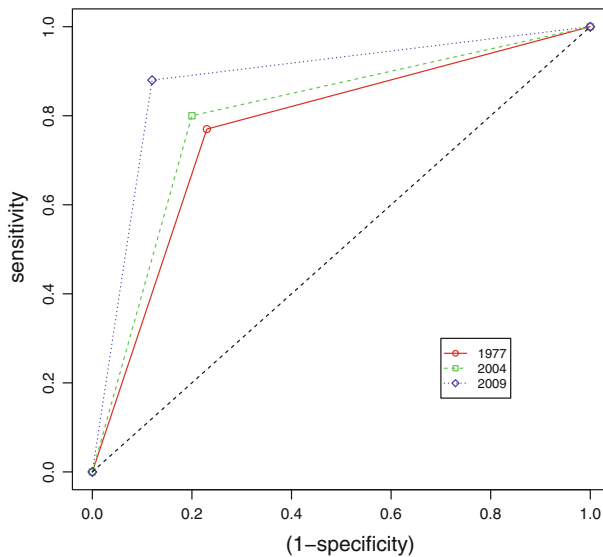


Fig. 4 ROC for three binary urban extent classifications from Landsat images

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. 3d. Further explanations and details can be found in (Wieland et al. 2012).

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. 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

Building 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 (Ramm et al. 2011). A complete manual digitization of a large town is often very time and resource intensive (Ehrlich et al. 2010). 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 (Aytekin et al. 2009; Durieux et al. 2008; Wieland et al. 2012). 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 4-week project, and 455 were imported from OpenStreetMap. The collected footprints' area amounts to about 1.5 km², roughly equal to 0.75 % of the total 2009 built-up area (estimated to be 207 km²) in Bishkek. Details on the selection and sizing of the sample areas can be found in (Wieland et al. 2012).

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 Sect. 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 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

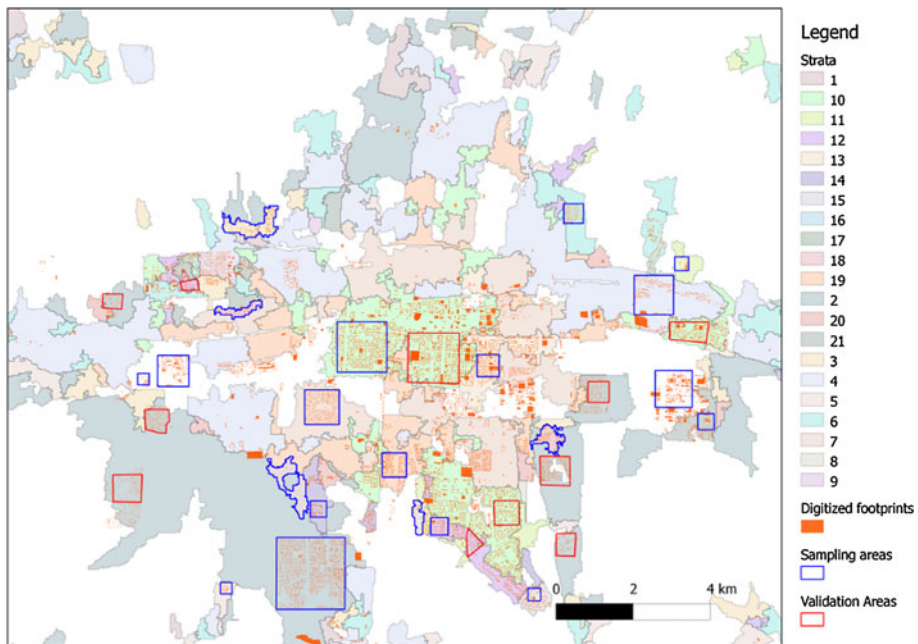


Fig. 5 Bishkek: manually digitized footprints, along with the stratification. Respectively in *blue* and *red* the considered sample and validation areas are shown

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.

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

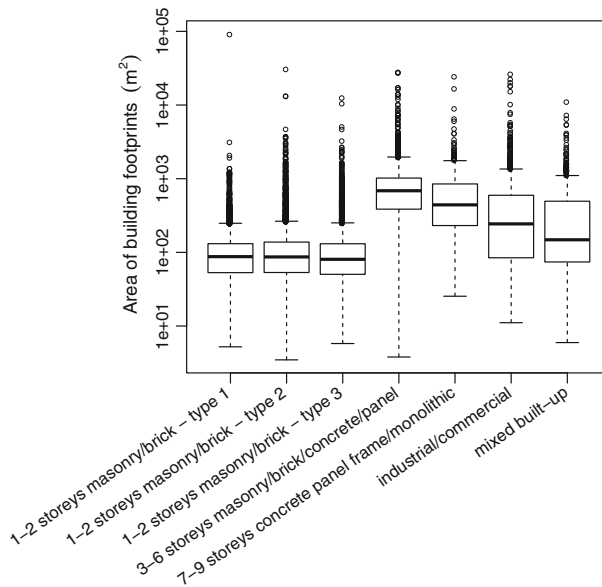


Fig. 6 Distribution of area of digitized building footprints in the different strata, grouped by land use/land cover. The LULC classes are clearly separated also in terms of area distribution, therefore indirectly confirming the quality of stratification

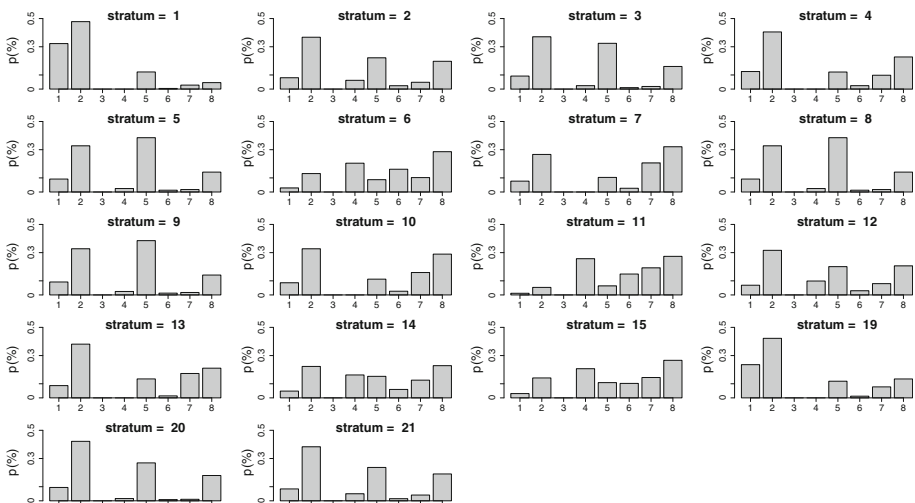


Fig. 7 Distribution of area of digitized building footprints in the different strata, grouped by age

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 (Cox and

Isham 1980). 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 $\{N(A)\}_{A \in \mathcal{A}}$ 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 $ci = \pm \frac{1.96 * \sigma_{sa}}{\sqrt{N_{ssa}}}$.

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, has 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.

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 $N_b = 182,000 \pm 14,000$.

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 (Tao and Li 2007). 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)

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)

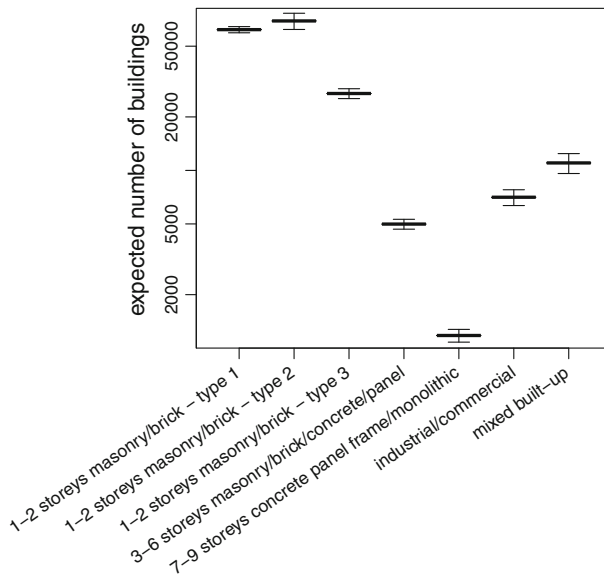
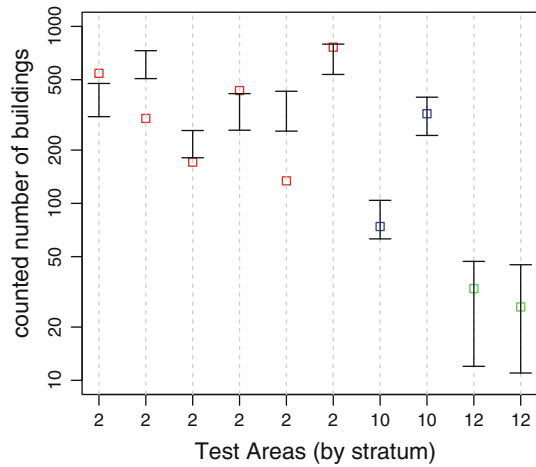


Fig. 9 Expected number of buildings, by land use/land cover, with SE for Bishkek

and microphones can be added, depending on the specific application. In this context, omni-directional imaging (Haggrén et al. 2004; Teller 1998), 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 km² 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 (Wieland et al. 2012), 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 (Torii et al. 2005). 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 m away from the observer location. Covering such a field of view with a conventional camera would imply capturing several images with different aims.

The above-described processing has been applied to an extended sample of collected 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 m 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-storey masonry brick/concrete/panel and LULC:7–9-storey 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).

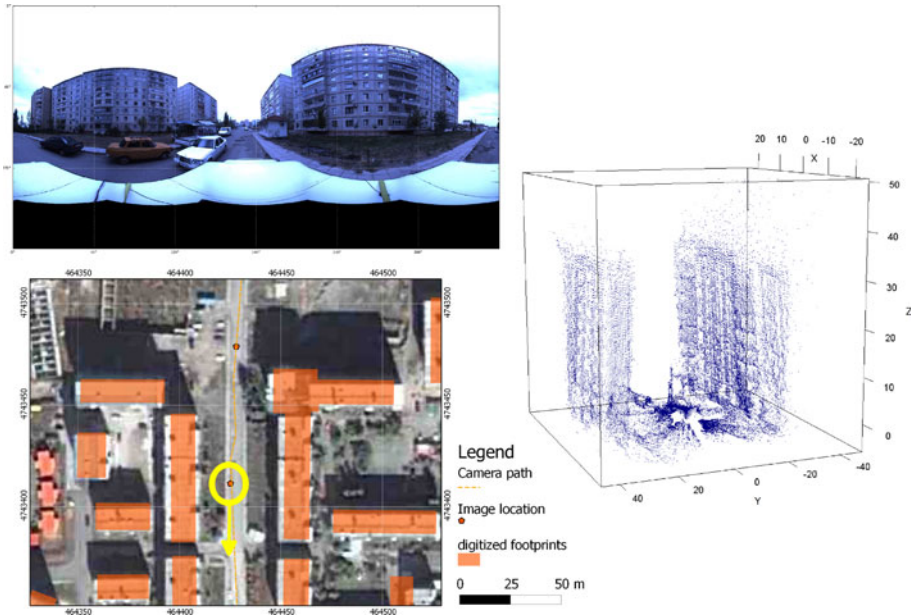


Fig. 10 Example, for Bishkek, of a 3D reconstruction based on the processing of omni-directional image stream. The map in the *lower left corner* shows the locations of the optical center of the camera when the images were shot. In *orange*, several digitized footprints are also shown. In the *upper 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

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 (10–40 m, also depending on its height), there should not be major visual occlusions (vegetation, poles, road signs), 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.

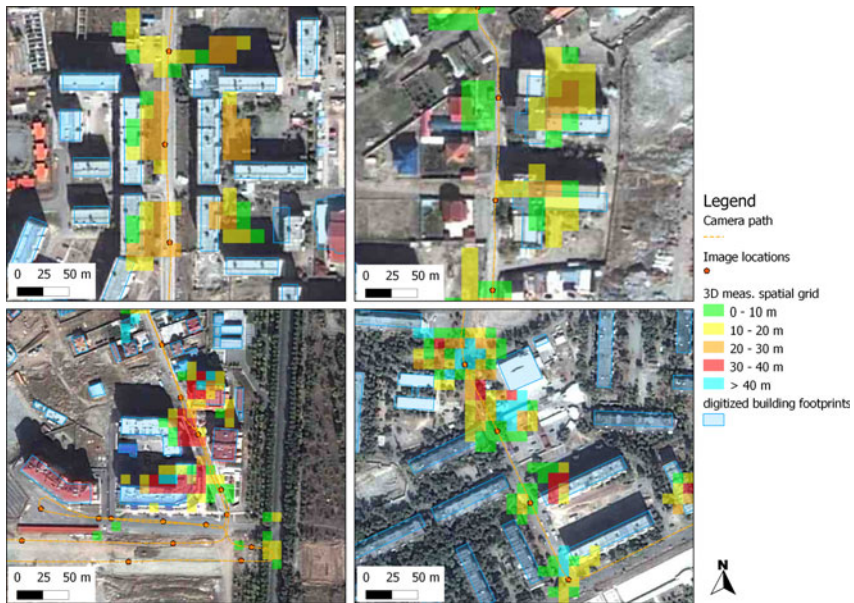


Fig. 11 Results of extracting 3D measurements from omni-directional image processing. *Colored cells* represent elements of the sparse measurements grid. Four different sub-areas of Bishkek are shown

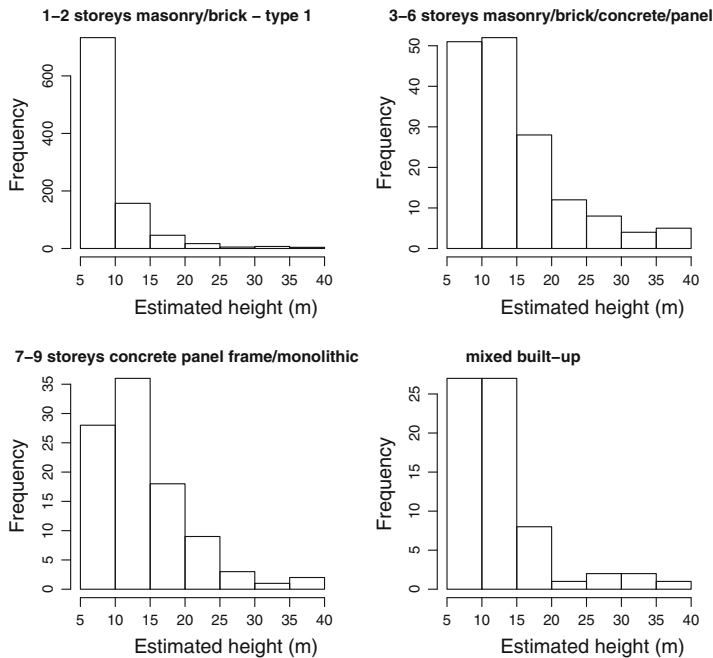


Fig. 12 Distribution of estimated building height in spatial segments classified with different land use/land cover labels. The upper histograms are in accordance with the expected height values, while the bottom histograms show a larger spread of the values

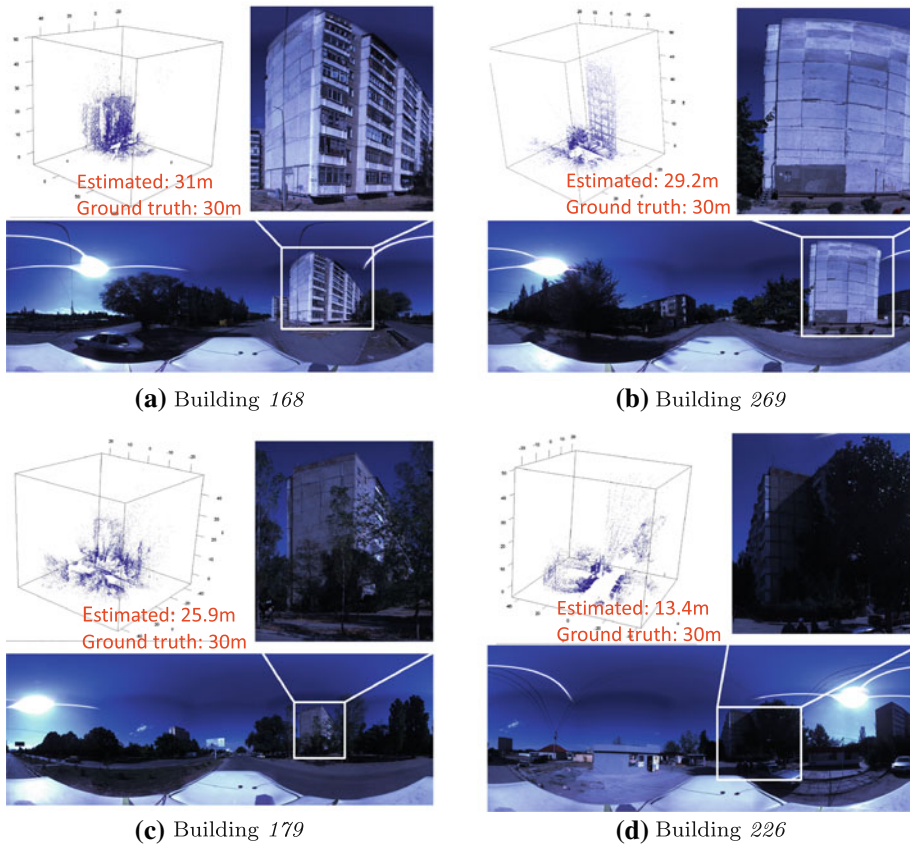


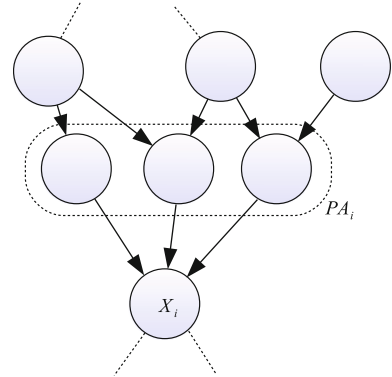
Fig. 13 Examples of building height estimation, compared with available ground truth information. The upper row cases **a**, **b** shows results obtained under optimal conditions, the lower row cases **c**, **d** shows the degraded quality of the assessment where non-optimal conditions are encountered

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 (Pearl 1985, 1991). Bayesian networks have been recently proposed, within the context of earthquake engineering, for the spatial modeling of seismic vulnerability (Li et al. 2010) for insurance pricing applications or to tackle the broader problem of managing uncertainties in seismic risk (Bayraktarli and Faber 2011). 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 variables can be expressed as:

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



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

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 (Heckerman et al. 1995)). 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 (Gelfand and Smith 1990). 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 Sect. 5), the most likely building type and its most likely structural vulnerability class according to EMS-98 (Grunthal and Commission 1998). 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 to 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, 5, 6, 7, 8. For each node without parents, the prior distribution is also provided. The prior for nodes A (age) and S (LULC) is 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 (<http://www.world-housing.net>). 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.

It's 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.

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 Sect. 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 = 30 m, storeys = 9, EMS-98 vulnerability = E (detailed information on the WHE building type is not available). As we can observe from

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

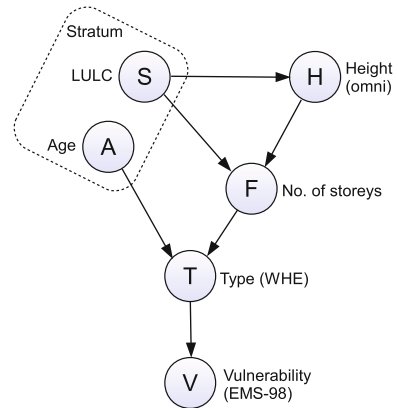


Table 3 Bayes network for vulnerability assessment. (summary of nodes)

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	A
2	B
3	C
4	D
5	E
6	F

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, Sect. 5.4)

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 Sect. 5, the BN described in Sect. 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 ($p(\text{state}_i) = \frac{1}{n_{\text{states}}}$) 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

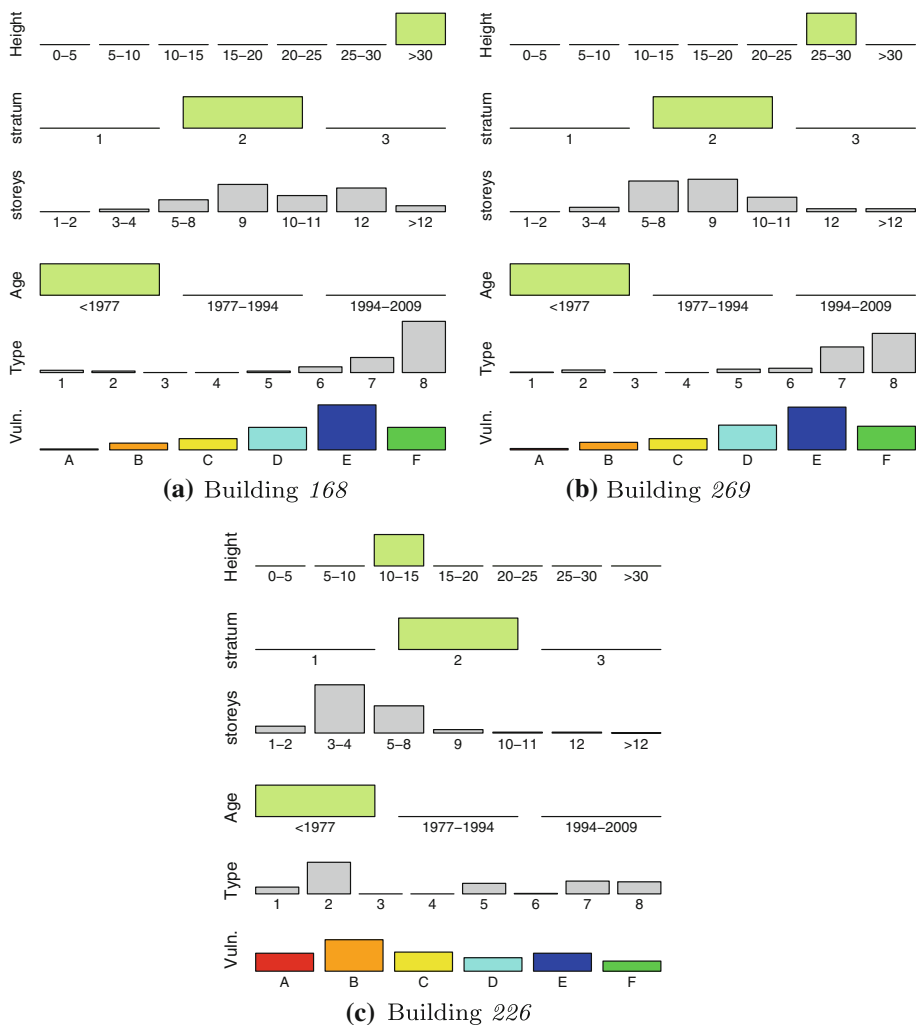


Fig. 16 Posterior probability distributions computed for the buildings depicted in Fig. 13. In *light green* are displayed the values inserted as evidence; the strongly peaked probabilities indicate a strong belief in the values

Table 5 Node: “building type”

State	Building type	WHE report
1	Houses with mud walls and thatch roofs	42
2	Two-storey 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

“commercial, industrial” are provided. A meaningful 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 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 Sect. 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 buildings counts aggregated with respect of building types is provided, considering the aggregated land use classification as a spatial reference. The uncertainty in the buildings count is significant and depends on the size of the sample used for the density

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 storeys, masonry, brick	0.829
2	3–6 storeys, masonry, reinforced concrete, panel	0.143
3	7–9 storeys, reinforced concrete, panel, monolithic	0.028

Table 7 Nodes relative to building height

(a) Node: "height"		
State	Height (m)	Prior
1	0–5	0.14286
2	5–10	0.14286
3	10–15	0.14286
4	15–20	0.14286
5	20–25	0.14286
6	25–30	0.14286
7	30–50	0.14286
(b) Node: "num. storeys"		
State	Num. storeys	
1	1–2	
2	3–4	
3	5–8	
4	9	
5	10–11	
6	12	
7	13–18	

estimation and on the simplistic model used in the modeling of the density distribution itself. In Fig. 19, the spatial distribution of the areas with aggregated land use attributes considered in Fig. 17 and Table 9 is shown.

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, have been computed by averaging over each considered building footprint. In Fig. 20, an overview of the probability

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

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

Aggregated land use	Type 1 ± σ	Type 2–4 ± σ	Type 5–6 ± σ	Type 7–8 ± σ	Total ± σ
1–2-storey masonry, brick	26,130 ± 9,645	68,773 ± 25,385	37,222 ± 13,739	26,261 ± 9,693	158,386 ± 58,464
3–6-storey masonry, brick, concrete, panel buildings	407 ± 132	1,816 ± 591	1,316 ± 429	1,450 ± 472	4,989 ± 1,625
7–9-storey concrete panel, frame + monolithic	65 ± 27	413 ± 174	228 ± 96	471 ± 198	1,177 ± 496

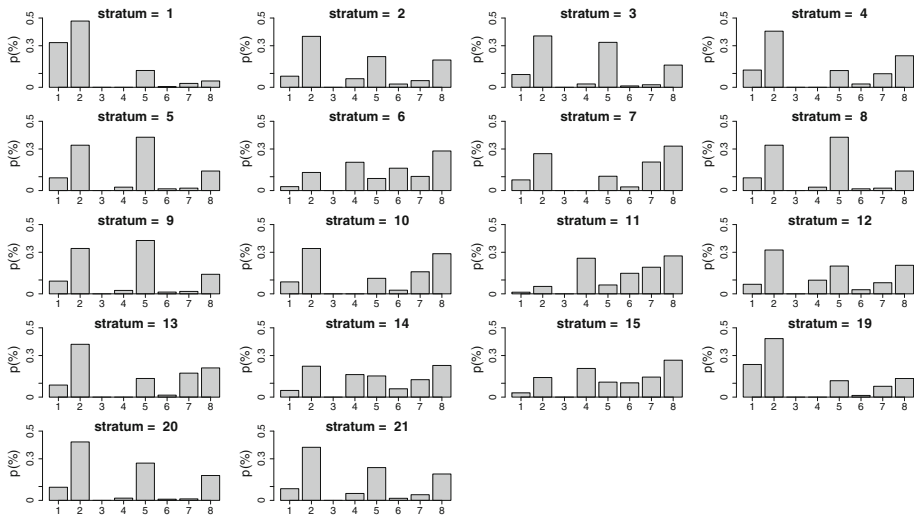


Fig. 17 Probability distribution of the considered building types according to the WHE taxonomy, for each stratum not classified as “commercial, industrial”

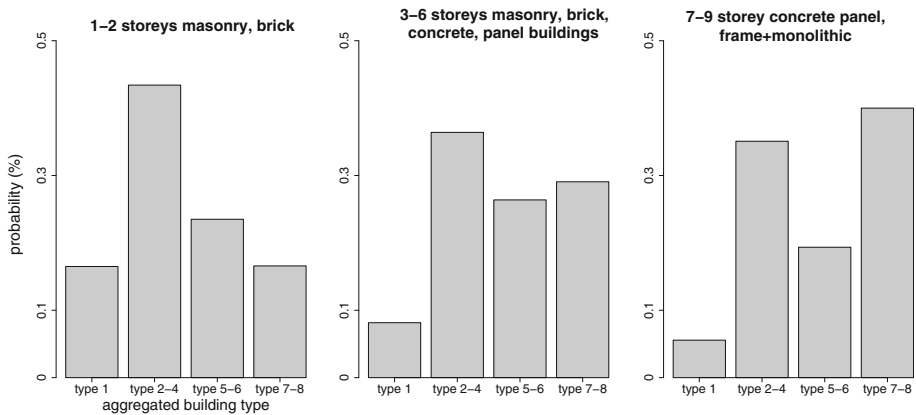


Fig. 18 Probability distributions of the aggregated building types, averaged based on three aggregated land use classes

distribution of the EMS-98 vulnerability states is provided for each stratum not classified as “commercial, industrial”.

Considering the aggregated inventory composition described in Sect. 7 and displayed in Fig. 18, we can compute the relative vulnerability probability distribution. Figure 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-storey 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-storey concrete panel, frame + monolithic”) is lower and with higher relative frequency of low-vulnerability classes such as *D*, *E*, and *F*.

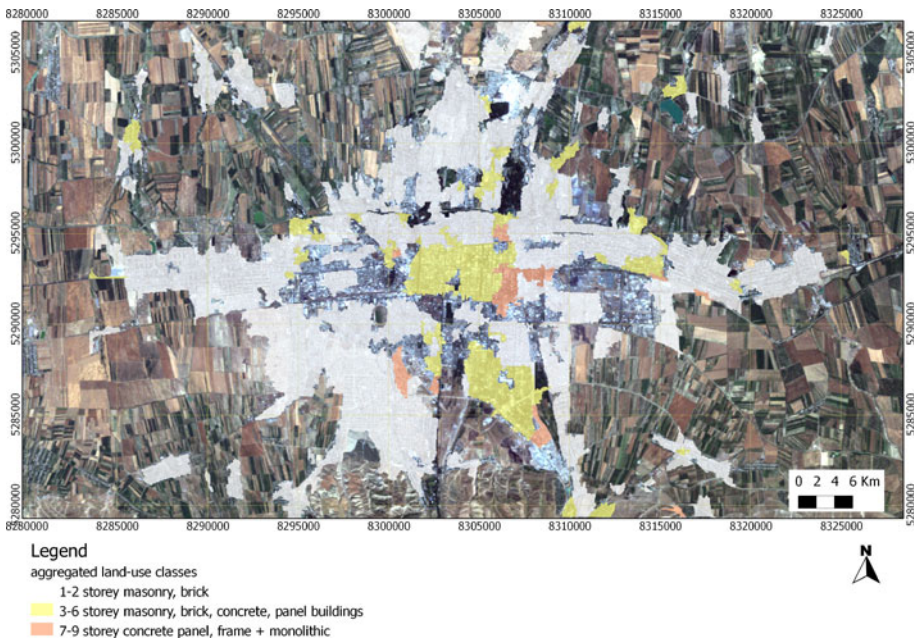


Fig. 19 Spatial distributions of areas with aggregated land use attributes considered in Fig. 17 and Table 9

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

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

where n is the number of possible vulnerability states (in our case $n = 6$), and $\{V_{i=0 \dots 5} = A, B, C, D, E, F\}$ 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 of 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.

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

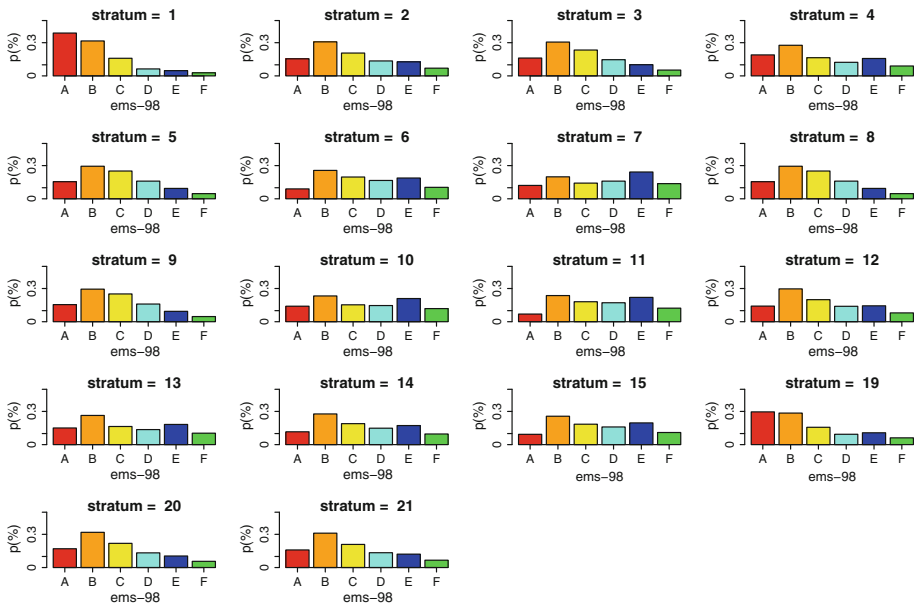


Fig. 20 Probability distribution of the vulnerability in EMS-98 scale for each stratum not classified as “commercial, industrial”

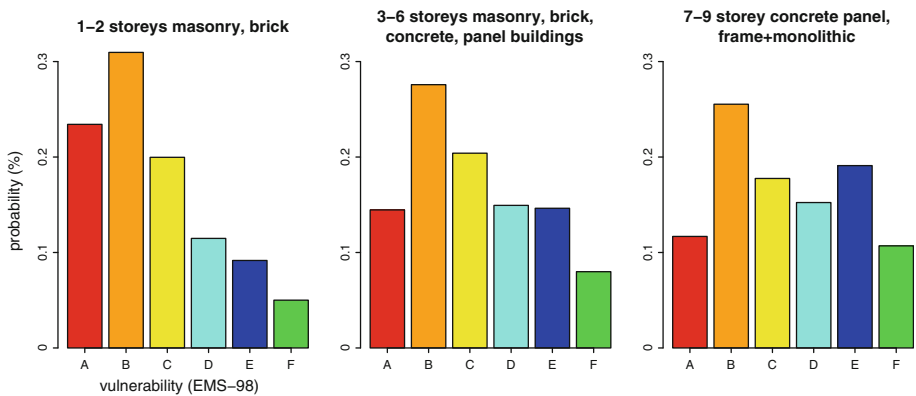


Fig. 21 Probability distribution of vulnerability in the EMS-98 scale for three aggregated land use attributes (see also Fig. 18)

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.

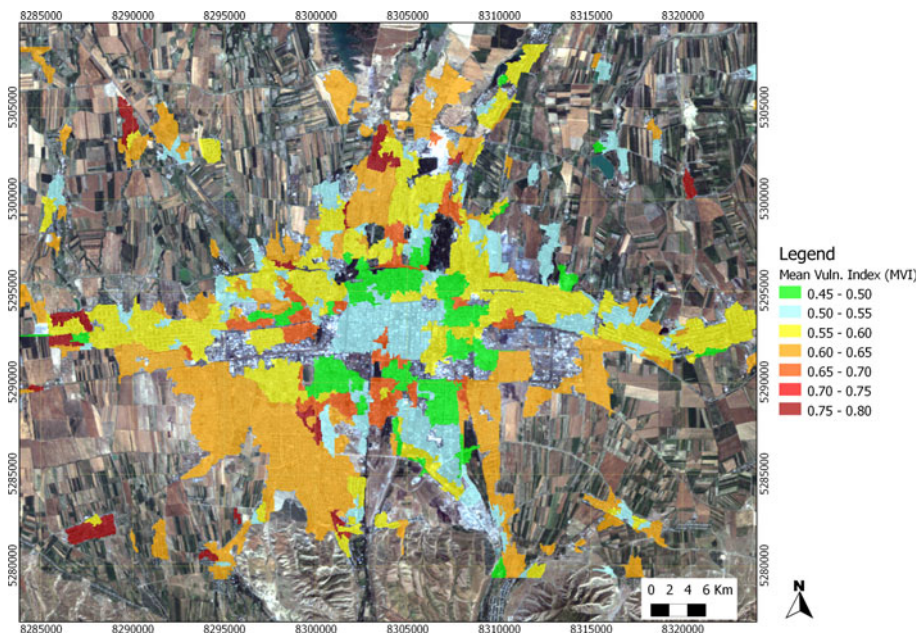


Fig. 22 Spatial distribution of mean vulnerability index (MVI) in Bishkek, Kyrgyzstan. A 2009 Landsat TM image is displayed in the background

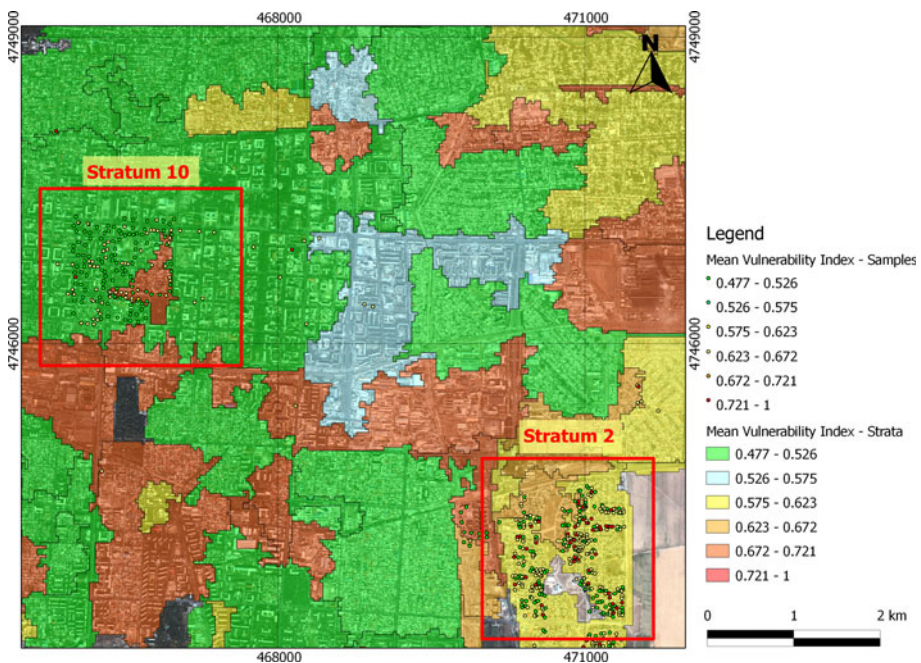


Fig. 23 Spatial distribution of sampled vulnerability and comparison with inferred values

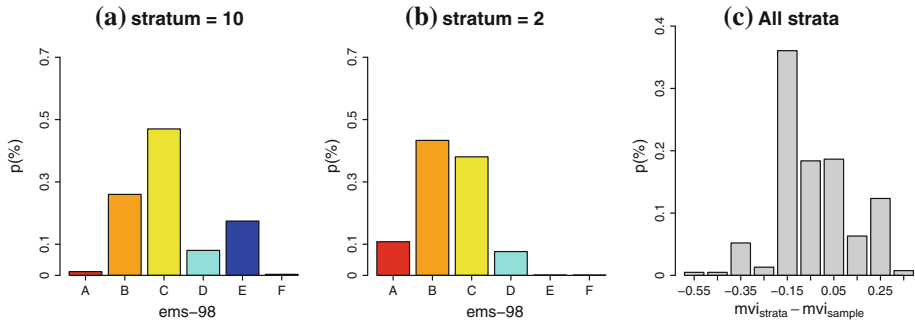


Fig. 24 **a, b** Empirical distributions of seismic vulnerability of sampled buildings in the two *strata* considered in Fig. 23. The distributions should be compared with those in Fig. 20. **c** Difference between average vulnerability of strata and vulnerability of sampled buildings, expressed as *mvi*

Figure 24a, 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 24c 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 with the stratification. Since the considered sample is covering only a subset of the stratification, only a preliminary and qualitative validation can be performed.

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 (Wieland et al. 2012). 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. This work was supported by PROGRESS (Georisiken im Globalen Wandel), Helmholtz-EOS (Earth Observation System) and CEDIM (Center for Disaster Management and Risk Reduction).

References

- Aytekin O, Ulusoy I, Erener A, Duzgun HSB (2009) 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

- Bayraktarli YY, Faber MH (2011) Bayesian probabilistic network approach for managing earthquake risks of cities. *Georisk Assess Manag Risk Eng Syst Geohazards* 5:2–24
- Bindi D, Mayfield M, Parolai S, Tyagunov S, Begaliev U, Abdrakhmatov K, Moldobekov B, Zschau J (2011) Towards an improved seismic risk scenario for bishkek, kyrgyz republic. *Soil Dyn Earthq Eng* 31:521–525
- Calvi GM, Pinho R, Magenes G, Bommer JJ, Restrepo-Velez LF, Crowley H (2006) Development of seismic vulnerability assessment methodologies over the past 30 years. *ISET J Earthq Technol* 43(472):75–104
- Coburn AW, Spence R, Pomonis A (1994) Vulnerability and risk assessment. UNDP Disaster Management Training Programme
- Cochran W (1977) Sampling techniques, 3rd edn. Wiley, New York
- Cox DR, Isham V (1980) Point processes. Chapman and Hall, London
- Durieux L, Lagabrielle E, Nelson A (2008) A method for monitoring building construction in urban sprawl areas using object-based analysis of spot 5 images and existing GIS data. *ISPRS J Photogramm Remote Sens* 63(4):399–408
- Ehrlich D, Zeug G, Gallego J, Gerhardinger A, Caravaggi I, Pesaresi M (2010) Quantifying the building stock from optical high-resolution satellite imagery for assessing disaster risk. *Geocarto Int* 25:281–293
- Erdik M, Rashidov T, Safak E, Turdukulov A (2005) Assessment of seismic risk in tashkent, uzbekistan and bishkek, kyrgyz republic. *Soil Dyn Earthq Eng* 25:473–486
- Felzenszwalb PF, Huttenlocher DP (2004) Efficient graph-based image segmentation. *Int J Comput Vis* 59(2):167–181
- Gelfand AE, Smith AFM (1990) Sampling-based approaches to calculating marginal densities. *J Am Stat Assoc* 85:398–409
- Grossi P, Kleindorfer PR, Kunreuther H (1999) The impact of uncertainty in managing seismic risk: the case of earthquake frequency and structural vulnerability. Wharton Financial Institution Center WP 99-03-26, Risk Management and Decision Processes Center, The Wharton School, Philadelphia, PA
- Grunthal G, Commission ES (1998) European macroseismic scale 1998 : EMS-98. Luxembourg: European seismological commission subcommission on engineering seismology working group macroseismic scales, 2nd edn
- Haggrén P, Hyypä H, Jokinen O, Kukko A, Nuikka M, Pitkänen T, Pöntinen P, Rönholm P (2004) Photogrammetric application of spherical imaging. In: Panoramic photogrammetry workshop, Dresden, 19/22 February
- Heckerman D, Geiger D, Chickering D (1995) Learning bayesian networks: the combination of knowledge and statistical data. *Mach learn* 20(3):197–243
- Li L, Wang J, Leung H (2010) Using spatial analysis and bayesian network to model the vulnerability and make insurance pricing of catastrophic risk. *Int J Geogr Inf Sci* 24:1759–1784
- Michael AJ (2011) Random variability explains apparent global clustering of large earthquakes. *Geophys Res Lett* 38:L21301. doi:[10.1029/2011GL049443](https://doi.org/10.1029/2011GL049443)
- Pearl J (1985) bayesian networks: a model of self-activated memory for evidential reasoning. In: Proceedings of seventh annual conference of the cognitive science society
- Pearl J (1991) Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Publishers, San Mateo
- Ramm F, Topf J, Chilton S (2011) OpenstreetMap: using and enhancing the free map of the world. UIT Cambridge, Cambridge
- Rapoport A (1980) Neighborhood heterogeneity or homogeneity: the field of man-environment studies. *Archit Behav* 1(1):65–77
- Shi J, Malik J (2000) Normalized cuts and image segmentation. *IEEE Trans Pattern Anal Mach Intell* 22(8):888–905
- Steimen S, Fh D, Giardini D, Bertogg M, Tschudi S (2004) Reliability of building inventories in seismic prone regions. *Bull Earthq Eng* 2(3):361–388
- Steiniger S, Lange T, Burghardt D, Weibel R (2008) An approach for the classification of urban building structures based on discriminant analysis techniques. *Trans GIS* 12:31–59
- Tao CV, Li J (2007) Advances in mobile mapping technology. Taylor & Francis, London
- Teller S (1998) Toward urban model acquisition from geo-located images. In: Proceedings of pacific graphics, Singapore pp 45–51
- Tobler WR (1970) A computer movie simulating urban growth in the detroit region. *Econ Geogr* 46(2):234–240
- Torii A, Imiya A, Ohnishi N (2005) Two-and three-view geometry for spherical cameras. In: Proceedings of the sixth workshop on omnidirectional vision, camera networks and non-classical cameras

- Vapnik V (2000) The nature of statistical learning theory, 2nd edn. Springer, New York
- Wieland M, Pittore M, Parolai S, Zschau J (2012) Exposure estimation from Multi-Resolution optical satellite imagery for seismic risk assessment. ISPRS Int J Geo-Inf 1:69–88
- Wieland M, Pittore M, Parolai S, Zschau J, Moldobekov B, Begaliev U (2012) Estimating building inventory for rapid seismic vulnerability assessment: towards an integrated approach based on multi-source imaging. Soil Dyn Earthq Eng