Article:

Ettinger, S, Mounaud, L, Magill, C et al. (11 more authors) (2016) Building vulnerability to hydro-geomorphic hazards: Estimating damage probability from qualitative vulnerability assessment using logistic regression. Journal of Hydrology, 541. pp. 563-581. ISSN 0022-1694

https://doi.org/10.1016/j.jhydrol.2015.04.017

Building vulnerability to hydro-geomorphic hazards: estimating damage probability from qualitative vulnerability assessment using logistic regression

Susanne Ettinger [a], Loïc Mounaud [b], Christina Magill [c], Anne-Françoise Yao-Lafourcade [b], Jean-Claude Thouret [a], Vern Manville [d], Caterina Negulescu [e], Giulio Zuccaro [f], Daniela De Gregorio [f], Stefano Nardone [f], Juan Alexis Luque Uchuchoque [g], Anita Arguedas [h], Luisa Macedo [i], Nélida Manrique Llerena [i]

[a] Clermont Université, Université Blaise Pascal, Laboratoire Magmas et Volcans, CNRS UMR6524, IRD-R163 and ClerVolc, 5 rue Kessler, F-63000 Clermont-Ferrand

[b] Université Blaise Pascal, Laboratoire de Mathématiques, CNRS UMR 6620, Campus des Cézeaux, 63171 Aubière Cedex, France

[c] Risk Frontiers, Macquarie University, NSW2109, Australia

[d] School of Earth and Environment, University of Leeds, Leeds LS29JT, United Kingdom

[e] BRGM, 3 avenue C. Guillemin, BP 36009, 45060 Orléans Cedex 2, France

[f] Plinivs Study Centre, Universita Federico II, Naples, Italy

[g] Defensa Civil Paucarpata and NGO PREDES, Av. Unión Nro. 200, Urb. César Vallejo, Paucarpata, Arequipa, Peru

[h] Defensa Civil, Región Arequipa, Municipalidad Paucarpata, Arequipa, Peru

[i] INGEMMET, Instituto Nacional de Geologia, Minero y Metalurgico, Av. Dolores (Urb. Las Begonias B-3), José Luis Bustamante y Rivero, Arequipa, Peru

Corresponding author:

Susanne ETTINGER

Current address:
Université Blaise Pascal, Laboratoire Magmas et Volcans, CNRS UMR6524, IRD R163,
5 rue Kessler, 63038 Clermont-Ferrand, France
Telephone: +33 (0)4 73 34 67 99
Email: susanne.ettinger@gmail.com

Abstract

The focus of this study is an analysis of building vulnerability through investigating impacts from the 8 February 2013 flash flood event along the Avenida Venezuela channel in the city of Arequipa, Peru. On this day, 124.5mm of rain fell within 3 hours (monthly mean: 29.3mm) triggering a flash flood that inundated at least 0.4km^2 of urban settlements along the channel, affecting more than 280 buildings, 23 of a total of 53 bridges (pedestrian, vehicle and railway), and leading to the partial collapse of sections of the main road, paralyzing central parts of the city for more than one week.

This study assesses the aspects of building design and site specific environmental characteristics that render a building vulnerable by considering the example of a flash flood event in February 2013. A statistical methodology is developed that enables estimation of damage probability for buildings. The applied method uses observed inundation height as a hazard proxy in areas where more detailed hydrodynamic modeling data is not available. Building design and site-specific environmental conditions determine the physical vulnerability. The mathematical approach considers both physical vulnerability and hazard related parameters and helps to reduce uncertainty in the determination of descriptive parameters, parameter interdependency and respective contributions to damage. This study aims to (1) enable the estimation of damage probability for a certain hazard intensity, and (2) obtain data to visualize variations in damage susceptibility for buildings in flood prone areas. Data collection is based on a post-flood event field survey and the analysis of high (sub-metric) spatial resolution images (Pléiades 2012, 2013). An inventory of 30 city blocks was collated in a GIS database in order to estimate the physical vulnerability of buildings. As many as 1103 buildings were surveyed along the affected drainage and 898 buildings were included in the statistical analysis. Univariate and bivariate analyses were applied to better characterize each vulnerability parameter. Multiple corresponding analyses revealed strong relationships between the “Distance to channel or bridges”, “Structural building type”, “Building footprint” and the observed damage. Logistic regression enabled quantification of the contribution of each explanatory parameter to potential damage, and determination of the significant parameters that express the damage susceptibility of a building. The model was applied 200 times on different calibration and validation data sets in order to examine performance. Results show that 90% of these tests have a success rate of more than 67%. Probabilities (at building scale) of experiencing different damage levels during a future event similar to the 8 February 2013 flash flood are the major outcomes of this study.

Keywords

Flash flood; Vulnerability; Logistic regression; Damage probability; Risk; Arequipa

1. Introduction

On February 8 2013, heavy rainfall (124.5mm within 3 hours versus a monthly mean of 29.3mm) triggered a flash flood event along the Avenida Venezuela channel in the city of

Arequipa, Peru. On this day, more than 280 buildings and 23 of a total of 53 bridges (pedestrian, vehicle and railway) were affected; the partial collapse of sections of a major road paralyzed central parts of the city for more than one week. Previous risk assessment studies in Arequipa did not include the Avenida Venezuela channel due to its smaller size and largely confined channel course. The high recurrence rate of hydro-geomorphic hazards (Martelli, 2011; Thouret et al., 2013, 2014), and apparent locally high vulnerability of buildings and critical infrastructure in Arequipa, are major motivations for this study.

Risk in the context of disaster risk management is commonly defined as a potential loss for a given probability function (Crichton, 1999; Kaplan and Garrick, 1981). In the standard conceptual framework, risk is the product of hazard, vulnerability and exposure (Cardona, 2004; Carreno et al., 2006). While the hazard is generally described by its severity, e.g. inundation height for a given return return period, exposure relates to the number and value of elements potentially affected (Hiete and Merz, 2009). Many different definitions, concepts and methods to systemize vulnerability exist in the current literature (Birkmann, 2006; Cutter, 2003; Wisner et al., 2004; Thywissen, 2006; IPCC, 2007; Bründl et al., 2009). In this study we follow the definition for physical vulnerability proposed by Glade (2003) and Villagran de Leon (2006) as the predisposition of an element or system to be affected or susceptible to damage as the result of the natural hazard's impact. Vulnerability assessment for hydro-geomorphic hazards such as dilute floods, debris flows, hyperconcentrated flows etc. is inherently complex, mainly as a result of the following factors (Gaume et al., 2009): (i) lack of accurate or real-time observational data necessary for reliable hazard analysis; (ii) only substantial damage information is generally recorded and accurate information on failure characteristics is often missing (Fuchs et al., 2007b; Papathoma Köhle et al., 2011); (iii) different time and geographical scales involved (Gruntfest, 2009; Marchi et al., 2010); (iv) natural adjustments of the environment to return to a state of equilibrium; (v) rapid intervention by technical services to restore functionality of urban infrastructure reduces the time frame for damage assessment in the field; (vi) site-specific triggering processes and upstream-downstream evolution of debris-flow phenomena (Di Baldassarre and Montanari, 2009). If field investigations are conducted to study and record structural damage following a hazard event, these data are then generally correlated to the process intensity, frequently derived from deposition height or inundation height, in order to develop empirical fragility curves (Fuchs et al., 2007a,b; Holub and Fuchs, 2008). These curves are then employed within risk assessments to estimate structural damage in future events. The lack of high-quality observational evidence and uniform, i.e. non site-specific, approaches to data collection, implies that the majority of fragility curves are still developed using expert judgment (Papathoma Köhle et al., 2012; Totschnig and Fuchs, 2013). The compilation of field data for different sites in the European Alps, Taiwan etc. published in recent studies (Totschnig et al., 2011; Holub et al., 2012; Papathoma Köhle et al., 2012; Totschnig and Fuchs, 2013) has helped to develop vulnerability functions applicable within the framework of risk management for specific regions (Totschnig and Fuchs, 2013). If the required input data are available, the method is transferable to other alpine regions. However, data availability remains a major constraint in many countries (e.g., Douglas, 2007; Jakob et al., 2012; Lo et al., 2012; Totschnig and Fuchs, 2013). For Latin America, very few case studies have been published with a focus on physical vulnerability analysis. In contrast to many sites monitored and equipped in the European Alps, areas prone to hydro-geomorphic hazards in the Andes are rarely monitored, and in the worst case, not even mapped. It therefore becomes difficult to apply methods derived from European experience in the same or a similar way. In addition there is a critical lack of observational data collected in the immediate aftermath of disasters. For the study presented here, data to apply existing vulnerability assessment methods were not available, although alternative information could be collected.

Flash floods are common in semi-arid areas, such as Arequipa, and can, despite their infrequent nature, have a devastating effect in both geomorphological and human terms (Gaume et al., 2009; Jonkman and Vrijling, 2008; Martínez Ibarra, 2012). The occurrence of flash floods is highly variable, both spatially and temporally, most occurring as the result of localized intense storms (Graf, 1988; Reid and Frostick, 1992; Hooke and Mant, 2000). Several important factors arise as a result of these characteristics. First, areas prone to flash floods need to be adequately prepared. Because events usually occur unexpectedly, warning and preparation are essential (Montz and Gruntfest, 2002; Collier, 2007; Borga et al., 2008; Gaume et al., 2009); however, because events are typically rare, the motivation to invest time and resources into such activities may be lower than for more frequent hazards (Gruntfest and Handmer, 2001). Flash floods usually affect relatively small areas and losses resulting from them do not always generate much long-term response, unless there is high loss of life. However, losses per unit (acre, square mile, or kilometer) of area affected tend to be high compared to other events such as riverine floods or hurricanes due to locally high intensity (Gaume et al., 2009; Martínez Ibarra, 2012).

Vulnerability indicators for flash flood hazard are at present too site-specific to render the use of vulnerability assessment broadly operational. Additionally, building structures differ regionally and nationally and channel settings vary locally. The general approach to assess structural vulnerability focuses on impact intensity and structural susceptibility of elements at risk, assigning probabilities to different damage thresholds, from no damage through to complete destruction. From this technical point of view, and as a general rule, vulnerability assessment is based on the evaluation of parameters and factors such as building type, construction materials and techniques, state of maintenance, and presence of protection structures (Fell et al., 2008). For this reason, vulnerability values describe the susceptibility of elements at risk, facing different process types, with various spatial and temporal distributions of hazard intensity (e.g. flow depth, accumulation height, flow velocity and/or pressure, Fuchs et al., 2007a,b; Holub and Fuchs, 2008).

Several recent studies (Martelli, 2011; Santoni, 2011; Ettinger et al., 2014a,b; Thouret et al., 2013, 2014) examined the physical vulnerability of buildings and critical infrastructure in the city of Arequipa, Peru. Thouret et al. (2014) established vulnerability indicators for buildings based on experiences by Zuccaro and Ianniello (2004), Zuccaro et al. (2008) and Zuccaro and De Gregorio (2013) that were calibrated on-site in Arequipa. Our research builds on this work and analyzes the relationships between these parameters and their significance in terms of the susceptibility of a building to experience damage. The present study aims to develop a methodology for rapid estimation of potential damage of existing structures facing natural hazards, in particular flood-hazard. It can be useful, in particular for developing countries, in the case of inadequate hazard information, i.e. in areas where there have been no surveyed hazard events or hydraulic modeling. The objectives of this research are fourfold: to (1) map and characterize channel morphology in natural and built sections; (2) determine and quantify the relationships between building vulnerability parameters; (3) identify the weight of each parameter; and (4) apply mathematical models to calculate the damage probability for each building.

2. Geographical setting

Arequipa, with a population of c. 900,000, is the second largest city in Peru, located at c. 2,300m above sea level, at the foothill of three summits of the Peruvian Andes: El Misti volcano (5,821m asl) to the northeast, flanked by Mounts Chachani (6,075m asl) to the North and Pichu Pichu (5,664m asl) to the East. The high altitude and semi-arid climate are responsible for scarce vegetation cover in both low and high altitudes. Abundant unconsolidated volcanic debris is therefore exposed on steep mountain slopes. Mean annual precipitation does not exceed 150mm and rainfall occurs mainly in the form of low frequency-high intensity rainstorms. These events often trigger flash floods, which sweep through the city of Arequipa following one or more of the numerous channels draining the flanks of El Misti volcano. Since the 1940s, the city's population has quadrupled, occupying at present a built surface of approximately 5,025ha (Fig. 1). The Río Chili valley is a geographical barrier separating the city in two parts; urbanization is extending the city area to the West but also to the East, colonizing the lower flanks of El Misti volcano. Intense urbanization is exposing an increasing number of people and built environment to flash flood hazards.

On 8 February 2013 the La Pampilla meteorological station, located close to the city centre (Fig. 1A), recorded c. 123mm of rain over 3 hours (SENAMHI, 2013); compared with a mean February total of 29.3mm. Since the beginning of pluviometric records in the 1960's, the

February 2013 rainfall was the highest for that month (SENAMHI, 2013; Cacya et al., 2013). The high intensity of this particular rainstorm generated a flash flood that affected several districts of the city (INDECI, 2013; Cacya et al., 2013). Previously conducted risk assessment studies in Arequipa (Martelli, 2011; Thouret et al., 2013, 2014) considered major drainages such as the Río Chili, San Lazaro and Huarangal. However, the 2013 rainfall event rainfall event affected in particular the smaller Avenida Venezuela secondary drainage channel.

Two tributaries drain a c. 7.8km^2 catchment characterized by abundant non-consolidated debris and feed the Avenida Venezuela channel (Fig. 1B): (1) the northern tributary drains watersheds to the Northeast, upstream of the Cooperativa 14 to La Galaxia urban areas; and (2) the southern tributary drains watersheds to the Southeast, upstream of the Mariano Bustamante and Joven Vencedores del Cenepa urban areas. Before joining the main Río Chili valley to the West, the Avenida Venezuela channel crosses the city from NE to SW. Over a total length of 5.2km, the channel depth ranges from 1.3 to 6.3m, with channel widths from 1.63 to 20.64m.

Figure 1A: Geographical setting and location of Arequipa city, Peru. B: The study area Avenida Venezuela channel and six zones that will serve to illustrate observations in the following.

3. Methods

The general methodological approach proposed in this study benefits from data and insights gained from previous exposure and vulnerability assessments carried out in Arequipa (Santoni, 2011; Martelli, 2011; Thouret et al., 2013, 2014). Additional data, in particular, concerning the flood hazard, Avenida Venezuela channel characteristics and surrounding built environment, which had not been studied before, was collected during field work in March 2013 and compiled in an extensive GIS database.

The choice of parameters to be considered for the statistical analysis was motivated by the following reasons:

(i) Information for each parameter, potentially describing vulnerability, needed to be available for all of the individual buildings considered in the study.

(ii) Thouret et al. (2014) observed most vulnerable city blocks to be located within c. 100 m of river channels or in proximity to tributary confluences. Past flood events and flow extension are frequently associated with overbank flow and occasions where bridges acted as obstacles to flow evacuation downstream (Martelli, 2011; Thouret et al., 2014). The distance from the channel and from bridges was therefore considered to be an important parameter to investigate.

(iii) Previous studies (Santoni, 2011; Martelli, 2011; Quan Luna et al., 2011) examined vulnerability at the city block-scale and highlighted the importance of city block shape, building density, and soil impermeability for flow propagation/deviation or velocity, both in downstream direction and laterally.

(iv) The structural type of buildings as well as the number of storeys has been demonstrated as decisive for survival and resistance to flow impact by numerous studies (Papathoma Köhle et al., 2011; Zuccaro and Ianniello, 2004; Zuccaro et al., 2008; Jenkins et al., 2014) and was therefore considered.

(v) The building footprint was included in order to determine its dependency to other building related parameters and susceptibility to damage.

Buildings were selected for sampling as a function of accessibility and willingness of owners to grant access and document damage. Systematically, all accessible buildings in a block (city block) were sampled. Surveyed characteristics regarding building design and environmental characteristics were adapted from previous studies (Thouret et al., 2013, 2014).

3.1 Data collection and processing

As rainfall data was the most reliable information available regarding the origin of the hazard, and too few additional parameters were known to realize numerical simulations of the flood event, this study essentially relies on data acquired from high resolution satellite images, field surveys and GIS (ArcGIS, QGIS) data processing. Pleiades satellite images from 2012 and

2013 at sub-metric resolution were used in ArcGIS to map the channel and built environment affected by the flash flood, both before (2012) and after (2013) the hazardous event.

A field survey was carried out on site a five weeks after the flash flood. The survey was particularly aimed at collecting data regarding inundation height and damage characteristics to buildings, bridges, and training walls caused by the 8 February 2013 flood, but also helped to validate and complete imagery-based measurements and mapping of post-flood channel characteristics (width, depth, wetted section, river bank erosion, etc.). Measuring tape and laser distance meter enabled mapping of bridge opening heights, channel depth and width. Laser measurements were used in particular at sites where channel access from both riversides was not possible (e.g. where building foundations represent the channel wall). GPS (Trimble, handheld) data was simultaneously collected.

Additional data (photography, eyewitness accounts, Civil Defense reports, etc.) were also gathered and compiled in the GIS database. Media images and video footage (professional and social), freely available on the Internet, were invaluable in assessing hazard intensity, flow impacts, damage types, affected sites and deposit types or height. Images taken the day after the event were particularly useful to estimate the immediate aftermath of the inundation. This complimentary data allowed us to monitor impacts in near real-time, identify areas where impact assessments would be most informative and to map the spatial extent of affected areas and occurrence of damage. More than 300 photographs and 15 newspaper articles were studied to extract qualitative and semi-quantitative information regarding damage and flow characteristics. Where flood marks were still visible at the time of our field survey, or where inhabitants were willing to share their experiences, damage level and inundation height were verified along the channel. Run-up measurements could be realized at the tributary confluence in the upstream sector (Fig. 1), at one channel bend in the intermediate section and on a building in the downstream sector. The use of *Chow's (1959)* formula $(v = 2\,g\Delta h)^{0.5}$ or Wigmosta (1983), $(v = 1.2\ g\Delta h)^{0.5}$ enabled estimation of flow velocities based on the measured runup height (in both formulae v is flow velocity, g is gravitational acceleration and Δh the difference in mudline elevation).

Data describing the footprints of buildings were gathered from digital cadastral maps (mostly city block scale), downscaled using the HSR images and cross-checked, where possible, with Google Street View.

For the damage assessment of buildings, survey forms were conceived for masonry and reinforced concrete structures (see Appendices A and B) following experiences from previous studies concerning natural hazard impact (Zuccaro et al., 2008; Zuccaro and De Gregorio, 2013; Jenkins et al., 2014). The survey scheme followed in these forms relies on detailed predefined categories describing different damage intensities and impact types. This procedure was based here on standardized characteristics in order to optimize repeatability of the survey and minimize operator bias.

Once integrated into the GIS database, all surveyed buildings were attributed, in preparation for the statistical data analysis, one of the following four categories describing the observed damage intensity: "1" for no (structural) damage, "2" for light damage, "3" for moderate damage, and "4" for serious damage.

3.2 Statistical data analysis

Building data was statistically analyzed in order to: (i) visualize and quantify relationships between vulnerability parameters; (ii) improve threshold estimates for the different parameter levels; (iii) define the weight of each parameter; (iv) discard or add parameters as a function of their significance; (v) determine significant parameters that are likely to determine whether damage occurs or not; and (vi) calculate a damage probability for each building. Data processing was conducted using R software packages. In order to conduct a statistical analysis on the relationship between the parameters, building data first underwent a selection process to eliminate all elements with one or more unknown parameter and to remove all duplicates. From 1103 buildings surveyed, 898 were therefore extracted for a comparative analysis.

All of the nine considered parameters were initially qualitative, i.e. observational or descriptive (Table 1). Four of them ("Distance from channel" and "Distance from bridge", "Number of storeys” and “Building density in a city block”) were rendered quantitative, i.e. calculated or measured, in order to increase the statistical model performance. Qualitative parameters include the “Inundation height”, the “Soil (im-)permeability”, the “Structural building type” and the “Shape of the city block”. These explanatory parameters were then related to the dependent parameter observed “Damage”. Parameters are either binary (e.g. soil permeability either “permeable” or “impermeable”) or described with up to 5 value categories (levels). An increased number of parameters could not be differentiated as this would have reduced the total sample size of buildings to an extent where the number of cases in each corresponding damage class would have been too low for a robust probabilistic assessment.

Table 1: Vulnerability parameters concerning building characteristics, building environment and flood hazard with their respective levels as defined for this study.

A first step assessed the frequency distribution of buildings for each of the different parameter levels: for example, concerning the parameter “Distance from channel”, the frequency distribution represents the number of buildings which are part of level 1, 2, … or 5. The thresholds delimiting each parameter level were determined in order to respect a minimum of 45 buildings per level (5% of the total building data), necessary to render the level significant. The frequency distribution was then graphically displayed in 2D histograms with the abscissa representing the number of buildings and the ordinate the parameter levels of the examined parameter.

In a second step, correspondence analysis (CA) was conducted in order to bring to light relationships among the different levels of each parameter and among several parameters.

The CA summarizes the relationships between the different parameter levels as scores in contingency tables and enables graphical representation of the latter in several 2D-plots. Each graphical representation is naturally based on two axes (dimensions), each of which expresses a certain percentage of information (inertia) of the contingency table. The dimensions are ranked as a function of their contribution to global inertia (=100%) of the contingency table. In our study, dimensions 1 and 2 have the highest contributions (from 50 to 98.9, with most of that attributed to the first dimension) compared with dimension 3 (< 10).

The CA also provides the coordinates for each parameter level on each dimension. Here, the coordinates are illustrated by (i) small boxes (for results of the simple CA) or (ii) dots (for results of the multiple CA) with dimension 1 being the abscissa and dimension 2 the ordinate.

In graphs plotted for results of the simple CA, two parameters are opposed to each other and each box represents one parameter level. The closer the boxes are, the more similar is the behavior of the buildings that are part of the respective parameter levels.

Graphs illustrating results of the multiple CA are presented as scatter plots and individual buildings are represented as dots. The color of each dot indicates the parameter level (from 1 to 5) that the building belongs to. Ellipses are drawn to help identify buildings that are part of the same parameter level. When the ellipse centers are close to each other, they are strongly related, i.e. the buildings within these groups behave in a similar way. Ellipse centers that are far from each other indicate opposing behavior of the respective group members.

On the basis of the relationships identified between the different parameters by the CA, the final objective was determining the contribution of these parameters to potential building damage. This was achieved using a multinomial logistic regression. Logistic regression was adopted instead of classical linear regression due to the dependent parameter “Damage” being qualitative. As “Damage” levels decline with more than two parameters (level 1 to 4), the logistic regression is referred to as multinomial.

By the use of the following equations, the multinomial dependent parameter “Damage” is related to several other explanatory parameters (e.g. distance to channel, structural type of building, building footprint, etc.). Numerical outputs are probability scores representing the predicted values of damage related to these parameters.

With the hypothesis that explanatory variables are independent, we obtain an additive model, i.e. a model without interactions that is expressed as follows:

$$\log it\left(P\left(Y \leq k \middle| X_1, \cdots, X_j\right)\right) = a_0 + a_1X_1 + \cdots + a_jX_j \qquad \text{with k = 1,…K} \qquad (1)$$

where a_0 is the model constant, Y the dependent parameter (Damage), k the level of the dependent parameter (Damage level), K the highest possible level of damage, j the number of explanatory parameters, X*i, …, Xj* the respective parameter level of each explanatory parameter. The applied logit function is:

$$\log it(p) = \ln\left(\frac{p}{1-p}\right) \qquad (2)$$

For ease of presentation in the following, we set:

$$S = \log it\left(P\left(Y \leq k \middle| X_1, \cdots, X_j\right)\right) \qquad (3)$$

This implies that:

$$P\left(Y \leq k \middle| X_1, \cdots, X_j\right) = \frac{e^S}{1+e^S} \qquad (4)$$

Using the calculated coefficients and parameter level values proper to each building in Eq. (4), we can therefore define the probability of a building to experience damage at damage level k. In order to obtain the damage probability at the precise level 1, 2, 3 or 4, we use

$$P\left(Y = k \middle| X_1, \cdots, X_j\right) = P\left(Y \leq k \middle| X_1, \cdots, X_j\right) - P\left(Y \leq k-1 \middle| X_1, \cdots, X_j\right) \qquad (5)$$

Finally, based on the Bayesian Information Criterion (BIC), the optimal model is selected. The BIC enables elimination of non-significant parameters and reduces the model to the significant parameters determining damage.

3.3 Model validation

Formal model validation was realized in two steps. First, using a calibration and validation data set. Hereby, the original data set was split into a calibration and a validation data set. From the 898 totally sampled buildings, the validation data set contained 300 arbitrarily selected buildings. The model was calibrated using the 598 remaining buildings and then run to test the validation data set. Calculated damage probabilities obtained by the model for buildings in the validation data set was then compared to observed damage in the field.

Second, based on the results of the validation data set, the “good classification rate” was determined. This rate describes the performance of the model indicating the percentage of buildings in which predicted damage corresponds to observed damage.

4. Results

4.1 Channel characteristics

Three main cross section types were observed: (1) unconfined, natural; (2) confined, both sides; and (3) confined, unilateral. Type 1 is typical in upstream and downstream channel segments with channel widths from 5 to 16m, characterized by predominantly natural channel bed and walls, as well as frequent terraces along either left or right channel walls. Type 2 (confined, concrete) is characterized by narrow and straight channel sections, especially in the intermediate sections. Type 3 generally corresponds to the largest channel widths (> 15m) and is transitional in character between types 1 and 2, i.e. semi-natural. In confined sections, either concrete (reinforced or not), or mixed material (volcanic rocks, mainly andesite or ignimbrite, brick, metal tubes, etc.) are used to stabilize channel banks (table 2). On both channel sides, more than 70% of counted sections have their start or end point within 15m of a bridge. Of note is that only 23 of a total of 181 channel sections distinguished along the Venezuela drainage are still natural and, of which, the majority are located on the right riverside (in downstream direction). This corresponds to a length of c. 1.5km out of 13.8km in total (table 3). Mixed material bank stabilization is shown in 48% of total sections, employed more frequently than concrete constructions (47%) but on shorter section lengths (table 3). While concrete constructions extend over c. 58% of total section length, mixed material reaches approximately 23%. This is a result of recent channel confinement work especially in the intermediate section where major road works were ongoing in 2013.

Table 2: Material types characterizing channel banks along the left and right riverside of the Avenida Venezuela channel. Numbers indicate the frequency distribution of channel sections in each material category and their respective distance to a bridge. The total number of sections located at a certain distance from a bridge is illustrated in bold with, to its right, the corresponding percentage of total channel bank length.

Table 3: Section lengths of channel banks as a function of construction material and location on left and right riverside.

The mean slope of the channel from its upstream confluence to joining the Río Chili downstream is 12.54% on a recently calculated 5m-DEM based on high resolution satellite imagery (Pléiades-data) versus 4.67% on a previously utilised 30m-DEM derived from SPOT5 images. Channel width ranges from a minimum of 1.63m to a maximum of 20.64m. Large channel widths are mostly tied to a natural bed type (gravel, sand; Fig. 2, dark gray color), while narrow reaches correspond to confined concrete channel sections (Fig. 2, medium gray). Between the shopping mall La Negrita and the Villa Militar Salaverry area (Fig. 1B), flow velocities for the February 2013 event could be estimated, based on run-up measurements, to be between 7.6 and 10.9m/s (Chow, 1959) and between 5.9 and 8.5m/s (Wigmosta, 1983), respectively.

Figure 2: Longitudinal channel profile (black line) with channel width (green dots), and sections in which erosion occurred (orange bars). The gray scale bar represents channel bed type, i.e. natural gravel, sand (dark gray), natural with occasional concrete steps (white) and concrete (medium gray); the channel wall material is represented by concrete (red), mixed

material (concrete, brick, boulders; yellow) and natural (blue). For complementary information see also table 2 to 4.

Peak flow discharge for the February 2013 event was estimated using channel cross section measurements and average flow velocity to be 123.4m^3s^{-1} in the upstream reach and c. 425m^3s^{-1} in the middle reach (in proximity to the La Negrita shopping mall, fig. 1B). A clear delineation of areas inundated by intense surface runoff or by overbank flow of the Venezuela channel was not possible in many parts along the channel, which prevents a more precise estimate of the flow volume. An attempt at mapping the inundation extent resulting from overbank flow was made on the basis of field survey, including observations of erosion marks along the channel, flood marks on built infrastructure and eye witness accounts (Fig. 3). Comparing this map with flow simulations published in previous studies (Martelli, 2011; Oehler et al., 2014), the flow volume of the February 2013 flash flood can be estimated between 50,000 and 100,000m^3. There appeared to be no general rule to where erosion occurred; natural bed types were affected to the same extent as concrete beds. Concerning the channel wall material, concrete and mixed material sections seemed to be affected more often by erosion than natural sections. For concrete sections, erosion is most likely at channel contractions or expansions or where bed or channel wall materials change. When erosion occurs in natural sections, the proximity to bridges with low opening height frequently determines whether erosion occurs or not.

Figure 3: Field-survey-based mapping of inundation extent resulting from overbank flow along the Avenida Venezuela channel.

For a flash flood event of relatively small volume such as the one of February 2013, obstacles such as bridges play a major role in terms of flow propagation, extent of inundated area, and the type and intensity of damage. Impact forces of this particular flood were strong enough to completely erode one pedestrian bridge at the upstream border of the Palomar market (Figs.

2 and 5). In the case of 17 other damaged bridges along the Venezuela channel, openings were either not large enough for an increased discharge or were obstructed by boulders. In both cases, the consequence was similar: the flow front overtopping the bridge and caused overbank flow to both sides of the obstacle (Fig. 4, zone 1 and zone 5-6). Additionally, partially confined sections with an abrupt change in channel direction (e.g. close to 90° angle) were particularly prone to overbank flow (Fig. 4, Zone 1).

Figure 4: Three examples of particular channel courses and resulting damage.

4.2 Flood impact on channel banks

Generally, an increase in the damage degree was observed from upstream to downstream reaches with 64% of right channel banks (in downstream direction) affected. Plotting erosion and bridge location on the longitudinal channel profile shows that there seems to be a positive relationship between the presence of a bridge and the occurrence of erosion in its proximity (Fig. 5, table 4). Field observations suggest that erosion in the immediate downstream region of bridges is more likely than in upstream parts. However, especially in the intermediate channel reach, the distance between bridges is so close that it is difficult to determine whether erosion is a consequence of flow turbulence immediately upstream or downstream of a bridge. Concerning the channel width, erosion was frequently related to a change from narrow to wide channel sections (Fig. 2, 5). This is likely linked to increasing flow turbulence at the transition from harder trained confined sections to unconfined sections and a decrease in flow velocity with lateral flow expansion.

Material types of training walls were regrouped into 5 major classes (Fig. 5): (1) concrete (reinforced); (2) rock piles; (3) gabion meshes; (4) mixed material or (5) natural banks. While rock piles appear to be affected with similar proportions in all damage categories, concrete dominates the category with the heaviest observed damage (Fig. 6). Mixed material channel walls are present in light damage categories (1 and 2) as are gabions. The latter represent a small percentage in damage category 3. When relating the spatial distribution of damaged material types to the location of bridges, as expected, damage occurs preferentially within 100m of a bridge. Only natural channel banks and mixed material banks were damaged at greater distances (up to 200m). This observation also confirms that erosion preferentially occurred downstream of bridges rather than upstream.

Figure 5 Left: Damage level observed for different material types of retaining walls. Right: Material types of retaining walls relative to the proximity of bridges.

Table 4: Damaged channel bank sections represented as the percentage of the total length of either the left or right channel side. Sections are attributed to one of six groups (A to F) depending on the closest distance to a bridge of either the start or end point of the section.

4.3 Flood impact on buildings

Inundation height could be measured at almost 300 sites and flood marks along building walls indicated minimum heights of 0.2m and maxima of 0.7m. While 611 sampled buildings were not affected by the flood (68%), 287 buildings were inundated, among which 11% were below and 21% above a water level of 0.2m.

Four damage levels were defined for the building stock (Fig. 6):
(1) inundated, without any structural damage; (2) inundated, light damage, building still fit for habitation after cleaning; (3) inundated, significant damage, rooms livable only after refurbishment; (4) inundated, heavy damage, structural refitting required. Among the 287 inundated buildings, 144 experienced significant to heavy damage, and 143 buildings were slightly damaged.

Figure 6: Observed damage levels from left to right (4) inundated, heavy damage, (3) inundated, significant damage; (2) inundated, light damage; and (1) inundated, no structural damage.

4.4 Linking vulnerability parameters to observed damage

Each vulnerability parameter was plotted separately on the basis of the respective contingency table containing the number of levels (Table 1), the number of buildings at each level and their distribution frequency (Fig. 7). Along the channel, 30 randomly selected city blocks of variable size (from 531m^2 to 10.57ha), and with rather compact and regular shape, were studied. They contained 1103 buildings with footprints ranging from 4 to 2,185m^2; the majority of buildings of commercial, industrial or agricultural use were larger than 80 m^2 and grouped in size categories 4 and 5, while primarily residential buildings represented about 45% of all those analyzed. Building density per hectare ranged from 3 to more than 11. However, the majority of the sampled city blocks were characterized by a relatively low building density (< 6 buildings per hectare), which was the result of the relatively large footprints of non-residential buildings.

Figure 7: Results of univariate analysis summarizing the number of buildings per category. Grayscales from the lowest parameter level 1 (white) to the highest level 5 (dark gray) are the same for all figures.

In terms of the presence or absence of relationships, the results of the correspondence analysis show that couples incorporating the parameter "Damage" have very strong relationships (Fig. 8). Graphic plots show proximities of parameter levels (Fig. 8, DBR5 and DO4 in the green circle) and oppositions to other levels (Fig. 8, DBR1, DBR2 and DO1 in the blue circle). For the presented example in figure 9, this data projection suggests that buildings more than 50m from a bridge behave in a similar way and are less exposed to experiencing damage than buildings within 5m of a bridge. In our data set, 69.1% of all buildings less than 15m and 39.6% between 15 and 30m from a bridge were damaged.

Figure 8: *Plot of parameter couple "Distance from bridge" and "Damage" at respective levels.* Note the strong relationship between buildings located close to a bridge (DBR5) and damage level 4 (DO4; right side of vertical axis) compared to buildings far from a bridge (DBR1 and 2) that have damage level 2 (DO2, left side of vertical axis). Eigenvalues represent 96.54% for axis 1, 2.76% for axis 2, and 0.7% for axis 3.

While some relationships are expected and strong (e.g. damage versus inundation), others have a weak relationship (damage versus number of storeys or city block shape) or are a direct consequence of the characteristics of the data set (damage versus soil impermeability). Some particularly interesting observations are outlined in the following (Fig. 9):

- Damage versus distance from channel: inundated buildings without damage dominate at distances beyond 60m of the channel. The closer to the channel, the more intense the damage, e.g. while at distances of 5 to 10m, damage categories 2 to 4 are almost equally present, damage category 4 becomes most important at distances lower than 5 m.
- Damage versus distance from bridge: similar to the previous observations, inundation without damage occurs preferentially at distances > 90m from bridges. Slight and significant damage also appear at this distance. Heavy damage is significantly less important beyond 90m, but still present. Damage category 4 dominates, however, in distances up to 30m from a bridge.
- Damage versus structural type: overall, for the data set studied here, damage categories 1, 2 and 3 mostly occur with buildings of type 3 (masonry of terra cotta with reinforced concrete roof, 1 or 2 storeys). The heaviest damage is preferentially observed in non-residential buildings.
- Damage versus inundation height: slight or significant damage is the main consequence at intermediate inundation heights while maximum inundation is related to the highest damage.

- Damage versus building footprint: generally, damage intensity appears to be independent of building size, i.e. all damage levels have been observed for buildings larger than 20m^2. Only footprint category 4 (80 – 150m^2) experienced more damage than other groups, particularly in the intermediate damage level (2 and 3).
- Damage versus building density: city blocks of low building density are more likely to suffer damage. That is, at building densities of less than 6 buildings per hectare (category 4 and 5), buildings were more often damaged than in city blocks exhibiting high building density. City blocks of higher densities were more often affected by inundation without damage.

Figure 9: Results of the bivariate analysis. Damage level is displayed in different gray *shades, the abscissa (1 to 5) displays the categories of the respective parameter “Distance from channel”, “Distance from bridge”, etc.*

These observations are hypotheses essentially based on graphically plotting the results for all parameter couples of the correspondence analysis (Fig. 8). In order to confirm or reject these hypotheses, it is therefore necessary to further verify using contingency tables and plots from the multiple correspondence analysis (Fig. 9 and 10).

While simple correspondence analysis examines the relationship between two parameters, multiple correspondence analysis generalizes the comparison by including as many qualitative parameters as available. Scatter plots in this context graphically represent the relationships between the different levels of each parameter (Fig. 10); they enable the comparison between individual buildings, their position among others at certain parameter levels to be determined and finally reveal behavioral tendencies of building groups showing similar characteristics. The position of each parameter level is determined by a bagplot (bivariate boxplot), displaying an ellipse representing 67.5% of buildings within the respective level. Some parameters such as “Inundation height”, “Impermeability of soil” and “Structural building type” display good sorting, i.e. buildings of the same characteristic are well grouped. For other parameters, the sorting is much less evident, dispersion is high and overlapping ellipses representing the weighted center of each parameter group are the result. Hence, one can recognize similarities in the distribution of parameter levels (from 1 to 5, light blue and red, respectively in Fig. 10): buildings of level 4 and 5 are preferentially located to the right side, while levels 1 and 2 remain close to the central axis with a tendency to the left side. Distributions of parameter levels concerning “distance from channel”, “distance from bridge” and “density of buildings” are observed to be very close; the same can be seen for the parameters “damage category” and ‘inundation height”. For these two latter parameters, the scatter plots are very similar. Their relationship is therefore direct as illustrated previously by the bivariate analysis: the higher the inundation height, the higher the damage level. Generally, the scatter plots illustrate that buildings located in the vicinity of the channel and / or a bridge tend to be 1-storey constructions of structural type 3 to 5. They are commonly located in areas of low building density (category 4 or 5). The parameters “shape of city block”, “impermeability of soil” and “building footprint” are not directly related to the previous groups as buildings of varying characteristics, i.e. several parameter levels, are present on the right side of the plot. The distribution pattern of building points follows two general tendencies: firstly, a horse-shoe shape where the distribution of buildings representing a particular parameter level produces an arch pattern spanning from level 1 (light blue; left of vertical axis) to level 5 (red; right of vertical axis); secondly, a random distribution to vertical clustering of all with level 4 and 5 (orange and red, respectively) mostly located in the opposed direction of level 1 to 3 (light blue, dark blue, green; fig. 10).

Figure 10: Scatter plots representing results of the correspondence analysis. Each point represents a building. Ellipses colored from light blue to red represent parameter levels (1 to 5, respectively) as bagplots (bivariate boxplots). Each bagplot represents 67.5% of the buildings defining each level.

Plotting all parameter levels along with the number of city blocks enables us to easily relate different levels to each other. As for the scatterplots, the position of the parameter level is defined by the bagplot with 67.5% of the building points defining each level. Again, it becomes clear that buildings that have experienced the highest inundation are localized very close to bridges and have also suffered the most damage; this is particularly within city block n°20 (Fig. 11). This group of buildings contrasts with those that did not experience any damage and those that were inundated temporarily without experiencing damage. The latter were located both far away from the channel and any bridge (Fig. 11, city block 3).

Figure 11: Projection of parameter levels (color) and city blocks as a result of the bivariate analysis. The position of each square is defined by the bagplot representing 67.5% of the buildings defining each level. The circles indicate city blocks of similar characteristics and thus behavior. City block numbers are plotted to allow comparison but are not included in the bivariate analysis. The number of buildings per city block and the respective percentage is detailed in the histogram to the right.

Logistic regression

Logistic regression was applied in order to directly analyze the link between the qualitative parameter "Damage" and one or more other parameters and to calculate damage probabilities for all buildings. For this part of the analysis, the parameter "Inundation height" was not considered, as it was the only data measured after the event and hence strongly related to the observed damage, which is the parameter to explain. The eight remaining parameters were: Distance to channel, distance to bridge, shape of city block, impermeability of soil, number of stories, structural type of building, building footprint and building density. Isolating the most significant parameters that determine the damage likelihood progressively reduces the number of vulnerability parameters from 8 to 5. By eliminating non-significant parameters, the contribution of the maintained parameters is constantly recalculated so that the remaining significant ones also reflect the non-significant parameters. Consequently, the "Damage" parameter can be expressed as a function of its relationships with the parameters “Distance from channel”, “Distance from bridge”, “Shape of city block”, “Structural building type” and “Building footprint”. To illustrate the different steps of calculation and associated model outputs, we chose one of 200 scenarios that were realized using different calibration and validation data sets. Equations (1) and (2) enable us to obtain the respective contributions from the different parameters (table 5).

Table 5: Contributions of each parameter level to damage probability based on the calibration data set (598 buildings).

The following example illustrates the way in which the values of table 5 were used to calculate damage probability for a specific building at a particular damage level.

The applied logistic regression model with all remaining significant parameters is presented as follows:

$$S = \begin{matrix} 1|2 & 4.59 \\ 2|3 & 5.99 \\ 3|4 & 7.30 \end{matrix} + \begin{matrix} DC1 & 1.25 \\ DC2 & 0.57 \\ DC3 & 0.20 \\ DC4 & 1.36 \\ DC5 & 0.27 \end{matrix} + \begin{matrix} DBR1 & 4.79 \\ DBR2 & 0.03 \\ DBR3 & 1.02 \\ DBR4 & 1.51 \\ DBR5 & 2.22 \end{matrix} + \begin{matrix} SH2 & 1.03 \\ SH3 & 1.42 \\ SH4 & 0.38 \end{matrix} + \begin{matrix} S2 & 2.52 \\ S3 & 0.08 \\ S4 & 1.48 \\ S5 & 0.96 \end{matrix} + \begin{matrix} A1 & 10.33 \\ A2 & 1.99 \\ A3 & 2.91 \\ A4 & 2.62 \\ A5 & 2.82 \end{matrix} \quad (5)$$

Values in this equation (5) were taken from table 5 and completed following the logistic regression constraint that the sum of coefficients of each parameter must be equal to zero. That is, parameter “Distance to channel” has been completed by DC1, parameter “Distance to bridge” by DBR1 and parameter “Building footprint” by A1.

We consider for example a building with the following parameter levels: Distance to channel = 4, Distance to bridge = 5, Shape of city block = 4, Structural building type = 5, building footprint = 5. In order to obtain the probability of experiencing damage (e.g. at level 3) for this building, we fill in the previous equation using results of table 5 as follows:

S = 5.99 - 1.36 – 0.27 + 0.38 - 0.96 - 2.82 = 0.96 (6)

The damage probability for this specific building at intensity level lower or equal to 3 can then be calculated using equation (3) and the value obtained for S in equation (6):

$$P(Y \quad 3) = \frac{e^{0.96}}{1+e^{0.96}} \qquad (7)$$

This particular building therefore has a probability of 72.3% (Eq. 7) of experiencing damage at level 3 (significant damage) or lower in the case of a future flood event of similar intensity to the one of February 2013. In order to calculate the probability for a different damage category, one has to change the intercept constant and proceed in the same way.

For the 300 buildings in the validation data set for this scenario, the probability to experience damage at different intensity levels was calculated (table 5). With a success rate of 74%, the model performs well predicting for almost three quarters of sampled buildings a damage probability that corresponds to field observations.

As a result, 27.7% of buildings of the validation data set have a 100% chance of experiencing damage of level 2 (slight damage) or more for an event similar to February 2013. For damage levels 3 and 4, 6.3 and 9% of buildings, respectively, have a 100% chance of being damaged. Various maps could then be drawn representing each building in the data set, whether sampled in the field or not, and their probability of experiencing damage at a certain level (Fig. 12). A comparison of these results with maps illustrating damage levels assessed during field work correlates well: the distribution of the damage probability is overall coherent with field observations of damaged buildings and of the measured flood extent. Results for both calibration and validation data sets are close, not exceeding a difference of 5% for the number of buildings attributed to each damage level. The city block examples presented for the six areas within the study (Fig. 12) enable us to highlight several main points: (1) calculating damage probability by incorporating field data enables expansion of the analysis to buildings not sampled in the field or for which only some of the required parameters were identified; (2) calculated results for the validation data classify a lower number of buildings (c. 15%) in the predicted damage category “2 or more”, which is equivalent to the observed damage level 2, 3 and 4, than observed in reality (28%). This is mainly due to the fact that in reality, parameters not considered in our analysis also seem to be important; these include the height of channel retaining walls (Fig. 12, zone 2), the presence of increased surface run-off not coming from the channel (Fig. 12, zone 4), the height and width of bridge openings, the building position upstream or downstream from a bridge, etc.; (3) especially for buildings not affected in 2013 (cf. Fig. 12, zone 2, 5 and 6), the calculation of damage probability allows us to identify and visualize graphically areas where mitigation measures such as refitting may be necessary to avoid serious damage during future events; and (4), although rare, some buildings with significant damage observed in the field (Fig. 12, zone 1 and 5) appear to have lower calculated damage probabilities than expected.

The absence of a logical pattern in the distribution of these damage probabilities reflects the fact that several parameters were considered, each parameter with an individual weight. This considerably improves the damage assessment compared to methods where simple buffer zones along the channel borders are used to determine vulnerability, essentially as a function of distance to the channel.

Figure 12: Damage probabilities calculated for damage levels 1,2, 3 and 4 (series A) and all observed damage levels (1 to 4) in the field (series B) using calibration and validation data sets (898 buildings) of the selected test scenario presented in the manuscript.

Finally, concerning the model performance, one scenario with a success rate of 74% was chosen to illustrate the methodology. The model was applied in total 200 times on different calibration and validation data sets in order to examine general performance. Results show that 90% of these tests have a success rate of more than 67%.

5. Discussion

A quantitative analysis of the uncertainty of results has been beyond the focus of this study. However, it is important to be aware that both data and applied methods introduce uncertainties. Major sources of uncertainty stem from imprecise or ambiguous measurements (both in the field and from remote acquisition). This may be, for example, a consequence of surveyed data being based on standardized characteristics or deformation due to spatial georeferencing etc. Certain parameters related to the structural behavior of buildings to flow impact, varying impact forces as a function of the impact angle and the grain size of the sediment in the flow, etc. may also not have been assessed in sufficient detail. Another source of uncertainty may stem from the selection of vulnerability parameters taken into account for the mathematical analysis. The fact that we considered a minimum of 5% of the total sample size of buildings needed to be part of a parameter level in order to make it eligible for the statistical analysis was a constraint for data sets where little information was available and adjustments were not possible.

For this study however, we could reduce the uncertainty related to the latter aspect to a minimum by limiting the conditions for parameters to the following: (1) contain a minimum number of buildings per parameter level to be significant (5% of total sample number); (2) data either existent from field work or able to be collected from calculation, satellite data or other sources; and (3) represent a single piece of information in order to avoid repetition or overlapping. Parameters that did not account for these conditions were reconsidered. For example, the parameter “Structural type”, originally intended to represent different building characteristics such as construction material, the type of roof, and the number of storeys. To make this parameter eligible, “Number of storeys” was extracted as a separate parameter. The ”Type of roof” was not identified for enough buildings in order to be considered and was therefore eliminated from the analysis. “Structural type” therefore only represents the construction material in this context.

Parameters containing similar information, such as “distance from channel” and “distance from bridge” were voluntarily kept in order to obtain elements of response for particular questions. In reality, a building far from the channel cannot be close to a bridge, while the inverse is possible. Despite a certain interdependency between these parameters, keeping both helps to better estimate the role of bridges in terms of damage probability, which would not be possible if the distance from a bridge was not considered individually.

Interpretation of the results requires careful comparison of the links between several parameters, particularly if no previous knowledge from real field conditions is available. A few aspects brought to light by scatter plots and logistic regression are pointed in the following.

Scatter plots

- In this study, building density was calculated as the number of buildings per unit area. High building density in Arequipa generally corresponds to dense habitat where buildings share one or more walls, reducing the space between them and thus reducing the risk of inundation in back rows. Lower building density, in particular for residential buildings, implies gaps between buildings, which creates hydraulic roughness and resistance to flow. This reduces flow velocities on buildings in back rows, but potentially increases them in front rows. This fits to observations from the scatter plots showing that city blocks of lower density exhibit more damaged buildings than those of higher density. However, field observations suggest another reason for increased damage in low-density areas: larger footprints of industrial, commercial and agricultural buildings that, especially for the latter, have more vulnerable structural characteristics than residential buildings. To clarify this possibility may require improvements in how building density is assessed in relation to the structural building type and/or its use.

- Comparing building footprint and density of buildings, it appears that the largest buildings (mainly agricultural or industrial, and more rarely commercial) are: (i) more vulnerable than smaller buildings because of their very larger openings, and fewer load-carrying structures such as columns, etc.; and (ii) typically located in city blocks with low building density, which additionally increases the probability of being damaged since the screening effect from adjacent buildings is absent. For footprint categories 1 and 2 (small size), the building density per city block frequently remains low enough to potentially put them in damage category 4.
- Relating damage level 3 (significant damage) to the structural type of buildings, it becomes obvious that buildings of all types can be damaged, but in this flood event, mostly buildings of category 2 (masonry of terra cotta or ignimbrite with mortar and metal sheet roof, 1 storey) and 3 (masonry of terra cotta with reinforced concrete roof, 1 or 2 storeys) were affected; this coincides also with the highest inundation measured and with the smallest distance from the channel and bridges. However, damage categories 2 and 3 group buildings of structural types 6, 7, 8 and 4, 5, respectively. These are all structural types of higher quality, which are estimated to be less vulnerable than structural types 1, 2, 3 grouped in damage category 4. Previous studies regarding physical vulnerability of buildings related to landslide and debris flow hazard in Austria (Fuchs et al., 2007; Papathoma Köhle et al., 2012), Germany (Kaynia et al., 2008), Italy (Aleotti et al., 2004; Luino, 2005; Galli and Guzzetti, 2007) and the United States (FEMA: HAZUS-MH, 2010), seem to confirm an increasing vulnerability with decreasing construction quality (either due to construction material or structural characteristics). We therefore deduce for our results that the relationship statistically calculated to be strong between structural type and damage is, in this particular case,

the result of the strong influence of the 611 buildings that have not experienced any damage.

- “Inundation” versus “number of storeys” shows that buildings of 2 and 3 storeys behave in the same way and have been affected similarly, inundated but less severely damaged than buildings of 1 storey – this is partly due to the fact that the 1-storey-buildings are more frequently close to the channel in the building sample considered for this study. In this context, another parameter has been observed in other studies to be important (Fuchs et al., 2007; Lo et al., 2012; Papathoma Köhle et al., 2012): the presence of windows and other openings that allow material to enter the building may also affect the degree of damage experienced by a building. While for some buildings, such information has been collected in Arequipa, it was not available for enough sampled buildings to be considered for this analysis. Further research should take this aspect into account.

Logistic regression

Logistic regression in addition to the scatter plots, reflects well the issue of an unbalanced data set: buildings without any damage are largely over-represented in our data. In the applied mathematical analysis, this group gains in weight because of their high number leading to calculated damage likelihoods that are greater than those derived from expert judgment. For example, the logistic regression analysis favors buildings without damage as a way to rank buildings of higher quality (structural types 4 to 7) to be more likely to experience damage level 3 than buildings supposedly more vulnerable (structural types 1-3). This should improve by incorporating additional damage data from future potential flood events, which will lead to a more balanced data set.

Scope and limits

In the field following the February 2013 event, damage was observed at buildings that were identified in the calculations to have a 10-20% probability of significant damage (intensity 3).

On the maps (Fig. 12), this is illustrated by the color code representing the probability to experience a certain damage level for each building. Hence, for buildings colored red, the probability of experiencing heavy damage is higher than 30%, while a dark green building has a more than 30% chance of not experiencing structural damage. This may seem underestimated, however, as when taking damage levels 3 and 4 together, the likelihood to be seriously affected reaches 57%, which is rather considerable. At the same time, 57% of significant to heavy damage implies roughly 43% of slight damage. Since the ultimate goal of risk assessment studies is to avoid the occurrence of future damage, it is therefore important to interpret the damage probabilities with the damage intensity scale in mind. A first interpretation should therefore examine the probability of experiencing structural damage (i.e. damage level 2 to 4) versus no structural damage (damage level 1). And then, in a subsequent step, differentiate the probability of structural damage for informing local risk mitigation strategies. This is all the more important given that this study is considering damage potential for a relatively small flood event. Higher damage probabilities for flood events of larger volumes are therefore to be expected. The comparison of calculated damage probabilities and observed damage emphasizes the potential for using damage probabilities to identify areas with special need for mitigation measures in order to avoid future damage. It also enables us to understand that local parameters, especially those related to channel morphology or topography, such as the height of retaining walls, the direction of the channel course (Fig. 12, zone 2) or local slope, that have not yet been considered in this approach, also influence the damage likelihood.

6. Conclusion and perspectives

This study was carried out following a damage survey in the field early after a flash flood event in February 2013. Observed damage intensities were overall low and only few buildings suffered serious damage. Results from the proposed statistical data analysis validate the method as an operational tool to calculate damage probabilities for a flash flood event of similar intensity. However, the lack of damage documentation, in particular for highest damage categories, is at present a constraint to further develop the model for a larger range of hazard types and magnitudes. Ideally, the method will be tested on a database including many case studies. This would allow validation of the methodology not only for small scale, site-specific analysis, but also for broad scale generalized assessment of damage probability for different hazard types and magnitudes.

In the close future, several possibilities should be statistically explored:

(1) Although an initial analysis of the interactions between parameter couples did not render significant results, parameter interdependency needs further examination in order to be accounted for fully in equations calculating damage probability. This aspect may become more important in the context of a possible extension of the data set in the future. Given additional flood events and new damage data, vulnerability parameters or thresholds of parameter levels may need to be adjusted. This possibly implies an increase in either quantitative or qualitative parameters, in which case the method requires interactions to be accounted for in order to guarantee a similar degree of effectiveness.

(2) Analysis of the present data enabled a critical evaluation of parameters considered for vulnerability; however, some of the parameters originally identified were not considered in analysis because of partially missing data and/or an insignificant number of buildings in each parameter level. Other parameters have emerged and should be taken into account, such as the location – upstream or downstream – of the closest bridge, a more detailed series of structural building types for non-residential constructions etc.

(3) Given the relatively small size of the event, damage categories were established accordingly in order to ensure a maximum of damage data being recorded; for a larger event, damage categories may need to be defined differently. In order to take into account different event scenarios for the flood, the method needs to evolve towards an additive or cumulative approach. This is important in the case of a building hit by different flood volumes, which implies repetitive measurements of a single data point, an independent analysis of the parameters determining its vulnerability is no longer possible.

(4) If the dependent variable is qualitative, the lognormal law may need to be considered instead of the log logistic approach.

One of the major advantages of the method outlined is that damage probability can be estimated and mapped even for buildings that have not been sampled in the field as long as some of their characteristics are known or are able to be assessed from remote sensing data. Especially in a context where few damage data are available or where access to the field in the aftermath of an event is difficult, this technique helps to assess and project damage potential to non-sampled areas. This is as useful for both loss estimation and risk prevention, by contributing to the planning of mitigation measures such as refitting or risk management, or by evacuation planning in the case of disaster.

Acknowledgements

The authors would like to thank the Municipality of Arequipa, the Arequipa Civil Defence and INGEMMET Arequipa and Lima for providing logistic and technical support during field work. This research was financed by the French Government Laboratory of Excellence initiative n°ANR- 10-LABX-0006, the Région Auvergne and the European Regional Development Fund. This is Laboratory of Excellence ClerVolc contribution number 140.

References

Aleotti, P., Ceriani, M., Fossati, D., Polloni, G., Pozza, F., 2004. I fenomeni franosi connesi alle precipitazioni del Novembre 2002 in Valtellina (Soil slips and debris flows triggered by the november 2002 storms in Valtellina, Italian Central Alps), International Symposium Interpraevent 2004, Riva/Trient, Italy, 11–20, 2004.

Birkmann, J., 2006. Measuring vulnerability to promote disaster resilient societies. Conceptual frameworks and definitions, in: Birkmann, J. (Ed.), Measuring vulnerability to natural hazards—towards disaster resilient societies, United Nations University Press, Tokyo, pp 9–54.

Borga, M., Gaume, E., Creutin, J.D., Marchi, L., 2008. Surveying flash floods: gauging the ungauged extremes. Hydrological Processes 22(18), 3883-3885.

Bründl, M., Romang, H.E., Bischof, N., Rheinberger, C.M., 2009. The risk concept and its application in natural hazard risk management in Switzerland. Natural Hazards and Earth System Science 9, 801-813.

Cacya, L., Meza, P., Carlotto, V., Mamani, L., 2013. Aluvión del 8 de Febrero del 2013 en la ciudad de Arequipa. Foro internacional Peligros geológicos, Arequipa, 14-16 Octubre 2013, GRM, OVI Ingemmet, Libro de Resúmenes, pp. 195-203.

Cardona, O.D., 2004. The Need for Rethinking of the Concepts of Vulnerability and Risk from a Holistic Perspective. A Necessary Review and Criticism for Effective Risk Management, in: Bankoff, G., Frerks, G., Hilhorst, D. (Eds.), Mapping Vulnerability, Disasters, Development and People, Earthscan Publishers, London.

Carreno, M.-L., Cardona, O.D. and Barbat, A. (2006) Urban seismic risk evaluation: a holistic approach. Natural Hazards, 40, 137-172.

Chow, V.T., 1959. Open Channel Hydraulics. McGraw-Hill, New York, 680 p.

Collier, C., 2007. Flash flood forecasting : what are the limits of predictability ? Quarterly Journal of the Royal Meteorological Society 133(622A), 3-23.

Crichton, D., 1999. The Risk Triangle, in: Ingleton, J., (Ed.), Natural Disaster Management, London, Tudor Rose, pp. 102–103.

Cutter, S.L., Boruff, B.J., Shirley, W.L., 2003. Social Vulnerability to Environmental Hazards. Social Science Quarterly 84(2), 242–261

Di Baldassarre, G., Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. Hydrology and Earth System Sciences 13(6), 913-921.

Douglas, J., 2007. Physical vulnerability modeling in natural hazard risk assessment. Natural Hazards and Earth System Science 7, 283-288.

Ettinger, S., Manrique Llerena, N., Talibart, C., Mounaud, L., Yao-Lafourcade, A.-F., Thouret, J.-C., 2014a. Using HSR imagery to downscale vulnerability assessment of buildings and local infrastructure facing hazards from floods and hyperconcentrated flows. Geophysical Research Abstracts, vol. 16, EGU2014-13978, EGU General Assembly 2014.

Ettinger, S., Zeghdoudi, M., Manrique Llerena, N., Yao-Lafourcade, A.-F., Thouret, J.-C., 2014b. L'apport de l'imagérie Pléiades à la cartographie des enjeux et de leur vulnérabilité face aux crues torrentielles : La ville d'Arequipa, Pérou. Revue Française de Photogrammétrie et de Télédétection. In Press, 7 p.

Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., Savage, W.Z., 2008. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. Engineering Geology 102, 85-98.

Fuchs, S., Heiss, K., Hübl, J., 2007a. Towards an empirical vulnerability function for use in debris flow risk assessment. Natural Hazards and Earth System Sciences 7, 495-506.

Fuchs, S., Thöni, M., McAlpin, M.C., Gruber, U., Bründl, M., 2007b. Avalanche hazard mitigation strategies assessed by cost effectiveness analysis and cost benefit analysis – evidence from Davos, Switzerland. Natural Hazards 41, 113-129.

Galli, M., Guzzetti, F., 2007. Landslide Vulnerability Criteria: A Case Study from Umbria, Central Italy, Environ. Manage. 40(4), 649–665.

Gaume, E., Bain, V., Bernardara, P., Newinger, O., Barbuc, M., Bateman, A., Blaškovičová, L., Blöschl, G., Borga, M., Dumitrescu, A., Daliakopoulos, I., Garcia, J., Irimescu, A., Kohnova, S., Koutroulis, A., Marchi, L., Matreata, S., Medina, V., Preciso, E., Sempere-Torres, D., Stancalie, G., Szolgay, J., Tsanis, I., Velasco, D., Viglione, A., 2009. A compilation of data on European flash floods. Journal of Hydrology 367 (1-2), 70-78.

Glade, T., 2003. Vulnerability assessment in landslide risk analysis. Die Erde 134, 123–146.

Graf, W., 1988. Fluvial processes in dryland rivers. Springer, Berlin, 329 p.

Gruntfest, E., Handmer, J., 2001. Dealing with flash floods: contemporary issues and futures possibilities, in: Gruntfest, E., Handmer, J. (Eds.), Coping with Flash Floods. Kluwer Academic Publishers, Dordrecht, pp. 3-10.

Gruntfest, E., 2009. Editorial. Journal of Flood Risk Management 2, 83-84.

HAZUS-MH, 2010. HAZUS – FEMAs Methodology for Estimating Potential Losses from Disasters, available at: http://msc.fema.gov/portal/resources/hazus.

Hiete, M., Merz, M., 2009. An Indicator Framework to Assess the Vulnerability of Industrial Sectors against Indirect Disaster Losses, Proceedings of the 6th International ISCRAM conference, Gothenburg, Sweden, May 2009, Paper no 131.

Holub, M., Fuchs., S., 2008. Benefits of local structural protection to mitigate torrent-related hazards, in: Brebbia, C., Beriatos, E. (Eds.), Risk analysis VI. WIT transactions on information and communication technologies 39. WIT, Southampton, pp. 401-411.

Holub, M., Suda, J., Fuchs, S., 2012. Mountain hazards: reducing vulnerability by adapted building design. Environmental Earth Sciences 66, 1853-1870.

Hooke, J.M., Mant, J.M., 2000. Geomorphological impacts of a flood event on ephemeral channels in SE Spain. Geomorphology 34 (3-4), 163-180.

INDECI, 2013. Estado Situacional de la Emergencia: Informe sobre los eventos occasionados por las intensas lluvias del mes de febrero 2013. Base de datos del Instituto Nacional de Defensa Civil, Arequipa, Peru.

IPCC, 2007. Summary for policymakers, in: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.) Climate change 2007: the physical science basis, Contribution of Working Group I to the 4th Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge

IPCC Working Group I, 2007. Contribution of Working group I to the 4th Assessment Report of the Intergovernmental Panel on Climate Change 2007, Cambridge University Press, Cambridge, p. 996

Jakob, M., Stein, D., Ulmi, M., 2012. Vulnerability of buildings to debris flow impact. Natural Hazards 60, 241-261.

Jenkins, S.F., Spence, R.J.S., Fonseca, J.F.B.D., Solidum, R.U., Wilson, T.M., 2014. Volcanic risk assessment: Quantifying physical vulnerability in the built environment. Journal of Volcanology and Geothermal Research 276, 105-120. Doi.org/10.1016/j.jvolgeores.2014.03.002

Jonkman, S.N., Vrijling, J.K., 2008. Loss of life due to floods. Journal of Flood Risk Management 1, 43-56. Doi:10.1111/j.1753-318x.2008.00006.x.

Kaplan, S., Garrick, B., 1981. On the Quantitative Definition of Risk. Risk Analysis 1, 11–27.

Kaynia, A.M., Papathoma Köhle, M., Neuhäuser, B., Ratzinger, K., Wenzel, H., Medina-Cetina, Z., 2008. Probabilistic assessment of vulnerability to landslide : Application to the village of Lichtenstein, Baden-Württemberg, Germany. Engineering Geology 101(1-2), 33-48.

Lo, W.-C., Tsao, T.-C., Hsu, C.-H., 2012. Building vulnerability to debris flows in Taiwan: a preliminary study. Natural Hazards 64, 2107-2128.

Luino, F., 2005. Sequence of instability processes triggered by heavy rainfall in the northern Italy. Geomorphology 66(1–4), 13–39.

Marchi, L., Borga, M., Preciso, E., Gaume, E., 2010. Characterisation of selected extreme flash floods in Europe and implications for flood risk management. Journal of Hydrology 394 (1-2), 118-133. Doi:10.1016/j.jhydrol.2010.07.017.

Martelli, K., 2011. The physical vulnerability of urban areas facing the threat of inundation from lahars and flash floods: application to the case study of Arequipa, Peru. (Unpublished) PhD dissertation, Université Blaise Pascal, Clermont-Ferrand, 245 p.

Martínez Ibarra, E., 2012. A geographical approach to post-flood analysis: The extreme flood event of 12 October 2007 in Calpe (Spain). Applied Geography 32, 490-500.

Montz, B.E., Gruntfest, E., 2002. Flash flood mitigation: recommendations for research and applications. Environmental Hazards 4, 15-22.

Oehler, J.-F., Thouret, J.-C., Solikhin, A., Ettinger, S., 2014. High resolution DEMs based on HSR PLEIADES images: Applications to the Merapi volcano (Indonesia) and the city of Arequipa close to the Misti volcano (Peru). Cities on Volcanoes 8 meeting, 9-13 September, Yogjakarta, Indonesia. Abstract and poster.

Papathoma-Köhle, M., Kappes, M., Keiler, M., Glade, T., 2011. Physical vulnerability assessment for alpine hazards: state of the art and future needs. Natural Hazards 58(2), 645-680.

Papathoma-Köhle, M., Keiler, M., Totschnig, R., Glade, T., 2012. Improvement of vulnerability curves using data from extreme events: debris flow event in South Tyrol. Natural Hazards 64, 2083-2105.

Quan Luna, B., Blahut, J., Van Westen, C.J., Sterlacchini, S., Van Asch, T.W.J., Akbas, S.O., 2011. The application of numerical debris flow modelling for the generation of physical vulnerability curves. Natural Hazards and Earth System Science 11: 2047–2060.

Reid, I., Frostick, L.E., 1992. Channel form, flows and sediments in deserts, in: Thomas, D.S.G. (Ed.), Arid Zone Geomorphology, Belhaven Press, Rotterdam, pp. 117-135.

Santoni, O., 2011. Etude de la vulnérabilité de la ville d'Arequipa, Pérou, vis à vis du volcan 1049 El Misti, au moyen d'un SIG. M.Sc. Thesis, Université Blaise Pascal, Clermont-Ferrand, 51 p (in French).

SENAMHI, 2013. Periodo de lluvias 2013 en la région Arequipa y su relación con el Cambio Climático. Foro Regional Desastre, Reconstrucción y desarrollo en escenario de cambio climático Arequipa – emergencia 2013.

Thouret, J.-C., Ettinger, S., Guitton, M., Martelli, K., Santoni, O., Magill, C., Revilla, V., Charca, J.A., 2014. Assessing physical vulnerability and risk in large cities exposed to debris flows and flash floods: the Arequipa (Peru) case study. Natural Hazards, DOI 10.1007/s11069-014-1172-x, 47 p.

Thouret, J.C., Enjolras G., Martelli, K., Santoni, O., Luque J.A., Nagata, M., Arguedas, A., Macedo, L., 2013. Combining criteria for delineating lahar- and flash-flood-prone hazard and risk zones for the city of Arequipa, Peru. Nat. Hazards Earth Syst. Sci. 13, 339-360.

Thywissen, K., 2006. Components of risk. A comparative glossary. Published as SOURCE No. 2—2006 of the Institute of Environment and Human Security of the United Nations University. Available at: http://www.ehs.unu.edu/file/get/4042

Totschnig, R., Fuchs, S., 2013. Mountain torrents: Quantifying vulnerability and assessing uncertainties. Engineering Geology 155, 31-44.

Totschnig, R., Sedlacek, W., Fuchs, S., 2011. A quantitative vulnerability function for fluvial sediment transport. Natural Hazards 58, 681-703.

Villagran de Leon, J.C., 2006. Vulnerability - A Conceptual and Methodological Review, SOURCE - Publication Series of UNU-EHS n°4/2006, United Nations University UNU-EHS, UN Campus, Bonn, Germany, 90 p.

Wigmosta, M.S., 1983. Rheology and flow dynamics of the Toutle debris flows from Mt. St. Helens. M.Sc. Thesis, University of Washington, Seattle, 184 p.

Wisner, B., Blaikie, P., Cannon, T., Davis, I., 2004. At risk: Natural hazards, people's vulnerability, and disasters, 2nd edn. Routledge, London

Zuccaro, G., Ianniello, D., 2004. Interaction of pyroclastic flows with building structures in an urban settlement: a fluid-dynamic simulation impact model. Journal of Volcanology and Geothermal Research 133, 345-352.

Zuccaro, G., Cacace, F., Spence, R.J.S., Baxter, P.J., 2008. Impact of explosive eruption scenarios at Vesuvius. Journal of Volcanology and Geothermal Research 178, 416-453.

Zuccaro, G., De Gregorio, D., 2013. Time and space dependency in impact damage evaluation of a sub-Plinian eruption at Mount Vesuvius. Natural Hazards 68(3), 1399-1423.

Table caption

Table 1: Vulnerability parameters concerning building characteristics, building environment and flood hazard with their respective levels as defined for this study.

Table 2: Material types characterizing channel banks along the left and right riverside of the Avenida Venezuela channel. Numbers indicate the frequency distribution of channel sections in each material category and their respective distance to a bridge. The total number of sections located at a certain distance from a bridge is illustrated in bold with, to its right, the corresponding percentage of total channel bank length.

Table 3: Section lengths of channel banks as a function of construction material and location on left and right riverside.

Table 4: Damaged channel bank sections represented as the percentage of the total length of either the left or right channel side. Sections are attributed to one of six groups (A to F) depending on the closest distance to a bridge of either the start or end point of the section.

Table 5: Contributions of each parameter level to damage probability based on the calibration data set (598 buildings).

Figure caption (black and white reproduction in print is intended)

Figure 1A: Geographical setting and location of Arequipa city, Peru. B: The study area Avenida Venezuela channel and six zones that will serve to illustrate observations in the following.

Figure 2: Longitudinal channel profile (black line) with channel width (green dots), and sections in which erosion occurred (orange bars). The gray scale bar represents channel bed type, i.e. natural gravel, sand (dark gray), natural with occasional concrete steps (white) and concrete (medium gray); the channel wall material is represented by concrete (red), mixed material (concrete, brick, boulders; yellow) and natural (blue). For complementary information see also table 2 to 4.

Figure 3: Field-survey-based mapping of inundation extent resulting from overbank flow along the Avenida Venezuela channel.

Figure 4: Three examples of particular channel courses and resulting damage.

Figure 5 Left: Damage level observed for different material types of retaining walls. Right: Material types of retaining walls relative to the proximity of bridges.

Figure 6: Observed damage levels from left to right (4) inundated, heavy damage, (3) inundated, significant damage; (2) inundated, light damage; and (1) inundated, no structural damage.

Figure 7: Results of univariate analysis summarizing the number of buildings per category. Grayscales from the lowest parameter level 1 (white) to the highest level 5 (dark gray) are the same for all figures.

Figure 8: Plot of parameter co*uple “Distance from bridge” and “Damage” at respective levels.* Note the strong relationship between buildings located close to a bridge (DBR5) and damage level 4 (DO4; right side of vertical axis) compared to buildings far from a bridge (DBR1 and 2) that have damage level 2 (DO2, left side of vertical axis). Eigenvalues represent 96.54% for axis 1, 2.76% for axis 2, and 0.7% for axis 3.

Figure 9: Results of the bivariate analysis. Damage level is displayed in different gray shades, the abscissa (1 to 5) displays the categories of the respective parameter *“Distance from channel”, “Distance from bridge”, etc.*

Figure 10: Scatter plots representing results of the correspondence analysis. Each point represents a building. Ellipses colored from light blue to red represent parameter levels (1 to 5, respectively) as bagplots (bivariate boxplots). Each bagplot represents 67.5% of the buildings defining each level.

Figure 11: Projection of parameter levels (color) and city blocks as a result of the bivariate analysis. The position of each square is defined by the bagplot representing 67.5% of the buildings defining each level. The circles indicate city blocks of similar characteristics and thus behavior. City block numbers are plotted to allow comparison

but are not included in the bivariate analysis. The number of buildings per city block and the respective percentage is detailed in the histogram to the right.

Figure 12: Damage probabilities calculated for damage levels 1,2, 3 and 4 (series A) and all observed damage levels (1 to 4) in the field (series B) using calibration and validation data sets (898 buildings) of the selected test scenario presented in the manuscript.

APPENDIX caption

Appendix A and B: Survey forms for the damage assessment of buildings, conceived for masonry (A) and reinforced concrete (B) structures following experiences from previous studies concerning natural hazard impact (Zuccaro et al., 2008; Zuccaro and De Gregorio, 2013; Jenkins et al., 2014).

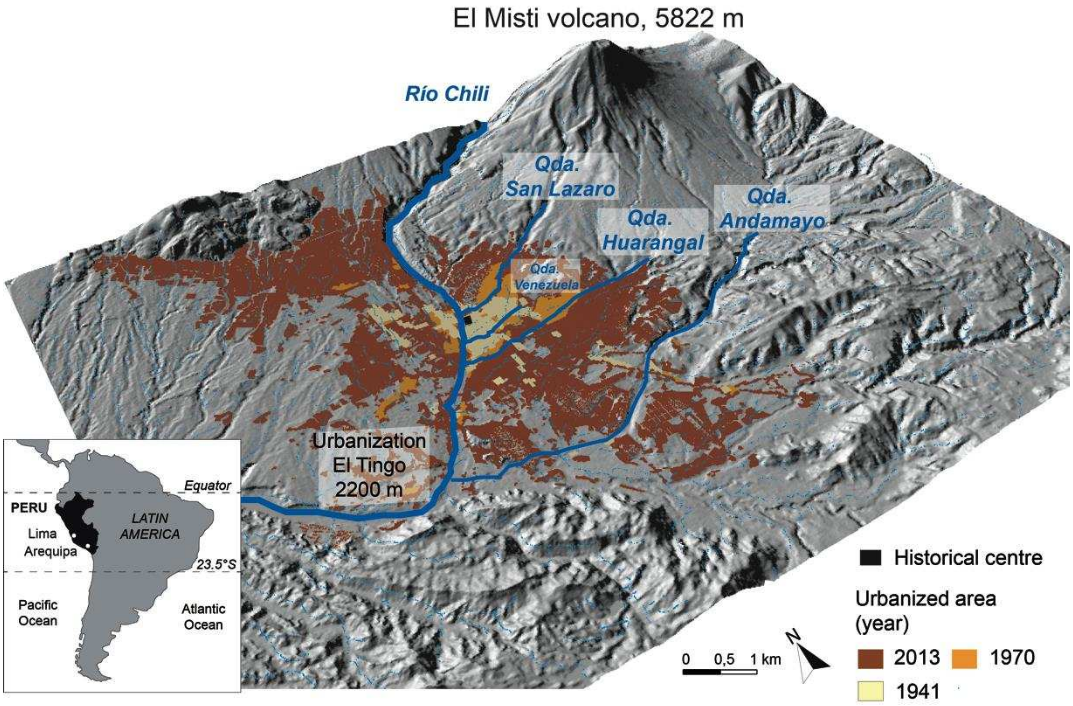

Figure 1: Geographical setting and location of Arequipa city, Peru

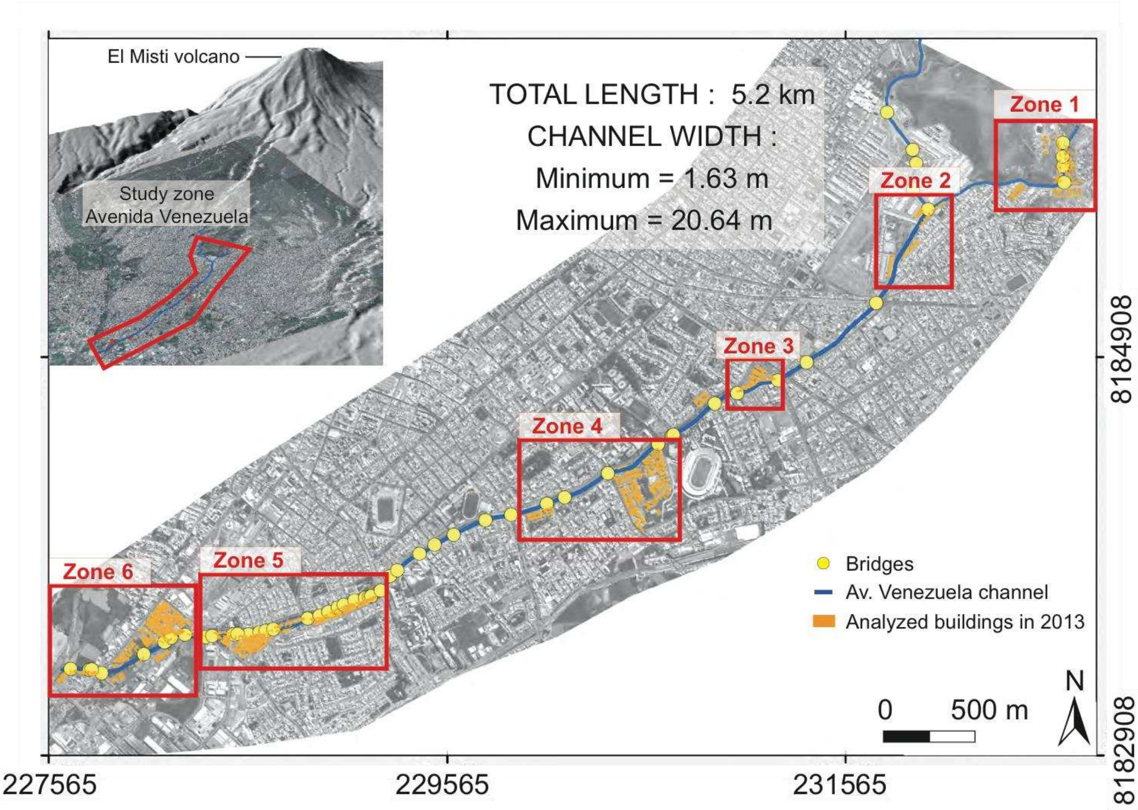

Figure 2: The study area Avenida Venezuela channel and six zones that will serve to illustrate observations in the following.

Building vulnerability parameter							
	Parameter Abbreviation	**Unit**	**Parameter level**				
			5	4	3	2	1
Distance from channel	DC	Meter	≤ 5	]5;10]	]10;25]	]25;60]	>60
Distance from bridge	DBR	Meter	≤ 15	]15;30]	]30;50]	]50;90]	>90
Shape of city block	SH	/	Complex	Irregular	Regular	Compact	Perfect
Impermeability of soil	IS	/	Permeable	/	/	/	Impermeable
Structural type of building	S	/	1,2,3	9,10,11,12	4,5	6,7,8	/
Number of storeys	NS	/	1	2	3	4 and more	/
Inundation	I	Meter	/	> 0.4	0.2-0.4	< 0.2	0
Building footprint	A	Square meter	> 150	]80-150]	]50-80]	]10-50]	≤10
Building density per city block	DE	Number per hectare	]0-40]	]40-60]	]60-80]	]80-110]	≥110
Damage parameter							
Observed damage (see fig. 7)	DO	/	Heavy	Moderate	Slight	Temporary inundation without damage	No damage

Figure 3: Building vulnerability parameters and their levels as defined for this study.

Figure 4: Longitudinal channel profile (blue line) with channel width (green dots), and sections in which erosion occurred (orange bars). The gray scale bar below represents channel bed type, i.e. natural gravel, sand (dark gray), natural with occasional concrete steps (light gray) and concrete (medium gray); the channel wall material is represented by concrete (dark brown), mixed material (concrete, brick, boulders; light brown) and natural (medium brown).

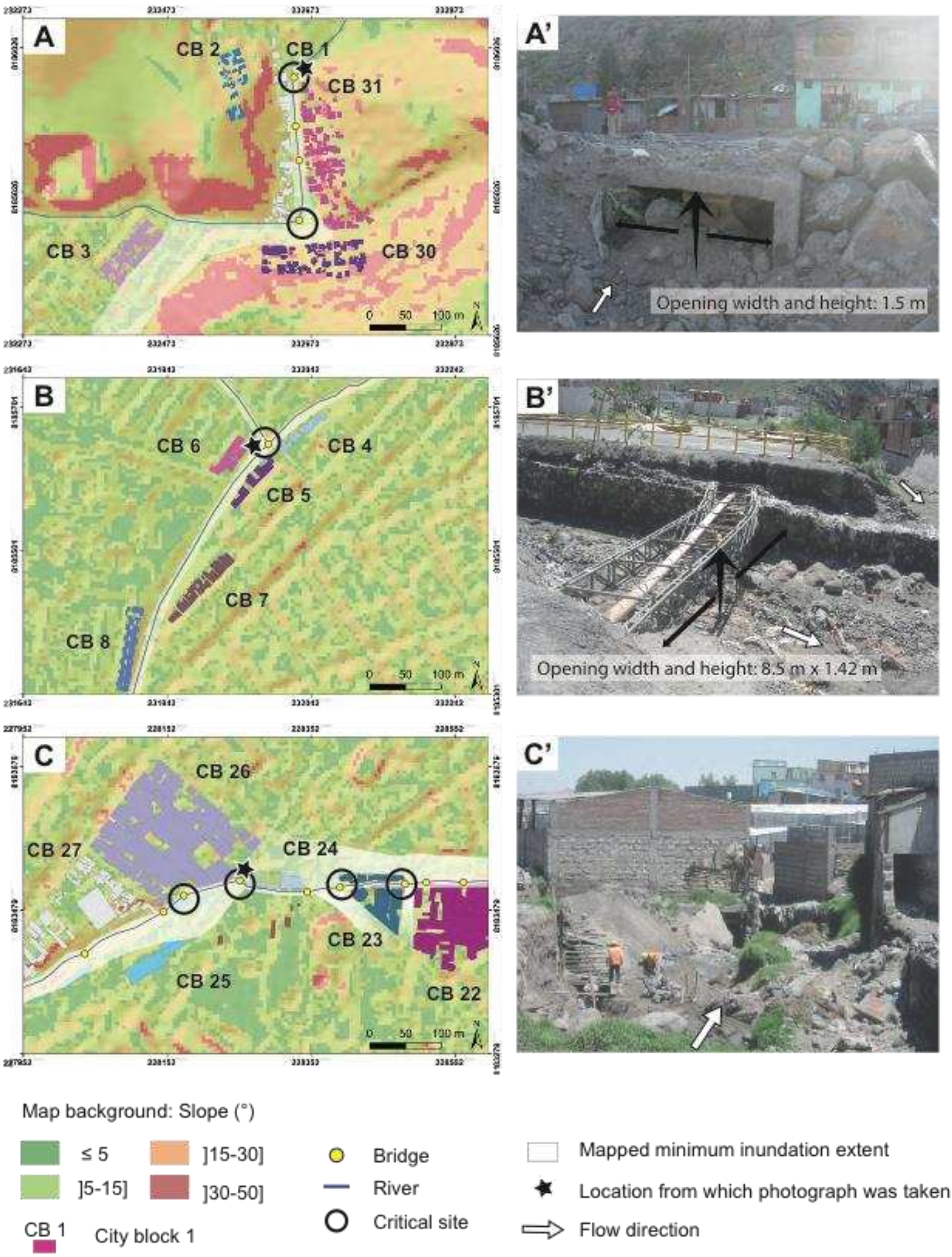

Figure 5: Three examples for particular channel courses and occurred damage.

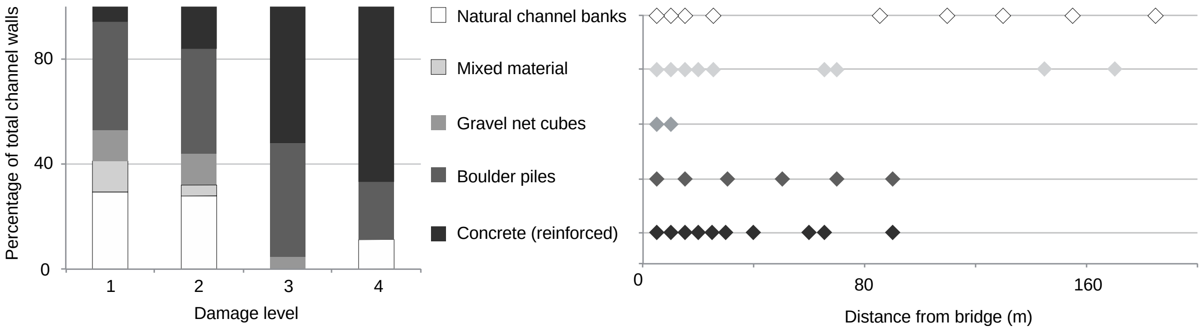

Figure 6 Left: Damage level observed for different material types of contention walls. Right: Material types of contention walls relative to the proximity to bridges.

Figure 7: Observed damage levels from left to right (4) inundated, heavy damage, (3) inundated, important damage; (2) inundated, light damage; and (1) inundated, without damage.

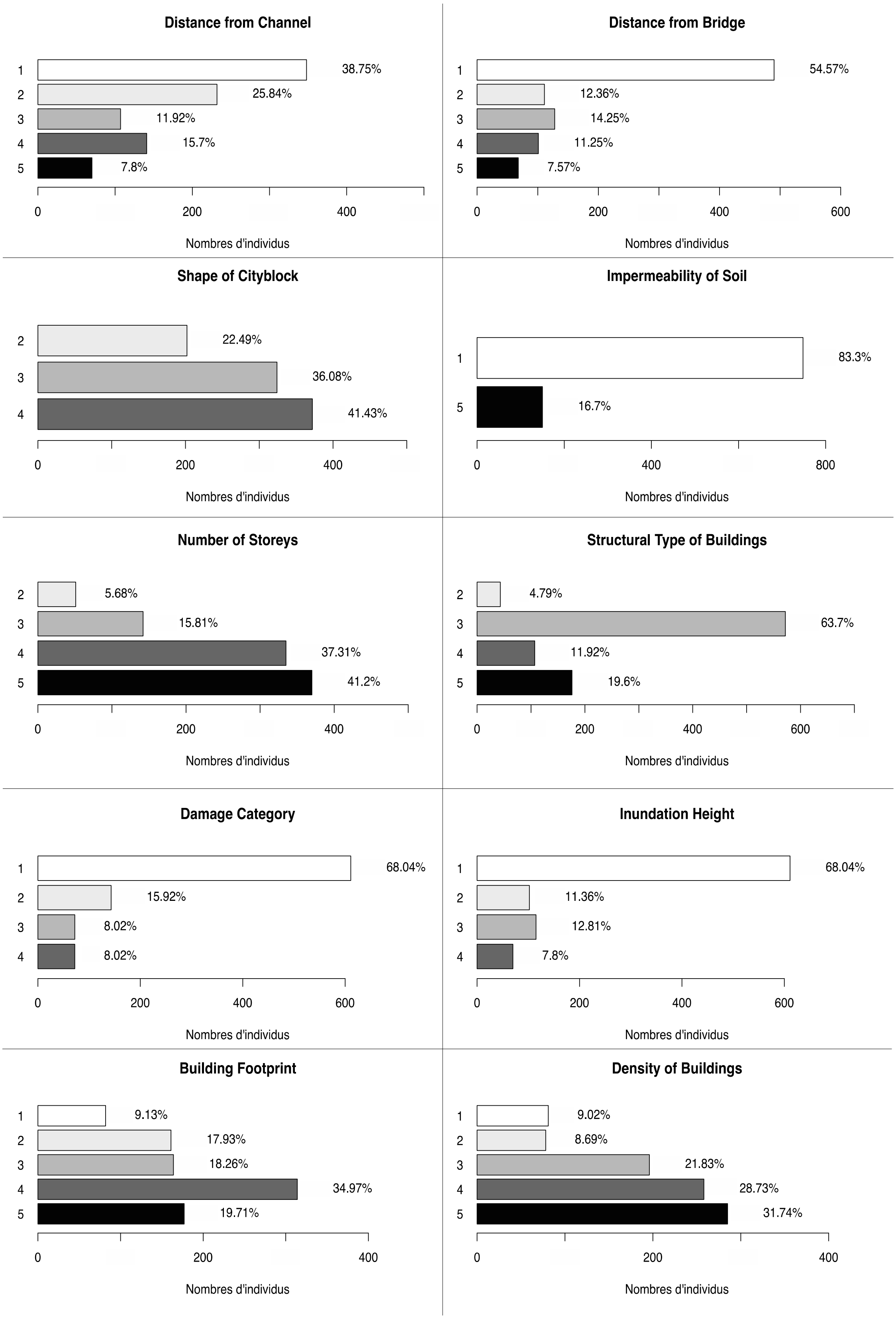

Figure 8: Results of univariate analysis summarizing the number of buildings per level. Grayscales from the lowest level 1 (white) to the highest level 5 (dark gray) are the same for all figures and represent the parameter levels (refer to fig. 3 for details on level characteristics).

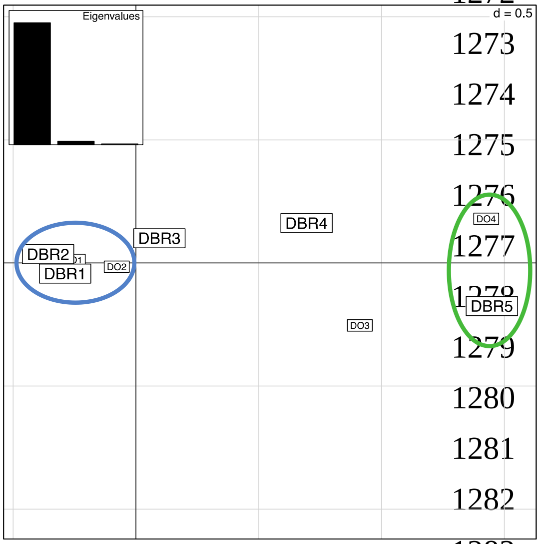

Figure 9: Plot of parameter *couple "Distance from bridge" and "Damage" with respective* levels. Note the strong relationship between buildings located close to a bridge (DBR5) and damage level 4 (DO4; right side of vertical axis) opposed to buildings far from a bridge (DBR1 and 2) that have damage level 2 (DO2, left side of vertical axis).

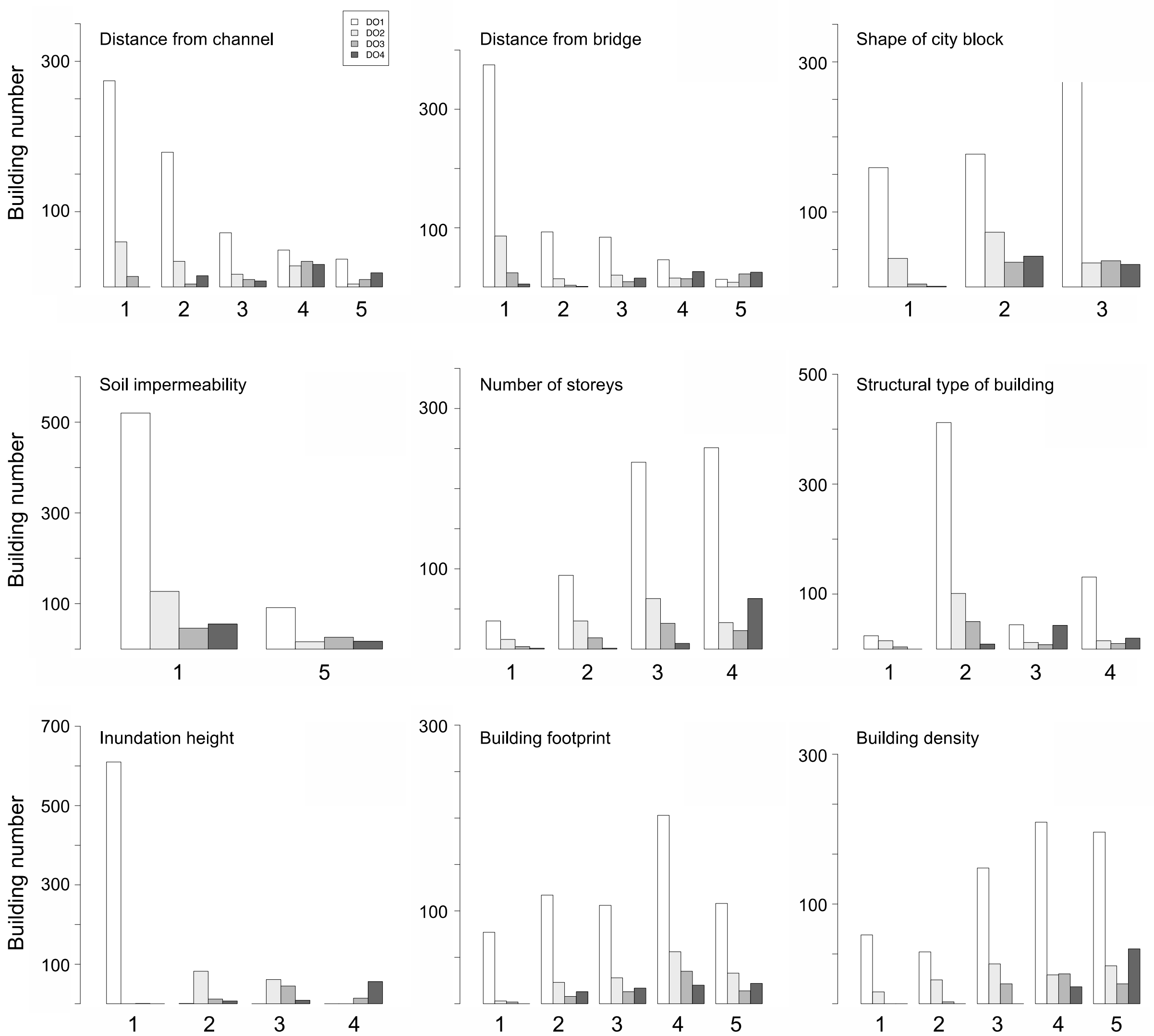

Figure 10: Results of the bivariate analysis. Damage level is displayed in different gray *shades, the abscissa (1 to 5) displays the categories of the respective parameter "Distance from channel", "Distance from bridge", etc.*

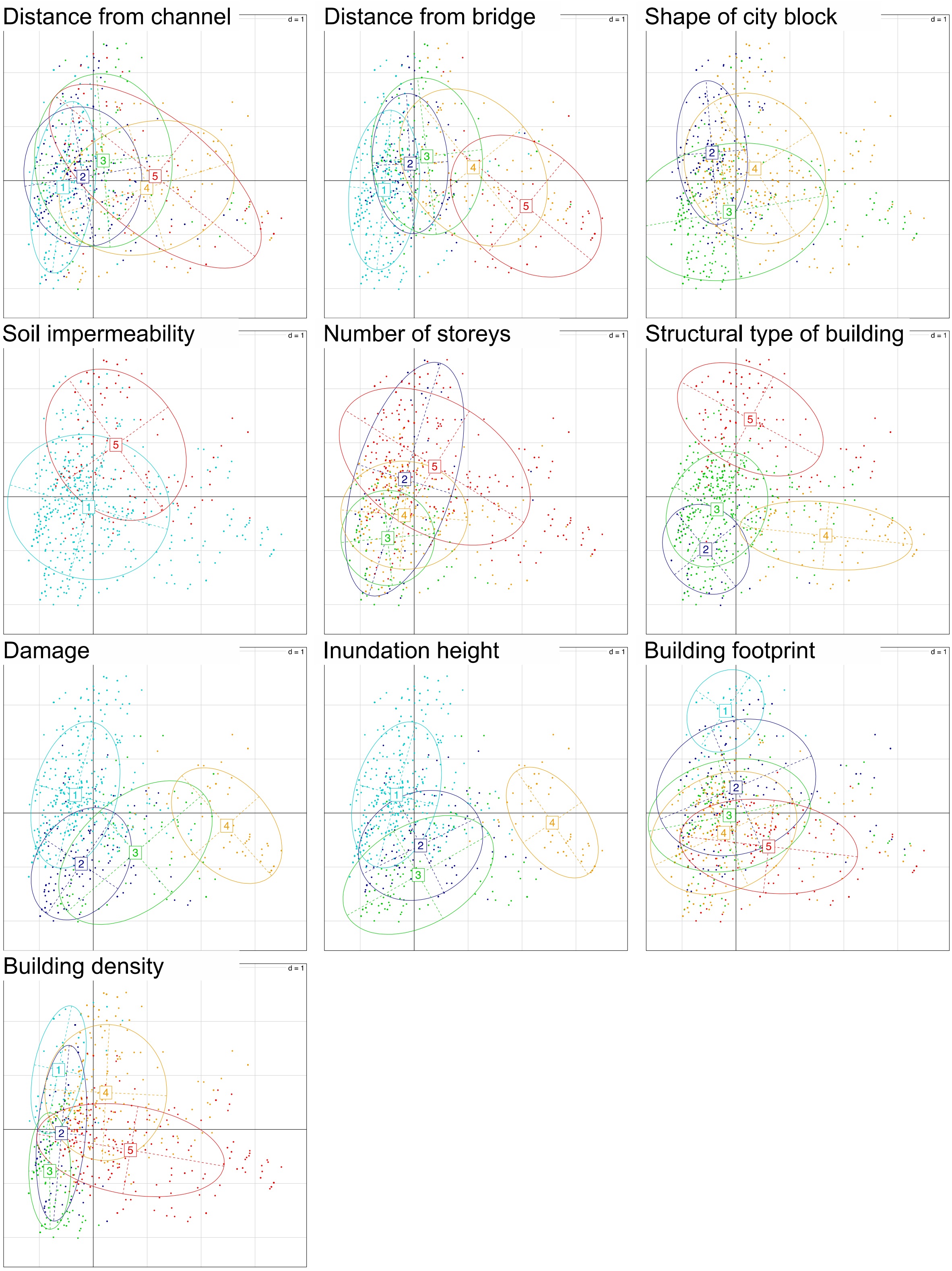

Figure 11: Scatter plots representing results of the bivariate analysis. Colors from blue to red represent parameter levels (1 to 5). The barycenter of the ellipse represents 67.5 % of the buildings defining each level.

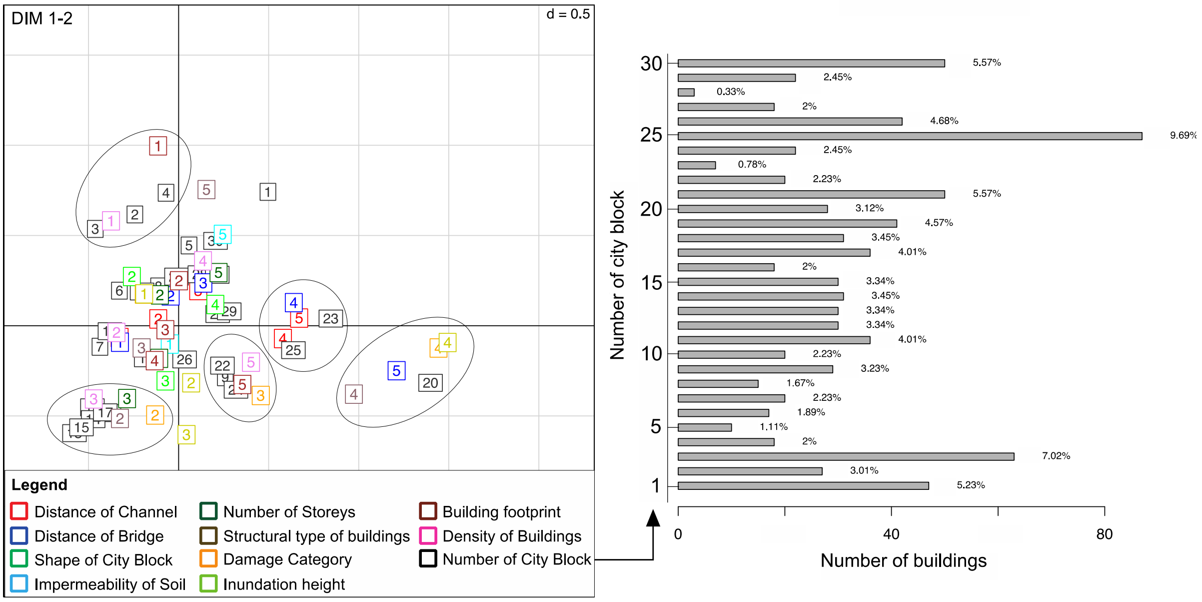

Figure 12: Projection of parameter levels (color) and city blocks as a result of the bivariate analysis. The position of each square is defined by the barycenter of the ellipse representing 67.5 % of the buildings defining each level. The circles indicate city blocks of similar characteristics and thus behavior. City block numbers are plotted to allow comparison but are not included in the bivariate analysis. The number of buildings per city block and the respective percentage is detailed in the histogram to the right.

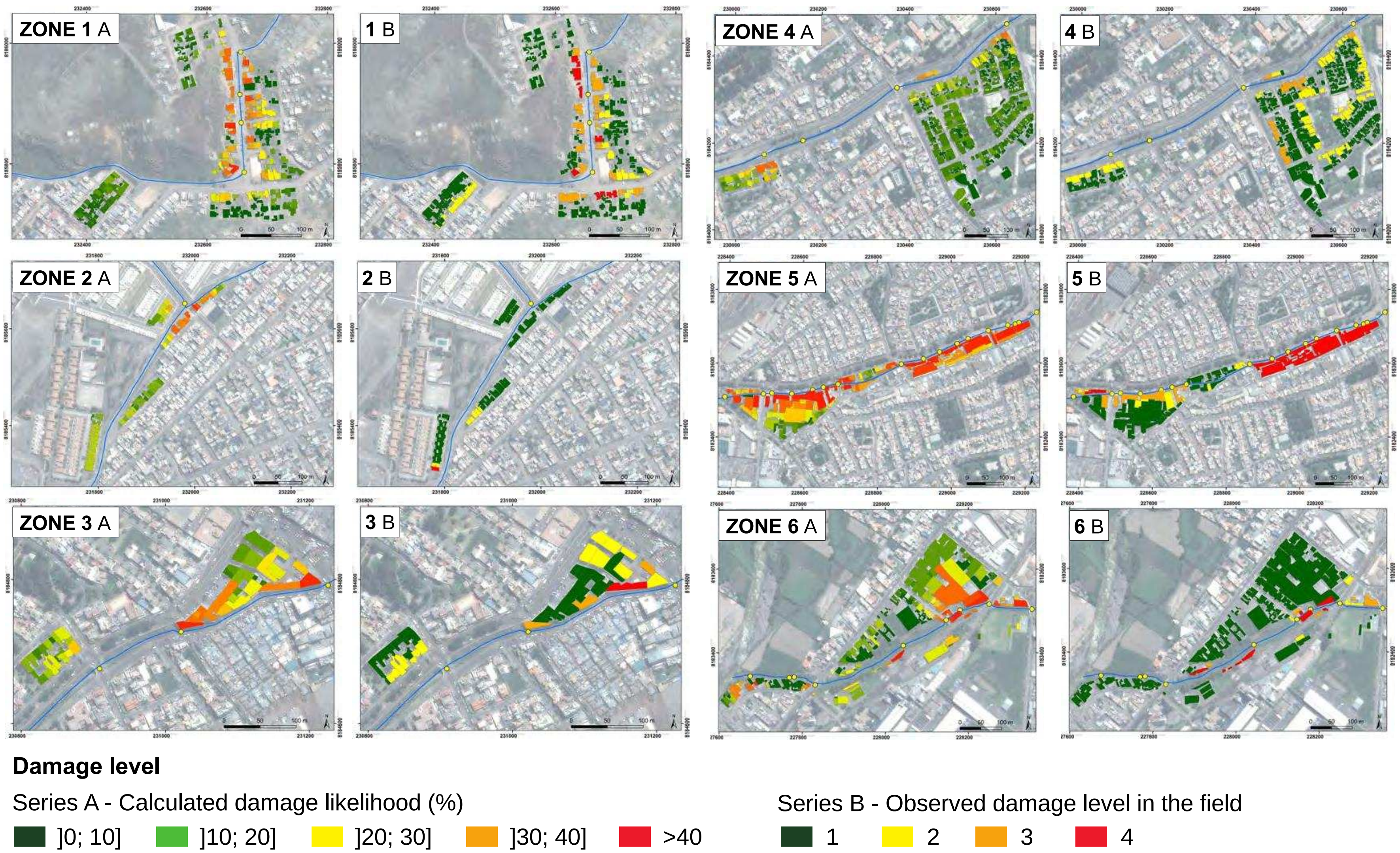
ZONE 1 A
1 B
ZONE 4 A
4 B
ZONE 2 A
2 B
ZONE 5 A
5 B
ZONE 3 A
3 B
ZONE 6 A
6 B
Damage level
Series A - Calculated damage likelihood (%)
]0; 10]
]10; 20]
]20; 30]
]30; 40]
>40
Series B - Observed damage level in the field
1
2
3
4

Figure 13: Damage probabilities calculated for damage level 3 and 4 (series A) and all observed damage levels (1 to 4) in the field (series B).