



**HAL**  
open science

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

Susanne Ettinger, Loïc Mounaud, Christina Magill, Anne-Françoise Yao-Lafourcade, Jean-Claude Thouret, Vern Manville, Caterina Negulescu, Giulio Zuccaro, Daniela de Gregorio, Stefano Nardone, et al.

### ► To cite this version:

Susanne Ettinger, Loïc Mounaud, Christina Magill, Anne-Françoise Yao-Lafourcade, Jean-Claude Thouret, et al.. Building vulnerability to hydro-geomorphic hazards: Estimating damage probability from qualitative vulnerability assessment using logistic regression. *Journal of Hydrology*, 2016, 541, pp.563 - 581. 10.1016/j.jhydrol.2015.04.017 . hal-01637127

**HAL Id: hal-01637127**

**<https://uca.hal.science/hal-01637127>**

Submitted on 26 Mar 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITY OF LEEDS

This is a repository copy of *Building vulnerability to hydro-geomorphic hazards: Estimating damage probability from qualitative vulnerability assessment using logistic regression*.

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/110292/>

Version: Accepted Version

---

**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>

---

© 2015, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International  
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

**Reuse**

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

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

3  
4 Susanne Ettinger <sup>a</sup>, Loïc Mounaud <sup>b</sup>, Christina Magill <sup>c</sup>, Anne-Françoise Yao-Lafourcade <sup>b</sup>,  
5 Jean-Claude Thouret <sup>a</sup>, Vern Manville <sup>d</sup>, Caterina Negulescu <sup>e</sup>, Giulio Zuccaro <sup>f</sup>, Daniela De  
6 Gregorio <sup>f</sup>, Stefano Nardone <sup>f</sup>, Juan Alexis Luque Uchuchoque <sup>g</sup>, Anita Arguedas <sup>h</sup>, Luisa  
7 Macedo <sup>i</sup>, Nélide Manrique Llerena <sup>i</sup>

8  
9 <sup>a</sup> Clermont Université, Université Blaise Pascal, Laboratoire Magmas et Volcans, CNRS  
10 UMR6524, IRD-R163 and ClerVolc, 5 rue Kessler, F-63000 Clermont-Ferrand

11 <sup>b</sup> Université Blaise Pascal, Laboratoire de Mathématiques, CNRS UMR 6620, Campus des  
12 Cézeaux, 63171 Aubière Cedex, France

13 <sup>c</sup> Risk Frontiers, Macquarie University, NSW2109, Australia

14 <sup>d</sup> School of Earth and Environment, University of Leeds, Leeds LS29JT, United Kingdom

15 <sup>e</sup> BRGM, 3 avenue C. Guillemin, BP 36009, 45060 Orléans Cedex 2, France

16 <sup>f</sup> Plinius Study Centre, Università Federico II, Naples, Italy

17 <sup>g</sup> Defensa Civil Paucarpata and NGO PREDES, Av. Unión Nro. 200, Urb. César Vallejo,  
18 Paucarpata, Arequipa, Peru

19 <sup>h</sup> Defensa Civil, Región Arequipa, Municipalidad Paucarpata, Arequipa, Peru

20 <sup>i</sup> INGEMMET, Instituto Nacional de Geología, Minero y Metalúrgico, Av. Dolores (Urb. Las  
21 Begonias B-3), José Luis Bustamante y Rivero, Arequipa, Peru

22  
23  
24  
25  
26  
27 **Corresponding author:**

28 Susanne ETTINGER

29

30 **Current address:**

31 Université Blaise Pascal, Laboratoire Magmas et Volcans, CNRS UMR6524, IRD R163,

32 5 rue Kessler, 63038 Clermont-Ferrand, France

33 **Telephone:** +33 (0)4 73 34 67 99

34 **Email:** susanne.ettinger@gmail.com

35

36

37 **Abstract**

38

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

46 This study assesses the aspects of building design and site specific environmental  
47 characteristics that render a building vulnerable by considering the example of a flash flood  
48 event in February 2013. A statistical methodology is developed that enables estimation of  
49 damage probability for buildings. The applied method uses observed inundation height as a  
50 hazard proxy in areas where more detailed hydrodynamic modeling data is not available.  
51 Building design and site-specific environmental conditions determine the physical vulnerability.  
52 The mathematical approach considers both physical vulnerability and hazard related  
53 parameters and helps to reduce uncertainty in the determination of descriptive parameters,  
54 parameter interdependency and respective contributions to damage. This study aims to (1)  
55 enable the estimation of damage probability for a certain hazard intensity, and (2) obtain data

56 to visualize variations in damage susceptibility for buildings in flood prone areas. Data  
57 collection is based on a post-flood event field survey and the analysis of high (sub-metric)  
58 spatial resolution images (Pléiades 2012, 2013). An inventory of 30 city blocks was collated in  
59 a GIS database in order to estimate the physical vulnerability of buildings. As many as 1103  
60 buildings were surveyed along the affected drainage and 898 buildings were included in the  
61 statistical analysis. Univariate and bivariate analyses were applied to better characterize each  
62 vulnerability parameter. Multiple corresponding analyses revealed strong relationships  
63 between the “Distance to channel or bridges”, “Structural building type”, “Building footprint”  
64 and the observed damage. Logistic regression enabled quantification of the contribution of  
65 each explanatory parameter to potential damage, and determination of the significant  
66 parameters that express the damage susceptibility of a building. The model was applied 200  
67 times on different calibration and validation data sets in order to examine performance. Results  
68 show that 90% of these tests have a success rate of more than 67%. Probabilities (at building  
69 scale) of experiencing different damage levels during a future event similar to the 8 February  
70 2013 flash flood are the major outcomes of this study.

71

## 72 **Keywords**

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

74

75

76

77

78

79

80

## 81 **1. Introduction**

82 On February 8 2013, heavy rainfall (124.5mm within 3 hours versus a monthly mean of  
83 29.3mm) triggered a flash flood event along the Avenida Venezuela channel in the city of

84 Arequipa, Peru. On this day, more than 280 buildings and 23 of a total of 53 bridges  
85 (pedestrian, vehicle and railway) were affected; the partial collapse of sections of a major road  
86 paralyzed central parts of the city for more than one week. Previous risk assessment studies  
87 in Arequipa did not include the Avenida Venezuela channel due to its smaller size and largely  
88 confined channel course. The high recurrence rate of hydro-geomorphic hazards (Martelli,  
89 2011; Thouret et al., 2013, 2014), and apparent locally high vulnerability of buildings and  
90 critical infrastructure in Arequipa, are major motivations for this study.

91 Risk in the context of disaster risk management is commonly defined as a potential loss for a  
92 given probability function (Crichton, 1999; Kaplan and Garrick, 1981). In the standard  
93 conceptual framework, risk is the product of hazard, vulnerability and exposure (Cardona,  
94 2004; Carreno et al., 2006). While the hazard is generally described by its severity, e.g.  
95 inundation height for a given return return period, exposure relates to the number and value of  
96 elements potentially affected (Hiete and Merz, 2009). Many different definitions, concepts and  
97 methods to systemize vulnerability exist in the current literature (Birkmann, 2006; Cutter, 2003;  
98 Wisner et al., 2004; Thywissen, 2006; IPCC, 2007; Bründl et al., 2009). In this study we follow  
99 the definition for physical vulnerability proposed by Glade (2003) and Villagran de Leon (2006)  
100 as the predisposition of an element or system to be affected or susceptible to damage as the  
101 result of the natural hazard's impact. Vulnerability assessment for hydro-geomorphic hazards  
102 such as dilute floods, debris flows, hyperconcentrated flows etc. is inherently complex, mainly  
103 as a result of the following factors (Gaume et al., 2009): (i) lack of accurate or real-time  
104 observational data necessary for reliable hazard analysis; (ii) only substantial damage  
105 information is generally recorded and accurate information on failure characteristics is often  
106 missing (Fuchs et al., 2007b; Papathoma Köhle et al., 2011); (iii) different time and  
107 geographical scales involved (Gruntfest, 2009; Marchi et al., 2010); (iv) natural adjustments of  
108 the environment to return to a state of equilibrium; (v) rapid intervention by technical services  
109 to restore functionality of urban infrastructure reduces the time frame for damage assessment  
110 in the field; (vi) site-specific triggering processes and upstream-downstream evolution of  
111 debris-flow phenomena (Di Baldassarre and Montanari, 2009). If field investigations are

112 conducted to study and record structural damage following a hazard event, these data are then  
113 generally correlated to the process intensity, frequently derived from deposition height or  
114 inundation height, in order to develop empirical fragility curves (Fuchs et al., 2007a,b; Holub  
115 and Fuchs, 2008). These curves are then employed within risk assessments to estimate  
116 structural damage in future events. The lack of high-quality observational evidence and  
117 uniform, i.e. non site-specific, approaches to data collection, implies that the majority of fragility  
118 curves are still developed using expert judgment (Papathoma Köhle et al., 2012; Totschnig  
119 and Fuchs, 2013). The compilation of field data for different sites in the European Alps, Taiwan  
120 etc. published in recent studies (Totschnig et al., 2011; Holub et al., 2012; Papathoma Köhle  
121 et al., 2012; Totschnig and Fuchs, 2013) has helped to develop vulnerability functions  
122 applicable within the framework of risk management for specific regions (Totschnig and Fuchs,  
123 2013). If the required input data are available, the method is transferable to other alpine  
124 regions. However, data availability remains a major constraint in many countries (e.g.,  
125 Douglas, 2007; Jakob et al., 2012; Lo et al., 2012; Totschnig and Fuchs, 2013). For Latin  
126 America, very few case studies have been published with a focus on physical vulnerability  
127 analysis. In contrast to many sites monitored and equipped in the European Alps, areas prone  
128 to hydro-geomorphic hazards in the Andes are rarely monitored, and in the worst case, not  
129 even mapped. It therefore becomes difficult to apply methods derived from European  
130 experience in the same or a similar way. In addition there is a critical lack of observational data  
131 collected in the immediate aftermath of disasters. For the study presented here, data to apply  
132 existing vulnerability assessment methods were not available, although alternative information  
133 could be collected.

134 Flash floods are common in semi-arid areas, such as Arequipa, and can, despite their  
135 infrequent nature, have a devastating effect in both geomorphological and human terms  
136 (Gaume et al., 2009; Jonkman and Vrijling, 2008; Martínez Ibarra, 2012). The occurrence of  
137 flash floods is highly variable, both spatially and temporally, most occurring as the result of  
138 localized intense storms (Graf, 1988; Reid and Frostick, 1992; Hooke and Mant, 2000). Several  
139 important factors arise as a result of these characteristics. First, areas prone to flash floods

140 need to be adequately prepared. Because events usually occur unexpectedly, warning and  
141 preparation are essential (Montz and Grunfest, 2002; Collier, 2007; Borga et al., 2008; Gaume  
142 et al., 2009); however, because events are typically rare, the motivation to invest time and  
143 resources into such activities may be lower than for more frequent hazards (Grunfest and  
144 Handmer, 2001). Flash floods usually affect relatively small areas and losses resulting from  
145 them do not always generate much long-term response, unless there is high loss of life.  
146 However, losses per unit (acre, square mile, or kilometer) of area affected tend to be high  
147 compared to other events such as riverine floods or hurricanes due to locally high intensity  
148 (Gaume et al., 2009; Martínez Ibarra, 2012).

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

161 Several recent studies (Martelli, 2011; Santoni, 2011; Ettinger et al., 2014a,b; Thouret et al.,  
162 2013, 2014) examined the physical vulnerability of buildings and critical infrastructure in the  
163 city of Arequipa, Peru. Thouret et al. (2014) established vulnerability indicators for buildings  
164 based on experiences by Zuccaro and Ianniello (2004), Zuccaro et al. (2008) and Zuccaro and  
165 De Gregorio (2013) that were calibrated on-site in Arequipa. Our research builds on this work  
166 and analyzes the relationships between these parameters and their significance in terms of  
167 the susceptibility of a building to experience damage. The present study aims to develop a



168 methodology for rapid estimation of potential damage of existing structures facing natural  
169 hazards, in particular flood-hazard. It can be useful, in particular for developing countries, in  
170 the case of inadequate hazard information, i.e. in areas where there have been no surveyed  
171 hazard events or hydraulic modeling. The objectives of this research are fourfold: to (1) map  
172 and characterize channel morphology in natural and built sections; (2) determine and quantify  
173 the relationships between building vulnerability parameters; (3) identify the weight of each  
174 parameter; and (4) apply mathematical models to calculate the damage probability for each  
175 building.

176

## 177 **2. Geographical setting**

178

179 Arequipa, with a population of c. 900,000, is the second largest city in Peru, located at c.  
180 2,300m above sea level, at the foothill of three summits of the Peruvian Andes: El Misti volcano  
181 (5,821m asl) to the northeast, flanked by Mounts Chachani (6,075m asl) to the North and Pichu  
182 Pichu (5,664m asl) to the East. The high altitude and semi-arid climate are responsible for  
183 scarce vegetation cover in both low and high altitudes. Abundant unconsolidated volcanic  
184 debris is therefore exposed on steep mountain slopes. Mean annual precipitation does not  
185 exceed 150mm and rainfall occurs mainly in the form of low frequency-high intensity  
186 rainstorms. These events often trigger flash floods, which sweep through the city of Arequipa  
187 following one or more of the numerous channels draining the flanks of El Misti volcano. Since  
188 the 1940s, the city's population has quadrupled, occupying at present a built surface of  
189 approximately 5,025ha (Fig. 1). The Río Chili valley is a geographical barrier separating the  
190 city in two parts; urbanization is extending the city area to the West but also to the East,  
191 colonizing the lower flanks of El Misti volcano. Intense urbanization is exposing an increasing  
192 number of people and built environment to flash flood hazards.

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

196 February 2013 rainfall was the highest for that month (SENAMHI, 2013; Cacya et al., 2013).  
197 The high intensity of this particular rainstorm generated a flash flood that affected several  
198 districts of the city (INDECI, 2013; Cacya et al., 2013). Previously conducted risk assessment  
199 studies in Arequipa (Martelli, 2011; Thouret et al., 2013, 2014) considered major drainages  
200 such as the Río Chili, San Lazaro and Huarangal. However, the 2013 rainfall event rainfall  
201 event affected in particular the smaller Avenida Venezuela secondary drainage channel.  
202 Two tributaries drain a c. 7.8km<sup>2</sup> catchment characterized by abundant non-consolidated  
203 debris and feed the Avenida Venezuela channel (Fig. 1B): (1) the northern tributary drains  
204 watersheds to the Northeast, upstream of the Cooperativa 14 to La Galaxia urban areas; and  
205 (2) the southern tributary drains watersheds to the Southeast, upstream of the Mariano  
206 Bustamante and Joven Vencedores del Cenepa urban areas. Before joining the main Río Chili  
207 valley to the West, the Avenida Venezuela channel crosses the city from NE to SW. Over a  
208 total length of 5.2km, the channel depth ranges from 1.3 to 6.3m, with channel widths from  
209 1.63 to 20.64m.

210

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

214

### 215 **3. Methods**

216 The general methodological approach proposed in this study benefits from data and insights  
217 gained from previous exposure and vulnerability assessments carried out in Arequipa (Santoni,  
218 2011; Martelli, 2011; Thouret et al., 2013, 2014). Additional data, in particular, concerning the  
219 flood hazard, Avenida Venezuela channel characteristics and surrounding built environment,  
220 which had not been studied before, was collected during field work in March 2013 and compiled  
221 in an extensive GIS database.

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

- 224 (i) Information for each parameter, potentially describing vulnerability, needed to be  
225 available for all of the individual buildings considered in the study.
- 226 (ii) [Thouret et al. \(2014\)](#) observed most vulnerable city blocks to be located within c.  
227 100 m of river channels or in proximity to tributary confluences. Past flood events  
228 and flow extension are frequently associated with overbank flow and occasions  
229 where bridges acted as obstacles to flow evacuation downstream ([Martelli, 2011](#);  
230 [Thouret et al., 2014](#)). The distance from the channel and from bridges was therefore  
231 considered to be an important parameter to investigate.
- 232 (iii) Previous studies ([Santoni, 2011](#); [Martelli, 2011](#); [Quan Luna et al., 2011](#)) examined  
233 vulnerability at the city block-scale and highlighted the importance of city block  
234 shape, building density, and soil impermeability for flow propagation/deviation or  
235 velocity, both in downstream direction and laterally.
- 236 (iv) The structural type of buildings as well as the number of storeys has been  
237 demonstrated as decisive for survival and resistance to flow impact by numerous  
238 studies ([Papathoma Köhle et al., 2011](#); [Zuccaro and Ianniello, 2004](#); [Zuccaro et al.,](#)  
239 [2008](#); [Jenkins et al., 2014](#)) and was therefore considered.
- 240 (v) The building footprint was included in order to determine its dependency to other  
241 building related parameters and susceptibility to damage.

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

246

### 247 3.1 Data collection and processing

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

252 2013 at sub-metric resolution were used in ArcGIS to map the channel and built environment  
253 affected by the flash flood, both before (2012) and after (2013) the hazardous event.

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

263 Additional data (photography, eyewitness accounts, Civil Defense reports, etc.) were also  
264 gathered and compiled in the GIS database. Media images and video footage (professional  
265 and social), freely available on the Internet, were invaluable in assessing hazard intensity, flow  
266 impacts, damage types, affected sites and deposit types or height. Images taken the day after  
267 the event were particularly useful to estimate the immediate aftermath of the inundation. This  
268 complimentary data allowed us to monitor impacts in near real-time, identify areas where  
269 impact assessments would be most informative and to map the spatial extent of affected areas  
270 and occurrence of damage. More than 300 photographs and 15 newspaper articles were  
271 studied to extract qualitative and semi-quantitative information regarding damage and flow  
272 characteristics. Where flood marks were still visible at the time of our field survey, or where  
273 inhabitants were willing to share their experiences, damage level and inundation height were  
274 verified along the channel. Run-up measurements could be realized at the tributary confluence  
275 in the upstream sector (Fig. 1), at one channel bend in the intermediate section and on a  
276 building in the downstream sector. The use of [Chow's \(1959\)](#) formula ( $v = 2 g\Delta h)^{0.5}$  or [Wigmosta](#)  
277 [\(1983\)](#), ( $v = 1.2 g\Delta h)^{0.5}$  enabled estimation of flow velocities based on the measured runup  
278 height (in both formulae  $v$  is flow velocity,  $g$  is gravitational acceleration and  $\Delta h$  the difference  
279 in mudline elevation).

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

283 For the damage assessment of buildings, survey forms were conceived for masonry and  
284 reinforced concrete structures (see Appendices A and B) following experiences from previous  
285 studies concerning natural hazard impact (Zuccaro et al., 2008; Zuccaro and De Gregorio,  
286 2013; Jenkins et al., 2014). The survey scheme followed in these forms relies on detailed  
287 predefined categories describing different damage intensities and impact types. This  
288 procedure was based here on standardized characteristics in order to optimize repeatability of  
289 the survey and minimize operator bias.

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

294

295

### 296 3.2 Statistical data analysis

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

306 All of the nine considered parameters were initially qualitative, i.e. observational or descriptive  
307 (Table 1). Four of them (“Distance from channel” and “Distance from bridge”, “Number of

308 storeys” and “Building density in a city block”) were rendered quantitative, i.e. calculated or  
309 measured, in order to increase the statistical model performance. Qualitative parameters  
310 include the “Inundation height”, the “Soil (im-)permeability”, the “Structural building type” and  
311 the “Shape of the city block”. These explanatory parameters were then related to the  
312 dependent parameter observed “Damage”. Parameters are either binary (e.g. soil permeability  
313 either “permeable” or “impermeable”) or described with up to 5 value categories (levels). An  
314 increased number of parameters could not be differentiated as this would have reduced the  
315 total sample size of buildings to an extent where the number of cases in each corresponding  
316 damage class would have been too low for a robust probabilistic assessment.

317

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

320

321 A first step assessed the frequency distribution of buildings for each of the different parameter  
322 levels: for example, concerning the parameter “Distance from channel”, the frequency  
323 distribution represents the number of buildings which are part of level 1, 2, ... or 5. The  
324 thresholds delimiting each parameter level were determined in order to respect a minimum of  
325 45 buildings per level (5% of the total building data), necessary to render the level significant.  
326 The frequency distribution was then graphically displayed in 2D histograms with the abscissa  
327 representing the number of buildings and the ordinate the parameter levels of the examined  
328 parameter.

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

331 The CA summarizes the relationships between the different parameter levels as scores in  
332 contingency tables and enables graphical representation of the latter in several 2D-plots. Each  
333 graphical representation is naturally based on two axes (dimensions), each of which expresses  
334 a certain percentage of information (inertia) of the contingency table. The dimensions are  
335 ranked as a function of their contribution to global inertia (=100%) of the contingency table. In

336 our study, dimensions 1 and 2 have the highest contributions (from 50 to 98.9, with most of  
337 that attributed to the first dimension) compared with dimension 3 (< 10).

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

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

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

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

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

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

362

363  $\text{logit}\left(P\left(Y = k \mid X_1, \dots, X_j\right)\right) = a_0 + a_1 X_1 + \dots + a_j X_j$  with  $k = 1, \dots, K$  (1)

364

365 where  $a_0$  is the model constant,  $Y$  the dependent parameter (Damage),  $k$  the level of the  
 366 dependent parameter (Damage level),  $K$  the highest possible level of damage,  $j$  the number of  
 367 explanatory parameters,  $X_1, \dots, X_j$  the respective parameter level of each explanatory  
 368 parameter. The applied logit function is:

369

370  $\text{logit}(p) = \ln \frac{p}{1-p}$  (2)

371

372

373 For ease of presentation in the following, we set:

374  $S = \text{logit}\left(P\left(Y = k \mid X_1, \dots, X_j\right)\right)$  (3)

375 This implies that:

376  $P\left(Y = k \mid X_1, \dots, X_j\right) = \frac{e^S}{1 + e^S}$  (4)

377 Using the calculated coefficients and parameter level values proper to each building in Eq. (4),  
 378 we can therefore define the probability of a building to experience damage at damage level  $k$ .

379 In order to obtain the damage probability at the precise level 1, 2, 3 or 4, we use

380

381  $P\left(Y = k \mid X_1, \dots, X_j\right) = P\left(Y = k \mid X_1, \dots, X_j\right) - P\left(Y = k - 1 \mid X_1, \dots, X_j\right)$  (5)

382

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

386

387 3.3 Model validation



388 Formal model validation was realized in two steps. First, using a calibration and validation data  
389 set. Hereby, the original data set was split into a calibration and a validation data set. From the  
390 898 totally sampled buildings, the validation data set contained 300 arbitrarily selected  
391 buildings. The model was calibrated using the 598 remaining buildings and then run to test the  
392 validation data set. Calculated damage probabilities obtained by the model for buildings in the  
393 validation data set was then compared to observed damage in the field.

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

397

## 398 **4. Results**

### 399 **4.1 Channel characteristics**

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

416 approximately 23%. This is a result of recent channel confinement work especially in the  
417 intermediate section where major road works were ongoing in 2013.

418

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

424

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

427

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

438

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

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

445

446 Peak flow discharge for the February 2013 event was estimated using channel cross section  
447 measurements and average flow velocity to be  $123.4\text{m}^3\text{s}^{-1}$  in the upstream reach and c.  
448  $425\text{m}^3\text{s}^{-1}$  in the middle reach (in proximity to the La Negrita shopping mall, fig. 1B). A clear  
449 delineation of areas inundated by intense surface runoff or by overbank flow of the Venezuela  
450 channel was not possible in many parts along the channel, which prevents a more precise  
451 estimate of the flow volume. An attempt at mapping the inundation extent resulting from  
452 overbank flow was made on the basis of field survey, including observations of erosion marks  
453 along the channel, flood marks on built infrastructure and eye witness accounts (Fig. 3).  
454 Comparing this map with flow simulations published in previous studies (Martelli, 2011; Oehler  
455 et al., 2014), the flow volume of the February 2013 flash flood can be estimated between  
456 50,000 and 100,000 $\text{m}^3$ . There appeared to be no general rule to where erosion occurred;  
457 natural bed types were affected to the same extent as concrete beds. Concerning the channel  
458 wall material, concrete and mixed material sections seemed to be affected more often by  
459 erosion than natural sections. For concrete sections, erosion is most likely at channel  
460 contractions or expansions or where bed or channel wall materials change. When erosion  
461 occurs in natural sections, the proximity to bridges with low opening height frequently  
462 determines whether erosion occurs or not.

463

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

466

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

471 2 and 5). In the case of 17 other damaged bridges along the Venezuela channel, openings  
472 were either not large enough for an increased discharge or were obstructed by boulders. In  
473 both cases, the consequence was similar: the flow front overtopping the bridge and caused  
474 overbank flow to both sides of the obstacle (Fig. 4, zone 1 and zone 5-6). Additionally, partially  
475 confined sections with an abrupt change in channel direction (e.g. close to 90° angle) were  
476 particularly prone to overbank flow (Fig. 4, Zone 1).

477

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

479

#### 480 **4.2 Flood impact on channel banks**

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

493

494 Material types of training walls were regrouped into 5 major classes (Fig. 5): (1) concrete  
495 (reinforced); (2) rock piles; (3) gabion meshes; (4) mixed material or (5) natural banks. While  
496 rock piles appear to be affected with similar proportions in all damage categories, concrete  
497 dominates the category with the heaviest observed damage (Fig. 6). Mixed material channel  
498 walls are present in light damage categories (1 and 2) as are gabions. The latter represent a

499 small percentage in damage category 3. When relating the spatial distribution of damaged  
500 material types to the location of bridges, as expected, damage occurs preferentially within  
501 100m of a bridge. Only natural channel banks and mixed material banks were damaged at  
502 greater distances (up to 200m). This observation also confirms that erosion preferentially  
503 occurred downstream of bridges rather than upstream.

504

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

507

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

511

#### 512 **4.3 Flood impact on buildings**

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

517

518 Four damage levels were defined for the building stock (Fig. 6):

519 (1) inundated, without any structural damage; (2) inundated, light damage, building still fit for  
520 habitation after cleaning; (3) inundated, significant damage, rooms livable only after  
521 refurbishment; (4) inundated, heavy damage, structural refitting required. Among the 287  
522 inundated buildings, 144 experienced significant to heavy damage, and 143 buildings were  
523 slightly damaged.

524

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

528

#### 529 4.4 Linking vulnerability parameters to observed damage

530 Each vulnerability parameter was plotted separately on the basis of the respective contingency  
531 table containing the number of levels (Table 1), the number of buildings at each level and their  
532 distribution frequency (Fig. 7). Along the channel, 30 randomly selected city blocks of variable  
533 size (from 531m<sup>2</sup> to 10.57ha), and with rather compact and regular shape, were studied. They  
534 contained 1103 buildings with footprints ranging from 4 to 2,185m<sup>2</sup>; the majority of buildings of  
535 commercial, industrial or agricultural use were larger than 80 m<sup>2</sup> and grouped in size categories  
536 4 and 5, while primarily residential buildings represented about 45% of all those analyzed.  
537 Building density per hectare ranged from 3 to more than 11. However, the majority of the  
538 sampled city blocks were characterized by a relatively low building density (< 6 buildings per  
539 hectare), which was the result of the relatively large footprints of non-residential buildings.

540

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

544

545 In terms of the presence or absence of relationships, the results of the correspondence  
546 analysis show that couples incorporating the parameter “Damage” have very strong  
547 relationships (Fig. 8). Graphic plots show proximities of parameter levels (Fig. 8, DBR5 and  
548 DO4 in the green circle) and oppositions to other levels (Fig. 8, DBR1, DBR2 and DO1 in the  
549 blue circle). For the presented example in figure 9, this data projection suggests that buildings  
550 more than 50m from a bridge behave in a similar way and are less exposed to experiencing  
551 damage than buildings within 5m of a bridge. In our data set, 69.1% of all buildings less than  
552 15m and 39.6% between 15 and 30m from a bridge were damaged.

553

554 *Figure 8: Plot of parameter couple "Distance from bridge" and "Damage" at respective levels.*

555 *Note the strong relationship between buildings located close to a bridge (DBR5) and damage*  
556 *level 4 (DO4; right side of vertical axis) compared to buildings far from a bridge (DBR1 and 2)*  
557 *that have damage level 2 (DO2, left side of vertical axis). Eigenvalues represent 96.54% for*  
558 *axis 1, 2.76% for axis 2, and 0.7% for axis 3.*

559

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

- 564 • Damage versus distance from channel: inundated buildings without damage dominate at  
565 distances beyond 60m of the channel. The closer to the channel, the more intense the  
566 damage, e.g. while at distances of 5 to 10m, damage categories 2 to 4 are almost equally  
567 present, damage category 4 becomes most important at distances lower than 5 m.
- 568 • Damage versus distance from bridge: similar to the previous observations, inundation  
569 without damage occurs preferentially at distances > 90m from bridges. Slight and  
570 significant damage also appear at this distance. Heavy damage is significantly less  
571 important beyond 90m, but still present. Damage category 4 dominates, however, in  
572 distances up to 30m from a bridge.
- 573 • Damage versus structural type: overall, for the data set studied here, damage categories  
574 1, 2 and 3 mostly occur with buildings of type 3 (masonry of terra cotta with reinforced  
575 concrete roof, 1 or 2 storeys). The heaviest damage is preferentially observed in non-  
576 residential buildings.
- 577 • Damage versus inundation height: slight or significant damage is the main consequence  
578 at intermediate inundation heights while maximum inundation is related to the highest  
579 damage.

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

588

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

592

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

597

598 While simple correspondence analysis examines the relationship between two parameters,  
599 multiple correspondence analysis generalizes the comparison by including as many qualitative  
600 parameters as available. Scatter plots in this context graphically represent the relationships  
601 between the different levels of each parameter (Fig. 10); they enable the comparison between  
602 individual buildings, their position among others at certain parameter levels to be determined  
603 and finally reveal behavioral tendencies of building groups showing similar characteristics. The  
604 position of each parameter level is determined by a bagplot (bivariate boxplot), displaying an  
605 ellipse representing 67.5% of buildings within the respective level. Some parameters such as  
606 “Inundation height”, “Impermeability of soil” and “Structural building type” display good sorting,  
607 i.e. buildings of the same characteristic are well grouped. For other parameters, the sorting is



608 much less evident, dispersion is high and overlapping ellipses representing the weighted  
609 center of each parameter group are the result. Hence, one can recognize similarities in the  
610 distribution of parameter levels (from 1 to 5, light blue and red, respectively in Fig. 10): buildings  
611 of level 4 and 5 are preferentially located to the right side, while levels 1 and 2 remain close to  
612 the central axis with a tendency to the left side. Distributions of parameter levels concerning  
613 “distance from channel”, “distance from bridge” and “density of buildings” are observed to be  
614 very close; the same can be seen for the parameters “damage category” and ‘inundation  
615 height”. For these two latter parameters, the scatter plots are very similar. Their relationship is  
616 therefore direct as illustrated previously by the bivariate analysis: the higher the inundation  
617 height, the higher the damage level. Generally, the scatter plots illustrate that buildings located  
618 in the vicinity of the channel and / or a bridge tend to be 1-storey constructions of structural  
619 type 3 to 5. They are commonly located in areas of low building density (category 4 or 5). The  
620 parameters “shape of city block”, “impermeability of soil” and “building footprint” are not directly  
621 related to the previous groups as buildings of varying characteristics, i.e. several parameter  
622 levels, are present on the right side of the plot. The distribution pattern of building points follows  
623 two general tendencies: firstly, a horse-shoe shape where the distribution of buildings  
624 representing a particular parameter level produces an arch pattern spanning from level 1 (light  
625 blue; left of vertical axis) to level 5 (red; right of vertical axis); secondly, a random distribution  
626 to vertical clustering of all with level 4 and 5 (orange and red, respectively) mostly located in  
627 the opposed direction of level 1 to 3 (light blue, dark blue, green; fig. 10).

628

629

630 **Figure 10: Scatter plots representing results of the correspondence analysis. Each point**  
631 **represents a building. Ellipses colored from light blue to red represent parameter levels (1 to**  
632 **5, respectively) as bagplots (bivariate boxplots). Each bagplot represents 67.5% of the**  
633 **buildings defining each level.**

634

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

643

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

650

### 651 **Logistic regression**

652 Logistic regression was applied in order to directly analyze the link between the qualitative  
653 parameter “Damage” and one or more other parameters and to calculate damage probabilities  
654 for all buildings. For this part of the analysis, the parameter “Inundation height” was not  
655 considered, as it was the only data measured after the event and hence strongly related to the  
656 observed damage, which is the parameter to explain. The eight remaining parameters were:  
657 Distance to channel, distance to bridge, shape of city block, impermeability of soil, number of  
658 stories, structural type of building, building footprint and building density. Isolating the most  
659 significant parameters that determine the damage likelihood progressively reduces the number  
660 of vulnerability parameters from 8 to 5. By eliminating non-significant parameters, the  
661 contribution of the maintained parameters is constantly recalculated so that the remaining  
662 significant ones also reflect the non-significant parameters. Consequently, the “Damage”

663 parameter can be expressed as a function of its relationships with the parameters “Distance  
 664 from channel”, “Distance from bridge”, “Shape of city block”, “Structural building type” and  
 665 “Building footprint”. To illustrate the different steps of calculation and associated model outputs,  
 666 we chose one of 200 scenarios that were realized using different calibration and validation  
 667 data sets. Equations (1) and (2) enable us to obtain the respective contributions from the  
 668 different parameters (table 5).

669

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

672

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

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

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

678

679 Values in this equation (5) were taken from table 5 and completed following the logistic  
 680 regression constraint that the sum of coefficients of each parameter must be equal to zero.

681 That is, parameter “Distance to channel” has been completed by DC1, parameter “Distance to  
 682 bridge” by DBR1 and parameter “Building footprint” by A1.

683 We consider for example a building with the following parameter levels: Distance to channel =  
 684 4, Distance to bridge = 5, Shape of city block = 4, Structural building type = 5, building footprint  
 685 = 5. In order to obtain the probability of experiencing damage (e.g. at level 3) for this building,  
 686 we fill in the previous equation using results of table 5 as follows:

687

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

689

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

692

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

694

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

699

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

704

705 As a result, 27.7% of buildings of the validation data set have a 100% chance of experiencing  
706 damage of level 2 (slight damage) or more for an event similar to February 2013. For damage  
707 levels 3 and 4, 6.3 and 9% of buildings, respectively, have a 100% chance of being damaged.  
708 Various maps could then be drawn representing each building in the data set, whether sampled  
709 in the field or not, and their probability of experiencing damage at a certain level (Fig. 12). A  
710 comparison of these results with maps illustrating damage levels assessed during field work  
711 correlates well: the distribution of the damage probability is overall coherent with field  
712 observations of damaged buildings and of the measured flood extent. Results for both  
713 calibration and validation data sets are close, not exceeding a difference of 5% for the number  
714 of buildings attributed to each damage level. The city block examples presented for the six

715 areas within the study (Fig. 12) enable us to highlight several main points: (1) calculating  
716 damage probability by incorporating field data enables expansion of the analysis to buildings  
717 not sampled in the field or for which only some of the required parameters were identified; (2)  
718 calculated results for the validation data classify a lower number of buildings (c. 15%) in the  
719 predicted damage category “2 or more”, which is equivalent to the observed damage level 2,  
720 3 and 4, than observed in reality (28%). This is mainly due to the fact that in reality, parameters  
721 not considered in our analysis also seem to be important; these include the height of channel  
722 retaining walls (Fig. 12, zone 2), the presence of increased surface run-off not coming from  
723 the channel (Fig. 12, zone 4), the height and width of bridge openings, the building position  
724 upstream or downstream from a bridge, etc.; (3) especially for buildings not affected in 2013  
725 (cf. Fig. 12, zone 2, 5 and 6), the calculation of damage probability allows us to identify and  
726 visualize graphically areas where mitigation measures such as refitting may be necessary to  
727 avoid serious damage during future events; and (4), although rare, some buildings with  
728 significant damage observed in the field (Fig. 12, zone 1 and 5) appear to have lower calculated  
729 damage probabilities than expected.

730

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

736

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

740

741 Finally, concerning the model performance, one scenario with a success rate of 74% was  
742 chosen to illustrate the methodology. The model was applied in total 200 times on different

743 calibration and validation data sets in order to examine general performance. Results show  
744 that 90% of these tests have a success rate of more than 67%.

745

## 746 **5. Discussion**

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

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

770 therefore eliminated from the analysis. “Structural type” therefore only represents the  
771 construction material in this context.

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

778  
779 Interpretation of the results requires careful comparison of the links between several  
780 parameters, particularly if no previous knowledge from real field conditions is available. A few  
781 aspects brought to light by scatter plots and logistic regression are pointed in the following.

782

783

784

785

786

787 Scatter plots

- 788 • In this study, building density was calculated as the number of buildings per unit area.  
789 High building density in Arequipa generally corresponds to dense habitat where  
790 buildings share one or more walls, reducing the space between them and thus reducing  
791 the risk of inundation in back rows. Lower building density, in particular for residential  
792 buildings, implies gaps between buildings, which creates hydraulic roughness and  
793 resistance to flow. This reduces flow velocities on buildings in back rows, but potentially  
794 increases them in front rows. This fits to observations from the scatter plots showing  
795 that city blocks of lower density exhibit more damaged buildings than those of higher  
796 density. However, field observations suggest another reason for increased damage in  
797 low-density areas: larger footprints of industrial, commercial and agricultural buildings

798 that, especially for the latter, have more vulnerable structural characteristics than  
799 residential buildings. To clarify this possibility may require improvements in how  
800 building density is assessed in relation to the structural building type and/or its use.

- 801 • Comparing building footprint and density of buildings, it appears that the largest  
802 buildings (mainly agricultural or industrial, and more rarely commercial) are: (i) more  
803 vulnerable than smaller buildings because of their very larger openings, and fewer load-  
804 carrying structures such as columns, etc.; and (ii) typically located in city blocks with  
805 low building density, which additionally increases the probability of being damaged  
806 since the screening effect from adjacent buildings is absent. For footprint categories 1  
807 and 2 (small size), the building density per city block frequently remains low enough to  
808 potentially put them in damage category 4.
- 809 • Relating damage level 3 (significant damage) to the structural type of buildings, it  
810 becomes obvious that buildings of all types can be damaged, but in this flood event,  
811 mostly buildings of category 2 (masonry of terra cotta or ignimbrite with mortar and  
812 metal sheet roof, 1 storey) and 3 (masonry of terra cotta with reinforced concrete roof,  
813 1 or 2 storeys) were affected; this coincides also with the highest inundation measured  
814 and with the smallest distance from the channel and bridges. However, damage  
815 categories 2 and 3 group buildings of structural types 6, 7, 8 and 4, 5, respectively.  
816 These are all structural types of higher quality, which are estimated to be less  
817 vulnerable than structural types 1, 2, 3 grouped in damage category 4. Previous studies  
818 regarding physical vulnerability of buildings related to landslide and debris flow hazard  
819 in Austria ([Fuchs et al., 2007](#); [Papathoma Köhle et al., 2012](#)), Germany ([Kaynia et al.,](#)  
820 [2008](#)), Italy ([Aleotti et al., 2004](#); [Luino, 2005](#); [Galli and Guzzetti, 2007](#)) and the United  
821 States ([FEMA: HAZUS-MH, 2010](#)), seem to confirm an increasing vulnerability with  
822 decreasing construction quality (either due to construction material or structural  
823 characteristics). We therefore deduce for our results that the relationship statistically  
824 calculated to be strong between structural type and damage is, in this particular case,



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

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

838

839 Logistic regression

840 Logistic regression in addition to the scatter plots, reflects well the issue of an unbalanced data  
841 set: buildings without any damage are largely over-represented in our data. In the applied  
842 mathematical analysis, this group gains in weight because of their high number leading to  
843 calculated damage likelihoods that are greater than those derived from expert judgment. For  
844 example, the logistic regression analysis favors buildings without damage as a way to rank  
845 buildings of higher quality (structural types 4 to 7) to be more likely to experience damage level  
846 3 than buildings supposedly more vulnerable (structural types 1-3). This should improve by  
847 incorporating additional damage data from future potential flood events, which will lead to a  
848 more balanced data set.

849

850 Scope and limits

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

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

873

## 874 **6. Conclusion and perspectives**

875 This study was carried out following a damage survey in the field early after a flash flood event  
876 in February 2013. Observed damage intensities were overall low and only few buildings  
877 suffered serious damage. Results from the proposed statistical data analysis validate the  
878 method as an operational tool to calculate damage probabilities for a flash flood event of similar  
879 intensity. However, the lack of damage documentation, in particular for highest damage  
880 categories, is at present a constraint to further develop the model for a larger range of hazard

881 types and magnitudes. Ideally, the method will be tested on a database including many case  
882 studies. This would allow validation of the methodology not only for small scale, site-specific  
883 analysis, but also for broad scale generalized assessment of damage probability for different  
884 hazard types and magnitudes.

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

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

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

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

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

909

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

918

### 919 **Acknowledgements**

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

925

926 **References**

927

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

932

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

937

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

940

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

944

945 Cacya, L., Meza, P., Carlotto, V., Mamani, L., 2013. Aluvión del 8 de Febrero del 2013 en la  
946 ciudad de Arequipa. Foro internacional Peligros geológicos, Arequipa, 14-16 Octubre  
947 2013, GRM, OVI Ingemmet, Libro de Resúmenes, pp. 195-203.

948

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

953

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

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

957

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

960

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

963

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

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

968

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

971

972 Ettinger, S., Manrique Llerena, N., Talibart, C., Mounaud, L., Yao-Lafourcade, A.-F., Thouret,  
973 J.-C., 2014a. Using HSR imagery to downscale vulnerability assessment of buildings  
974 and local infrastructure facing hazards from floods and hyperconcentrated flows.  
975 *Geophysical Research Abstracts*, vol. 16, EGU2014-13978, EGU General Assembly  
976 2014.

977

978 Ettinger, S., Zeghdoudi, M., Manrique Llerena, N., Yao-Lafourcade, A.-F., Thouret, J.-C.,  
979 2014b. L'apport de l'imagerie Pléiades à la cartographie des enjeux et de leur

980           vulnérabilité face aux crues torrentielles : La ville d'Arequipa, Pérou. *Revue Française*  
981           de Photogrammétrie et de Télédétection. In Press, 7 p.

982

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

986

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

989

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

993

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

996

997   Gaume, E., Bain, V., Bernardara, P., Newinger, O., Barbuc, M., Bateman, A., Blaškovičová,  
998           L., Blöschl, G., Borga, M., Dumitrescu, A., Daliakopoulos, I., Garcia, J., Irimescu, A.,  
999           Kohnova, S., Koutroulis, A., Marchi, L., Matreata, S., Medina, V., Preciso, E., Sempere-  
1000           Torres, D., Stancalie, G., Szolgay, J., Tsanis, I., Velasco, D., Viglione, A., 2009. A  
1001           compilation of data on European flash floods. *Journal of Hydrology* 367 (1-2), 70-78.

1002

1003   Glade, T., 2003. Vulnerability assessment in landslide risk analysis. *Die Erde* 134, 123–146.

1004

1005   Graf, W., 1988. *Fluvial processes in dryland rivers*. Springer, Berlin, 329 p.

1006

1007 Gruntfest, E., Handmer, J., 2001. Dealing with flash floods: contemporary issues and futures  
1008 possibilities, in: Gruntfest, E., Handmer, J. (Eds.), *Coping with Flash Floods*. Kluwer  
1009 Academic Publishers, Dordrecht, pp. 3-10.

1010

1011 Gruntfest, E., 2009. Editorial. *Journal of Flood Risk Management* 2, 83-84.

1012

1013 HAZUS-MH, 2010. HAZUS – FEMAs Methodology for Estimating Potential Losses from  
1014 Disasters, available at: <http://msc.fema.gov/portal/resources/hazus>.

1015

1016 Hiete, M., Merz, M., 2009. An Indicator Framework to Assess the Vulnerability of Industrial  
1017 Sectors against Indirect Disaster Losses, Proceedings of the 6<sup>th</sup> International ISCRAM  
1018 conference, Gothenburg, Sweden, May 2009, Paper no 131.

1019

1020 Holub, M., Fuchs, S., 2008. Benefits of local structural protection to mitigate torrent-related  
1021 hazards, in: Brebbia, C., Beriatos, E. (Eds.), *Risk analysis VI. WIT transactions on*  
1022 *information and communication technologies* 39. WIT, Southampton, pp. 401-411.

1023

1024 Holub, M., Suda, J., Fuchs, S., 2012. Mountain hazards: reducing vulnerability by adapted  
1025 building design. *Environmental Earth Sciences* 66, 1853-1870.

1026

1027 Hooke, J.M., Mant, J.M., 2000. Geomorphological impacts of a flood event on ephemeral  
1028 channels in SE Spain. *Geomorphology* 34 (3-4), 163-180.

1029

1030 INDECI, 2013. Estado Situacional de la Emergencia: Informe sobre los eventos  
1031 ocasionados por las intensas lluvias del mes de febrero 2013. Base de datos del  
1032 Instituto Nacional de Defensa Civil, Arequipa, Peru.

1033

1034 IPCC, 2007. Summary for policymakers, in: Solomon, S., Qin, D., Manning, M., Chen, Z.,



1035 Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.) Climate change 2007: the  
1036 physical science basis, Contribution of Working Group I to the 4th Assessment Report  
1037 of the Intergovernmental Panel on Climate Change, Cambridge University Press,  
1038 Cambridge  
1039  
1040 IPCC Working Group I, 2007. Contribution of Working group I to the 4th Assessment Report  
1041 of the Intergovernmental Panel on Climate Change 2007, Cambridge University Press,  
1042 Cambridge, p. 996  
1043  
1044 Jakob, M., Stein, D., Ulmi, M., 2012. Vulnerability of buildings to debris flow impact. Natural  
1045 Hazards 60, 241-261.  
1046  
1047 Jenkins, S.F., Spence, R.J.S., Fonseca, J.F.B.D., Solidum, R.U., Wilson, T.M., 2014.  
1048 Volcanic risk assessment: Quantifying physical vulnerability in the built environment.  
1049 Journal of Volcanology and Geothermal Research 276, 105-120.  
1050 Doi.org/10.1016/j.jvolgeores.2014.03.002  
1051  
1052 Jonkman, S.N., Vrijling, J.K., 2008. Loss of life due to floods. Journal of Flood Risk  
1053 Management 1, 43-56. Doi:10.1111/j.1753-318x.2008.00006.x.  
1054  
1055 Kaplan, S., Garrick, B., 1981. On the Quantitative Definition of Risk. Risk Analysis 1, 11–27.  
1056  
1057 Kaynia, A.M., Papathoma Köhle, M., Neuhäuser, B., Ratzinger, K., Wenzel, H., Medina-  
1058 Cetina, Z., 2008. Probabilistic assessment of vulnerability to landslide : Application to  
1059 the village of Lichtenstein, Baden-Württemberg, Germany. Engineering Geology 101(1-  
1060 2), 33-48.  
1061

1062 Lo, W.-C., Tsao, T.-C., Hsu, C.-H., 2012. Building vulnerability to debris flows in Taiwan: a  
1063 preliminary study. *Natural Hazards* 64, 2107-2128.  
1064

1065 Luino, F., 2005. Sequence of instability processes triggered by heavy rainfall in the northern  
1066 Italy. *Geomorphology* 66(1–4), 13–39.  
1067

1068 Marchi, L., Borga, M., Preciso, E., Gaume, E., 2010. Characterisation of selected extreme  
1069 flash floods in Europe and implications for flood risk management. *Journal of Hydrology*  
1070 394 (1-2), 118-133. Doi:10.1016/j.jhydrol.2010.07.017.  
1071

1072 Martelli, K., 2011. The physical vulnerability of urban areas facing the threat of inundation  
1073 from lahars and flash floods: application to the case study of Arequipa, Peru.  
1074 (Unpublished) PhD dissertation, Université Blaise Pascal, Clermont-Ferrand, 245 p.  
1075

1076 Martínez Ibarra, E., 2012. A geographical approach to post-flood analysis: The extreme flood  
1077 event of 12 October 2007 in Calpe (Spain). *Applied Geography* 32, 490-500.  
1078

1079 Montz, B.E., Grunfest, E., 2002. Flash flood mitigation: recommendations for research and  
1080 applications. *Environmental Hazards* 4, 15-22.  
1081

1082 Oehler, J.-F., Thouret, J.-C., Solikhin, A., Ettinger, S., 2014. High resolution DEMs based on  
1083 HSR PLEIADES images: Applications to the Merapi volcano (Indonesia) and the city of  
1084 Arequipa close to the Misti volcano (Peru). *Cities on Volcanoes 8 meeting*, 9-13  
1085 September, Yogyakarta, Indonesia. Abstract and poster.  
1086

1087 Papathoma-Köhle, M., Kappes, M., Keiler, M., Glade, T., 2011. Physical vulnerability  
1088 assessment for alpine hazards: state of the art and future needs. *Natural Hazards*  
1089 58(2), 645-680.

1090

1091 Papathoma-Köhle, M., Keiler, M., Totschnig, R., Glade, T., 2012. Improvement of  
1092 vulnerability curves using data from extreme events: debris flow event in South Tyrol.  
1093 *Natural Hazards* 64, 2083-2105.

1094

1095 Quan Luna, B., Blahut, J., Van Westen, C.J., Sterlacchini, S., Van Asch, T.W.J., Akbas, S.O.,  
1096 2011. The application of numerical debris flow modelling for the generation of physical  
1097 vulnerability curves. *Natural Hazards and Earth System Science* 11: 2047–2060.

1098

1099 Reid, I., Frostick, L.E., 1992. Channel form, flows and sediments in deserts, in: Thomas,  
1100 D.S.G. (Ed.), *Arid Zone Geomorphology*, Belhaven Press, Rotterdam, pp. 117-135.

1101

1102 Santoni, O., 2011. Etude de la vulnérabilité de la ville d'Arequipa, Pérou, vis à vis du volcan  
1103 1049 El Misti, au moyen d'un SIG. M.Sc. Thesis, Université Blaise Pascal, Clermont-  
1104 Ferrand, 51 p (in French).

1105

1106 SENAMHI, 2013. Periodo de lluvias 2013 en la región Arequipa y su relación con el Cambio  
1107 Climático. Foro Regional Desastre, Reconstrucción y desarrollo en escenario de cambio  
1108 climático Arequipa – emergencia 2013.

1109

1110 Thouret, J.-C., Ettinger, S., Guitton, M., Martelli, K., Santoni, O., Magill, C., Revilla, V., Charca,  
1111 J.A., 2014. Assessing physical vulnerability and risk in large cities exposed to debris  
1112 flows and flash floods: the Arequipa (Peru) case study. *Natural Hazards*, DOI  
1113 10.1007/s11069-014-1172-x, 47 p.

1114

1115 Thouret, J.C., Enjolras G., Martelli, K., Santoni, O., Luque J.A., Nagata, M., Arguedas, A.,  
1116 Macedo, L., 2013. Combining criteria for delineating lahar- and flash-flood-prone

1117 hazard and risk zones for the city of Arequipa, Peru. *Nat. Hazards Earth Syst. Sci.* 13,  
1118 339-360.

1119

1120 Thywissen, K., 2006. Components of risk. A comparative glossary. Published as SOURCE  
1121 No. 2—2006 of the Institute of Environment and Human Security of the United Nations  
1122 University. Available at: <http://www.ehs.unu.edu/file/get/4042>

1123

1124 Totschnig, R., Fuchs, S., 2013. Mountain torrents: Quantifying vulnerability and assessing  
1125 uncertainties. *Engineering Geology* 155, 31-44.

1126

1127 Totschnig, R., Sedlacek, W., Fuchs, S., 2011. A quantitative vulnerability function for fluvial  
1128 sediment transport. *Natural Hazards* 58, 681-703.

1129

1130 Villagran de Leon, J.C., 2006. Vulnerability - A Conceptual and Methodological Review,  
1131 SOURCE - Publication Series of UNU-EHS n°4/2006, United Nations University UNU-  
1132 EHS, UN Campus, Bonn, Germany, 90 p.

1133

1134 Wigmosta, M.S., 1983. Rheology and flow dynamics of the Toutle debris flows from Mt. St.  
1135 Helens. M.Sc. Thesis, University of Washington, Seattle, 184 p.

1136

1137 Wisner, B., Blaikie, P., Cannon, T., Davis, I., 2004. At risk: Natural hazards, people's  
1138 vulnerability, and disasters, 2nd edn. Routledge, London

1139

1140 Zuccaro, G., Ianniello, D., 2004. Interaction of pyroclastic flows with building structures in an  
1141 urban settlement: a fluid-dynamic simulation impact model. *Journal of Volcanology and*  
1142 *Geothermal Research* 133, 345-352.

1143

1144 Zuccaro, G., Cacace, F., Spence, R.J.S., Baxter, P.J., 2008. Impact of explosive eruption  
1145 scenarios at Vesuvius. *Journal of Volcanology and Geothermal Research* 178, 416-  
1146 453.

1147

1148 Zuccaro, G., De Gregorio, D., 2013. Time and space dependency in impact damage  
1149 evaluation of a sub-Plinian eruption at Mount Vesuvius. *Natural Hazards* 68(3), 1399-  
1150 1423.

1151

1152

1153

1154

1155

1156

1157 **Table caption**

1158 Table 1: Vulnerability parameters concerning building characteristics, building environment  
1159 and flood hazard with their respective levels as defined for this study.

1160 Table 2: Material types characterizing channel banks along the left and right riverside of the  
1161 Avenida Venezuela channel. Numbers indicate the frequency distribution of channel  
1162 sections in each material category and their respective distance to a bridge. The total  
1163 number of sections located at a certain distance from a bridge is illustrated in bold  
1164 with, to its right, the corresponding percentage of total channel bank length.

1165 Table 3: Section lengths of channel banks as a function of construction material and location  
1166 on left and right riverside.

1167 Table 4: Damaged channel bank sections represented as the percentage of the total length  
1168 of either the left or right channel side. Sections are attributed to one of six groups (A  
1169 to F) depending on the closest distance to a bridge of either the start or end point of  
1170 the section.

1171 Table 5: Contributions of each parameter level to damage probability based on the  
1172 calibration data set (598 buildings).

1173

1174 **Figure caption (black and white reproduction in print is intended)**

1175 Figure 1A: Geographical setting and location of Arequipa city, Peru. B: The study area  
1176 Avenida Venezuela channel and six zones that will serve to illustrate observations in  
1177 the following.

1178 Figure 2: Longitudinal channel profile (black line) with channel width (green dots), and  
1179 sections in which erosion occurred (orange bars). The gray scale bar represents  
1180 channel bed type, i.e. natural gravel, sand (dark gray), natural with occasional  
1181 concrete steps (white) and concrete (medium gray); the channel wall material is  
1182 represented by concrete (red), mixed material (concrete, brick, boulders; yellow) and  
1183 natural (blue). For complementary information see also table 2 to 4.

1184 Figure 3: Field-survey-based mapping of inundation extent resulting from overbank flow  
1185 along the Avenida Venezuela channel.

1186 Figure 4: Three examples of particular channel courses and resulting damage.

1187 Figure 5 Left: Damage level observed for different material types of retaining walls. Right:  
1188 Material types of retaining walls relative to the proximity of bridges.

1189 Figure 6: Observed damage levels from left to right (4) inundated, heavy damage, (3)  
1190 inundated, significant damage; (2) inundated, light damage; and (1) inundated, no  
1191 structural damage.

1192 Figure 7: Results of univariate analysis summarizing the number of buildings per category.  
1193 Grayscales from the lowest parameter level 1 (white) to the highest level 5 (dark gray)  
1194 are the same for all figures.

1195 Figure 8: Plot of parameter couple *“Distance from bridge”* and *“Damage”* at respective levels.  
1196 Note the strong relationship between buildings located close to a bridge (DBR5) and  
1197 damage level 4 (DO4; right side of vertical axis) compared to buildings far from a  
1198 bridge (DBR1 and 2) that have damage level 2 (DO2, left side of vertical axis).  
1199 Eigenvalues represent 96.54% for axis 1, 2.76% for axis 2, and 0.7% for axis 3.

1200 Figure 9: Results of the bivariate analysis. Damage level is displayed in different gray  
1201 shades, the abscissa (1 to 5) displays the categories of the respective parameter  
1202 *“Distance from channel”, “Distance from bridge”, etc.*

1203 Figure 10: Scatter plots representing results of the correspondence analysis. Each point  
1204 represents a building. Ellipses colored from light blue to red represent parameter  
1205 levels (1 to 5, respectively) as bagplots (bivariate boxplots). Each bagplot represents  
1206 67.5% of the buildings defining each level.

1207 Figure 11: Projection of parameter levels (color) and city blocks as a result of the bivariate  
1208 analysis. The position of each square is defined by the bagplot representing 67.5% of  
1209 the buildings defining each level. The circles indicate city blocks of similar  
1210 characteristics and thus behavior. City block numbers are plotted to allow comparison

1211 but are not included in the bivariate analysis. The number of buildings per city block  
1212 and the respective percentage is detailed in the histogram to the right.

1213 Figure 12: Damage probabilities calculated for damage levels 1,2, 3 and 4 (series A) and all  
1214 observed damage levels (1 to 4) in the field (series B) using calibration and validation  
1215 data sets (898 buildings) of the selected test scenario presented in the manuscript.

1216

1217 **APPENDIX caption**

1218 Appendix A and B: Survey forms for the damage assessment of buildings, conceived for  
1219 masonry (A) and reinforced concrete (B) structures following experiences from  
1220 previous studies concerning natural hazard impact (Zuccaro et al., 2008; Zuccaro and  
1221 De Gregorio, 2013; Jenkins et al., 2014).

1222

1223

1224

1225

1226

1227

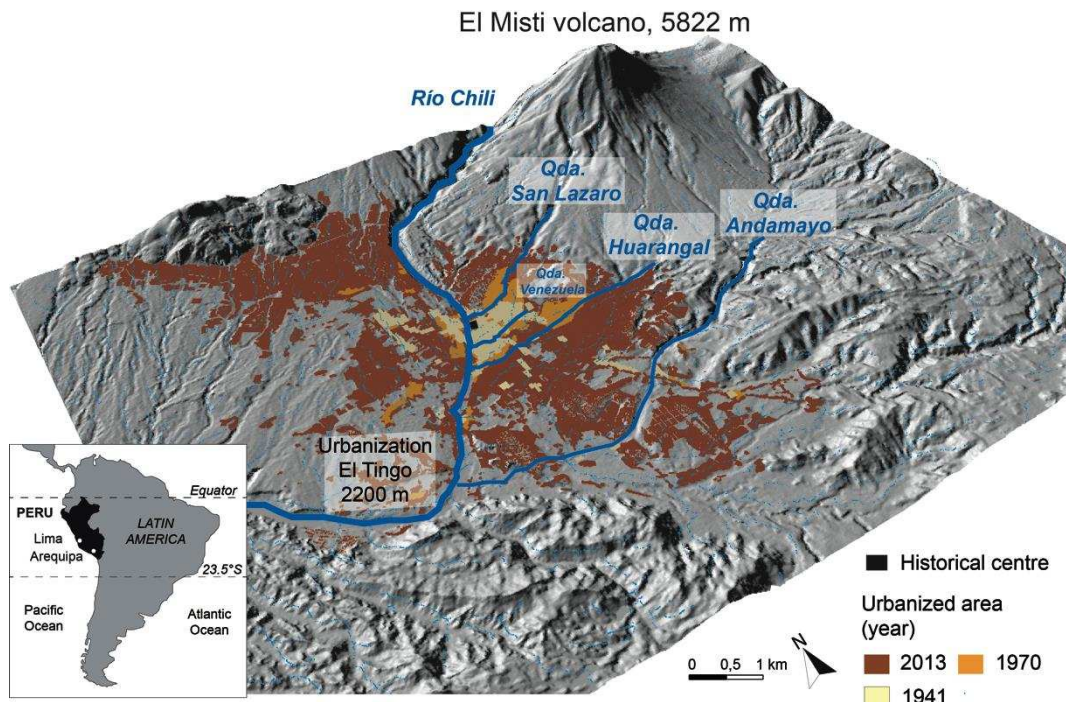
1228

1229

1230

1231





1232  
1233  
1234  
1235

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

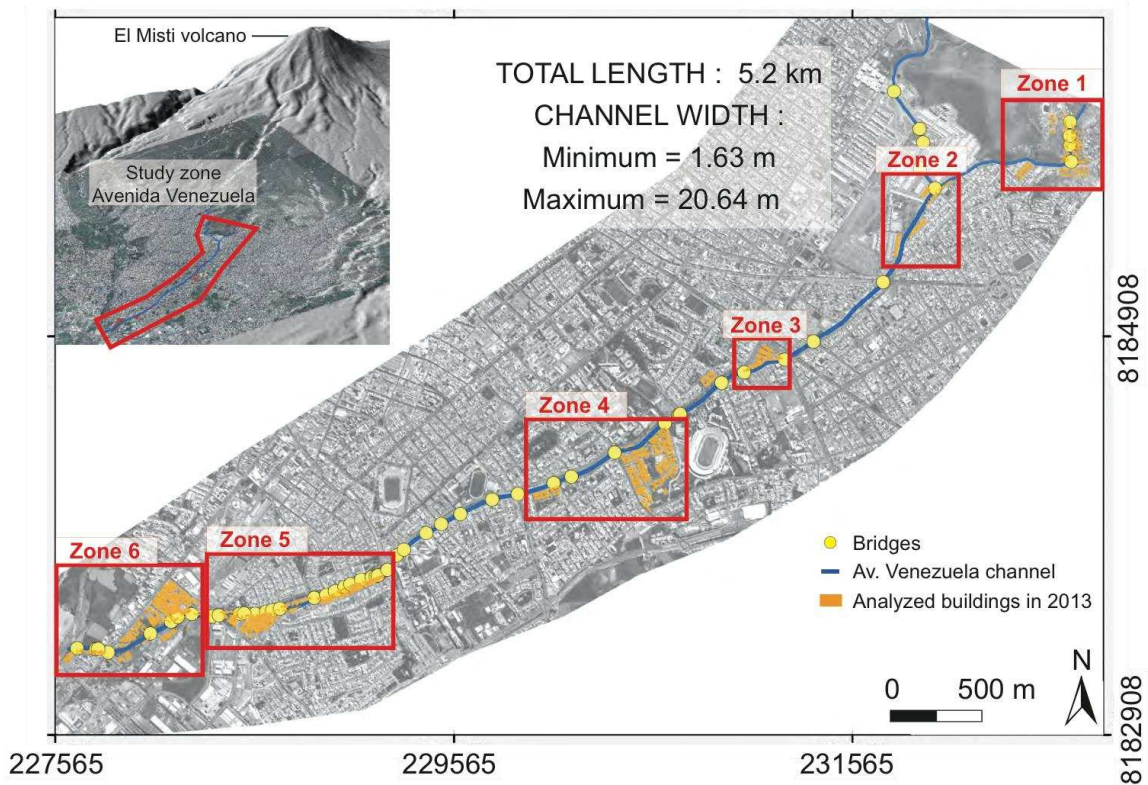


Figure 2: The study area Avenida Venezuela channel and six zones that will serve to illustrate observations in the following.

1236  
1237  
1238  
1239

1240

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

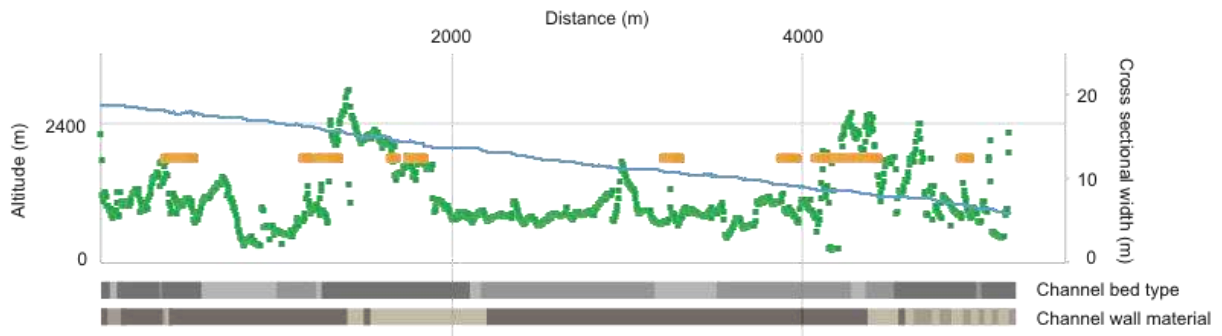
1241

1242

1243

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

1244



1245

1246

1247

1248

1249

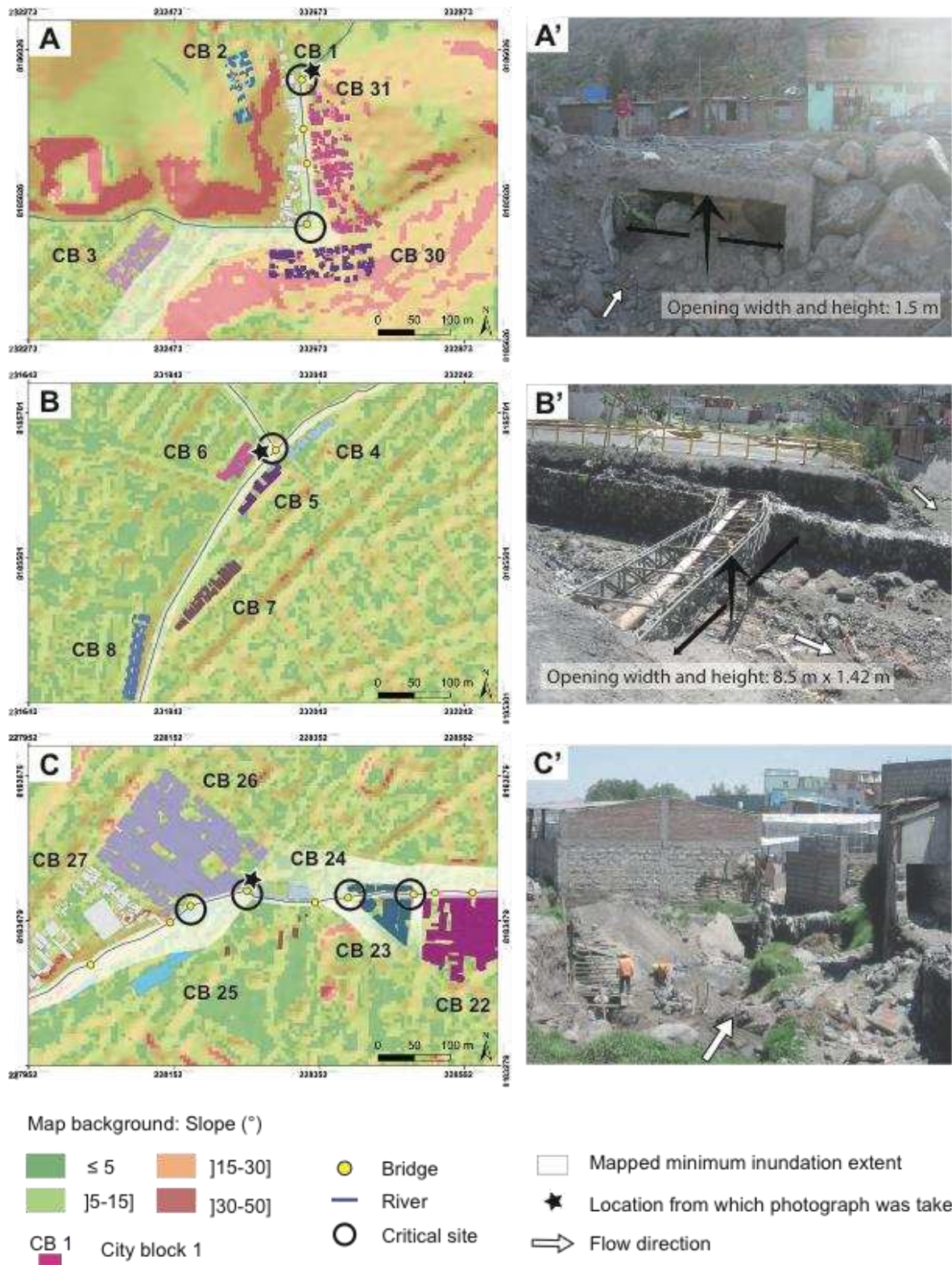
1250

1251

1252

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).

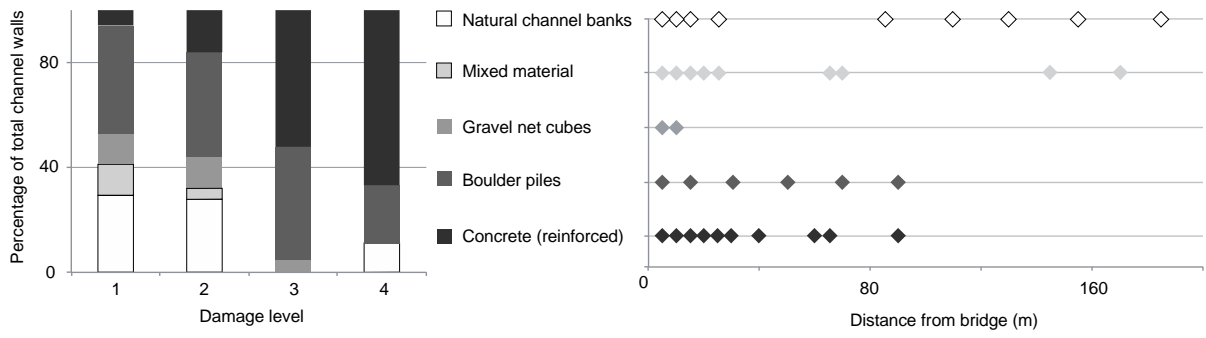




1253  
1254  
1255

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

1256



1257

1258

1259

1260

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.

1261



1262

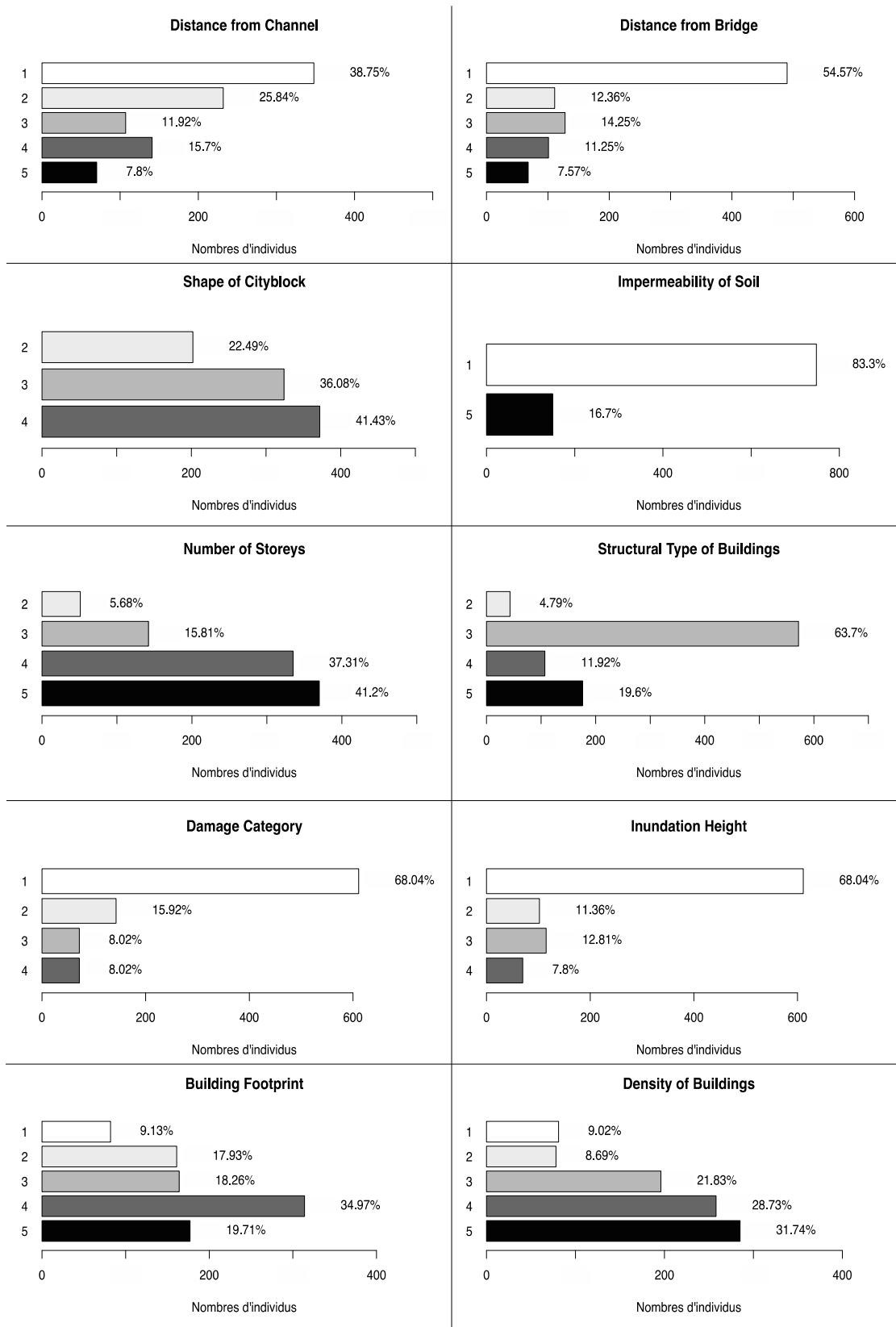
1263

1264

1265

1266

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.



1267  
 1268  
 1269  
 1270  
 1271

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).



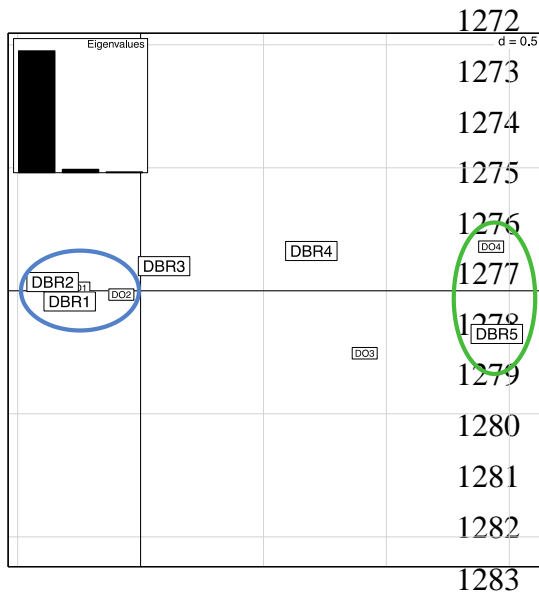
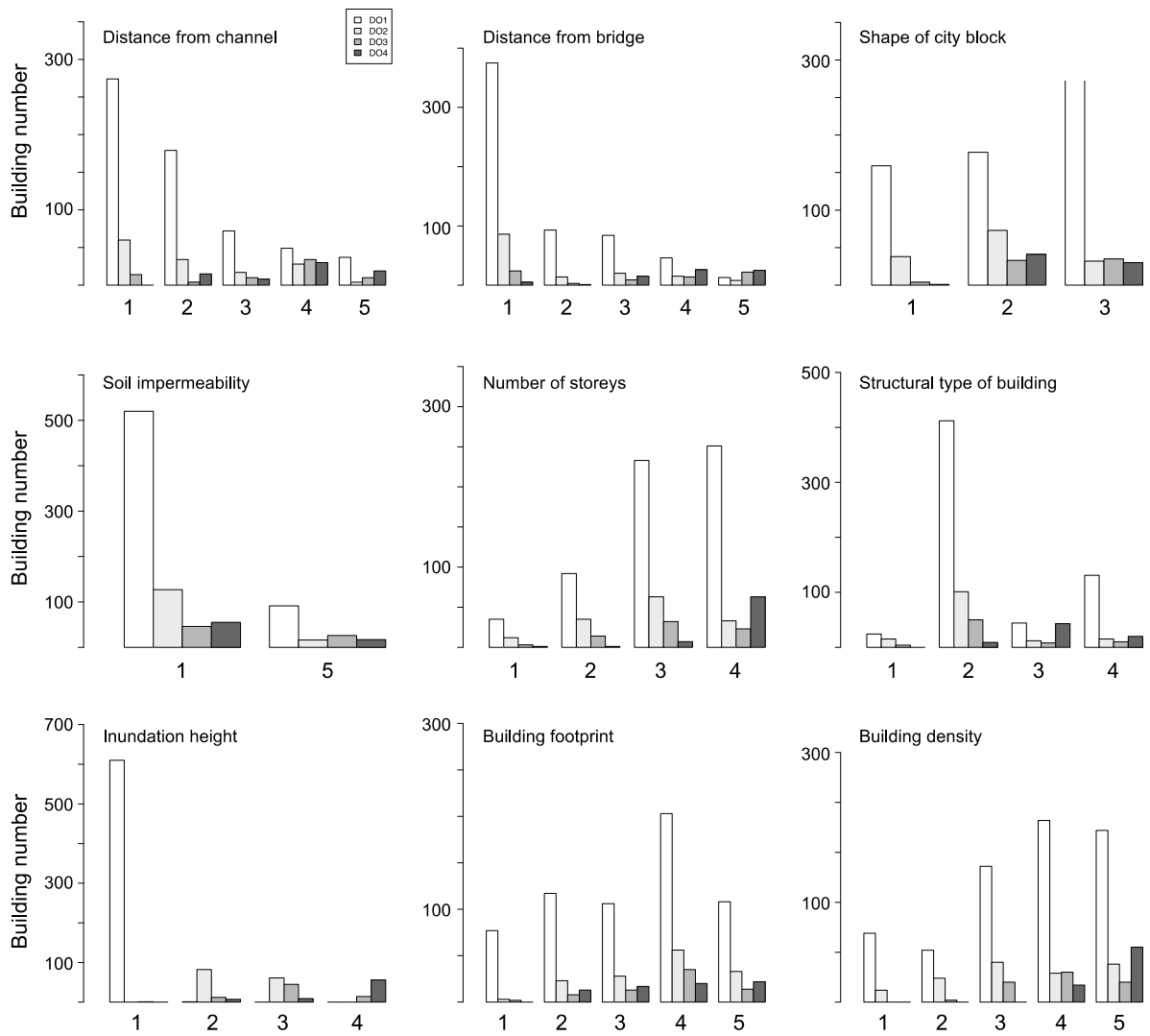


Figure 9: Plot of parameter couple “Distance from bridge” and “Damage” with respective levels. Note the strong relationship between buildings located close to a bridge (DBR5) and damage level 4 (DO4; right side of vertical axis) opposed to buildings far from a bridge (DBR1 and 2) that have damage level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



1284

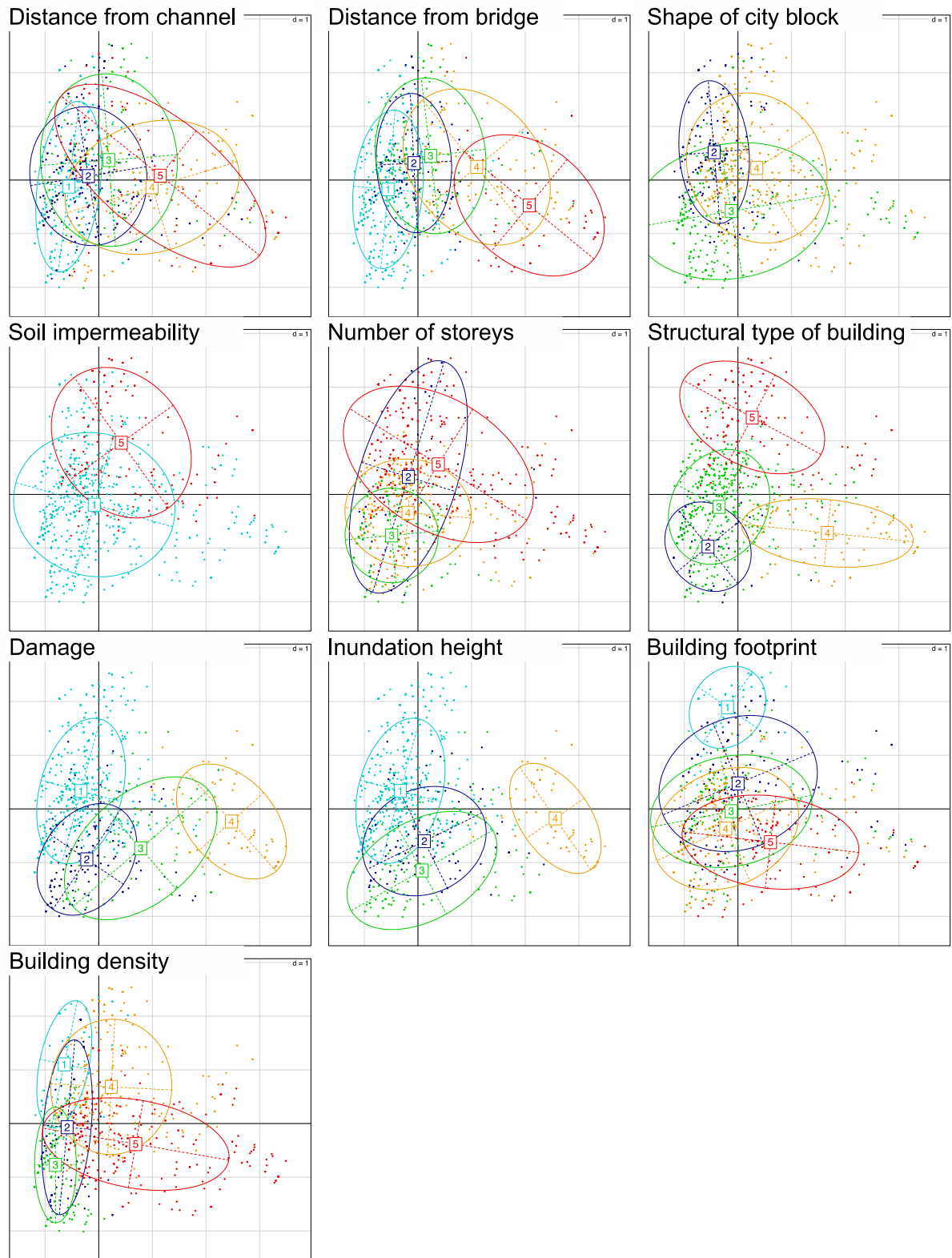
1285

1286

1287

1288

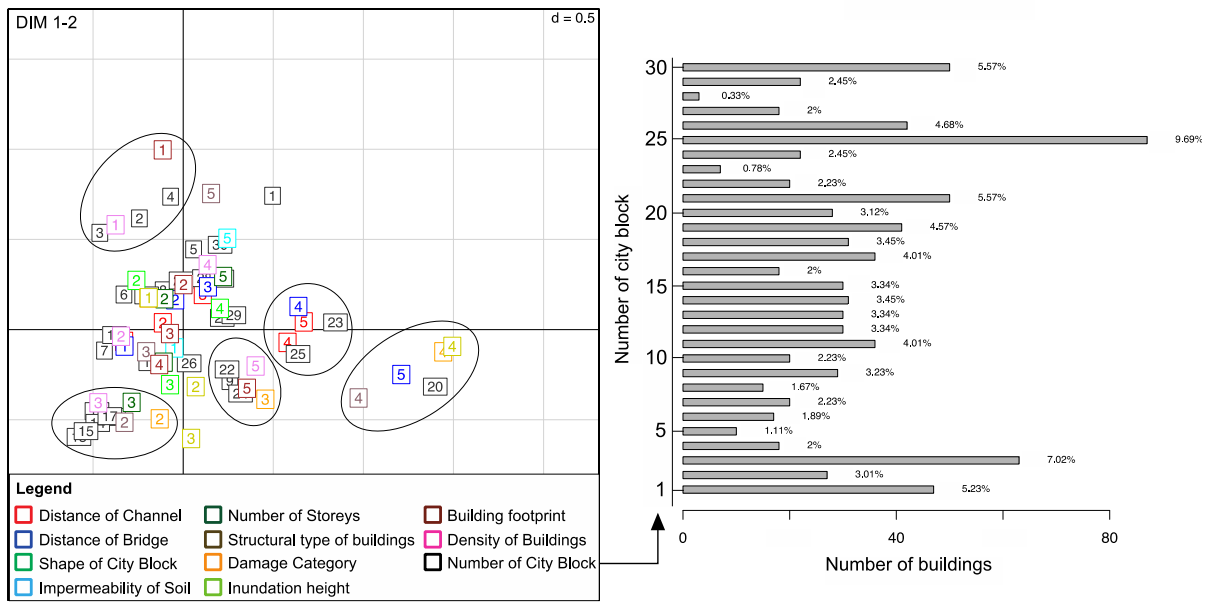
Figure 10: Results of the bivariate analysis. Damage level is displayed in different gray shades, the abscissa (1 to 5) displays the categories of the respective parameter “Distance from channel”, “Distance from bridge”, etc.



1289  
 1290  
 1291  
 1292  
 1293

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.

1294



1295

1296

1297

1298

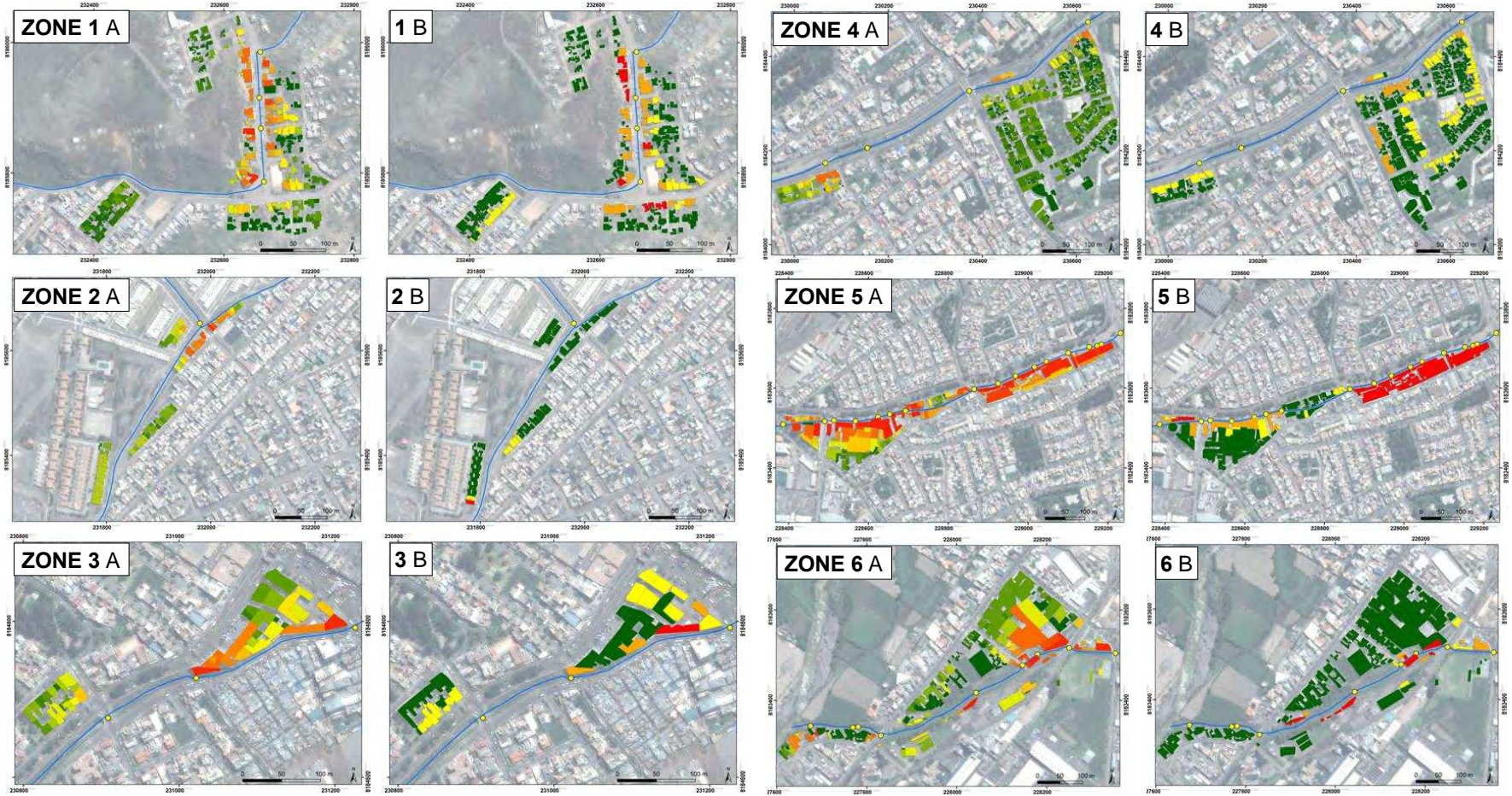
1299

1300

1301

1302

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.

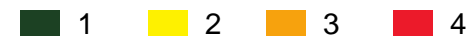


**Damage level**

Series A - Calculated damage likelihood (%)



Series B - Observed damage level in the field



1304 Figure 13: Damage probabilities calculated for damage level 3 and 4 (series A) and all  
1305 observed damage levels (1 to 4) in the field (series B).  
1306