Damage assessment in Braunsbach 2016: A data collection and analysis for an improved understanding of damaging processes during flash floods

Jonas Laudan[1], Viktor Rözer[2], Tobias Sieg[1/2], Kristin Vogel[1], Annegret H. Thieken[1]

[1]University of Potsdam, Institute of Earth and Environmental Science, Karl-Liebknecht-Strasse 24-25, 14476 Potsdam, Germany
[2]GFZ German Research Centre for Geosciences, Department of Hydrology, Telegrafenberg, 14473 Potsdam, Germany

Correspondence to: Jonas Laudan (jlaudan@uni-potsdam.de)

Abstract. Flash floods are caused by intense rainfall events and represent an insufficiently understood phenomenon in Germany. As an effect of higher precipitation intensities, flash floods might occur more frequently in future. In combination with changing land use patterns and urbanisation, damage mitigation, insurance and risk management in flash flood prone regions are becoming increasingly important. However, a better understanding of damage caused by flash floods requires ex-post collection of relevant but yet sparsely available information for research. At the end of May in 2016, very high and concentrated rainfall intensities led to severe flash floods in several south German municipalities, of which the small town Braunsbach stood as a prime example for the devastating potential of such events. Eight to ten days after the flash flood event, a damage assessment and data collection was conducted in Braunsbach by investigating about all affected buildings and their surroundings. To record and store the data on site, the open source software bundle "KoBoCollect" was used as an efficient and easy way to gather information. Since the damage driving factors of flash floods are expected to differ from those of riverine flooding, a post-hoc data analysis was performed, aiming to identify the influence of flood processes and building attributes on damage grades. The analyses include the application of Random Forest, a Random General Linear Model and multinomial logistic regression as well as the construction of a local impact map to reveal influences on the damage grades. Further, a Spearman's Rho correlation matrix was calculated. The results reveal that the damage driving factors of flash floods differ from those of riverine floods to a certain extent. Especially the exposition of a building in flow direction shows a strong correlation with the damage grade and has a high predictive power within the constructed damage models. Additionally, the results suggest that building material as well as various building aspects such as the existence of a shop window and the surrounding might have an effect on the resulting damage. To verify and confirm the outcomes as well as supporting future mitigation strategies, risk management and planning, more comprehensive and systematic data collection is necessary.

1 Introduction

Flooding is a common hazard in Central Europe, resulting in high economic losses (Munich Re, 2016). To promote and tailor local planning, flood risk management policies such as the European Floods Directive (2007/60/EC) set up framework conditions for member states to implement flood risk management on a national, regional and local level. Risk assessments
5 and policy decisions are expected to take different flood types into account (e.g. coastal floods, riverine floods, pluvial floods, flash floods), according to the local circumstances (BMUB, 2007). In Germany for instance, storm surge and river flooding are dominant and were therefore considered as risks of national importance. Due to recent severe riverine flooding in eastern and southern parts of Germany, particularly in August 2002 and June 2013, the flood risk management system and the relevant legislation have been substantially improved, among others, according to the EU Floods Directive (2007/60/EC)
10 and its implementation in the Federal Water Act of 2009 (e.g. Thieken et al., 2016). Supportingly, a large body of literature exists that addresses the topic of riverine flooding in Germany and its effects as well as demands on people, policy makers and general planning. In this regard, risk assessment strategies and effects of preparedness decisions are presented and extensively discussed with a strong focus on recent major riverine flood events. (e.g. Bubeck et al., 2013; Kienzler et al., 2015; Bracken et al., 2016; Osberghaus et al., 2016; Thieken et al., 2016; Kundzewicz et al., 2017). However, when
15 implementing the EU Floods Directive in Germany, surface water flooding and flash floods were not considered as risks of national importance and thus neglected. This assessment is currently questioned due to the destructive events in May and June 2016.

Intense rainfall at the end of May and beginning of June 2016 over Central Europe led to severe surface water flooding and flash floods, which were partly accompanied by mud and debris. Several municipalities mainly in the south of Germany
20 were hit, eleven people lost their life; infrastructure and buildings were heavily damaged (GDV, 2016). The insured losses of these events amounted to EUR 1.2 billion (GDV, 2016) and the overall loss was estimated to EUR 2.6 billion (Munich Re 2017), an extraordinary monetary loss caused by flash floods in Germany. The district of Schwäbisch Hall in Baden-Württemberg was particularly affected. Moreover, in the beginning of June 2016 the municipality of Rottal-Inn in southern Bavaria was hit by flash flooding, triggered by the same weather situation (GDV, 2016).
25 Especially a small village in Schwäbisch Hall, named Braunsbach, faced severe flash floods on May 29th, causing high damage on buildings and infrastructure. Braunsbach is counting about 1000 residents, yet due to the devastating character and abruptness of this event, the media attention was high and policymakers were attracted. The monetary losses for the municipality of Braunsbach (~2500 residents) were estimated to EUR 104 million, which is more than 90% of the estimated EUR 112 million of total damage in Schwäbisch Hall (Landkreis Schwäbisch Hall, 2016). The catchment of the creek
30 primarily responsible for the inundation in May 2016, the "Orlacher Bach", is only about 6 km² in size and characterized by steep slopes. Heavy rainfalls in the catchment area between 18:45 and 20:00 of May 29th resulted in an estimated accumulated precipitation of 100 mm (Ziese et al., 2017) - revised to estimates of up to 140 mm using the method of Jacobi and Heistermann (2016) (Bronstert et al., 2016) - finally leading to the severe flash flood, which was accompanied by

massive amounts of debris and rubble. Streets along the main runoff channel were blocked by layers of debris, up to a thickness of two to three meters while numerous houses in the area showed severe structural damage. Given the relation of town size, event duration and catchment area, the losses were extremely high. Eventually, these cases of severe flash flooding in Germany triggered a reassessment of local risk and revealed that the processes and impacts of flash floods are insufficiently understood, also showing that research on and management of this particular flood type need to catch up.

Flash floods are defined as rapid flood events as an effect of very intense, timely and concentrated precipitation, which are potentially enhanced by orographic features (Gaume et al., 2009; Borga et al., 2014). According to Gaume et al. (2009), flash floods can be triggered by diverse hydrological and meteorological processes and are, compared to riverine and pluvial flooding, associated with a higher number of fatalities. Whereas pluvial floods are related to urban areas and caused by sewage overflow and surface runoff (Maksimović et al. 2009), flash floods usually occur in mountainous regions, where they can trigger debris flows and/or hyperconcentrated flows. Debris flows and hyperconcentrated flows are characterized by the amount of transported and suspended sediment. With a sediment concentration between 60 and 80 volume percent, the quantity of solid material is often higher for debris flows than for hyperconcentrated flows (Gaume et al., 2009). Both flow types show a variation in grain size distribution and deposition characteristics as well: while debris flows potentially carry large debris, boulders and gravel, hyperconcentrated flows transport finer sediments. (Pierson et al. 1987; Gaume et al., 2009; Totschnig et al., 2011; Hungr et al., 2013; Borga et al., 2014).

Weather extremes in Europe are expected to occur more frequently, leading to strong storms, droughts and heavy precipitation in various regions (Beniston et al., 2007; Murawski et al., 2016; Volosciuk et al., 2016). More intense and concentrated rainfall in Central Europe might increase the risk of severe flash flood events not only in mountainous regions but uplands as well, affecting regions which were previously not perceived as flood prone. Further, an increased risk for higher damage due to flooding can be detected, which is mainly influenced by urbanisation, economic growth as well as changing land use patterns (Thieken et al., 2014; European Environment Agency, 2017). As an effect, flash floods are progressively perceived as a serious hazard in Central Europe. Yet, what implications on elements at risk in case of flash flooding might be expected? How to assess their vulnerability?

Vulnerability can be defined as the tendency of elements at risk to suffer negative effects and damage, if affected by a specific hazard (Cardona et al., 2012). Regarding flash floods, vulnerability and risk estimations were already conducted in several studies. For instance, Papathoma-Köhle (2016) pointed out that vulnerability assessments for flash floods or debris flows need to be reviewed and adjusted constantly. In her study, an indicator-based method was used for assessing the vulnerability of elements at risk which are exposed to debris flows in South Tyrol. In this regard, the relevance of building characteristics and location for vulnerability estimations were highlighted. Similarly, Fuchs et al. (2012) conducted a study which describes the vulnerability of elements at risk, based on clusters of similar damage ratios caused by flood events. This spatial approach revealed that higher damage ratios are not only an effect of stronger floods, debris flows or hyperconcentrated flows, but are as well dependent on land-use patterns and the characteristics of the elements at risk. With regard to non-alpine environments, Hlavčová et al. (2016) performed a post-hoc analysis of three strong flash flood events

which occurred between 1988 and 2004 in northern Slovakia, focusing on the hydrology as well as hydraulic and topographic properties of the catchment areas. They showed that the modelling of flash flood events goes along with major uncertainties due to the lack of data and overall nonlinear relationship between precipitation, runoff and catchment properties.

5   Regarding flash floods in Central Europe and particularly non-alpine environments, we are in the early stages to understand specific flood and subsequent damaging processes. Especially concerning vulnerability estimations of the elements at risk as well as damage driving factors, flash floods are insufficiently understood. Yet, it can be assumed that damage processes of flash floods differ from those of riverine floods, highlighting the need for elaborate research in this field. To obtain a better understanding of the damage processes of flash floods as well as of effective mitigation options, a comprehensive damage

10   data base that links process dynamics and intensities with damage and loss is needed, but currently not available. Consequently, we present Braunsbach as a case study, having collected and analysed data in order to add to the knowledge in this field. This research paper follows two major objectives. Using the flash flood in Braunsbach as a case study, it is aimed at identifying, analysing, comparing and discussing factors that govern damage caused by this event, applying different linear and nonlinear methods. As a second issue, the methods used for the ex-post damage data collection in Braunsbach and

15   the creation of this database are presented and discussed to demonstrate accompanying challenges as well as advantages during the field work.

2 Methods

Collecting and analysing data of structural and non-structural damage to buildings is valuable for understanding specific damage processes, helping to design and assess effective mitigation measures and creating damage models, which can be

20   used to estimate potential monetary losses ex-ante. Thus, a digital survey was designed to collect relevant information in Braunsbach which can be used for detailed post-hoc analysis. The type of recorded information is based on existing literature on flood damage surveys (e.g. Thieken et al., 2005; Schwarz and Maiwald, 2007; Merz et al., 2010; Molinari et al., 2014).

2.1 Contents of the survey

The creation of the survey as well as the data collection was implemented with "KoBoCollect", a self-explaining and

25   network-based open source software which was developed by the Harvard Humanitarian Initiative together with the Birmingham and Women's Hospital in 2014 (kobotoolbox.org, 2016). The software is designed for quick and reliable information collection after natural disasters or in humanitarian crises. Open source software, as a method for data collection and gaining of knowledge, is increasingly becoming important within the field of natural hazards (Eckle et al., 2016; Klonner et al., 2016). For instance, OpenStreetMap (OSM) and other volunteered geographic information helps to create

30   comprehensive databases of up to date geospatial data which also can be used for natural risk assessment (Schelhorn et al., 2014; Vaz and Arsanjani, 2015; Yang et al., 2016).

The gathered information includes an estimation of damage grades ranging from D1 (no structural damage, slight nonstructural damage) to D5 (very heavy structural damage, very heavy non-structural damage). For this classification, the scheme developed by Schwarz and Maiwald (2007) was adopted to obtain a consistent database and to ensure comparability with follow-up studies and with data of riverine flood damage. Since monetary losses could not be recorded shortly after the

5   event, this classification scheme further offers options for potential subsequent loss estimations. Additional information included the GPS coordinates, the address (for internal computational use only), the inundation depth at the building in cm, visible damage caused by debris, visible contamination by oil or sewage, the building material and type, specific precautionary measures at the building, the building usage (residential, commercial, public etc.), the number of storeys and types of outbuildings, the estimated year of construction, the perceived condition of the building before the event, existing

10 shop windows on the ground floor, the existence of a cellar, the sealing degree of the near surrounding as well as the exposition (of the building) in flow direction. All variables except for the address, inundation depth, storeys and the estimated year of construction were pre-coded with the option to record open answers or NA values, resulting in a nominal-, ordinal- as well as interval-scaled data structure. The complete survey with variable descriptions can be seen in table 1.

2.2 On-site data collection

15 The on-site damage assessment was carried out between June 6th and June 9th, 2016, i.e. 8 to 10 days after the event. The digital survey was conducted by a team of five researchers who investigated all buildings in Braunsbach affected by the flash flood, using mobile tablet computers with an integrated GPS function.

Some of the flooding characteristics such as flow velocities, grain size as well as the degree of erosion and amount of suspended material could not reliably be determined in the aftermath of the event. Hence, the exposition of the building was

20 used as a proxy instead. It is assumed that the degree of exposition can be related to flow velocities, hydrostatical forces and (to a certain extent) to sediment/debris load, which in turn leads to different erosion rates at the buildings' foundation. The exposition in flow direction describes the exposition of building walls, corners or parts to the direction and area of the main runoff channel. In this case, a high exposition means that at least one side of the building was fully exposed to water and potential debris flows. A medium exposition was assumed when parts of the building were exposed, sheltered buildings are

25 characterized by a low exposition.

A thermographic camera (model Testo 876, 160x120 pixels) was used to validate and to derive the inundation depth in such cases, where a reliable estimation through visible traces and marks was not possible. This was done by detecting the remaining moisture in the walls - caused by inundation - through slight differences in surface temperature. A second advantage of the thermographic camera was the detection of different building materials, which may be covered externally

30 (i.e. plastered half-timbered houses could still be identified as such).

5

2.3 Post-hoc data analysis

The data was pre-processed and analysed/prepared in R 3.3.1 and QGIS 2.14.3, using the R packages "randomForest", "randomGLM" and "nnet". Since our study aims to identify and analyse damage driving factors of flash floods, the following variables were considered and used as predictor variables for the damage grade:

5

Building material (binary-coded):
- masonry
- wood
- concrete
10 - half timbered

Precaution (binary-coded):
- Different (building) materials (of cellar and ground floor)
- Higher ground floor
- No precaution

15 External forces:
- Inundation depth
- Exposition in flow direction
- Contamination visible (binary-coded)

Resistance parameters:
20 - (Building) condition before event
- Estimated construction year

Various:
- Shop window present (binary-coded)
- Near surrounding sealed
25 - Having cellar (binary-coded)
- Outbuilding present (binary-coded)
- Private building usage (binary-coded)

The choice of the variables which were specifically analysed was based on both, own judgement (e.g. if the near surrounding
30 is sealed or if an outbuilding or shop window is present) as well as existing literature. Here, Thieken et al. (2005), Merz et al. (2010) and Maiwald and Schwarz (2015) give an overview of important damage influencing factors in case of (river) flooding, including building characteristics, precaution measures and contamination.

2.3.1 Models and correlation tests

Detecting nonlinear and nonmonotonic relationships within recorded data becomes increasingly important with regard to flood loss modelling and associated uncertainties (Kreibich et al., 2016). Consequently, a Random Forest model (RF) (Breiman, 2001) was chosen as a method of analysis due to its potential to display nonlinear relationships between variables.

5　The Random Generalized Linear Model (RGLM) (Song et al., 2013) was constructed as an alternative model to compare the results of the nonlinear RF to a method which implies linear variable coherences. Both models use the same predictor variables and the "damage grade" as dependent variable in order to identify potential damage driving factors (Figure 1 and 2).

The RF is calculated with a number of 500 trees and 4 random variables per split. The number of trees represents the default
10　settings of the algorithm. The number of variables per split corresponds to the square root of the total variable count (17 in this case, resulting in 4 random variables per split). The RGLM takes 100 iterations with 50 samples per iteration (bag) and a varying count of variables (2 to 17) per bag. Also the number of iterations within the RGLM represents the default settings. The number of samples per bag corresponds to 2/3 of the total number of observations in the dataset (74 in this case, which gives 50 observations selected by bootstrapping). The count of variables per bag is randomly chosen between 2 and the total
15　count of variables. The variable/feature importance of the RF is given by the Mean Decrease Gini, which describes the loss in model performance when permuting the feature values (Breiman 2001). A higher Mean Decrease Gini indicates a higher importance of the particular variable for the RF model prediction. The feature importance of the RGLM is expressed through the selection count of a variable for the model prediction. By using feature forward selection, a higher selection count of a particular variable indicates a stronger predictive power within the RGLM model (Song et al., 2013). The performance of
20　both models is given by the rate of false classifications, based on the out of bag predictions. The relative number of cases which were not recognized as the true class is hereby shown in percent (see section 3.2).

Categories with a nominal variable structure (i.e. building usage, building material and structural precaution measures) exist in a binary format, allowing for basic correlation tests. Thus, the identified feature importance from the models was compared to the results from a Spearman's rank correlation matrix. The Spearman's rank correlation was chosen due to the
25　advantage that this method is suitable to analyse variables with different scales of measurements and indicates the strength of monotonic relationships. Here, the same variables as in the RF and RGLM models were used (Figure 3). A complete list of variable correlations is attached in the appendix.

Furthermore, a multinomial logistic regression was applied to test variable coherences between the damage grades and a local impact indicator (Figure 4), which describes a combination of inundation depth and the buildings' exposition in flow
30　direction (the preparation of the local impact indicator is explained in section 2.3.2). By treating the damage grade as a categorical variable, the multinomial logistic regression model gives probabilities of category affinity, given a specific local impact. In order to obtain more data points per category and to reduce modelling uncertainties, the damage grades 3, 4 and 5

were combined. The inherent calculations are based on artificial neural networks (Ripley, 1996; Venables et al., 2002) which again do not require any model-specific assumptions such as linearity.

2.3.2 Derivation of an indicator "Local impact"

Maiwald and Schwarz (2015) give an up to date overview of factors which influence structural damage on buildings in case
5 of flooding. Especially the building material, condition (before the event) and the age are important factors related to a building's resistance potential. Factors such as inundation level, flow velocity, specific energy and contamination relate to "action" parameters and describe external forces (Maiwald and Schwarz, 2015). Thus, in our study, the inundation depth measured at the building and the building's exposition in flow direction were combined to create a local impact, which can be seen as a proxy for local flood related impact and hydrostatical forces at the building. Consequently, we chose a
10 combination of these factors where both contribute to equal extents. While the inundation depth has continuous values, which are roughly uniformly distributed between 2 and 360 cm, the exposition in flow direction is recorded in three classes (low, medium, high). To achieve comparable variable ranges, the exposition classes "low", "medium" and "high" are transformed into the mean values of the lower, middle and upper third of recorded water levels. The derived values 57, 133 and 230 fit into the range of observed water levels, enabling a combination of both attributes (Figure 5). The calculated local
15 impact corresponds to the sum of water level and transformed exposition value. Please note that the exposition values are not used to replace water levels, but are only transformed into a comparable range.

Furthermore, a local impact map was created in QGIS (Figure 6) by calculating Voronoi diagrams for the geocoded data points and solely displaying the area with affected houses. For simpler visual appearance and better distinction of the displayed data, the Voronoi diagrams were smoothed by a Gaussian filter. The local impact map is used for visualization and
20 comparison of the local impact indicator to the spatial distribution of the damage grades, since potential areas of similar local impact between and around the buildings are shown. However, it has to be noted that the local impact is measured on buildings itself and therefore hypothetical for the areas around.

3 Results and discussion

The flash flood in Braunsbach was accompanied by a considerable amount of sediment and building rubble, potentially
25 showing flow characteristics of debris flows such as defined by Totschnig et al. (2011), Hungr et al. (2013) and Borga et al. (2014). Yet, a clear distinction between flash floods and debris flows is not always straightforward and could not be reliably determined afterwards. Throughout our discussion and conclusion, we will therefore relate to the term "flash flood" only. The following section begins with a general reflection on the data collection process and data quality. Thereafter, the damage influencing factors are identified and discussed by applying different linear and nonlinear methods. Finally, features such as
30 the local impact and the damage grades are spatially visualised, helping to discuss our outcomes and illustrating the flash flood processes in Braunsbach.

3.1 Data collection and KoBoCollect

The in-field work load can be estimated to roughly 10 hours in which a team of five researchers was able to survey 96 buildings in Braunsbach, for each specifying 21 variables and taking a picture. Table 1 provides an overview of the data types and frequency distributions. One week after the event, the structural damage on buildings and building characteristics

5   were still assessable, since the main work within this period was focused on clearing the roads, establishing paths for large construction machinery as well as removing and cleaning the interior of affected buildings. The progress of the clean-up work was even beneficial for the damage assessment to a certain degree, as a thick layer of debris and rubble previously covered big parts of the damaged buildings. However, a few buildings could not be reliably examined, since debris and rubble were still considerably hampering the access.

10 When handling the thermographic camera it has to be pointed out that, even one week after the event, remaining moisture and visible traces could still be detected without problems. Yet, ascending humidity in the walls is a point to consider when using a thermographic camera for water level estimations. For that reason the thermal images were checked against estimations based on visible mud contamination and marks caused by water and transported debris as well. Since the thermally derived water levels matched well with visible traces, the inundation depth for buildings derived from thermal

15 images could be accepted without any correction.

Overall, the in-field data collection was greatly facilitated by the use of "KoBoCollect" in terms of speed, handling of the gathered data and efficiency of data processing and analysis. However, to create a uniform database and to maintain consistency among the different team members throughout the data collection process, objective criteria for items such as the structural damage had to be defined. Therefore, careful preparations and agreements were carried out prior to the field trip

20 off site as well as on site. In retrospect, we consider the data to be consistent in a way that the team members had very similar opinions e.g. on the damage grades or exposition in flow direction. Thus, a bias in the dataset due to personal variations in expert judgement is expected to be low. This assumption is further supported by the engineering analysis of Maiwald and Schwarz (2016), who applied the same damage classification system to assess the buildings structural damage in Braunsbach. Their report reveals that the distribution of the recorded damage grades after a second inspection (D1: 40,

25 D2: 43, D3: 5, D4: 7, D5: 3) is relatively similar to the distribution presented in this study (D1: 40, D2: 35, D3: 5, D4: 6, D5: 5) (Table 1). Although it is not known which damage grade was assigned to which building, it is likely that, even among people with different qualifications (experienced engineer, researcher or student), comparable results can be achieved and that data collection can be consistent. This offers interesting options for crowdsourced information collection using open source software such as "KoBoCollect", which can be helpful for scientific research.

3.2 Models and Correlation tests

First, the collected data were used to identify damage driving factors by creating a Random Forest model (RF) and a Random Generalized Linear Model (RGLM). In a next step, a Spearman's rank correlation matrix was constructed. In the following, the different model and correlation results are discussed and compared to each other.

5 The post-hoc data analysis revealed that the RF and RGLM both show a relatively poor model performance, based on the false classifications. Here, the percentages of false classifications for the RF are 40% for damage grade D1, 37.5% for D2 and 100% for D3 and higher. The RGLM performed slightly better with a false classification of 40% for D1, 45.8% for D2, 40% for D3, 80% for D4 and 100% for D5. However, trends and relations of predictor variables with the damage grade can be derived. Both models give the highest feature importance for the damage class to the inundation depth and the exposition

10 (of the building) in flow direction. Here, the Mean Decrease Gini for the RF was 11.2 and 7.0 (average: 2.6), whereas the RGLM feature selection count in 100 iterations was 90 and 89 respectively (average: 40) (Figures 1 and 2). Regarding the correlation tests, the highest positive (and significant) correlations can be seen between the damage class and the exposition of the building in flow direction as well as the damage class and inundation depth with a value of 0.65 and 0.69 respectively (Figure 3), which strongly confirm the results of the RF and RGLM. The detected strong link of the exposition of the

15 building in flow direction and the inundation depth to the caused damage makes sense, given the nature of the event and the mass of debris, water and mud flowing down the main channels within the village of Braunsbach. These results are confirmed by Maiwald and Schwarz (2016) as well, who identified the exposition of a building to the flow direction as an important parameter for potential structural damage.

The estimated construction year of a building displays a certain importance within the RF as well as the RGLM model with a

20 Mean Decrease Gini of 5.9 and a feature selection count of 41, respectively. In this case, the correlation analysis does not reveal any significant monotonic relationships between the estimated construction year and the damage grade (Figure 3). Additionally, the building condition before the event displays only slight importance within the RF and slight, non-significant correlation with the damage grade. However, it can be assumed that newer buildings or buildings in a better condition have a higher resistance to structural damage, which is in line with Maiwald and Schwarz (2015) who consider the

25 building age and condition to have an influence on the expected structural damage.

Further, a positive correlation, but of low significance, can be observed between the damage grade and the existence of a shop window with a value of 0.19 and a p-value of 0.11. Here, a trend towards higher damage grades caused by shop windows on ground level which - in case of breaking - open debris and water paths to the inside of the building can be assumed. Also Maiwald and Schwarz (2016) underline the fact that broken windows may allow water and debris to

30 accumulate inside the building, causing damage to sustaining building structures. Yet, our results might also be affected by the fact that buildings with shop windows mainly occur along the main street and city centre and are therefore located inside the main flow channels.

No (obvious) precaution at the property level indicates a slightly higher chance for higher damage as well by displaying a positive correlation of 0.15 with the damage class, although the significance is relatively low (p-value 0.19) and there is no remarkable importance within the RF and the RGLM models. Yet, this is in line with the negative correlation between the damage class and the precaution measure "different (building) materials (of cellar and ground floor)" of -0.19 which in fact

5   shows a low significance (p-value 0.1) but still allows for meaningful assumptions. This is supported by Thieken et al. (2005) and Merz et al. (2010) who claim that different precautionary measures significantly reduce the damage on buildings in case of flooding. Still, the question arises to which degree precautionary measures, which were effective at riverine flooding, are suitable to lower or mitigate structural damage on buildings in case of flash flooding.

The building material masonry seems to have a damage reducing effect by displaying a negative correlation of -0.12 (p-value

10   0.33) with the damage class whereas the building materials half-timbered and wood show almost no significant positive correlations with the damage. Interestingly, the RGLM model considers the building materials such as masonry or half-timbered as relatively important for the damage grade prediction, whereas the RF does not display a significant feature importance in this case. Although it is not clearly shown which distinct building material is related to lower structural damage, it can be assumed that if being hit by debris, half-timbered houses are more susceptible to structural damage than

15   houses made of masonry and concrete due to their lower structural stability (Schwarz and Maiwald, 2007).

Overall, when performing detailed analyses such as models and correlation tests it has to be considered that the database of 96 data points is rather small and assumingly insufficient for creating representative and universal results. This fact could also explain the low model performances and low significances in some of the cases discussed above. Nonetheless, it is important to point out the strong correlations in many cases of up to 0.69 (damage class and exposition in flow direction)

20   revealing obvious damage driving factors and showing as well that the data collection within the team of different researchers was consistent.

3.3 Multinomial logistic regression model

To evaluate the newly created local impact indicator, which serves as a proxy for hydrostatical and impact forces at the building (see section 2.3.2), a multinomial logistic regression was applied. By analysing the dependency between only the

25   local impact indicator and the damage grade, the influence of external forces on the damage grade can be observed separately, since resistance parameters and building characteristics are neglected. As can be seen in figure 4, there is a clear coherence of an increasing local impact and an increasing probability to belong to higher damage grades. However, figure 4 also reveals that external forces are not enough to explain the complex damage pattern, especially with regard to higher local impact values and higher damage grades. Here, the 95% confidence interval indicates that the probability of damage grade

30   affinity varies up to 45%. This means, that i.e. given a local impact of 550, the chance of belonging to class D2 ranges from 5% to 50% and for greater equal D3 from 50% to 95%. However, it needs to be considered that the greater variability for higher values can be caused by fewer data points for higher damage grades as well. An interaction of both factors (more complex damage driving factors for higher damage grades and fewer cases of high damage) might be most likely. Still,

figure 4 shows that a local impact indicator can be suitable to evaluate the exposure to this type of hazard, which, in addition to the characteristics of the element at risk, might allow vulnerability estimations such as performed by Papathoma-Köhle (2016).

3.4 Local impact map

5   The local impact map reveals that highly damaged buildings and a strong local impact occurred along the main runoff channels of water and debris during the event (Figure 6). Higher damage classes were also recorded in the lower-lying town regions, where the tributaries "Orlacher Bach" and "Schlossbach" flow into the river "Kocher", since debris and water accumulated in these areas and caused severe structural damage. Most of the higher damage classes are located in high local impact areas. Yet, especially in those areas, the degree of damage differs strongly, highlighting the complexity of damage

10   driving processes that cannot be explained by the local impact alone (section 3.3). The flow characteristics of debris and rubble during severe flash floods can be unforeseen and influenced by chaotic factors, changing sediment deposition as well as bedload processes (Totschnig et al., 2011; Hungr et al., 2013). Thus, it can be assumed that, during the flash flood in Braunsbach, chaotic factors and deposition of debris led to various damage patterns which remain inexplicable through quantitative analysis and modelling. This is strongly supported by the engineering report of Maiwald and Schwarz (2016),

15   who claim that chaotic flow processes at the building caused by rubble and debris can greatly influence the inundation depth. However, it is revealed that buildings with a high exposition in flow direction are more susceptible to severe structural damage, since the probability of large debris colliding with building walls is much higher and erosion of the foundation is more likely to happen. Also Maiwald and Schwarz (2016) stated that the recorded damage patterns differ from damage patterns caused by riverine flooding and appear to be more severe due to higher hydrodynamic stress and collision of debris

20   with the building. Conversely, some buildings can benefit from shadowing effects of neighbouring buildings, which retain debris and suspended material to a certain degree.

It can be summarised that, next to individual flow- and deposition processes of the flood, local factors, shadowing processes, and building characteristics shared a certain importance as damage driving factors in Braunsbach, highlighting the complexity of this event. This is supported by the findings of Fuchs et al. (2012) who revealed as well that damage on

25   buildings is not only caused by flood-inherent processes and intensities, but is also influenced by building characteristics and dependent on the general land use pattern.

4 Conclusion

The evaluation and data analysis in this study resulted in important information about the impacts (damage to buildings) of the flash flood event in Braunsbach. It is revealed that not only the water depth, which is often considered as only damage

30   driving factor in riverine flood loss modelling (see Merz et al., 2010), but also the exposition of a building in flow direction and susceptible building parts like e.g. shop windows seem to be risk factors in flash-flood prone regions. This result

considerably differs from investigations on damage caused by riverine floods (Kreibich et al., 2009). Yet, the damage driving as well as damage reducing factors of flash floods are complex, often unpredictable, contingent upon the surrounding as well as dependent on certain building characteristics.

Knowing processes of flash floods and their impacts can help to create awareness for future events and support strategic

5   planning with regard to similar emergencies. Concerning the European Floods Directive 2007/60/ EC and its implementation in Germany, implications according to the German Federal Water Act exist. The consideration of flash floods and surface water flooding as a "significant risk" would result in the obligation to create new nationwide hazard and risk maps. As a further consequence, the German Federal Water Act intends a building ban in all areas that are affected by a 100-year flood event, which would lead to serious consequences for local planning in flash flood prone regions.

10 Still, maps such as the presented local impact map could be a supportive first step. The estimation of a local impact could be used in integrated risk management and strategic planning of mitigation measures against future hazards in Braunsbach or similar villages in that region. Thus, the introduced concept may be beneficial for the identification of potentially vulnerable locations on a small scale and within case studies, helping to understand the potential future development of flash flood prone regions. However, further investigations are needed in order to verify the results and to obtain larger databases.

15 To facilitate data collection in the future, the case further demonstrates the potential of mobile devices and open-source applications. In the field, the simplicity, speed, quality and handling of information using the open source application "KoBoCollect" particularly stood out as a great advantage. Even in a short time and with a small team of researchers it was possible to gather a fair amount of useful information that could be further processed and analysed. The public availability of the software makes it a fast and ad-hoc tool for assessing different kind of questions, usable in various research fields and not

20 only for scientific but also for private uses. However, the question should be raised whether the quality of crowdsourced information is suitable for scientific investigations and how to approach and deal with possible limitations, security and copyright issues as well as uncertainties. Still, it can be concluded that open source data collection software for mobile use has great potential as a scientific tool to generate extensive valuable data under challenging conditions. It should be especially considered in time critical research applications such as ex-post-disaster analyses, as was demonstrated by the

25 presented case of Braunsbach.

Competing interests

The authors declare that they have no conflict of interest.

13

5

10

References

Beniston, M., Stephenson, D. B., Christensen, O. B., Ferro, C. A. T., Frei, C., Goyette, S., Halsnaes, K., Holt, T., Jylhä, K., Koffi, B., Palutikof, J., Schöll, R., Semmler, T., and Woth, K.: Future extreme events in European climate: an exploration of regional climate model projections, Climatic Change, 81, Suppl.1, 71-95, doi:10.1007/s10584-006-9226-z, 2007.

5

Borga, M., Stoffel, M., Marchi, L., Marra, F., and Jakob, M.: Hydrogeomorphic response to extreme rainfall in headwater systems: Flash floods and debris flows, Journal of Hydrology, 518, 194-205, 2014.

Bracken, L. J., Oughton, E. A., Donaldson, A., Cook, B., Forrester, J., Spray, C., Cinderby, S., Passmore, D., and Bisset, N.:
10 Flood risk management, an approach to managing cross-border hazards, Nat Hazards, 82, Suppl.2, 217-240, doi:10.1007/s11069-016-2284-2, 2016.

Breiman, L.: Random forests, Machine learning, 45, 1, 5-32, doi:382 10.1023/A:1010933404324, 2001.

15 Bronstert, A., Agarwal, A., Boessenkool, B., Fischer, M., Heistermann, M., Köhn, L., Moran, T., and Wendi, D.: Die Sturzflut von Braunsbach am 29. Mai 2016: Entstehung, Ablauf und Schäden eines „Jahrhundertereignisses". Teil 1: Meteorologische und Hydrologische Analysen, submitted to Hydrologie und Wasserbewirtschaftung, 2016 (in German).

Bubeck, P., Botzen, W. J. W., Kreibich, H., and Aerts, J. C. J. H.: Detailed insights into the influence of flood-coping
20 appraisals on mitigation behaviour, Global Environmental Change, 23, 5, 1327–1338, doi:10.1016/j.gloenvcha.2013.05.009, 2013.

Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit (BMUB): Directive 2007/60/EC of the European Parliament and of the Council - of 23 October 2007 - on the assessment and management of flood risks, Official Journal of
25 the European Union, L, 27-34, 2007.

Cardona, O.D., van Aalst, M. K., Birkmann, J., Fordham, M., McGregor, G., Perez, R., Pulwarty, R. S., Schipper, E. L. F., and Sinh, B.T.: Determinants of risk: exposure and vulnerability, In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation [Field, C.B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L.,
30 Mastrandrea, M. D., Mach, K. J., Plattner, G. -K., Allen, S. K., Tignor, M., and Midgley, P. M., (eds.)], A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC), Cambridge University Press, Cambridge, UK, and New York, NY, USA, 65-108, 2012.

Eckle, M., De Albuquerque, J. P., Herfort, B., and Zipf, A.: Leveraging OpenStreetMap to support flood risk management in municipalities: A prototype decision support system, ISCRAM 2016 Conference Paper – Rio de Janeiro, Brazil, May 2016.

European Environment Agency: Climate change, impacts and vulnerability in Europe 2016 - An indicator-based report, 1,
5   Luxembourg, doi:10.2800/534806, 2017.

Fuchs, S., Ornetsmüller, C., Totschnig, R.: Spatial scan statistics in vulnerability assessment: an application to mountain hazards, Nat Hazards, 64, 3, 2129-2151, doi:10.1007/s11069-011-0081-5, 2012.

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

15 GDV – Gesamtverband der Deutschen Versicherungswirtschaft: Naturgefahrenreport 2016, Berlin, 2016 (in German).

Grüning, H., and Grimm, M.: Unwetter mit Rekordniederschlägen in Münster, KW Korrespondenz Wasserwirtschaft, 8, 2, 88-93, doi:10.3243/kwe2015.02.001, 2015 (in German).

20 Hlavčová, K., Kohnová1, S., Borga, M., Horvát, O., Šťastný, P., Pekárová, P., Majerčáková, O., and Danáčová, Z.: Post-event analysis and flash flood hydrology in Slovakia, J. Hydrol. Hydromech., 64, 4, 304-315, doi:10.1515/johh-2016-0041, 2016.

Hungr, o., Leroueil, S., and Picarelli, L.: The Varnes classification of landslide types, an update, Landslides, 11, 2, 167-194,
25 doi:10.1007/s10346-013-0436-y, 2013.

Jacobi, S. and Heistermann, M.: Benchmarking attenuation correction procedures for six years of single-polarized C-band weather radar observations in South-West Germany, Geomatics, Natural Hazards and Risk, 7, 6, 1785-1799, doi:10.1080/19475705.2016.1155080, 2016.
30
Kienzler, S., Falter, D., Jentsch, S., and Thieken, A. H.: Reduktion von Hochwasserschäden im Zusammenspiel von staatlicher und privater Vorsorge - eine Fallstudie an der Mulde, KW Korrespondenz Wasserwirtschaft, 8, 11, 680-687, doi:10.3243/kwe2015.11.002, 2015 (in German).

Klonner, C., Marx, S., Usón, T., De Albuquerque, J. P., and Höfle, B.: Volunteered Geographic Information in Natural Hazard Analysis: A Systematic Literature Review of Current Approaches with a Focus on Preparedness and Mitigation, ISPRS Int. J. Geo-Inf., 5, 103, doi:10.3390/ijgi5070103, 2016.

5   KoBoToolbox: http://www.kobotoolbox.org/, last access: 15 November 2016.

Kreibich, H., Botto, A., Merz, B., and Schröter, K.: Probabilistic, Multivariable Flood Loss Modeling on the Mesoscale with BT-FLEMO, Risk Analysis, 37, 4, 774-787, doi:10.1111/risa.12650, 2016.

10   Kreibich, H., Piroth, K., Seifert, I. Maiwald, H., Kunert, U., Schwarz, J., Merz, B. and Thieken, A. H.: Is flow velocity a significant parameter in flood damage modelling?, Nat. Hazards Earth Syst. Sci., 9, 1679-1692, doi:10.5194/nhess-9-1679-2009, 2009.

Kundzewicz, Z. W., Krysanova, V., Dankers, R., Hirabayashi, Y., Kanae, S., Hattermann, F. F., Huang, S., Milly, P. C. D., 15   Stoffel, M., Driessen, P. P. J., Matczak, P., Quevauviller, P., and Schellnhuber, H. -J.: Differences in flood hazard projections in Europe – their causes and consequences for decision making, Hydrological Sciences Journal, 62, 1, 1-14, doi:10.1080/02626667.2016.1241398, 2017.

Maiwald, H., and Schwarz, J.: Die Sturzflut von Braunsbach - Ingenieuranalyse der Gebäudeschäden, Bautechnik, 93, 12, 20   925-932, doi:10.1002 / bate.201600087, 2016 (in German).

Maiwald, H., and Schwarz, J.: Damage And Loss Prognosis Tools Correlating Flood Action And Building's Resistance-type Parameters, International Journal of Safety and Security Engineering, 5, 3, 222-250, doi:10.2495/SAFE-V5-N3-222-250, 2015.

25

Maksimović, Č., Prodanović, D., Boonya-Aroonnet, S., Leitão, J. P., Djordjević, S., and Allitt, R.: Overland flow and pathway analysis for modelling of urban pluvial flooding, Journal of Hydraulic Research, 47, 4, 512-523, doi:10.1080/00221686.2009.9522027, 2009

30   Merz, B., Kreibich, H., Schwarze, R and Thieken, A. H.: Review article: Assessment of economic flood damage, Nat. Hazards Earth Syst. Sci., 10, 1697-2010, doi:10.5194/nhess-10-1697-2010, 2010.

Munich Re: Topics Geo 2016 - Natural catastrophes 2016 - Analyses, assessments, positions, München, 2017.

Murawski, A., Zimmer, J., and Merz, B.: High spatial and temporal organization of changes in precipitation over Germany for 1951–2006, International Journal of Climatology, 36, 6, 2582-2597, doi:10.1002/joc.4514, 2015.

Osberghaus, D., and Philippi, A.: Private Hochwasservorsorge und Elementarschadenversicherung - Moral Hazard, der Effekt von Informationskampagnen, und eine Versicherungsillusion, ZVersWiss, 105, 3, 289-306, doi:10.1007/s12297-016-0341-2, 2016 (in German).

Papathoma-Köhle, M.: Vulnerability curves vs. vulnerability indicators: application of an indicator-based methodology for debris-flow hazards, Nat. Hazards Earth Syst. Sci., 16, 1771-1790, doi:10.5194/nhess-16-1771-2016, 2016.

Pierson, T. C., and Costa, J. E.: A rheologic classification of subaerial sediment-water flows, Geological Society of America, doi:10.1130/REG7-p1, 1987.

Ripley, B. D.: Pattern recognition and Neural Networks, Cambridge University Press, Cambridge, 1996.

Schelhorn, S. J., Herfort, B., Leiner, R., Zipf, A., and De Albuquerque, J. P.: Identifying Elements at Risk from OpenStreetMap: The Case of Flooding, ISCRAM Conference Paper – University Park, Pennsylvania, USA, May 2014.

Schwarz, J. and Maiwald, H.: Prognose der Bauwerksschädigung unter Hochwassereinwirkung, Bautechnik, 84, 7, 450-464, 2007 (in German).

Song, L., Langfelder, P., and Horvath, S.: Random generalized linear model: a highly accurate and interpretable ensemble predictor, BMC Bioinformatics, 14, 5, 1471-2150, 2013.

Thieken, A. H., Kienzler, S., Kreibich, H, Kuhlicke, C. Kunz, M., Mühr, B., Müller, M., Otto, A. Petrow, T., Pisi, S. and Schröter, K.: Review of the flood risk management system in Germany after the major flood in 2013, Ecology & Society, 21, 2, 51, doi:10.5751/ES-08547-210251, 2016.

Thieken, A. H., Cammerer, H., Dobler, C., Lammel, J., and Schöberl, F.: Estimating changes in flood risks and benefits of non-structural adaptation strategies - a case study from Tyrol, Austria, Mitig Adapt Strateg Glob Change, 21, 3, 343-376, doi:10.1007/s11027-014-9602-3, 2014.

Thieken, A. H., Müller, M., Kreibich, H., and Merz, B.: Flood damage and influencing factors: New insights from the August 2002 flood in Germany, Water Resources Research, 41, 12, 1-16, doi:10.1029/2005WR004177, 2005.

Totschnig, R., Sedlacek, W., and Fuchs, S.: A quantitative vulnerability function for fluvial sediment Transport, Nat Hazards, 58, 2, 681-703, doi: 10.1007/s11069-010-9623-5, 2011.

5   Vaz, E., and Arsanjani, J. J.: Crowdsourced mapping of land use in urban dense environments: An assessment of Toronto, The Canadian Geographer / Le Géographe canadien, xx(xx), 1–10, doi:10.1111/cag.12170, 2015.

Unwetter-Schadensliste   Landkreis   Schwäbisch   Hall: http://www.lrasha.de/index.php?id=302?&id=302?&no_cache=1&publish[id]=457120&publish[start]=,   last   access: 16
10   August 2016.

Venables, W. N., and Ripley, B. D: Modern Applied Statistics with S - Fourth Edition, Springer, New York, doi:10.1007/978-0-387-21706-2, 2002.

15   Volusciuk, C., Maraun, D., Semenov, V. A., Tilinia, N., Gulev, S. K., and Latif, M.: Rising Mediterranean Sea Surface Temperatures Amplify Extreme Summer Precipitation in Central Europe, Scientific Reports, 6, 32450, doi:10.1038/srep32450, 2016.

Yang, A., Fan, H., Jing, N., Sun, Y., and Zipf, A.: Temporal Analysis on Contribution Inequality in OpenStreetMap: A
20   Comparative Study for Four Countries, ISPRS Int. J. Geo-Inf., 5, 5, doi:10.3390/ijgi5010005, 2015.

Ziese, M., Junghänel, T., and Becker, A.: Andauernde Großwetterlage Tief Mitteleuropa entfaltet Ihr Unwetterpotential in Deutschland,   DWD   online   resource,   available   at: https://www.dwd.de/DE/leistungen/besondereereignisse/niederschlag/20160603_starkregen_mai-
25   2016_meldung.pdf?__blob=publicationFile&v=3, 2016 (in German).

Table 1: Features of 96 buildings affected by flooding in Braunsbach, Germany, recorded between 6th and 9th June 2016, and their frequency of occurrence.

Variable	Characteristics	n
Damage grade	D1 (no structural damage, slight nonstructural damage)	40
	D2 (no to slight structural damage, moderate nonstructural damage)	35
	D3 (moderate structural damage, heavy nonstructural damage)	5
	D4 (heavy structural damage, very heavy nonstructural damage)	6
	D5 (very heavy structural damage, very heavy nonstructural damage)	5
	No damage	5
	NA	0
Inundation depth (cm)	"Integer value"	90
	NA	6
House type	Single-family house	47
	Apartment building	25
	Semi-detached house	3
	Terraced house	0
	NA	21
Building material	Masonry	57
	Half-timbered	26
	Wood	9
	Concrete	2
	Steel	0
	Rubber	0
	NA	2
Building usage	Residential	59
	Commercial	8
	Combined/Mixed	21
	Public services	7
	NA	1
Near surrounding sealed	Yes	66
	Mainly yes	21
	Mainly no	8
	No	0
	NA	1
Exposition in flow direction	High	34
	Medium	35
	Low	27
	NA	0
Damage caused by debris	Yes	56
	No	38

	NA	2
Building condition before event	Good	45
	Medium	48
	Bad	1
	NA	2
Outbuildings present	Yes	32
	No	0
	NA	64
Type of outbuilding	Garage	11
	Carport	1
	Barn	9
	Shed	7
	Summerhouse	1
	Greenhouse	0
	Conservatory	0
	Other	6
Number of storeys	"Integer value"	96
	NA	0
Shop window	Yes	18
	No	75
	NA	3
Having cellar	Yes	30
	No	58
	NA	8
Estimated construction year	"Integer value"	90
	NA	6
Structural precaution	Higher ground floor	17
	Different (building) materials (of cellar and ground floor)	12
	Protection of cellar duct	3
	Other	6
	No precaution	59
	NA	4
Contamination visible	Yes	79
	No	15
	NA	2
Contamination type	Oil	4
	Chemicals	0
	Sewage	0
	Mud	79
	Other	0
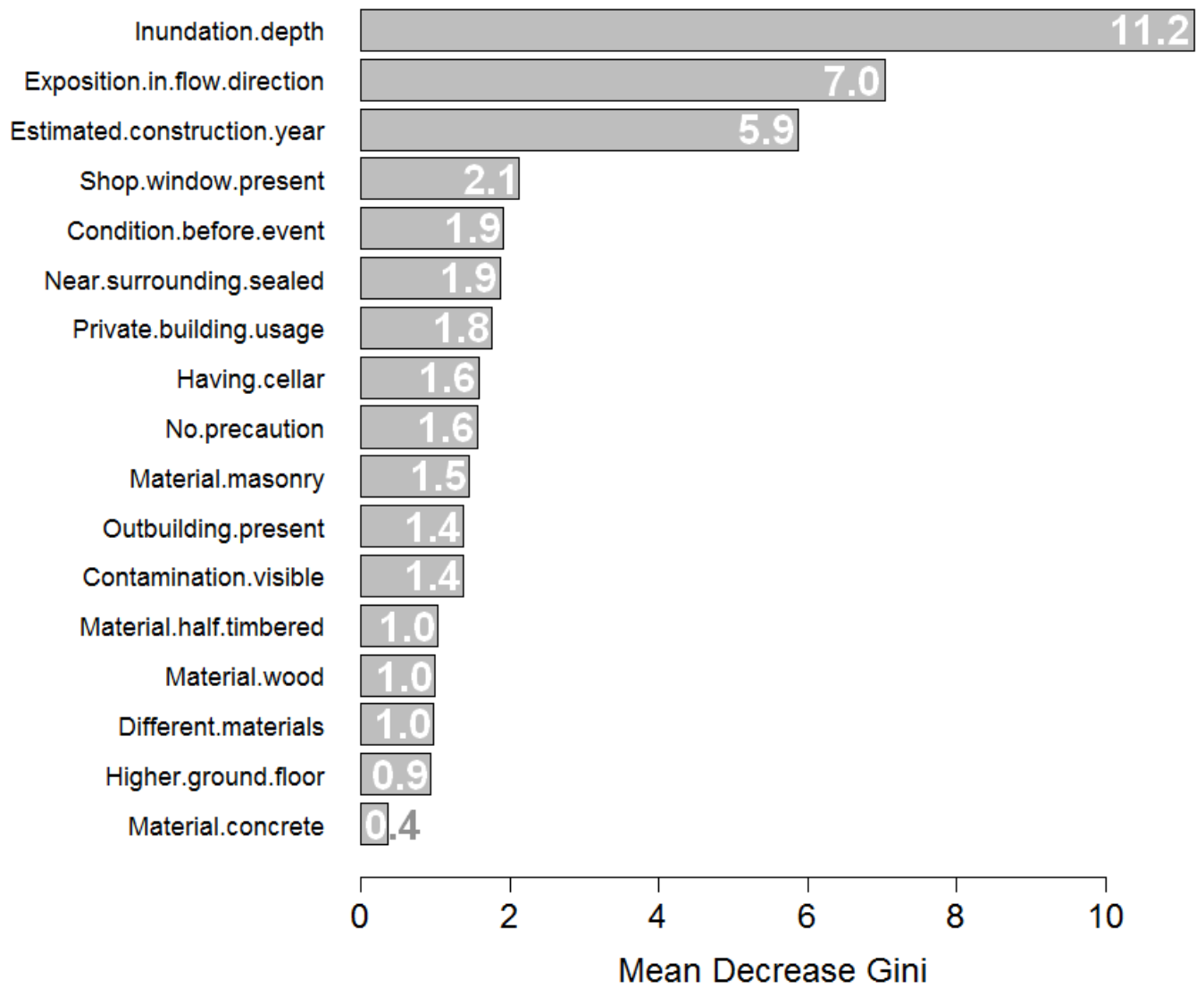
Figure 1: Random Forest feature importance (Mean Decrease Gini) for the damage grade prediction.
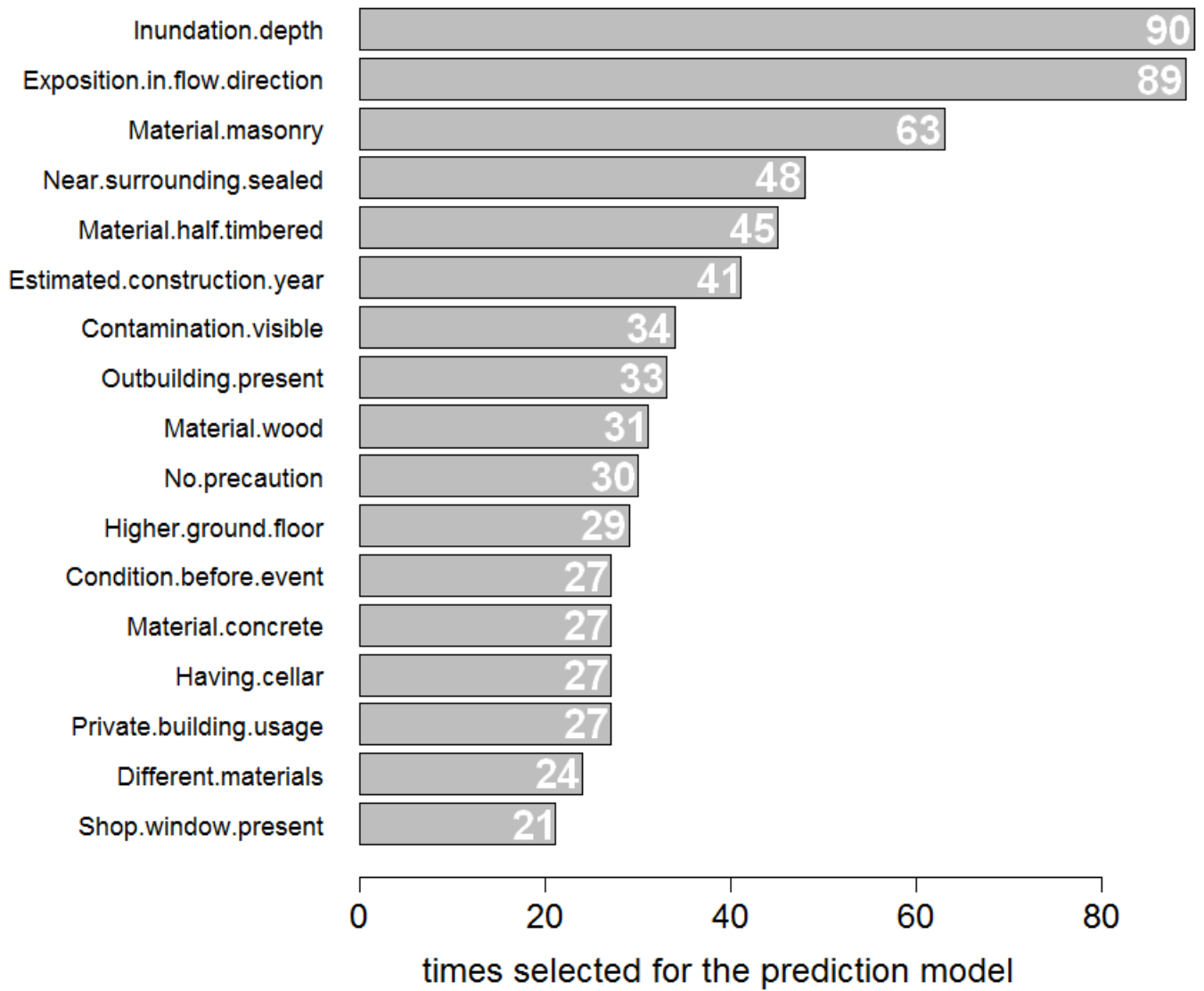
5

10

Figure 2: Random Generalized Linear Model feature importance (times selected) for the damage grade prediction.
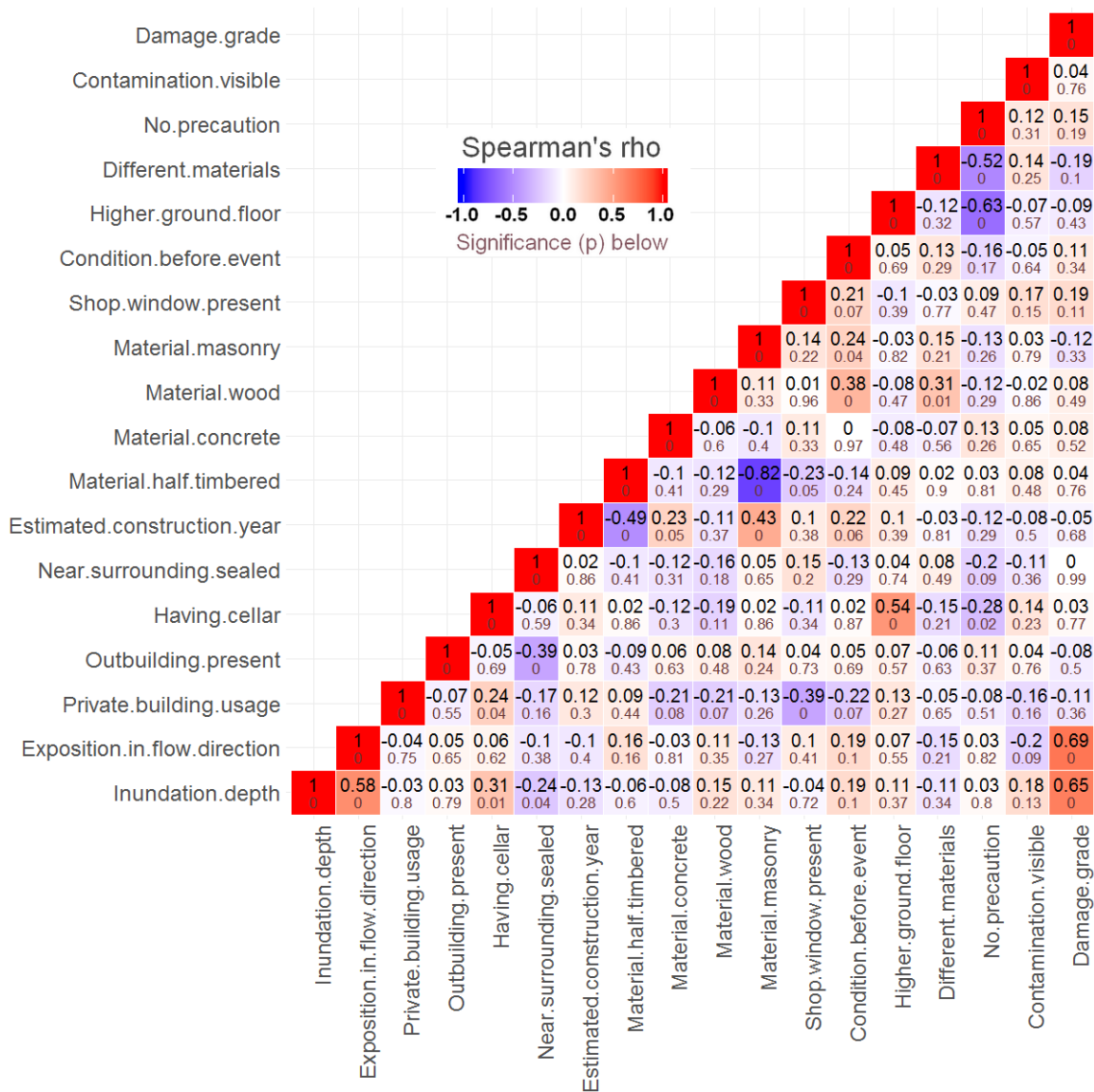
5

10

Figure 3: Spearman's rank correlation matrix and correlation significances of relevant variables (see Table 1 for a description of the variables).
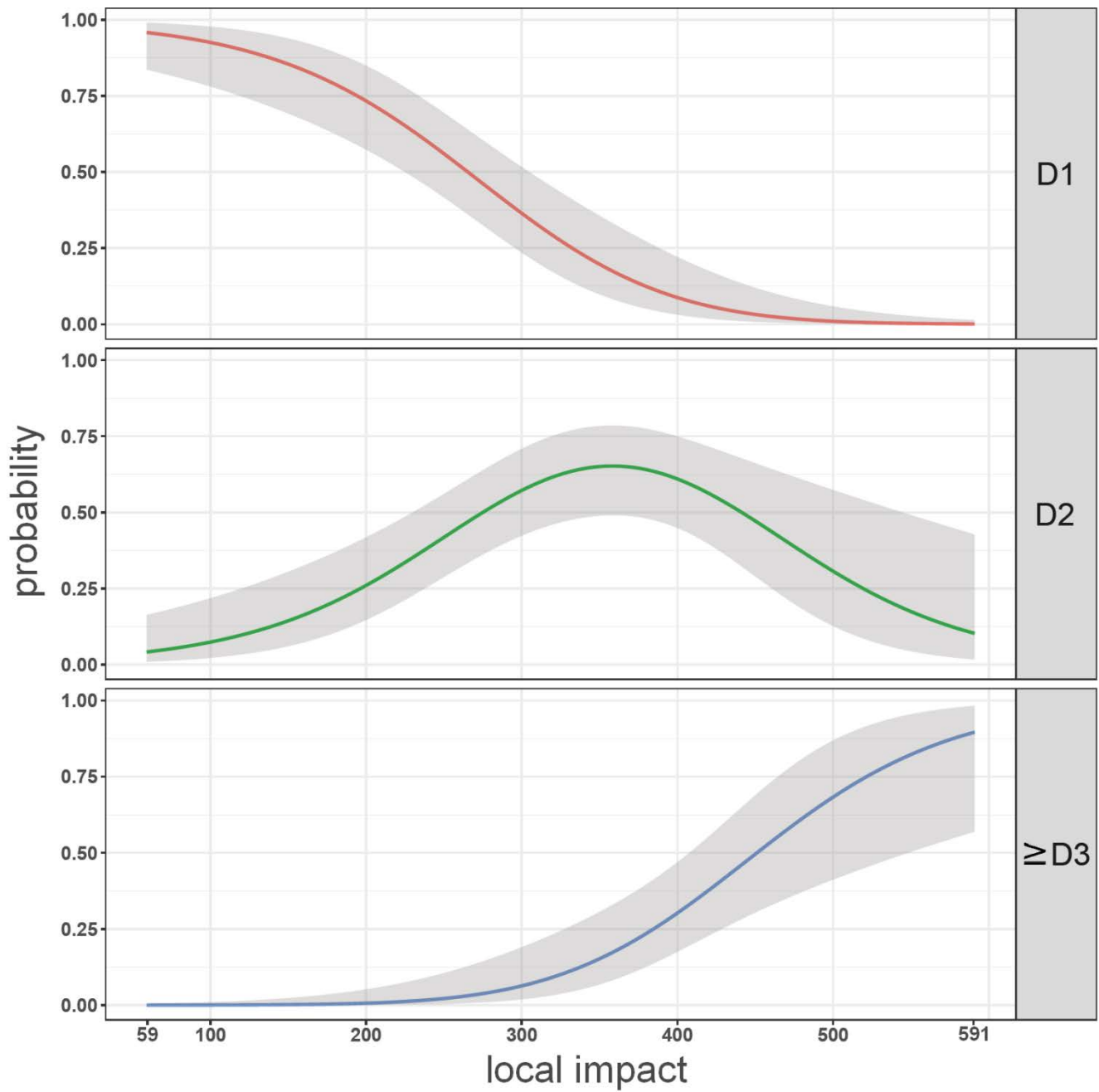
5

Figure 4: Probabilities of the damage grade predicted by the multinomial logistic regression model. It can be seen how the probability for a specific group affinity changes with an increasing local impact value and shifts towards higher damage grades.
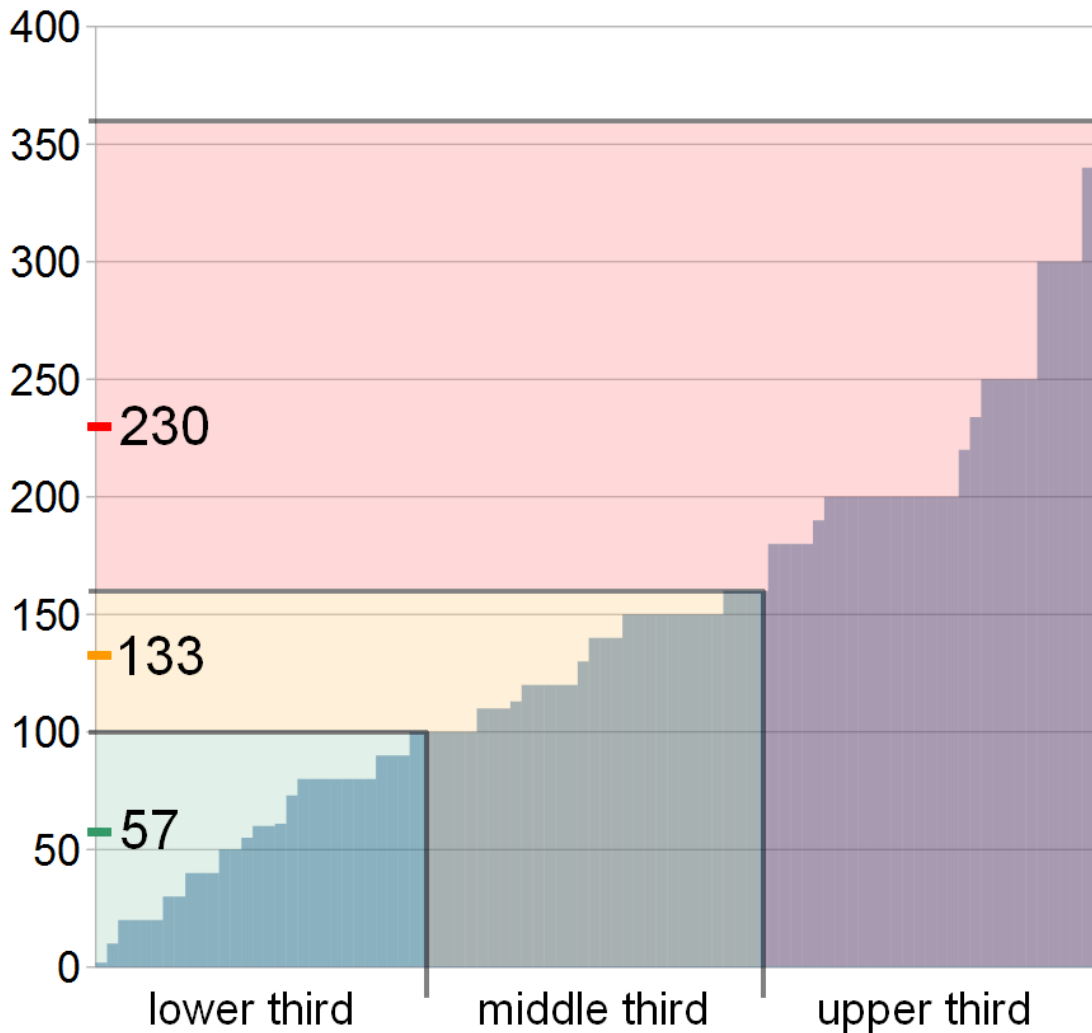
5

Figure 5: Step of deriving the local impact indicator. The recorded inundation depth is sorted in ascending order. By sorting, the relatively uniform distribution of the inundation values is shown, which allows the general procedure. On the left, the mean values of the lower, middle, and upper third of the sorted inundation depth (57, 133, 230) are given, which were used to replace the exposition classes "low" "medium" and "high". This step enables a comparable variable range and the derivation of an interval-scaled indicator for further analysis.
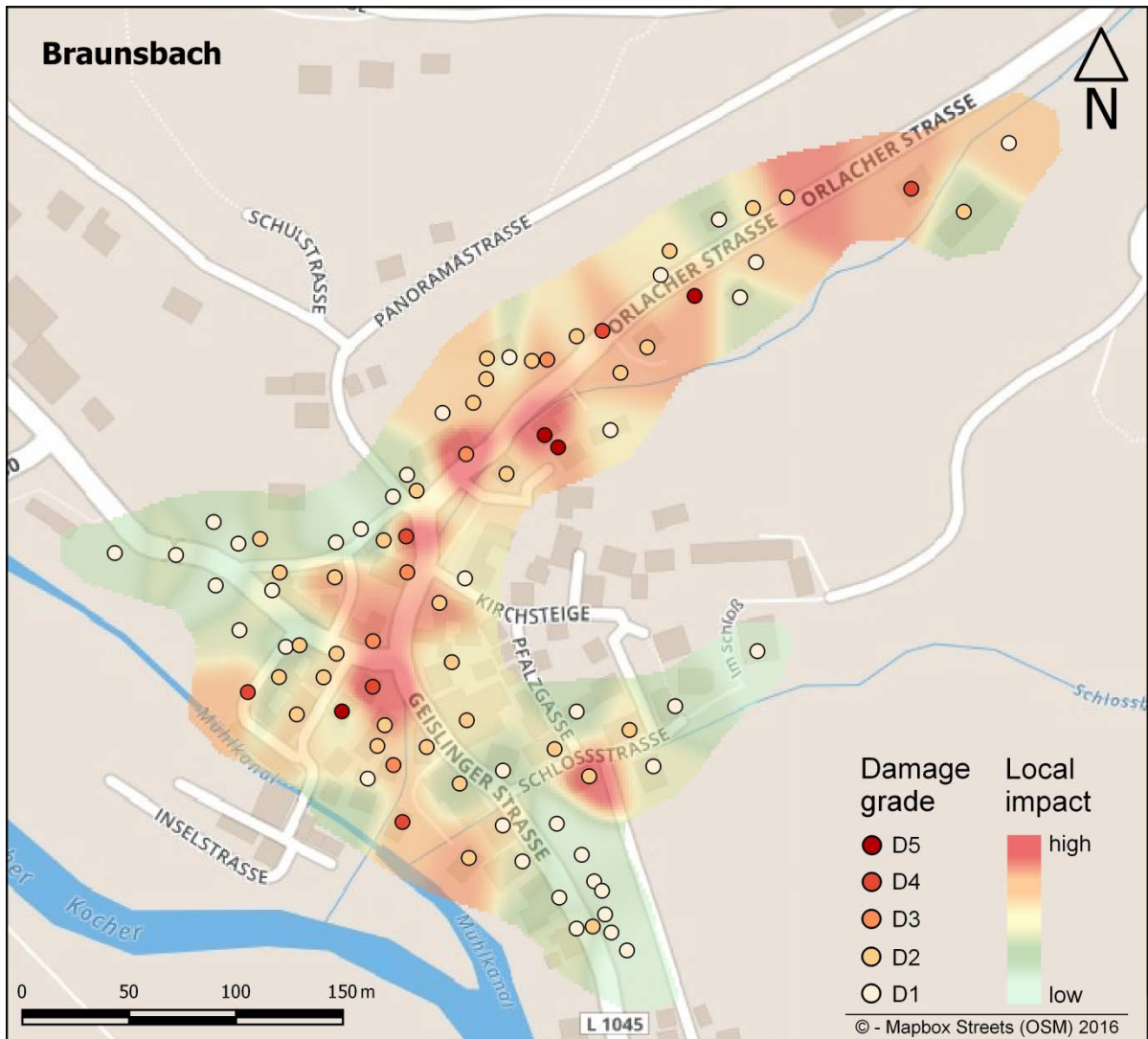
Figure 6: Map of the study area with the local impact, which is a combination of the inundation depth at the building and its exposition in flow direction (see text for further details). Further, the damage grades as recorded on site using the classification scheme of Schwarz and Maiwald (2007) are shown; see Table 1 for a verbal description of the damage grades.