International Journal of Disaster Risk Reduction 57 (2021) 102148

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

journal homepage: http://www.elsevier.com/locate/ijdrr

Review Article

Expert-based versus data-driven flood damage models: A comparative evaluation for data-scarce regions

Mark Bawa Malgwi [a,b,*], Matthias Schlögl [a,c], Margreth Keiler [a,b,d,e,f]

[a] University of Bern, Institute of Geography, Hallerstrasse 12, 3012, Bern, Switzerland
[b] University of Bern, Oeschger Centre for Climate Change Research, Hochschulstrasse 6, 3012, Bern, Switzerland
[c] University of Natural Resources and Life Sciences, Institute of Mountain Risk Engineering, Peter-Jordan-Str. 82, 1190, Vienna, Austria
[d] University of Bern, Mobiliar Lab for Natural Risks, Hallerstrasse 12, 3012, Bern, Switzerland
[e] University of Innsbruck, Department of Geography, Innrain 52f, 6020, Innsbruck, Austria
[f] Austrian Academy of Sciences, Institute of Interdisciplinary Mountain Research, Innrain 25, 6020, Innsbruck, Austria

ARTICLE INFO

Keywords:
Flood
Expert-based
Data-scarce
Data-driven
Damage prediction
Damage grades

ABSTRACT

The knowledge about potential flood damage is a key issue for disaster risk reduction. However, the scarcity of empirical data has limited flood damage modeling in several regions. As a result, studies in data-scarce regions have mostly been restricted to either building exposure assessment or identification of vulnerability indicators without a further linkage to probable damage. As expert-based approaches do not require empirical damage data, they have a high potential for flood damage modeling in data-scarce regions. In this study, we carried out a comparative assessment between an expert-based and a data-driven approach. The expert-based approach systematically combines the vulnerability indicator method and synthetic what-if analysis based on the knowledge of regional experts. The data-driven approach integrates empirical flood damage data in the analysis applying a multivariate random forest model. Flood damage data, collected through interviews after two flood events in 2017 and 2019 at separate locations in Nigeria, were used to evaluate the performance of both methods based on developed damage grades. Results from both methods showed i) a predictive accuracy of 30% and 38% for the expert-based and data-driven approaches respectively, ii) that distance to channel, wall material, building condition, and building quality are significant regional damage drivers, and iii) comparable model performance can be achieved even with a reduced number of variables. Furthermore, the study demonstrated how experts are likely to underestimate damage at low water depths and how a difference in conformity to building standards can add to challenges in flood damage prediction.

1. Introduction

Globally, flood disasters continue to give cause for growing concern. The increase in frequencies and severity of flood events likely driven by climate change [1] and changing exposure [2,3] has resulted in considerable human and economic losses [4]. The occurrence of flood events in vulnerable communities is even more critical due to low coping and adaptive capacities. Vulnerability relates to conditions that make communities prone or susceptible to disasters [5]. Large scale building damage caused by floods from Cyclones Idai and Kenneth in several regions in the southern part of Africa [6] further underlines the need to facilitate efforts for physical vulnerability assessment in vulnerable regions, which are mostly data-scarce [7]. For example, in the last decade, the observed increase in frequency and intensity of climate extremes in Africa [1,6] has consequently resulted in an increased number of fatalities and affected people [6]: a trend that is expected to continue given the impacts of climate change and the socio-economic development [1]. Flood hazards, in particular, accounted for over 64% of all 1164 hazard events recorded between 2000 and 2019 in Africa [6]. Despite huge losses to floods, studies on flood vulnerability for common building types typical in many African countries remain under-investigated. For example, while sandcrete block and clay buildings make up a high percentage of buildings in Africa [8], studies on their vulnerability to floods and damage patterns remain largely unknown [9]. Generally, studies within these regions have been limited to flood exposure assessment of the built environment (e.g. Refs. [10–13]) using globally available satellite date or identification of vulnerability indicators through literature reviews (e.g. Refs. [14,15]) without further relating

* Corresponding author. University of Bern, Institute of Geography, Hallerstrasse 12, 3012, Bern, Switzerland.
E-mail address: mark.malgwi@giub.unibe.ch (M.B. Malgwi).

https://doi.org/10.1016/j.ijdrr.2021.102148
Received 21 September 2020; Received in revised form 21 February 2021; Accepted 23 February 2021
Available online 26 February 2021

indicators with expected building damage.

In Nigeria, floods have become a yearly event with huge human and economic losses [9]. In 2012, Nigeria experienced one of the worst flooding on record with 28 out of the 36 states being affected [16]. The 2012 floods resulted in 364 fatalities, and over 3.8 million displaced people across the country [16]. An assessment of the 12 most-affected states from the 2012 flood showed that over 1.3 million houses were fully or partially destroyed with an estimated monetary value exceeding 6.8 billion USD [16]. The 2012 Nigerian flooding further exposed the considerable physical vulnerability of buildings and a stronger need for increased efforts in risk reduction targeted on the built environment in Nigeria. Such disaster risk reduction efforts are required even more urgently given observed trends in extreme events resulting from climate change and the current increase in population in exposed regions within the country [6].

Physical vulnerability assessment methods, such as flood damage models (stage-damage curves, multivariate models) and vulnerability indicators, explore the relationship between flood damage and corresponding damage influencing variables [17]. As a result, physical vulnerability assessment provides an important basis for physical resilience assessment and mitigation planning [18], evaluating economic losses [19], and cost-benefit analysis, which supports resource allocation for hazard protection [20,21]. In general, such efforts tailored to understanding and reducing vulnerability are considered important steps for disaster risk reduction [21]. Flood damage models either show the relationship between flood damage (or monetary loss) and water depth (referred to as stage-damage curves) or include other additional variables (referred to as multivariate models). While stage-damage curves use a continuous curve to relate water depth and damage state [22,23], multivariate models, which are commonly derived using empirical data, use different statistical approaches such as Bayesian network [24], regression and ensembles of bagged decision trees [25] or logistic regression [26]. Studies have shown that multivariate flood damage models have better prediction accuracy in terms of model transferability both spatially (between different regions) and temporarily (between different flood events) [27], and better explain the variance in damage data [28] compared to stage-damage curves that use only water depths. While most flood damage models are developed using empirical loss data (see Refs. [29–32]), scarcity of such data, especially in developing countries, has hindered the application of multivariate models, consequently resulting limiting the application of appropriate risk reduction strategies in such regions [33,34]. Although, synthetic stage-damage curves, developed using expert what-if analysis (for example [35–37]) provide a provisional alternative to empirical data, uncertainties related to damage prediction persist, especially as a consequence of using a single variable to predict flood damage (for further details, see Refs. [22,25,27,28,38]).

Recently, Englhardt et al. [34] presented one of the first studies that relate building vulnerability and absolute damage values for different building types within an urban and rural setting in Ethiopia. The study, which aims at large scale building damage assessment, uses a 15″ X 15” grid resolution database of building inventory from imageCAT (http://www.imagecatinc.com/) to reclassify buildings into vulnerability classes before applying stage-damage curves to estimate absolute damage using hazard maps with different return periods. The study by Englhardt et al. [34] represents an important step for understanding building vulnerability and damage assessment for selected building types in Africa. However, a limitation of the study by Englhardt et al. [34] is that out of 23 studies considered as a basis for generating the stage-damage curve, only two were from regions with comparable building types similar to those in many African countries. Such regional differences in building characteristics have been pointed to limit model transferability [17,39].

A combination of physical vulnerability methods was recommended by several studies [17,22,40] to systematically balance data-scarcity and model uncertainty. Recently, a method tailored to flood damage prediction in data-scarce regions was proposed [7]. This method is fully expert-based and relies on the deduction that buildings within the same resistance (vulnerability) class will incur comparable damage when impacted by the same level of hazard [39,41,42]. The method classifies buildings using an index generated from the vulnerability indicator approach (see Ref. [43]) and then implements a what-if analysis adapted for specific building vulnerability classes. The integration of the vulnerability indicator method allows the consideration of multiple damage influencing variables similar to the multivariate approach. Damage is assessed using damage grades that represent repeatedly observed damage patterns within a region [39] since they: i) allow temporal and spatial comparison of impacts between different regions more easily [44], ii) improve transferability of flood damage models [45], and iii) are comparable for similar building types [39].

In this paper, we implement and evaluate the performance of i) an expert-based flood damage model [7], and ii) a multivariate data-driven model using random forests. This paper aims to: i) assess the prediction accuracy of a fully expert-based model, ii) comparatively assess the expert-based and data-driven method, and iii) gain insights on main regional damage drivers typical for Nigeria. Both the expert-based and data-driven methods are evaluated using flood damage data collected from two separate study regions and flood events in Nigeria. This study provides one of the first attempts at using flood damage data for typical building types in Nigeria. The study demonstrates the potential of using expert-based methods for flood damage prediction in data-scarce areas and provides recommendations for improved performance.

2. Study regions and flood events

Two study regions are used for this study. Both are located in the north-central (study region 1) and north-eastern (study region 2) parts of Nigeria (Fig. 1). Although the regions are about 500 km apart, they share similar climatic characteristics typical of a guinea savannah with distinct dry and wet season [46,47]. Both regions are predominated by low-lying areas and small river channels. A light gray canvas base map provided by ESRI [48] is used to show the location of inspected buildings and river channels in both study regions.

Study region 1 is located between Suleja and Tafa in Niger State, Nigeria (Fig. 1 A). Located about 30 km north-west of the capital Abuja, the population in Suleja and Tafa is around 215,000 and 83,000 respectively [49]. The area has several river reaches passing through the settlement. Although some studies identified the region as non-vulnerable to floods based on terrain analysis (see Refs. [12,50]), recently, high-magnitude events occurred. Heavy rainfall was reported on 8 and June 9, 2017 resulting in severe flooding of settlements located in the area. The floodplains of all five river reaches were flooded (see Fig. 1). Several hundred people were affected and 18 fatalities claimed as well as substantial damage to hundreds of residential buildings and infrastructural facilities were reported. Losses from the flood event were mostly attributed to buildings constructed very close to the river channels and blockage of drainages as a result of transported material [13]. Google Earth satellite images from June 2016 (prior to the flood event) [51] show that the area is dominated by dense settlements and sparse grassland vegetation.

Study region 2 is located in Wuro-Jebbe in Yola-north, Adamawa State, Nigeria (Fig. 1 B). The population of Yola-north is 198,000 [49]. Wuro-Jebbe is about 4 km from River Benue, which is the second major rivers in Nigeria (see inset map, Fig. 1). The region is characterized by scattered grassland vegetation and a single river channel that passes through the settlements and flows downstream into river Benue (Fig. 1 B). High-intensity rainfall on August 1, 2019 resulted in the flooding of many regions in Yola-north [52]. Available reports compiled for all affected areas showed that 15 persons were reported to have died from the flood and 5000 persons were displaced [52]. Within the selected study region (Wurro-Jebbe), hundreds of houses were reported to have been damaged as a result of the flood. Locals in Wurro-Jebbe reported

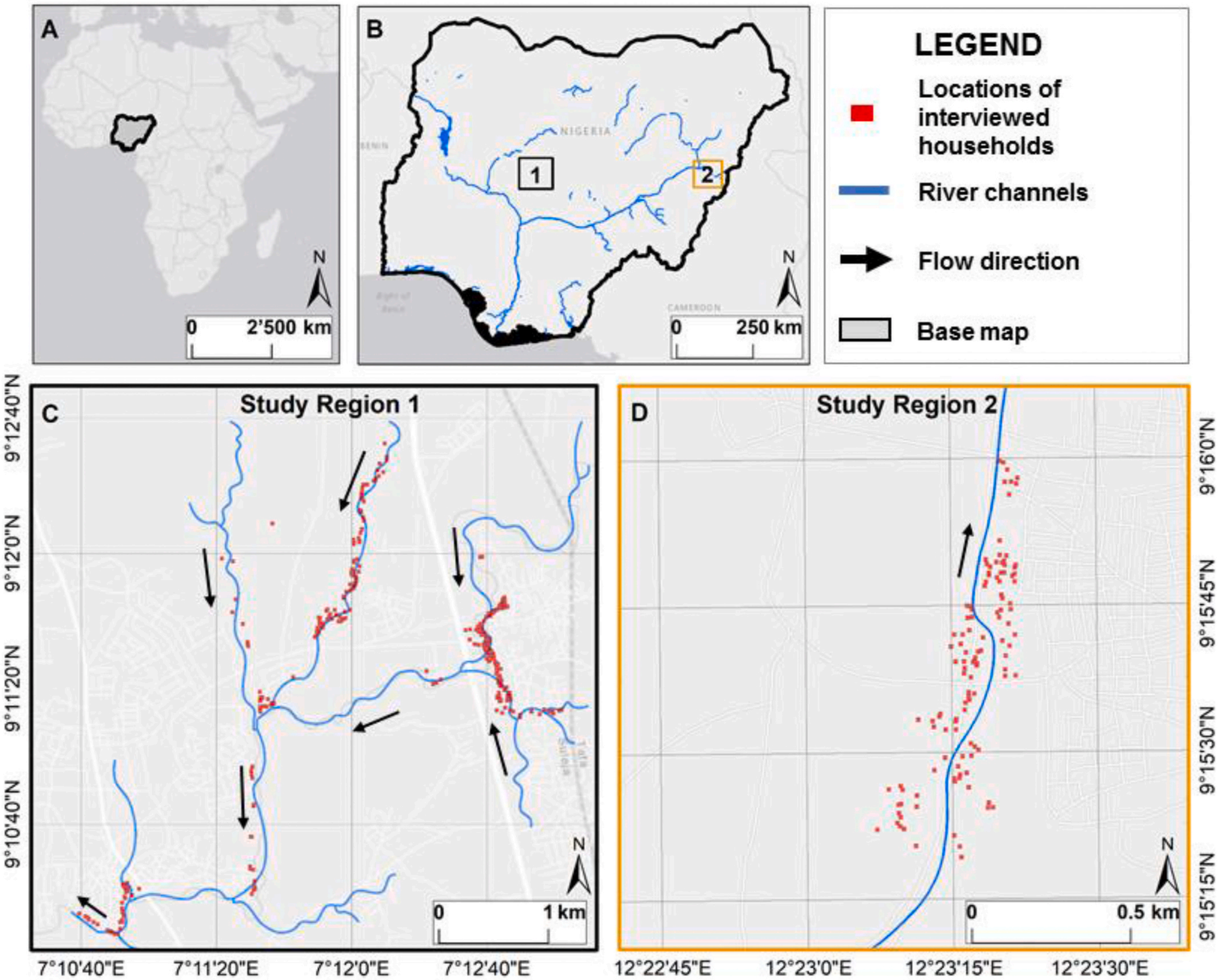


Fig. 1. (A, B) Inset maps showing the location of Nigeria, (C) Study region 1, and (D) study region 2. A base map (global light gray canvas satellite image) provided by ESRI [48] is used to show the location of inspected buildings and river channels.

that the rainfall lasted for about 5 h (between 12:00 and 17:00) and resulted in high flow velocities. Google Earth satellite image [51], taken in April 2019, shows that buildings are relatively sparse and very likely a more recent settlement compared to study region 1.

3. Methods

This study uses a combination of flood damage data from field interviews, expert interviews, and statistical analysis to gain insights into regional drivers and how they contribute to model performance. In this section, an overview on damage drivers and damage grades are first provided. Thereafter, data collection, pre-processing and analysis are further elaborated. Fig. 2 shows a detailed flow chart consisting of all components of the procedure adopted for model development.

3.1. Damage drivers

A general categorization of damage drivers into i) impact (or action), and ii) resistance variables, as proposed by several studies (see Refs. [53–55]) was adopted as shown in Table 1. A combination of literature review and expert knowledge was used to deduce damage drivers for the case study regions. The definition of all variables in Table 1 is presented in the Appendix. Action variables are related to flood hazard characteristics such as flood depth. Resistance variables relate to building and exposure characteristics that influence the degree of hazard impact on a building. In this study, we further categorized resistance variables into susceptibility, local protection, and exposure variables [7] (Table 1). Susceptibility relates to the inherent structural characteristics of a building without considering the measure for flood protection. Local protection variables refer to features of a building that directly or indirectly served to reduce the impacts of the flood. Local protection includes additional building features that are not necessarily required for the functionality (or stability) of a building but helps to reduce flood impact [56]. Exposure variables relate to the characteristics of the natural or built environment that can reduce or exacerbate the impact of floods on a building.

3.2. Damage grades

All observed damage patterns were systematically classified using an ordinal interval into a six-class damage grade based on an increasing level of severity (Table 2). The damage grades were adapted after Schwarz and Maiwald [57] with some modifications tailored for regional building characteristics. A pictorial representation of each

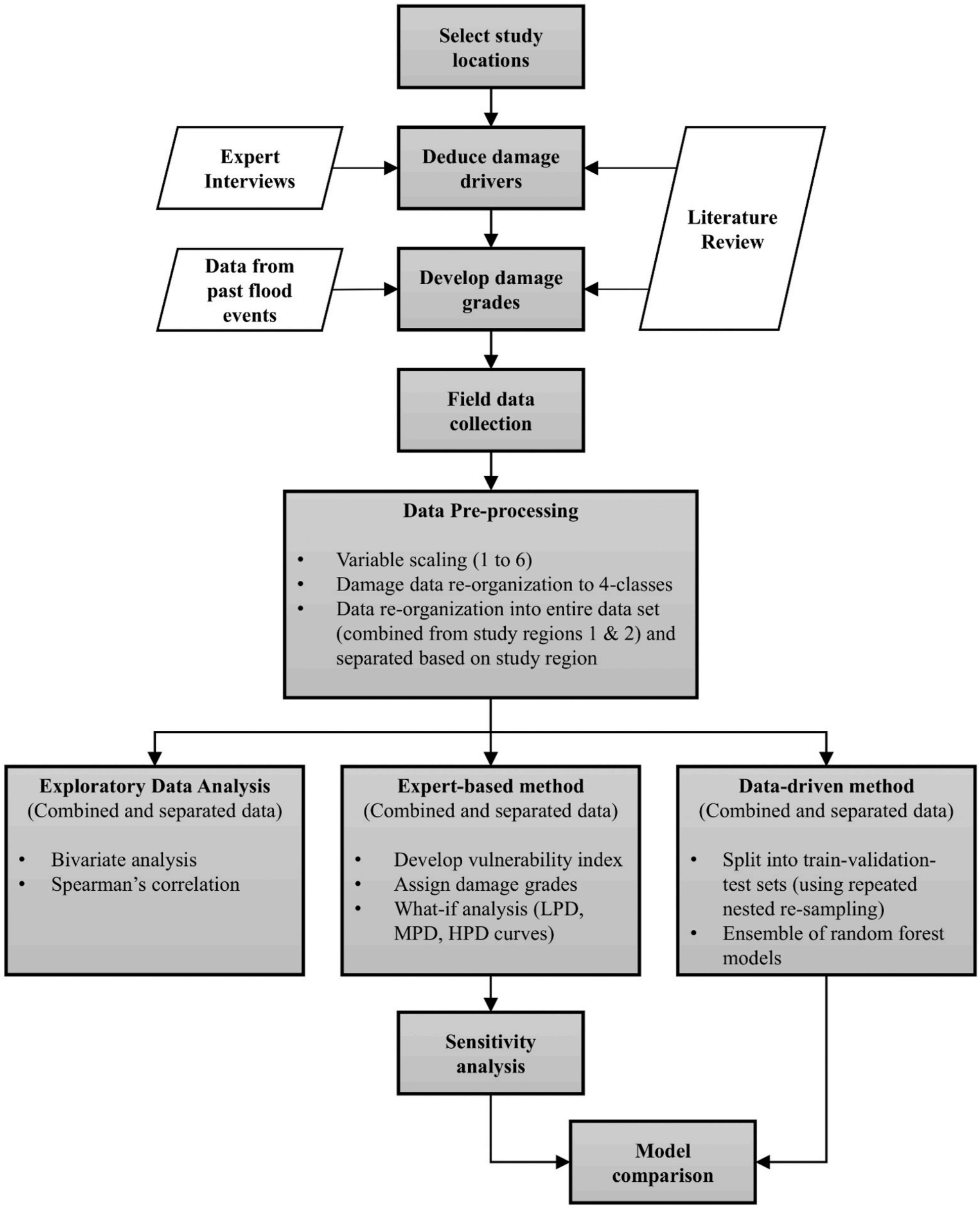


Fig. 2. Flow chart of model development.

damage pattern, compiled using literature review and assessment of flood damage data from past events, is presented in Table 2. Damage grade class 1 represents moisture defects or non-structural (only on finishes or plaster) cracks resulting from short or moderate contact with water. Moisture defects are associated with dampness and a weakening of surface material (finishes) (Table 2). Damage grade class 2 indicates de-bonding or peeling-off (falling away) of building finishes (Table 2) resulting from an extensive weakening of the surface material which can be caused by prolonged inundation. While damage grade class 3 represents light structural cracks, damage grade class 4 represents severe cracks on the building (Table 2). Both damage grade class 3 and 4 are structural cracks, hence, extend beyond the finishes and occur on the main building material. Damage grade classes 3 and 4 can occur vertically (resulting from movement of the soil), or horizontally, resulting from differential water pressure. In addition to severe structural cracks, damage grade class 4 includes damage to the ground flood material, which can occur due to compaction problems from the foundation or soil layer underneath (Table 2). Damage grade class 5 represents a partial collapse of the building usually resulting from overstress or increased weakness of the building. Usually, damage classified as 5 includes a collapse of less than or about one-third of the entire building (Table 2). Damage grade class 6 represents the collapse of the entire building or more than two-thirds of the structure.

3.3. Data collection

Field data collection consisted of a house-to-house interview using structured questionnaires. Before the data collection, preliminary assessments of the study regions and flood events using media reports, photos, and videos from different sources was carried out. Information gathered from preliminary assessments were used to locate and map affected areas using Google Earth satellite images. A sampling of buildings for interviews was limited to i) houses affected by the floods, and ii) availability of household owners (or a community representative) to provide the required information. Questionnaires were developed such that interview questions covered i) building and exposure characteristics, ii) flood characteristics, and iii) damage sustained and/or repairs done after the flood (see supplementary 1 (S1), questionnaire

Table 1
Summary of damage influencing parameters.

Category	Sub-category	Variable	Sub-variable	Abbreviation
Action variables		Flood depth		wat_hei
Resistance variables	Susceptibility variables (s)	Construction material	Wall material	wall_mat
			Wall thickness	wall_thic
		Building condition		buil_con
		Building quality		buil_qua
		Height of opening		hei_open
		Building footprint		buil_footprint
	Local protection variables (l)	Wall Finishes/cover	Wall plaster material	wallplas_mat
			Extent wall plaster	wallplas_ext
		Fencing	Fencing material	fenc_mat
			Extent of fencing	fenc_ext
		Ground floor elevation		grofloor_ele
	Exposure Variables (e)	Distance to channel		dis_chan
		Functional drainage		drainage
		Natural barriers		nat_barr
		Sheltering		shelt

Table 2
Damage grades for buildings in Nigeria (modified after Schwarz and Maiwald [57]).

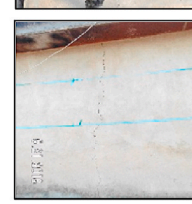
Class	Description	Example	
6	**Collapse** Complete collapse of the entire building or more than 2/3 of the building		
5	**Very Heavy** Partial collapse of building element		
4	**Heavy** Heavy structural cracks on building elements (walls, floors, beams, Columns) Settlement of ground floor material		
3	**Moderate** Slight to moderate cracks on building elements (walls, floors, beams, Columns)		
2	**Slight** De-bonding and (or) peeling-off of building finishes		
1	**Negligible** Water contact Moisture defects Surface cracks on floor or wall finishes		

Table 3
Building resistance variables with corresponding variable levels as defined for this study.

Variable	Variable levels (y)					
	6	5	4	3	2	1
Wall material (s_1)	Sandcrete		Mixed		Clay	
Wall thickness (s_2)	22.5			15		
Building condition (s_3)	3.5–4 (Very good)	2.6–3.5 (good)		1.6–2.5 (moderate)		1–1.5 (poor)
Building quality (s_4)	2.6–3 (Very good)	2.1–2.5 (good)		1.6–2 (moderate)		1–1.5 (poor)
Height of opening (cm) (s_5)	above 120		81–120			0–80
Building footprint (s_6)	rectangle	L-shape		Irregular		
Wall plaster material (l_1)	Cement/sand			Clay		None
Extent wall plaster (l_2)	complete			Partial		None
Fencing material (l_3)	Sandcrete	Clay or mixture of sandcrete and clay			Zinc/thatched	none
Extent of fencing (l_4)	complete			Partial		none
ground floor elevation (cm) (l_5)	81–100	61–80	41–60	21–40	5–20	0
Distance to channel (m) (e_1)	above 60		41–60	21–40		0–20
Functional channel (e_2)	yes			no		
Natural barriers (e_3)	yes			no		
Sheltering (e_4)	direct			Partial		none

1). Flood depths were either measured directly (where flood marks exist) or based on personal recollections from residents. A Garmin etrex 2000 handheld GPS device was used to track spatial location (longitude, latitude, and elevation above sea level) to allow for georeferencing of each inspected building (Fig. 1). Exposure variables such as distance to the river channel and sheltering were extracted (or supported) by information from Google Earth satellite maps. In study region 1, data collection was carried out between 1 March – May 31, 2018 (three months), i.e eight months after the flood event. In study region 2, the data collection started three months after the flood event and was conducted between 7 December – January 15, 2019 (approx. 1 month).

Observations from the field show that both study regions share a comparable building, floodplain, and channel characteristics. Buildings in both regions are predominantly one-story constructed from either sandcrete block, clay, or a mixture of both sandcrete and clay material. The river channels in both regions are semi-natural and shallow with channel bed elevation mostly between 0.5 m and 1.5 m. In both regions, the cross-sectional width of the river channels highly varies between 4 m and 18 m. Typical for guinea savannah, the rivers are dry in the dry season and surface run-off begins at the start of the rainy season (around April).

Damage data was documented using a questionnaire (c.f. S1, questionnaire 1) based on i) visual observations of damage (where the repair was not yet carried out), or ii) deduced from visual observations of repairs that were already carried out. Where new patterns were observed, which were not included in the preliminary damage list, an update is made. For example, ground floor settlement was included based on observations after the data collection in study region 1. Documented damage was specifically related to the building structure itself and not for external components such as fencing walls.

3.4. Data pre-processing and analysis

The overall data set consists of observations from study regions 1 and 2. Prior to using the data for analysis, few pre-processing steps were taken. Firstly, to maintain a consistent range across all variables (recommended in Ref. [59]), we implement a variable level (scaling) as shown in Table 3. A variable level is a form of internal weighting for sub-variables whereby low or high scores are assigned based on deduced influence on building vulnerability to floods [26,59,60]. In this study, a combination of literature review and author expert knowledge were used to assign variable levels using a score range between 1 and 6. The variable level is implemented in such a way that sub-variables, which are indicative of high vulnerability were given lower scores and sub-variables indicative of lower vulnerability are given higher scores (see Table 3). Secondly, a data reorganization was carried out which involves merging damage grade classes 1 and 2 (Class 1 + 2), as well as classes 5 and 6 (Class 5 + 6), while damage grade classes 3 and 4 are maintained as observed on the field. Merging damage grade classes were necessary given the low number of observations in each of the merged classes and low variation in explanatory variables between these neighbouring classes. In addition, merging the initial six-factor levels into just four levels facilitated model training, as this resulted in an almost perfectly balanced data set, thereby limiting bias of the data-driven model toward damage grade classes with higher observations. Merging the damage classes was also warranted from a contextual point of view, as they exhibit very similar damage patterns: Damage classes 1 and 2 represent minor damage including moisture defects or damage to wall finishes, and damage classes 5 and 6 represent severe damage including the partial or complete collapse of the building.

Furthermore, we carried out an Exploratory Data Analysis (EDA) to better understand the distribution of the data and empirical relationships between observed variables. Given that the scale of measurement of most variables is categorical, Spearman's ρ rank correlation coefficient was computed for all variables using a variable scale based on Table 3.

For the data analysis, two approaches were used: i) all observations were considered together and model training/validation/testing was performed on the full data set, and ii) the full data set was stratified by study region, using one subset for training/validation and the second subset for testing to check transferability from one case study to another. The combination of data from study regions 1 and 2 was plausible given the highlighted similarity in (i) channel and flood plain geomorphology (sec. 2 and 3.1), and (ii) reported inundation sequence in both regions. Combining data from both study regions is beneficial since it alleviates the problem of overfitting a single event [45]. Two methods were used to analyze the data sets: an expert-based and a data-driven approach. A confusion matrix was used to enable a systematic comparison between predicted and observed damage grades: it enables a simple visualization for the categorical response variable whereby the diagonal of the matrix represents correct predictions. Finally, from the confusion matrix, the predictive accuracy is computed as the ratio between the number of correct predictions and the total number of predictions (total number of buildings) such that a predictive accuracy of 1 represents 100% correct predictions. The predictive accuracy used in this study is very conservative such that only correctly predicted damage classes are considered for computing model accuracy.

3.4.1. Expert-based method

The expert-based method was based on three phases used to develop the model including i) development of vulnerability index, ii) development of damage grade, and iii) synthetic what-if analysis [7]. Experts selected to participate in the study were chosen based on their background in the area of building vulnerability to floods in Nigeria. Selected

Table 4
Table of influence for indicator weighting, ranging from slight influence of an indicator (1) to extreme influence (9) (modified after Saaty [58]).

1	2	3	4	5	6	7	8	9
Slight influence	Slight to moderate influence	Moderate influence	Moderate to strong influence	Strong influence	Strong to very strong influence	Very strong influence	Very strong to extreme influence	Extreme influence

experts were also from different geographical regions of the country (northeast, north-central, and south-west), and disciplines (such as geography, building or civil engineering, and environmental studies) so that evaluations received are representative from different expert communities within Nigeria.

In phase 1, the selection, weighting, and aggregation of indicators to form an index were carried out. Firstly, a literature review was used to provide a preliminary list of vulnerability indicators for Nigeria. Thereafter, we develop a questionnaire in which experts were requested to comment on the preliminary selected indicators and suggest other indicators that were not included. A second questionnaire was developed using a compiled list of indicators from different experts for indicator weighting (S1, questionnaire 2). Both the indicator selection and weighting was carried out by seven experts. A table of influence was provided (Table 4), containing a scale of influence, to maintain consistent weighting across different experts. Experts were requested to assign a weight for each indicator (variable) and indicator category: the weight is a quantitative value that represents the extent an indicator influences flood damage based on definitions from Table 4 adapted after Saaty [58]. Mean expert weights were calculated and a threshold of weight 3 (moderate influence), representing the first quartile of the influence indicator weighting scale (c.f. Table 4) was chosen to enable a final indicator selection for aggregation. A lower threshold value was not suitable given that it would result in selecting variables that on average were either having only slight (for a threshold of 1) or slight to moderate (for a threshold of 2) influence on vulnerability. Indicator aggregation was carried out using a simple weighted additive method (equation (1)) to form a building resistance index (BRI). The BRI sums the product between the weight of each variable and the variable levels (Table 3) for all exposure (e_i, y_i), susceptibility (s_j, y_j), and local protection (l_k, y_k) variables. Mean weights of variables representing exposure (e_i), susceptibility (s_i) and local protection (l_i) are shown deduced from expert interview. Corresponding variable levels (ranging between 1 and 6 on Table 3) are represented by y_i, y_j, and y_k for variables belonging to exposure, susceptibility and local protection respectively. Each variable category (exposure, susceptibility and local protection) is divided by the number of variables (n, p, k) (equation (1)) so that results are not biased towards categories with higher number of variables. Lastly, each variable category is multiplied by the corresponding weights ($Exposure_{weight}$, $Susceptibilty_{weight}$, $Local\ protection_{weight}$) (equation (1)). The BRI measures the resistance to flooding that a building can offer given its susceptibility, local protection, and exposure. A maximum-minimum normalization (see Ref. [59]) was implemented to confine the upper and lower bounds of the BRI between 0 and 100 (equation (2)). The maximum-minimum normalization uses the BRI calculated from equation (1) and a minimum and maximum BRI (BRI_{min}, BRI_{max}) values which are computed using the minimum variable level ($y_{\min}$) and maximum variable level ($y_{\max}$) for all variables. The formulae for calculating the BRI_{min}, BRI_{max} are shown in the appendix in equations A4 and A5 respectively. Using a quartile classification, the normalized BRI values (BRI_{norm}) are classified into buildings with poor (low), moderate (average) and good (high) resistance to floods.

$$BRI = \frac{Exposure_{weight}}{n}\sum_{i=1}^{n}(e_i y_i) + \frac{Susceptibility_{weight}}{p}\sum_{j=1}^{p}(s_j y_j) + \frac{Local\ protection_{weight}}{r}\sum_{k=1}^{r}(l_k y_k) \quad (1)$$

$$BRI_{norm} = 100\left(\frac{BRI - BRI_{min}}{BRI_{max} - BRI_{min}}\right) \quad (2)$$

Phase 2 relies on the classification of commonly observed damage patterns into damage grades. Building damage patterns used in this study (sec. 3.2.1) were developed from a combination of literature review, building damage reports, and evaluation of field data. Additional details on damage grades are given in section 3.2.2 and Table 3.

In phase 3, a synthetic what-if analysis for the three BRI classes was carried out. For selected representative buildings in each BRI class, seven experts were asked to predict expected building damage patterns using synthetic flood depths 1–5 m at 1 m intervals (see S1, questionnaire 3). For each flood depth interval, experts predicted three damage states; i) Low Probable Damage (LPD), ii) Most Probable Damage (MPD), and iii) High Probable Damage (HPD). While the LPD and HPD define the lowest and highest likely damage expected, the MPD defines the most likely damage. Predicted damage from different experts was used to compute a single mean damage grade class per flood depth interval for each BRI class. The mean damage grade represents an average expected damage given a vulnerability class (BRI class) and specific water depth interval [57]. A Bayesian ordered logistic regression model [61] was used to fit mean damage grades to water depth. The model fit was implemented by leveraging the 'arm' package in R. Specifically, the function 'bayespolr' was used, employing a logistic link function to assign class probabilities for each damage grade through a maximum likelihood approach. Resulting damage grade probabilities for synthetic water depth intervals represent damage curves for the three damage states MPD, LPD, and HPD for all BRI classes.

A model performance check was carried out using the MPD curves given that they define the most likely damage expected. The MPD curves for each BRI class is used to predict damage grade class for the full data set (combined data from study regions 1 and 2). Damage grade with the highest probability is assigned as the predicted class for the building. Separate confusion matrices were generated the three BRI classes, which were later merged using a matrix addition to generate one confusion matrix.

3.4.2. Data-driven method

The data-driven approach is based on ensembles of random forest models [62] carried out on the full data set (combined from study regions 1 and 2). To obtain robust models and unbiased estimates of model generalization performance, a modeling strategy featuring repeated nested resampling was pursued. Since honest model quality assessment is only possible if all elements of model building are included in the resampling procedure, models that require hyperparameter optimization necessitate two nested resampling loops. The outer resampling, which provides information for model performance assessment, was realized by means of five-fold cross-validation. In the inner resampling, which is targeted at hyperparameter tuning, out-of-bag predictions were used for evaluation. Hyperparameter tuning was carried out by means of model-based optimization [63,64]. In order to minimize the variance of random partitioning within folds, the whole nested resampling procedure was repeated ten times.

3.5. Further analysis

To additionally evaluate collective outputs of both the expert-based and data-driven method, few additional steps were taken. Firstly, the selected threshold for indicator importance was chosen at weight 3

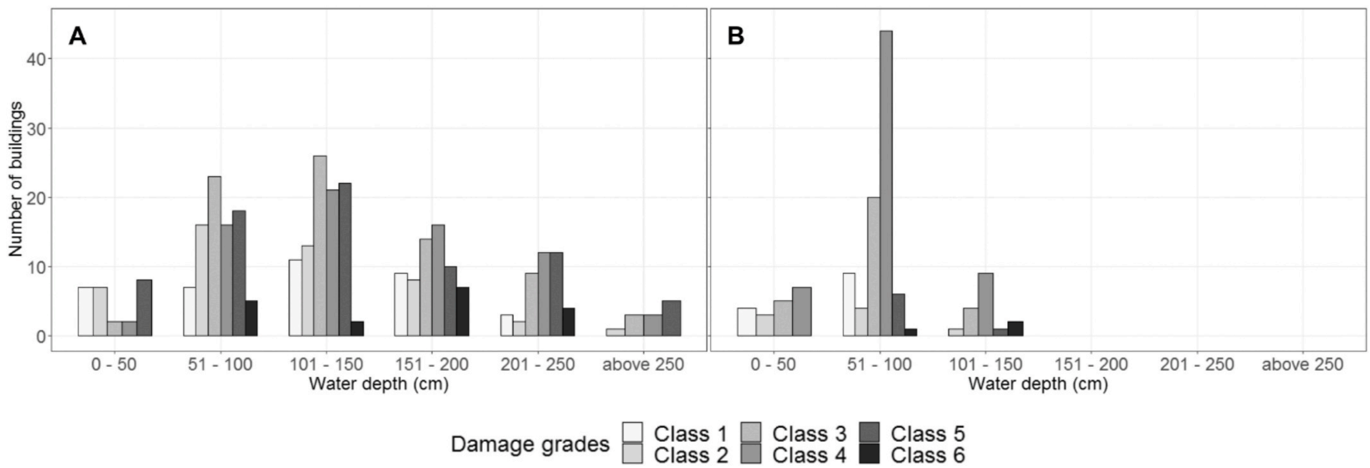


Fig. 3. Bivariate plots for water depth and damage grade classes for (A) study region 1, and (B) study region 2.

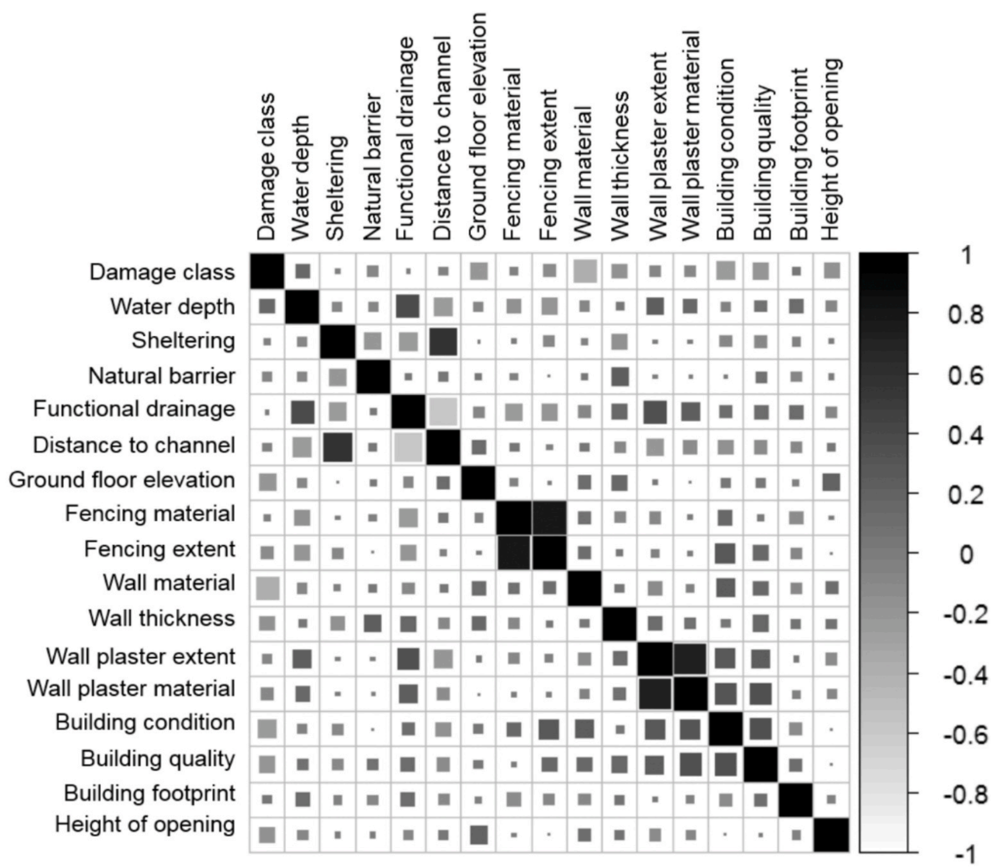


Fig. 4. Spearman's correlation for the entire data. Areas of squares represent absolute values of correlation coefficients.

(Table 4) to ensure that variables with low mean weights were not included in the final indicator selection and index aggregation. To evaluate the uncertainty of the selected threshold was optimal, we carried out a sensitivity analysis for different threshold values and re-evaluate model performance. To do this, we varied the threshold between 4, 5, 6, and 'no threshold', and in each case, we recalculated the BRI value, reclassified each building into a BRI class using the quartile classification and predicted damage grades using the MPD curve. The sensitivity analysis was carried out using both the combined and separate data from the two study regions. Secondly, to further examine the transferability of the data-driven model, we use data from study region 1 to train a random forest model using the same procedure described in section 3.3.2. The prediction accuracy of the model is evaluated thereafter using data from study region 2. Furthermore, we carried out a critical comparison between both methods, focusing on variable importance and model performance.

4. Results and discussion

4.1. Exploratory data analysis

A total of 324 observed buildings are located in study region 1 and 120 buildings in study region 2. The maximum inspected flood depth in study region 1 is 333 cm compared to study region 2 with 147 cm.

Bivariate analysis showed that water depth correlates rather poorly with damage grades (Fig. 3). For example, in study region 1, over 40% of buildings with a damage grade 6 (complete collapse) were from the lowest category of water depths (0–50 cm) (Fig. 3A). Also, out of 12 buildings with water depths above 250 cm, none sustained a damage grade 6 (Fig. 3A). In study region 2, most of the buildings incurred damage grade 4 especially at water depths between 51 and 100 cm (Fig. 3A). In particular, damage grade 4 for in study region 2 is mostly as a result of damage to the ground floor (Table 2). Spearman correlation coefficient for the combined data showed a high positive correlation between sheltering and distance to channel (Fig. 4). The correlation can partly be explained by the fact that the more distanced a building is from a river channel, the more likely it will be surrounded by other buildings. Other high correlations observed were between wall plaster with wall plaster extent and fencing material with fencing plaster extent. However, these are directly related since only buildings with wall plaster (or fencing material) have features for wall plaster extent (or fencing extent). A summary of the bivariate analyses for all resistance variables for both study regions and the combined data is presented in supplementary 2 (Figure S2-1 and S2-2). Results show that sandcrete block buildings are predominant in both regions followed by the mixed building (sandcrete and clay) and clay buildings. Another interesting result of the EDA was that while building condition and building quality were positively correlated in study region 1, they were however negatively correlated in study region 2. The reason for such difference is not completely clear, however, the negative correlation in study region 2 might be related to the fact that it is relatively a new settlement (10–20 years) (as visualized from Google Earth satellite image from 2004). As a result, while many buildings were identified to be in a good condition (recently built), their construction quality was not well according to standard specified by NBC [65]. Similar problems relating to building

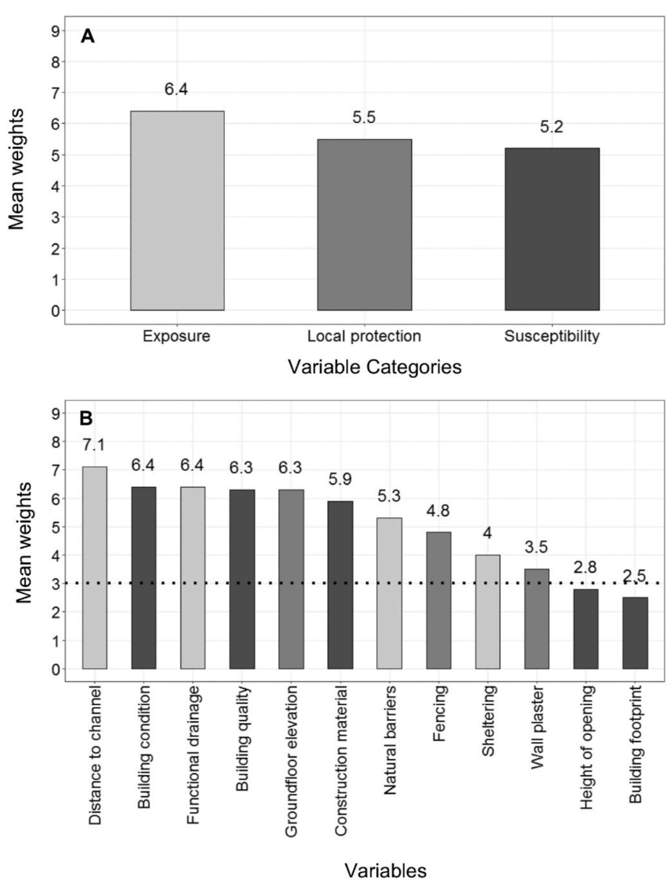


Fig. 5. Mean expert weights for (A) variable categories, and (B) variables (or indicators). Dotted line in plot A indicates the selected threshold.

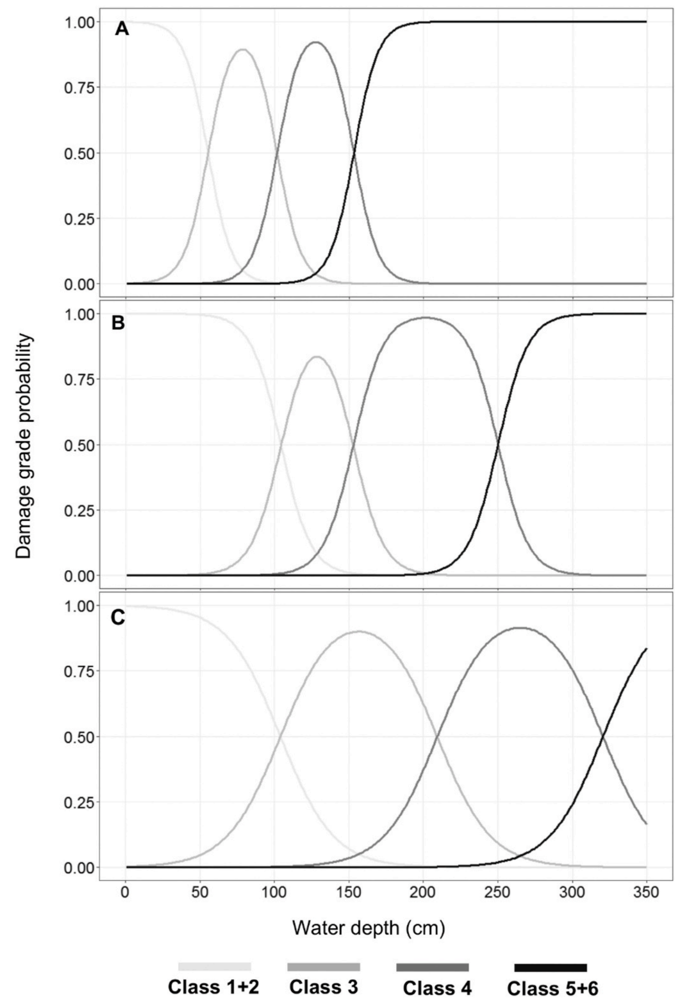


Fig. 6. Most Probable Damage (MPD) curves for (A) poor, (B) moderate, and (C) good BRI class.

quality were identified in studies by FGN [16] where over 60% of inspected buildings were found to have built their houses without using the services of formal institutions.

The distribution of building damage grades in the data is shown in Figure S2-3. Damage grade 6 was generally less represented in both data sets (Figure S2-3A): this was partly because in many completely damaged buildings, residents were no longer available or no community representative could ascertain the flood depth at the building location. Merging damage classes 1 and 2 (class 1 + 2), and classes 5 and 6 (5 + 6) (Figure S2-3B) resulted in a distribution that limits bias from over (or under) representation especially for the combined data set.

4.2. Expert-based method

The list of indicators selected by experts is given in Table 1. Mean indicator weights, representing a ranking of variable and variable category importance are presented in Fig. 5. The results for mean weights assigned to variables categories (exposure, susceptibility, and local protection) are shown in Fig. 5A. The exposure variable was identified to have the highest influence on damage followed by local protection and susceptibility. An additive aggregation of indicators resulted in the calculation of the BRI value for each building (equation (1)). For the variables (Fig. 5B), distance to channel has been identified by experts to have the highest influence (weight of 7.1) on building damage to floods in Nigeria. Distance to channel is followed by building condition and functional drainage (weight of 6.4) and building quality

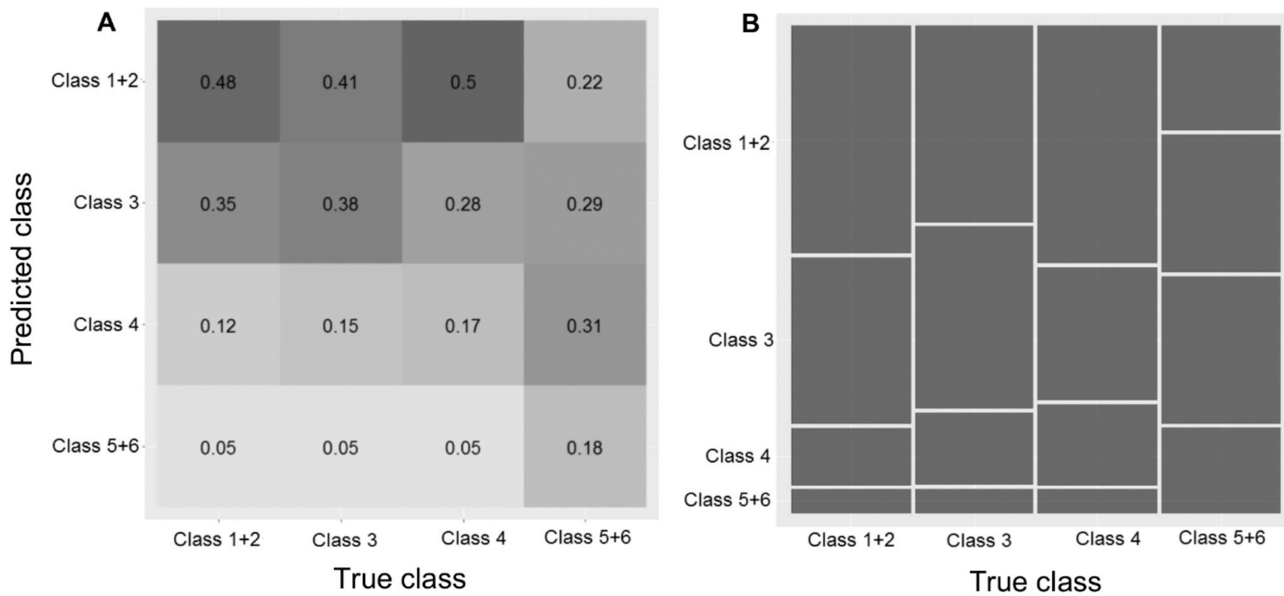


Fig. 7. Confusion matrix for observed and predicted classes using the MPD curve. The heatmap (A) indicates the percentage of values in each cell with respect to the total number of instances in the true class (i.e., columns sum up to 1). The mosaic plot (B) is a graphical illustration of the conditional relative frequency for each combination. The area of the tiles is proportional to the number of observations exhibiting the respective combination of factor levels, i.e. the corresponding joint frequency.

and ground floor elevation (weight of 6.3). A threshold (cut-off) for final indicator selection (Fig. 5B, dotted line) shows that two variables, height of opening and building footprint, were on average considered to have only a slight influence on damage. As a result, they were not included for indicator aggregation. Variables included in the final indicator selection (Fig. 5B) represent one of the first attempts at compiling a comprehensive list of vulnerability indicators to floods for Nigeria. The additive aggregation allows compensation such that lower values for one variable can be compensated by another variable with a higher value. Results of the quartile classification for the normalized indices for all data sets are shown in Figure S2-4. The quartile classification categorizes all buildings into four equal classes with the lower and upper quartile reassigned as poor and good BRI classes respectively. The buildings in the interquartile range are reassigned as moderate BRI class. The normalized BRI has a range between 10 and 70 with a mean value of 40.

Synthetic flood damage curves for most probable damage (MPD) are shown in Fig. 6 for BRI classes poor, moderate, and good. The curves were generated using probabilities estimated by the Bayesian ordinal logistic regression model.

A maximum of 350 cm was used for the curves given that buildings in Nigeria are predominantly one story [49]. The damage curve shows that a damage grade of class 3 and above is expected for buildings with poor

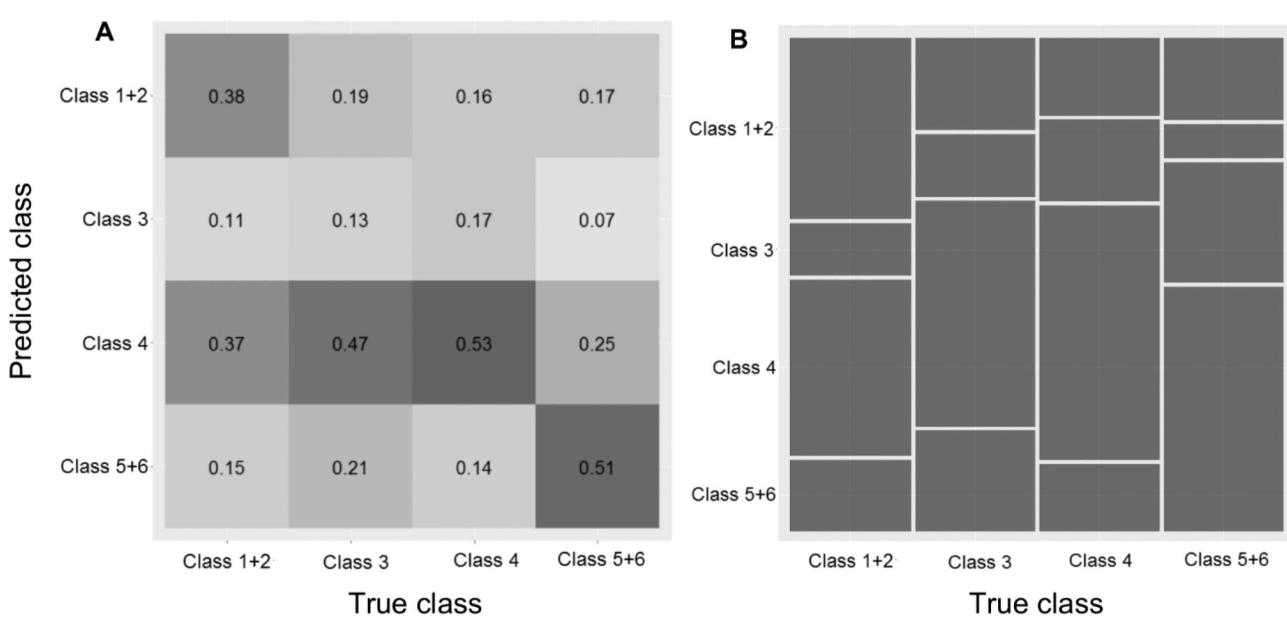


Fig. 8. Confusion matrix for observed and predicted classes using random forests. The heatmap (A) indicates the percentage of values in each cell with respect to the total number of instances in the true class (i.e., columns sum up to 1). The mosaic plot (B) is a graphical illustration of the conditional relative frequency for each combination. The area of the tiles is proportional to the number of observations exhibiting the respective combination of factor levels, i.e. the corresponding joint frequency.

BRI and water depths above 50 cm (Fig. 6A). Buildings identified as BRI moderate and good are expected to have damage grades class 1 + 2 for water depths between 0 and 100 cm (Fig. 6 B, C). While the range for damage grades 3 and 4 are small for the category BRI poor (50–150 cm), the range is wider for BRI moderate (100–250 cm), and good (100–300 cm). Complete or partial collapse (damage grade 5 + 6) is expected for water depths above 150 cm for building categorized as poor BRI: the same damage grade is expected at water depth above 250 cm (for BRI category moderate) and 320 cm (for BRI category good).

The multiclass prediction accuracy for the expert-driven method on the full data set is 0.30. Heatmap and mosaic plots (Fig. 7) generated from the confusion matrix showed that in general, the MPD curve performed relatively well for low damage classes (1 + 2, 3) compared to higher damage classes (4, 5 + 6) (Fig. 7A).Correct predictions for damage class 1 + 2 were relatively high at 48%, while that of damage class 3 is at 38%. Prediction accuracy dropped to 17% for damage class 4 and 18% for damage class 5 + 6. The MPD curve incorrectly predicted 41% of damage class 3 as class 1 + 2 (Fig. 7A). In addition, half of the buildings with damage class 4 were incorrectly predicted as damage class 1 + 2. Further evaluation of the individual confusion matrix generated for BRI classes poor, moderate, and good (Figure S2-5) showed that a substantial part of the misclassifications was resulting from buildings observed as damage classes 3 and 4 being predicted as class 1 + 2: this is especially high for BRI classes moderate and good with about 50–60% misclassification. In BRI poor, the main misclassification is from observed damage classes 1 + 2 and 4 being predicted as class 3. Relatively high accuracies in the classification of low damage grades are likely because a high percentage of the observed water depths were less than 150 cm (Fig. 3), and within this range, the damage probabilities of the MPD curve is high for low damage (see Fig. 7). Conversely, poor accuracy in predicting high damage grades is related to the underestimation of damage grade classes at low water depths. From the MPD curve (Fig. 7) higher damage was only assigned for high water depths and low damage for low water depths. However, as seen from the bivariate analysis (Fig. 3), higher damage occurs even at lower water depths. The use of a water depth range between 1 and 500 cm for the expert what-if assessment might have influenced the results given that experts become likely to associate higher water depth (400–500 cm) with the higher damage grades (class 6).

Results for the mean low probable (LPD) and high probable (HPD) damage are shown in Appendix, Figure A1. Generally, LPD and HPD aim to accommodate variations in building characteristics within each BRI category. For example, for buildings classified as BRI poor, the lowest damage expected for a 200 cm flood depth is damage class 4, while the highest damage expected is a damage grade 5 + 6 (Figure A1 A). For the same flood depth (200 cm), we expect the lowest damage of class 4 for BRI moderate or class 3 for BRI good (Figure A1 C, E). The highest probable damage for a 200 cm flood is a class 4 or Class 5 + 6 (both have equal probabilities) and a class 4 for BRI good (Figure A1 D, F). Between 150 and 250 cm water depth range, the LPD and HPD for BRI moderate overlap, both predicting damage class 4 (Figure A1 B). Generally, for most of the LPD and HPD curves, a change in damage grades class consistently occurs at around 100 cm. The change might be related to the height at which water starts entering the building: in Figure S2-1O, over 60% of the entire data have a height of openings between 81 and 120 cm.

4.3. Data-driven method

The data-driven approach results in a multiclass prediction accuracy of 0.38. However, results vary across the different classes as depicted by the heatmap and mosaic plot resulting from the confusion matrix (Fig. 8). Generally, the model performed well for severe damage (Classes 4, 5 + 6) compared to lower minor or moderate damage (class 1 + 2, 3).

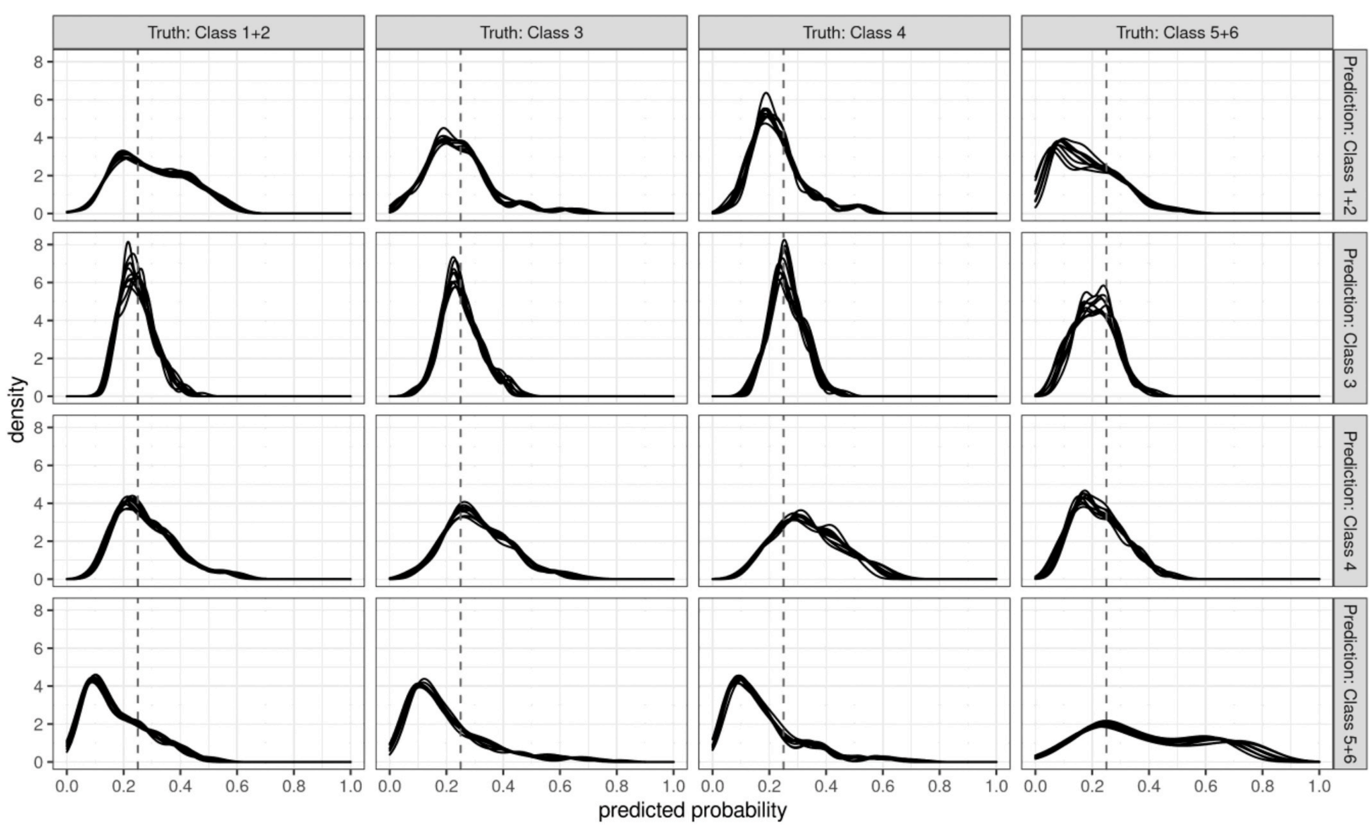


Fig. 9. Densities of predicted class probabilities for all class combinations. Results are presented for all ten models obtained via the repeated nesting resampling procedure. The vertical dashed line at 0.25 indicates the threshold for random guessing.

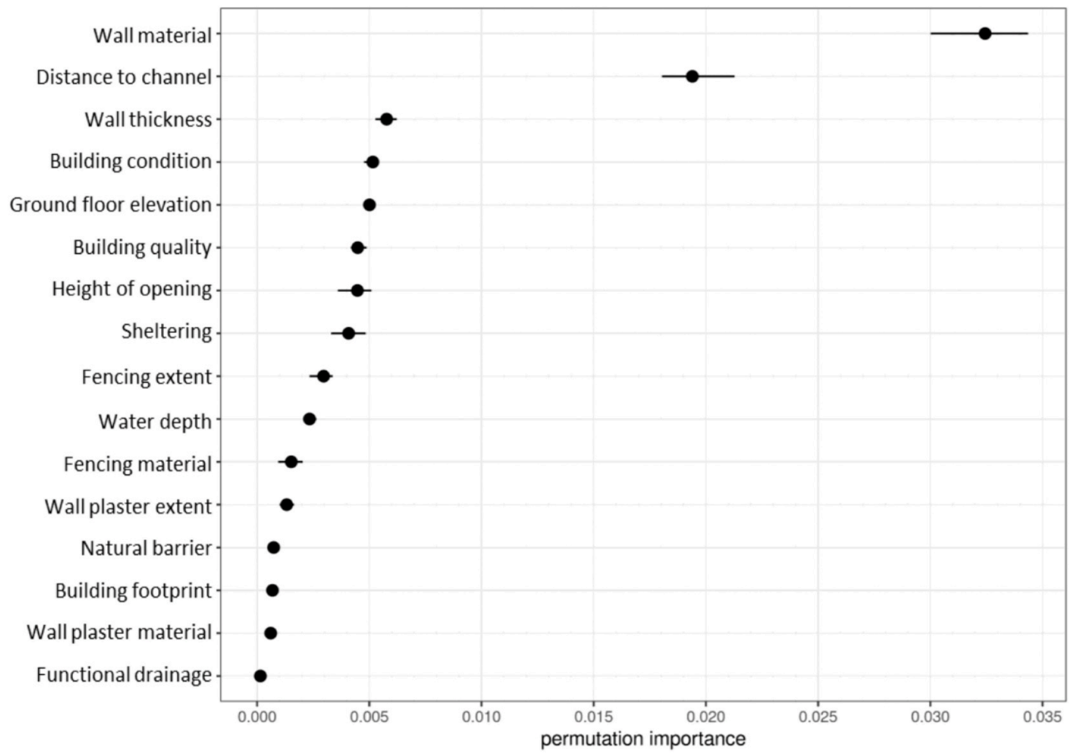


Fig. 10. Random forest ensemble variable importance. Both the mean (black dot) and the range across the ten-member ensemble are depicted.

While more than 50% of all instances of classes 4 and 5 + 6 are predicted correctly, accuracy drops to 38% for class 1 + 2 and down to only 13% for class 3 (Fig. 8). About half of the buildings observed as class 3 (47%) were predicted as class 4. Also, high number of class 1 + 2 is predicted as class 4 as well. Density estimates (Fig. 9) of predicted class probabilities for all combinations of the confusion matrix show interesting patterns. Some correctly classified instances of class 5 + 6 values are predicted with relatively high certainty. Predictions for class 4 exhibit a mean clearly above 0.25 (which would correspond to random guessing given a classification problem with four classes) (Fig. 8), with predicted class probabilities up to almost 0.7. Class 3 exhibits a very balanced prediction across all four true classes, with distinct peaks around or slightly below 0.25, reflecting the overall low prediction of class 3 instances. Among all class 3 predictions, most belong to observed class 4. Predicted class probabilities for class 1 + 2 exhibit the best result for true class 1 + 2 observations, with class probabilities up to 60%, but also show between 15% and 20% false predictions in each of the other classes with single cases exhibiting similar class probability of up to 60%.

In terms of variable importance, wall material clearly prevails as the most important variable (Fig. 10). Distance to channel emerges as the second most important variable, exhibiting a large importance gap not only to wall material but also to wall thickness, which is ranked third (Fig. 10). The identification of wall material and distance to channel as the two most important damage influencing features was not entirely unexpected, given that i) buildings with sandcrete block have relatively consistent stability after contact with water compared to with clay material (especially unburnt clay), and ii) buildings constructed closer to the channel have higher water depths on average and are less sheltered by other buildings.

as shown by a high correlation between the two variables (Fig. 4). Building condition and quality were also ranked to be relatively important variables. Both wall plaster material and extent were identified to have low variable importance. Given their importance in delaying the intrusion of water into the primary wall material, it remains unclear why they do not significantly contribute to explaining the variance in building damage. The low variability in factor levels of the wall material and extent (Figure S2-1) is a probable suspect to their performance. Functional drainage was identified to be the least important variable. The poor performance of functional drainage was unexpected given its identification as important damage influencing variables in previous studies (e.g. Refs. [13,14,66]). However, the distribution of the data for functional drainage (see Figure S2-1D) shows that most observed buildings (about 80%) do not have drainages or the drainages were blocked (both classified as 'no functional drainage'), consequently limiting variability in the variable and hence its performance in the model.

Concerning predicted class probabilities, the results of the repeated nested resampling approach present a consistent picture across all ten models (Fig. 9). Density plots of predicted probabilities show a favorable distribution for correct class 5 + 6 predictions, with some instances being predicted with a probability of up to almost 90%. Densities of class 5 + 6 predictions for all other classes show modes of around 0.1. Predicted probabilities for observations predicted as class 3 and class 4 are more ambiguous. While densities for class 4 predictions indicate comparatively high predicted probabilities not only for true class 4 instances but across all other classes as well, class 3 exhibits narrower densities with lower-class probabilities. True class 1 + 2 predictions exhibit a very similar pattern to predicted class 4 probabilities for class 1 + 2 instances, thereby reflecting the percentage of 38% true class 1 + 2 predictions and 37% of class 1 + 2 erroneously predicted as class 4.

4.4. Model comparison

Interesting implications can be drawn from a comparison of the expert-based and data-driven method. Firstly, the variable importance between the two methods reveals both similar and contrasting deductions. In both methods, distance to channel was consistently identified as significant damage influencing variable. The consistency comes from its (i) repeated performance across the ten models of the random forest, and (ii) high mean weights from multiple experts compared to other variables. Wall material, although identified as the most important in the data-driven method, was moderately important in the expert method. Both methods rank building condition as more important than building quality, hence suggesting that a building with a low quality of construction is likely to perform well when it is properly maintained. Functional drainage, ranked second most important variable in the expert-based method, is the least important variable in the data-driven approach. Given the limited variability of functional drainage in our data (Figure S2-1D), this variable may require further evaluation. Furthermore, while the height of opening was identified to be relatively important in the data-driven method, it was averagely weighted as only 'slightly important' by experts. Consequently, it was removed from the analysis based on the selected threshold (Fig. 5A).

Model performance was measured using the percentage of correct predictions in each damage class. Multiclass prediction accuracy for the data-driven method (38%) was higher than that of the expert-based approach (30%) by 8%. Better performance of the data-driven method was not unexpected given that the implemented approach (random forest) is a supervised learning method that uses the observed damage for model training. On the other hand, the expert-based method was 'unsupervised' given that no data on observed damage grade classes were used: it relies on a prior classification of buildings into resistance (vulnerability) classes based on building and exposure characteristics, and a what-if analysis based on expert knowledge. The data-driven method performed well especially for higher damage grades (classes 4, 5 + 6) with predictions at over 50% accuracy (Fig. 8). The accuracy of the data-driven method dropped for low damage grades, especially for damage class 3 with the least accuracy at 17%. Apparently, class 3 and class 4 are difficult to distinguish based on the available data, implying that the collected variables do not have enough explanatory power to better separate between these two classes. The low accuracy for class 3 is mainly because observed damage class 3 instances are mostly overestimated and predicted as damage class 4.

Conversely, the expert-based method performed well at low damage grades (class 1 + 2, 48% and class 3, 38%) compared to high damage grades (class 4, 17% and class 5 + 6, 18%) (Fig. 7). However, this can be attributed to the fact that the expert-based method is biased towards underestimating the damage class in general. Since low damage classes are predicted much more frequently, the number of correct predictions seems to be high. At the same time, the number of buildings incorrectly predicted to belong to low damage classes is high as well. These results suggest that the data-driven method outperforms the expert-based method, especially for predicting damage classes 4 and 5 + 6. Results for low damage classes have to be interpreted carefully. While the data-driven method generally overestimates low damage classes, the expert-based method clearly underestimates high damage classes. The highest misclassifications in the data-driven approach were related to 47% of class 3 being predicted as class 4 (data-driven method). As highlighted earlier, the small difference between moderate cracks (class 3) and heavy crack (class 4), might result in difficulty for the model to distinguish the two classes. In the expert method, 50% of class 4 and 41% of class 3 were predicted as class 1 + 2, further underlining how experts underestimated damage for low water depths. In both methods, while class 1 + 2 has the highest cumulative prediction accuracy considering both methods (38% in the data-driven and 48% in the expert method) class 3 had the worst performance (13% in the data-driven and 40% in the expert method). In general, given limited research on typical sandcrete, clay, and mixed buildings used in this study, and the unsupervised approach adopted by the expert-based method, a 30% prediction accuracy is considered satisfactory and provides provisional alternative for typical data-scarce areas.

4.5. *Model transferability*

To ensure that the spatial extent of model applicability is not entirely limited to the selected study regions, some steps were taken regarding data used and methods applied. For example, the use of a merged data set (section 3.3) reduces model overfitting to a single event. In the expert-based method, experts were chosen from different locations within Nigeria (section 3.3.1), so that variable (indicator) selection, weighting, and what-if analysis were reflective of different regional damage drivers, physical geomorphology, and hazard characteristics. On the other hand, the data-driven model implemented a nested resampling technique (sec 3.3.2) to allow for unbiased performance estimates on the combined data. Model results for training a random forest model using data from study region 1 and testing on study region 2 show the same multiclass accuracy of 0.38 similar to the model developed from the combined data set. Heatmap and mosaic plot (Figure S2-5) generated from the confusion matrix of true and predicted damage classes show the highest prediction accuracy for class 5 + 6 at 70%. All the other classes show low accuracies: class 1 + 2 (29%), class 3 (24%), and class 4 (32%).

Given the highlighted steps and the multiclass prediction accuracy of the model (0.38) on a different flood event and spatial location, the transferability of both models is highly plausible particularly in regions with similar building characteristics. In general, we further recommend the application of both methods in regions with comparable building characteristics so that model transferability can be further evaluated. Alternatively, spatial resampling such as spatial cross-validation can be used to alleviate the problem of spatial autocorrelation, which might lead to overoptimistic results. Spatial autocorrelation structures stem from the fact that geographic data are often not statistically independent, since observations tend to be more similar the closer they are to each other [67]. When employing spatial partitioning, model performance results are evaluated on spatially disjointed subsets and do thus exhibit a lower bias [68].

4.6. *Sensitivity analysis*

The results of the sensitivity analysis are shown in Figure S2-6. Generally, percentage correct predictions for varying threshold values were marginally comparable across each data set. The highest variation in prediction accuracy within the same data was within a 5% difference as observed in data set 2 between a threshold value of 5 and 6. Observed low sensitivity of varying threshold values might be related to two reasons. Firstly, variables with low mean weights do not significantly contribute to the BRI, hence their removal results in only slight (or no) changes on the BRI classification. Secondly, the quartile classification of the BRI was relatively conservative since it maintains the same number of observations between each quartile. Also, for the quartile classification, the classification of BRI moderate is relatively large (interquartile range). As a result, it limits the ease with which small changes in BRI will alter a buildings' class, except for buildings at class boundaries. Generally, the low sensitivity of different threshold values suggests that even with a low number of variables (e.g. only 5 variables at a threshold value of 6), relatively comparable model performance can be achieved. Hence efforts required in undertaking field data collection can be drastically reduced and methods for rapid building vulnerability assessment can be further enhanced. Figure S2-6 indicates that data from region 1 maintained a performance accuracy above 31% but region 2 showed a consistent low performance with an average of around 20%. The reason for the low performance in region 2 is most likely attributable to a higher proportion of damage grade 4 at lower water depths (51–100 cm) (Fig. 3). Field observations showed that a high percentage of the buildings (in region 2) experienced either (i) a ground floor settlement or (ii) disturbance to the compacted soil material directly below the ground floor - in both cases, resulting in damage to the ground floor cover material (class 4). The frequency in damage class 4 for region 2 may be related to the soil properties and will need further investigation. The general low water depths observed in study region 2 contributed to the low performance since the MPD curve underestimates damage for low water depths. The standard deviation for the percentage correct prediction between the BRI classes shows that study region 2 had the highest variation of percentage accuracy between BRI classes compared to study region 1 or the combined data. The high variation in study region 2 shows that poor prediction accuracy is related to one or two BRI classes showing significant low performance.

4.7. Model uncertainties and outlook

Several uncertainties regarding data collection and analysis exist. In this section, we discuss these uncertainties and provide recommendations for future studies.

A basic input for the study is the field data collected by interviews. As a result, model output relies on the accuracy of such data, which in turn partly relies on personal recollections by the building residents during field interviews. Such personal reflections present some uncertainties especially if the field surveys were conducted long after the flood event. In study region 1, field data collection was carried out eight months after the flood. While data such as susceptibility, local protection, and exposure could be directly observed on the field and satellite imagery, water depths are not directly deductible. Yet, a large percentage of the water depths used in the study were based on personal reflections by building residents and not on measured watermarks. In study region 2, field surveys were conducted three months after the flood event, hence, residents have a higher tendency to remember flood depths with good accuracy. In general, floods are traumatic events [69–71] and people affected by them do not easily forget details about the events. However, where possible, data collections should be carried out as soon as possible after flood occurrence to generally ensure higher accuracy in water depths.

Bivariate analysis showed that the distribution of some variables is skewed and the data-driven model may favor factor levels that are over-represented. For example, variables such as functional drainage, natural barrier, wall thickness are characterized by imbalanced factor levels, with one dominating manifestation of the variable. However, since the random forest algorithm is a non-parametric classifier, there are no prerequisites with respect to the distribution assumptions of input data. The repeated nested resampling strategy was applied to obtain a robust model with honest performance estimates. Models fitted using a simple single train-test split might still be prone to suffer from a slight bias caused by overfitting. For some variables, the disproportionate distribution of the data (e.g. sandcrete block in wall material) is representative of the actual situation (see data from Ref. [49]).

Additional uncertainties relating to non-inclusion of other variables that may influence building damage exists (i.e. unobserved heterogeneity). For example, two resistance variables, building age and number of floors, initially selected by experts had to be removed. The building age was highly incomplete because residents could either not remember the year of construction or simply have no knowledge about it since they were not the first to reside in the house. The number of stories was removed due to a lack of variability: out of the entire data comprising 444 buildings, only one building was two-story while all the others were one story (only ground floor). Other hazard variables that could have been interesting for the study were flood duration and velocity. Many residents qualitatively described how the flow approached at 'high' speed. Other residents, especially those residing in clay buildings, suggested that their houses were damaged due to longer durations of exposure to water. Given high uncertainties in translating qualitative and quantitative reports on flood velocities and duration, both variables were excluded from the data. For example, many residents have stated that during flooding, it is very difficult to keep track of time. Hence, in some cases reported durations for two buildings next to each were more than 6 h apart. A method for hydrodynamic modeling for data-scarce regions without hydrological data has been proposed in Ref. [72]. Such methods provide a pathway for flood data extrapolation (flood duration and velocity) to complement current efforts for flood damage modeling in data-scarce regions.

Predicting damage grades presents a rather challenging task which is also evident from the mediocre performance observed in the random forest model. Such difficulties are even more pronounced in regions where policies on building standards are less well implemented, which results in substantial variation in building quality and value [7,16,34]. The variations in building quality contribute to uncertainty in damage prediction since buildings within the same BRI category (poor, moderate, good) may not incur comparable damage even when impacted by the same flood depth. The LPD and HPD curves (Figure A1) are developed to reduce uncertainties inherent in each BRI class. We recommend a further sub-classification within each BRI class such that MPD is only used for buildings that are more comparable to the representative building. LPD and HPD are then applied to buildings with characteristics (or calculated BRI value) deviating from the representative building. Where high variations in building standards do not exist, the recommended additional sub-setting in each category will not be necessary. In general, we recommend further performance assessment of the MPD, LPD, and HPD curves in different data-scarce regions with comparable building characteristics.

Expert interviews are generally subjective, hence present some uncertainties. As a result, a high number of experts are required so that results are representative. In practice, getting a high number of experts is not always feasible, especially in regions where the required expertise is limited. In Nigeria, the challenge regarding the low number of experts in flood damage and vulnerability assessment has been previously pointed out by Komolafe et al., [9]. Our study included seven experts for indicator selection, weighting, and what-if analysis. Selected experts were chosen from different geographical regions and fields of study, which generally influences how they carry out the assessment. For example, in the what-if analysis, while experts with an engineering background included 'ground floor settlement' in their assessment, it was partly challenging for few others (with other backgrounds) to relate the variable with a water depth range. We generally recommend that for future studies, the what-if analysis should be conducted within the framework of a workshop during which all relevant information regarding damage states, water depth classification, and representative buildings from BRI classes are properly discussed. Such workshops will reduce consequent uncertainties that arise from knowledge gaps based on expert background. As highlighted above, the general poor performance of expert-based assessment for higher damage may be related to using water depth ranges up to 500 cm. It would be interesting for future studies to re-evaluate the outcome of such what-if analysis using a maximum water depth of 350 cm to evaluate the influence of maximum water depths on damage grade estimates.

Furthermore, future studies should consider upscaling the approach from a micro-scale to a regional scale to support disaster management at city level. The availability of free regional scale building data-sets such as ImageCAT [73] which provides georeferenced data on building characteristics for many African countries provides a good potential for such regional models. In addition, another important recommendation for future studies is to link the developed damage grades to repair cost so that model application can be extended to evaluating monetary loss.

Both the expert-based and data-driven method have applications for disaster management. For example, calculated values of the BRI is useful for identifying buildings that are highly vulnerable to floods. This information on vulnerable buildings is important for recommending mitigation or local protection measures to house owners during disaster preparedness phase. In addition, both the expert-based and data-driven models can be used to identify buildings with a high probability to incur severe damage during a flood event. Such information is important for emergency planners for rescue operations during disaster response. More so, the developed damage grades are simplistic and provide an easy communication tool to both decision makers and community residents for creating awareness on flood disaster especially for community residents located in high risk areas.

5. Conclusion

The development of flood damage models represents an important step towards flood risk disaster reduction given its multiple applications in mitigation and emergency planning, economic loss evaluation, the cost-benefit analysis for flood protection measures. However, the unavailability of well documented empirical data has so far limited the application of flood damage models in several data-scarce regions. In addition, current methods have either been limited to exposure assessment, identification of vulnerability indicators, or not well representative of regional building types.

In this study, we carried out a comparative assessment of a data-driven with an expert-based approach that does not require empirical data. The comparative assessment aimed at evaluating the prediction accuracy of the expert-based approach as well as gain understanding into regional damage drivers for typical building types. Flood damage data, collected from two different regions in Nigeria, was used to evaluate model performance. Data from the two study regions were used either combined to reduce model overfitting to a single event or separate to enable an evaluation of model transferability.

Several conclusions can be drawn from the study:

i. Generally, experts underestimated the damage to lower water depths. The MPD curves predict high damage (class 4, 5 + 6) only for high water depths (above 100 cm). However, the bivariate analysis showed that high damage can occur at lower flood depths, for example, some buildings categorized under BRI low incur a damage grade 6 from a water depth of 50 cm. The underestimation of damage for low water depths consequently resulted in multiclass prediction accuracy of 30% by the expert-based model. However, given i) limited research on sandcrete and clay building types, ii) observed variation in building standards, a 30% performance accuracy is considered satisfactory and provides a provisional alternative for typical data-scarce regions. The bivariate analysis supports other studies that demonstrate high uncertainty in predicting building damage using only water depths.

ii. Damage grade prediction presents several challenges especially in regions with relatively high variation in building standards. Consequently, the data-driven method had an average performance accuracy of 38%.

ii. Both methods suggest that even at a reduced number of variables, comparable model performance can be achieved. Hence, efforts and time spent on field data collection can be reduced or better targeted to the most important variables. Similar conclusions were deduced in a recent study by Papathoma-Köhle et al. [40].

iv. Buildings within the BRI classes showed considerable differences for similar water depth range. This suggests that achieving better performance with the expert-based method will require i) a re-evaluation of the variables weights or classification scheme used or ii) incorporating additional variables that were not considered in the study.

v. Both the expert-based and data-driven methods suggest that distance to channel, wall material, building condition, and building quality are important variables to be considered for physical vulnerability assessment in typical regions.

vi. Experts assessments, in particular the what-if analysis, might require a formal discussion (e.g. workshop) to bridge knowledge gaps that arise especially when experts are from different fields of study.

vii. The combination of different physical vulnerability assessment methods shows good potential for adapting flood damage models to regional situations typical for data-scarce areas. The inclusion of local experts allows the model to be tailored specifically to regional situations.

Financial support

This research has been supported by the Swiss Government Excellence Scholarship (grant no. 2017.1027).

Data availability

The data used in this study can be acquired based on request to the corresponding author.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This study is carried out within the framework of a PhD scholarship funded by the Swiss Government Excellence Scholarships for Foreign Scholars (ESKAS). The authors would like to especially thank all the experts that volunteered to participate in the expert evaluation: Dr. A. A. Komolafe, Dr. M. K. Kawu, Dr. O. S. Rafiu, Dr. A. Richard, Dr. O. Samsideen, Prof. O. D. Jimoh, Prof. I. O. Adelekan, Dr. O. P. Uchenna, Dr. T. Akukwe, Dr. A. B. Ismail. In addition, we would like to thank Bala Inuwa and Ambrose Patrick for guidance and assistance during data collection in study regions 1 and 2 respectively.

Appendix 1

Variable definition and measurement

Flood depth

The height of floodwater, measured from the ground level (not floor level) at a building location. These were measured directly if flood marks were still visible or residents were asked based on personal recollections.

Construction material

The material used for erecting the walls of the building. The construction material was further classified into wall material and thickness.

- Wall material: Three building materials are considered i) sandcrete block, ii) clay (burnt and unburnt), and iii) Mixed (a combination of sandcrete block and clay). Usually, for the mixed class, sandcrete block is used to a height of about 100 cm and the rest are completed using clay. The mixed system is mainly used in regions that are exposed to floods providing a relative balance between safety and construction cost.
- Wall thickness: This refers to the size (width) of the wall unit. Common widths for wall thickness found in the study regions are 15 cm and 24 cm.

Building condition

Assesses the maintenance status or state of the individual components of a building. In our study region, we carried out this assessment using scoring for different building components: i) 4 - very good (as new), ii) 3 - good (light deterioration), iii) 2 - moderate (average or increased deterioration), and iv) 1 - poor (severe deterioration) [74,75]. Building components included in this assessment are: i) walls, ii) doors and windows, iii) finishes or plastering, iv) flooring (part of the building the room stands on), and v) roofing. The building condition is computed as an average of the assigned score overall building components as shown in equation A1. The range of the building quality lies between 1 and 4.

$$Building\ Quality = \frac{1}{5}\left(external\ wall_{score} + doors\ and\ windows_{score} + flooring + finishes_{score} + roofing_{score}\right) \tag{A1}$$

Building quality

Relates to the conformity of the building to country standards. In order to be consistent with established standards outlined in the Nigerian building code [65], building standard is assessed based on three considerations: aesthetics, durability, and functionality. A scoring system is used to qualify conformity to standards using i) 3 - good (high), ii) 2 - moderate (above average), and iii) 1 - poor (below average) conformity. Four building components used for evaluating building standard includes; external walls, doors and windows, roofing, and finishes or plastering. The building quality for each component is computed as an average of the score for aesthetic, durability, and functionality. An example for external wall is given in equation A1. The building quality score is thereafter computed as the average of the score for all four building components as shown in equation (A3). The range of the building quality lies between 1 and 3.

$$External\ wall_{score} = \frac{1}{3}\left(aesthetic_{score} + durability_{score} + functionality_{score}\right) \tag{A2}$$

$$Building\ Quality = \frac{1}{4}\left(external\ walls_{score} + doors\ and\ windows_{score} + finishes_{score} + roofing_{score}\right) \tag{A3}$$

Height of opening

The height of opening is the distance measured from the ground level to the lowest part of the window. If the window heights are different, the lowest window is used. The lower the height of opening, the faster flood water can gain entrance into a building.

Building footprint

Taken as the external form (or outline) of a building. In our study regions, predominant building footprints have been categorized into i) L-shape, ii) Rectangular, iii) irregular.

Ground floor elevation

The elevation of the building ground level (floor) or entrance partly influences the amount and time flood water can gain access inside a building. In certain communities where floods are relatively common, residents either raise the height of the building floor level or erect a small barrier at the doors as a local protection measure.

Wall finishes/plaster

Materials used as cover to the main wall unit are commonly referred to as wall finishes. The use of wall finishes varies from protection and aesthetics. While the recommended practice is the use of plaster on the entire building [65], this is not necessarily the case in many buildings. To accommodate these differences, we further categorize wall finishes/plaster into i) plaster material, and ii) plaster extent.

- Plaster material: The material used for plaster categorized into i) cement and sand, ii) clay, and iii) none (no plaster).
- Plaster extent: The proportion of wall that is plastered categorized into i) complete, ii) partial, and iii) none (no plaster).

Fencing

Refers to a walling unit erected around a building. In Nigeria, the practice of building a fence is relatively common. Although fences can primarily be for controlling access, they reduce the direct impact of the flood on buildings. We further classify fencing into i) fencing material, and ii) fencing extent.

- Fencing material: The material used for the fencing influences how much pressure it can withstand from the flood. The fencing material is classified into i) sandcrete blocks, ii) clay, iii) mixture of sandcrete blocks and clay, iv) others (Zinc metal and thatched), and iv) none (no fencing erected)
- Fencing extent: Fencing can be erected around the entire building or allowed to cover only part of a building. Three classes of fencing extent used in this study include i) complete (fencing covers the entire building), ii) partial (fencing covers only a part of the building), and iii) none (no fencing erected)

Distance to river channel

The distance between a building and the river channel is expressed using the variable distance to river channel. Distance to channel is usually limited to the flood extent since only buildings affected by the flood are included in this assessment.

Functional drainage

Drainages function to provide channelization of excess runoff. Where such drainages are not available or are non-functional (e.g., blocked by debris), excess runoff can quickly result in floods. In Nigeria, some studies (see Refs. [13,14,66]) have shown that the availability (or functionality) of drainages have remarkably contributed to flood occurrence and subsequent building damage. Here, we categorize functional drainage into two i) yes (available), and ii) no (not available or non-functional).

Natural barrier(s)

The presence of vegetation (grasses, trees) around a building is expected to influence the velocity of flood water as it approaches the building. Here, each building is classified into to i) yes (existence of a natural barrier), and ii) no (no natural barrier) around the building.

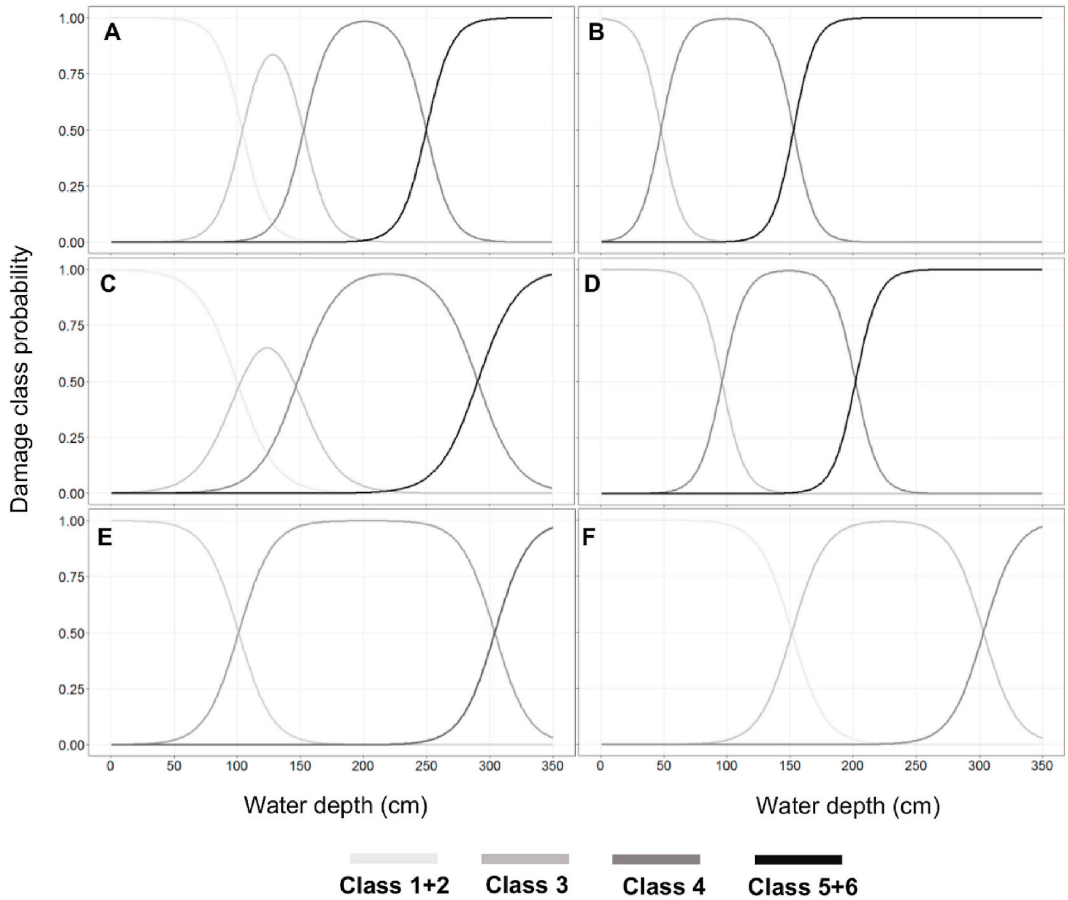


Fig. A1. Lower and higher probable damage curves for (A, B) poor, (C, D) moderate, and (E, F) good BRI classes respectively.

Sheltering

The availability of a structure in between a building and a flood source or preferable water path (e.g., river channel or roads) influences the impacting force [76]. Sheltering refers to the relative protection of one building by another. Such that a direct impact is reduced or avoided. In our study, similar to Maiwald and Schwarz [76], two considerations are used to classify a buildings' sheltering status: spatial location of a building relative to the river channel and other buildings, and direction of river flow. Buildings are classified into i) direct (complete), ii) partial (moderate), and iii) no (none) sheltering.

Equations for maximum-minimum normalization

$$BRI_{min} = \frac{E_{weight}}{n}\sum_{i=1}^{n}(e_i y_{min}) + \frac{S_{weight}}{p}\sum_{j=1}^{p}(s_j y_{min}) + \frac{L_{weight}}{r}\sum_{k=1}^{r}(l_k y_{min}) \tag{A4}$$

$$BRI_{max} = \frac{E_{weight}}{n}\sum_{i=1}^{n}(e_i y_{max}) + \frac{S_{weight}}{p}\sum_{j=1}^{p}(s_j y_{max}) + \frac{L_{weight}}{r}\sum_{k=1}^{r}(e_k y_{max}) \tag{A5}$$

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijdrr.2021.102148.

References

[1] O. Hoegh-Guldberg, D. Jacob, M. Taylor, M. Bindi, S. Brown, I. Camilloni, A. Diedhiou, R. Djalante, K.L. Ebi, F. Engelbrecht, J. Guiot, Y. Hijioka, S. Mehrotra, A. Payne, S.I. Seneviratne, A. Thomas, R. Warren, G. Zhou, Special report on global Warming of 1.5 °C - chapter 3: impacts of 1.5° C global warming on natural and human systems, glob. Warm. 1.5°C. An IPCC spec. Rep. Impacts glob. Warm. 1.5°C above pre-industrial levels relat. Glob. Greenh. Gas emiss. Pathways, context strength, Glob. Response Threat Clim. Chang. (2018) 175–311, https://doi.org/10.1002/ejoc.201200111.

[2] S. Fuchs, M. Keiler, A. Zischg, A spatiotemporal multi-hazard exposure assessment based on property data, Nat. Hazards Earth Syst. Sci. 15 (2015) 2127–2142, https://doi.org/10.5194/nhess-15-2127-2015.
[3] V. Röthlisberger, A.P. Zischg, M. Keiler, Identifying spatial clusters of flood exposure to support decision making in risk management, Sci. Total Environ. 598 (2017) 593–603, https://doi.org/10.1016/j.scitotenv.2017.03.216.
[4] (Centre for Research on the Epidemiology of Disasters) CRED, Disaster Year in Review 2019, 2020.
[5] (United Nations International Strategy for Disaster Reduction) UNISDR, Terminology on Disaster Risk Reduction, United Nations International Strategy for Disaster Reduction Geneva, Geneva, Switzerland, 2009. https://www.undrr.org/publication/2009-unisdr-terminology-disaster-risk-reduction. accessed March 15, 2020.
[6] (Centre for Research on the Epidemiology of Disasters) CRED, Disasters in Africa: 20 Year Review (2000-2019, 2019.
[7] M.B. Malgwi, S. Fuchs, M. Keiler, A generic physical vulnerability model for floods: review and concept for data-scarce regions, Nat. Hazards Earth Syst. Sci. 20 (2020) 2067–2090, https://doi.org/10.5194/nhess-20-2067-2020.
[8] P. Gasparini, Analysis and Monitoring of Environmental Risk: CLUVA Final Report, Clim. Chang. Urban Vulnerability Africa, 2013, pp. 1–26. http://cordis.europa.eu/docs/results/265137/final1-cluva-final-publishable-summary-report.pdf. (Accessed 1 January 2019).
[9] A.A. Komolafe, S.A.A. Adegboyega, F.O. Akinluyi, A review of flood risk analysis in Nigeria, Am. J. Environ. Sci. 11 (2015) 157–166, https://doi.org/10.3844/ajessp.2015.157.166.
[10] P.C. Nwilo, D.N. Olayinka, A.E. Adzandeh, Flood modelling and vulnerability assessment of settlements in the Adamawa state floodplain using GIS and cellular framework approach, Global J. Hum. Soc. Sci. 12 (2012) 11–20.
[11] M. Zumo, Delineation of built-up areas liable to flood in Yola, Adamawa State, Nigeria using remote sensing and geographic information system technologies, FUTY J. Environ. 8 (2014) 20–30.
[12] A.B. Ndanusa, Z. Dahalin, A. Ta, Topographic-based framework for flood vulnerability classification: a case of Niger state, Nigeria, J. Inf. Syst. Technol. Manag. 3 (2018) 27–38.
[13] B. Adeleye, A. Popoola, L. Sanni, N. Zitta, O. Ayangbile, Poor development control as flood vulnerability factor in Suleja , Nigeria, T. Reg. Plan (2019) 23–35, https://doi.org/10.18820/2415-0495/trp74i1.3.
[14] P.U. Okoye, F.O. Ezeokoli, J. Ugochukwu, Building development practice in flood prone area: case of ogbaru council area of anambra state Nigeria, Int. J. Eng. Res. Afr. 5 (2015) 30–40.
[15] T.I. Akukwe, C. Ogbodo, Spatial Analysis of Vulnerability to Flooding in Port Harcourt Metropolis, Nigeria, vol. 5, SAGE Open, 2015, https://doi.org/10.1177/2158244015575558.
[16] (Federal Government of Nigeria) FGN, Nigeria, Post-disaster Needs Assessment - 2012 Floods, 2013. https://www.gfdrr.org/sites/gfdrr/files/NIGERIA_PDNA_PRINT_05_29_2013_WEB.pdf. (Accessed 1 January 2019).
[17] M. Papathoma-Köhle, B. Gems, M. Sturm, S. Fuchs, Matrices, curves and indicators: a review of approaches to assess physical vulnerability to debris flows, Earth Sci. Rev. 171 (2017) 272–288, https://doi.org/10.1016/j.earscirev.2017.06.007.
[18] M. Papathoma-Köhle, M. Kappes, M. Keiler, T. Glade, Physical vulnerability assessment for alpine hazards: state of the art and future needs, Nat. Hazards 58 (2011) 645–680, https://doi.org/10.1007/s11069-010-9632-4.
[19] A. Blanco-Vogt, J. Schanze, Assessment of the physical flood susceptibility of buildings on a large scale - conceptual and methodological frameworks, Nat. Hazards Earth Syst. Sci. 14 (2014) 2105–2117, https://doi.org/10.5194/nhess-14-2105-2014.
[20] S. Fuchs, Susceptibility versus resilience to mountain hazards in Austria-paradigms of vulnerability revisited, Nat. Hazards Earth Syst. Sci. 9 (2009) 337–352, https://doi.org/10.5194/nhess-9-337-2009.
[21] (United Nations International Strategy for Disaster Reduction) UNISDR, Sendai framework for disaster risk reduction 2015–2030, in: Proc. 3rd United Nations World Conf. DRR, Sendai, Japan, 2015, p. 32. https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030. (Accessed 26 February 2020).
[22] S. Fuchs, M. Keiler, R. Ortlepp, R. Schinke, M. Papathoma-Köhle, Recent advances in vulnerability assessment for the built environment exposed to torrential hazards: challenges and the way forward, J. Hydrol. 575 (2019) 587–595, https://doi.org/10.1016/j.jhydrol.2019.05.067.
[23] S. Fuchs, M. Heiser, M. Schlögl, A. Zischg, M. Papathoma-Köhle, M. Keiler, Short communication: a model to predict flood loss in mountain areas, Environ. Model. Software 117 (2019) 176–180, https://doi.org/10.1016/j.envsoft.2019.03.026.
[24] K. Vogel, C. Riggelsen, B. Merz, H. Kreibich, F. Scherbaum, Flood damage and influencing factors: a bayesian network perspective, Proc. 6th Eur. Work. Probabilistic Graph. Model. (PGM 2012 625 (2012) 314–354.
[25] B. Merz, H. Kreibich, U. Lall, Multi-variate flood damage assessment: a tree-based data-mining approach, Nat. Hazards Earth Syst. Sci. 13 (2013) 53–64, https://doi.org/10.5194/nhess-13-53-2013.
[26] S. Ettinger, L. Mounaud, C. Magill, A.F. Yao-Lafourcade, J.C. Thouret, V. Manville, C. Negulescu, G. Zuccaro, D. De Gregorio, S. Nardone, J.A.L. Uchuchoque, A. Arguedas, L. Macedo, N. Manrique Llerena, Building vulnerability to hydro-geomorphic hazards: estimating damage probability from qualitative vulnerability assessment using logistic regression, J. Hydrol. 541 (2016) 563–581, https://doi.org/10.1016/j.jhydrol.2015.04.017.
[27] K. Schröter, H. Kreibich, K. Vogel, C. Riggelsen, F. Scherbaum, B. Merz, How useful are complex flood damage models? Water Resour. Res. 50 (2014) 3378–3395, https://doi.org/10.1002/2013WR014396.
[28] B. Merz, H. Kreibich, A. Thieken, R. Schmidtke, Estimation uncertainty of direct monetary flood damage to buildings, Nat. Hazards Earth Syst. Sci. 4 (2004) 153–163, https://doi.org/10.5194/nhess-4-153-2004.
[29] B. Merz, H. Kreibich, R. Schwarze, A. Thieken, Review article "assessment of economic flood damage, Nat. Hazards Earth Syst. Sci. 10 (2010) 1697–1724, https://doi.org/10.5194/nhess-10-1697-2010.
[30] B. Jongman, H. Kreibich, H. Apel, J.I. Barredo, P.D. Bates, L. Feyen, A. Gericke, J. Neal, J.C.J.H. Aerts, P.J. Ward, Comparative flood damage model assessment: towards a European approach, Nat. Hazards Earth Syst. Sci. 12 (2012) 3733–3752, https://doi.org/10.5194/nhess-12-3733-2012.
[31] M.J. Hammond, A.S. Chen, Urban flood impact assessment : a state-of-the-art review, Urban Water J. 12 (2015) 14–29, https://doi.org/10.1080/1573062X.2013.857421.
[32] T. Gerl, H. Kreibich, G. Franco, D. Marechal, K. Schröter, A review of flood loss models as basis for harmonization and benchmarking, PloS One 11 (2016) 1–22, https://doi.org/10.1371/journal.pone.0159791.
[33] I. Niang, O.C. Ruppel, M.A. Abdrabo, A. Essel, C. Lennard, J. Padgham, P. Urquhart, Africa, in: V.R. Barros, C.B. Field, D.J. Dokken, M.D. Mastrandrea, K. J. Mach (Eds.), Clim. Chang. 2014 Impacts, Adapt. Vulnerability Part B Reg. Asp. Work. Gr. II Contrib. To Fifth Assess. Rep. Intergov. Panel Clim. Chang., Cambridge University Press, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2015, pp. 1199–1266, https://doi.org/10.1017/CBO9781107415386.002.
[34] J. Englhardt, H. de Moel, C.K. Huyck, M.C. de Ruiter, J.C.J.H. Aerts, P.J. Ward, Enhancement of large-scale flood risk assessments using building-material-based vulnerability curves for an object-based approach in urban and rural areas, Nat. Hazards Earth Syst. Sci. 19 (2019) 1703–1722, https://doi.org/10.5194/nhess-19-1703-2019.
[35] E. Penning-Rowsell, S. Tunstall, S. Tapsell, J. Morris, J. Chatterton, G. Colin, The Benefits of Flood and Coastal Risk Management: a Manual of Assessment Techniques, 2005, https://doi.org/10.1596/978-0-8213-8050-5. ISBN 19047.
[36] M. Neubert, T. Naumann, C. Deilmann, Synthetic water level building damage relationships for GIS-supported flood vulnerability modeling of residential properties, in: Flood Risk Manag. Res. Pract. Proc. Eur. Conf. Flood Risk Manag. Res. Into Pract. FLOODrisk, 2008, p. 294.
[37] T. Naumann, J. Nikolowski, G. Sebastian, Synthetic depth-damage functions – a detailed tool for analysing flood resilience of building types, Road Map Towar. a Flood Resilient Urban, Environ. Proc. Final Conf. COST Action C. (2009), https://doi.org/10.1093/CERCOR/8.4.321.
[38] A. Pistrika, G. Tsakiris, I. Nalbantis, Flood depth-damage functions for built environment, Environ. Process. 1 (2014) 553–572, https://doi.org/10.1007/s40710-014-0038-2.
[39] H. Maiwald, J. Schwarz, Damage and loss prognosis tools correlating flood action and building's resistance-type parameters, Int. J. Saf. Secur. Eng. 5 (2015) 222–250, https://doi.org/10.2495/SAFE-V5-N3-222-250.
[40] M. Papathoma-köhle, M. Schlögl, S. Fuchs, Vulnerability indicators for natural hazards: an innovative selection and weighting approach, Sci. Rep. 9 (2019) 1–14, https://doi.org/10.1038/s41598-019-50257-2.
[41] J. Schwarz, H. Maiwald, Damage and loss prediction model based on the vulnerability of building types, in: 4th Int. Symp. Flood Def. Manag. Flood Risk, Reliab. Vulnerability, vol. 9, 2008, https://doi.org/10.13140/2.1.1358.3043.
[42] H. Maiwald, J. Schwarz, Unified damage description and risk assessment of buildings under extreme natural hazards, Eur. J. Mason. – Mauerw. 23 (2019) 95–111, https://doi.org/10.1002/dama.201910014.
[43] F. Dall'Osso, M. Gonella, G. Gabbianelli, G. Withycombe, D. Dominey-Howes, A revised (PTVA) model for assessing the vulnerability of buildings to tsunami damage, Nat. Hazards Earth Syst. Sci. 9 (2009) 1557–1565.
[44] R. Blong, A review of damage intensity scales, Nat. Hazards 29 (2003) 57–76, https://doi.org/10.1023/A:1022960414329.
[45] D. Wagenaar, J. De Jong, L.M. Bouwer, Multi-variable flood damage modelling with limited data using supervised learning approaches, Nat. Hazards Earth Syst. Sci. 17 (2017) 1683–1696, https://doi.org/10.5194/nhess-17-1683-2017.
[46] A. Ayanlade, Seasonal rainfall variability in Guinea Savanna part of Nigeria: a GIS approach, Int. J. Clim. Chang. Strateg. Manag. 1 (2009) 282–296, https://doi.org/10.1108/17568690910977492.
[47] O.M. Eludoyin, I.O. Adelekan, The physiologic climate of Nigeria, Int. J. Biometeorol. 57 (2013) 241–264, https://doi.org/10.1007/s00484-012-0549-3.
[48] H.E.R.E. Esri, Garmin, OpenStreetMap, GIS-Community, Light gray canvas base map, 2020. https://services.arcgisonline.com/ArcGIS/rest/services/Canvas/World_Light_Gray_Base/MapServer.
[49] (National Bureau of Statistics), NBS, Annual Abstract of Statistics, Federal Republic of Nigeria, 2012 https://doi.org/nigerianstat.gov.ng.
[50] I. Mayomi, M.S. Kolawole, A.K. Martins, Terrain analysis for flood disaster vulnerability assessment: a case study of Niger state, Nigeria, Am. J. Geogr. Inf. Syst. 3 (2014) 122–134, https://doi.org/10.5923/j.ajgis.20140303.02.
[51] Google Earth Pro, Google Earth, (n.d.). earth.google.com/web/.
[52] (United Nations), Office for the coordination of humanitarian affairs) OCHA, Nigeria: floods in borno , Adamawa and yobe, situation report No. 2. https://reliefweb.int/sites/reliefweb.int/files/resources/ocha_nga_floodoverview_13092019.pdf, 2019.
[53] A.H. Thieken, M. Müller, H. Kreibich, B. Merz, Flood damage and influencing factors: new insights from the August 2002 flood in Germany, Water Resour. Res. 41 (2005) 1–16, https://doi.org/10.1029/2005WR004177.
[54] J. Schwarz, H. Maiwald, Prognose der Bauwerksschdigung unter Hochwassereinwirkung, Bautechnik 84 (2007) 450–464, https://doi.org/10.1002/bate.200710039.

[55] H. Kreibich, I. Seifert, B. Merz, A.H. Thieken, Development of FLEMOcs – a new model for the estimation of flood losses in the commercial sector, Hydrol. Sci. J. 55 (2010) 1302–1314, https://doi.org/10.1080/02626667.2010.529815.
[56] M. Holub, S. Fuchs, Benefits of local structural protection to mitigate torrent-related hazards, WIT Trans. Inf. Commun. Technol. 39 (2008) 401–411.
[57] J. Schwarz, H. Maiwald, Berücksichtigung struktureller Schäden unter Hochwassereinwirkung, Bautechnik 84 (2007) 450–464.
[58] T.L. Saaty, The Analytical Hierarchy Process, Planning, Priority, Resour. Alloc. RWS Publ. USA, 1980.
[59] (Joint Research Centre-European Commission and OECD) JRC, Handbook on Constructing Composite Indicators: Methodology and User Guide, Joint Research Centre-European Commission and OECD, 2008, https://doi.org/10.1787/9789264043466-en.
[60] F. Dall'Osso, M. Gonella, G. Gabbianelli, G. Withycombe, D. Dominey-Howes, A revised (PTVA) model for assessing the vulnerability of buildings to tsunami damage, Nat. Hazards Earth Syst. Sci. 9 (2009) 1557–1565.
[61] G. Andrew, S. Yu-Sung, G.P. Maria, Arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R Package Version 1.11-1, 2020. https://cran.r-projkect.org/package/arm.
[62] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[63] P. Probst, M.N. Wright, A.-L. Boulesteix, Hyperparameters and tuning strategies for random forest, WIREs Data Min. Knowl. Discov. 9 (2019), e1301, https://doi.org/10.1002/widm.1301.
[64] B. Bischl, J. Richter, J. Bossek, D. Horn, J. Thomas, M. Lang, mlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions, ArXiv Mach. Learn, 2017.
[65] (Nigerian Building Code) NBC, National Building Code, Federal Republic of Nigeria, 2006.
[66] H.M. Nda Musa, M.Y. Usman, H. Abdul, L.M. Sanni, An assessment of flood vulnerability on physical development along drainage channels in Minna, Niger State, Nigeria, Afr. J. Environ. Sci. Technol. 9 (2015) 38–46, https://doi.org/10.5897/AJEST2014.1815.
[67] H.J. Miller, Tobler's first law and spatial analysis, Ann. Assoc. Am. Geogr. 94 (2004) 284–289, https://doi.org/10.1111/j.1467-8306.2004.09402005.x.
[68] R. Lovelace, J. Nowosad, J. Muenchow, Geocomputation with R, CRC Press, 2019.
[69] P. Verger, M. Rotily, C. Hunault, J. Brenot, E. Baruffol, D. Bard, Assessment of exposure to a flood disaster in a mental-health study, J. Expo. Sci. Environ. Epidemiol. 13 (2003) 436–442, https://doi.org/10.1038/sj.jea.7500290.
[70] V. Mason, H. Andrews, D. Upton, The psychological impact of exposure to floods, Psychol. Health Med. 15 (2010) 61–73, https://doi.org/10.1080/13548500903483478.
[71] A. Fontalba-Navas, M.E. Lucas-Borja, V. Gil-Aguilar, J.P. Arrebola, J.M. Pena-Andreu, J. Perez, Incidence and risk factors for post-traumatic stress disorder in a population affected by a severe flood, Publ. Health 144 (2017) 96–102, https://doi.org/10.1016/j.puhe.2016.12.015.
[72] M.B. Malgwi, J.A. Ramirez, A. Zischg, M. Zimmermann, S. Schürmann, M. Keiler, Flood reconstruction using field interview data and hydrodynamic modelling: a method for data-scarce regions, Nat. Hazards, *Submitted*.
[73] ImageCat, CIESIN, K. Porter, Africa Disaster Risk Financing Phase 1 – Result Area 5, Exposure Development for 5 Sub-saharan African Countries, 2017. Ethiopia, Kenya, Uganda, Niger, Senegal.
[74] T. Manager, Maintenance Management Framework Guidline: Building Condition Assessment, 2017. Brisbane, Queensland, https://www.hpw.qld.gov.au/_data/assets/pdf_file/0019/3277/mmfbca.pdf.
[75] A. Straub, Dutch standard for condition assessment of buildings, Struct. Surv. 27 (2009) 23–35, https://doi.org/10.1108/02630800910941665.
[76] H. Maiwald, J. Schwarz, Damage and loss prediction model considering inundation level, flow velocity and vulnerability of building types, WIT Trans. Ecol. Environ. 159 (2012) 53–65, https://doi.org/10.2495/FRIAR120051.