**Natural Hazards
and Earth
System Sciences**

Validation of a social vulnerability index in context to river-floods in Germany

A. Fekete

United Nations University, Institute for Environment and Human Security (UNU-EHS), UN Campus,
Hermann-Ehlers-Str. 10, 53113 Bonn, Germany

Abstract. Social vulnerability indices are a means for generating information about people potentially affected by disasters that are e.g. triggered by river-floods. The purpose behind such an index is in this study the development and the validation of a social vulnerability map of population characteristics towards river-floods covering all counties in Germany. This map is based on a composite index of three main indicators for social vulnerability in Germany – fragility, socio-economic conditions and region. These indicators have been identified by a factor analysis of selected demographic variables obtained from federal statistical offices. Therefore, these indicators can be updated annually based on a reliable data source. The vulnerability patterns detected by the factor analysis are verified by using an independent second data set. The interpretation of the second data set shows that vulnerability is revealed by a real extreme flood event and demonstrates that the patterns of the presumed vulnerability match the observations of a real event. It comprises a survey of flood-affected households in three federal states. By using logistic regression, it is demonstrated that the theoretically presumed indications of vulnerability are correct and that the indicators are valid. It is shown that indeed certain social groups like the elderly, the financially weak or the urban residents are higher risk groups.

1 Introduction

Developing vulnerability indices at sub-national level is a common approach that is increasingly applied in countries like the United States of America (Clark et al., 1998; Cutter et al., 2000; Wu et al., 2002; Chakraborty et al., 2005; Olfert et al., 2006; Rygel et al., 2006; Kleinosky et al., 2007;

Yarnal, 2007) the United Kingdom (Tapsell et al., 2002), Spain (Weichselgartner, 2002), Latin America (Hahn et al., 2003; Cardona, 2005), Australia (Dwyer et al., 2004) the Philippines (Acosta-Michlik, 2005) or generally for regions worldwide (Nakamura et al., 2001). In Germany there are yet only few attempts that either capture only one federal state (Kropp et al., 2006) or capture social vulnerability by very few variables (Meyer et al., 2007). There is still no comprehensive profile of social vulnerability or sub-national index map for whole river channels in Germany. This gap is to be filled by the Social Vulnerability Index (SVI) in context to river-floods.

The research results are integrated in a multi-disciplinary "Disaster Information System for Large-Scale Flood Events Using Earth Observation" (DISFLOOD, Damm et al., 2006). The project is a platform for multi-disciplinary and multi-institutional research. It is a joint effort of the German Aerospace Centre (DLR), the German Research Centre for Geosciences Potsdam (GFZ) and the United Nations University – Institute for Environment and Human Security (UNU-EHS). The project as financed by the Helmholtz society started in late 2005 and runs until the end of 2008. This platform targets extreme river floods in Germany by assessing hazard and vulnerability parameters. At the same time DSFLOOD is a pilot study to combine different methods like remote sensing, hydraulic hazard models, and economic damage models with social and ecological vulnerability indicators. The outcome is an online information system that will be available on the Natural Disaster Network web site NaDiNe (http://nadine.helmholtz-eos.de/projects/disflood/disflood_de.html). The prime target groups of NaDiNe are experts working on flood protection, regional planners and scientists. Moreover, the interactive hazard and vulnerability maps will also be accessible by the public to a certain degree.

While social vulnerability indices are increasingly developed there are yet few attempts to validate these indices.

Correspondence to: A. Fekete
(fekete@ehs.unu.edu)

Validation can be achieved by using an independent second data set, preferably at a finer spatial level of resolution. Alternatively, the internal reliability of the model can be tested on sensitivity and uncertainty by statistical tests. This shows that there are two sides to the validity of an index, conceptual and methodological validity. Conceptual validity of the content of social vulnerability is more important yet more difficult to achieve and estimate. It demands terminological clarity, empirical evidence and a theoretical framework of social vulnerability.

Attempting to validate a social vulnerability index meets several constraints. First, it is difficult to find empirical evidence about social vulnerability itself. Social vulnerability is often hidden, complex and nested in various human aspects and contingencies bound to different levels of society. Second, vulnerability as a concept is conceived in at least two major ways. On the one hand it is perceived as a holistic and generic concept, encompassing many complex interrelations. On the other hand it is seen as a more single-dimensional concept, focusing on one specific item to a specific hazard, for example, evacuation needs in context to river-floods. Depending on each conceptualisation, vulnerability is better apt for experimental tests. Third, social vulnerability is difficult to estimate for methodological reasons. Indicators and indices are indirect numerical surrogates of real phenomena. Quantitative assessments of qualitative phenomena are subject to generalisations in order to achieve computation and comparability.

Due to these challenges it is not surprising that most social vulnerability index assessments are occupied mostly with the generation of the index and not with an additional validation. The creation of an index is largely constrained by the availability of data (King, 2001) and independent second data sets are scarce. Technical validation of such indices is increasingly demanded, for example by using sensitivity and uncertainty analysis (Wu et al., 2006) or random simulation tests of the index robustness by Monte Carlo tests (Gall, 2007). Validation at other spatial levels typically focuses on finding empirical evidence of social vulnerability. This approach is applicable, when the hazard context is the same, as well as the spatial and cultural context (O'Brien et al., 2004). The feasibility and success of such a validation at several scales is even more likely when the same theoretical framework is applied. For river-floods in Germany the BBC-framework (Birkmann, 2006) has been used to compare social vulnerability at county level to household level (Fekete et al., 2009). Still, there remains a need to first develop an index for social vulnerability and secondly, to find means of validating this index.

2   Objective

This paper will focus on the case of a Social Vulnerability Index (SVI) in the context of river-floods in Germany. The

SVI captures characteristics of certain social groups that render them exposed, susceptible or adaptive to disaster risk. The spatial extent of the research area comprises whole river channels in Germany. The objective behind the SVI in the context of river-floods is to identify, quantify, rank and validate social vulnerability in Germany. The input parameters to this index are selected according to the special focus of river floods in Germany, but are not directly dependent on flood related information. Rather, this index consists of population characteristics that could also be applied to other natural hazards. The selection criteria and thresholds for the SVI are for this case study explicitly targeted on a river flood context, for example, capturing elderly people above a certain age as being potentially vulnerable due to increased risk of fragility. This index can be principally applied to all potential flooding areas in Germany, to some extent even for coastal areas. This index bears no direct dependence on hazard parameters which is necessary in order to capture a different dimension of disaster causation than typically derived by flood risk approaches.

The Social Vulnerability Index is to show the potential vulnerability of a county, as a profile of general typified demographic profiles, settlement patterns and infrastructure information. It is not a target to capture profiles of single individuals or single buildings at this level. Rather, variables like the number of unemployed people per county point at general characteristics that are attributed to the average of unemployed persons per county. While many unemployed individuals might not suffer from financial shortcomings, this is however a most likely assumption for the general group of unemployed people per county.

The objective behind the validation is to find evidence whether the construction of a Social Vulnerability Index without direct relation to disaster impact or hazard parameters is valid. That means that first, test categories have to be found, which allow probing for revealed social vulnerability. Second, the independent variables that are the input data for social vulnerability indicators have to be checked on validity. Third, the methodology of grouping variables to indicators has to be checked. Only then conclusions can be drawn on the construction of an index composed of the single indicators and the patterns of social vulnerability that are indicated by such an index for spatial regions such as counties.

3   Social vulnerability factors

The data used for capturing the social vulnerability characteristics at county level in Germany are standard census data of the Federal Statistical Office in Germany (BBR, 2007; Destatis, 2006a). This data is first analysed by a factor analysis and in a second step validated by analysing an independent second data set with a logistic regression model. The scope of the factor analysis is to extract profiles of social groups regarding certain characteristics like income, gender

or age that can be linked to a certain extent to measurable variables like building type, urban or rural context, and medical care as a key element of coping capacities. The aim of using factor analysis as a method lies in variable reduction in order to derive a set of variables that summarise social vulnerability characteristics. Further, underlying (latent) structures of variables groups can be elicited to build a social vulnerability profile.

Methodology: Factor analysis is a multivariate analysis technique used to identify information packaging considering the interdependencies between all variables (Bernard, 2006: 495). The factor analysis is computed in SPSS version 14.0 with a principal component analysis for data reduction and in order to identify variable groupings. The methodology of the factor analysis follows standard procedure (e.g. Nardo et al., 2005). The principal component analysis aims at finding a linear combination of variables that accounts for as much variation in the original variables as possible. A Varimax rotation with Kaiser Normalisation is applied to the component matrix in order to ease the interpretation (Schneiderbauer, 2007: 55) by rotating the axes of the components perpendicular to each other. This step places the respective components as much apart from each other as possible. The extracted communalities are all above 0.5 which indicates that the extracted components represent the variables well. For the interpretation, only eigenvalues greater than one are regarded and absolute loading values below 0.30 suppressed (Nardo et al., 2005: 40, 43; Bühner, 2006: 200, 211; Bernard, 2006: 677). The eigenvalue is the standardised variance associated with a particular factor. The scree plot serves as another criterion to limit the number of factors. The factors on the steep slope up to the "scree elbow" in the curve are especially able to explain the most of the data (Fig. 1).

The factor analysis follows the principle of variance maximisation, wherein those factors are sought-after that explain most of the variance of all items (Bühner, 2006: 182). The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) of 0.905 indicates that the variable selection is suitable for factor analysis. The KMO explains the proportion of variance in the variables that might be caused by underlying factors. KMO-values above 0.60 indicate an acceptable level, and from 0.80 a good level of compatibility of the variables with the test (as cited in Bühner, 2006: 207). The value below 0.05 of the Bartlett's test of sphericity rejects that the variables are unrelated and therefore unsuitable for structure detection.

The factor analysis of an input of 41 variables (Table 1) uncovers seven latent factors that describe relationships between all variables to 76.6% of the cumulative variance. From these seven factors (or components), only the first three factors contain more than two loading values that get marked per column (Table 1). Marked are those values per variable which load highest within the seven factors matrix. Major latent groupings are identified by the factor analysis within the value loadings. These factors can be used for the com-
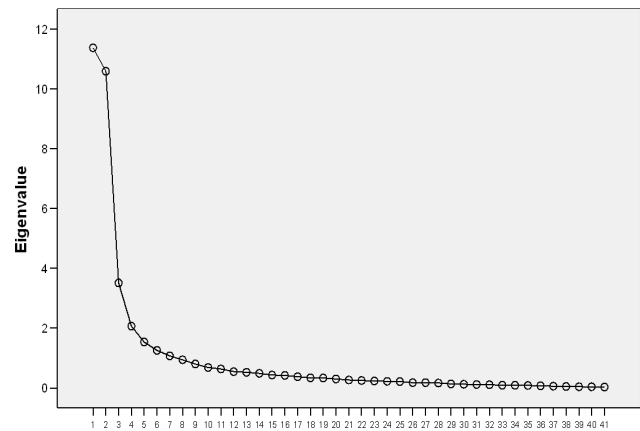

Fig. 1. Scree plot of the factor analysis showing the eigenvalues (y-axis) explained by the resulting factors (x-axis).

putation of the composite index. The first component can be named region as it captures variables connoted with regional urban or rural aspects like population density and housing type (Table 1). This factor consists of variables that contain both negative and positive characteristics of vulnerability towards river-floods, like density of medical care or population density. The second component depicts socio-economic conditions by financial deficiencies or income resources related to other variables that allow for interpretation of certain age and employment groups. The third component identifies fragility of physical fitness or need of assistance by the elderly as opposite to middle-aged adults.

Factor analysis is sensitive to the exclusion or selection of the variables for input. Yet, the selection set of 41 variables (Table 1) is internally sound which is revealed by testing of the factor analysis with stepwise exclusion of variables of the full-model. Additionally, the factor analysis is repeated stepwise by starting with two variables and stepwise inclusion of one more variable. By stepwise testing the patterns observed have been verified. However, this result bears no direct interrelation with hazard or disaster parameters. An index derived by factor analysis requires for either internal statistical reliability analysis, for example, Cronbach's alpha, or, even better, the validation by an independent second data set.

3.1 Validation of social vulnerability by flood impact analysis

The result of the factor analysis typically represents the endpoint of data analysis for the creation of a social vulnerability index. However, in this paper the results of the factor analysis are validated by testing the patterns of the factor analysis on a real case event. For the purpose of testing the social vulnerability profiles on a real flood event, a second data set is selected. The research question is whether a real extreme flood event reveals some of the potential social vulnerability that is expected from the vulnerability factors

Table 1. Rotated component matrix of the factor analysis showing the computed value loadings.

Input variables with presumed direction towards vulnerability: − more vulnerable, + more capacities	Component						
	1	2	3	4	5	6	7
− Residents below age 6		0.773	−0.423				
+ Residents from age 30 to 50			−0.850				
− Residents age 65 and older		−0.318	0.882				
− Persons in need of care			0.586		0.377		
− Handicapped unemployed		0.629					
− Female gender	0.632		0.545				
+ Income per hh		0.767			−0.343		
− Unemployment		−0.830	0.330				
+ Female employed	0.821						
+ Foreign employed		0.705					
+ High qualification employed	0.737	−0.307				−0.329	
− Foreign females		0.828					
− Social welfare recipients	0.433				0.655		
− Rent subsidies		−0.811					
− Graduates without basic education		−0.415			0.380		0.540
+ Graduates with high school graduation	0.740	−0.337					
+ University students	0.719						−0.454
− Foreigners	0.597	0.618					
− Residents per doctor	−0.829						
+ Hospital beds	0.707					0.348	
+ Rural population	−0.724			0.303			
− Population per settlement area	0.833					−0.358	
+ Open space	−0.735				−0.383		
+ Building land prices	0.634	0.484					
− Commuters in	0.734						
+ New apartments		0.350	−0.681				
+ One and two family homes	−0.819						
− Small apartments	0.824			0.378			
+ Living space pp	−0.351	0.583				0.444	
− Persons per hh	−0.756		−0.376				
− New residents	0.697	0.340	−0.369				
− Municipality debts per resident					0.567		
− Tourist overnight stays				0.904			
+ GDP per labour force		0.637					0.396
− Key funds allocation		−0.800					
+ Fixed investments	−0.375	−0.613			−0.359		
− Day-care centre		−0.866					
− Rehabilitation centres per Resident				0.840			
− Elementary Schools per Resident	−0.649						
− Medical care centres			0.451			0.618	
− Population projection age 60+		−0.736	0.580				

Interpretation:							
Positive value loadings	Urban	Young, income, foreigners	Old, fragile	Tourism	Welfare, debts	Care centres	Low education
Negative value loadings	Rural	Financial deficiencies	Mid-age, home owners				
Percent variance explained Total=76.6%	26.1%	22.1%	10.9%	5.4%	4.7%	4.6%	2.8%
Factor name	Region	Socio-economic conditions	Fragility				

Abbreviations: hh = household, pp = per person. Varimax rotation, PCA, *N*=439.

in Table 1. Due to the lack of data on extreme event evidence in Germany, validation is difficult. The German Research Centre for Geosciences Potsdam (GFZ) and Deutsche Rückversicherung kindly provided a data set on a household survey after the extreme floods of 2002 in Germany (GFZ Potsdam und Deutsche Rückversicherung AG, 2003). While the scope of this survey is mainly on flood damage characteristics of buildings and properties (Kreibich et al., 2005a), it also deals with flood preparedness and recovery (Thieken et al., 2007). Additionally, the survey captures demographic categories which are of special interest for validating the social vulnerability profiles. The survey covered three major regions, the River Elbe and the lower Mulde River; the Erzgebirge (Ore Mountains) and the River Mulde in Saxony; and the Bavarian Danube catchment (Thieken et al., 2007: 1020). In each region about the same number of interviews was conducted, altogether 1697 households, who were all affected by floods in 2002. One variable, urbanity, is additionally calculated by the author according to the definition of rural areas in the respective variable of the first data set of the Federal Offices (BBR, 2007). Rural areas are regions with up to 150 persons per km^2 per municipality. These areas are calculated in the Geographic Information System (GIS) using the federal statistical data and provide important additional information about population density.

In order to find evidence whether the presumed social vulnerability concept and -profiles play a role in the outcome of disaster, a testing category has to be identified. For the purpose of this study, the question "Did you have to leave your home due to the flood?" is identified as a discriminator of people severely affected by the flood in terms of social vulnerability. This question is not focusing on the economic perspective only, but captures a broader scope of exposure, susceptibility and capacities. The people who had to leave their home (765 of 1697) were especially exposed to floods, had to cope with finding an interim shelter and needed a recovery phase after the flood. For this they needed financial resources but also social networks like relatives or friends.

A certain amount of those people (N=765) who had to leave their home, had to seek emergency shelter (N=70). This is an especially interesting sub-group, because it can be assumed that these persons lacked alternative social networks or financial resources. Since the questionnaire contains no questions about the exact reasons for each decision of the single individuals in the survey, these are mere assumptions. They can be however compared to findings on social vulnerability in evacuation groups (Cutter et al., 2003; Chakraborty et al., 2005). Therefore, "people forced to leave their home" and "people who had to seek emergency shelter" are apt test categories for eliciting different social group profiles. It permits comparing those who had to leave and those who could stay in their homes, despite being affected by the flood.

The third test category is taken from the question "are you satisfied with the status of damage regulation?". The answers were expressed in a positive to negative range from one to

Table 2. Dependent and independent variables used for all three logistic regressions.

Dependent variables	Independent variables or sub-variables	
leave_home	age	unemployed
emergency_shelter	gender	pensioner
regulation	high_school_degree	residents up to 14 years
	elementary_school	persons per household
	income_very_high	rooms
	income low	home ownership
	high_qual_employed	urbanity

N=[765; 960]

six. This range is transformed into binary coding for bivariate comparison. Indirect financial needs and satisfaction with administration are to a certain degree identified by this dependent variable. This type of susceptibility measure therefore supplements the other two dependent variables, which capture exposure and evacuation needs.

Methodology: The logistic regression is computed with the second data set for the three binary dependent variables individually. Each dependent variable is analysed with the same pre-selected sub-set of independent variables (Table 2). The main purpose of the logistic regression is to show whether there exists a significant difference in the independent variables. The independent variables contain demographic vulnerability characteristics (e.g. age of persons) and are checked against dependent variables that contain binary yes/no cases. For example, independent variables like age are checked within the full logistic regression model against the dependent variable leave_home, whether age is a factor that characterises human groups as more vulnerable.
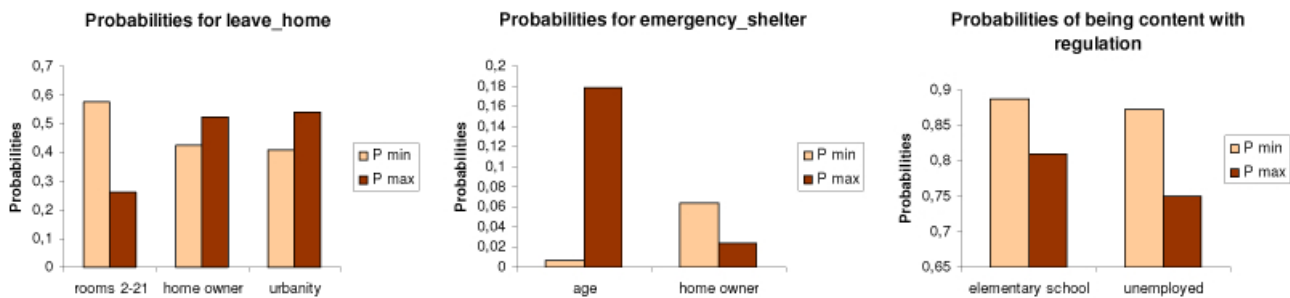
The logistic regression provides two types of measurement that are of interest here. First, the regression model indicates which independent variables are significant within the full model; only these are selected for calculating the probabilities (marked in Table 3). Second, the probabilities calculated for the minimum and maximum values per independent variable predict the direction of impact of the dependent variable (Table 3 and Fig. 2). This direction can be positive or negative, meaning that flood impact either rises with increasing values like higher income or is inversely related to it.

3.2 Results

For the regression with the dependent variable leave_home, the number of rooms, home ownership and degree of urbanity are capable to explain the distribution of those who had to leave their home and those who had not (Table 3 and Fig. 2). The probability for the dependent variable (leave_home, answer "yes") can increase or decrease for each independent variable. Therefore, the variable rooms (number of rooms from 2–21) shows an inverse relationship of

Table 3. Calculated significances, probabilities and confidence intervals for the three dependent variables.

Dependent variable	Independent variable	Significances				Probabilities		
		Sig.	95.0% lower CI	95.0% upper CI	P min	P max	95% CI change min	95% CI change max
leave_home	rooms [2; 21]	0.024	0.877	0.991	0.5755	0.2624	−0.5506	−0.0756
	home ownership	0.019	1.066	2.053	0.4272	0.5245	0.0167	0.1779
	urbanity	0.000	1.272	2.261	0.4091	0.5399	0.0608	0.201
emergency_shelter	age	0.012	1.010	1.081	0.0067	0.1785	−0.0537	0.3974
	home ownership	0.003	0.175	0.707	0.0636	0.0233	−0.0752	−0.0052
regulation	elementary school	0.019	0.325	0.905	0.8873	0.8102	−0.1453	−0.0089
	unemployed	0.020	0.221	0.881	0.8719	0.7502	−0.2457	0.0024

Fig. 2. Minimum and maximum probabilities of the three logistic regressions.

probability (Fig. 2). The higher the number of rooms, the lower is the quota of those in the group who had to leave their home. In other words, people living in apartments with fewer rooms had to leave their home more often. The higher the number of rooms, the lower is the quota of those in the group who had to leave their home. The higher the number of home owners in comparison to tenants the more likely it has been that these households had to leave their home due to the flood. Persons in rural areas (up to 150 km^2) are less affected than residents in urban areas.

For the regression with the dependent variable emergency_shelter, age (from 16–95) and home ownership are apt to explain the distribution of those who had to seek emergency shelter and those who had not (Table 3 and Fig. 2). Higher age was a reason to seek emergency shelter. The higher the number of home owners in comparison to tenants the more likely it has been that these households had not to seek emergency shelter due to the flood. This contradicts the prediction direction of ownership in the dependent variable leave_home.

For the regression with the dependent variable regulation, elementary school and unemployment explain the distribution of satisfaction with damage regulation (Table 3 and Fig. 2). Persons with low education background (elementary school) are more dissatisfied with damage regulation. The same observation is made for unemployed people.

The quality of the statistical model is supported by the Hosmer and Lemeshow Test which describes the model-goodness of fit of the input data for values with significance values above 0.05 (Backhaus et al., 2006: 457). The respective confidence intervals for the variables are observed whether they are either both below or above value one (Table 3), which supports that these independent variables deliver a valid explanation (Fromm, 2005: 24). Error margins are indicated by the quality tests described above, or by the confidence intervals. Multi-collinearities are avoided, which is shown by the variance inflation factors lower than three (Nardo et al., 2005). Outliers that could distort the model are identified by z residuals and removed. Monte Carlo tests like the Jacknife replication tests (Backhaus et al., 2006: 454) and bootstrapping with 1000 repetitions (Moore, 2006: 14–27) underscore the stability of the results for the dependent variables leave_home and emergency_shelter, while regulation was not stable with the current variable set in the bootstrap test.

The probabilities are used here only for the identification of the direction of influence of flood impact. The probabilities are not used for weighting or relative ranking of the variables since this is only an explorative approach. Further uncertainty analyses and additional confirmative analyses of flood impact cases would be a requisite for justifying the use of exact numerical values for weighting and ranking.

4 Creation of a validated social vulnerability index

The validation procedure comprises two steps; first the independent variables of the factor analysis (Table 1) are checked for validity by using the independent variables of the independent second data set and running a logistic regression model. Since the second data set did not capture exactly the same demographic variables, only a few independent variables of the first data set are at the same time available from the second data set (Table 4). The logistic regression analysis reveals that six independent variables of the second data set are able to discriminate vulnerability. These six variables capture demographic as well as spatial parameters that are also captured by nine independent variables of the first data set. That means that in the first step, certain variables have been validated as having a significant effect in determining vulnerability.

This knowledge can be transferred to the subset of nine variables out of the 41 variables of the first data set. It would be unsafe to suggest that the full model of 41 variables is validated by this process. However, at least nine variables of the first data set can be assumed to describe vulnerability. The remaining 32 independent variables are not significant within the regression model or can not be tested as they are not contained in the second data set. Of course, this does not imply that they can't be significant within another model or are not meaningful.

In the second step of the validation, the factor analysis is repeated with the subset of nine independent variables of the Federal Statistics. The objective behind this second step of validation is to check whether the factors (or social vulnerability indicators) obtained without any direct disaster-relation are similarly revealed by the reduced set of nine validated variables.

As a result of the validation, the three factors derived from the nine variables as grouped in the rotated component matrix (Table 5) display the same factors that have been identified with the full variable set of 41 variables with the full federal statistics set (Table 1). This result reveals that the factors are generally valid.

The factors derived from the factor analysis are the basis for the selection and aggregation of the Social Vulnerability Index. Each factor delivers one indicator, the aggregated indicators make up the index. In order to enable positive and negative indications of vulnerability to the same degree, each factor must have the potential to indicate both directions equally. The variables are first standardised to equal intervals from zero to one. Missing values are replaced with the average value of the variable so that in the average of all either negative or positive variables they do not invoke a trend. The resulting direction of vulnerability is different for each county. The factors are used as the three indicators of social vulnerability.

Table 4. Comparison of the nine variables of the Federal Statistics with the according variables of the logistic regression.

Variables of the logistic regression	Variables of the factor analysis from the first data set
urbanity (urban areas have more than 150 persons per km^2 per municipality)	Population per settlement area
home ownership	One and two family homes
urbanity (rural areas have less than 150 persons per km^2 per municipality)	Rural population
rooms [2; 21]	Small apartments
age	Residents from age 30 to 50
age	Residents age 65 and older
unemployed	Unemployment
rooms [2; 21]	Living space pp
elementary school	Graduates with only elementary education
Data source: GFZ and Deutsche Rück household survey (2002); urbanity definition after BBR (2007)	Data source: Federal Statistics (2006)

$$\text{Indicator} = \frac{\text{positivevar1} + \text{positivevar2} + \dots \text{varx}}{\text{number\_of\_var}}$$
$$- \frac{\text{negativevar1} + \text{negativevar2} + \dots \text{varx}}{\text{number\_of\_var}}$$

SVI = Indicator1 + Indicator2 + Indicator3

4.1 Result: the Social Vulnerability Index (SVI) per county in Germany

Description: The SVI (Fig. 3) identifies demographic patterns of susceptibility, capacities and potential exposure towards river-floods. It is computed as the simple sum of three indicators: fragility, socio-economic conditions and region. The indicator fragility consists of the ratio of elderly residents (>64 years); the indicator socio-economic conditions consists of living space per person, (un)employment ratio and education type; the indicator region consists of population density and housing type. Census data of the Federal Statistical Office in Germany are used and the results are displayed in ratios per county as equal intervals from zero to one.

Low SVI counties (Fig. 3) are characterised by strengths towards river-floods. These strengths are prevailing capacities for river-flood mitigation, for example, financial capacities for private preparedness measures and recovery from floods by high income sources. These counties lack indications for potential exposure to floods like high population density. Susceptibility like physical fragility of elderly

Table 5. Rotated Component Matrix of the nine variables of the federal statistics that are validated by the logistic regression.

Component	1	2	3
Population per settlement area	−.951		
One and two family homes	.856		−.358
Rural population	.831		
Small apartments	−.788		
Residents from age 30 to 50		−.935	
Residents age 65 and older		.913	
Unemployment		.383	.853
Living space pp	.416		−.716
Graduates with only elementary education			.697
Factor name	Region	Fragility	Socio-economic conditions

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalisation. Rotation converged in 4 iterations. Total variance explained: 78.8%. Kaiser-Meyer-Olkin Measure of Sampling Adequacy: 0.7.

citizens is also typically low. Counties with high SVI are characterised by predominating weaknesses towards river-floods. These weaknesses are lack of capacities, high degrees of susceptibility and indications for exposure potential.

Hazard context: The SVI detects potential strengths and weaknesses of counties, not the actual river-flood hazard or -risk. The SVI contains no hazard information and therefore no actual exposure. However, the SVI is not an index for any kind of natural hazard, since the variables are selected and aggregated only after flood impact evidence. The input variables for the indicators are created after verified flood impact to different social groups and settlement types. Counties have distinct profiles of social vulnerability, composed of demographic characteristics, population density and settlement type. The strength of the SVI is its independence from direct hazard information. It identifies key aspects of flood impact and –risk not identified by hazard assessments.

5 Discussion

The standard statistical methods of factor analysis and logistic regression have been selected for their ability to consider multivariate interdependencies between the selected variables. This seems an appropriate measure for capturing the multiple interdependencies in the case of social vulnerability and allows for elicitation of hitherto unclear interrelations between demographic and spatial variables on county scale in Germany. While factor analysis is open to shifts in the factor composition by the selection of input variables, it permits a transparent and comprehensible interpretation through the display of the matrix with the value loadings. The logistic re-
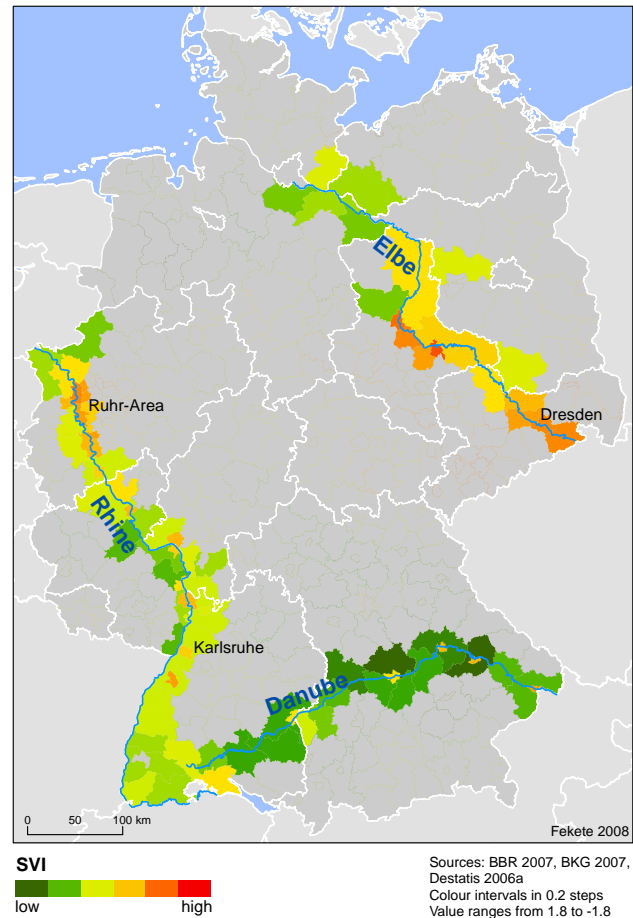

Fig. 3. Main result of the social vulnerability assessment, the map of the Social Vulnerability Index (SVI) per county.

gression is able to analyse multiple interdependencies of the input variables against certain test categories of social vulnerability outcome. As in the factor analysis, this technique is also sensitive to the selection of input parameters. However, for both techniques the stepwise inclusion and exclusion of input variables as well as the sensitivity tests underscored that the general pattern of interdependencies is stable. This combination of techniques successfully demonstrates that it is possible to identify latent interdependencies of social vulnerability, to group them to indicators and to find a method to validate these indicators. The validation shows that the selected variables are valid in identifying social vulnerability as a decisive factor in context to river-floods; second, that composite indicators can be created without direct relation to hazard parameters.

Certain conditions and limitations of validation must be emphasised. Although the research area for this questionnaire is relatively large and covers three federal states, it still is difficult to generalise the results for whole river channels. More case studies are necessary to cover other regions in Germany. The questionnaire contains vital data categories, but was not intentionally invented for the purpose of

validating a social vulnerability indicator or the data of this study. Therefore not all variables can be covered for validation.

The choice of the dependent variables is based on the assumption, that the fact that someone has to leave his home or seek emergency shelter is a severe impact. Although this type of measure is used in literature to identify social vulnerability (Chakraborty et al., 2005) it is not sufficiently explored to which extent it reveals social vulnerability in Germany. For the explorative scope of this pilot study it is though an excellent opportunity with a lack of comparable attempts. Regulation satisfaction is chosen to extract coping problems of an indirect economic, administrative and perception character. The selection of variables, the exclusion of sub-variables and the setting of thresholds is dependent on assumptions and decisions of the author.

The results of the analyses show that home owners had more often to leave their home for a certain period yet less often to seek emergency shelter in comparison to all interviewees affected by the flood. This could be interpreted as home owners living more exposed, yet having more financial and social resources at hand in such an evacuation situation, compared to the average. It could also be thought that home owners may have more financial potential to privately prepare for risks, however, this relation is not visible from the data presented in this paper and therefore such an interpretation should be avoided. Urban residents had more often to leave their home than rural residents, so it might be inferred that population density plays a role. Many elderly residents had to seek emergency shelter, which could be assumed to relate to lack of financial or social resources for alternative shelter or accommodation at relatives, friends, or hotels. Persons with low education background and unemployed people are more dissatisfied with damage regulation. This might accord to observations in other social vulnerability studies that people with low qualification have less access to damage compensation which might itself be based on a couple of other reasons. As with all interpretations here, as previously stated, one has to be very careful, since the underlying reasons for providing each individual answer to the interview are not recorded. Therefore this semi-quantitative approach allows for identifying patterns of social vulnerability, yet aims not at finding causal explanations.

Besides all these necessary disclaimers it is satisfying that the overall picture complies with a great extent to the findings of previous studies in Germany (Steinführer and Kuhlicke, 2007; Birkmann et al., 2008) and other countries (Morrow, 1999; Tapsell et al., 2002; Cutter et al., 2003; Schneiderbauer and Ehrlich, 2006; Simpson and Katirai, 2006; Masozera et al., 2007). Therefore the directions of impact of Table 3 and Fig. 2 are useful to justify the use of certain variables that characterise age, settlement and apartment type, education and financial deficiencies for the construction of a social vulnerability index.

6 Conclusions

The main result of the vulnerability assessment is the SVI (Fig. 3), composed of three indicators derived from the factor analysis of census data and validated by an independent second data set. While most common sub-national quantitative vulnerability assessments stop at the creation of vulnerability indicators, this paper demonstrates further validation steps. In the first step, the vulnerability factors as identified by factor analysis are tested in a historic real flood event by impact analysis. The logistic regression reveals that indeed certain demographic vulnerability characteristics discern between the affected and less affected groups. With the subset of those variables that could be validated, the factor analysis is rerun. The same factors emerge which additionally supports the validity of the approach. Three main factors are identified, and are used as indicators of social vulnerability: region, social conditions and fragility.

The SVI identifies counties in Germany with a potentially strong or weak social vulnerability to floods. Since social vulnerability is regarded independent of the individual river flood hazard, this index contains no hazard information. This index delivers the prerequisite for an integrated risk assessment, which spans both hazard and vulnerability analyses. The outcome of a coupled hazard-vulnerability assessment can, for example, result in a disaster risk index. The creation of such an index and the applications for disaster risk reduction are outlooks for further research.

References

Acosta-Michlik, L.: Intervulnerability Assessment, Shifting Foci from Generic Indices to Adaptive Agents in Assessing Vulnerability to Global Environmental Change (A Pilot Project in the Philippines), University Catholic of Louvain (UCL), Department of Geography, Louvain-la-Neuve, Belgium, 2005.

Backhaus, K., Erichson, B., Plinke, W., and Weiber, R.: Multivariate Analysemethoden, 11th edn., Springer, Berlin, 2006.

BBR: INKAR 2006, Indikatoren, Karten und Graphiken zur Raum- und Stadtentwicklung in Deutschland und in Europa, BBR – Bundesamt für Bauwesen und Raumordnung (Federal Office for Building and Regional Planning), Bonn, 2007.

Bernard, H. R.: Research Methods in Anthropology, Qualitative and Quantitative Approaches, 4th edn., Altamira Press, Oxford, 2006.

Birkmann, J.: Measuring vulnerability to promote disaster-resilient societies: Conceptual frameworks and definitions, in: Measuring Vulnerability to Natural Hazards: towards disaster resilient societies, edited by: Birkmann, J., United Nations University Press, Tokyo, 9–54, 2006.

Birkmann, J., Krings, S., and Renaud, F.: Assessment of Vulnerability to Floods at Local Level With a Special Focus on Human-Environmental Systems and Critical Infrastructures, Proceedings of the 4th International Symposium on Flood Defence 6–8 May 2008, Managing Flood Risk, Reliability & Vulnerability, Toronto, Ontario, 2008,

Bühner, M.: Einführung in die Test- und Fragebogenkonstruktion, 2nd edn. Pearson Studium, München, 2006.

Cardona, O. D.: System of Indicators for Disaster Risk Management, Program for Latin America and The Carribean, Main Technical Report, Inter-American Development Bank, Sustainable Development Department, Washington, D.C., 2005.

Chakraborty, J., Tobin, G. A., and Montz, B. E.: Population Evacuation: Assessing Spatial Variability in Geophysical Risk and Social Vulnerability to Natural Hazards, Natural Hazards Review, 23–33, 2005.

Clark, G. E., Moser, S. C., Ratick, S. J., Dow, K., Meyer, W. B., Emani, S., Jin, W., Kasperson, J. X., Kasperson, R. E., and Schwarz, H. E: Assessing the Vulnerability of Coastal Communities to Extreme Storms: The Case of Revere, MA, USA, Mitigation and Adaptation Strategies for Global Change, 3, 59–82, 1998.

Cutter, S. L., Mitchell, J. T., and Scott, M. S.: Revealing the Vulnerability of People and Places: A Case Study of Georgetown County, South Carolina, Ann. Assoc. Am. Geogr., 90, 713–737, 2000.

Cutter, S. L., Boruff, B. J., and Shirley, W. L.: Social Vulnerability to Environmental Hazards, Soc. Sci. Quart., 84, 242–261, 2003.

Damm, M., Fekete, A., Uhlemann, S., and Zwenzner, H.: Development of an Information System for Large-Scale Flood Events Supported by Remote Sensing (Project DISFLOOD), Beiträge zur Konferenz "Strategien und Instrumente zur Verbesserung des vorbeugenden Hochwasserschutzes" 23–25 November 2005 in Tangermünde, 133–141, 2006.

Dwyer, A., Zoppou, C., Nielsen, O., Day, S., and Roberts, S.: Quantifying Social Vulnerability: A methodology for identifying those at risk to natural hazards, 92 pp., 2004.

Fekete, A., Damm, M., and Birkmann, J.: Scales as a challenge for vulnerability assessment, Natural Hazards, in review, 2009.

Fromm, S.: Binäre logistische Regressionsanalyse, Eine Einführung für Sozialwissenschaftler mit SPSS für Windows, Bamberg, 2005.

Gall, M.: Indices of social vulnerability to natural hazards: A comparative evaluation, PhD dissertation, Department of Geography, University of South Carolina, 2007.

Hahn, H., Villagrán De Leon, J. C., Hidajat, R., Schaef, T., and Bollin, C.: Component III: Indicators and other disaster risk management instruments for communities and local governments, Inter-American Development Bank, Washington, D.C., 2003.

King, D.: Uses and Limitations of Socioeconomic Indicators of Community Vulnerability to Natural Hazards: Data and Disasters in Northern Australia, Natural Hazards, 24, 147–156, 2001.

Kleinosky, L. R., Yarnal, B., and Fisher, A.: Vulnerability of Hampton Roads, Virginia to Storm-Surge Flooding and Sea-Level Rise, Natural Hazards, 40, 43–70, 2007.

Kreibich, H., Thieken, A. H., Petrow, Th., Müller, M., and Merz, B.: Flood loss reduction of private households due to building precautionary measures – lessons learned from the Elbe flood in

August 2002, Nat. Hazards Earth Syst. Sci., 5, 117–126, 2005a, http://www.nat-hazards-earth-syst-sci.net/5/117/2005/.

Kropp, J. P., Block, A., Reusswig, F., Zickfeld, K., and Schellnhuber, H. J.: Semiquantitative assessment of regional climate vulnerability: the North-Rhine Westphalia study, Climatic Change, 76, 265–290, 2006.

Masozera, M., Bailey, M., and Kerchner, C.: Distribution of impacts of natural disasters across income groups: A case study of New Orleans, Ecol. Econ., 63, 299–306, 2007.

Meyer, V., Haase, D., and Scheuer, S.: GIS-based Multicriteria Analysis as Decision Support in Flood Risk Management, Department of Economics, Leipzig, 2007.

Moore D. S. and McCabe, G. P.: Introduction to the practice of statistics, Freeman, New York, 2006.

Morrow, B. H.: Identifying and Mapping Community Vulnerability, Disasters, 23, 1–18, 1999.

Nakamura, T., Hutton, C., Kunbao, X., and Gavidia, J.: Assessment of vulnerability to flood impacts and damages, UNCHS (Habitat), 39 pp., 2001.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., and Giovannini, E.: Handbook on constructing composite indicators: methodology and user guide, 2005.

O'Brien, K., Leichenko, R., Kelkar, U., Venema, H., Aandahl, G., Tompkins, H., Javed, A., Bhadwal, S., Barg, S., Nygaard, L., and West, J.: Mapping vulnerability to multiple stressors: climate change and globalization in India, Global Environ. Chang., 14, 303–313, 2004.

Olfert, A., Greiving, S., and Batista, M. J.: Regional multi-risk review, hazard weighting and spatial planning response to risk – Results from European case studies, in: Geological Survey of Finland, Special Paper 42, Natural and Technological Hazards and Risks Affecting the Spatial Development of European Regions, edited by: Schmidt-Thomé, P., 125–151, 2006.

Rygel, L., O'Sullivan, D., and Yarnal, B.: A Method for Constructing a Social Vulnerability Index, Mitigation and Adaptation Strategies for Global Change, 11, 741–764, 2006.

Schneiderbauer, S. and Ehrlich, D.: Social levels and hazard, (in)dependence in determining vulnerability, in: Measuring Vulnerability to Natural Hazards: towards disaster resilient societies, edited by: Birkmann, J., United Nations University Press, Tokyo, 78–102, 2006.

Schneiderbauer, S.: Risk and Vulnerability to Natural Disasters – from Broad View to Focused Perspective, Doctoral thesis, FU Berlin, 2007.

Simpson, D. M. and Katirai, M.: Indicator Issues and Proposed Framework for a Disaster Preparedness Index (DPi), University of Louisville, 49 pp., 2006.

Steinführer, A. and Kuhlicke, C.: Social vulnerability and the 2002 flood, Country Report Germany (Mulde River), Leipzig, 2007.

Tapsell, S. M., Penning-Rowsell, E. C., Tunstall, S. M., and Wilson, T. L.: Vulnerability to flooding: health and social dimensions, Philos. T. Roy. Soc. A, 1511–1525, 2002.

Thieken, A., Kreibich, H., Müller, M., and Merz, B.: Coping with floods: preparedness, response and recovery of flood-affected residents in Germany in 2002, Hydrolog. Sci. J., 52, 1016–1037, 2007.

Weichselgartner, J.: About the capacity to be wounded: the need to link disaster mitigation and sustainable development, Extreme Naturereignisse – Folgen, Vorsorge, Werkzeuge, DKKV, Bonn,

150–158, 2002.

Wu, S.-Y., Yarnal, B., and Fisher, A.: Vulnerability of coastal communities to sea-level rise: a case study of Cape May County, New Jersey, USA, Climate Res., 22, 255–270, 2002.

Yarnal, B.: Vulnerability and all that jazz: Addressing vulnerability in New Orleans after Hurricane Katrina, Technol. Soc., 249–255, 2007.